DEVELOPMENT OF A SUBTRACTIVE HYBRIDIZATION TECHNIQUE AND THE CLONING OF THE MOUSE EPSILON ISOFORM 14-3-3 PROTEIN, A GENE INVOLVED IN FETAL KIDNEY DEVELOPMENT.

Jane E. McConnell

Ph.D.
University of Edinburgh
1994

ABSTRACT

The kidney and its ontogenesis has been studied since the turn of the century as a model system for the complex cellular inductive interactions and the mesenchyme to epithelial transition typical of mammalian organogenesis. Transplantation and micro-dissection experiments in many vertebrates have created a depth of knowledge of the cellular basis of nephrogenesis. Conversely, the events at the molecular level have only recently started to be explored and are still poorly understood. Obviously, there are crucial molecular interactions that guide proper kidney development but the genes responsible have not yet been found and properly analysed.

The aim of this project was to devise a simple and efficient method to identify cDNAs expressed specifically in the mouse fetal kidney which might be important in kidney development.

A 14.5 day post coitum fetal kidney cDNA library was made which has a depth of 1.2×10^6 original clones and an average insert size of 1.5 kilobases. Using a novel subtractive hybridization technique which identifies genes expressed in the fetal kidney but not in the adult liver, a subtracted 14.5 day post coitum fetal kidney cDNA pool was obtained. By screening the library with the subtracted fetal kidney cDNA, ten clones were identified. Six, when analysed by sequencing and northern blot analysis, were shown to contain B1 embryonic repeats and were not analysed further. Another was not characterised. Three proved to be the most interesting. Further characterisation identified the clones as three independent isolates of a cDNA encoding the ε member of the 14-3-3 gene family, implicated in the regulation of protein kinase C activity and of exocytosis. Northern blot analysis of RNA identified a 2.0 kilobase transcript present in many tissues including the fetal kidney, but not in adult liver. More detailed *in situ* hybridization analysis showed that, although the gene is expressed in most 12.5 day post coitum embryonic mouse tissues, the level of expression varies. In general, it is low in epithelial tissues and high in mesenchymal cells. As the mesenchyme in bone kidney gut and lung differentiates, the level of

expression drops. This observation argues that the 14-3-3 ε isoform protein plays a role in mesenchyme differentiation including acting as the kidney mesenchyme epithelialises.

The project has created a valuable fetal kidney library and developed a rapid screen for important genes in kidney development. The screen performed has shown that subtraction against fetal mRNA is advisable to minimise enrichment for B1 repeats. The cloning of the ε 14-3-3 protein shows that genes that may be important in differentiation can be found using the existing subtracted cDNA pool.

ABBREVIATIONS

Amp	ampicillin
aa	amino acid
APS	ammonium persulphate
ATP	adenosine triphosphate
bis-acrylamide	N-N'-methylenebisacylamide
bp	base pair
CAM	cell adhesion molecule
cAMP	cyclic adenosine monophosphate
cDNA	complementary deoxyribonucleic acid
CTP	cytidine triphosphate
Ci	curie
Da	dalton
DEAE	diethylaminoethyl
DEPC	diethyl pyrocarbonate
dH_2O	single distilled water
DNA	deoxyribonucleic acid
dNTP	deoxy nucleotide triphosphate
d.p.c.	days post coitum (post conception)
dpm	disintegrations per minute
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratories
g	acceleration due to gravity ($9.81 m/s^2$)
GBM	glomerular basement membrane
GTP	guanosine triphosphate
HEPES	4-(2-hydroxyethyl) -1- piperazine -N' -2- ethane-sulfonic acid
HPLC	high pressure/performance liquid chromatography
IPTG	isopropyl-β–galactothiopyranoside
k	kilo= (1×10^3)
kb	kilobases
kDa	kilo daltons
l	litre
μ	micro

M	moles per litre
m	milli (x 10^{-3})
min	minute
MOPS	4-morpholine-propane-sulfonic acid
mRNA	messenger RNA
N-CAM	neural cell adhesion molecule
NGFR	nerve growth factor receptor
OD	optical density
o/n	overnight
PAB	photo activatable biotin
pc	post coitum (post conception)
PCR	polymerase chain reaction
pfu	plaque forming unit
RNA	ribonucleic acid
rpm	revolutions per minute
RTK	receptor tyrosine kinase
SDS	sodium dodecyl sulphate
Tet	tetracycline
Tris	2-amino-2-hydroxymethyl-propane-1,1-diol
TTP	thymidine triphosphate
UTP	uridine triphosphate
UV	ultraviolet
UWGCG	University of Wisconsin's Genetics Computer Group
V	volt
v/v	volume per unit volume
w/v	weight per unit volume
W	watt
WT-1	Wilms' tumour susceptibility gene
X-Gal	5-bromo-4-chloro-indoyl-β-galactopyranoside

ONE LETTER ABBREVIATIONS USED IN DNA SEQUENCE

T	thymidine
A	adenosine
C	cytidine
G	guanosine
N	unknown nucleotide

ONE LETTER ABBREVIATIONS OF AMINO ACIDS

A	alanine
R	arginine
N	asparagine
D	aspartic acid
C	cysteine
G	glycine
E	glutamic acid
Q	glutamine
H	histidine
I	isoleucine
L	leucine
K	lysine
M	methionine
F	phenylalanine
P	proline
S	serine
T	threonine
W	tryptophan
Y	tyrosine
V	valine
X	unknown amino acid

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank :

Peter for stopping me from quitting when the unit drove me to suicidal lows and for the scientific advice and being the sounding board for most of the ideas that inspired this thesis. And perhaps most of all for holding me together during the long months of writing this epic when everyone else had let me down!

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 Aims of the thesis

The project was initiated to explore nephrogenesis at the molecular level. Kidney development has been studied at the morphological level for the last century, and the precise physical details have been observed and explored using transplantation, deletion experiments and electron-microscopy. The molecular events that drive these processes have only recently started to be explored and the initial results are beginning to create a more intimate understanding of the process. There are, however, still many pieces of the jigsaw for which the molecular driving forces have yet to be elucidated.

This is not going to be an easy task as the importance of any gene which may be shown to have specific expression in the nephron must be evaluated within the context of the morphological details of nephrogenesis. The expression pattern must also be reviewed in context of the genes already known to be involved in kidney development to try to understand the interactions that may be occurring at the expression level of the gene. While this thesis concentrates on the molecular aspects of nephrogenesis and the search for new genes integral to the process, any results must be interpreted within the context of normal kidney development. In order to understand the type of genes which may be important and at which stages the critical genes may start to exert their effect, a review of the normal development of the nephron and the kidney is needed to explain the context and aims of this thesis.

1.2 Evolution of the nephron

The nephron is the functional unit of vertebrate excretory system; its physiology and structure are remarkably similar in cyclostomes and mammals. The fact that a similar structure can be found in fossils of the oldest known vertebrates, the ostracoderms (Torrey, 1965) and in modern mammals suggests that the basic structure must have evolved early in the vertebrate lineage. In higher vertebrates, the excretory system has evolved in such a way that three sets of excretory organs develop sequentially (see Figure 1.1), first the pronephros, then the mesonephros and

finally the metanephros which will form the functional excretory system. All three systems develop from two identical blocks of mesoderm that form along the dorsal wall of the embryo which form from intermediate mesenchyme. This mesoderm is called the nephrogenic cord in the cervical and thoracic regions and the urogenital ridge in the coelomic cavity where it widens.

1.3 The nephric duct

The (pro)nephric duct is the central component of the excretory system throughout embryonic development. The duct provides the drainage channel of the functional (when applicable) pronephros and mesonephros and an out-growth from the duct, the ureteric bud, induces the formation of the metanephric mesenchyme. The distal region of the mesonephric duct ultimately contributes to the male genital system forming the ductus deferens (Tiedemann, 1976). The nephric duct's development is independent of the pronephros (Torrey, 1954), as surgical deletion of the pronephric anlage allows perfect formation of the nephric duct (Holtfreter, 1944) except in rabbits where the distal region of the pronephric tubules meld to form the first rudiment of the nephric duct (Toivonen, 1945). The duct primordium segregates from the dorsal portion of the lateral plate mesoderm and extends along the ventrolateral border of the somites. Duct growth occurs by cell multiplication and elongation through cell streaming and migration as shown by *in vitro* experiments (Overton, 1959). In some amphibia, duct elongation occurs by cell rearrangements which cause the duct to thin and the number of cells across the duct's diameter to reduce (Poole & Steinberg, 1981, 1984). In others, such as the axolotl, the migrating cells are guided by an adhesion gradient provided by the mesoderm, especially the vascular endothelium with which the epithelial cells are in contact through thin philipodia (Poole & Steinberg, 1982). The adhesive gradient is regulated in relation to the craniocaudal wave of differentiation (Zackson & Steinberg, 1987). The duct elongates until it reaches and eventually opens into the cloaca.

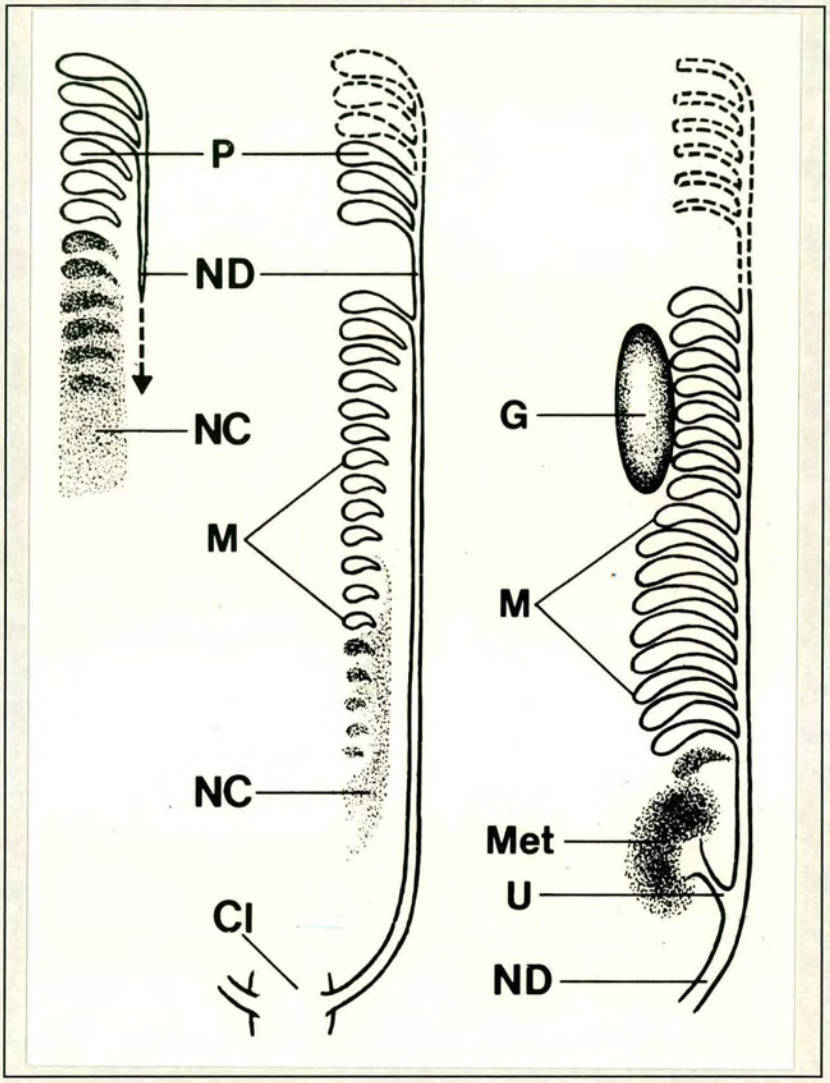


Figure 1.1

Schematic drawing of the development of a) the pronephros, b) the mesonephros and c) the metanephros. P, pronephros; ND, nephric duct; NC, nephrogenic cord; G, gonad; M, mesonephros; Met, metanephros; U, ureteric bud; Cl, cloaca (after Saxén 1987).

1.4 The pronephros

The pronephros is the first nephric system to develop and is the most ancestral of the excretory systems. Although the pronephros develops in higher vertebrates during early embryogenesis (8 days p.c. in mouse and 22 days p.c. in man) only very rudimentary development occurs and it has no actual function in any of the higher vertebrates apart from some teleosts [bony fish] (Saxén, 1987). The pronephros only develops into a functional excretory system in lower vertebrates, and then only during embryonal and larval stages. The functionality of the pronephros in amphibians was first shown when unilateral pronephroctomy in *Triturus* larva caused compensatory hypertrophy of the remaining pronephros (Fox, 1956). In all vertebrates, it regresses as the embryo develops. The pronephros consists of vaguely defined condensations, vesicles or grooves, the number of which varies in different species. The tubules open through the nephrostome (the nephric cavity) into the coelomic cavity and the tubule free ends can fuse to form the most anterior part of the pronephric duct. In mouse, the transitory pronephric nephrotomes are located between the somitic mesenchyme and the lateral plate caudal to the third somite.

1.5 The mesonephros

The mesonephros is the second of the excretory systems to develop in vertebrates and most of the study on its formation and structure has been done in amphibians and birds (Burns, 1955; Croisille, 1976) and in some domestic mammals. The number of tubules and the functional maturity of the mesonephros varies among different species; there are 30 in man and 50 in sheep and pig (Tiedemann, 1976). There is an inductive interaction between the nephric cord mesenchyme (the presumptive mesonephros) and the caudally extending nephric duct. The presumptive mesonephros only becomes terminally determined when the nephric duct comes in contact with it, suggesting that the morphogenesis of the two cell lineages are integrated. The

mesonephric mesenchyme however already has a predetermined kidney bias that was created during earlier development.

The first mesonephric tubules are found at the rostral end of the unsegmented nephric blastema and new tubules are added caudally. When the most caudal tubules are being formed, the most rostral tubules have already started to regress. In female embryos, the mesonephros regresses completely but in males the mesonephric duct is retained and contributes to genital development as the efferent ductules and the vasa deferentia (reviews Marin-Padilla, 1964; Du Bois, 1969; Tiedemann, 1976). There is no direct proof of the functional maturity of the human mesonephric nephrons and a positive view is merely based on structural findings (Silverman, 1969). In the mouse, the nephric ducts make contact with the cloaca by day 10 p.c. but do not drain into it until day 14 when the mesonephros has almost completely regressed, suggesting that the mesonephros is non-functional here in the mouse (Armstrong, 1992).

Mesonephric morphogenesis starts with the formation of mesenchymal condensates which develop into renal vesicles and **S**-shaped bodies very like metanephric development (see later). The nephric body then connects to the pronephric (Wolffian) duct; at this point the mesonephric corpuscle, the prospective glomeruli, and the proximal tubules, the distal tubules and the collecting tubules can all be distinguished at the ultrastructural level. The early condensation and **S**-shaped body development and further maturation of the mesonephric nephron are similar to the same events in metanephric development, although ultrastructural differences between the two have been observed (De Martino & Zamboni, 1966). The juxtaglomerular apparatus is missing from the mesonephros and the mesangial component is conspicuous and rich in extracellular matrix. The fenestration of the capillaries is poor in the mesonephros, but there is a compensatory abundance of capillaries. The glomerular basement membrane in the mesonephros is immature, and frequently splits into an epithelial and endothelial component and resembles the glomerular basement membrane of an

embryonic metanephric glomerulus (see reviews by Berton, 1965, De Martino & Zamboni, 1966, Tiedemann & Egerer, 1984).

1.6 The metanephros

1.6.1 The early developmental stages

The metanephros is the permanent functional kidney in higher mammals which derives from the metanephric blastema and the epithelial Wolffian duct (originally derived from mesoderm). From the most caudal part of the Wolffian duct, a bud emerges and grows dorsally towards the metanephric blastema. This occurs at day 10 p.c. in mouse and week 5 p.c. in man. The bud invades the mesenchyme and starts to proliferate and bifurcate, an event induced and regulated by the mesenchyme and local environmental factors. From *in vitro* experiments it has been shown that the first branching events are induced by the nephrogenic mesenchyme (Saxén, 1987). The metanephric mesenchyme starts to express *WT-1* at this point (Armstrong, 1992; Armstrong *et al.*, 1992). *WT-1* is (at least one of) the Wilms' tumour susceptibility gene(s) which has been shown to play a crucial role in nephrogenesis and is discussed in far greater detail later. The expression of *WT-1* in the mesenchyme can be assumed to be a marker of competence; at this stage, the mesenchyme is composed of stem cells that can undergo two different fates according to whether they are induced or not by the ureteric bud.

The very early developmental changes of the metanephric mesenchyme after invasion of the ureteric bud can be visualised when the kidney is viewed by time lapse photography (using *in vitro* culture systems) but can not be readily detected in fixed samples. After the duct has branched once, the mesenchyme has created a longitudinal condensation with indistinct borders (Saxén *et al.*, 1965; Vainio *et al.*, 1989). The primary condensate surrounds the entire ureteric epithelium creating a sheath around the duct, the maximal thickness of this layer of condensed cells is the same as the deduced transmission distance of the inductive stimulus (Saxén &

Karkinen-Jääskeläinen, 1975). When the tip of the ureter bifurcates, the solid condensate splits into two identical parts surrounding the growing tips. As the duct continues to split again and again the mesenchymal condensate continues to fractionate further, the continual splitting and packing of cells of the condensate helps lead to the formation of distinct renal vesicles.

1.6.2 The branching of the collecting duct

The ureteric bud as it invades the mesenchyme, branches several times forming a set of collecting ducts, each of which can either branch symmetrically or asymmetrically. If a duct branches symmetrically then the later branching patterns of each daughter branch will be similar, whereas if it branches asymmetrically the daughter branches will have different branching patterns later. Variation in the balance between elongation and branching determines the shape of the collecting duct and of the final kidney, such differences result in the murine kidney being unilobed whereas the human kidney is a multi-lobed structure.

The branching pattern that the ureter follows has been studied most extensively in human kidneys where elegant studies in microdissection have followed the precise pattern of divisions that occurred in several samples of developing kidneys (Osathanondh & Potter, 1963). The study categorises the branching of the ureter into four stages of branching (see figure 1.2).

During the first stage the collecting system branches symmetrically several times in succession, with branching occurring at the tips of the duct. The earliest condensations form adjacent to the duct tips and the prospective nephrons soon join the collecting system. In the second stage, the duct has undergone terminal branching and is not programmed to branch again, but still retains the ability to induce the mesenchyme to condense and form nephrons. At this point, the duct can induce the formation of further nephrons which will fuse onto the same parent duct and the older nephron shifts its point of attachment to the connection tubule of the younger

Figure 1.2

Branching pattern of the human kidney as Osathanondh & Potter, 1963.

a) Formation of nephron arcades showing how a single ureteric bud can induce as many as four nephrons to form alongside the first nephron.

b) The arrangement of human nephrons at the time of birth illustrating periods 2 and 3 (as designated by Osathanondh & Potter, 1963) of the development and branching of the collecting duct. The four deeper nephrons constitute one arcade formed during period 2, while the six cortical nephrons are directly connected to the collecting duct as is characteristic of period 3.

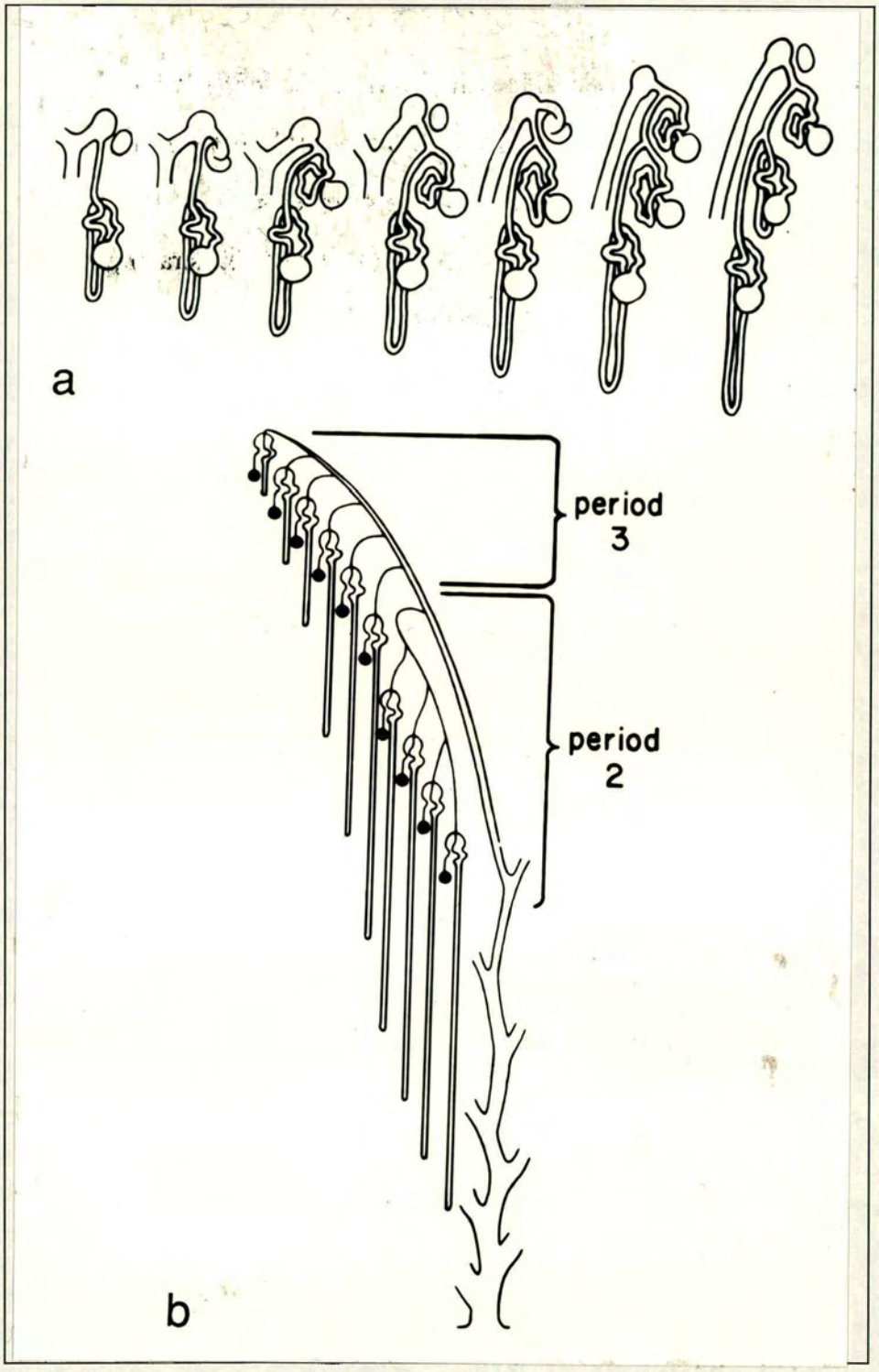


Figure 1.2

nephron. New nephrons form and repeat this pattern joining the same collecting duct until 4-7 nephric buds contribute to the collecting duct forming an arcade of nephrons. This seems to be a universal morphogenetic feature in the kidney as similar arcade structures of 3-4 nephrons have been seen in micro-dissections of the rat kidney (Neiss, 1982). In the third stage, the collecting duct grows past the attachment point of the arcade, and during this growth period will induce a set of terminal subcapsular nephrons to form, five to seven in the man and two in rat, and these nephrons join the collecting duct directly. In the last period of development terminal ampullae disappear, new nephric vesicles stop forming and the existing nephrons mature.

1.6.3 The role of the stem cells

As the kidney develops, the metanephric mesenchyme stem cells still capable of differentiating become restricted to the outer edge of the kidney, on the top of the t-shaped ureter epithelium close to the condensate. The pattern of branching situates the tips of the ureter epithelium in the upper cortex of the growing kidney separated from the capsule of the kidney by a layer of 5-15 cells of un-induced mesenchyme (still competent-stem cells). The new condensates are only induced from the mesenchyme close to the tip of the ureter and never from mesenchyme close to the stalk of the duct and only form on the medullary side of the tip, even though the condensate seems to surround all sides of tips of the branching ureter. It appears that the stem cells around the medulla produce signals stimulating growth and branching creating an exclusive region of directed growth in the cortical layer. These signals have yet to be defined. A consequence arising from the continual addition of new nephrons at the tip of the ureter while the older ones continue to develop is that from as early a stage in kidney development as 14 d.p.c., embryonic kidneys show nephrons at a wide range of different developmental stages. From 2 days later, it is possible to examine all stages from the initial mesenchymal condensations to the fully mature nephron in a single kidney.

1.6.4 Nephron development

When the metanephric blastema forms, the mesenchyme is composed of cells which are not directly attached to each other. The 'loose' cells, which express N-CAM a neural cell adhesion molecule on their surfaces (Klein *et al.*, 1988a), are surrounded by a extracellular matrix which contains types I and III collagen and fibronectin. The ureteric bud develops and grows towards and invades the mesenchymal blastema. As the ureteric bud invades the metanephric mesenchyme it initiates a complex set of reciprocal inductive interactions between the two tissues (Figure 1.3). As previously elaborated, the mesenchyme induces and regulates the branching pattern of the ureter. Experiments have shown (Saxén, 1987) that the ureter in turn induces the mesenchyme to condense and start to form pre-tubular aggregates (or renal vesicles). There are two types of mesenchyme at this point that can be clearly distinguished, the induced mesenchyme that will become epithelium and the mesenchyme that was not induced which instead becomes the stromal cell compartment. Both types of mesenchyme have already started to differentiate and neither should be considered identical to the stem cells of the immediately previous stage discussed previously. Tenascin, an extracellular matrix glycoprotein, starts to be expressed in the surrounding mesenchyme as the tubule epithelia forms; it was first thought to be implicated in epithelial-mesenchyme interactions (Chiquet-Ehrismann *et al.*, 1986) but closer examination of its expression pattern now suggests that it is a marker for the mesenchymal cells that will form the stromal cells (Aufderheide *et al.*, 1987).

As the ureteric bud enters the mesenchyme, syndecan, a cell surface proteoglycan, starts to be expressed around the invaginating bud (Vainio *et al.*, 1989) and this glycoprotein acts as a receptor for the interstitial matrix material contributing to the local aggregation of the mesenchyme. B1 and B2 chains of laminin are expressed in the non-polar mesenchyme, although the A chain required to form the main cruciform

Figure 1.3

Schematic illustration of the development of the metanephros from primary induction to the formation of the glomerulus (from Ekblom, 1992).

Loose Mesenchyme The ureteric bud invades the metanephric mesenchyme and inducing the mesenchyme to condense around it, a reciprocal induction (indicated by arrows) in which the mesenchyme induces the ureteric bud to branch.

Condensation The ureteric bud has branched and the mesenchyme is condensing around the tips of the ureter forming a cap of dense mesenchymal cells. The five cell lineages found in nephrogenesis are indicated: 1)the ureter epithelium, 2) uninduced stem cells, 3) condensed mesenchyme differentiating into epithelium, 4) uncondensed mesenchyme differentiating into interstitium, 5) blood vessels.

Comma-shape First, the condensing cells form into a renal vesicle (not shown) then a slit forms on the distal side of the vesicle forming a comma-shaped body. At this point vascularisation can be seen (arrow).

S-shape Formation of the S-shaped body by formation of a second slit. At this point the cells of the future distal and proximal tubules can be distinguished. The vesicle starts to elongate.

Tubule elongation The **S**-shaped body elongates and fuses to the collecting duct whilst endothelial cells form a capillary bed or glomerulus within the Bowman's capsule.

Podocyte folding The glomerulus and podocytes become highly convoluted. This is the final stage of nephron maturation.

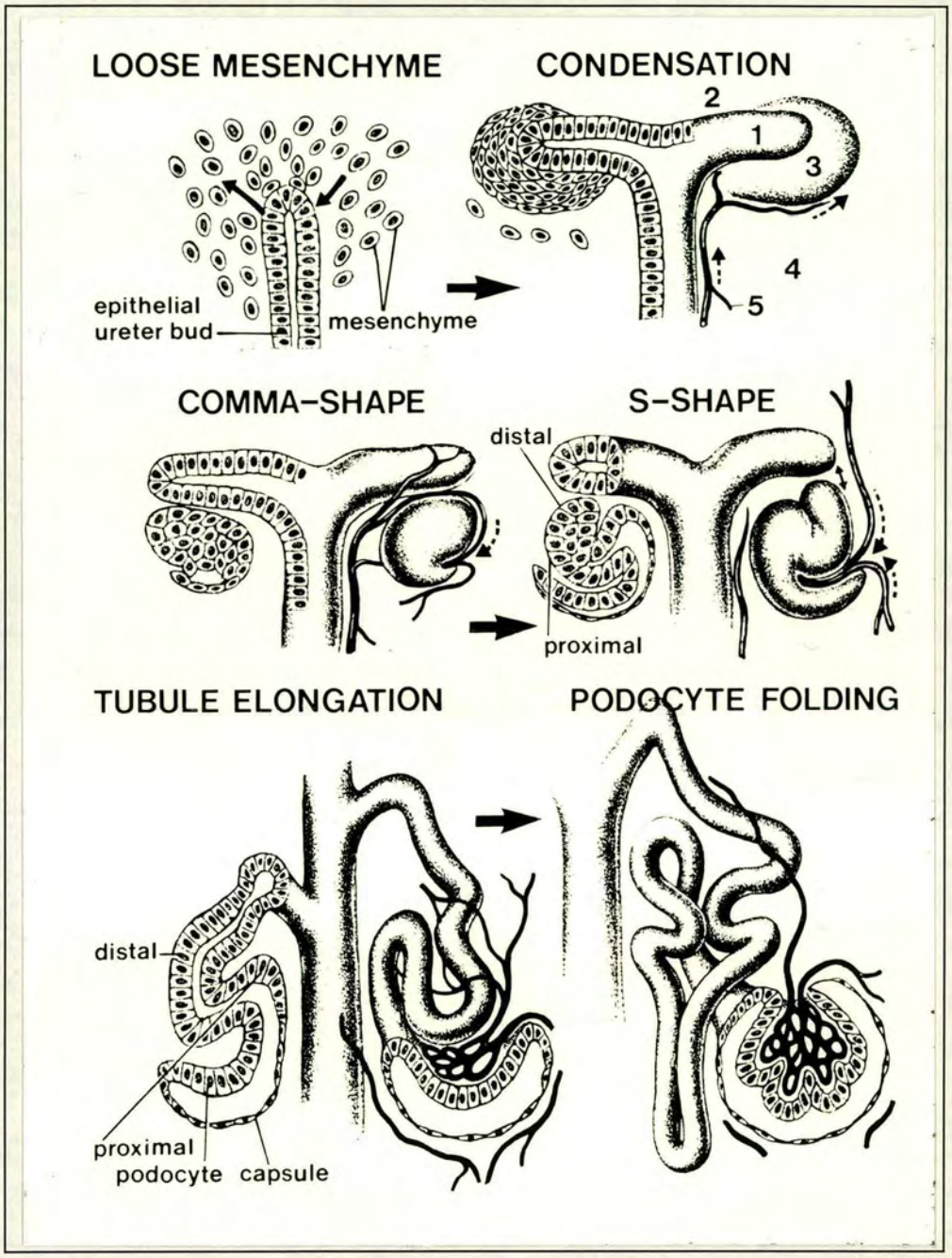

LOOSE MESENCHYME
CONDENSATION
epithelial
ureter bud
mesenchyme
1
2
3
4
5
COMMA-SHAPE
S-SHAPE
distal
proximal
TUBULE ELONGATION
PODOCYTE FOLDING
distal
proximal
podocyte
capsule

Figure 1.3

is not. A variant isoform of laminin is therefore found, a 'small' form which consists of laminin B1 and B2 chains only (Klein *et al.*, 1988b).

As the mesenchyme condenses and starts to form renal vesicles the expression of types I and III collagen and fibronectin fall (Ekblom *et al.*, 1981a). The condensing mesenchymal cells form organised condensates (renal vesicles) whose cells have a uniform ultra-structure. The cells that are trapped in the condensate move more slowly than the cells outside and it is postulated that the cells in the condensate are retarded by factors in the extracellular matrix that are directing cell behaviour (Saxén, 1987). Immunohistological study of the intermediate filaments of the cytoskeleton at the early stages of nephron development have shown that during the formation of the renal vesicle the mesenchymal cells lose their vimentin type fibres and start to express cytokeratins in a fibrillar manner (Lehtonen *et al.*, 1985). N-CAM is still expressed as the mesenchyme aggregates and starts to epithelialise. Uvomorulin then starts to be co-expressed and for a brief period both cell adhesion molecules are expressed by the condensing cells. Then, as the condensate starts to polarise and undergo vigorous proliferation, N-CAM expression disappears.

After the initial growth phase, the condensate assumes a comma shape and the first diversification of cell types is seen. The cells furthest from the collecting duct (at the proximal pole) become elongated, their nuclei change position and some cells become funnel shaped. This is the first sign of development of the glomerular crevice which divides the presumptive tubular epithelium from the presumptive Bowman's capsule (Jokelainen, 1963; Dørup & Maunsbach, 1982). The endothelial cells of the capillary system will migrate into this crevice. The condensate then starts to express type IV collagen, heparin sulphate proteoglycan and the laminin A chain, all proteins associated with epithelialization and polarization of cells. The three laminin chains can now form the laminin cruciform protein, which is associated with the formation of basement membranes.

Electron microscopy of the early renal vesicles, both in whole kidneys and in experimentally induced mesenchyme, has shown that a basement membrane accumulates around the aggregate soon after its formation and is completed by early **S**-shape phase when the membrane extends into the forming slit (Jokelainen, 1963; Abrahamson, 1985). This has been shown using antibodies against the main constituents of the basement membrane, laminin (the A chain), collagen IV and heparin sulphate proteoglycan. All of these proteins have previously been shown to be co-expressed, allowing the expression pattern of all three to be surmised by studying the expression of any one of the proteins. If laminin expression is examined, a non-uniform low expression can be seen at first, but, as the comma-shaped body starts to form expression becomes confined to the basement membrane area around the condensate. New protein is laid down during the growth and reshaping of the condensate; so that, when the **S**-shaped body has formed, a continuous layer of laminin surrounds the main body. At the distal pole of the **S**-shape the laminin is still granular suggesting an active synthesis in the developmentally younger portion of the **S**-shape (Ekblom, 1981; Ekblom *et al.*, 1990). This type of result suggests that the basement membrane may act as a scaffold for attachment of the epithelial cells allowing polarisation and shaping of the cells (Ekblom, 1984; Hay, 1984).

The aggregate undergoes cell growth, a second slit is gradually formed at the opposite (distal) pole, and the tubular condensate assumes the **S**-shape before it joins the collecting duct. At this stage the ultra-structure of the **S**-shaped body starts to diversify, the presumptive distal and proximal tubules, the visceral epithelium (the presumptive podocytes) and the parietal epithelium of the glomerulus can all be distinguished from each other by their lectin-binding capacity. More diversity has been shown in an immunohistological study of the intermediate filaments of the cytoskeleton which showed that there is cell heterogeneity at the late **S**-shape stage as

both the cells of the presumptive Bowman's capsule and the presumptive podocytes are devoid of cytokeratin (Lehtonen *et al.*, 1985).

Morphologically, the cells of the visceral layer of the S-shaped body are more columnar than those of the parietal layer, and have been shown to have a different lectin-binding capacity (Ekblom *et al.*, 1981b). The visceral layer binds *Maclura pomifer* agglutin, wheatgerm agglutin and peanut agglutin, whilst the rest of the nephron does not. The visceral layer cells start as a typical polarised epithelial layer with apical occluding and adhering junctions. As the glomerulus matures, the junctions migrate down the cell surface and gradually disappear allowing the visceral cells to form the podocyte layer. The intercellular spaces become open and bridged together by slit diaphragms that form a continuous 3-dimensional structure attaching the podocytes to each other. These podocytes give rise to secondary processes, pedicels or foot processes, that also form part of the 3-dimensional network. The podocytes express ZO-1, a protein associated with tight junctions, at a higher level than any other cells of the nephron suggesting that tight junctions play a role in connecting the cells of the podocytes (Schnabel *et al.,* 1990).

A basal lamina, the glomerular basement epithelium (GBM) separates the podocytes from the endothelial cells; this membrane provides the main filtration barrier of the kidney. When murine metanephric mesenchyme is cultured on quail chorio-allantoic membrane, the GBM is found to be derived from both the epithelium and the endothelium indicating that the GBM initially forms by fusion of an endothelial and epithelial cells (Sariola *et al.,* 1984).

Finally, the centre of the **S**-shaped body elongates and forms the long convoluted tubules of the distal and proximal tubules and the loop of Henlé becomes extremely convoluted. At the same time, the epithelium of the glomerulus becomes highly folded and the cells of the presumptive Bowman's capsule flatten (Dørup & Maunsbach,

1982). The nephrons will have fully formed before birth occurs but the tubules continue to mature and the cells that control filtration postnatally.

1.6.5 Vascularisation of the metanephros

The adult kidney comprises 1% of the total body weight and approximately 20-25% of the cardiac output flows through the kidney indicating just how well vascularised the kidney is. The endothelial cells of the vascular system comprise the third cell type of the metanephros. These cells form a capillary bed, the glomerulus, within the Bowman's capsule, and together these form the renal corpuscle. The development of blood vessels in the metanephros and vascularisation of the nephron is of major importance in nephrogenesis. The first sign of vascularisation can be seen just as the ureteric bud has branched for the first time and the first ureteric branches have formed (Ekblom, 1981).

For a long time the origin of the vascularisation of the nephron was controversial, but Potter (1965) suggested, the now widely-held belief, that the blood vessels originate from outgrowths from the general vascular system. It had been shown indirectly that when early embryonic kidney rudiments were grown in culture and the kidney developed normally even though there was no endothelial component. More direct proof was obtained when Sariola *et al.* (1984) cultured early embryonic kidney rudiments on the vascular quail chorio-allantoic membrane. The resulting nephron showed proper vascularisation and development of the endothelial cells of the glomeruli. Quail cells can be distinguished from murine cells by Feulgen staining of the distinctive quail nucleolar structure (Le Douarin, 1973), allowing Sariola *et al.*, (1984) to demonstrate that the vascular component of the nephron had arisen from the quail tissue. It was concluded that the capillaries were derived from the host tissue and thus had an external origin (Ekblom *et al.,* 1982). The kidney is thought to stimulate and guide nephron vascularisation. The development of blood vessels is a major event and is initiated very early in nephrogenesis being found as soon as the

first branches of the ureter have formed (Ekblom, 1981). The development of the vessels is dependent on the formation of epithelial sheets which apparently secrete angiogenesis factors that guide the vessels to the appropriate end point (Ekblom, 1981, Ekblom, 1992). Recently, a soluble heparin-binding growth factor of 18 kDa has been purified from murine embryonic kidneys that stimulates the proliferation of capillary endothelial cells and induces angiogenesis (Risau & Ekblom, 1986). Further study will tell if this factor is important in the vascularisation of the kidney.

1.6.6 In Vitro *organ culture of the embryonic kidney.*

One of the most advantageous points in studying kidney development is the well-defined *in vitro* organ culture that allows early kidney rudiments to be dissected from the embryo and to develop and differentiate in culture. The culture system allows nephrogenesis and the requirement of various factors to be studied closely. Embryonic kidneys can also be cultured in other localisations of the embryo or *in vitro*. Mouse kidneys are most commonly used for culture but other species have been used including human, rat and chick kidneys (Crocker, 1973; Saxén, 1983).

Two types of *in vitro* culture systems have been used, the whole-organ and the trans-filter culture; both of which were developed by Grobstein (1953a, 1953b, 1955, 1956) and subsequently modified and improved (Saxén, 1987).

In whole organ culture, 11.5 d.p.c. embryonic kidneys are cultured intact with their ureters and placed on porous filters kept at the air-liquid interface of a dish containing the appropriate media. Surface tension keeps the tissue covered with a thin layer of media. Kidney rudiments survive well under these conditions for up to 12 days (Ekblom, 1992) but good differentiation of the kidney can normally be seen after 6 days. When kidneys grown under these conditions are examined, well advanced morphogenesis and tubulogenesis is apparent.

In the trans-filter culture system, the un-induced metanephric rudiments are dissected from 11 d.p.c. embryos and the mesenchyme and the ureteric bud separated

either enzymatically or mechanically. The mesenchyme is placed on the upper side of a Nucleopore filter to which an inducer (e.g. neural tube fragments) is fixed to the lower side using agarose. The filter is then cultured in similar conditions as the whole organ system. Culturing the mesenchyme with heterologous inducers next to it has allowed the induction to be studied. Other tissues have been shown to be able to induce the mesenchyme to form nephrons, but no other mesenchyme can be induced to produce nephrons (Saxén, 1970). The conclusion of these experiments was that the metanephric mesenchyme has a pre-determined kidney bias and that the induction is permissive rather than directive (reviewed by Gurdon, 1987; Saxén, 1977). When used in direct contact with the metanephric mesenchyme salivary gland, jaw mesenchyme, rostral somites, neural teratomas, brain and spinal cord can act as inducers but not heart, liver and lung (Sariola *et al.*, 1989). When trans-filter culture was used, only spinal cord, forebrain and midbrain could induce differentiation, and the ureteric bud, the natural inducer, could not induce differentiation under trans-filter conditions.

When successfully induced cultures were studied, close cytoplasmic contacts between the neural tissue and the mesenchyme could be seen. The neurofilaments take 16-20 hours to invade the mesenchyme, and not unsurprisingly this is the minimum period of time that the neural tissue must be left in contact with the mesenchyme to allow differentiation. After that point the inducer can be removed without affecting mesenchyme differentiation. Such experiments suggest that neurons play an important role in nephrogenesis. Under trans-filter culture conditions, normal three dimensional histogenesis of the nephron occurs that is comparable to the events that occur *in vivo* including development of proximal tubules with well-defined brush borders, distal tubules, renal-corpuscle-like bodies and podocytes but no endothelial cells develop (Ekblom *et al.*, 1980, 1981a).

1.6.7 Innervation of the metanephros

As discussed previously, nerves are important in kidney development and seem to be required for the induction of tubular differentiation. The exact role that nerves play in the differentiation has yet to be well understood but the absolute requirement in the process has been proved. Preliminary studies using antibodies against nerves prevented kidney morphogenesis, showing that the neural involvement was essential (Sariola *et al.,* 1989). Further studies on neural development during nephrogenesis, have used two sets of antibodies that reveal neuron cell bodies and neurites or neurofilaments. Whole-mount cultures had to be used, as cut sections do not allow the long cell processes to be studied. Early in nephrogenesis, nerves can be seen around the periphery of the rudiment and ganglion beside the Wolffian duct, but very few around the tip of the ureteric bud. The nerves then grow into the stroma and into the condensates that will form the secretory tubules. The invasion of the nerves preceded angiogenesis and morphological differentiation of the metanephric mesenchyme. Nerve growth factor receptor (NGFR) has also been shown to be transiently expressed by the embryonic kidney but to disappear upon terminal differentiation. Antisense oligonucleotides to NGFR mRNAs inhibit the gene expression and prevent kidney morphogenesis and epithelial tubule formation (Sariola *et al.*, 1991). The innervation and angiogenesis may be temporally related but *in vitro* cultures have shown that they are clearly independent and separable developmental events.

1.7 The molecular basis of nephrogenesis

1.7.1 Introduction

The cellular events that occur during kidney development have just been described. The aim of this project has been to find un-characterized genes that will help to further the understanding of nephrogenesis at the molecular level. Therefore it is helpful to consider the types of genes that have been shown to have pivotal roles in

other major developmental fields. From this point we can consider the subset of these genes that have already been identified as being expressed in the developing kidney, their general role in embryogenesis and their specific role in nephrogenesis.

There are several types of genes that are important in defining cellular identity and regulating the important events that must take place in the course of morphogenesis. The first are the genes for extracellular signals such as neurotrophic factors, cytokines, hormones, the extracellular matrix and cell-cell adhesion molecules that the cells need to give signals to and to receive signals from the neighbouring or distant cells. The second class of genes implicated are for the receptors that respond to extracellular signals and internalise them allowing the cell to respond. These include receptor tyrosine kinases (RTKs), G-protein coupled receptors and gated ion channels. The third type of genes are responsible for the second messenger pathways that allow the cell to respond to the signals that the receptors transmit, and are likely to include the genes for kinases, phosphatases, phospholipases, *ras* and ras like proteins, G-proteins and calmodulin. The nuclear proteins, transcription factors and DNA-binding proteins are also likely to be important, allowing the proper transcriptional response. These tissue-specific transcription factors are likely to include the proteins like the HOX and *Pax* gene family, the POU proteins, zinc-finger proteins and nuclear binding proteins such as *myc* and *max*.

1.7.2 Growth factors

The Insulin like growth factors (IGF) I and II are expressed in the fetal kidney and anti-IGF I or II antibodies prevent the growth of the metanephros in culture (Rogers *et al.*, 1991). IGF II stimulates the growth and differentiation of kidney explants and is a target for WT-1 repression (discussed later). Therefore the IGFs seem to be absolutely necessary for kidney development. The receptors for IGF I and II are present in the glomeruli and proximal tubule of the developing kidney and stop being expressed after birth in rodents but are maintained in human kidneys after birth

(reviewed by Hammermann *et al.*, 1993). Hepatocyte Growth Factor (HGF or scatter factor) has also been shown to play a role in nephrogenesis. HGF is the ligand for the *c-met* RTK (Naldini *et al.*, 1991a, b) and both are expressed in the fetal kidney. The interaction between the two proteins is thought to play an important role in mesenchyme epithelial interactions during nephrogenesis (Sonnenberg *et al.*, 1993). Nerve growth factor (NGF) has also been implicated as playing a vital role in kidney development. Antisense oligonucleotides to the NGF Receptor mRNA inhibits the expression of the receptor and prevent kidney morphogenesis and epithelial tubule formation (Sariola *et al.*, 1991) as previously described. This would imply that NGF may play an important role in nephrogenesis.

1.7.3 Receptor tyrosine kinases

The receptor (-like) tyrosine kinases (RTKs), are a family of proteins that play a major role in specifying cell fates in the early embryo and later in development as the tissues form (reviewed in Pawson & Bernstein, 1990; Egan & Weinberg, 1993). The RTKs have an extracellular, ligand-binding domain, a trans-membrane domain, and an intra-cellular tyrosine kinase domain. It is thought that the super-family has evolved from gene duplication and divergence permitting increasingly complex cellular interactions. The RTKs fall into four classes based on the structure of their extracellular domains. Class I includes the human and *Drosphila* epidermal growth factor receptors, whose extracellular domain contains two or three cysteine-rich domains. There is strict conservation of the spacing of the cysteine residues and the interspersed regions. Class II receptors includes the insulin receptor and IGF-I receptor, and are thought to function as heterotetrameric structures with 2 α and 2 β subunits connected by disulphide bonds. The α subunits contribute to the formation of the ligand binding domain whilst the β subunits pass through the membrane and carry the kinase domains. There are also class III receptors, the *c-kit*-like family characterized by immunoglobulin-like repeats, and class IV receptors containing

fibronectin-type-III repeats. The ligands for these receptors where known are a variety of growth factors and hormones, but many have not had their ligand identified. The universal feature of RTKs with known ligands is that their tyrosine kinase activity is physiologically dependent on the ligand binding. The metabolic and mitogenic effect of the ligand is mediated through the tyrosine kinase activity of their receptor probably through specific phosphorylation of cytosolic proteins. There is currently evidence to suggest that the RTKs are important in controlling cellular differentiation and cell-cell interactions, particularly during embryogenesis. A number of RTKs are only expressed in specific cell types, suggesting they may have a tissue-specific function. Some of the genes encoding the RTKs are highly conserved in evolution and are very similar where homologies have been identified, for instance between *Drosophila* and humans.

The receptor tyrosine kinase family plays a major role in early kidney development. *c-met* (Chan *et al.*, 1988), *c-ret* (Tahira *et al.*, 1988; Takahashi *et al.*, 1988, Schuchardt *et al.*, 1992, Pachnis *et al.*, 1993) and *c-ros* (Tessarollo *et al.,* 1992) have all been shown to be expressed during embryonic kidney development as well as in other areas including the nervous system and enteric system. The human homologue of *c-ret* has been shown to be mutated in MEN-2A (multiple endocrine neoplasia-2A) [Mulligan *et al.*, 1993] which has a range of symptoms including dismorphic kidneys. The murine c-*ret* has been mutated to give a truncated non-functional product with no trans-membrane region and reintroduced into transgenic mice. The homozygous offspring carrying this mutation are non-viable. The lack of viability is probably due to the absolute requirement of the gene product in its three main areas of expression: the nervous system, the excretory system and the enteric system (Pachnis *et al.*, 1993). Looking at the kidneys in new-born mutant mice, there is either complete agenesis or severe hypoplastic dysplasia of the kidneys and the mice die within a few days of birth. *c-ret* is normally expressed in the pronephros, mesonephros, the nephric duct

and the ureteric bud. Its predicted protein structure and the consequences of mutation suggest that it is involved in the signal transduction between the ureteric bud and the metanephric mesenchyme, and the mutation that interrupts the signal pathway disrupts the interaction between the two tissues.

The mammalian *c-ros* gene encodes an unusually large (2438 aa) protein with a large (1862 aa) extracellular domain. It is the murine homologue of the *Drosophila* tyrosine-kinase receptor, *sevenless* (Hafen *et al.*, 1987). The sevenless protein is a trans-membrane protein tyrosine-kinase that is absolutely required for the development of the *Drosophila* eye. It acts at the beginning of a signal transduction cascade transferring information from the membrane to the nucleus in a pathway that has almost been fully elucidated. It is sensible to consider that most RTKs act in similar signal transduction pathway.

The *sevenless* ligand is the *boss* (*bride of sevenless*) protein (Hart *et al.*, 1990), a trans-membrane G protein-coupled receptor suggesting that activation of both the sevenless and boss receptors requires direct cell-cell interaction and that the *boss*-expressing cell responds through the G-protein second messenger pathway. *sevenless* is now known to interact with *drk* (*downstream of receptor kinase*) which binds to *sos* (*son of sevenless*) and enters the *ras* second messenger pathway [Simon *et al.*, 1993; Olivier *et al.*, 1993).

c-ros expression is coincident with the proliferation of epithelium. It is expressed at the terminal ends of the branching ureteric tubules and expression disappears as the primary induced epithelium differentiates into the comma-shaped body. It is involved in the obligate cell-cell interactions of the induction and proliferation of the epithelial cells of the kidney and work on the *Drosophila* homologues suggests it is involved in reciprocal signalling between two cell types. As deletion of the *c-ret* gene is not compensated for by *c-ros, c-ros* and *c-ret* must act independently without functional redundancy.

The hepatocyte growth factor (HGF)/scatter factor has been shown to be the ligand for the *c-met* receptor (Bottaro *et al.,* 1991; Naldini *et al.*, 1991a, b). *c-met* is expressed in the epithelial cells of various developing organs, including the kidney, whilst HGF is expressed in the mesenchymal cells close to *c-met* expressing cells. *c-met* is found in all epithelia of the developing kidney except for glomerular epithelium and HGF is expressed only in the induced mesenchyme (Sonnenberg *et al.*, 1993). Therefore the interaction of the receptor and ligand has been implicated as a signal exchange required for nephrogenesis to occur.

1.7.4 Homeobox genes

The study of *Drosophila* homeotic mutants that change the developmental fate of one tissue or structure into that of another has led to the isolation, cloning and sequencing of a large number of genes encoding nuclear proteins, putative transcription factors that have interesting expression patterns. A number of these genes were found to map together in two large gene clusters, the Bithorax complex (Bx-C) and the Antennapedia complex (ANT-C). The map position of a mutation within the Bx-C cluster is reflected in the position along the body axis at which the mutation had its effect (Lewis, 1978) and the expression domains (Harding *et al.*, 1985). The two gene families arose from gene duplication and amplification of a common ancestral gene, and then divergence of the genes and their functions. The chromosomal location of each of these genes reflects the expression boundaries along the anterior posterior axis of the embryo, suggesting a progressional modification of expression patterns as the genes diverged.

Sequence analysis of the genes in the two clusters identified a common 180 base pair consensus sequence, the homeobox, encoding a basic 60 amino acid motif found in the members of these gene families and of other important developmental *Drosophila* genes (McGinnis *et al.*, 1984). The homeobox is found in three types of developmental genes, the maternal effect genes (*e.g. bicoid)*, the segmentation genes

(*e.g. fushi tarazu)* and the homeotic genes (*e.g. Antp)* The homeobox is generally found in the 3' end of genes and encodes a peptide with a helix-turn-helix motif. The peptide has homology to a region the *cro* and repressor proteins of bacteriophage λ, and the CAP protein of *E. coli* all of which act as DNA-binding domains, and so by inference the homeobox was proposed to be a DNA-binding motif.

It is now known that homeobox containing genes can encode transcription factors that can autoregulate their own expression (Zappavigna *et al.*, 1991, Hoey & Levine, 1988, Kuziora & McGinnis, 1989) and regulate the expression of other homeobox proteins as well as other not yet identified targets. Homeobox genes have been found in many diverse species now including *Xenopus* (Xhox), chick (Ghox) human (HOX) and *C. elegans*. Many genes have been found in mouse by virtue of homology to the homeoboxes and most show similarity to the *Antp* homeobox. The clustered mouse genes also are expressed in domains that reflect their relative position within the cluster (Duboule & Dolle, 1989). In the mouse there are 4 large clusters of genes, all of which span more than 100kb. These are the Hox A [previously Hox-1] cluster on chromosome 6, Hox B [previously Hox-2] on chromosome 11, Hox C [previously Hox-3] on chromosome 15 and Hox D [identified as Hox-5 then called Hox-4] on chromosome 2.

Several murine genes have been found with homology to more divergent *Drosophila* homeobox sequences: homologues of the *engrailed* genes- (*En-1* and *En-2;* Joyner & Martin, 1987), homologues of *even-skipped* (*Evx-1* and *Evx-2;* Bastian & Gruss, 1990) and homologues of the *caudal* box (*Cdx-1;* Duprey *et al.*, 1988). Other families of genes containing divergent homeobox sequences are the homologues of the muscle specific box (*msh*) *Hox 7, 8* and *9,* recently renamed *Msx-1, 2* and *3* (Hill *et al.*, 1989: Robert *et al.*, 1989); and the recently cloned *Otx* genes, homologues of the *orthodenticle* box (Simeone *et al.*, 1992a, 1993); and the *Emv* genes, *empty spiracles* homologues (Simeone *et al.*, 1992b).

A number of *Antp*-like homeobox genes are expressed in the developing kidney (reviewed by Holland & Hogan, 1988a); some are expressed in a non-specific manner whilst others have a restricted pattern of expression. *Hox B5 (2.1)* is highly expressed in the epithelial tubules of the developing kidney (Holland & Hogan, 1988b) whilst *Hox C9 (3.2)* is expressed in the fetal kidney, mainly in the cortex, and probably in the ureter or genital ridge (Erselius *et al.*, 1990). *Hox B7 (2.3)* is expressed in the early stages of ureter formation and maintained in all duct derivatives (Kress *et al.*, 1990). *Hox A3 (1.5), C9 (3.2)* and *6.1* are all expressed in mesonephros and *Hox A3, A5, C9* and *6.1* in the metanephros. The mouse homeobox genes LFB1 (HNF-1/HNF-1α/APF/HP) and LFB3 (vHNF-1/HNF-1β) play a critical role in the conversion of the nephric mesenchyme into polarised epithelium (Lazzaro *et al.*, 1992).

The homeobox gene *Hox C9 (Hox 3.2)* is the most 5' member of the Hox complex on chromosome 15 and is expressed in the posterior of 8.5-15.5 d.p.c. embryos. The expression is seen in the neural tube, posterior pre-vertebrae and the hindlimb buds. It shows a regionalised expression pattern in the developing kidney and is also expressed in the adult kidney. It is expressed in the developing kidney starting at day 12.5 and by day 14.5-16.5 the expression is mainly in the cortex, with reduced expression in the stromal mesenchyme. Strong expression was also seen in either the ureter or genital ridge (but it was unable to be discerned which tissue) (Erselius *et al.*, 1990). Generally *Hox C9* is expressed in mesoderm and ectoderm derived structures but not endoderm. Its paralogs, *Hox B9* and *D9* are expressed in the mesonephros and metanephros and *Hox A9* is also expressed in the fetal kidney.

Hox B5 is expressed in an anteroposterior restricted domain extending from the hindbrain through the length of the spinal cord, predominantly in the dorsal region. It is expressed in the spinal cord (occipital and cervical regions) and in the dorsal root ganglia, lung, stomach, mesonephros and the metanephros (Jackson *et al.*, 1985). It is strongly expressed in all the mesonephric tubules identified and initially has general

expression in the metanephros (12 d.p.c.). By day 16 p.c. *Hox B5* expression is localized to the epithelial cells of the developing tubules and collecting ducts.

The *Hox B7* (Deschamps *et al.*, 1987) expression pattern was first studied in detail using a *lac-Z* fusion transgene. The gene is expressed in the spinal cord, the somitic lateral plate (the stomach primordium), mesonephric mesoderm and at a lower level in the spinal ganglia. The gene was also very highly expressed in the mesonephric duct and tubule epithelium. By 12.5 d.p.c., the mesonephric duct and the metanephros show strong expression whilst the degenerating mesonephric tubules show a lowering of expression. Transcripts are localised to the epithelial tubules, the ureter and ureteric tree. By 14.5 d.p.c., there is high expression only in the collecting tubules which persists and can be detected in the adult kidney (Kress *et al.*, 1990).

Both the homeobox genes LFB1 and LFB3 have been isolated in a number of different contexts and hence are known by at least three separate names each. LFB1 was originally cloned as HNF-1, hepatic nuclear factor-1 which was cloned as a homeobox gene encoding a protein involved in the transcriptional regulation of several liver-specific genes (Courtois *et al.*, 1987). The gene encodes a distinct type of homeodomain, characterized by 21 extra amino acids. LFB3 is another member of the LFB family, also known as vHNF-1 (Rey-Campos *et al.*, 1991) or HNF-1β (Mendel *et al.*, 1991). It encodes a transcriptional activator with an enlarged homeodomain that forms heterodimers with LFB1 both *in vitro* and *in vivo* (Mendel *et al.*, 1991). The two proteins are expressed in liver, kidney and the digestive tract. All of these tissues develop a specialised polarised epithelium, where these proteins have been shown by *in situ* hybridization to be expressed (De Simone *et al.*, 1991; Ott *et al.*, 1991). The expression pattern of the two genes has been studied in detail in the developing rat kidney. LFB3 is expressed in mesenchymal condensations as soon as they are induced to differentiate (Lazzaro *et al.*, 1992). After the nephron has formed, it is expressed in the proximal and distal tubules and in the collecting ducts, but not in

the glomerulus. Its expression is maintained after birth. It is thought LFB3 may be involved in the early induction stages of nephrogenesis. LFB1 is expressed later in the post-inductory period again in the proximal and distal tubules but not in the collecting ducts or the glomerulus. The two genes are hypothesized to be involved in supplying the information to start differentiation and maintain to stay differentiation.

1.7.5 Pax genes

The *Pax* genes are the mammalian homologues of the *Drosophila paired* genes. The *paired* gene contains a homeobox distinctive to the gene family, as well as another motif called the *paired* box (Bopp *et al.*, 1986). Both peptides are predicted to fold into helix-turn-helix structures and to be DNA-binding domains. The paired box is conserved between *Drosophila* and vertebrates. Eight murine genes, *Pax 1-8*, constituting the *Pax* gene family, have now been found by homology with the *paired* box, all contain the *paired* box but only three (*Pax 3*, *6* and *7*) also contain a *paired* type homeobox. The *Pax* genes have recently been shown to be responsible for several classical mouse and human mutations. *Pax-1* is mutated in *undulated* mouse (Balling *et al.*, 1988), *Pax-3* mutations cause the *splotch* mutation in mouse (Epstein *et al.*, 1993; Goulding *et al.,* 1993) and Waardenburg's syndrome in humans (Tassabehji *et al.*, 1992), whereas *Pax-6* mutation causes small eye (*Sey*) in mouse (Hill *et al.,* 1991) and aniridia in humans (Ton *et al.*, 1991).

Pax-2 and *Pax-8*, both of which contain the paired box but not the *paired* type homeobox, are expressed in the developing kidney as well as other regions of the embryo. *Pax-2* is expressed in the kidney as well as the developing central nervous system. *Pax-2* is expressed in the pro-nephric tubules and extending nephric duct at day 9. The mesonephros also expresses *Pax-2* but in a non-specific manner. In the metanephros, *Pax-2* is expressed in the nephric duct, the ureteric bud, the condensing mesenchyme and the epithelial derivatives of the condensed mesenchyme (Dressler *et al.*, 1990). There is strong expression in the **S**-shaped body which is down-regulated

as the epithelium differentiates. *Pax 2* is still actively expressed in the perimeter of the growing kidney; the more mature renal tubules in the interior express the gene less and expression is not found at all in the adult kidney.

While mutation of the *Pax-2* gene has not been implicated in any classic mouse developmental mutants, insertion of a transgene which deleted *Pax-2*, produced mice which exhibited unilateral renal agenesis (Dr. M. Meisler, Mouse Genome Mapping Conference, Buffalo 1992). Constitutive expression of the *Pax-2* gene from the human cytomegalovirus (CMV) promoter in transgenic mice causes severe kidney abnormalities, suggesting that *Pax-2* must be repressed for terminal differentiation of renal epithelium and normal kidney development (Dressler *et al.*, 1993). The transgenic mice have abnormal and dysfunctional renal epithelium and show symptoms similar to human congenital nephrotic syndrome.

Conversely, the use of anti-sense oligonucleotides to prevent the expression of *Pax-2* in *in vitro* culture systems also causes kidney defects. Ureter branching is inhibited as is the differentiation of mesenchymal condensates to epithelial cells (Rothenpieler & Dressler, 1992, 1993). This indicates that *Pax-2* is needed for the initiation of kidney tubule morphogenesis, probably in the information transduction pathway during cell polarization and cell condensation. But for later kidney development *Pax-2* must be repressed, which probably involves *WT-1* expression. The expression of Pax-2 has been studied in human Wilms' tumours, where *WT-1* expression has been lost, for instance by deletion of both alleles. *Pax-2* expression persists at an elevated level in the undifferentiated epithelium of Wilms' tumours (Dressler & Douglass, 1992; Eccles *et al.*, 1992).

Pax-2 expression has also been studied in mice with deletions of the *WT-1* gene (discussed later) and although normal *Pax-2* expression can be seen in the Wolffian duct there is no expression in the metanephric mesenchyme. These observations suggest that *WT-1* is required both for the activation of expression of *Pax-2* in the

metanephros and for the repression of *Pax-2* expression at the later stages (Dressler *et al.*, 1992).

Pax-8 is expressed in the kidney and thyroid gland (Plachov *et al.*, 1990). *In situ* hybridization analysis shows that *Pax-8* is expressed in the mesenchymal condensations and the resulting epithelial structures, the **S**-shaped bodies. *Pax-8* maps to mouse chromosome 2 in the same region as the mouse mutant Danforth's Short-tail (*Sd*), a mutant which has abnormal kidney development (Dunn *et al.*, 1940; Glueckson-Schoenheimer, 1945). There are opposing views on the possibility of *Pax-8* being a candidate gene for this mutant on the basis of chromosomal location (Koseki *et al.*, 1993) and expression pattern of the *Pax-8* gene.

1.7.6 Zinc-finger genes

The zinc-finger proteins have also been shown to play an important role in development. The motif was originally recognised in *Xenopus laevis* in the RNA Polymerase III transcription factor TFIIIA (Miller *et al.*, 1985). Each finger contains two invariant pairs of cysteines and hystidines that co-ordinate the zinc ion. They fall into two classes: the Cys_2-His_2 class, of which TFIIIa is the prototype, and the Cys_2-Cys_2 class which is characteristic of the receptor family of zinc-fingers. The fingers act in a modular fashion and the individual fingers do not interact precisely with adjacent nucleotides but rather clusters of fingers bind to three functionally important regions of the binding site. The number of zinc-fingers can vary from as little as two to as many as thirty. Zinc-finger proteins bind specifically to DNA and both the zinc-fingers and zinc are necessary for the protein to bind to DNA and act as a transcriptional activator.

Perhaps the best known of the zinc-fingers genes is *WT-1*, found as a candidate gene for Wilms' tumour, a childhood cancer of the kidney caused by persistence of stem cells in the kidney leading to de-regulated growth. The gene was identified by reverse genetics (positional cloning) and when sequenced was found to contain four

zinc-fingers and hence proposed to be a DNA-binding protein (Call *et al.*, 1990; Rose *et al.*, 1990). Looking at the gene in a sporadic unilateral Wilms' tumour, the gene was found to have an internal deletion of the gene suggesting that it may be a tumour-suppressor gene whose removal is responsible for the tumorigenesis (Haber *et al.*, 1990).

In situ hybridization analysis has shown that the gene is expressed in a wide variety of mesothelial tissues during embryogenesis including the heart, liver, mesonephros, metanephros and the testes at a high level of expression (Pritchard-Jones *et al.*, 1990; Armstrong, 1992; Armstrong *et al.*, 1992). The *WT-1* transcripts are alternatively spliced at two sites, one splicing variant includes a separate exon encoding 17 amino acids, that are inserted between the proline-rich amino terminus and the zinc-finger domains (Haber *et al.*, 1991) but there seems to be no functional difference between the proteins containing this peptide or not. The second alternate splice inserts in three extra amino acids, lysine, threonine and serine (+KTS), between zinc-finger 3 and 4 (Haber *et al.*, 1991). RNAase protection analysis demonstrates that the most prevalent splice variant in both human and mouse contains both alternative splices, whereas the least common transcript is missing both splices (Bickmore *et al.*, 1992). The +KTS and -KTS WT1 protein variants bind to DNA with different sequence specificities. Therefore, alternative splicing within the DNA binding domain of a transcription factor can generate proteins with distinct DNA binding specificities and probably different physiological targets.

The *WT-1* gene binds to the EGR1 (early growth response genes) consensus DNA binding site (Rauscher *et al.*, 1990; Bickmore *et al.*, 1992). A mutation in a zinc-finger first found in a Wilms' tumour abolishes the DNA-binding activity suggesting that the loss of ability to bind DNA may be responsible for the Wilms' tumour. *WT-1* has also been shown to interact with the p53 gene, a tumour-suppressor gene with ubiquitous expression, in transfected cells (Maheswaran *et al.*, 1993). The two

proteins can interact changing their ability to transactivate their respective targets. In the absence of p53, the WT1 protein is a potent transcriptional activator of the early growth response gene 1 (EGR1) site, rather than a transcriptional repressor (Madden *et al.*, 1991). In contrast, the WT1 protein exerts a co-operative effect on the p53 protein enhancing its ability to transactivate the muscle creatine kinase promoter (Maheswaran *et al.*, 1993). The *WT-1* gene is also mutated in Denys-Drash syndrome, another human condition in which severe urogenital aberrations result in renal failure, pseudohermaphroditism, and Wilms' tumour (nephroblastoma) (Pelletier *et al.*, 1991). The WT-1 protein is a transcriptional repressor of the IGF II gene (Drummond *et al.*, 1992; Werner *et al.*, 1993) and the platelet-derived growth factor A-chain (PDGFA) gene (Wang *et al.*, 1992; Gashler *et al.*, 1992) both of which have been shown to have highly elevated levels of expression in Wilms' tumours. Further studies have suggested that the WT-1 protein has domains that can both act as a transcriptional activator and repressor (Wang *et al.*, 1993).

A transgenic mouse has recently been made with a mutation in the WT-1 gene (Kriedberg *et al.*, 1993). The mutation, when homozygous, results in embryonic lethality between day 13 and 15 p.c. but the heterozygote is not affected. The lethality is thought to be due to defects in the heart causing cardiac dysfunction and hemodynamic failure and ultimately death. The metanephros fails to form and no ureteric bud can be found at the appropriate stages. Although the metanephric blastema can be found, when this is used in the *in vitro* culture system the mesenchyme can not be induced even when the inducer is spinal cord, the most effective inducer (Kriedberg *et al.*, 1993).

Looking at *Pax-2* gene expression in these mice, the gene is expressed normally in the Wolffian duct but not at all in the mutant metanephric mesenchyme. This suggests that the WT-1 protein is required for *Pax-2* gene expression in the blastema. It has already been speculated that the WT-1 protein represses *Pax-2* gene expression as

WT-1 has been shown to be a transcriptional repressor (Drummond *et al.*, 1991) and binds to a EGR-like sequence (Early growth response gene) (Rauscher *et al.*, 1990, Bickmore *et al.*, 1992) copies of which have been found in regions upstream of the *Pax-2* gene (Dressler and Douglass, 1992). But this is not the only important role hypothesised for WT-1, as in Wilms' tumours there is also increased levels of *N-myc* (Nisen *et al.*, 1986). Hence it might be postulated that WT-*1* is acting in the kidney as a repressor of DNA-binding proteins (*Myc* and *Pax-2*), the new evidence suggested by the mutant also suggests that the WT1 protein is necessary early in metanephric development to allow the transcription of other key genes in nephrogenesis. Therefore presence of the WT-1 protein in the mesenchyme can be assumed to be a marker for the competence of the mesenchyme to be induced.

1.7.7 Other nuclear genes implicated in nephrogenesis

Other nuclear proteins expressed in the developing fetal kidney include the myc family and the products of the limb deformity (*ld*) gene. The *myc* family of proto-oncogenes are presumed to be master regulators of cell growth and differentiation (reviewed by Cole, 1986). All three myc genes, *N-myc*, *c-myc*, and *L-myc* are expressed in the developing kidney with distinct expression patterns (Mugrauer & Ekblom, 1991). *N-myc* expression increases in the mesenchyme as it condenses and forms the comma and **S**-shaped bodies. Expression is confined to early differentiation stages and decreases as the epithelium forms and polarises (Mugrauer *et al.*, 1988; Hirvonen *et al.*, 1989). Likewise the *c-myc* gene is expressed in the un-induced mesenchyme, becoming restricted to the newly forming epithelium, but no expression in the loose mesenchyme and at the onset of terminal differentiation *c-myc* is turned off in the tubules. *L-myc* gene by contrast is expressed at a high level throughout nephrogenesis in the ureter, renal pelvis, papilla and collecting ducts increasing as the nephron develops to a maximal level in the adult kidney. As previously stated, *N-myc*

expression remains elevated at times when it should have been repressed in Wilms' tumours suggesting that WT-1 may be involved in *N-myc* repression.

The limb deformity gene (*ld*) was found by an insertion of a transgene into the coding region truncating transcripts and disrupting both limb and kidney development (Woychik *et al.*, 1985). The gene was cloned and found to be a member of the formin family. The protein is a phospho-protein and preliminary evidence suggests it may be a DNA-binding protein (Vogt *et al.*, 1993). The protein is expressed in the ureter and ureteric bud and in the various different alleles of the *ld* mouse the ureteric bud has a variable success rate of outgrowth and invasion of the kidney. Therefore it suggests that *ld* is involved in the signal for the ureter to produce the ureteric bud and for it to grow towards and invade the mesenchyme.

1.8 Thesis outline

This introduction has introduced what is currently known about the cellular and molecular basis of nephrogenesis and described briefly the developmental systems that have been used to elucidate this information. As can be seen from the description of the genes known to be involved in nephrogenesis, there is a great paucity of knowledge about the genes that regulate kidney development. Therefore, the aim of this thesis was to use modern molecular techniques to search for genes important in nephrogenesis. Rather than use a candidate gene approach or searching using motifs (as discussed in chapter 3), it was decided to use an approach that would isolate genes expressed in the fetal kidney by subtractive hybridization. Such an approach has no preconceptions of the kind of genes to be isolated and hence allow a greater spectrum of genes to be found. This thesis describes the approach taken and the analysis of the cDNAs isolated. It also evaluate the success of this approach. The first results chapter describes a proportion of clones isolated using the subtractive hybridization approach and evaluates these cDNAs. The other two results chapters describe the cloning of the mouse 14-3-3 epsilon isoform cDNA, a gene identified by the subtractive

hybridization, the characterization of the cDNA and the *in situ* hybridization analysis of this protein. This is summarised by an analysis of the 14-3-3 family and their role in the organism.

The mouse 14-3-3 epsilon isoform gene is shown to be very highly expressed in undifferentiated mesenchyme, and highly expressed in the fetal kidney and not expressed in the adult liver. This suggests that the gene satisfies the requirements of the subtractive hybridization, as do the other seven isolated cDNAs. The clones isolated show that a highly successful technique has been designed, applied and evaluated. The gene discussed here at length is a member of a very intriguing family of proteins that are implicated in activation of proteins in the signal transduction pathway, the catecholamine pathway and the family have also functionally been proved to be involved in Ca^{2+} regulated exocytosis and these points are considered at length in the final discussion.

Chapter 2

Materials and Methods

2.1. Bacterial cell culture: Media and additives

LB (Luria-Bertoni) broth: 10 g bacto-tryptone (Difco), 5 g yeast extract (Difco), 10 g NaCl, 4.6 g of $MgSO_4$ per litre of distilled water, adjusted to pH 7.2. (The $MgSO_4$ was omitted from broth used with tetracycline as Mg^{2+} is an antagonist of tetracycline.)

L agar: 15 g agar to 1 litre of LB broth as above.

Cy agar: Per litre 10 g casamino acids (Difco), 5 g yeast extract, 3 g NaCl, 10g agar.

Cy top agar: As above, except 6.5 g agarose in place of the agar.

2 x Ty broth: Per litre, 16 g Na_2HPO_4 10 g yeast extract, 5 g NaCl.

Terrific broth: Per 800 ml 12 g tryptone, 24 g yeast extract, 4 ml glycerol, 90 ml of KPO_4 (see below).

KPO_4: 170 mM KH_2PO_4: 720 mM K_2HPO_4

SOB broth: Per litre 20 g bacto-tryptone, 5 g bacto yeast extract, 0.05 g NaCl, 10 ml of 250 mM KCl, 5 ml of 2 M $MgCl_2$.

SOB agar: 1.5% agar to SOB broth as above.

2.1.1. Additives

Ampicillin dissolved in water was added to media and agar to give a final concentration of 50 µg/ml.

Tetracycline dissolved in 100% ethanol was added to media and agar to give a final concentration of 10 µg/ml.

IPTG (isopropyl-β–galactothiopyranoside) dissolved in water, was added to agar to give a final concentration of 48 µg/ml.

X-Gal (5-bromo-4-chloro-indoyl-β-galactopyranoside) dissolved in dimethyl-formamide, was added to agar to give a final concentration of 40 µg/ml.

2.1.2 E. coli *(Escherichia coli) strains*

JM101: *supE, thi-1,* Δ*(lac-proAB),* F' *traD36, proAB, lacI*^{q}*ZΔM15,* restriction: (*rk*+, *mk*+), *mcrA*+.

XL1-Blue: *recA1, endAI, gyrA96, thi-1, hsdR17, supE44, relA1, lac [*F' *proAB lacI*^{q}*ZΔM15*, Tn*10*(tet^{r}).

2.1.3 Phage strains

λZAPII: λ*sbhI*λ*1° chiA131* [T amp *ColE1 ori lacZ'* T3 promoter- polycloning site-T7 promoter I] *srI*λ*3°, cIts857, srI*λ*4°, nin5, srI*λ*5°* (Short *et al.,* 1988).

2.2 DNA manipulation

2.2.1 Preparation of competent cells

One colony per 10 ml of broth was picked from a freshly streaked SOB agar plate and dispersed in SOB media by vortexing and was grown at 37°C with vigorous agitation. The bacteria grew to an OD_{550} of 0.45-0.55. The cells were incubated on ice for 15 min, sedimented at 3000 rpm, 4°C for 15 min. The cells were thoroughly drained and resuspended gently in one third of the original volume of RF1 (100 mM rubidium chloride, 45 mM manganese chloride, 30 mM potassium acetate, 10 mM calcium chloride and 15% w/v glycerol, pH 5.8) and incubated on ice for 15 min. The cells were sedimented as before and resuspended in two twenty-fifths of the original volume of RF2 (50 mM sodium acetate, 5 mM EDTA [ethylenediaminetetraacetic acid], 10 mM rubidium chloride, 75 mM calcium chloride, 15% w/v glycerol, and 10 mM MOPS [morpholinopropanesulfonic acid], pH 6.8) and incubated on ice for 15 min before flash freezing 200 µl aliquots in liquid nitrogen and storing at -70°C (Hanahan, 1983).

2.2.2 Transformation of competent cells

200 µl aliquots of frozen competent cells were slowly thawed, put on ice and DNA added (10 ng of supercoiled plasmid or DNA from a ligation). The cells were mixed

and incubated on ice for 30 min. The cells were heat-shocked at 42°C for exactly 2 min, placed on ice, and 200 µl of 2 x Ty broth added. The cells were placed at 37°C for 30 min, and spread on L agar plates with the appropriate antibiotic. The plates were incubated, inverted, at 37°C overnight.

2.2.3 Small scale preparation of plasmid DNA

A colony from a freshly streaked plate or from a transformation was used to inoculate 3 ml of L Broth plus required antibiotics and shaken at 37°C overnight. 1.5 ml of the culture was added to an Eppendorf tube and sedimented in a microcentrifuge and the supernatant discarded. The cells were resuspended in 100 µl of solution A (50 mM glucose, 10 mM EDTA, 25 mM Tris·HCl, pH 8.0) and incubated at room temperature for 10 min. 200 µl of solution B (0.2 M NaOH, 1% SDS) was added, and the tube inverted to mix the solution and incubated on ice for 5 min. 150 µl 5 M sodium acetate, pH 4.8 was added and the solution incubated on ice for 30 min. The denatured protein and chromosomal DNA was sedimented in a microcentrifuge and 400 µl of the supernatant removed to a fresh Eppendorf tube. 200 µl of phenol (buffered with TE [1 mM EDTA, 10 mM Tris·HCl, pH 8.0]) was added, 200 µl of chloroform: isoamyl alcohol (24:1) added, the tube vortexed and centrifuged in a microcentrifuge for 10 min. The supernatant was carefully removed to a fresh Eppendorf tube and the phenol/chloroform extraction repeated. The second supernatant was removed and placed in a fresh Eppendorf tube, 1 ml of ice cold 100% ethanol added, and incubated on ice for 5 min to precipitate the DNA. The Eppendorf tube was centrifuged in a microcentrifuge at 4°C for 15 min. The supernatant was removed and the DNA pellet washed with 200 µl of 70% ethanol, dried under vacuum and resuspended in 100 µl TE with 2µl of 10 mg/ml RNAase (Birnboim & Doly, 1979).

2.2.4 Large scale preparation of plasmid DNA

A colony was used to inoculate 10 ml L broth with required antibiotics and incubated overnight. This starter culture was added to 400 ml terrific broth and grown for a further night. The bacteria were sedimented at 8000 rpm, 4°C for 10 min, the pellet drained and resuspended in 20 ml solution A (as above) and incubated at room temperature for 10 min. 40 ml of solution B (as above) was added, mixed and incubated on ice for 5 min. 20 ml of 5 M sodium acetate, pH 4.8 was added, and the mixture incubated on ice for 30 to 60 min. Denatured chromosomal DNA and protein were sedimented at 8000 rpm, 4°C for 10 min. The supernatant was filtered through muslin and 0.6 volumes of isopropanol added. The DNA was sedimented at 10,000 rpm at 4°C for 20 min. The pellet was washed with 70% ethanol, air-dried and dissolved in 3 ml TE, with 1 g/ml CsCl and 0.1 mg/ml ethidium bromide. The solution was sealed in tubes and centrifuged at 80,000 rpm for 16 hours to equilibrium. The plasmid DNA was removed from the tube with a 19 gauge needle. The ethidium bromide was removed by repeated extraction with equal volumes of NaCl-saturated isopropanol. Two volumes of water and six volumes of ethanol were added, the solution incubated at -70°C for 20 min, and centrifuged at 12,000 rpm, 4°C for 30 min. The DNA pellet was washed with 70% ethanol, dried under vacuum, and redissolved in 1 ml TE, pH 8.0.

2.2.5 Preparation of genomic DNA

Genomic DNA was prepared from frozen mouse tissues. A 1 cm^3 piece of tissue was thawed and chopped as finely as possible before being homogenized in a dounce homogenizer in 10 ml of resuspension buffer (150 mM NaCl, 10 mM EDTA, 10 mM Tris·HCl, pH 7.8). 0.5 ml of 20% SDS was added and the solution incubated at 4°C overnight, 50 µl of RNAase A (10 mg/ml) was added and the solution incubated at 37°C for 60 min, with intermittent gentle mixing. Another 100 µl of 20% SDS and 10 mg of proteinase K was added and the mixture incubated at 37°C overnight. The

DNA was extracted once with phenol (buffered with TE, pH 8.0), then repeatedly with phenol: chloroform: isoamyl alcohol (25:24:1) until the interface was clear, then extracted once with chloroform: isoamyl alcohol (24:1). The DNA was precipitated with 2.5 volumes 24:1 ethanol: 3 M sodium acetate, pH 5.2 and spooled out using a flame-sealed glass pipette. The DNA was resuspended in 1 ml dH_2O.

2.2.6 Amplification of DNA by the polymerase chain reaction

Polymerase chain reactions were performed on either plasmid DNA or crude lysates of phage or mouse embryonic material. For plasmid and genomic DNA, 100 ng was used per amplification.

Agar plugs containing phage plaques were cored into 50 µl of dH_2O, boiled for 5 min, sedimented for 30 seconds in a microcentrifuge, and amplification performed on 40 µl of the supernatant.

Crude cytoplasmic extracts were prepared from mouse embryo lysates according to Kawasaki (1990). 7.5 d.p.c. and 8.5 d.p.c. embryos were homogenized in 20-40 µl of ice-cold IHB (140 mM NaCl, 15 mM $MgCl_2$, 10 mM Tris·HCl, pH 8.0) containing 0.5% (v/v) TWEEN and 0.01% (v/v) diethylpyrocarbonate (depc). The nuclei and debris were removed by centrifugation for 10 min in a microcentrifuge at 4°C. The supernatant was transferred to a fresh tube, and incubated at 37°C for 20 min and at 90°C for 10 min. The solution was centrifuged to remove any debris and the supernatant containing the crude cytoplasmic extract transferred to a fresh tube.

PCR reactions were carried out in a 100 µl mix containing typically 100 ng of template DNA (as discussed as above), 0.2 mM dNTPs, 100 ng of each primer, 1 x PCR buffer (50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris·HCl, pH 8.3) and 2 units of Taq polymerase (Promega), and the reaction mix overlaid with 35 µl liquid paraffin. The typical program was 94°C for 2 min, followed by 30 cycles of 92°C for 0.5 min, 55°C for 1 min, 72°C for 2 min. This program was modified emperically according to the primers used and the length of product required but gave results for most products

required. 10 µl of the reaction was visualised by gel electrophoresis and ethidium bromide staining.

RNA-PCR was performed on 10 µl of the cytoplasmic extracts from 7.5 d.p.c. and 8.5 d.p.c. embryos, or on 1 µg of purified total cellular RNA. RNA was denatured, with 100 ng of either an internal oligonucleotide to the sequence of interest or random hexamers (N_6), by incubation at 75°C for 10 min and snap-cooling on ice. 2 µl (15 units) of RNAase inhibitor (Boehringer Mannheim), and 15 units of M-MuLV reverse transcriptase was added and the volume adjusted to 20 µl with RNAase-free water before incubation at 37°C for 90 min. The reaction volume was adjusted to 100 µl containing the PCR reagents as above and amplified as above.

2.2.7 Isolation of DNA fragments

Plasmids were cut with restriction enzymes to release appropriate cDNA fragments for subcloning or random-hexamer labelling. The digestion reaction was separated in 1.5-2.0% low melting point agarose gel, at 40-60 V until the fragment could be seen to be separated from the vector. The DNA was then visualised under long-wave UV, to prevent nicking of the DNA, and excised from the gel. The agarose was melted at 65°C, then cooled at 40°C for 10 min, then 1 unit of agarase (New England Biolabs) was added per 100 µl of melted agarose. The solution mixed and incubated at 40°C for 4-6 hours. The Eppendorf tube was then placed on ice for 5 min and briefly centrifuged to precipitate undigested agarose, the supernatant was removed to a fresh Eppendorf tube and 2.5 volumes of 24:1 ethanol:3 M sodium acetate, pH 5.2 added. The tube was then incubated at -70°C for 15 min, and centrifuged for 15 min at 4°C in a microcentrifuge. The supernatant was removed, the DNA pellet washed with 70% ethanol, and dried under vacuum, then resuspended in 20-30 µl of TE depending on the amount of DNA retrieved.

2.2.8 Dephosphorylation of vector DNA

Plasmid DNA was restriction-enzyme digested to give either compatible ends for the fragment to be sub-cloned or blunt ends to allow any fragment to be ligated in. Part of the digestion was run on a test gel to check that the digestion had gone to completion. When ligation was to be optimised, the vector was gel purified as above and dephosphorylated to prevent self-annealing of the vector. The DNA was incubated with 5 units of calf intestinal alkaline phosphatase (CIAP) in the buffer supplied by the manufacturer (Boehringer Mannheim) at 37°C for 30 min. The reaction was phenol/chloroform extracted twice and the DNA precipitated by the addition of 2.5 volumes of 24:1 ethanol: 3 M sodium acetate, pH 5.2. The DNA was sedimented for 15 min at 4°C in a microcentrifuge and resuspended to give a DNA concentration of 100 ng/µl.

2.2.9 Ligation of DNA fragments and vectors

For most ligation reactions 200 ng of vector DNA was used and an appropriate amount of fragment DNA to give a 1:1 or a 2:1 fragment to vector molar end ratio. For most ligations, multiple reactions were set up in parallel with different ratios of fragment to vector, with vector alone and with fragment alone to assess ligation efficiency and vector reannealing. Ligations were set up in a 10 µl reaction with both vector and fragment DNA being added in as small a volume as possible in 1 x ligation buffer (66 mM $MgCl_2$, 10 mM ATP, 66 mM Tris·HCl, pH 7.5,) with 2 units of T4 DNA ligase (Boehringer Mannheim). The reaction mix was incubated at 12°C for 16-20 hours before being transformed into competent cells.

2.210 Restriction enzyme digestions

DNA was routinely enzymatically digested by restriction enzymes according to the manufacturers instructions. Plasmid DNA (200-500 ng) was usually digested for 1-2 hours with 2-4 units of enzyme, while genomic DNA (10 µg) was digested for 12-16 hours with 10-15 units of enzyme. Genomic digests were checked by running an

aliquot of the digest on a mini-gel, if digestion was not complete a further 10 units of enzyme were added and the DNA incubated a further 8-10 hours to allow complete enzyme digestion.

2.2.11 Southern blotting

DNA was transferred from agarose gels to nylon membranes using the method of Southern (1975).

A photograph was taken of the gel visualised by UV light. The gel was denatured in 1.5 M NaCl , 0.5 M NaOH for 40 min and neutralised in 1.5 M NaCl, 1 mM EDTA 0.5 M Tris·HCl, pH 5.5 for 30 min. The gel was then briefly washed in 2 x SSC and the DNA transferred onto Hybond N (Amersham) by capillary transfer (Southern, 1975) using 20 x SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0) as a transfer medium. The gel was blotted for 16-20 hours, the filter was marked so as to allow accurate orientation, washed in 2 x SSC before being baked at 80°C under vacuum for 2 hours.

2.2.12 Random hexamer labelling of DNA

This technique was used to radioactively label DNA and is based on the method of Feinberg and Vogelstein (1983). The DNA (50-100 ng) was denatured by boiling for 5 min, snap-cooled on ice and random hexamers in a 1 x reaction buffer, 0.025 mM of cold dATP, dGTP and dTTP added, before the addition of 50 μCi of $\alpha[^{32}P]$-dCTP and 2-4 units of Klenow polymerase. The reaction was incubated at room temperature for 12-16 hours. An approximation of the efficiency of the labelling reaction could be determined by taking a small aliquot of the final reaction spotting it onto a Whatman GF/A filter and monitoring the radiation both before and after washing the disc with 10% TCA (trichloroacetic acid). Washing the filter with TCA removes un-incorporated nucleotides while retaining DNA molecules larger than 20 bp, this allows an accurate estimation of the efficiency of the reaction. Probes were labelled at 60-80% efficiency and if a probe did not labelled to this extent it was discarded and

the reaction repeated. The probe was purified by centrifugation through a pre-spun 1 cm high column of sephadex-G50. The probe was denatured by heating at 100°C for 10 min before being added to a pre-hybridized filter in hybridization solution.

2.2.13 Hybridization of Southerns and northern filters.

Filter hybridization was generally done in sealed bags in a shaking waterbath, or occasionally in rotating bottles in a controlled temperature oven (Hybaid). In each case a maximum 3 filters were placed in each bag/bottle to allow efficient hybridization. When multiple filters were hybridized in bottles, nylon cloth was placed between each filter. The temperature of hybridization was dependent on the stringency required. For the experiments described in this thesis all hybridizations were performed at high stringency, so pre-hybridizations, hybridizations and washes were performed at 68°C.

Filters were pre-wetted in 2 x SSC, then pre-hybridized in 0.5 M sodium phosphate, pH 7.2, 7% SDS (Church & Gilbert, 1984) at 68°C for at least 1 hour. The denatured probe was added and rolled over the filters several times to allow efficient initial mixing and then incubated at 68°C for 16-20 hours to allow maximal hybridization. [^{32}P] Probes used were of a specific activity of 2 x 10^9 dpm/µg and usually added to 20-40 ml of hybridization solution to give 3- 6 x 10^3 dpm/ml of hybridization solution.

The filters were removed from the hybridization solution and rinsed in wash buffer (40 mM sodium phosphate, pH 7.2, 1% SDS, 40 mM EDTA), fresh wash buffer added and shaken at 68°C for 30 min. After two further washes of 30 min in the same buffer the filters were monitored with a Geiger counter. If background radiation of the filter was low, the filters were partially dried and wrapped in cling-film to prevent complete drying. The filters were placed in film cassettes and exposed to XR-Film (Kodak) at -70°C for an appropriate length of time.

This protocol was used to hybridize cDNA probes to cloned DNA and genomic DNA Southerns, phage lift filters and northern blots.

2.2.14 Stripping filters

Filters were stripped of hybridized probe by adding them to a large volume of 0.1% SDS, boiling for 2 min and allowing them to cool to room temperature. Stripping blots allowed northern and Southern blots to be hybridized to different probes to a maximum of three times.

2.2.15 Preparation of oligonucleotides

Oligonucleotides were made on the MRC unit's oligonucleotide synthesiser machine by Ms Agnes Gallacher, and provided after de-protection as an ammonium hydroxide solution. 350 µl of the concentrated stock was precipitated by the addition of 770 µl of ice-cold ethanol and 35 µl of 3 M sodium acetate and precipitated at -20° C for 30 minutes. The oligonucleotide was pelleted in a microcentrifuge at 4°C for 15 min. The pellet was washed with 500 µl of 70% ethanol twice and air dried. The pellet was resuspended in 200 µl of TE their concentration determined by absorbance at 260 nm (using the assumption that 1 OD_{260} =33 µg/ml) and the concentration adjusted to 1 mg/ml.

2.2.16 Sequencing

Sequencing was performed by the dideoxy chain termination technique (Sanger *et al.*, 1977) using a 'Sequenase' sequencing kit (USB United States Biochemicals). Both single stranded and double stranded templates were successfully used.

Double stranded DNA sequencing required more template and primer, and the DNA had to be denatured prior to sequencing. 2 µg of plasmid DNA in 24 µl was incubated with 6 µl of 2 M NaOH at 37°C for 15 min. 70 ng of primer was added to the denatured template, the solution vortexed, and 9 µl of 3 M sodium acetate, pH 5.2 added. 240 µl ethanol was added and the mixture frozen on dry ice for 20 min. The DNA was sedimented in a microcentrifuge for 15 min at 4°C, the supernatant

removed and the pellet washed with 70% ethanol, air dried and used in the Sequenase protocol. Single strand templates were used directly in the Sequenase protocol.

2.2.17 Polyacrylamide gel electrophoresis

Denaturing polyacrylamide/urea gels were used to resolve the sequencing reactions. The gel plates were cleaned with acetone, then ethanol. The notched plate was treated with acrylease (Stratagene) to prevent the gel sticking to the plate. The two plates were taped together separated by 0.05 cm thick plastic spacers to cast a 0.05 cm thick gel. A 8% acrylamide, (20:1 acrylamide:bis-acrylamide [N, N'-methylene-bis-acrylamide]), 6 M urea, 0.5 x TBE gel (50 ml) was polymerized with 450 µl of 10% ammonium persulphate and 45 µl of TEMED (N,N,N',N'-tetramethylethylenediamine) added just before pouring. The acrylamide solution was poured between the two plates and the back of a comb was pushed 0.7 cm into the top of the gel before it set, to form a ledge for sample loading with a 'shark's tooth' comb. The gel was left to set for 30 min before being pre-run at 20 W for 15 min in 1 x TBE buffer. The samples were heated at 75°C for 2 min and loaded immediately. The samples were run on the gel at 20 W for two or five hours. The notched plate was removed and the gel fixed in 10% methanol, 10% acetic acid for 15 min. The gel was transferred to a piece of 17 mm Whatman filter paper and dried on a gel drier under vacuum for 45 min before being exposed to X-ray film overnight at room temperature.

2.3 RNA manipulation

2.3.1 RNA extraction

All adult tissue samples were taken from pregnant female Swiss mice of various ages except the testes samples which were taken from male Swiss mice. RNA was made by the acid-phenol method of Chomczynski & Sacchi (1987). The tissue was dissected, flash frozen in liquid nitrogen and stored at -70°C until RNA/DNA

extraction was to be performed. A 1 g piece of tissue was thawed into 15 ml ice cold 4 M guanidine isothiocyanate, 25 mM sodium citrate, 0.5% sarcosyl, 0.1 M β-mercaptoethanol, cut into small pieces and homogenized in a 7 ml glass dounce homogenizer (Jencons). The homogenate was passed through a 19 gauge needle to shear chromosomal DNA. To the homogenate 1/10 th volume 2 M sodium acetate, pH 4.0, 1 volume water saturated phenol and 2/10 th volume chloroform were added sequentially with the mixture being inverted after each addition. The homogenate was shaken vigorously for ten seconds, cooled on ice for 15 min and centrifuged at 10,000 rpm, 4°C for 20 min. The supernatant containing the RNA was removed to a fresh tube and an equal volume of isopropanol added, the solution mixed and precipitated overnight at -20°C. The RNA was sedimented at 10,000 rpm, 4°C for 30 min, the supernatant discarded and the pellet washed with 70% ethanol 30% RNAase-free dH_2O. The RNA pellet was dried under vacuum, and resuspended in an appropriate volume of RNAase-free dH_2O to give approximately 1 mg/ml of RNA. The concentration was checked by taking OD readings at 260 nM. For RNA an OD_{260} absorbance of 1 corresponds to 40 µg/ml of RNA.

2.3.2 Messenger RNA selection

mRNA selection by oligo (dT) cellulose chromatography was done according to Aviv & Leder (1972). 1 ml volume of oligo (dT) was resuspended in 0.2 M NaOH and poured into a plastic column washed to neutral pH with sterile TE, pH 8.0 and equilibrated with 1 x column loading buffer (one half volume 2 x column loading buffer;- 1 M NaCl, 2 mM EDTA, 1% Sarcosine, 40 mM Tris·HCl, pH 8.0). The RNA was heated to 65°C for 10 min, snap-cooled on ice and NaCl added to 0.5 M, an equal volume of 2 x loading buffer was added to the column. The RNA was applied to the column and the effluent collected and reheated at 65°C for 10 min and added to the column again. The column was washed with 7 ml of 1 x loading buffer. The poly $(A)^+$ RNA was eluted with 1 ml of loading buffer with no sarcosine. 200 µl fractions

from the column were collected and the OD_{260} of each fraction taken to identify the fractions which contained mRNA. These fractions were pooled and ethanol precipitated using 2 µg/100 µl of glycogen as a carrier to facilitate precipitation. The RNA was resuspended at 2 µg/µl and stored at -70°C.

2.3.3 Denaturing formaldehyde gels

RNA samples were run according to Sambrook *et al.*. (1990), a 100 ml 1 % formaldehyde gel was made by dissolving 1 g of agarose in 72 ml of water, cooling to 60°C and adding 10 ml of 10 x MOPS buffer (50 mM sodium acetate, 5 mM EDTA, 200 mM MOPS, pH 7.0), 4 µl of 10 mg/ml ethidium bromide and 17.5 ml of formaldehyde (37-40% w/v to a final concentration of 7%) and poured immediately into a gel casting tray.

7-10 µg of total RNA was loaded per lane in 5 µl of RNAase-free dH_2O. To this was added an equal volume of Formaldehyde Sample Buffer [FSB 100 µl of 10 x MOPS buffer (as above), 200 µl of formamide (99%) and 120 µl of formaldehyde (37-40% w/v)]. The sample was heated at 65°C for 10 min and snap-cooled on ice. A half volume of loading buffer (98% formamide, 10 mM EDTA [pH 8.0], 0.025% xylene cyanol, 0.025% bromophenol blue) was added and the samples loaded immediately. The samples were run in 1 x MOPS buffer at 100 volts for 10 min before being run overnight (16-20 hours) at 30 volts. The RNA was visualised by ethidium bromide staining and a photograph taken after electrophoresis, allowing a evaluation of both the quality of the RNA and the accuracy of RNA quantitation between samples.

The gel was incubated in dH_2O for 20 min, in 10 x SSC for 10 min, and the RNA transferred from the gel onto a nylon filter by northern blotting (Southern 1977) using 10 x SSC as the transfer medium. The northern blot was generally left for 16-24 hours before being dismantled, the filter was removed washed in 2 x SSC and visualised on

a UV box. The position of the 18S and 28S rRNAs and any RNA size markers could be marked directly on the filter to allow accurate sizing of any positive signal.

2.4 cDNA and phage manipulation

2.4.1 Preparation of a cDNA library

8 µg of mRNA was used in a commercial cDNA kit (Pharmacia), the first strand synthesis being made more efficient by supplementing the reaction with 40 units of AMV reverse transcriptase (BCL). The kit makes oligo (dT) primed cDNA with linker adapters on each end of the cDNA. The linker adapters contain a rare cutter restriction site (*Not I*) in the linker and have *Eco RI* compatible ends. One fifth of the total cDNA was precipitated, air dried and ligated to 1 µg of Eco RI digested and phosphatased λ ZAPII (Stratagene; Short *et al.,* 1988) by incubation in 1 x ligation buffer (66 mM $MgCl_2$, 10 mM ATP, 66 mM Tris·HCl, pH 8.0) with 2 units ligase at 16°C overnight.

2.4.2 Preparation of plating cells.

XL1-Blue bacterial cells were streaked out on a L -Tet agar plate (L agar with tetracycline at 10 µg/ml of tetracycline) grown at 37°C overnight. A single colony used to inoculate 10 ml of L-Broth, 0.2% maltose, 10 µg/ml of tetracycline. After an overnight incubation at 37°C the cells were sedimented at 3000 rpm, 4°C for 10 min and resuspended in half the original volume of 10 mM $MgSO_4$. The cells were then stored at 4°C until use, plating cells were most efficient on the day of preparation but could be used for up to one week after that.

2.4.3 Packaging of the cDNA library

The ligation was packaged using the Gigapack Gold kit (Stratagene) according to the manufacturers instructions. The ligation was added to the freeze/thaw extract, mixed and 15 µl of sonic extract added, mixed gently, and the solution incubated at room temperature for two hours. 500 µl of Phage Suspension Buffer [PSB] (5.8 g

NaCl, 2g $MgSO_4$, 0.02 M Tris·HCl, pH 7.6, 0.01% gelatin per litre) was added, and gently mixed. The packaged phage library was amplified by addition to 800 µl of plating cells and incubated at 37°C for 15 min. An appropriate amount of molten Cy top agar (50°C) was added, mixed and poured onto prewarmed and dried Cy agar plates and incubated at 37°C for 6-8 hours. The amplified library was recovered by adding 20 ml of PSB to the surface of the agar and shaking the plates on a suspension mixer at 4°C overnight, then decanting the liquid containing the adsorbed phage library, into a sterile container. Chloroform was added to 0.01% v/v to prevent the growth of bacteria. The library was diluted with PSB to an appropriate concentration and plated as above when analyzed further. These plates were used for taking filter lifts and could be stored at 4°C for 2-3 weeks while phage of interest were identified and removed for further analysis.

2.4.4 Plating cDNA phage libraries

Phage libraries usually had a titre of 10^{10}-10^{12} per ml. To screen a phage library on a 20 cm by 20 cm plate the phage were plated at a density of 5-10,000 phage as above. Secondary screening of isolated phage plaques were plated as above at a density of 1-200 phage per 9 cm plate.

2.4.5 Filter lifts from phage plates

Plates were incubated at 4°C for at least two hours to allow the top agar to solidify, which prevented the top agar sticking to filters when filter lifts were taken. A Hybond-N filter was cut to the size of the plate, placed dry onto the surface for 1 min and marked with India ink using a needle to allow orientation later. The filter was floated in denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 2 min, in neutralizing solution (2 M NaCl, 1 M Tris·HCl, pH 5.5) for 2 min, briefly washed in 2 x SSC, air dried then baked in a vacuum oven at 80°C for two hours. Replica filters could be taken from the same plate by the same technique but the initial incubation was

doubled for each subsequent plate to allow equivalent amounts of DNA to transfer to each filter, i.e. 1 min for the first filter, 2 min for the second, 4 min for the third.

2.5 Methods used for the subtractive hybridization

2.5.1 RNA biotinylation

10 µg aliquots of poly $(A)^+$ RNA were mixed with 50 µg (50 µl) of photoactivatable biotin (Super-XX-PAB Clontech) and irradiated on ice with a sun lamp (InVitrogen) at 10 cm above the sample for 15 min. Unreacted PAB was removed by adding a tenth volume 1 M Tris·HCl, pH 9.0 and extracting with TE-saturated 2-butanol until the butanol was clear. RNA was chloroform extracted and precipitated by the addition of 2.5 volumes of 24:1 ethanol: sodium acetate, pH 5.2. The biotinylation was repeated to increase the density of biotin attached to the RNA.

2.5.2 Subtractive hybridization

cDNA used for subtractive hybridization was primed with a tagged oligo (dT) oligonucleotide:- JM1, GGTCGACGGTACCGAATTCT$(T)_{17}$ using 40 units of AMV reverse transcriptase (Boehringer Mannheim) and incubated at 42°C for 90 min. The cDNA was passed over a Size-Select-400 column (Pharmacia) to remove free nucleotides and cDNAs smaller than 400 base pairs. The effluent was precipitated and an oligo (dC) tail added by incubation with terminal transferase in 1 x terminal transferase buffer (according to the manufacturers instructions, Boehringer Mannheim) with dCTP for 2 min at 37°C. The tailed cDNA was heated at 65°C for 10 min and phenol extracted to remove the enzyme and precipitated by the addition of 2.5 volumes of 24:1 ethanol: sodium acetate, pH 5.2 .

2.5.3 cDNA/PAB-RNA hybridization

Single-stranded cDNA and 20 µg PAB-RNA were hybridized at 10-fold excess of PAB-RNA to cDNA, in 7 µl of (HB) (50 mM HEPES [N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid], pH 7.6, 2 mM EDTA, 500 mM NaCl, 0.2% SDS), overlaid

with mineral oil, heated to 95°C for 2 min, and hybridized at 65°C for 48 hours to reach a R_0t of approximately 700 (Hastie & Bishop, 1976).

2.5.4 Separation of the hybrids from single-stranded cDNA.

100 µl HB without SDS was added to the hybridization mix, 5 µg of streptavidin (BRL) was added and the mixture incubated at room temperature for 5 min. The reaction was extracted with TE-saturated phenol:chloroform:iso-amyl-alcohol (25:24:1) (P/C). The aqueous layer containing the unhybridized single-stranded cDNA was removed to a new tube and the organic phase back-extracted to prevent loss of single-stranded cDNA. The entire extraction was repeated twice more, the aqueous phases were chloroform extracted and the cDNA precipitated by the addition of 2.5 volumes of 24:1 ethanol: sodium acetate, pH 5.2 with 2 µl of 2 mg/ml glycogen added as a carrier. The subtraction was repeated to ensure that all common mRNAs had been depleted.

The second strand of the cDNA was primed with 200 pg of JM-2+(dG): AGAATTCGGTACCGTCGACC$(G)_{18}$ using *Taq* polymerase at 95°C for 1 min, 45°C for 1 min, then 1 hour at 72°C. Lone linker PCR (Ko *et al.*, 1990) was performed with 2 µg of oligonucleotide JM1: GGGTCGACGGTACCGAATTCT, by incubation at 95°C for 2 min, then 40 cycles of 94°C for 0.5 min, 55°C for 1 min, 72°C for 2.5 min, followed by 72°C for 10 min.

2.5.5 In Vivo *excision of plasmids from λZAPII*

Phage plugs that had been identified by hybridization were cored from the agar plate and transferred to 500 µl PSB in an Eppendorf tube, vortexed and incubated at room temperature for 2 hours to release the phage particles. 200 µl of the phage stock (about 10^5 phage) was put in a falcon tube with 200 µl of an overnight XL1-Blue culture and 1 µl of VCSM13 helper phage (titre of 10^{10} pfu) (Stratagene) incubated at 37°C for 15 min before adding 5 ml 2 x Ty or terrific broth and shaking at 37°C for 3 hours The tubes were heated at 70°C for 20 min, the cell pellet centrifuged at 6000

rpm for 5 min, the supernatant decanted into a sterile tube and stored at 4°C as a phagemid stock. The plasmid was rescued by adding 10 µl of the phagemid stock to 200 µl of saturated culture of XL1-Blue cells, incubating at 37°C for 15 min, and spreading dilutions LB/amp/X-Gal/IPTG plates. The plates were incubated at 37°C for 16-20 hours before colonies could be picked to grow in LB-amp for further usage.

2.5.6 Single-strand rescue

Single-stranded DNA was prepared according to Katayama (1990). The bacteria were streaked out on a fresh LB-amp plate and 1 colony was used to inoculate 5 ml of 2 x Ty with amp at 75 µg /ml and 1 µl/ml of VCSM13 and grown with vigorous aeration for 2 hours. Kanamycin was added to a final concentration of 70 µg/ml and the culture was grown for a further 12 hours. The VCSM13 phage provides kanamycin resistance to infected cells allowing selection for infected bacteria. The cells were sedimented at 6000 rpm for 5 min, the supernatant was decanted to a fresh tube, 0.15 of the original volume of 20% PEG, 2.5 M NaCl added, and the supernatant cooled on ice for 15 min. The phage were sedimented by centrifugation in a microcentrifuge for 15 min, resuspended in 400 µl 0.3 M sodium acetate, 1 mM EDTA and extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1). The DNA was precipitated by the addition of 2.5 volumes of 24:1 ethanol:3 M sodium acetate, pH 5.2, incubation at -20°C for 30 min, sedimented in a microcentrifuge for 15 min at 4°C, washed with 70% ethanol, dried under vacuum and resuspended in TE.

2.6 *In situ* hybridization

2.6.1 Radioactive and non-radioactive in situ *hybridization*

Embryos for radioactive *in situ* were kindly prepared, sectioned and slides prepared by Mrs Allyson Ross. Embryos for non-radioactive whole mount *in situ* were kindly provided by Mrs Elizabeth Graham. The probes for both techniques were

in vitro transcribed by myself but Dr J. F. Armstrong performed the radioactive *in situ* hybridization experiments and Mrs Elizabeth Graham performed the non-radioactive *in situ* hybridization experiments discussed in this thesis. The radioactive *in situ* experiments were performed as Davidson *et al.* (1988) and the non-radioactive experiments as Wilkinson (1992).

2.6.2 Preparation of a template for in vitro *transcription*

The plasmid containing the mouse epsilon 14-3-3 cDNA was cut with *Hind III* or *Bam HI* to give a template that allowed transcription of an antisense or sense RNA probe by T7 or T3 polymerase respectively., 1 µg of cut DNA was phenol-chloroform extracted twice and ethanol precipitated to provide a clean template for the RNA *in vitro* transcription.

2.6.3 RNA labelling with digoxigenin (DIG)-UTP

1 µg of the template DNA in 5 µl of TE was incubated with 2 µl 10 x transcription buffer (Boeringher Mannheim), 13 µl dH_2O, 2 µl rNTP labelling mixture containing ATP (10 mM CTP, 10 mM GTP, 6.5 mM UTP, 3.5 mM DIG-UTP in Tris·HCl, pH 7.5), 2 µl T3 (or T7) polymerase, 1 µl RNAase Inhibitor, centrifuged briefly and incubated for 2 hours at 37°C. 2 µl of RNAase-free DNase 1 was then added and incubated at 37°C for 15 min to remove the DNA template. 2 µl of 0.2 M EDTA, pH 8.0 was added to stop the enzyme reaction, and 2.5 µl of 4 M LiCl and 75 µl of ice-cold ethanol added, the solution was mixed and incubated for 30 min at -70°C or 2 hours at -20°C to precipitate the RNA. The RNA was sedimented in a microcentrifuge for 15 min, washed with 70% ethanol, and dried under vacuum.

The RNA pellet was redissolved in 100 µl RNAase-free H_2O, and 1 µl RNAase Inhibitor by incubating the pellet at 37°C for 30 min. 30 µl of the RNA solution was then hydrolysed by incubation with 30 µl sodium bicarbonate at 60°C for 50 min. The alkaline digestion generates random fragments of approximately 150 bp, the optimal size to permeate cells. 30 µl ammonium acetate and 400 µl cold 100% ethanol was

added, the probe mixed and incubated at -20°C overnight. The probe was sedimented in a microcentrifuge for 15 min, washed with 70% ethanol, dried under vacuum, and resuspended in 25 µl TE, 2 µl was used to assay the labelling of the probe immunologically. DNA cut for sense and antisense transcriptions and a control template were transcribed for each set of experiments.

2.6.4 Immunological detection of DIG probe

All incubations (except for the colour reaction) were performed at room temperature with shaking. 1 µl dilutions of the probe, the control reaction and control labelled RNA (provided in the Boeringher Mannheim DIG detection kit) were spotted onto a piece of Hybond-N nylon membrane. Dilutions of 1, 1:10, 1:100, 1:1000, 1:10000, 1:100000 and 1: 1000000 were made. The filter was allowed to dry then UV irradiated for 3 min. The filter was washed 1 min in Buffer 1 (0.15 M NaCl, 0.1 M maleic acid, pH 7.5) then incubated for 30 min in buffer 2 (1% blocking reagent in buffer 1). The buffer was incubated for 30 min in anti-DIG-AP (digoxigenin-alkaline phosphate) conjugate diluted 1:5000 in buffer 2. The unbound antibody was removed by washing the filter in buffer 1 for 2 x 15 min. The membrane was then equilibrated for 2 min in buffer 3 (100 mM NaCl, 50 mM MgCl2, 100 mM Tris·HCl, pH 9.5). The equilibrated buffer was incubated in colour-substrate solution (45 µl NBT [75 mM nitroblue tetrazolium salt in dimethylformamide] and 35 µl X-phosphate [50 mg/ml 5-bromo-4-chloro-3-indoyl phosphate, toluidinium salt in dimethylformamide] per 10 mls of buffer 3) in the dark until the desired spots are detected (typically 2-3 hours). The filter is then incubated in TE to stop the reaction. The filter can be photocopied for permanent reference and the concentration of the probe estimated by comparison with the control labelled RNA.

2.6.5 RNA labelling with ^{35}S UTP

1 µg of the transcription template in 5 µl of TE was incubated with 6 µl of 5 x transcription buffer, 1 µl 10 mM rATP, 1 µl 10 mM rCTP, 1 µl 10 mM rGTP, 1 µl 1M DTT, 12 µl ^{35}S rUTP (>1 mCi/100 µl),1.2 µl RNAase Inhibitor, 0.8 µl T7 or T3 Polymerase and 3 µl H_2O, at 37°C for 25 min. Then another 0.8 µl T7 or T3 polymerase was added and the mixture was incubated at 37°C for a further 25 min. 2 µl tRNA (10 mg/ml) or rRNA and 1 µl RNAase-free DNase was added and the reaction incubated at 37°C for 10 min. 2 µl of 100 mM EDTA was added to stop the reaction and TE with 50 mM DTT was added to a final volume of 100 µl. The reaction was phenol extracted twice, then extracted once with phenol:chloroform 1:1, 10 µl of 3 M Na acetate, pH 5.2 and 240 µl of ice-cold ethanol was added and the reaction precipitated at -20°C overnight.

The reaction was pelleted in a microcentrifuge at 4°C for 30 min and the pellet washed twice with 80% ethanol/50 mM DTT and once with 100% ethanol. The pellet was centrifuged for 2 min to collect any remaining supernatant and the pellet dried under vacuum. The RNA was redissolved in 100 µl RNAase-free H_2O/50 mM DTT. 190 µl of sodium bicarbonate was added and the mixture incubated at 60°C for 70 minutes. 100µl 3 M ammonium acetate, pH 5.2, 1 µl of tRNA (10mg/ml), and 900 µl of ice-cold ethanol was added. The RNA was then precipitated overnight at -20°C. The RNA was pelleted in a microcentrifuge at 4°C for 30 min, the supernatant removed and the RNA washed in 80% ethanol/50 mM DTT in TE twice then pelleted in a microcentrifuge for 10 min. The RNA was then washed in 100% ethanol, spun for 5 min, the supernatant removed and the RNA dried to near completion, then redissolved in 50 µl of TE/50 mM DTT. To estimate the incorporation of the radiolabel, a 1µl aliquot of the transcription reaction was added to 19 µl of TE and divided between two glass-fibre filters (Whatman GF/A filters) One was air-dried while the other was washed twice with ice-cold 10% TCA and once with ethanol and

air-dried. The dried filters were placed in scintillation vials with 10 ml Aquasol (DuPont) and the activity measured using a scintillation counter (Packhard 1500) standardized for ^{35}S emissions. The reaction was then diluted 1:9 in TE/50 mM DTT:hybridization solution (50% formamide, 10% dextran sulphate, 1 x Denhardts solution (50 x: 5 g ficoll, 5 g polyvinylpyrrolidone, 5 g BSA), 20 mM Tris·HCl, pH 8.0, 0.3 M NaCl, 5 mM EDTA, 10 mM NaH_2PO_4, 0.5 mg/ml yeast RNA and 50 mM DTT added just prior to use) at a final concentration of 1 x 10^5 dpm/µl.

2.7 Plasmids

2.7.1 List of plasmids used

pActin 1 kb of mouse β-actin cloned into the *Pst I* site of pGEM1 (Minty *et al.*, 1981)

2.7.2 List of plasmids made

pB2- mouse epsilon isoform of the 14-3-3 protein (in pBluescriptSK-)
pC3 mouse epsilon isoform of the 14-3-3 protein (in pBluescriptSK-)
pD1 mouse epsilon isoform of the 14-3-3 protein (in pBluescriptSK-)
pC1 clone C1 from the first subtraction (in pBluescriptSK-)
pC2 clone C2 from the first subtraction (in pBluescriptSK-)
pC4 clone C4 from the first subtraction (in pBluescriptSK-)
pC5 clone C5 from the first subtraction (in pBluescriptSK-)
pD2/pD4 clones D2 and D4 B1 repeats (in pBluescriptSK-)
pD3 clone D3 from the first subtraction (in pBluescriptSK-)

2.8 Oligonucleotide primers

2.8.1 Oligonucleotides used

B761	GTAAAACGACGGCCAGT	M13 FORWARD
B762	AACAGCTATGACCATG	M13 REVERSE
D299	TCTAGAACTAGTGGATCC	SK
D300	CGAGGTCGACGGTATCGA	KS
D302	ATTAACCCTCACTAAAGG	T3
D304	AATACGACTCACTATAG	T7
C942	TTTTTTTTTTTTTTTTT	BIOTINYLATED OLIGO DT
C519	TGCTGAAGTGTTGCTGAA	FOR M14-3-3
C520	TTCAGCAACACTTCAGCA	FOR M14-3-3

C521	ATAGCAGATTGGACGTG	FOR M14-3-3
C634	CACGTCCAATCCCTGCTAT	FOR M14-3-3
C635	TGAAAGCCATTAGACTTCT	FOR M14-3-3 REV OF C636
C636	AGAAGTCTAATGGCTTTCA	FOR M14-3-3 C520 EXT
C637	ATGTACTGGACAAACACCTC	FOR M14-3-3 EXT OF C521
C638	GAGGTGTTTGTCCAGTACAG	FOR M14-3-3
C662	GGACACGCTGAGTGAAGAA	FOR M14-3-3 CONT OF C637
C663	TTCTTCACTCAGCGTGTCC	FOR M14-3-3 REV OF C662
C664	CAGGGCAGGAACCTAAAT	FOR M14-3-3 CONT OF C636
C665	AATTTAGGTTCCTGCCCTG	FOR M14-3-3 REV OF C664
C859	CCACGTCCAATCCCTGCTA	FOR M14-3-3 C634 ALT
C842	TAGCTGCTGGAATGAGGTG	FOR M14-3-3 C638 ALT

2.8.2 Subtraction primers

C628	GGGTCGACGGTACCGAATTCT	JM1
C629	GGTCGACGGTACCGAATTCT$(T)_{17}$	JM1+d(T)
D327	AGAATTCGGTACCGTCGACC$(G)_{18}$	JM2+d(G)

Chapter 3

Making a fetal kidney cDNA library and subtraction strategies

3.1 Introduction

The aim of the project was to isolate genes with a temporally- and spatially-restricted kidney expression pattern. These genes are likely to be important in the inductive stages of early nephrogenesis. To isolate such genes first a mouse fetal kidney cDNA library had to be made and a strategy devised to isolate the genes that satisfied this criterion. To evaluate the most efficient strategy it is helpful to review the different methods that have been used to isolate developmentally important genes in other systems.

The study of genetic mutations which perturb embryogenesis has helped understand how normal development occurs. Mutations have been generated and the mutants collected for many years. Some of these mutations have arisen spontaneously, while others have been created by radiation or chemical mutagenesis of the parent or developing embryo and a large effort has focused on identifying those genes responsible for these developmental mutations.

3.2 Isolating developmental genes

There are several different approaches that have helped gather the information that is currently known. One method is 'reverse genetics' positional cloning by mapping and then cloning the gene responsible for particular developmental mutants. This approach has been used extensively in attempts to clone human disease loci. Improved genetic mapping techniques have led to the cloning of the genes responsible for a number of classical mutations and, the role of many of these genes have shown an interesting expression profile in normal development. These genes include the *Drosophila* homeotic and segmentation genes (*i.e. Antennapedia, dicephalic, engrailed, fused, fushi-tarazu, hedgehog, hunchback, Kruppel*) [reviewed by Akam 1987]; classical mouse mutations such as *small eye* (Hill *et al.*, 1991) *splotch* (Goulding *et al.,* 1993), *undulated* (Balling *et al.*, 1988); and human genes mutated in

such diseases as Wilms' tumour (Call *et al.*, 1990, Rose *et al.*, 1990) and aniridia (Ton *et al.*,1991).

Other developmentally important genes have been found by selecting a gene known to be important in one species and looking for the homologous gene in different species and for related genes in the same species. This approach has been used with much success to find homologues of many *Drosophila* genes, especially those encoding conserved protein motifs such as the homeo-box, paired box and zinc fingers.

Another approach is to try and clone genes for particular functions. Sequence analysis of the many thousands of cloned genes have shown that the DNA encoding regions of proteins known to perform important functions such as DNA-binding (zinc fingers, the homeobox, paired boxes and Pax genes) and protein phosphorylation, (serine/threonine kinases, tyrosine kinases, the receptor tyrosine kinases and src genes) has been conserved throughout evolution as recognisable 'motifs'. The level of similarity allows oligonucleotides to such regions to be predicted and these can either be used directly as probes or can be used for degenerate PCR to amplify cDNA regions that can then be used as probes to find novel cDNAs. Strategies such as this have been used to find many genes important in development (Freeman *et al.*, 1992; Gilardi-Hebenstreit *et al.*, 1992; Hromas *et al.*, 1993; Wolda & Moon, 1992).

An elegant method for finding temporally and spatially regulated genes is the use of enhancer and promoter traps. Enhancer traps typically use the *E. coli lac Z* gene cloned into a vector with a weak promoter in such a manner that when the vector is transfected into the host's genome; in *Drosophila* embryos by P-element transposition, and in mouse into embryonic stem cells (ES cells), the *lac Z* gene will only be expressed if the vector integrates into the genome in a chromosomal region that provides enhancer elements necessary for expression (O'Kane & Gehring, 1987). The expression pattern in such cases are dictated by the enhancer elements in the genome,

which presumably regulate nearby genes in the same pattern. The expression pattern can then be studied by whole-mount embryo staining as the β-galactosidase acts on a chromogenic substrate to give a blue stain. It has been shown that this strategy can be used efficiently in mice to find coding regions of genes that have a temporal and spatially restricted pattern of expression (Korn *et al.*, 1992).

Retroviral promoter trap vectors encodes a fusion protein with both β-gal and neomycin phosphotransferase activities [neo] (Friedrich & Soriano, 1991). The vectors are transfected into ES cells and insertion events neighbouring a promoter and enhancer active in ES cells can be selected for by acquired resistance to the antibiotic G418. The ES cells can then be used to create chimeric transgenic mice. The spatial and temporal expression pattern of the trapped genomic promoter can then be visualised by the β-gal stain as described above. However this technique only targets genes which are expressed in ES cells.

One advantage of both strategies is that they target random chromosomal regions allowing targeting of previously unknown genes. These targeted genes can be identified by cloning the DNA surrounding the insertion site, while the expression pattern of the gene can be visualized with the *lac Z* marker (von Melchner *et al.*, 1990, 1992; Friedrich & Soriano, 1991; Friedrich *et al.*, 1992). The insertion may have a phenotypic effect when bred to homozygosity and several strategies have been devised that select for insertion events that are lethal when bred to homozygosity, indicating that the gene disrupted is vital during development (Chen *et al.*, 1992).

3.3 Differential screening

These approaches are not directly applicable to looking for genes that have a temporally and spatially restricted expression in the developing fetal kidney. Therefore other approaches that do allow these questions to be answered were considered. Differential screening was one of the first techniques to select for genes expressed solely in one tissue type (Dworkin & Dawid, 1980a, b; Sargent & Dawid, 1983). The

technique involves screening a cDNA library with two probes made from RNA from two different stages or tissues and then comparing the resultant hybridization signals. Plaques or colonies that hybridise to both probes are disregarded as they are assumed to be transcriptionally active in both tissues or stages and likely to contain cDNAs for 'housekeeping' (metabolic pathway) genes. Whereas plaques or colonies that hybridise to one probe and not the other should contain cDNAs that are tissue/stage specific. A major problem with differential screening is that the technique has poor sensitivity in detecting and selecting low abundance transcripts. The low sensitivity requires that large numbers of plaques or colonies must be screened, and the high density of screening this necessitates makes the isolation of any clones of interest complicated and difficult. The cause of this problem is the large number of 'house-keeping' genes that are expressed in all tissues that are essential for the cell but have no role in determining its fate. Some of these genes are expressed at a high level and will obscure signals from cDNAs that are less abundant and expressed at a lower level. Unfortunately, most of the cDNAs that are being looked for in differential screening would be expected to fall into the less abundant category and only a small subset of temporally and spatially restricted genes can be detected at this level of sensitivity. Differential screening can be made more sensitive by screening subtracted libraries (as discussed below) to allow more sensitive detection of transcripts.

3.4 Subtractive hybridization

Subtractive hybridization is a highly sensitive technique and hence more successful in finding low level transcripts. In subtractive hybridization the strategy is to lower background signal from genes of little interest, e.g. the housekeeping genes, by physically removing the transcripts. The technique involves hybridizing the transcripts from the tissue of interest with transcripts from a subtracting 'driver' tissue of choice and removing both the hybrid and the driver transcripts. The subtracting tissue is called the 'driver' as the huge excess of the RNA cDNA drives the kinetics of

hybridization towards the formation of hybrids of common messages. The cDNA from the tissue of interest, (*i.e.* the brain, the kidney, the eye, or a certain developmental stage such as 8.5 d.p.c. embryo) is hybridized to a 40-100 fold excess of either mRNA or cDNA from the driver tissue. The messages present in both samples hybridize together and the multifold excess of driver allows the hybridization to go to completion. The hybrids and excess driver are then removed.

Hydroxyapatite (HAP) columns separate single and double-stranded nucleic acids and are often used to separate and isolate both the cDNAs in common and the cDNAs specific to the tissue of interest. This approach has many technical difficulties that can often lead to the loss of the subtracted cDNA. HAP columns are run in a high phosphate buffer and the subtracted cDNA and hybrids are eluted in this buffer. Elution in this high phosphate buffer presents technical difficulties in subsequent steps as nucleic acids cannot be ethanol-precipitated from high phosphate solutions. The difficulties of the HAP column were generally exacerbated in the early subtractive hybridization protocols when radioactively labelled cDNA was subtracted and used directly for library screening, necessitating the use of large amounts of high specific activity cDNA.

Clones that have been found using hydroxyapatite columns include genes involved in controlling the circadian clock in Neurospora (Loros *et al.*, 1989) and *Myo D*. *Myo D* was also originally found by the subtractive hybridization of 5-azacytidine-induced fibroblast cells against uninduced cells as 5-azacytidine induces fibroblasts to differentiate into myoblasts. The expression of this single cDNA induces the fibroblasts to start to undergo myoblast differentiation (Davis *et al.*, 1987).

Recently a number of less technically demanding methods have been devised that use streptavidin and its strong affinity to biotin. The driver mRNA can be labelled with biotin and, after the hybridization, the hybrids and excess driver mRNA can be removed/subtracted by virtue of the biotin-streptavidin affinity. The subtraction

reaction can be passed over a streptavidin affinity column allowing only subtracted cDNAs to be eluted from the column which can then be cloned. A candidate gene for the classical mouse mutation *rd* retinal degeneration was cloned using this method to purify the subtracted cDNA (Bowes *et al.*, 1989). The hybrids and excess driver RNA can also be complexed to streptavidin-linked magnetic beads and removed using a magnet leaving the subtracted cDNA in solution. Perhaps the simplest method of removing the unwanted cDNAs and excess driver mRNA is the technique first suggested by Sive & St John (1988). Streptavidin is added to form streptavidin:biotin complexes, which can be removed by phenol-chloroform extraction. The protein complex remains on the liquid interface, allowing the subtracted cDNA to be separated by simply removing the aqueous layer. Back-extraction of the phenol allows complete recovery of all subtracted cDNAs (Sive & St John, 1988). Using this technique several cDNAs specific to cement gland development in *Xenopus* have been cloned and these cDNAs have been used as cement-gland-specific markers that allow the analysis of cement gland and adjacent neural ectoderm induction (Sive *et al.*, 1989). A *nanos*-like zinc finger gene localized to the vegetal cortex of *Xenopus* oocytes was also isolated using a subtracted cDNA probe generated by this technique (Mosquera *et al.*, 1993). It has also been used to clone a tissue-specific extinguisher-TSE-1 (Jones *et al.*, 1991), and F-spondin, a gene that promotes neural cell adhesion and neurite extension (Klar *et al.*, 1992). This elegant method was the subtraction method of choice for this project.

The major problem in subtractive hybridization is the minute amount of cDNA recovered which is difficult to handle without further loss. The use of PCR in subtractive hybridization techniques (Timblin *et al.*, 1990; Duguid & Dinaeur, 1990; Hla & Maciag, 1990) allows the amplification of the subtracted cDNA, so increasing ease of handling and cloning the material and subsequently the chance of isolating genes of interest.

In addition, the use of phage libraries has supplanted the use of plasmid libraries as there is greater efficiency in transformation (greater than 16-fold), screening and storing of phage libraries compared to plasmid libraries. The use of the new generation of λ vectors allows the efficiency of λ to be coupled to the advantages of plasmids for manipulating DNA. λZAP and λZAPII (Stratagene) are λ vectors with pBluescript plasmid cloned into the inessential 'stuffer' region of λ allow *in vivo* excision of the plasmid. When the λZAP is grown in the presence of an M13-derived helper phage, the pBluescript vector is excised from the λ vector, packaged in infectious M13 particles, and secreted into the media. The supernatant can then be used to infect bacteria.

λ phage can often rearrange or mutate the insertion site and restriction enzyme recognition sites around the insertion site preventing the subcloning of the desired cDNA, but *in vivo* excision circumvents these traditional problems. All the above points made λZAP II (Stratagene) the vector of choice for making a fetal kidney library.

Strategy

3.5 Making the 14.5 d.p.c. fetal kidney library

It was decided to make a total cDNA library from fetal kidney rather than a subtracted library. This should increase the likelihood of cloning large cDNA inserts, thus increasing the probability of isolating full length cDNAs. λZAPII was chosen as a vector for the reasons discussed above and also as pBluescript has many advantages as a plasmid. pBluescript is a multicopy plasmid with both T3 and T7 promoters which allows *in vitro* transcription of the insert on both strands. The plasmid is also complementary to a number of oligonucleotides primers for PCR amplification and sequencing of any cDNA inserts: the T7 promoter/primer, the M13 reverse universal primer and SK primer sequences are 5' to the cloning site on one strand and 5' on the other strand there are the T3 promoter/primer, the M13 forward universal primer and

KS primer sequences. The λZAP vector also allows immunoscreening of the library as the pBluescript vector it contains fuses the cDNA onto the *lac Z* gene, allowing translation of a proportion of the cDNAs open reading frame as *lac Z* fusion proteins. Hence λZAPII has the best features from λgt10 and λgt11 as it allows both DNA homology screening and antibody screening using the same library.

The mRNA used to make cDNA for the library and the subtraction was made from 14.5 d.p.c. fetal kidneys taken from a colony of Swiss mice maintained at the Western General Hospital Animal Breeding Facilities. This stage was chosen as fetal kidneys at this stage have all the stages of early nephron development including undifferentiated stem cells but do not yet express the genes needed for the physiological functions of the kidney. There was also the consideration of how much material each stage provided balanced against the fact that the study was primarily looking for genes important for the earlier inductive steps rather than the terminal differentiation stages of the proximal and distal tubules. To consider using earlier stages may have increased the number of fetal kidneys required by three to ten times. Using 14.5 d.p.c embryos, RNA was extracted from 600-700 14.5 d.p.c. fetal kidneys to isolate sufficient mRNA to make cDNA for both the library and one subtraction.

The cDNA library was made with 8 μg of poly A^+mRNA primed with oligo d(T) and the second strand made by the procedure of Gubler and Hoffman (1983) in which RNAse H nicks the RNA strand of the RNA:cDNA duplex formed by the first strand synthesis and DNA polymerase uses the nicks to replace RNA with DNA by nick translation. The cDNA was made blunt ended with the Klenow fragment of DNA polymerase and phenol:chloroform extracted. It was then purified and size-selected on a spun Sephacryl S-400 column (Pharmacia) that retains DNA molecules smaller than 400 bases. An *Eco RI/Not I* linker adaptor was ligated onto either end of the cDNA.

Linker

```
5' AATTCGCGGCCGC
       GCGCCGGCGp-5'
```

This linker has two main advantages. First, a *Not I* site in the linker allows cloned DNAs to be recovered intact by digestion with *Not I,* even if the cDNAs have internal *Eco RI* sites (*Not I* sites are extremely rare in mammalian cDNAs). Second, only one adapter can ligate to each cDNA terminus and the only product the adapter can form is dimers not multimers. The dimers can be removed from the cDNA by size selection, allowing efficient purification. The cDNA was then kinased with T4 polynucleotide kinase and phenol/chloroform extracted. The cDNA was size selected by another S-400 Sephacryl spin column then ligated to 1 μg of *Eco RI* restricted and phosphatased λZAPII. The ligation was then packaged using a Gigapack Gold reaction (Stratagene) as manufacturer's instruction and *E. coli* XL1-Blue were infected and plated on two 20cm by 20cm Cy agar plates, dilutions were made to allow the size of the library to be efficiently estimated.

The library had a complexity of 1.15 x 10^6 original clones. Plasmids from 12 random clones were *in vivo* excised to analyse their inserts and 12 phage plaques were picked into water, boiled, the agar precipitated and the phage lysate used for PCR amplification. For these clones, the average size of insert was 1.5 kb and the size range was from 2.2 kb -500 bases.

1 x 10^5 phage were plated on a large 20 cm by 20 cm Cy agar plate and filter lifts taken. The filters were screened with the mouse *Gap DH* (glyceraldehyde-3-phosphate dehydrogenase) cDNA, and 150 clones were positive. This demonstrates that the library had representative amounts of expected gene products in a random plating of the library.

3.6 The subtractive hybridization

The subtraction protocol is illustrated schematically in Figures 3.1 and 3.2. The cDNA synthesis was primed with a tagged oligo (dT) JM1 (Section 2.8.2), the RNA hydrolysed with NaOH and tailed with oligo (dC) using terminal transferase. An excessively long dCTP tail would cause false priming and termination of 2 nd strand

synthesis. The terminal transferase was calibrated in a control experiment to add $\alpha[^{32}P]$-dCTP to a DNA fragment of known size for a number of time points and the resultant samples were run alongside radioactively-labelled ϕ x *Hae III* markers on a non-denaturing polyacrylamide gel. From the results, the time required to allow the terminal transferase to add a dCTP tail of 20-25 nucleotides long was calculated.

mRNA from adult Swiss mice liver was used as driver RNA in the subtraction and biotinylated using photo-activatable biotin (PAB) as Sive & St John (1988). 30 µg of biotinylated adult liver mRNA was added to the tailed first-strand cDNA resuspended in 9 µl of 0.2% SDS, 2 mM EDTA, 500 mM NaCl, 50 mM HEPES, pH 7.6 (HB). The mixture was heated to 95°C at 2 min and hybridized at 65°C for 48 hours until a R_0t of 700 was reached. The R_0t is a measurement of the equilibrium and efficiency of the reaction (Hastie & Bishop, 1976). At this point there are two species formed, the cDNA which is common to the biotinylated adult liver mRNA and has formed hybrids, and the unique cDNA which remains as unhybridized single-stranded cDNA (Figure 3.1).

Streptavidin was added to the hybridization reaction to form a streptavidin:biotin complex and the reaction was phenol extracted. The streptavidin:biotinylated RNA:cDNA hybrid complexes are sequestered into the phenol interphase and the unique unhybridized cDNA in the interphase can be removed. The whole subtraction was repeated to maximise the efficiency of the subtraction. This should provide single-stranded kidney specific (mRNAs not present in the adult liver mRNA) cDNA.

The second strand of the cDNA was primed with 200 pg of oligonucleotide JM-2+dG (Section 2.8.2) and synthesised by incubation with Taq polymerase. Lone linker PCR (Ko *et al.*, 1990) was performed with 2 µg of oligonucleotide JM1 (Section 2.8.2) to amplify the subtracted material (Figure 3.2).

Figure 3.1

The figure shows a schematic illustration of the subtractive hybridization strategy described in the preceding pages. It shows how the mRNA from 14.5 d.p.c. fetal kidneys was used to make both a cDNA library (in λZAPII) and to make first strand cDNA which was then used for the subtraction protocol. The cDNA was synthesized using a oligo d(T) primer that also contained a known sequence. The RNA was removed by hydrolysis with sodium hydroxide and the first strand cDNA had a poly d(C) tail added using terminal transferase. The RNA chosen to subtract against was biotinylated using photoactivatable biotin and then hybridized with the first strand cDNA for 48 hours. This should produce hybrids of first strand cDNA and biotinylated mRNA and unique single-stranded cDNA.

In the figure:-

A squiggly line represents mRNA.

———— represents cDNA

represents the biotinylation of the RNA

represents the primer added to the cDNA

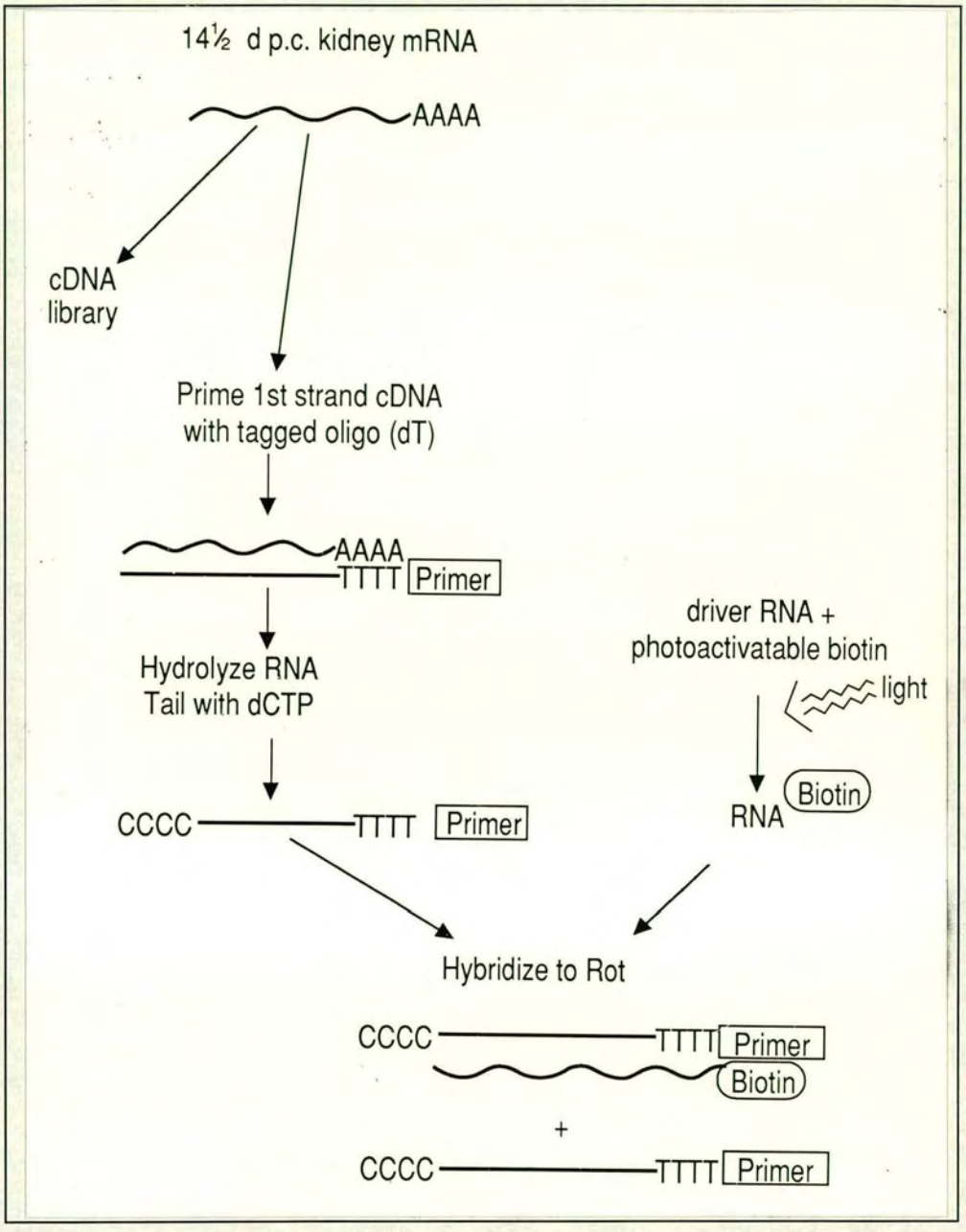

14½ d p.c. kidney mRNA
AAAA
cDNA
library
Prime 1st strand cDNA
with tagged oligo (dT)
AAAA
TTTT
Primer
Hydrolyze RNA
Tail with dCTP
CCCC
TTTT
Primer
driver RNA +
photoactivatable biotin
light
RNA
Biotin
Hybridize to Rot
CCCC
TTTT
Primer
Biotin
+
CCCC
TTTT
Primer

Figure 3.1

Figure 3.2

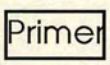
The figure shows a schematic illustration of the subtraction strategy described in the preceding pages. Streptavidin was added to allow a streptavidin/biotin complex to form. Phenol extractions then selectively remove the protein complexes, leaving the single stranded kidney specific cDNA. This subtraction was repeated twice to enhance the level of subtraction. A second cDNA strand was synthesized using oligo d(G) plus a known primer sequence. The cDNA was then amplified using the polymerase chain reaction before being used to identify potentially interesting clones from the cDNA library made previously.

In the figure:-

A squiggly line represents mRNA.

——— represents cDNA

represents the biotinylation of the RNA



represents the primer added to the cDNA

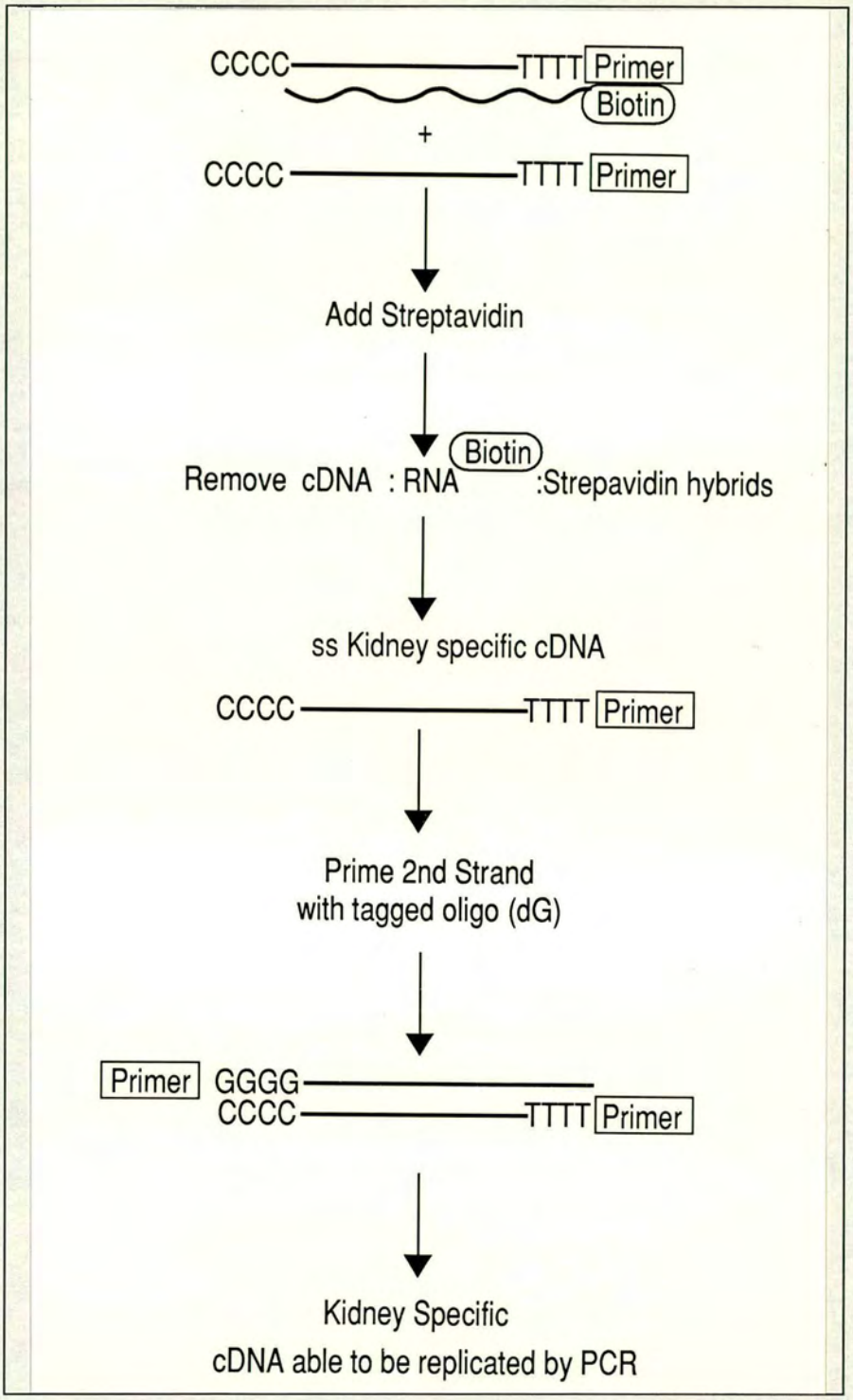

CCCC
TTTT
Primer
Biotin
+
CCCC
TTTT
Primer
Add Streptavidin
Biotin
Remove cDNA : RNA :Strepavidin hybrids
ss Kidney specific cDNA
CCCC
TTTT
Primer
Prime 2nd Strand
with tagged oligo (dG)
Primer
GGGG
CCCC
TTTT
Primer
Kidney Specific
cDNA able to be replicated by PCR

Figure 3.2

3.7 Analysis of the subtracted fetal kidney cDNA

The subtracted cDNA pool is renewable by PCR, hence alleviating the need to repeat the subtraction for subsequent uses to screen libraries or to categorize the composition of the cDNA pool. When analyzed by electrophoresis through an agarose gel and visualised by ethidium bromide staining (Figure 3.3a) it could be seen that the pool is very heterogeneous (as indicated by a smear of products) and includes material of relatively high molecular weight. When an aliquot of the subtracted cDNAs was radioactively labelled by random priming and hybridized to RNA from both fetal kidney and adult liver, the probe only hybridized weakly to a single band in the adult liver RNA, which was present at a much higher level in the fetal kidney RNA, but hybridizes to many different messages in the fetal kidney RNA (Figure 3.3b and c).

The above results suggest that the mouse fetal kidney library has a large number of original clones that have a large average insert size, and so this library was used to isolate fetal kidney specific clones. The results also suggest that the subtracted fetal kidney cDNA contains a wide range of cDNAs that preferentially hybridize to fetal kidney RNA in comparison to adult liver RNA. There was a low level of hybridization signal detected in the adult liver mRNA when the northern was analysed. The subtractive hybridization was originally performed in the belief that characterizing cDNA clones would be a relatively rapid procedure. If this had proven to be the case, the protocol would have been repeated to make the subtraction strategy more efficient. Original opinions on the most appropriate method of doing this were to choose a more specific tissue or tissues to subtract against. It may also have been benificial to increase the numbers of subtractions but the possibility of loss of cDNA preventing efficient recovery of the subtracted material tended to make this an unatractive proposition. Therefore the subtracted cDNA pool generated was used to screen the fetal kidney cDNA library to identify genes that are spatially and temporally regulated within the kidney.

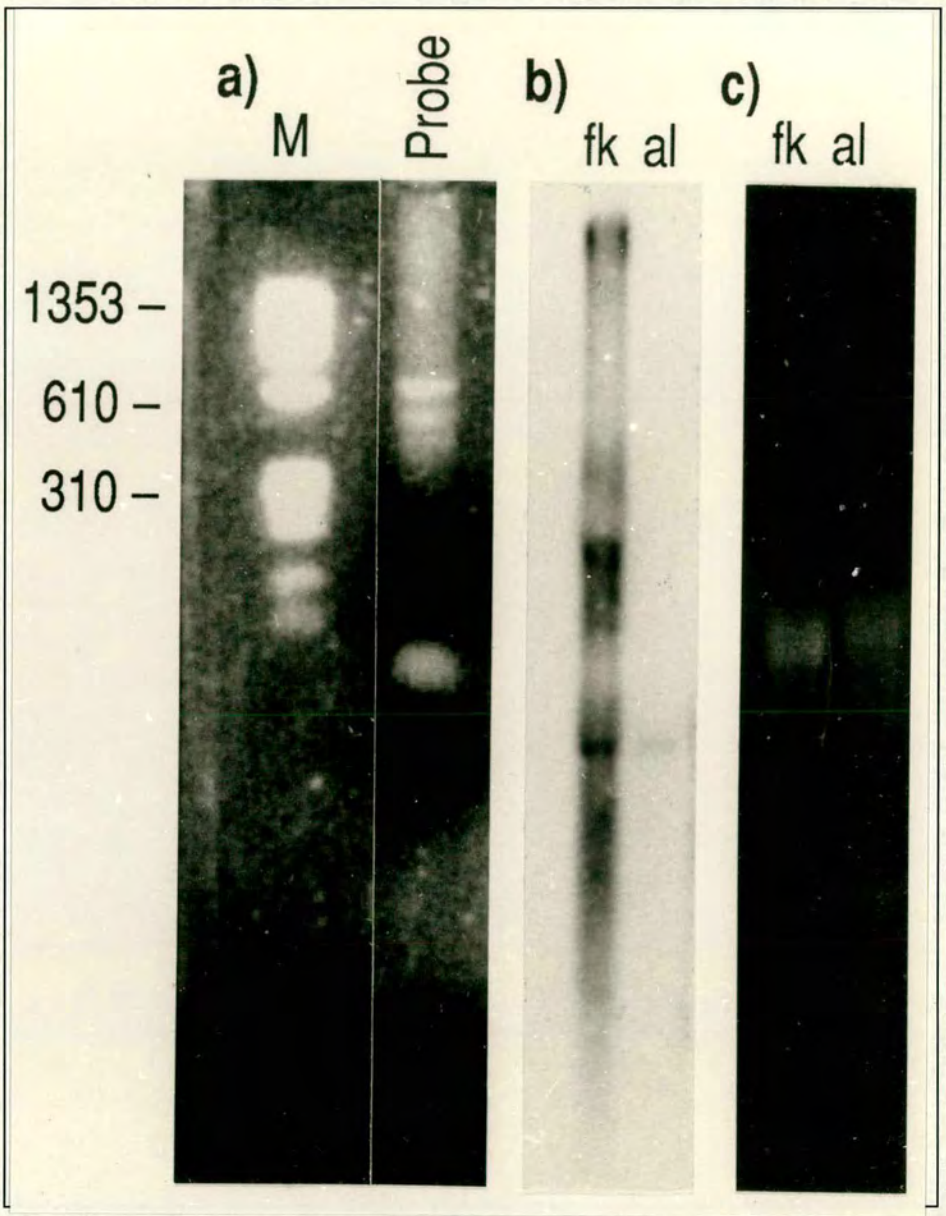


Figure 3.3

The PCR amplification products from the subtracted cDNA a) separated on an agarose gel alongside ϕx Hae III markers (M) and b) used to probe a northern blot of RNA from fetal kidney (f.k.) and adult liver (a.l.) compared to c) the ethidium bromide staining of the RNA gel.

Chapter 4

Isolation and characterization of subtracted clones

4.1 Screening with the subtracted probe

An aliquot of the subtracted cDNA was labeled by random priming and used to screen 5000 clones from the fetal kidney library plated at low density. 1250 phage plaques were screened on each of four 20 cm by 20 cm Cy agar plates (A-D). Ten positive clones were found and selected for further analysis. The clones were named after the phage plate they were isolated from and numbered sequentially as they were isolated *i.e.* clone B1, B2, C1, C2, C3. The size of the inserts contained in the positive phage was ascertained by PCR amplification of the phage insert using primers flanking the cDNA insertion site. The amplified inserts were visualized on an agarose gel by ethidium bromide staining (Figure 4.1) and ranged from 0.2 to 2.2 kilobases in size. Clone B1 proved to have no insert in the phage, the 'signal' on the film did not look encouraging but the phage was isolated to ensure all potentially interesting clones were studied.

4.2 Characterization of subtracted clones

Table 4.1 Table of phage isolated in the subtractive screen and the size of cDNA insert in each case.

Phage clone	Size of insert (kb)
B1	no insert
B2	1.7
C1	0.6
C2	1.1
C3	1.7
C4	1.3
C5	1.5
D1	1.7
D2	0.224
D3	2.2
D4	0.264

Each clone was sequenced using a primer from either side of the vector allowing a preliminary sequence analysis of the 5' and 3' sequences of each clone. The sequence

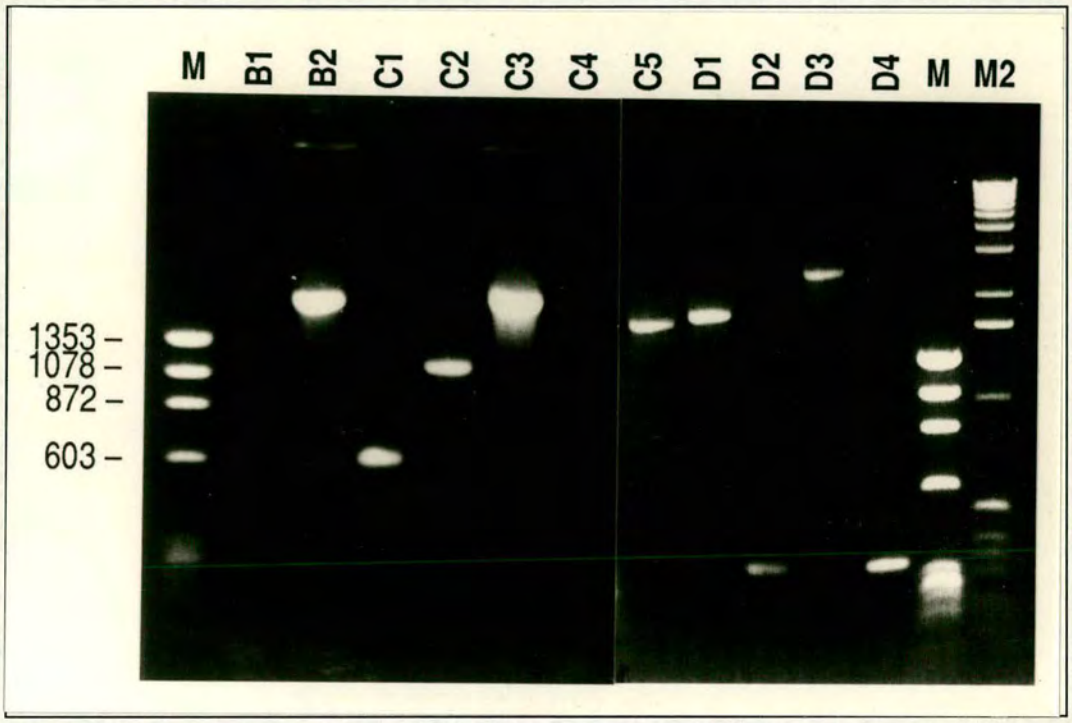


Figure 4.1

The PCR-amplified inserts from the positive clones B1-D4 identified by screening the fetal kidney library with the subtracted cDNA, separated on an agarose gel and visualized by ethidium bromide staining. M, ϕx *Hae III*-cut DNA; and M2, 1 kb markers (BRL). The photograph was taken from two separate gels, the first showing inserts B1 to C4 and the second showing C5 to D4. Both sections have been magnified to the same scale and aligned using the same markers on the two separate gels, as shown in the photograph.

obtained was analyzed using the UWGCG (University of Wisconsin Genetics Computer Group) programme FASTA to identify if the cDNA or a closely related sequence was in the GENBANK and EMBL databases (Devereux *et al.*, 1984). Comparisons were made of the DNA sequences as the preliminary sequence of some clones did not have an unambiguous open reading frame that would allow comparisons at the amino acid level. The FASTA programme compares the sequence with all the sequences currently in the databases; it then produces a list of the 50 highest similarities and the level and length of sequence similarity. As the number of sequences in the databases increases exponentially, the need to decide whether a homology is significant also increases. In this study in addition to the criteria of sequence similarity imposed by the UWGCG programmes, a number of criterion were considered in evaluating a sequence comparison.

- If a homology between the clone and another sequence was found at the DNA level, the predicted protein sequences of the clone and the sequence identified were then compared to see if the level of homology seen at the DNA level was also present at the amino acid level. The same or higher levels of protein homology should be seen if the homology is actually significant.

- Looking at the similarities in sequence at the DNA level the pattern of similar bases was examined, a pattern of 2 identical bases and one mismatched often means that the similarity may be significant. This pattern will allow protein homologies due to the third base redundancy of codons.

- Sequences from evolutionarily closer species are always more similar than sequences between evolutionary distant species. Therefore, any real similarity to an insect or yeast gene for which there were mammalian homologues would be expected to be lower than to the mammalian homologue.

- Finally, if the clone is similar to a coding region of another sequence, then the homology is more meaningful than if it had been similar to the 5' or 3 region of the sequence. Coding sequences are positively selected against changing, while non-coding regions are able to mutate and change their sequence with less effect on the organism. Therefore, homology to non-coding regions can happen by chance and are more likely to be spurious.

The cDNAs were also radioactively labeled and used to probe northern blots with total RNA from a variety of tissues. All the clones discussed in this chapter, C1, C2, C5, D2, D3 and D4 were shown to hybridize to a large number of RNA species giving a smeared signal on northern blots, indicating that these clones contain repetitive elements; however the signal was not present in adult liver mRNA. Initial sequence analysis of clones D2 and D4 indicated that the cDNAs were separate cloning events of the same parental mRNA and as such are discussed together in the following text; the complete cDNA had been sequenced for these clones, albeit on one strand only. All the northerns shown on this chapter were performed in the first half of the project using suboptimal RNA extraction methods. Consequently, in the following northern blots the loadings were not always equal, but the adult liver sample was always one of the more heavily loaded samples. Also the ovary sample is slightly degraded and the fetal liver, testes and spleen were badly degraded. Despite these samples being degraded and the uneven loadings, it became clear that these cDNAs were not giving a clear signal on northerns blots. As the other cDNAs seemed more interesting the work done with the cDNAs described in this chapter was minimal, the northerns were not repeated nor were genomic Southern blots performed. Clones B2, C3, and D1 were shown to encode the mouse 14-3-3 protein ε isoform and the analysis of these three clones will be described in detail in Chapter 5, while the 14-3-3 protein family to which they belong is reviewed in Chapter 6.

When Clone C4 was isolated the cDNA was PCR amplified and the cDNA insert was used to probe a northern blot with RNA from many different tissues (Figure 4.2). The expression profile looked very interesting with expression only in the fetal kidney and liver and the adult ovary. The ethidium bromide staining of this northern shows that whilst there was a lower amount of adult liver mRNA compared to the other samples, the cDNA also fails to hybridize to the brain RNA sample which contains an equally high amount of RNA compared to the ovary sample and higher than the other samples including the positive ones. This suggests that the hybridization is specific and not a concentration-dependent artifact. The ethidium bromide stain also shows that the smearing of the signal in the ovary and fetal liver samples is probably due to the presence of degraded RNA. Unforeseen circumstances led to the loss of the original phage plaque and the PCR-amplified insert. Attempts to perform *in vivo* excision yielded a plasmid which later proved to be a contaminant being worked on elsewhere in the laboratories. Unfortunately all that could be found from this clone was a tantalizingly interesting northern blot.

Sequence from the clones

The table below details each clone and the length of sequence in nucleotides which was obtained with each primer (SK and KS) to give sequence information for clone identification.

Table 4.2

Clone	Length of sequence (bp)	
	Primer SK	Primer KS
C1	222	253
C2	184	209
C5	210	199
D2	218	205
D3	172	210
D4	196	203

Figure 4.2

Northern blot of 10 μg of RNA from various mouse tissues probed with the PCR-amplified insert from clone C4.

14d f.k., 14.5 d.p.c. fetal kidney; 17d f.k., 17.5 d.p.c. fetal kidney; kidney, adult kidney; liver, adult liver; ovary; brain; Sal gland, salivary gland; heart; lung; s.i., small intestine; EB26, EB26/10a, embryonic carcinoma cell line; fetal liver; testes; spleen; M, RNA markers (BRL).

All tissues are adult unless stated otherwise.

The lower picture shows the 18S and 28S rRNA from each sample visualized by ethidium bromide staining. The lanes correspond to the samples in the upper blot. From the ethidium bromide staining it can be seen that the adult liver sample is comparatively under loaded (on the order of two times lower than the other samples) but a detectable amount of RNA is present. The ethidium bromide staining also shows that the ovary and fetal liver samples contain degraded RNA which accounts for the smeared signal seen in the ovary and fetal liver samples.

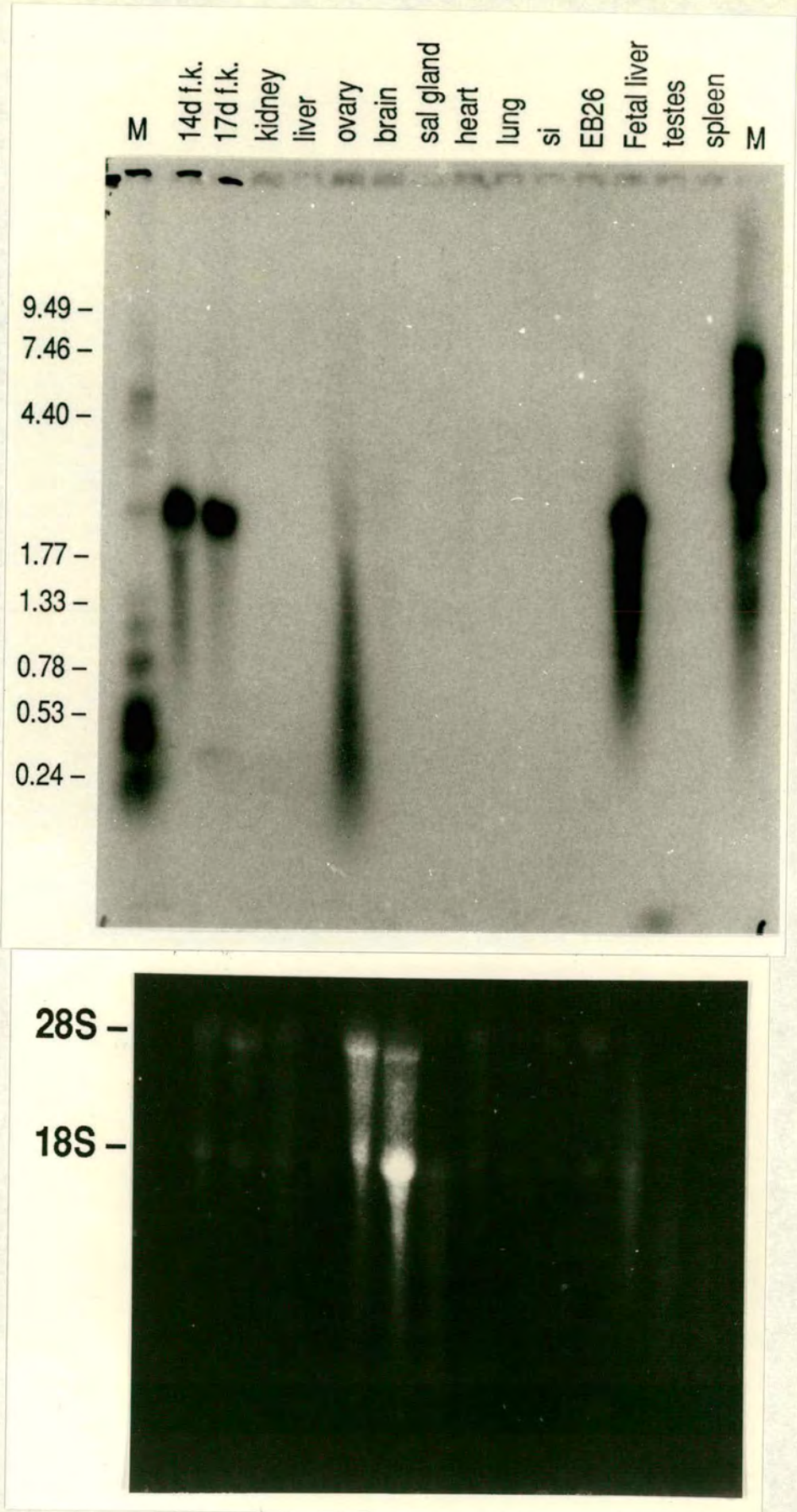

M
14d f.k.
17d f.k.
kidney
liver
ovary
brain
sal gland
heart
lung
si
EB26
Fetal liver
testes
spleen
M
9.49 –
7.46 –
4.40 –
1.77 –
1.33 –
0.78 –
0.53 –
0.24 –
28S –
18S –

Figure 4.2

4.3 Clone C1

Sequence of clone C1 using the SK primer

```
  1  CGGCCGCCTA GAAATGGTTA CATCACTAGA CTTTAAAAAC TTTGTATCTT
 51  ACCCAAAAAG AATTGTTAAT AAACAAAATT GGGTGCCTGT TGGCTTCATC
101  TGTTCTGTCA AAACTTTATG ATACAATGAT GGGATTTTAT TCTAAATCAG
151  AATGTGGCAT TATACGTTCC AATTTGTGTG GAGGAGTTGC TAAATTTATT
201  ATTAGGTATT GATAATCTAA AAA
```

Sequence of clone C1 using the KS primer

```
  1  TTCGCGGCCG GTTTTTTTCT TGGTTTGGTT TGGTTTGGTT TGGTTTGGTT
 51  TTTTGAGACA GCGTTTCTCT GTGTAGCCCT GGCTGTCTTG GAACTCTGTA
101  GACCAGGCTG GCCTTGAACT CAGAAATCTG CCTGCCTCTG CCTCCCAAGT
151  GCTAGGATTA AAGGCGTGTG NCACTACTGC CCAGTGAAAA TTTCTTTTTA
201  AATATAAAAT AATAAGTAGT GAGGTGATGG ATTTGTAAAG AAGTAGTGTT
251  ATA
```

The sequence obtained using the SK primer was analysed by the FASTA programme. No homologies of any significance were found. From the other end of the cDNA, sequenced using the KS primer, homology to the B1 repeat consensus sequence can be found (Figure 4.8). but no other significant homologies. When the insert was radioactively labelled and used to probe a northern blot with a selection of mouse tissues, a smeared signal was present in fetal kidney, ovary, EB26/10a (an embryonic carcinoma cell line), fetal liver and testes RNA, and at a lower level in brain heart, lung and small intestine RNA but no signal was present in adult liver (Figure 4.3). Examination of the ethidium bromide staining of the 18S and 28S rRNA shows that the RNAs are loaded relatively equally and that the fetal liver, testes and spleen samples were degraded as can also be seen in the hybridization pattern.

Figure 4.3

Northern blot of 10 μg of RNA from various mouse tissues probed with the PCR-amplified insert from clone C1. The 18S and 28S r RNA positions are shown.

14 d f.k.,14.5 d.p.c. fetal kidney; 17d f.k., 17.5 d.p.c. fetal kidney; a.k., adult kidney; a.l., adult liver; ov, ovary; brain; s.g., salivary gland; heart; lung; s.i., small intestine; EB26/10a, embryonic carcinoma cell line; f. liver, fetal liver; testes; spleen; M, RNA markers (BRL).

All tissues are adult unless stated otherwise.

The lower photograph shows the 18S and 28S rRNA from each sample visualised by ethidium bromide staining. The ethidium bromide staining shows that the RNA samples are relatively equally loaded and that the fetal liver, testes and spleen RNA samples were degraded.

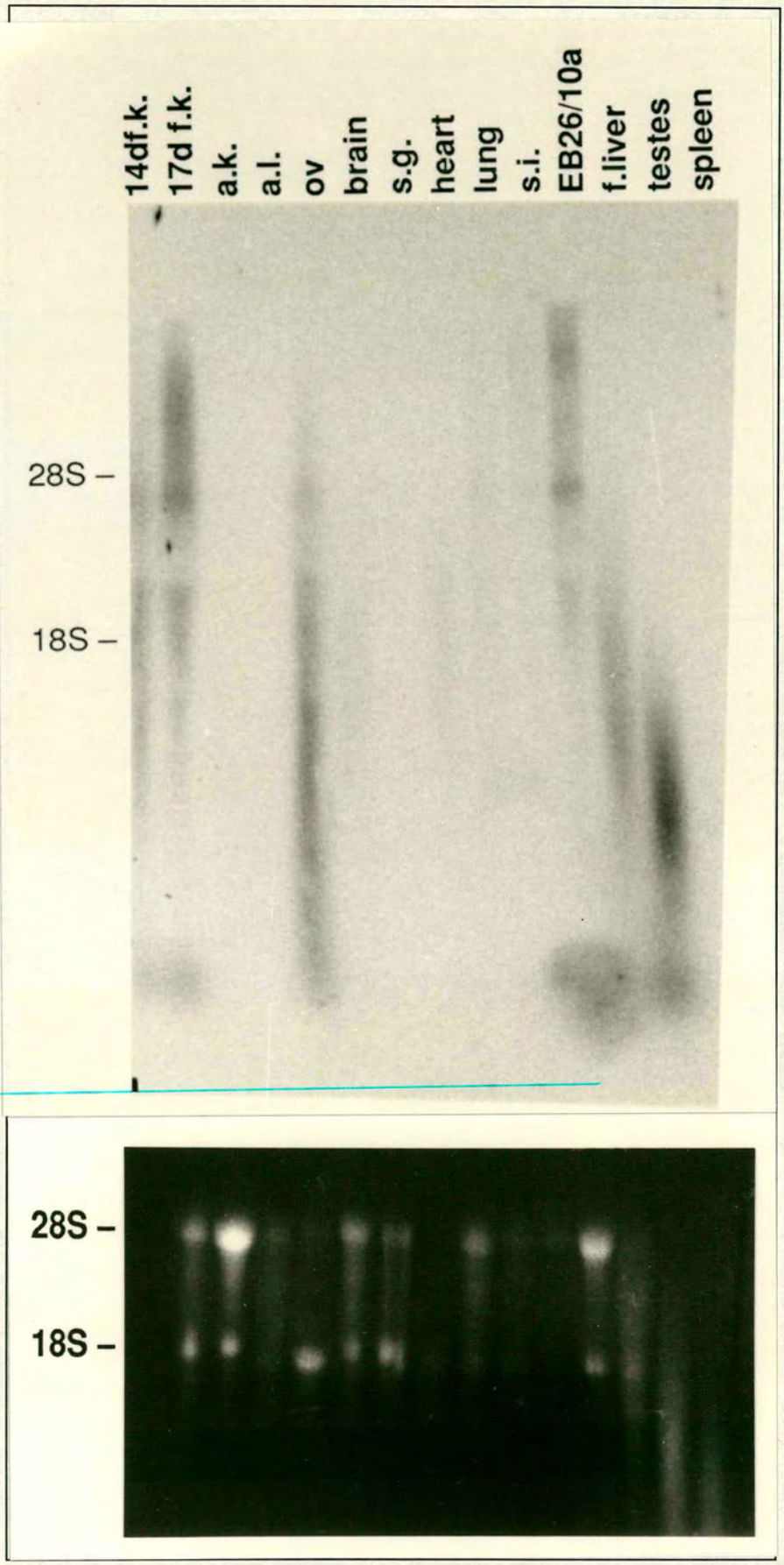

14df.k.
17d f.k.
a.k.
a.l.
ov
brain
s.g.
heart
lung
s.i.
EB26/10a
f.liver
testes
spleen
28S –
18S –
28S –
18S –

Figure 4.3

4.4 Clone C2

Sequence of clone C2 using the SK primer

```
  1  GGCCGCTAGA AGCACTGGAC AGCACGAGGA CTCTAGTAAC ACTACTGTCT
 51  TGAATGTGCA CTTGTTTCAC TATGGTAACC TTAGTTACCA GAAAAACCTA
101  TTTAAAAAGA AAGAAATCAA TACCAGGCAT GGTAGTGTAC TCCCTTAATC
151  CAGCAGTCAG CAGGCAGAGA TCTATGTCAG GTCG
```

Sequence of clone C2 using the KS primer

```
  1  CAGAGTCTCG CTCTGTAGCC TTGNTGTCCT GGAACTCACT CTGTAGACCA
 51  GCTGNCCTCC AACTCTCAGA GTCTCCTNCC TCTACCTCCC AAGTGCTAGC
101  ATCAAAGGTG TGTGCCACGA CCCCCTTCAA ATAGTTAACN GAACATACTA
151  AACCAACATT TAAAATTGAG TTGTAGCTTN NTCCCCACTT AGGTTTTAGA
201  AGNGCCCC
```

The sequence of clone C2 using the SK primer has an imperfect B1 repeat throughout the majority of the sequence and no homology of any significance to any other sequence. The sequence from the KS primer contains a B1 repeat in the first 125 bases. The B1 repeat sequence was then removed from the sequence and the remaining 83 bases analysed again. The further analysis revealed no significant homology along the whole sequence and the homology found (*i.e.* 67% identity in 73 bases overlap with hsc71 [major heat-shock-protein cognate-like gene] from rainbow trout testes and 66.7% identity in 75 bases with *Paracentrotus lividus* mitochondrial DNA) was not convincing and was disregarded. The cDNA was radioactively labelled and used to probe a northern blot and found to hybridize to RNA in fetal kidney, fetal liver, ovary, the cell line EB26/10a and salivary gland and testes at a lower level but not to RNA in adult liver (Figure 4.4). Examination of the ethidium bromide staining of the 18S and 28S rRNA shows that the salivary gland and testes mRNA was degraded, but all samples were relatively equally loaded. The fetal liver and ovary sample levels were five to ten fold higher than the other samples which

may explain the higher signal seen for these samples whilst salivary gland and testes samples are about two fold lower than the other samples. After taking into consideration the minor differences in loading levels the above observations on the expression levels are still relevant and it is still clear that the cDNA is not expressed in adult liver and satisfies the selection criterion. This northern blot was technically imperfect but it confirmed that the clone C2 contained a B1 repeat and did not hybridize to a single mRNA, and, therefore, provided enough information to decide to discard this clone on the criterion of it containing a repetitive sequence.

Figure 4.4

Northern blot of 10 μg of RNA from various mouse tissues probed with the PCR-amplified insert from clone C2. The 18S and 28S rRNA positions are shown.

14 d f.k.,14.5 d.p.c. fetal kidney; 17d f.k., 17.5 d.p.c. fetal kidney; a.k., adult kidney; f. liver, fetal liver; a.l., adult liver; ov, ovary; brain; s.g., salivary gland; testes; s.i., small intestine; EB26/10a, embryonic carcinoma cell line; M, RNA markers (BRL). All tissues are adult unless stated otherwise.

The lower photograph shows the 18S and 28S rRNA from each sample visualised by ethidium bromide staining. The first six samples were loaded satisfactorily but the next two samples, are supposed to be brain and salivary gland RNA but did not load into the gel and can be considered as empty wells. The next two samples, are salivary gland and testes RNA and can be seen to be degraded. The next two lanes were broken and so not used and are not labelled in the upper photograph. The last two samples correspond to the last two lanes of the upper photograph.

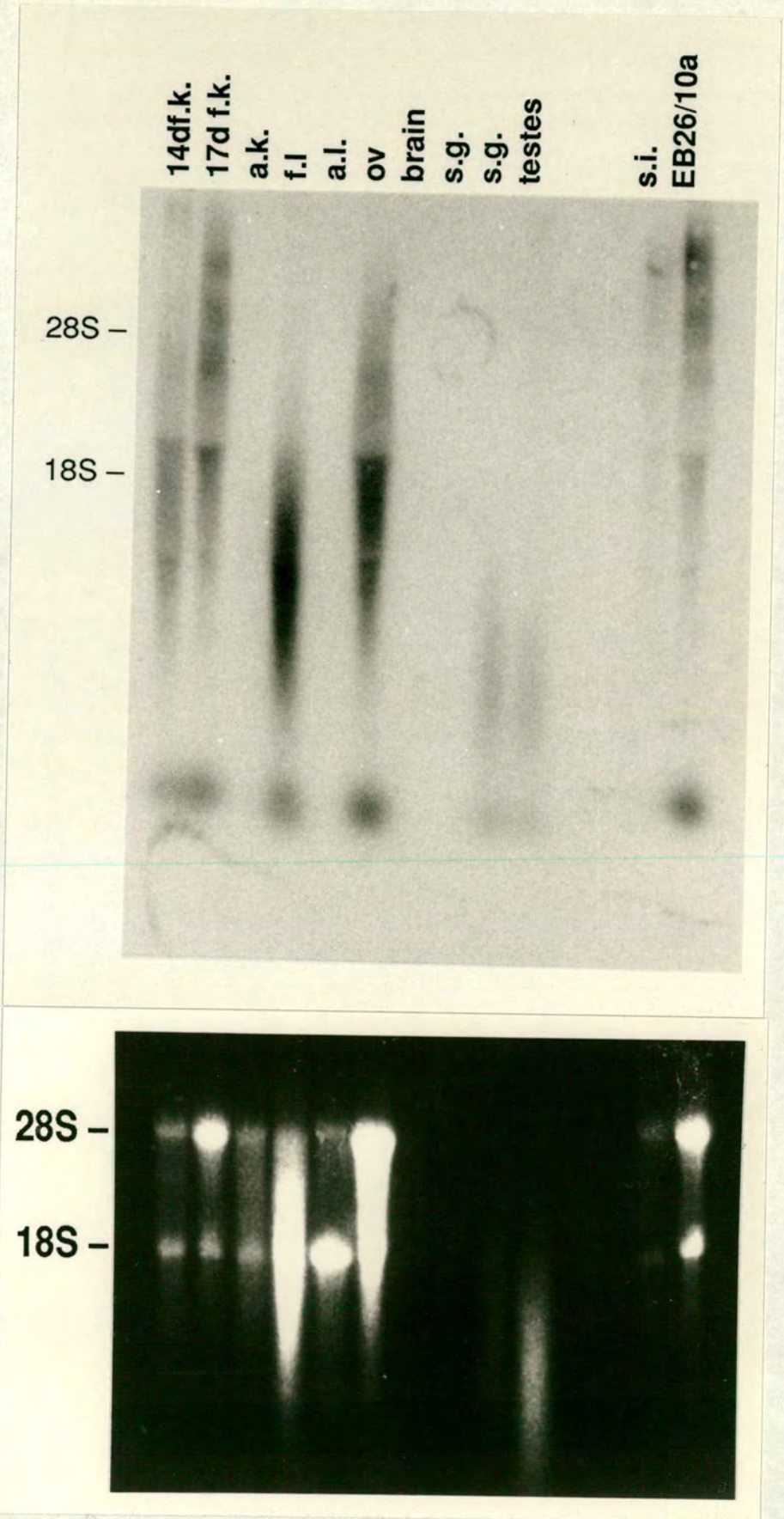

14df.k.
17d f.k.
a.k.
f.l
a.l.
ov
brain
s.g.
s.g.
testes
s.i.
EB26/10a
28S –
18S –
28S –
18S –

Figure 4.4

4.5 Clone C5

Sequence of clone C5 using the SK primer

```
  1  CGGCCTTGCC CAGGAACGTT TCAAATGAGT CAAAAGTTAA CTCCCGGGCC
 51  TAAGGCTAGA GGTGCAGTTG AGGTGGAGCT TAGCTGGCAT GCAGGAGGCT
101  CTGGGTTTGT CCTCAGCTCT GGGAGAAGGG GGAAAGGCAG AATCAAACAT
151  GGCTTGTAAG AAGCCTATTA TAGTATAGTG AGATGCAGGT TCCGGAACTA
201  GCCAGAGAA
```

Sequence of clone C5 using the KS primer

```
  1  CGCTCTCTGT GTAGCCCTGG CTGTCCTGGA ACTCACTTTG TAGACCAGGC
 51  TGGCCTTGAA CTCAGAAATC TGCCTGCCTC TGCCTCCCAA GTGCTGGGAT
101  TAAAGGCCCC ACCACTGCCC AGCTACCAGC AGTTTCTTAC TGGCATGCAG
151  CCATCAAGTT CTGGTGGAAG CCACTGTGTT CTGTGAGTTA AAGGAAAGG
```

The sequence of C5 from the SK primer gives no meaningful homologies. The sequence with the KS primer contains 125 bases of a B1 repeat, this was removed and the rest of the sequence was checked against the data-base. The search still gave no meaningful homologies. When the cDNA was radioactively labelled and used to probe a northern blot it hybridized to RNA in fetal kidney, ovary, the cell line EB26/10a, testes, spleen and fetal liver as well as brain, heart, lung and small intestine at lower levels but not to RNA in adult liver (figure 4.5). Examination of the ethidium bromide staining of the 18S and 28S rRNA shows that the salivary gland and testes mRNA was degraded, but all samples were relatively equally loaded allowing estimation of expression levels directly from the autoradiograph.

4.6 Clone D2 and D4

Sequence of clone D2 using the SK primer

```
  1  TTGGAAAGGC CAAAAGACTT TTGACTTTTA CTGAACAAAA CAAAAGTAAG
 51  TAATGTAACA CCAAACAAAA ACAACACAGT ATCACAAACA ATGGTTCGTC
101  AGCAGCAGTC GCTATGCCTT TAATCCCAGC ACTCGGGAGG AAGAGGCAAG
151  CGGATTTCTG AGTTGGAGGC AGCCTAGTCT ACAAAGTGAG TTCCAGGACA
201  GCCAGGGCTA TACAGAGAA
```

Sequence of clone D2 using the KS primer

```
  1  GGTTTCTCTG TATAGCCCTG TCCTGGAACT CACTTTGTAG ACTAGGCTGA
 51  CCTCCAACTC AGAAATCCGC TTGCCTCTGC CTCCCGAGTG CTGGGATTAA
101  AGGCATGCGC TGCTGCTGCG AACCATTGTT TGTGATACTG TGTTGTTTTT
151  GTTTGGTGTT ACATTACTTA CTTTTGTTTT GTTCAGTAAA AGTCAAAAGC
201  TTTTT
```

The two sequences from either end of the clone D2 could be compared to give the complete sequence cloned. The sequences were 98% identical (with sequencing errors causing the 2% error rate) and a final sequence of 223 bases was found of which 209 bases of which had been sequenced on both strands.

Sequence of D2

```
  1  TTGGAAAGGC CAAAAAGACT TTTGACTTTT ACTGAACAAA ACAAAAGTAA
 51  GTAATGTAAC ACCAAACAAA AACAACACAG TATCACAAAC AATGGTTCGT
101  CAGCAGCAGT CGCTATGCCT TTAATCCCAG CACTCGGGAG GTAGAGGCAA
151  GCGGATTTCT GAGTTGGAGG TCAGCCTAGT CTACAAAGTG AGTTCCAGGA
201  CAGCCAGGGC TATACAGAGA AACC
```

Sequence of clone D4 using the SK primer

```
  1  GCGGCCGCCT AGAAAGGTAC TGGCATAGGG ATCACTGGAA AATGGAAAGG
 51  CCAAAAGACT TTTGAGCTTT TACTGAACAA AACAAAAGTA ATGTAAGTAA
101  CACCAAACAA AAACAACACA GTATCACAAA CAAGGTTCGC AGCAGCAGCG
151  CATGCCTTTA ATGGCCCAGC AGCTCGGGAG GCAGAGGCAA GCGGATTTCT
```

Sequence of clone D4 using the KS primer

```
  1  GGTTTCTCTG TATAGCCCTG GCTGTCCTGG AACTCACTTT GTAGACTAGG
 51  CTGACCTCCA ACTCAGAAAT CCGCTTGCCT CTGCCTCCCG AGTGCTGGGA
101  TTAAAGGCAT GCCTGCTGCT GCGAACCATT GTTTGTGATA CTGTGTTGTT
151  TTTGTTTGGT GTTACATTAC TTACTTTTGT TTTGTTCAGT AAAAGTCAAA
201  AGTCTTT
```

The sequence from both ends of the D4 clone were compared and shown to overlap so that the complete sequence of the D4 clone could be shown. The sequences were 98 % identical (with sequencing errors causing the 2% error rate) and a final sequence of 265 bases was found of which 165 bases had been sequenced on both strands.

Sequence of clone D4

```
  1  GCGGCCGCCT AGAAAGGTAC TGGCATAGGG ATCACTGGAA AATGGAAAGG
 51  CCAAAAGACT TTTGAGCTTT TACTGAACAA AACAAAAGTA ATGTAATGTA
101  ACACCAAACA AAAACAACAC AGTATCACAA ACAACGGTTC GCAGCAGCAG
151  GCATGCCTTT AATGGCCCAG CAGCTCGGGA GGCAGAGGCA AGCGGATTTC
201  TGAGTTGGAG GTCAGCCTAG TCTACAAAGT GAGTTCCAGG ACAGCCAGGG
251  CTATACAGAG AAACC
```

When clones D2 and D4 sequences were compared it was obvious that they were the same sequence. They were 98.6% identical over 228 nucleotides.

Comparison of clones D2 and D4

```
D2      2 TGGAAAGGCCAAAAAGACTTTTGA.CTTTTACTGAACAAAACAAAAGTAA 50
          |||||||||| |||||||||||| |||||||||||||||||||||||||
D4     43 TGGAAAGGCC.AAAAGACTTTTGAGCTTTTACTGAACAAAACAAAAGTAA 91
                   .          .          .          .          .
       51 .GTAATGTAACACCAAACAAAAACAACACAGTATCACAAACAATGGTTCG 99
           |||||||||||||||||||||||||||||||||||||||||| ||||||
       92 TGTAATGTAACACCAAACAAAAACAACACAGTATCACAAACAACGGTTCG 141
                   .          .          .          .          .
      100 TCAGCAGCAGTCGCTATGCCTTTAAT..CCCAGCA.CTCGGGAGGTAGAG 146
           ||||||||| |   |||||||||||  ||||||| |||||||| ||||
      142 .CAGCAGCAGGC...ATGCCTTTAATGGCCCAGCAGCTCGGGAGGCAGAG 187
                   .          .          .          .          .
      147 GCAAGCGGATTTCTGAGTTGGAGGTCAGCCTAGTCTACAAAGTGAGTTCC 196
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      188 GCAAGCGGATTTCTGAGTTGGAGGTCAGCCTAGTCTACAAAGTGAGTTCC 237
                   .          .
      197 AGGACAGCCAGGGCTATACAGAGAAACC 224
          ||||||||||||||||||||||||||||
      238 AGGACAGCCAGGGCTATACAGAGAAACC 265
```

The preceding comparison aligns the complete sequence of clone D2 (top sequence) to the sequence of clone D4 and shows that clone D2 is completely enclosed within clone D4. The D2 and D4 sequences show homology to the B1 repeat along most of the length of their sequences. The D2 cDNA was radioactively labelled and used to probe a northern blot with RNA from different tissues, it hybridized to RNA in fetal kidney, fetal liver, ovary, EB26/10a, an embryonic carcinoma cell line and salivary gland at a lower level but not to RNA in the adult liver (Figure 4.6). Examination of the ethidium bromide staining of the 18S and 28S rRNA shows that the salivary gland and testes mRNA was degraded, but all samples were relatively equally loaded. The fetal liver and ovary sample levels were five to ten fold higher than the other samples which may explain the higher signal seen for these samples whilst salivary gland and testes samples are about two fold lower than the other samples. After taking into consideration the minor differences in loading levels the above observations on the expression levels are still relevant and it is still

clear that the cDNA is not expressed in adult liver and satisfies the selection criterion. This northern blot was technically imperfect but it confirmed that the clone D2 (and D4) contained B1 repeats and did not hybridize to a single mRNA, and, therefore, provided enough information to decide to discard these clones on the criterion that they contained repetitive sequences.

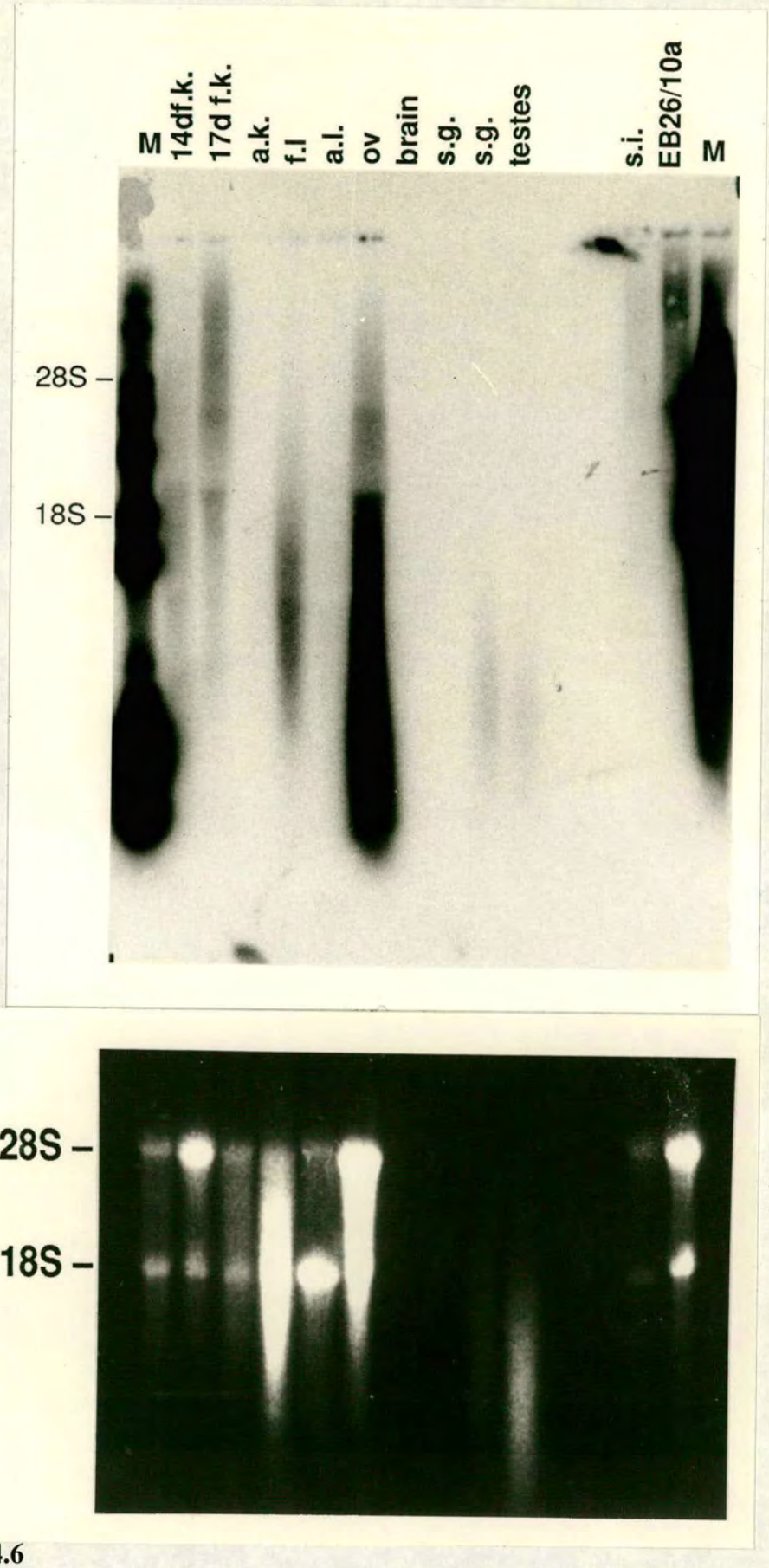

M
14df.k.
17d f.k.
a.k.
f.l
a.l.
ov
brain
s.g.
s.g.
testes
s.i.
EB26/10a
M
28S –
18S –
28S –
18S –

Figure 4.6

4.7 Clone D3

Sequence of clone D3 using the SK primer

```
  1  CGAAGCAGTA CACAAGGCCA TGTAGCCATA ATCGATTAGT GTGGAGGTTA
 51  CTACTACTAC ATGGCCTTGT GTACTGCTTC TGGGTACATA CATTACTGTC
101  TAGGCATGAT CTGAACAAAG ATGACACCAG CAAGCATGCT TAAGGTCACG
151  GGAGACCARA TAAGACGTCT AG
```

Sequence of clone D3 using the KS primer

```
  1  GCGATTGCGG CCGAAACAAA GCACCAGAGA AAGATACTCC TAACTTATTA
 51  GCAGGAGAAG CCTGGAGGAG CCAGGAGGAG GCAAAGAGCC AGACTTTAGG
101  GAAAGTAATG TTGAAGACAC GGGGTTGGTG ATGGGTTGAG GTTCAGAAGT
151  CAGAGAAATG GACAAAAAGC AATGGTCTGA GAGGGTGACG GTAACTNAGG
201  AAAAGGAAAC
```

The sequence from the SK primer shows no significant homology to any sequence in the database. The only homology of possible significance is a similarity of 57.4% along 155 bases to a expressed sequence tag in rice, (D15951), but the protein that this cDNA encodes is unknown. The sequence from the KS primer similarly shows no significant homology to any sequence in the database. Although the preliminary sequence of clone D3 did not find any homology to repeat sequences, the lack of a specific signal when the cDNA was radioactively labelled and used to probe a northern suggests that there is probably a repeat sequence in this cDNA. When the clone was used to probe a northern blot it hybridized to a smear of RNA in fetal kidney, fetal liver, ovary, the cell line EB26/10a, testes, and small intestine but not to RNA from adult kidney, adult liver, brain, heart and lung (Figure 4.7). The ethidium bromide staining of the gel shows that the gel was evenly loaded (allowing direct comparison of expression levels from the autoradiograph) and that the fetal liver and testes RNA samples were partially degraded.

Figure 4.7

Northern blot of 10 μg of RNA from various mouse tissues probed with the PCR-amplified insert from clone D3. The 18S and 28S rRNA positions are shown.

14 d f.k.,14.5 d p.c. fetal kidney; 17 d f.k., 17.5 d.p.c. fetal kidney; a.k., adult kidney; a.l., adult liver; ov, ovary; brain; s.g., salivary gland; heart; lung; s.i., small intestine; EB26/10a, embryonic carcinoma cell line; f. liver, fetal liver; testes.

All tissues are adult unless stated otherwise.

The lower photograph shows the 18S and 28S rRNA from each sample visualised by ethidium bromide staining. The ethidium bromide staining shows that the RNA samples are relatively equally loaded and that the fetal liver, testes and spleen RNA samples were degraded.

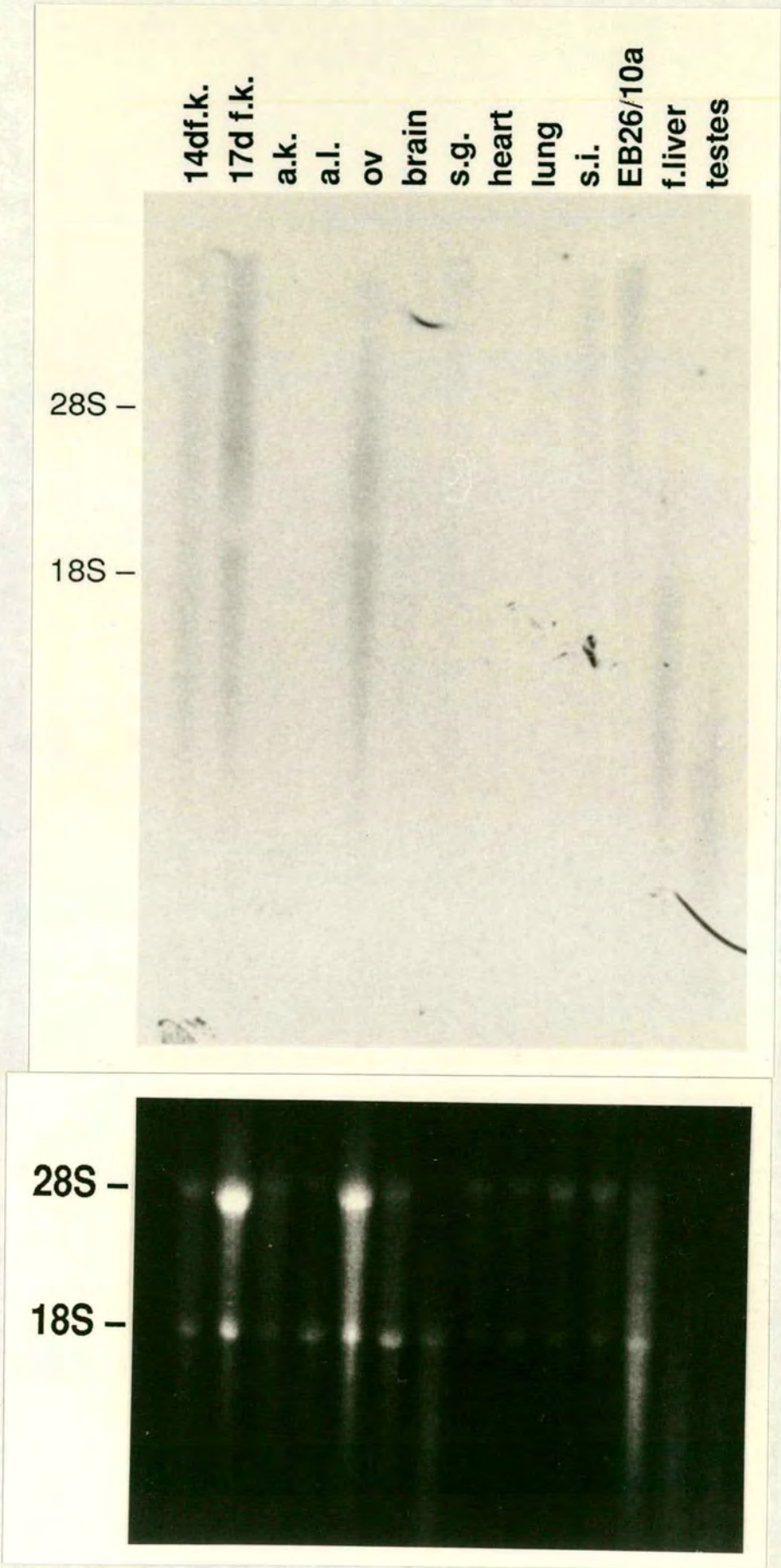

14df.k.
17d f.k.
a.k.
a.l.
ov
brain
s.g.
heart
lung
s.i.
EB26/10a
f.liver
testes
28S –
18S –
28S –
18S –

Figure 4.7

Comparisons of the B1 repeat element

Clones C1, C2, C5,and D2 all contained regions of homology to the B1 repeat sequence. The level of similarity is shown in Figure 4.8 below that compares all the sequences to the consensus B1 repeat.

Figure 4.8

```
          1                                                  50
   c1ks          TAG TGNCACACGC CTTTAATCCT AGCACTTGGG AGGCAGAGGC
   c2sk  CCAGGCATGG TAGTGTACTC CCTTAATCC* AGCAGTCAGC AGGCAGAG
   c2ks    GGGGGTCG TGGCACACAC CTTTGATGCT AGCACTTGGG AGGTAGAGGC
   c5ks                 GGGGC CTTTAATCCC AGCACTTGGG AGGCAGAGGC
   d2sk                   ATGC CTTTAATCCC AGCACTCGGG AGGAAGAGGC
   d2ks               GCGCATGC CTTTAATCCC AGCACTCGGG AGGCAGAGGC
   d4ks                 GGCATGC CTTTAATCCC AGCACTCGGG AGGCAGAGGC
 b1conc  CCGGGCATGG TGGTGCATGC CTTTAATCCC AGCACTCGGG AGGCAGAGGC

          51                                                 100
   c1ks  AGGCAGATTT CTGAGTTCAA GGCCAGCCTG GTCTACA*** *GAGTTCCAA
   c2ks  AGGAGACTCT GAGAGTTGGA GGCCAGCCTG GTCTACAGAG T*AGTTCCAG
   c5ks  AGGCAGATTT CTGAGTTCAA GGCCAGCCTG GTCTACAAAG TGAGTTCCAG
   d2sk  AAGCGGATTT CTGAGTTGGA GG*CAGCCTA GTCTACAAAG TGAGTTCCAG
   d2ks  AAGCGGATTT CTGAGTTGGA GGTCAGCCTA GTCTACAAAG TGAGTTCCAG
   d4ks  AAGCGGATTT CTGAGTTGGA GGTCAGCCTA GTCTACAAAG TGAGTTCCAG
 b1conc  AGGCGGATTT CTGAGTTCGA GGCCAGCCTG GTCTTCAGAG TGAGTTCCAG

          101                          130
   c1ks  GACAGCCAGG GCTACACAGA GAAACGCTGTCT
   c2ks  GACAGNCAAG GCTACA**AG CAACTCT
   c5ks  GACAGCCAGG GCTACA**CA GAGAC
   d2sk  GACAGCCAGG GCTATA**CA GAGAA
   d2ks  GA****CAGG GCTATA**CA GAGAAACC
   d4ks  GACAGCCAGG GCTATA**CA GAGAAACC
 b1conc  GACA*CCAGG GCTACA**GA GAAACCCTGTCT
```

4.8 Discussion

Analysis of the 6 clones discussed in this chapter gave sequences which contained regions with high homology to the B1 repeat. The B1 sequence exists in the mouse genome as transcribed monomers. The B1 repeat belongs to the *Alu* family of SINEs, short interspersed elements, which also include the human *Alu* and murine B2 repeats. The *Alu* repeat is a 280 bp imperfect dimer of the B1 sequence implying the two repeats probably have a common evolutionary origin. The B1 SINE has undergone amplification since the evolutionary divergence of rodents and there are now 40,000-80,000 copies of the B1 sequence scattered throughout the mouse genome. The amplification of the sequence is thought to have involved a RNA intermediate (Maraia, 1991). It has been argued that the lack of poly A or 3' oligo U tracts means that the B1 RNA found and sequenced are unlikely to represent transposition intermediates (Maraia, 1991). It had previously been thought that *Alu* sequences, the human equivalent of the B1 repeat, were no longer active as retrotransposons but recent studies have shown that *Alu* sequences are still active. A young subfamily of *Alu* repeats, has recently transposed within the human genome (Batzer *et al.*, 1990) and an *Alu* repeat was found in the coding region of the neurofibromatosis 1 (NF1) gene causing neurofibromatosis (Wallace *et al.*, 1991). It is possible that the B1 repeat is also still acting as a retrotransposon.

4.8.1 Discovery of the B1 repeat

The B1 sequence was first identified by virtue of its ability to form long double-stranded (ds) regions of RNA. Nuclear pre-mRNA (hnRNA [heterogeneous]) can form long double stranded regions which can be isolated from nuclear RNA preparations by extensive digestion with a mixture of single-strand specific ribonucleases (RNAases T1 and A) (Kramerov *et al.*, 1977). Long ds RNA which remains undigested can be separated into two fractions by gel electrophoresis. One fraction, the dsRNA-A contains RNAs of 300-800 base pairs long which are

incapable of folding back after strand separation. The other fraction, dsRNA-B, contains RNAs of 100-200 base pair long fragments which are capable of folding back after pre-mRNA melting (Kramerov *et al.*, 1977). dsRNA-B is transcribed from reiterated DNA sequences and reassociation kinetics have shown that it represents a small number of different sequences. Sequence analysis of this RNA fraction showed that it contains two main families of sequences, the B1 and B2 sequences, which account for 30% and 20% of all dsRNA-B respectively but which did not cross-hybridize. Three different B1 sequences were found, the a, b and c copies, which vary by less than 8%, allowing a consensus sequence of 129 or 130 base pairs to be identified.

4.8.2 Expression of B1 repeats

Young B1 sequences contain an internal split promoter that directs transcription by RNA polymerase (Pol III) and are followed by a poly(A) tract of approximately 40 nucleotides (Quentin, 1989). Studies have shown that B1 repeats are abundantly transcribed during very early mouse development (cleavage stages, the two cell stage) (Murphy *et al.*, 1983, Taylor & Pikó, 1987) and generally in fetal and adult germ-line tissue (Maraia, 1991, Adeniyi-Jones & Zasloff, 1985). As a result of their dispersal through the genome, B1 repeats are also present within pre-mRNA introns and untranslated regions (UTRs) of high molecular weight RNA (Bennett *et al.*, 1984, Kramerov *et al.*, 1982) as the B1 sequence is normally found close to poly A regions (Tokarskaya *et al.*, 1980).

Previous studies have found that B1 sequences are not transcribed in adult liver at a detectable level (Maraia, 1991). The high fetal kidney expression and the low liver expression of the B1 repeat sequence will lead to enrichment for B1 repeat sequences in the pool of fetal kidney cDNAs subtracted with adult liver mRNA. Two of the clones isolated in this screen (D2 and D4) contain only the B1 repeat sequences. Four of the clones described in this chapter, C1, C2, C5 and D3 have the B1 repeat within

their 5' or 3' UTRs. As these clones did not give a clear band on northern blots no attempts were made to get accurate expression levels in different tissues by optimizing northern blots. Therefore the northern blots presented in this thesis can only give a rough estimation of expression levels in tissues. Each northern blot did show no expression in the adult liver and that the clone contained a repeat sequence. These cDNAs may be interesting, but they were probably selected only because they contained a copy of the B1 repeat and so were not considered further in this project.

4.8.3 Conclusions from the subtracted screening

When the results are carefully considered and the subtractive hybridization evaluated in the light of the clones which were identified, the success of the subtraction can be seen. All of the clones that were isolated in the subtractive screening of the fetal kidney library satisfied the selection criterion, suggesting that the experimental design was a success. The screen identified one, possibly two, unique cDNAs, clone C4 that was lost and the mouse epsilon isoform of the 14-3-3 protein (clones B2, C3 and D1) which turned out to show an interesting pattern of expression in kidney development (discussed in chapter 5).

A major problem with subtractive hybridization is the choice of material to subtract against, and this choice is crucial to the procedures outcome. The choice of tissue to subtract against is difficult as only in retrospect can it be seen exactly where genes important in the system of interest are co-expressed. If, for instance, a subtraction of fetal kidney cDNA had chosen either fetal heart or testes (fetal or adult) it would not have selected *WT-1*, the Wilms' tumour susceptibility gene. The WT-1 gene is crucial in kidney development, but it would not have been immediately obvious that fetal heart and testes were a bad choice of material to subtract against to isolate important genes for fetal kidney development. Adult liver has several advantages as a subtracting material: it would not be expected to express genes that are important in kidney development as it is a different tissue type, but it will express

many of the house-keeping genes expressed in fetal kidney. Therefore this choice was expected to provide a large subtracted cDNA pool for a preliminary screen to isolate cDNAs expressed in the fetal kidney. The adult liver also has advantages as it is plentiful and easy to obtain - this is not a trivial point as 80 μg of mRNA is used in a subtraction.

The screen showed that mouse adult liver does not express the B1 repeat sequence. In early papers it was stated that adult liver mRNA had an abnormally low amount of these RNAs compared to most tissues (Krayev *et al.*, 1980). This study has provided valuable information to be considered for future subtractive hybridization strategies. Subtractive hybridization projects that use an embryonic tissue cDNA, should either not use adult liver to subtract against or B1 transcripts should be added to the subtracting material. The subtractive screen did identify one gene, the mouse epsilon isoform of 14-3-3 protein family, which has an interesting developmental expression profile. This clone is discussed in the next chapter.

Chapter 5

Characterization of mouse epsilon 14-3-3 isoform

This chapter considers the final three clones, B2, C3 and D1, that were isolated by the subtractive screen described in Chapter 3. Initial sequence analysis gave sequence of the 5' and 3' extremes of the cDNAs. Using the UWGCG programme to compare the sequences and by comparing the sequences visually it could easily be seen that clones B2 and C3 were the same sequence in the same orientation, but that clone B2 started fifteen bases further 3' than clone C3. Further analysis showed that clone D1 was also the same cDNA but cloned into pBluescript SK(-) in the opposite orientation. Clones B2 and C3 are cloned 3' to 5' relative to the polylinker orientation of *Sac I* to *Kpn I* 5' to 3', with clone D1 being 5' to 3' in the same relative manner. All three clones were approximately 1.7 kb long.

5.1 Sequence analysis of the mouse ε 14-3-3 cDNA.

Sequencing was done on both single-stranded and double-stranded DNA on two of the clones, either B2 or C3 to give the anti-sense sequence, and D1 to give the sense sequence. New primers of 20 nucleotides were synthesized according to the most extreme sequence produced and used to generate further sequence. All the sequences were compared and aligned manually using the UWGCG LINEUP programme until the sequence overlapped in the centre. The three cDNAs were sequenced to give a complete sequence in aggregate for both strands for the whole length of the cDNA. Regions that gave ambiguous sequence were re-sequenced, sometimes using an alternative primer, until the ambiguity was resolved and the anti-sense sequence complimented the sense sequence. A single open reading frame of 765 base pairs encoding 255 aa was found, starting at nucleotide 84 (Figure 5.1). The proposed initiator methionine codon lies in a region of strong homology to the consensus translation initiation sequence (CCGCAATGG, Kozak, 1984). It has no poly(A) tail but has the consensus sequence for polyadenylation at the 3' end of the cDNA, AAUAAA (Proudfoot & Brownlee, 1976; Proudfoot, 1991). Analysis of the deduced amino acid sequence predicted that the size of the protein is 29.2 kDa with

Figure 5.1

The nucleotide and deduced amino acid sequence of the mouse 14-3-3 ε isoform. The amino acid encoded is shown below the respective codon using the one letter code for amino acids. The nucleotide sequence is numbered from the start of the cDNA and the amino acids from the predicted initiator methionine. The EMBL databank accession number is Z19599.

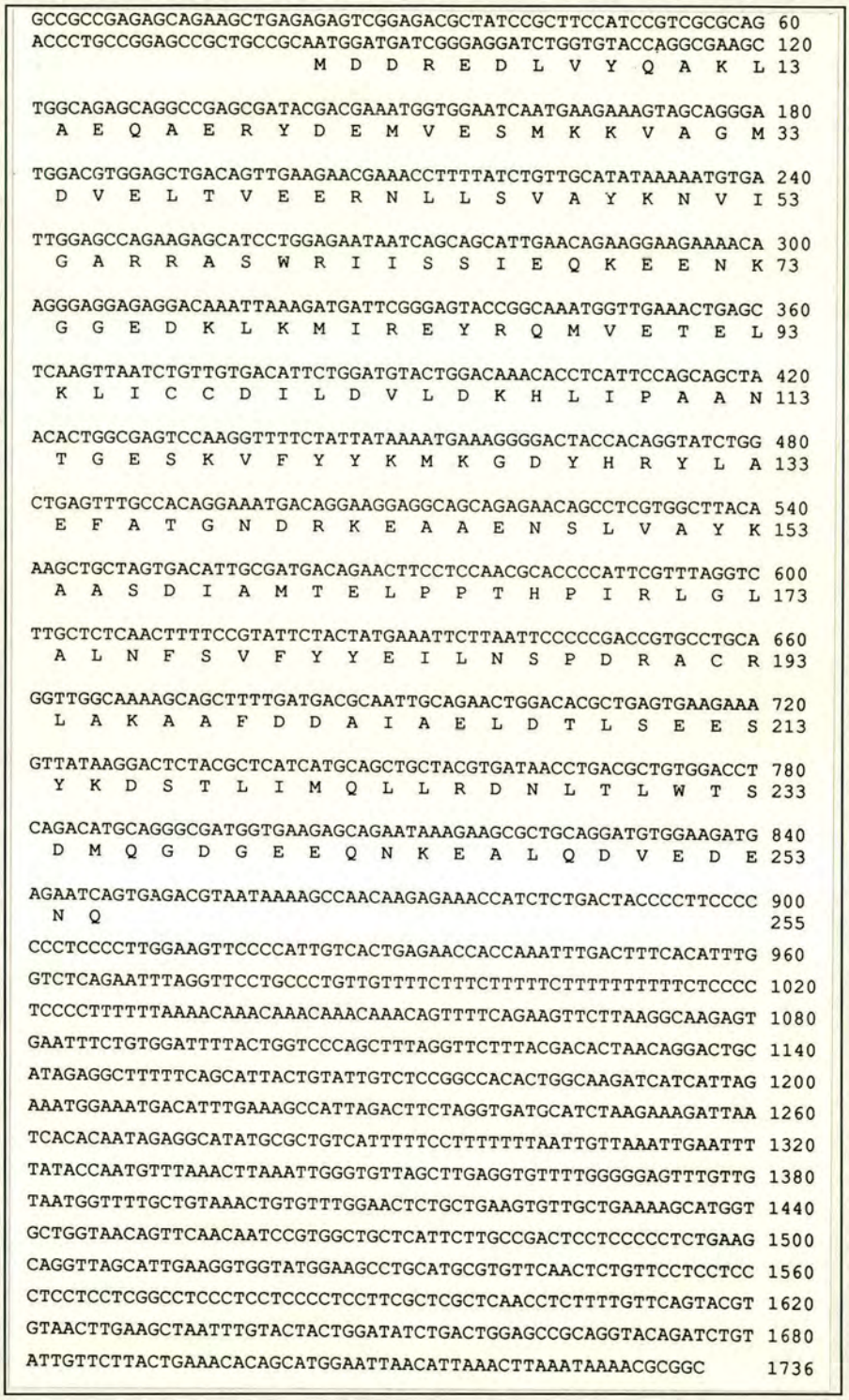

GCCGCCGAGAGCAGAAGCTGAGAGAGTCGGAGACGCTATCCGCTTCCATCCGTCGCGCAG 60
ACCCTGCCGGAGCCGCTGCCGCAATGGATGATCGGGAGGATCTGGTGTACCAGGCGAAGC 120
M D D R E D L V Y Q A K L 13
TGGCAGAGCAGGCCGAGCGATACGACGAAATGGTGGAATCAATGAAGAAAGTAGCAGGGA 180
A E Q A E R Y D E M V E S M K K V A G M 33
TGGACGTGGAGCTGACAGTTGAAGAACGAAACCTTTTATCTGTTGCATATAAAAATGTGA 240
D V E L T V E E R N L L S V A Y K N V I 53
TTGGAGCCAGAAGAGCATCCTGGAGAATAATCAGCAGCATTGAACAGAAGGAAGAAAACA 300
G A R R A S W R I I S S I E Q K E E N K 73
AGGGAGGAGAGGACAAATTAAAGATGATTCGGGAGTACCGGCAAATGGTTGAAACTGAGC 360
G G E D K L K M I R E Y R Q M V E T E L 93
TCAAGTTAATCTGTTGTGACATTCTGGATGTACTGGACAAACACCTCATTCCAGCAGCTA 420
K L I C C D I L D V L D K H L I P A A N 113
ACACTGGCGAGTCCAAGGTTTTCTATTATAAAATGAAAGGGGACTACCACAGGTATCTGG 480
T G E S K V F Y Y K M K G D Y H R Y L A 133
CTGAGTTTGCCACAGGAAATGACAGGAAGGAGGCAGCAGAGAACAGCCTCGTGGCTTACA 540
E F A T G N D R K E A A E N S L V A Y K 153
AAGCTGCTAGTGACATTGCGATGACAGAACTTCCTCCAACGCACCCCATTCGTTTAGGTC 600
A A S D I A M T E L P P T H P I R L G L 173
TTGCTCTCAACTTTTCCGTATTCTACTATGAAATTCTTAATTCCCCCGACCGTGCCTGCA 660
A L N F S V F Y Y E I L N S P D R A C R 193
GGTTGGCAAAAGCAGCTTTTGATGACGCAATTGCAGAACTGGACACGCTGAGTGAAGAAA 720
L A K A A F D D A I A E L D T L S E E S 213
GTTATAAGGACTCTACGCTCATCATGCAGCTGCTACGTGATAACCTGACGCTGTGGACCT 780
Y K D S T L I M Q L L R D N L T L W T S 233
CAGACATGCAGGGCGATGGTGAAGAGCAGAATAAAGAAGCGCTGCAGGATGTGGAAGATG 840
D M Q G D G E E Q N K E A L Q D V E D E 253
AGAATCAGTGAGACGTAATAAAAGCCAACAAGAGAAACCATCTCTGACTACCCCTTCCCC 900
N Q 255
CCCTCCCCTTGGAAGTTCCCCATTGTCACTGAGAACCACCAAATTTGACTTTCACATTTG 960
GTCTCAGAATTTAGGTTCCTGCCCTGTTGTTTTCTTTCTTTTTCTTTTTTTTTTCTCCCC 1020
TCCCCTTTTTTAAAACAAACAAACAAACAAACAGTTTTCAGAAGTTCTTAAGGCAAGAGT 1080
GAATTTCTGTGGATTTTACTGGTCCCAGCTTTAGGTTCTTTACGACACTAACAGGACTGC 1140
ATAGAGGCTTTTTCAGCATTACTGTATTGTCTCCGGCCACACTGGCAAGATCATCATTAG 1200
AAATGGAAATGACATTTGAAAGCCATTAGACTTCTAGGTGATGCATCTAAGAAAGATTAA 1260
TCACACAATAGAGGCATATGCGCTGTCATTTTTCCTTTTTTTAATTGTTAAATTGAATTT 1320
TATACCAATGTTTAAACTTAAATTGGGTGTTAGCTTGAGGTGTTTTGGGGGAGTTTGTTG 1380
TAATGGTTTTGCTGTAAACTGTGTTTGGAACTCTGCTGAAGTGTTGCTGAAAAGCATGGT 1440
GCTGGTAACAGTTCAACAATCCGTGGCTGCTCATTCTTGCCGACTCCTCCCCCTCTGAAG 1500
CAGGTTAGCATTGAAGGTGGTATGGAAGCCTGCATGCGTGTTCAACTCTGTTCCTCCTCC 1560
CTCCTCCTCGGCCTCCCTCCTCCCCTCCTTCGCTCGCTCAACCTCTTTTGTTCAGTACGT 1620
GTAACTTGAAGCTAATTTGTACTACTGGATATCTGACTGGAGCCGCAGGTACAGATCTGT 1680
ATTGTTCTTACTGAAACACAGCATGGAATTAACATTAAACTTAAATAAAACGCGGC 1736

Figure 5.1

an isoelectric point of 4.46. Analysis of the deduced amino acid sequence using the TFASTA program with the GENBANK database identified homologous proteins that belong to the 14-3-3 family which are predicted to play a role in signal transduction (Aitken *et al.,* 1992) and identified the cDNA as the ε isoform. Since the cloning of the mouse 14-3-3 ε cDNA the number of these proteins reported has increased exponentially. The newly reported sequences include the rat and sheep ε isoforms as well as human expressed sequence tags with homology to the ε cDNA.

Later analysis found that the predicted peptide sequence was identical to the N-acetylated partial peptide sequence of the sheep ε isoform (Toker *et al.,* 1992), confirming the assignment of the mouse protein's N-terminus. Recently the sheep and rat ε isoform of the 14-3-3 protein cDNAs have been entered into the EMBL database with accession numbers L07914 and M84416 respectively (Roseboom *et al.*, 1993), and the sheep sequence was also published in a review of the 14-3-3 proteins (Aitken *et al.,* 1992). Comparisons of the three DNA sequences (Figure 5.2) show that the mouse and rat sequences are very similar, with the figures being shown in Table 5.1. Comparisons of the ε sequence to the other mammalian isoforms (Figure 6.2) and also the plant isoforms (Figure 6.3) shows that the start of the putative open reading frame is very similar to the other mammalian isoforms whilst the plant isoforms have open reading frames that extended 5' to the mammalian ORF.

Table 5.1

The table below shows the level of homology between the 5' and 3' untranslated regions (UTR) of the mouse and rat and the open reading frame (ORF) of all three species.

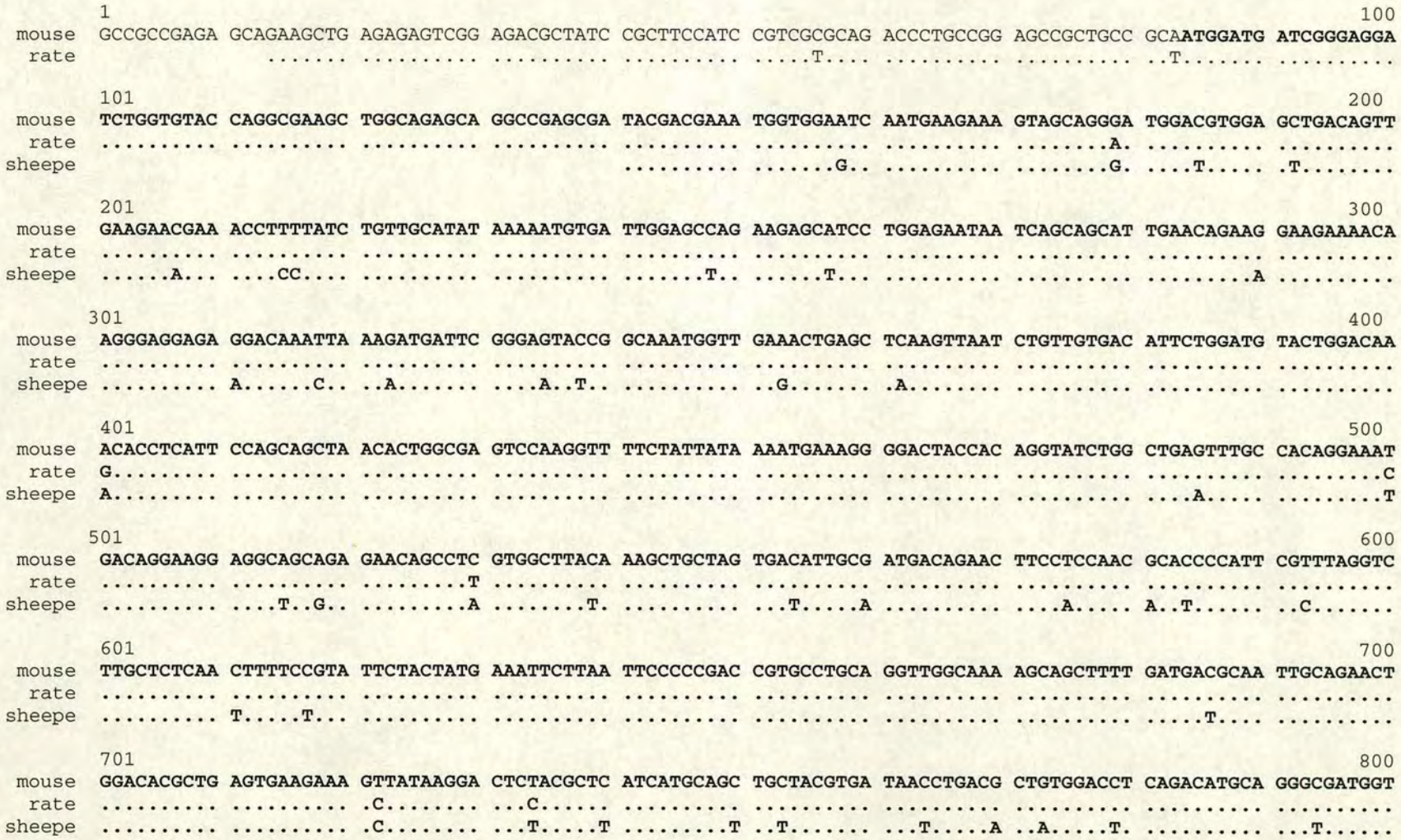
Species	rat			mouse
	5'	ORF	3'	ORF
mouse	98.6	99.1	97.6	
sheep		93.5		93.8

Figure 5.2

The BESTFIT comparison of the DNA sequences of the 14-3-3 ε isoform from rat, mouse and sheep. Identical nucleotides are indicated by a dot and nucleotides that differ are shown. The coding region is shown in bold.

mouse, mouse 14-3-3 ε cDNA; rate, rat 14-3-3 ε cDNA; sheepe, sheep 14-3-3 ε cDNA (only the coding region has been cloned).

Figure 5.2


1 100
mouse GCCGCCGAGA GCAGAAGCTG AGAGAGTCGG AGACGCTATC CGCTTCCATC CGTCGCGCAG ACCCTGCCGG AGCCGCTGCC GCAATGGATG ATCGGGAGGA
rate ....T.... ....T..... .....
101 200
mouse TCTGGTGTAC CAGGCGAAGC TGGCAGAGCA GGCCGAGCGA TACGACGAAA TGGTGGAATC AATGAAGAAA GTAGCAGGGA TGGACGTGGA GCTGACAGTT
rate ....A. ....
sheepe ....G.. ....G. ....T..... .T.....
201 300
mouse GAAGAACGAA ACCTTTTATC TGTTGCATAT AAAAATGTGA TTGGAGCCAG AAGAGCATCC TGGAGAATAA TCAGCAGCAT TGAACAGAAG GAAGAAAACA
rate ....
sheepe ......A... ....CC.... ....T.. ......T... ....A .....
301 400
mouse AGGGAGGAGA GGACAAATTA AAGATGATTC GGGAGTACCG GCAAATGGTT GAAACTGAGC TCAAGTTAAT CTGTTGTGAC ATTCTGGATG TACTGGACAA
rate ....
sheepe ..... A......C.. ..A..... ....A..T.. ....G..... .A..... ....
401 500
mouse ACACCTCATT CCAGCAGCTA ACACTGGCGA GTCCAAGGTT TTCTATTATA AAATGAAAGG GGACTACCAC AGGTATCTGG CTGAGTTTGC CACAGGAAAT
rate G..... C
sheepe A..... ....A..... .....T
501 600
mouse GACAGGAAGG AGGCAGCAGA GAACAGCCTC GTGGCTTACA AAGCTGCTAG TGACATTGCG ATGACAGAAC TTCCTCCAAC GCACCCCATT CGTTTAGGTC
rate ....T ....
sheepe ....T..G.. .....A .....T. ....T.....A ....A..... A..T..... ..C.....
601 700
mouse TTGCTCTCAA CTTTTCCGTA TTCTACTATG AAATTCTTAA TTCCCCCGAC CGTGCCTGCA GGTTGGCAAA AGCAGCTTTT GATGACGCAA TTGCAGAACT
rate ....
sheepe ..... T.....T... ....T.... .....
701 800
mouse GGACACGCTG AGTGAAGAAA GTTATAAGGA CTCTACGCTC ATCATGCAGC TGCTACGTGA TAACCTGACG CTGTGGACCT CAGACATGCA GGGCGATGGT
rate .... .C..... ...C..... ....
sheepe .... .C..... ...T.....T .....T ..T..... ...T.....A ..A.....T. ....T.....

```
        801                                                                                                  900
mouse   GAAGAGCAGA ATAAAGAAGC GCTGCAGGAT GTGGAAGATG AGAATCAGTG AGACGTAATA AAAGCCAACA AGAGAAACCA TCTCTGACTA CCCCTTCCC
 rate   .......... .......... .......... .......... .......... .........  .......... .......... .......... ....AC...A
sheepe  .......... .......... C..C.....C .....

        901                                                                                                 1000
mouse   CCCCTCCCCT TGGAAGTTCC CCATTGTCAC TGAGAACCAC CAAATTTGAC TTTCACATTT GGTCTCAGAA TTTAGGTTCC TGCCCTGTTG TTTTCTTTCT
 rate   .....G.... .......... .......... .......... .......... .......... .......... .......... .......... .... .C.T.

        1001                                                                                                1100
mouse   TTTTCTTTTT TTTTTCTCCC CTCCCCTTTT TTAAAACAAA CAAACAAACA AACAGTTTTC AGAAGTTCTT AAGGCAAGAG TGAATTTCTG TGGATTTTAC
 rate   C...T.C... .......... .......... ...              ..... .......... .......... .......... .......... ..........

        1101                                                                                                1200
mouse   TGGTCCCAGC TTTAGGTTCT TTACGACACT AACAGGACTG CATAGAGGCT TTTTCAGCAT TACTGTATTG TCTCCGGCCA CA CTGGCAA GATCATCATT
 rate   .......... .......... .......... .......... .......... .......... .......... .......... ..TG...... ..........

        1201                                                                                                1300
mouse   AGAAATGGAA ATGACATTTG AAAGCCATTA GACTTCTAGG TGATGCATCT AAGAAAGATT AATCACACAA TAGAGGCATA TGCGCTGTCA TTTTTCCTTT
 rate   .......... .......... .G........ .......... .......... .......... .......... .......... .......... ..........

        1301                                                                                                1400
mouse   TTTTAATTGT TAAATTGAAT TTTATACCAA TGTTTAAACT TAAATTGGGT GTTAGCTTGA GGTGTTTTGG GGGAGTTTGT TGTAATGGTT TTGCTGTAAA
 rate   .......... .......... .......... .......... .......... .......... .......... .........C G......... .........C

        1401                                                                                                1500
mouse   CTGTGTTTGG AACTCTGCTG AAGTGTTGCT GAAAAGCATG GTGCTGGTAA CAGTTCAACA ATCCGTGGCT GCTCATTCTT GCCGACTCCT CCCCCTCTGA
 rate   G......... .......... ...G.C.... ........C. .......... ....C..... G....C.... .......... ..........   ........

        1501                                                                                                1600
mouse   AGCAGGTTAG CATTGAAGGT GGTATGGAAG CCTGCATGCG TGTTCAACTC TGTTCCTCCT CCCTCCTCCT CGGCCTCCCT CCTCC   CC TCCTTCGCTC
 rate   C......... .......... .......... .......... ..........   ......T. .......... .......... .A........ ..........

        1601                                                                                                1700
mouse   GCTCAACCTC TTTTGTTCAG TACGTGTAAC TTGAAGCTAA TTTGTACTAC TGGATATCTG ACTGGAGCCG CAGGTACAGA TCTGTATTGT TCTTACTGAA
 rate   .......... .......... ..T....... .......... .......... .......... .........A ...A...... .......... ..........

        1701
mouse   ACACAGCATG GAATTAACAT TAAACTTAAA TAAAACGCGG C.........  ......
 rate   .......... .......... .......... ......AAAC..TAAATTAAA AATGCC.
```

The mouse cDNA was used to probe, at high stringency, a northern blot, with a range of mouse tissue RNAs on it. A transcript of 2.0 kb was found in the fetal liver and kidney, and in adult kidney, heart, brain, ovary, eye, testes, small intestine and EB26/10a (an embryonic carcinoma cell line) (Figure 5.3), but, as expected of the subtractive hybridization selection, no hybridization was seen in mRNA from adult liver. A second smaller transcript of 1.4 kb could be seen in the testes mRNA (Figure 5.3a); this smaller transcript could be an alternative transcript from the same gene, or it may be due to cross-hybridization to other isoforms. The RNA samples were not loaded in equal amounts as can be seen by looking at the ethidium bromide staining of the 18S and 28S rRNA from each sample (Figure 5.3b). The nylon membrane was also stained with methylene blue to show the amount of RNA that had transferred to the membrane, this showed an identical pattern to the ethidium bromide stain but the staining was too weak to be detected when photographed. However during the project all filters were visualized by UV illumination after blotting, which allowed an estimation of blotting efficiency and all the blots showed representative transfer to all regions of the nylon membrane.

Comparison of expression levels can not be accurately evaluated from this northern blot because of the differences between the amount of RNA in each lane but a general picture can be made and this can be compared to the results obtained by *in situ* hybridization studies. Taking the different loading levels into consideration it seems that the expression is higher in the adult kidney than it is in the fetal kidney (as a similar signal was obtained for the two samples although the adult kidney mRNA was present in at least two fold lower level than the fetal kidney mRNA). In comparison the gene appears to be expressed in moderate amounts in the fetal liver and not at all in the adult liver. (This observation shows the cDNA satisfies the selection criteria.) This is an interesting observation as the selection strategy selects for mRNAs that are down-regulated in the adult liver and that are highly expressed in the fetal kidney.

Figures 5.3

Northern blot a) of 10 μg of RNA from various mouse tissues probed with the mouse 14-3-3 ε isoform cDNA at high stringency. The lower photograph b) shows the 18S and 28S rRNA from each sample on the gel, visualized by ethidium bromide staining. The lower photograph shows the samples in the same loading order as the above photograph but between samples 8 (lung) and 9 (eye) the marker lane has not been cut out of the photograph. The marker lane was cut out of the upper photograph as the PCR amplified ε cDNA containing small regions of the Bluescript vector (50-100 bases) cross-hybridizes to the RNA markers provided by BRL presumably as they are transcribed from a pUC-derived plasmid. The adult kidney lane in both gels was originally located on the far right of the gel and has been repositioned to allow comparison of fetal and adult kidney RNA samples. The northern blot shows in which tissues the mouse ε mRNA is expressed and also gives a crude approximation, when both the autoradiograph and ethidium bromide staining are compared, of the relative expression levels in each tissue.

Abbreviations used and lanes of the ethidium bromide stained gel that the RNA correlates to:- fk, 17.5 d.p.c. fetal kidney (lane 1); ak, adult kidney (lane 2); fl, 17.5 d.p.c. fetal liver (lane 3); al, adult liver (lane 4); heart (lane 5); brain (lane 6); ovary (lane 7), lung (lane 8); eye (lane 9); spleen (lane 10); testes (lane 11); EB26/10a (an embryonic carcinoma cell line) (lane 12) and si, small intestine (lane 13).

All tissues are adult unless otherwise stated.

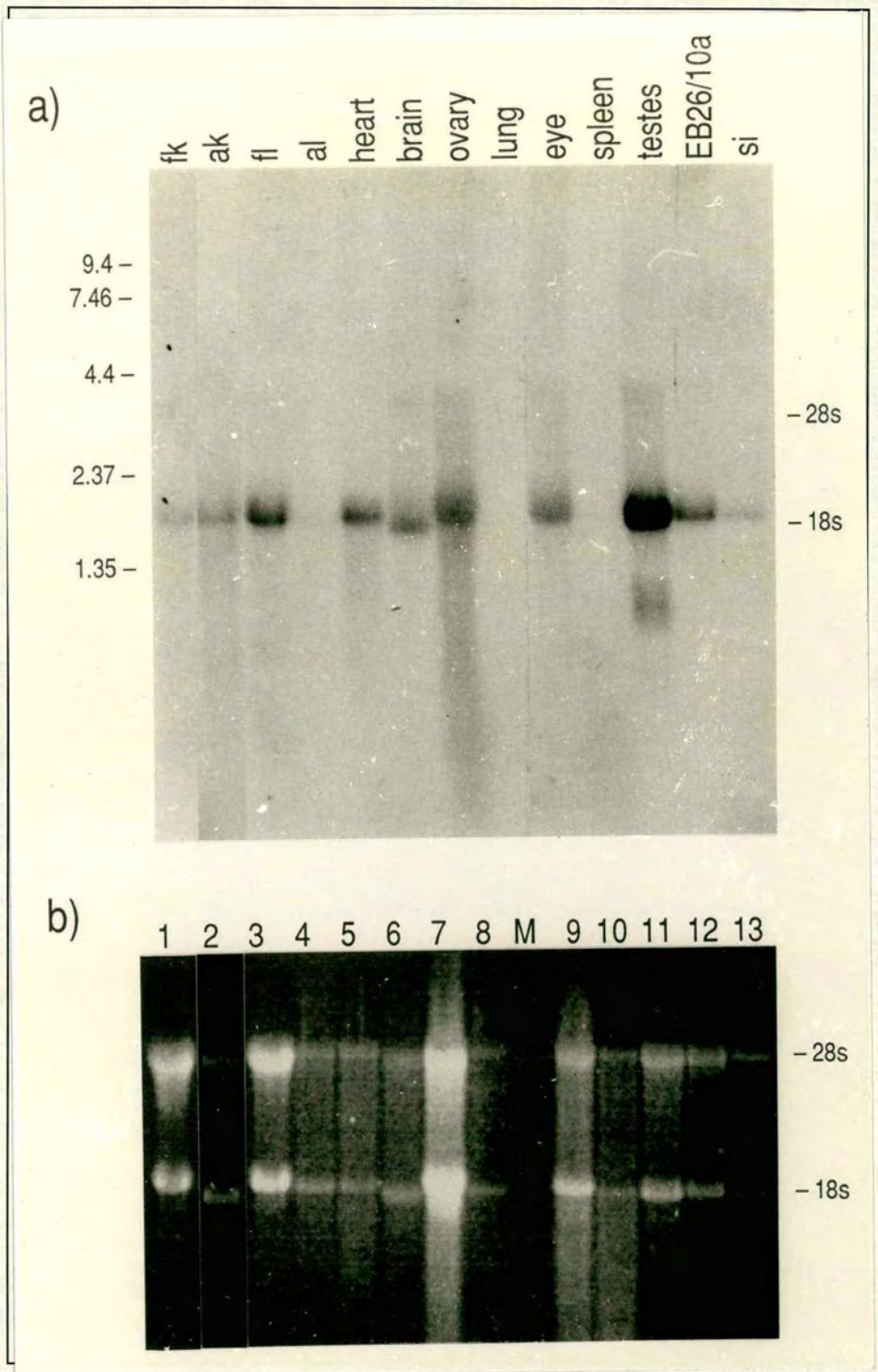

a)
fk
ak
fl
al
heart
brain
ovary
lung
eye
spleen
testes
EB26/10a
si
9.4 –
7.46 –
4.4 –
2.37 –
1.35 –
– 28s
– 18s
b)
1 2 3 4 5 6 7 8 M 9 10 11 12 13
– 28s
– 18s

Figure 5.3

The relative amounts of mRNA in the liver and kidney have been closely examined in this study due to their choice as the library material and subtractor in the subtraction strategy. The relevance of whether the fluctuations in expression levels reflect changes in cellular requirements for the 14-3-3 protein, perhaps reflecting changing requirements for a signal transduction pathway, would have to be carefully examined as another isoform may be replacing the ε isoform when it is down regulated and functionally replacing it.

The expression of the 14-3-3 ε gene in the early developing kidney is highest in the undifferentiated mesenchyme and seems to fall as the kidney matures. *In situ* hybridization of adult kidney tissue does not show signal confined to specific areas. Therefore it must be assumed that the gene is expressed in specific regions of the developing kidney at high levels and later as the kidney matures the expression falls but becomes more widespread throughout the whole structure (data not shown as the expression is diffuse and is not easily visualized in any one area). This makes the expression level in the overall kidney rise, and explains the rise in expression levels when the kidney homogenate is looked at by northern analysis compared to observing the specific structures by *in situ* hybridization. By comparison, in fetal liver, when studied throughout development by *in situ* hybridization, expression is generally low to moderate with a maximum expression level at day 11.5 d.p.c. of development. The expression falls thereafter until there is no detectable expression in the adult liver. The fetal kidney and liver for northern analysis were both taken from 17.5 d.p.c. embryos.

From northern analysis, it can also be seen that expression levels are extremely high in the adult testes , the number of transcripts detected in this RNA sample is not clear but there is at least one other transcript at 1.4 kb as well as the major transcript at 2.0 kb. Although there is no fetal testes sample on the northern blot, the *in situ* hybridization results that are discussed in the next section show that the genital tubercule, the region that will form the genitalia of the male and female, of the 12.5

and 13.5 d.p.c. embryo, can be seen to be expressing the mRNA at a high level. It can also be seen that the gene is highly expressed in the early seminiferous tubules of the testes in the 15.5 d.p.c. embryo. But it should also be noted that the ovary expresses the mRNA at a moderate but not high level. This therefore suggests the mRNA has a sex-dependent expression in at least one area of the adult mouse, an interesting observation as the subtractive screen was not planned to search for cDNAs with sex-dependent expression. The lack of expression in the adult lung, an epithelial organ, is an example of the general principle that expression levels of the mouse ε gene are low or non-detectable in epithelial tissues and higher in mesenchyme.

The cDNA was also used to probe mouse genomic Southern blots (Figures 5.4a and b). Genomic DNA from 3 different inbred strains of mice (Swiss, C57/BL/6 and DBA/2) was cut with several enzymes to look for strain-specific polymorphisms. It was shown that, at high stringency, the cDNA hybridizes to only one or two DNA fragments, suggesting that the hybridization is only to a single gene. The genomic blots show that the cDNA hybridizes to one DNA fragment when cut with the restriction enzyme *Bam HI* , two DNA fragments when cut with the restriction enzymes *Eco RI*, *Hind III* and *Pst I*, and one DNA fragment when cut with *Pvu I*.

As the sequence of the cDNA cloned does not contain any *Eco RI* or *Hind III* sites then the most obvious assumption is that the gene is probably located on two exons with the restriction site being found in the intron. Of course there is no proof that this gene is not multi-exonic but as the cDNA only picks up two DNA fragments the most intuitive observation is that this gene is on two or three exons, and most probably two. To investigate this properly would require cloning of the genomic DNA containing this region. PCR amplification of genomic DNA using primers made for the cDNA can sometimes detect small introns by amplifying a larger

Figure 5.4 a

A Southern blot of Swiss mouse genomic DNA cut with *Bam HI*, *Eco RI*, *Hind III*, *Pst I* and *Pvu I*, probed with the mouse 14-3-3 ε cDNA at high stringency. λ DNA cut with *Hind III* was run as a marker lane (M) to provide size markers for calculation of sizes of restriction fragments.

Figure 5.4b

A Southern blot of mouse genomic DNA from C57/BL/6 and DBA/2 mice cut with *Bam HI* and *Eco RI* restriction enzymes. λ DNA cut with *Hind III* was run as a marker lane (M) to provide size markers for calculation of sizes of restriction fragments.

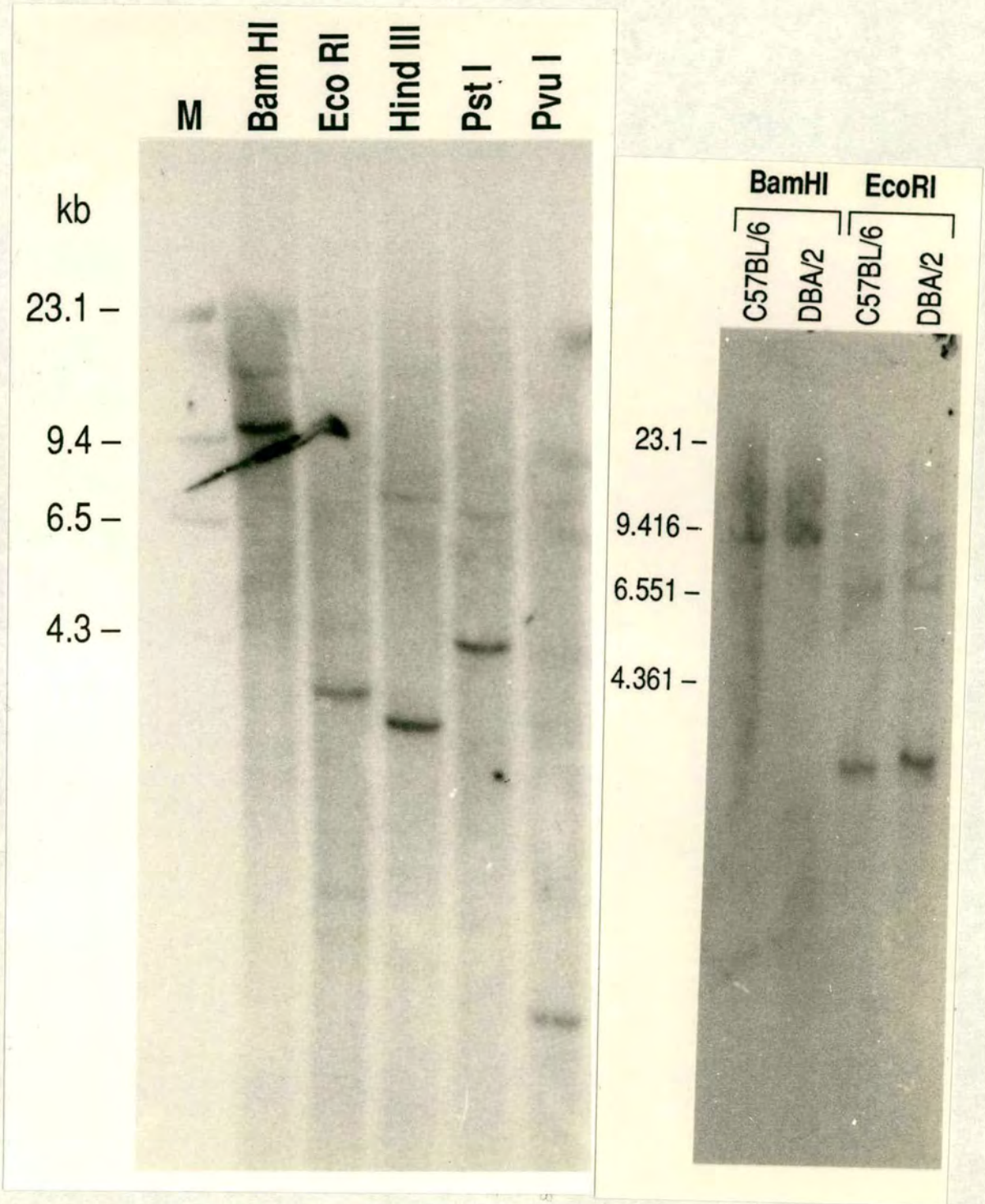


Figure 5.4a and b

genomic fragment than cDNA fragment, however amplification of the ε gene from genomic DNA either gave no product, possibly due to the presence of a intron too large to amplify or a similar sized fragment to the cDNA amplicon. Therefore the genomic organization of this gene can only be suggested.

Preliminary work was done to find the chromosomal location of the ε isoform using hamster-mouse hybrid cell lines. The hybrid cell lines have a full complement of hamster chromosomes but each cell line has a different selection of additional mouse chromosomes. Mapping the gene to a chromosome requires distinguishing between the mouse and hamster genes (by a difference in PCR fragment length or restriction-enzyme-digest fragment size) which allows the detection of the presence or absence of the mouse-specific band (the hamster band will be present in all cell lines). It is then possible to correlate the presence of the mouse-specific polymorphism in the hamster-mouse hybrid cell lines with the presence of a particular mouse chromosome. When primers designed for the mouse 14-3-3 ε cDNA were used to amplify mouse and hamster genomic DNA the PCR products showed no difference in size. Southern blots of the mouse and hamster PCR products when hybridized with radioactively labeled mouse 14-3-3 ε cDNA show that the mouse cDNA probe hybridizes to the hamster PCR products as strongly as to the mouse PCR products even when the hybridization and washes were performed under high stringency conditions (65°C, 0.1 x SSC) preventing differential identification of mouse and hamster PCR products. With no polymorphism the hybrid cell lines could not be used to assign a chromosomal location to the ε isoform. The level of sequence similarity between rat and mouse (Table 5.1) explains why attempts to find a PCR length polymorphism between hamster and mouse DNA did not prove successful.

The cDNA was then used for fluorescent *in situ* hybridization (FISH) (work kindly performed by Dr. John Inglis and Ms Muriel Lee) and the mouse ε isoform

Figure 5.5

Fluorescent *in situ* hybridization (FISH) of a metaphase spread of mouse spleen cells using the mouse 14-3-3 ε isoform cDNA as a probe showing hybridization to chromosome 11.

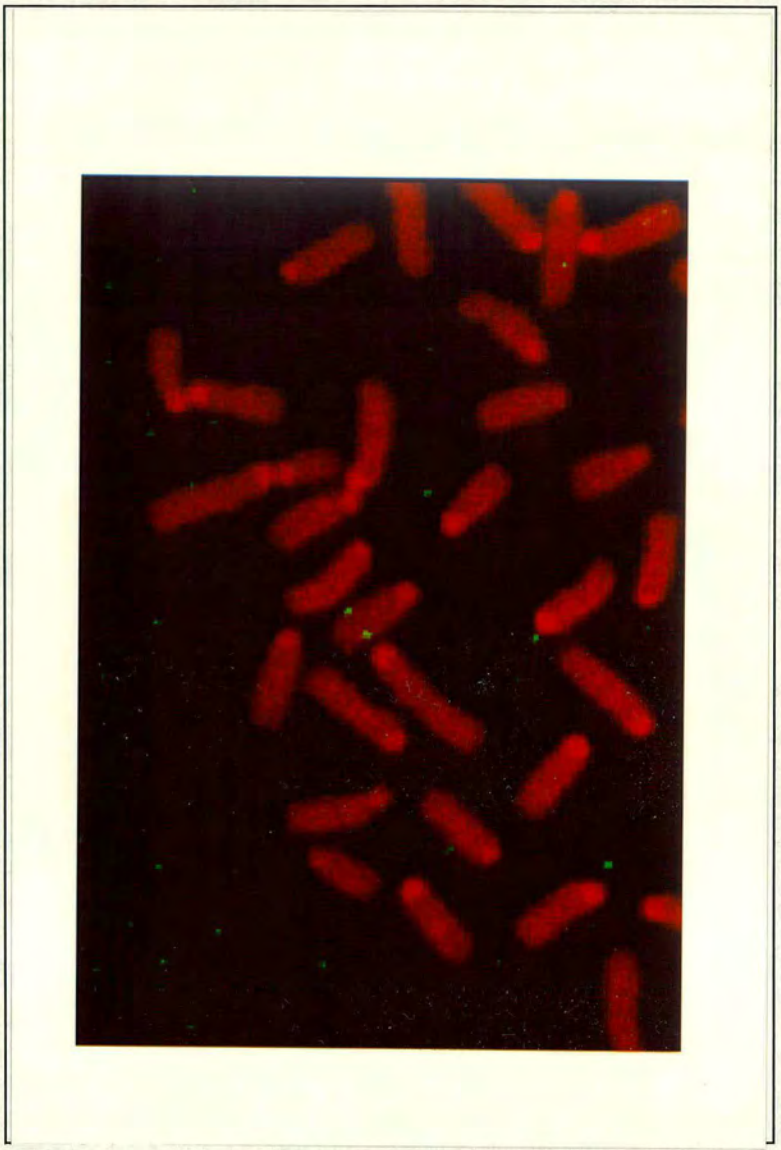


Figure 5. 5

was shown to map to mouse chromosome 11, region B5 or C (Figure 5.5). No mouse mutants are currently mapped to this chromosomal region.

To examine when the 14-3-3 ε mRNA was first expressed, RNA PCR was performed on cytoplasmic extracts from 7.5 d p.c. and 8.5 d.p.c. embryos, and purified RNA from 9.5, 10.5 and 15 d.p.c. embryos. First-strand cDNA synthesis was primed from either random hexamers (N_6) or from a mouse 14-3-3 ε cDNA-specific oligonucleotide, C520 (see section 2.8.1). Two internal 14-3-3 ε cDNA-specific oligonucleotides, C636 and C637, were then used to amplify the cDNA giving a specific PCR product of 750 bases from all samples tested. This band (plus a lower band of circa 600 bases) was present in all cases (Figure 5.6). The lower band may be due to amplification of the same region of other 14-3-3 isoforms due to cross-annealing of primers to similar sequences. This result shows that the embryo expresses the gene as early in development as 7.5 d.p.c. and up to and including 15.5 d.p.c. This data is reaffirmed by the *in situ* hybridization analysis discussed next. To assess levels of mRNA present in each sample from this experiment is possible but as reviewed by Foley *et al.* (1993) the application of RNA PCR to quantification of nucleic acids is highly problematic. It has been shown that minimal differences between adjacent wells of a thermocycler can dramatically affect the yields of the amplicon from different reactions. In order to quantify the amount of relative mRNA between different samples, the absolute amounts of RNA in the original samples must be standardized and it is necessary to contain internal PCR controls that show that the amplification has been quantitatively similar between samples. It could be suggested that the levels of the secondary amplicon found at 600 bp is low in the 7.5 d.p.c. embryo, high in the 9 and 10 d.p.c. embryo before falling in the later samples from 15 d.p.c. embryo and 17 d.p.c. kidney. However it does appear that the PCR may have reached plateau levels in the 9 and 10 d.p.c samples which would make estimations of levels of magnitude of increase in expression difficult to calculate. It

Figure 5.6

RNA PCR was performed on cytoplasmic extracts from 7.5 d p.c. and 8.5 d.p.c. embryos, and purified RNA from 9.5, 10.5 and 15 d.p.c. embryos. First strand cDNA synthesis was reverse transcribed using either a) random hexamers (N_6) or b) a mouse 14-3-3 ε cDNA-specific oligonucleotide, C520 (see section 2.8.1) as primers. Where it is not designated random hexamers were used. Two internal 14-3-3 ε cDNA-specific oligonucleotides, C636 and C637, were then to amplify the cDNA. Controls were the same reaction with no reverse transcriptase added.

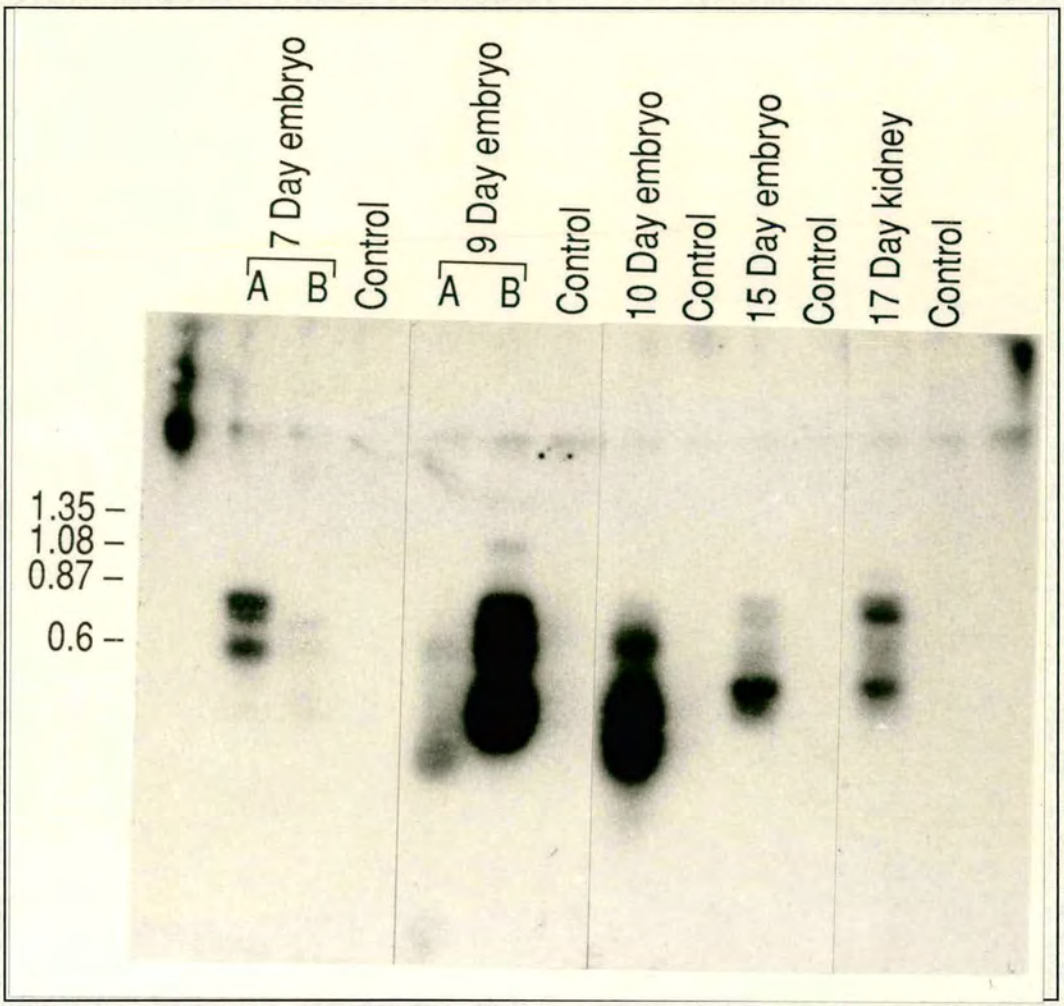

A
B
7 Day embryo
Control
A
B
9 Day embryo
Control
10 Day embryo
Control
15 Day embryo
Control
17 Day kidney
Control
1.35
1.08
0.87
0.6

Figure 5.6

would also seem that the 'real' amplicon of 750 bases is also at the highest levels in the 9 and 10 d.p.c. embryo. Comparing these results to the *in situ* hybridization expression data, the expression is low in the 7.5 and 8.5 d.p.c. embryo (in both the *in situ* hybridization and the RNA-PCR), and the levels in the 9.5 and 10.5 d.p.c. embryo is moderately high whilst the level of expression has fallen in the 15.5 d.p.c. *in situ* hybridization and fallen in the 15.5 and 17.5 d.p.c. kidney as the RNA-PCR would suggest. Therefore although the RNA-PCR results can not be taken on their own without further controls, both the detection of expression and the expression level data are in agreement with the *in situ* hybridization data and reaffirm the results obtained.

5.2 In situ hybridization of the mouse ε 14-3-3 isoform.

In situ hybridization to 8.5, 9.5, 10.5, 11.5, 12.5, 13.5 and 15.5 d.p.c. embryos was done to analyze the expression of the mouse 14-3-3 ε gene. This selection of ages was chosen to gives a 7 day window during a period of mouse development that covers most of the key events of post-implantation embryogenesis. During this time period a very wide range of tissues start their differentiation and the expression of the gene can be studied as their development occurs.

The data discussed here is summarized in a table format in section 5.3 which attempts to show how the levels of the mouse 14-3-3 ε isoform varies in various tissues throughout development.

A pilot set of *in situ* hybridization experiments were performed for the mouse ε 14-3-3 gene, where both the sense and anti-sense RNA strands were synthesized and hybridized to mouse embryo sections. When the results were analyzed it became clear that the anti-sense strand was giving a specific signal, localized on the slide over the embryo sections and concentrated in specific tissues. In contrast, the slides hybridized to the sense strand were negative; the signal over the embryo sections was not markedly different from surrounding areas of the slide and the signal was not

Figure 5.7

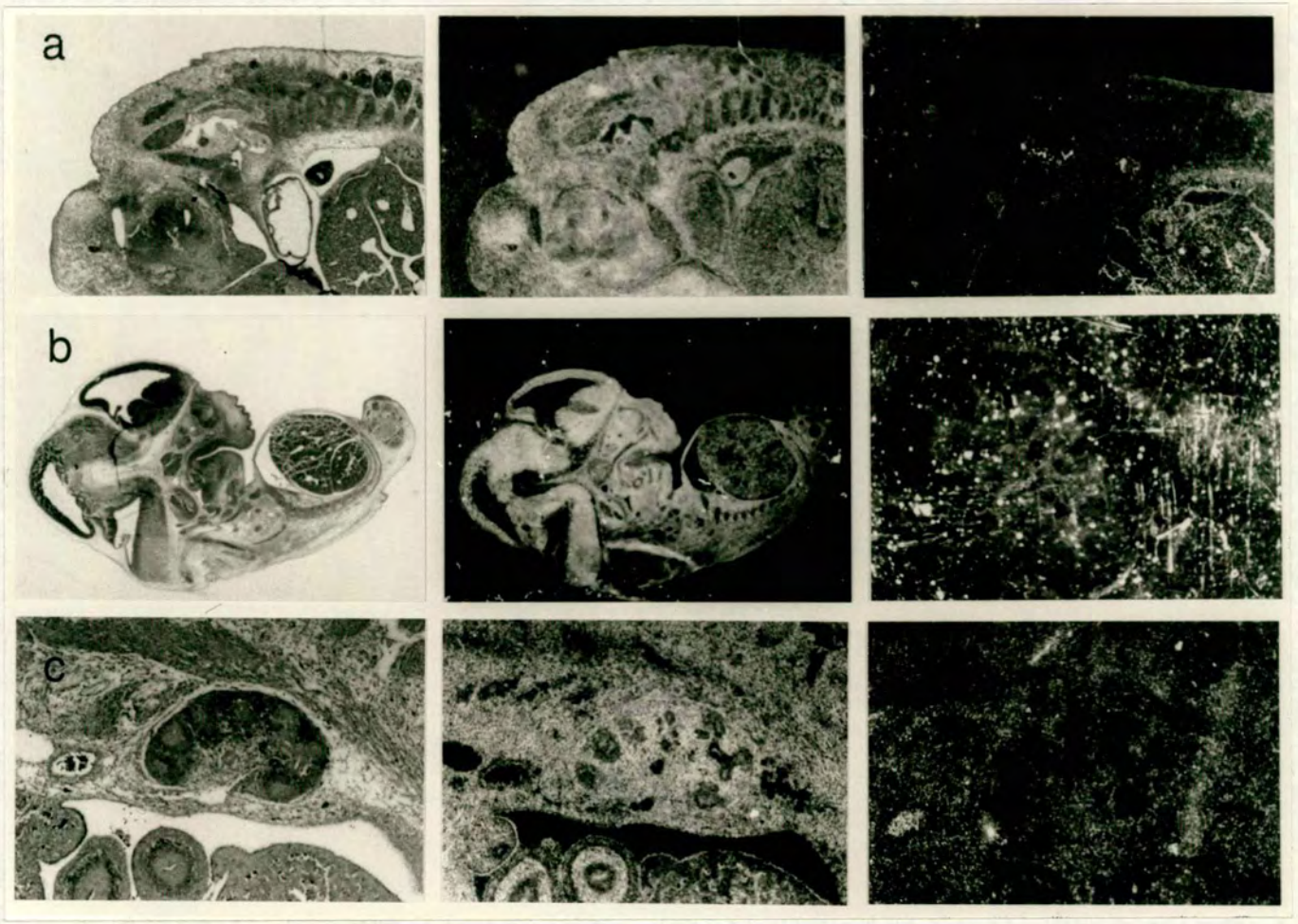
Panel of photographs showing *in situ* hybridization of the mouse 14-3-3 ε isoform sense and antisense transcripts to adjacent sections of the same embryo.

a, left) Bright field photograph of a parasagital section of a 12.5 d.p.c. embryo (16x), a,centre) Dark field photograph of hybridization with the anti-sense transcript, and a,right) Dark field photograph of hybridization with the sense transcript.

b,left) Bright field photograph of a parasagital section of a 13.5 d.p.c. embryo (7.5x), b,centre) Dark field photograph of hybridization with the anti-sense transcript, and b,right) Dark field photograph of hybridization with the sense transcript.

c,left) Bright field photograph of a parasagital section of a 13.5 d.p.c. kidney (45x), c,centre) Dark field photograph of hybridization with the anti-sense transcript, and c,right) Dark field photograph of hybridization with the sense transcript .

These photographs show a selection of sections hybridized to sense and anti-sense transcripts from the mouse 14-3-3 ε isoform. The lack of hybridization of the sense transcript shows that the expression pattern detected by the anti-sense transcript is specific.



Bright Field **Anti-Sense** **Sense**

Figure 5.7

localized to specific tissues. Figure 5.7 shows several examples of this. Sense controls are not shown for every *in situ* hybridization experiment in this chapter and for one experiment the sense strand synthesis was not hybridized to slides (due to lack of adequate transcription) but all experiments have shown that the sense strand of the mouse 14-3-3 ε cDNA does not hybridize to mouse embryos when used in *in situ* hybridization experiments. Therefore after Figure 5.7 the *in situ* hybridization results are shown only for the anti-sense strand hybridization.

In 8.5 d.p.c. embryos, the expression level is uniform in the epithelium, mesenchyme and the neural tube (indicated by an arrow in Figure 5.8a) (Figure 5.8a, b, c). The foregut (fg) [epithelium], forebrain (fb) [neural tissue] and heart (H) [mesenchyme] are indicated in the figures as examples of the different tissue types and all show similar expression levels.

In the 9.5 d.p.c. embryo, a differential pattern of expression can be seen. While the levels of expression remain the same in the epithelia and neural tube (indicated by an arrow, figure 5.9c), the level of expression rises in the undifferentiated mesenchyme (M in Figure 5.9c). Examination of the mesonephros, (indicated by Ms in figure 5.9c) where the mesenchyme has started to differentiate and partially epithelialize shows lower expression than the regions of undifferentiated mesenchyme. The non-somitic mesoderm is more heavily labeled than the somitic mesenchyme (the somites are indicated by thick arrows in figure 5.9a), with the general level of expression in the brain, neural tissue and somites being lower than in the undifferentiated mesenchyme (Figure 5.9a, b and c). Expression is also high in the presumptive dorsal root ganglia (indicated by a thin arrow in figure 5.9a). The whole-mount *in situ* hybridization to embryos of the same age shows that there is widespread expression at this stage (Figure 5.10) this can be seen as a diffuse staining pattern over the whole .

In the 10.5 d.p.c. embryo, the gene is expressed throughout the whole embryo. The level of expression in the epithelia had dropped to low levels and was very low in the

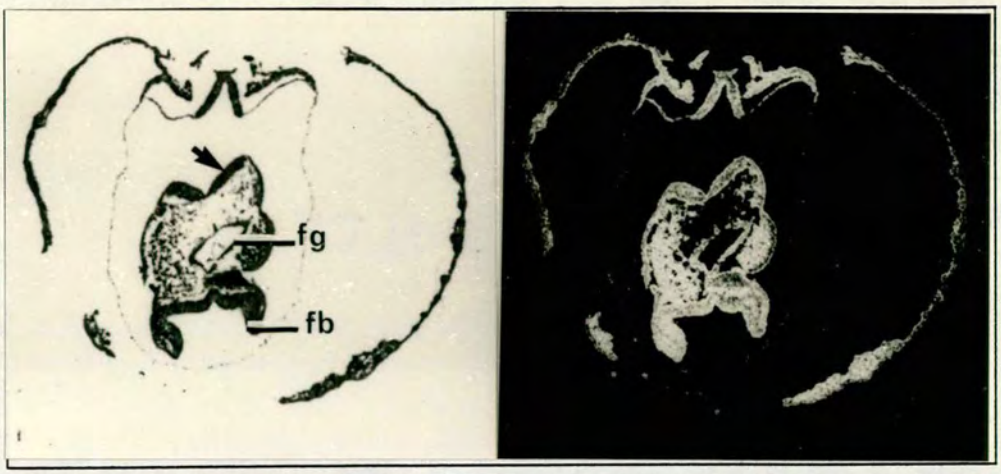
Figure 5.8

In situ hybridization photographs of a 8.5 day mouse embryo (Theiler stage 13) (x125) showing a) a bright field section and b) dark field sections. a, b. This shows a transverse section through the embryo. The positions of the foregut (fg), forebrain (fb) and neural folds (arrow) are indicated.

c) Colour photograph of *in situ* hybridization of a 8.5 d.p.c. mouse embryo (x140). The positions of the foregut (fg), heart (H) and neural folds (arrow) are indicated.

The *in situ* hybridization photographs show expression levels are uniformly low in epithelium (foregut), mesenchyme (heart) and the neural tissue (neural folds).

The plane of section in these photographs is shown below on the diagram of a 8.5 d.p.c. embryo (adapted from Kaufman, 1992) and labeled a,b and c corresponding to the figure panels.



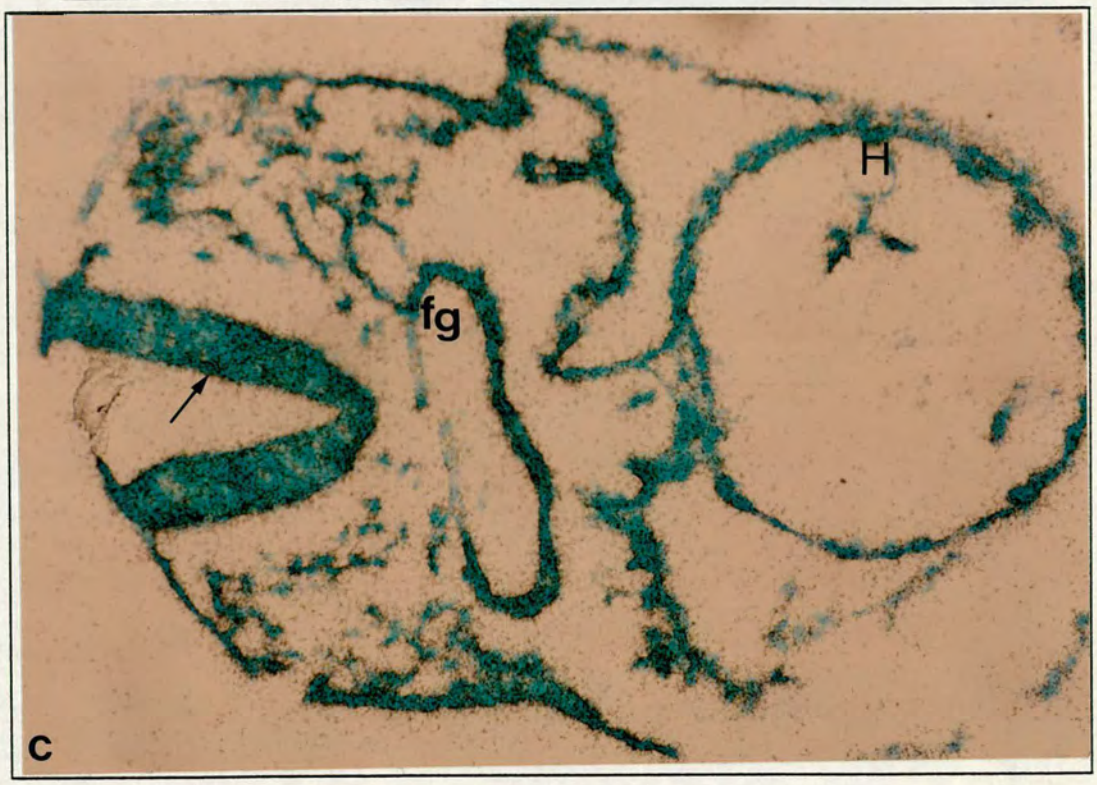


Figure 5.8a, b, c

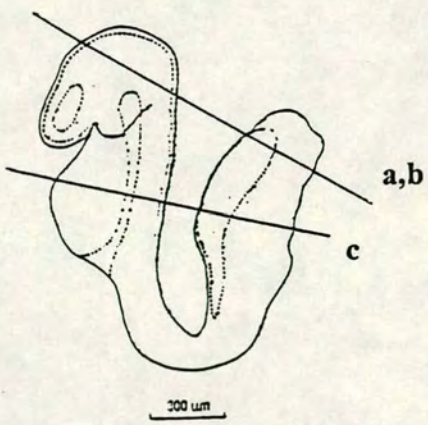


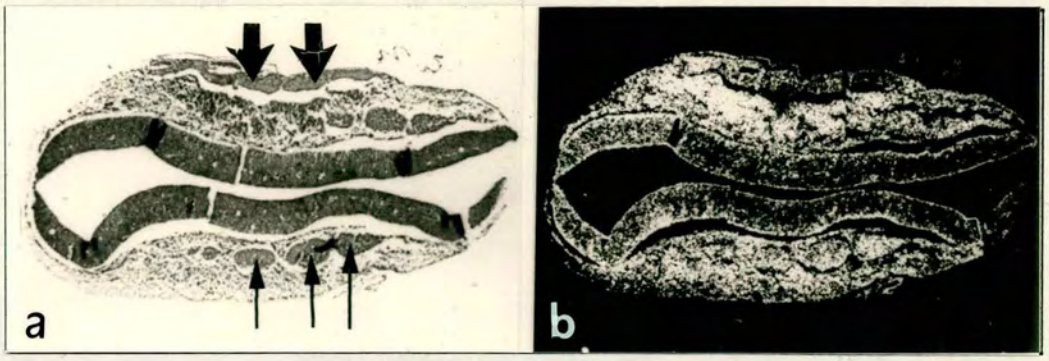
Figure 5.9

In situ hybridization photographs of a 9.5 d.p.c. mouse embryo (Theiler stage 15) (x100) showing a) bright field section and b) dark field section. a, b) This shows a transverse section through the trunk of the embryo. The thin arrows show the position of the dorsal root ganglia and the thick arrows indicate the somites which are already starting to break up at this stage in the embryo. This photograph shows the non-somitic mesenchyme is more highly labeled than the somitic mesenchyme and that the presumptive dorsal root ganglia are also highly labeled.

c) Colour photograph of *in situ* hybridization of a 9.5 day mouse embryo (x85). The positions of the mesonephros (Ms), undifferentiated mesenchyme (M) and neural tube (arrow) are indicated.

This photograph shows the increased expression in the undifferentiated mesenchyme (M), the lower expression in the mesonephros which has started to differentiate and partially epithelialise. The neural tube (arrow) can also be seen to be labeled at a low level.

The planes of section in these photographs are indicated below on the diagram of a 9.5 d.p.c. embryo (adapted from Kaufman, 1992) and labeled a, b and c corresponding to the figure panels.



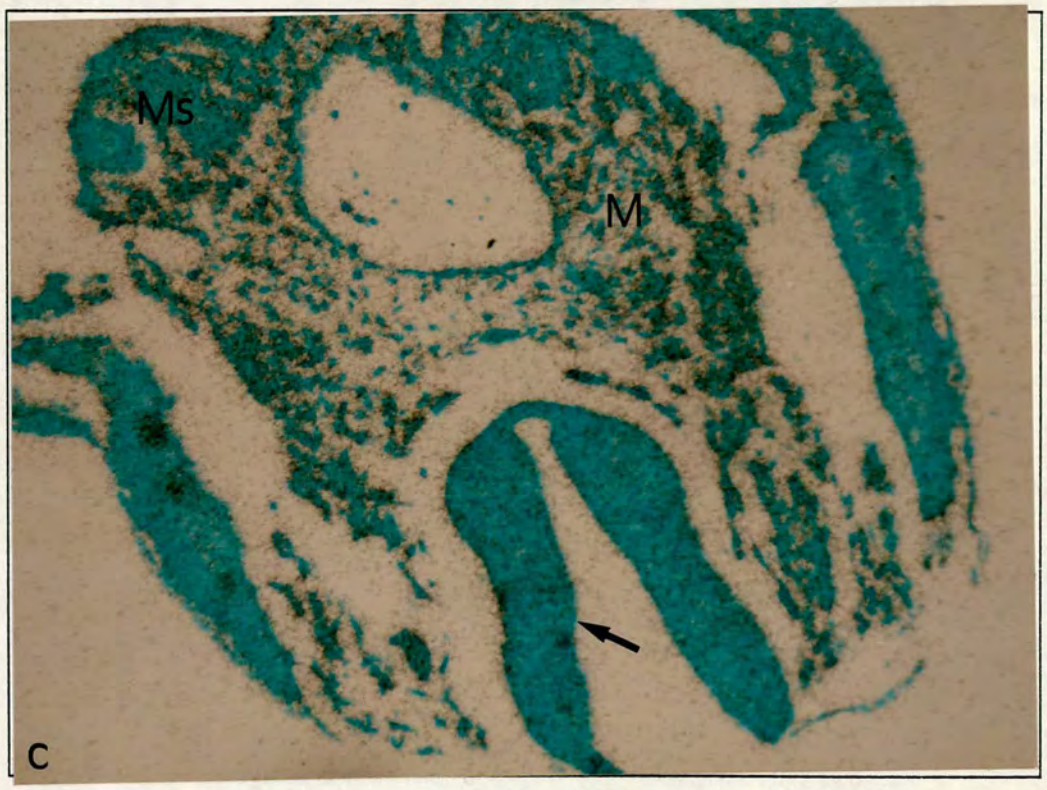


Figure 5.9a, b, c

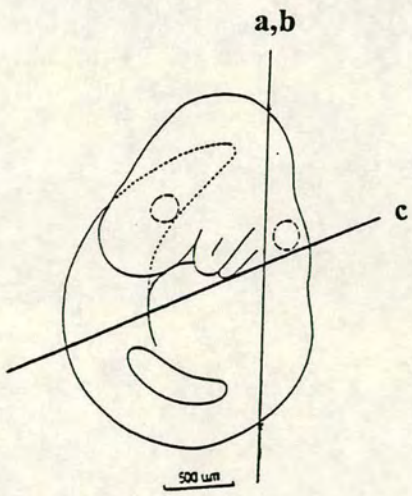


Figure 5.10

Whole-mount *in situ* hybridization photograph of a 9.5 d.p.c. embryo probed with DIG- labeled anti-sense transcript from the mouse 14-3-3 ε isoform cDNA. This photograph shows that there is widespread expression at this stage giving a diffuse overall stain to the embryo.

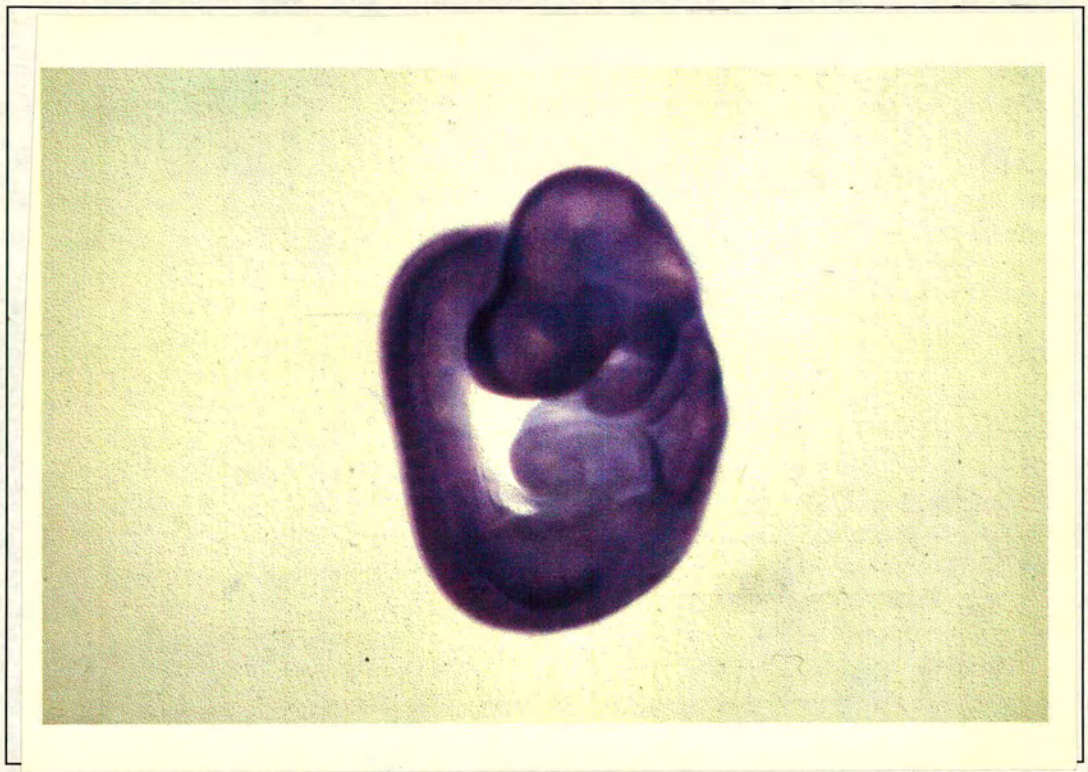


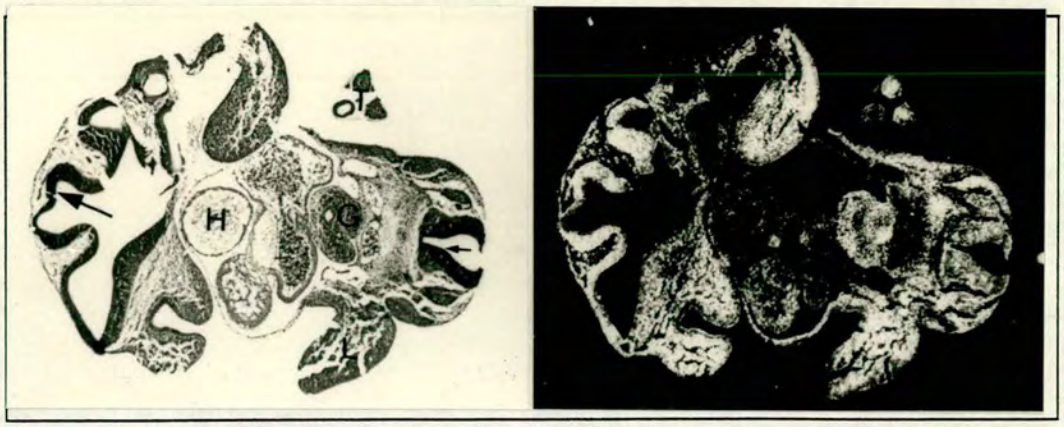
Figure 5.10

Figure 5.11

In situ hybridization photographs of a 10.5 day mouse embryo (Theiler stage 16) (x64) showing a) bright field section and b) dark field section. a, b) This shows a transverse section through the embryo. The positions of the gut (G), heart (H), tail (T), a limb bud (L) and neural folds (arrow) are indicated. This shows that the expression in the nervous system has become very pronounced, expression is still high in the mesenchyme but is very low in the epithelium. This can be seen in the limb (L) where the mesenchyme is strongly labeled whilst the outer area is less strongly labeled. This can also be seen in the early gut (G) where the epithelial regions, the presumptive serosal and submucosal epithelia expressed the gene at a low level, while the presumptive muscle layer showed high levels of expression. It can also be seen that the heart muscle (H) showed a relatively low level of expression.

c) Colour photograph of *in situ* hybridization of a 10.5 day mouse embryo (x140). The positions of the neural tube (N), ventral horn (Vh) and a dorsal root ganglia (DRG) are indicated. This photograph shows the strong expression in the neural tube except in the ventral horn region where gene expression is at a lower level.

The planes of section in the photographs are shown below on a diagram of a 10.5 d.p.c. embryo (adapted from Kaufman 1992) and labeled a, b and c corresponding to the figure panels..



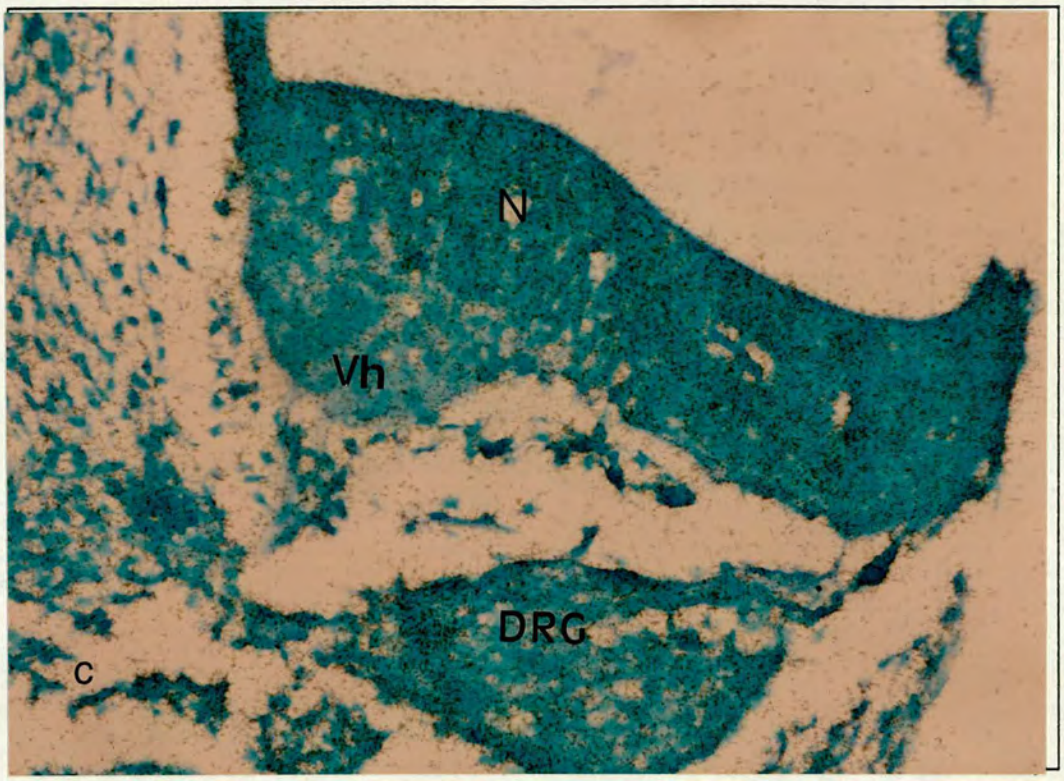


Figure 5.11 a, b, c

a,b,c

liver and red blood cells (data not shown). Expression in the brain is strong, with the diencephalon area showing particularly strong expression, as do all parts of the neural tube (indicated by an arrow in Figue 5.11a and by an N in Figure 5.11c) except the ventral horn region (indicated by a Vh in Figure 5.11c) where gene expression is at a lower level (Figure 5.11c). The level of the expression in the eye is high but lower than mesenchymal levels (data not shown).

All undifferentiated mesenchyme showed high expression and the difference between epithelial and mesenchymal expression levels could be seen clearly in the early gut (indicated by a G in Figure 5.11a) and lung buds [data not shown but similar to day 11.5 d.p.c. i.e. Figure 5.13] where the epithelial (presumptive serosal and submucosal for the gut and bronchi for the lung) expressed the gene at a low level, while the presumptive muscle layer and lung mesenchyme showed high levels of expression. However the heart muscle (indicated by a H in figure 5.11a) showed a relatively low level of expression.

The whole-mount *in situ* hybridization (for both 10.5 and 12.5 d.p.c. embryos) was performed alongside a whole-mount *in situ* hybridization using an anti-sense transcript of the mouse *Hox-7.1* as a probe. The regionalization of the signal for *Hox-7.1* and the similar quickly developing and tissue-specific staining pattern for the mouse 14-3-3 ε cDNA suggests that the whole-mount *in situ* hybridization was specific. Later experiments to repeat these whole-mount *in situ* hybridization experiments have included a sense control for the mouse ε cDNA (data not shown). Hybridization with sense transcripts gave a weak signal, only after prolonged incubation in the developing solution, primarily trapping of the probe in the diencephalon which is typical of other sense controls performed in the laboratory.

Figure 5.12

Whole-mount *in situ* hybridization photograph of a 10.5 d.p.c. embryo probed with DIG-labeled mouse 14-3-3 ε isoform anti-sense transcripts (right) alongside another embryo of the same age probed with a control *in situ* hybridization probe (*Hox-7.1*).

The whole-mount staining shows expression on the edge of the developing fore limbs and also in the hind limbs in the undifferentiated mesenchyme. Expression in the maxillary and mandibulary processes, and the front nasal processes is pronounced with a general pattern of being strongest at the edges of the processes. There is also strong expression in the tail, general weak expression along the neural tube, and weak expression in the genital ridge.

The staining pattern seen here has been reiterated by the *in situ* hybridization to sections of embryos of the same age. The expression in limb buds is also seen in figure 5.11a, whilst the neural expression is also seen in figures 5.11a, b and c.

The staining of the genital ridge is probably an early sign of the high expression that is seen in the genital tubercule of the 12.5 and 13.5 d.p.c. embryo, the early seminiferous tubules of the 15.5 d.p.c. embryo and the adult testes (detected by northern analysis).

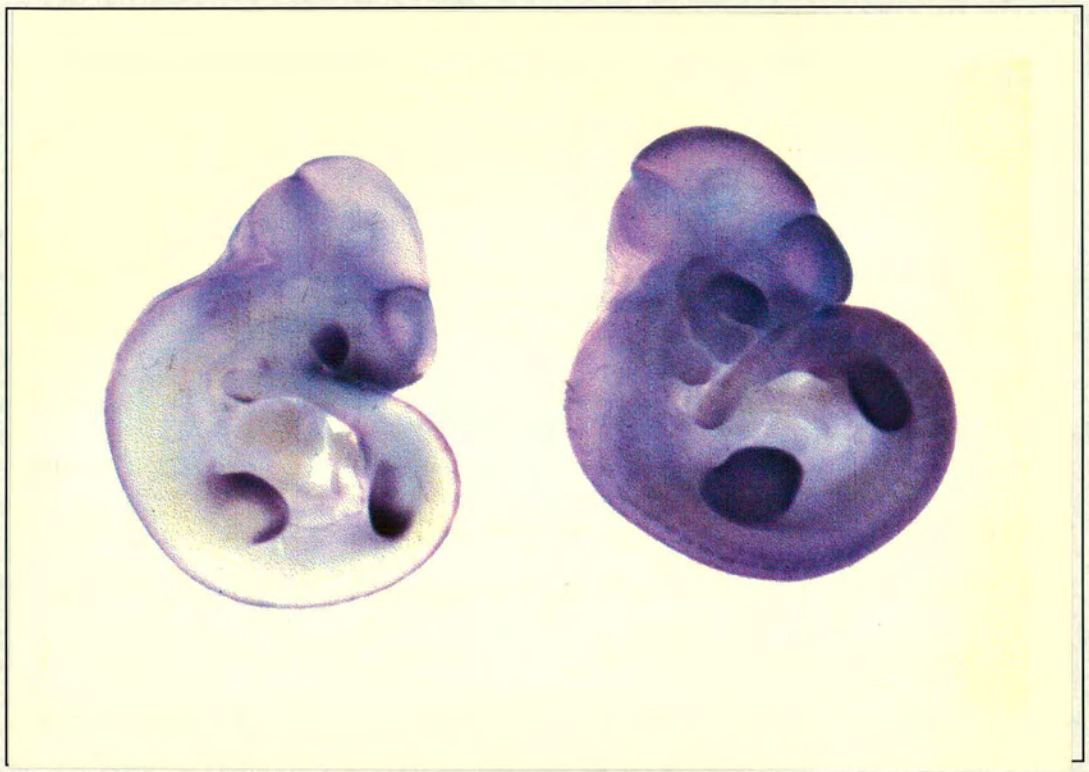


Figure 5.12

The whole-mount staining showed expression on the edge of the developing fore limbs and also in the hind limbs in the undifferentiated mesenchyme. This staining pattern is confirmed by the signal seen in sections of the limb buds in Figure 5.11a. Expression in the maxillary and mandibulary processes, and the front nasal processes was pronounced with a general pattern of being strongest at the edges of the processes. There was also strong expression in the tail, general weak expression along the neural tube, and weak expression in the genital ridge (Figure 5.12). Again this staining pattern is confirmed by the strong staining seen in the neural tissue in Figure 5.11a, b and c. The staining of the genital ridge is probably an early sign of the high expression that is seen in the 12.5 and 13.5 d.p.c. genital tubercule, the 15.5 d.p.c. early seminiferous tubules and the adult testes.

In general, the gene was most highly expressed in the 11.5 d.p.c. embryo. Expression in most neural tissues such as the brain, dorsal root ganglia and the neural tube was high, an exception being the ventral horn of the neural tube where the expression levels were low. This is similar to the situation a day earlier shown in Figure 5.11c. All mesoderm-derived tissues such as somites, the notochord and mesenchyme showed expression, with levels being higher in the lung bud than the adjacent heart (shown in Figure 5.13, L indicates lung bud and H heart). Levels of expression in the epithelia were low but positive unlike in the red blood cells where no expression could be seen (indicated by RBC in Figure 5.13).

In 12.5 d.p.c. embryos the pattern of expression was generally similar to expression in embryos a day earlier (see Section 5.3). Although the level of expression in epithelial tissues was very low, neural tissue such as cerebral cortex and dorsal root ganglia showed intermediate levels and undifferentiated mesenchyme very high levels. The differences between epithelial and mesenchymal cells were seen most clearly in ducted glands such as the lung where the epithelial linings of the

segmental bronchioles have low levels of expression while the interstitial mesenchyme showed a high level of expression (Figure 5.15c and d).

As the mesenchyme differentiates, expression levels often decline. This effect is clearly seen in three tissues, gut, bone and kidney. The situation in the gut is complex (Figure 5.15a and b, 5.17b) the gene is strongly expressed in the intermediate layer of mesenchyme which becomes muscle (indicated by an arrow in figure 5.15a and by im in Figure 5.17), weakly expressed in the epithelia layer, but absent in the gastric, mucosal lining which is of endodermal derivation. In developing bone, the level of expression in cartilage is lower than that of the surrounding mesenchyme and even lower in the peripheral regions that were ossifying. In the early rib primordia, for example, the gene is expressed in the intercostal region, but not by the differentiated cartilage of the vertebrae (Figure 5.15e arrow). In this forming primordia, the gene is expressed in the inner, undifferentiated region, but not in the outer, ossifying layer (Figure 5.15e arrow head).

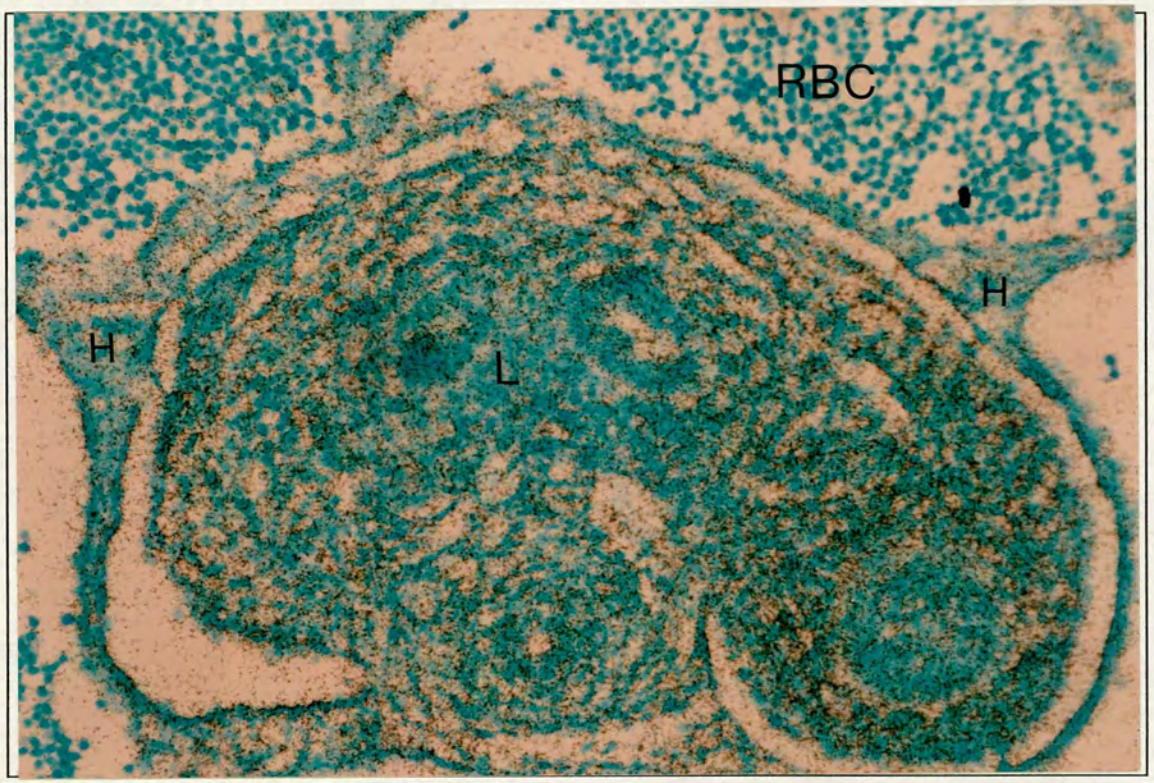
The different structures of the kidney can be clearly identified by day 12 and high expression can be found in the blastema and mesenchymal cells (Figures 5.15g, h and 5.16), while the epithelial collecting ducts and nephrons [which are derived from mesenchyme (Saxen 1987)] showed much lower levels of expression. In common with most other epithelial tissues the ureteric bud derivatives (Figure 5.15g arrow head) and the differentiating nephrons (indicated by an arrow in figure 5.15g and by N in figure 5.16), both of which are epithelial, show very low expression. The gene is expressed in tooth primordia but at a lower level than the adjacent tongue and jaw mesenchyme. The tooth primordia form from an infolding of the ectodermally-derived, stratified, surface epithelium which can be seen in Figure 5.17a and the expression is concentrated within the infolding. This is a rare example of epithelium showing moderate expression but it can be seen that the expression level is lower

Figure 5.13

Colour *in situ* hybridization photograph of a 11.5 d.p.c. mouse lung and surrounding tissue (x140). The positions of the lung (L), heart (H) and red blood cells are indicated .

The photograph shows the high expression of the 14-3-3 ε mRNA, with levels being higher in the lung bud (L) than the adjacent heart (H). Levels of expression in the epithelia are low shown in the epithelial regions of the lung but were positive unlike the red blood cells (RBC) where no expression could be seen.

The plane of section of this photograph is indicated on a diagram of a 11.5 d.p.c. embryo (adapted from Kaufman, 1992).



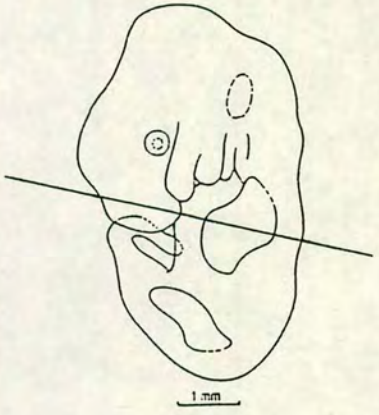


Figure 5.13

Figure 5.14

Whole-mount *in situ* hybridization photograph of a 12.5 d.p.c. embryo probed with DIG-labeled mouse 14-3-3 ε isoform antisense transcripts (right) alongside another embryo probed with a control *in situ* hybridization probe (*Hox-7.1*).

The hair follicles in the face show expression; comparison of *in situ* hybridization sections containing hair follicles show that the expression is concentrated underneath the ridges that are the basis of the whisker rudiments in the mesenchyme that will form the follicles. There is weak expression in the eyes, and the expression in the limbs has become more restricted to interdigital regions. This may be indicative of staining of the undifferentiated mesenchymal regions of the limb. The expression is still strong in the tail and has increased to a high level in the genital area, the umbilicus is also very strong.

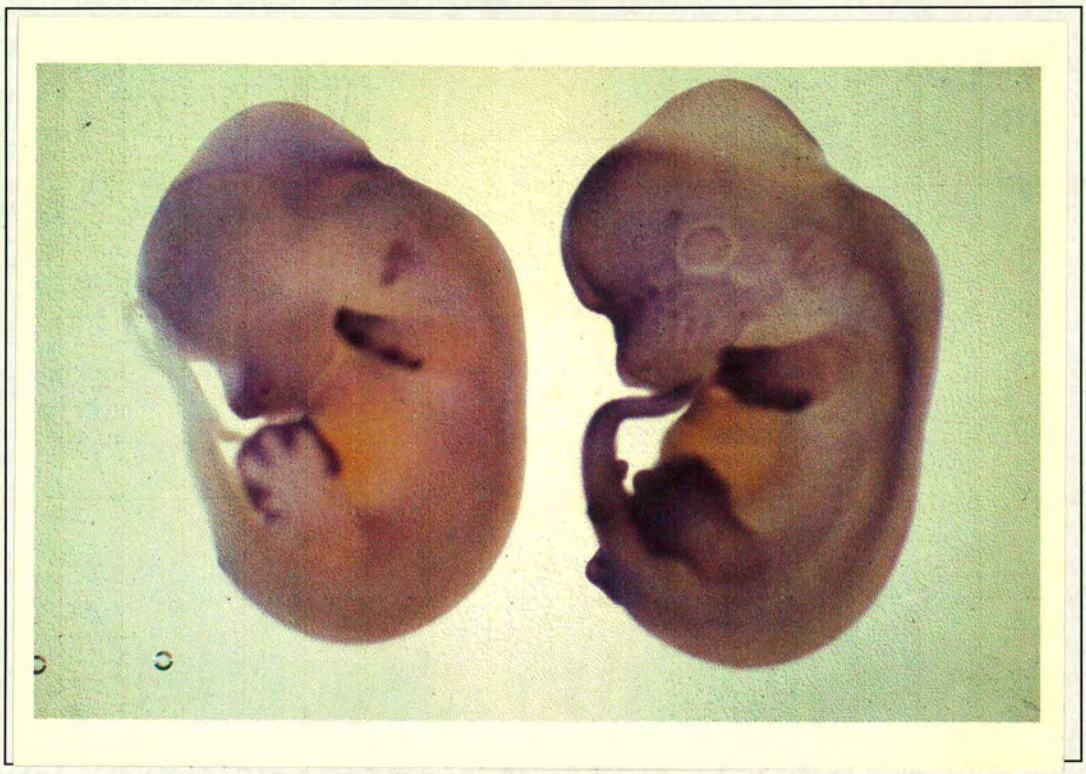


Figure 5.14

Figure 5.15

In situ hybridization photographs of parasagital sections (bright field sections a, c, e, g and dark field sections b, d, f, h) of a 12.5 day mouse embryo (Theiler stage 21).

a, b) The gut (bar 200 μm; x50). The gene is strongly expressed in the intermediate layer of mesenchyme which becomes muscle (arrow), weakly expressed in the epithelia layer, but absent in the gastric, mucosal lining which is of endodermal derivation.

c, d) The lung (bar 200 μm; x44). The lung shows the differences between epithelial and mesenchymal cells most clearly, the epithelial linings of the segmental bronchioles have low levels of expression while the interstitial mesenchyme show a high level of expression.

e, f) Early rib primordia (bar 200 μm; x42). In the early rib primordia, the gene is expressed in the intercostal region, but not by the differentiated cartilage of the vertebrae (arrow). In this forming primordia, the gene is expressed in the inner, undifferentiated region, but not in the outer, ossifying layer (arrow head). This expression pattern is generally observed in developing bone, the level of expression in cartilage being lower than that of the surrounding mesenchyme and even lower in the peripheral regions that are ossifying.

g, h. The metanephros (bar 200 μm; x80). High expression can be seen in the blastema and mesenchymal cells (figures g, h and 5.16), while the epithelial collecting ducts and nephrons (which are originally derived from mesenchyme) show much lower levels of expression. In common with the other epithelial tissues the ureteric bud derivatives (arrow head) and the differentiating nephrons (arrow), both of which are epithelial, show very low expression.

The plane of section in the photographs is indicated below on a diagram of a 12.5 d.p.c. embryo (adapted from Kaufman, 1992).

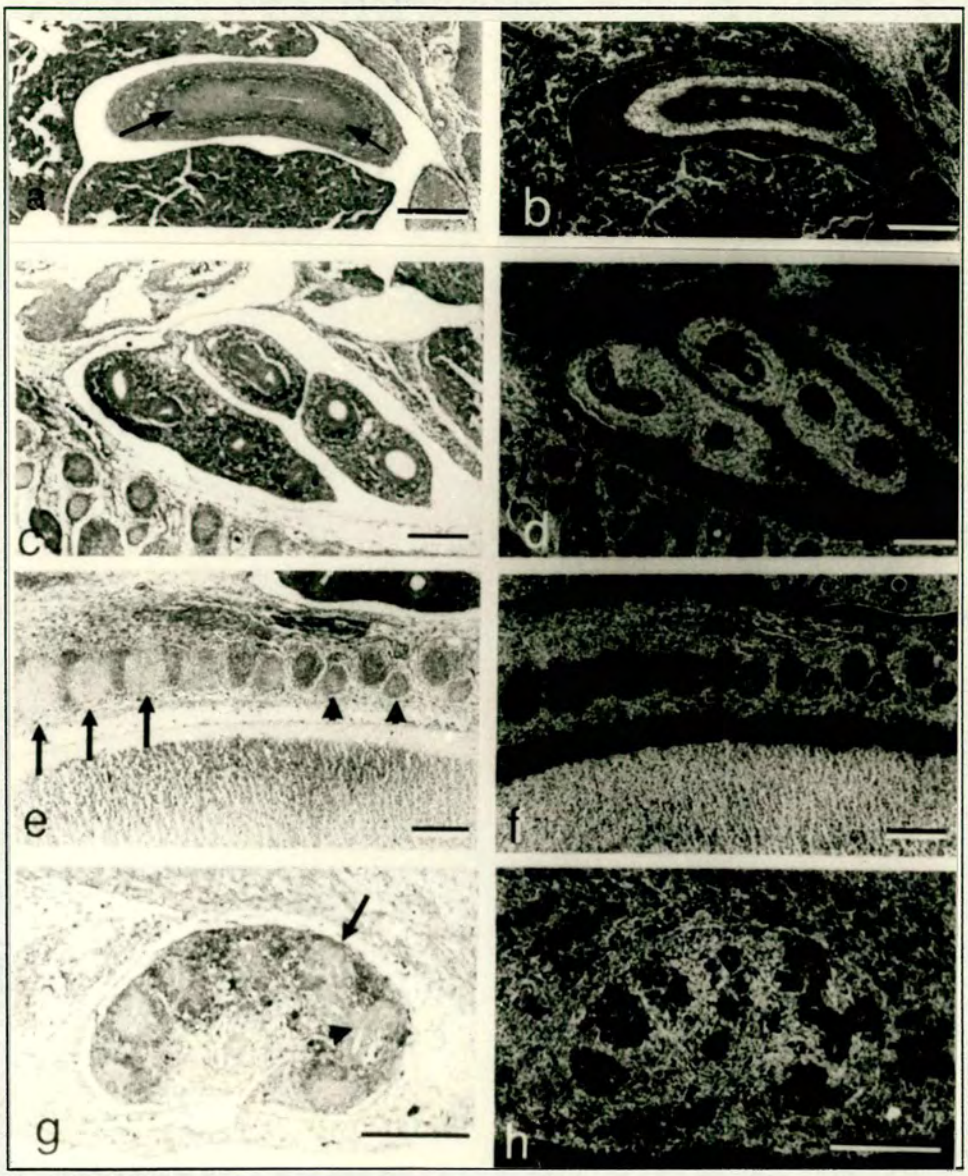


Figure 5.15

a-h

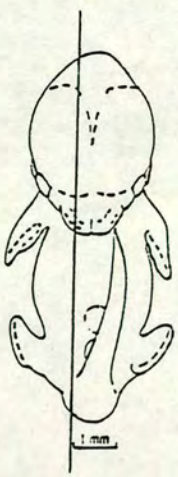


Figure 5.16

Colour *in situ* hybridization photographs of the 12.5 d.p.c. fetal kidney shown in the previous figure (5.15g and h) photographed at a higher magnification (x140), The mesenchyme (m), stem cells (s), a nephron (n) and a duct (d) are indicated.

This photograph shows the higher expression in the undifferentiated mesenchyme (m) and blastema compared to the lower expression in the ducts (d) and the differentiating epithelialised nephron (n).

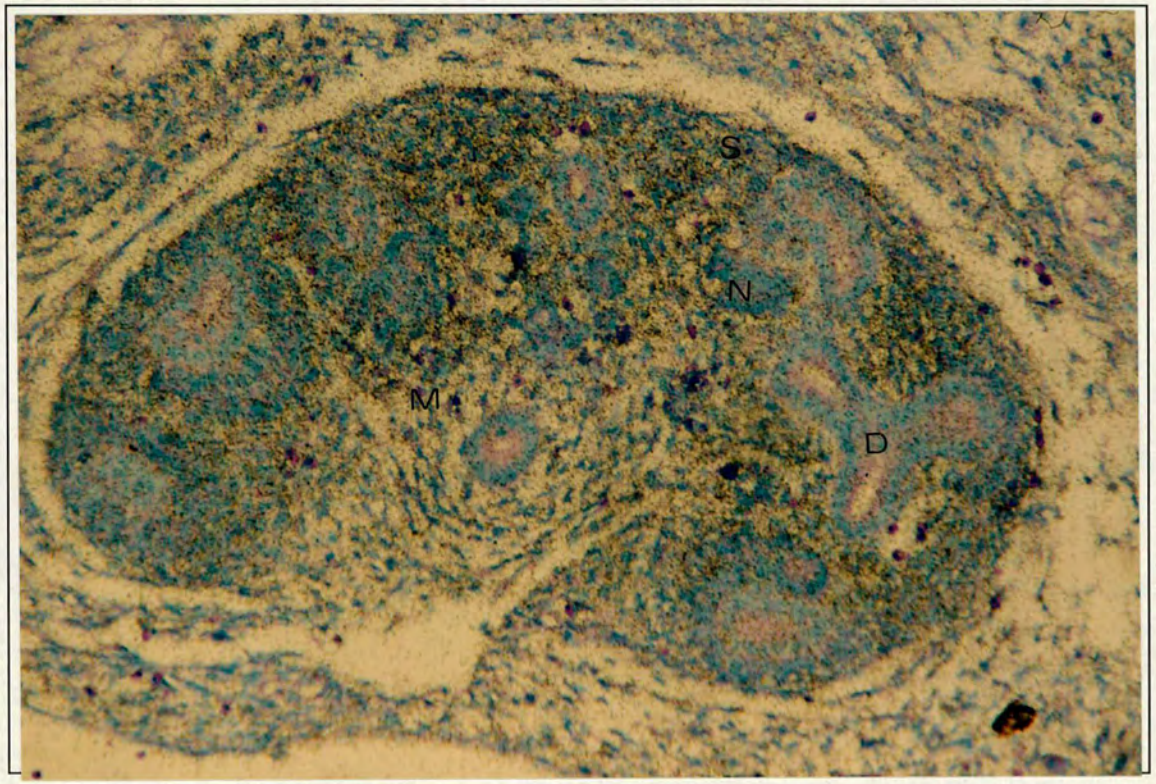


Figure 5.16

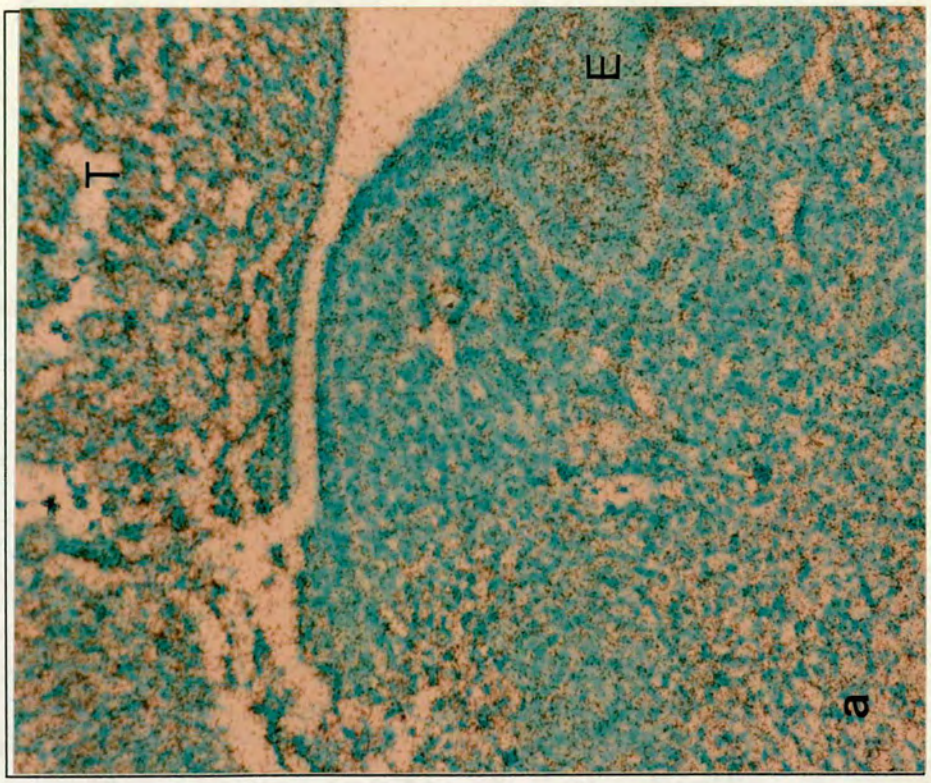
Figure 5.17

a) Colour *in situ* hybridization photograph of a 12.5 d.p.c. embryo section through the jaw area (x140).

The tongue (T) and the invaginating ectodermally-derived, stratified, surface epithelium (E) that will form the tooth are indicated. The gene is expressed in tooth primordia but at a lower level than the adjacent tongue and jaw mesenchyme. This shows a rare example of epithelium with moderate expression but it can be seen that the expression level is still lower than mesenchymal expression levels (the tongue and jaw).

b) Colour *in situ* hybridization photograph of a 12.5 d.p.c. gut section (x140). The intermediate mesenchyme (im) is indicated. The gene is strongly expressed in the intermediate layer of mesenchyme which becomes muscle (im), and weakly expressed in the epithelia layer, but absent in the gastric, mucosal lining which is of endodermal derivation. This photograph shows that expression of the gene in the gut is concentrated in the intermediate mesenchyme layer.

The plane of section in the photographs is indicated on a diagram of a 12.5 d.p.c. embryo (adapted from Kaufman,1992).



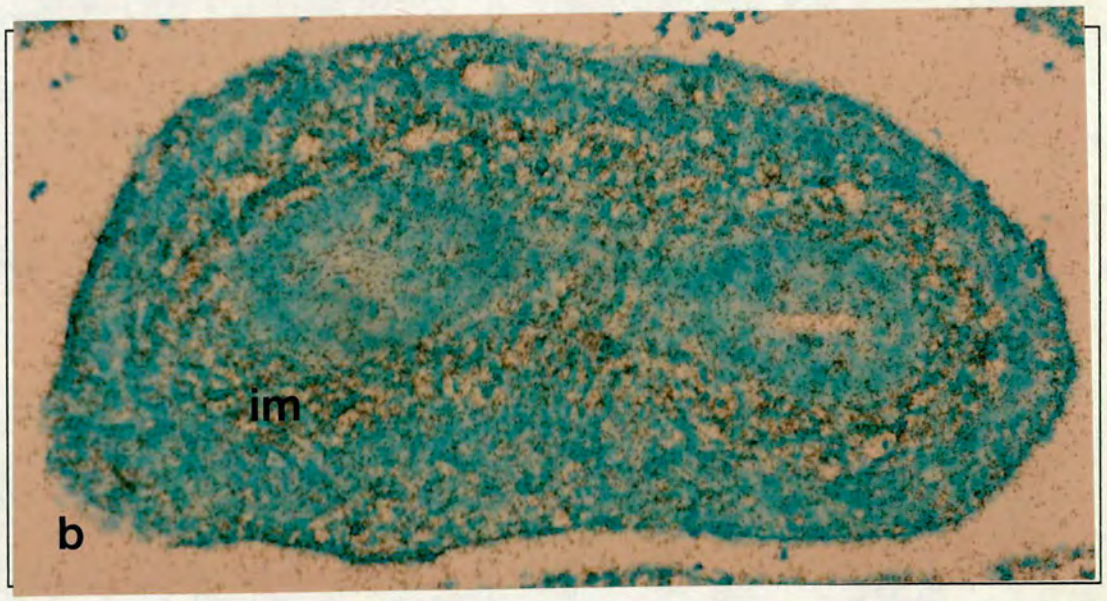


Figure 5.17

Figure 5.18

In situ hybridization photographs (bright field and dark field sections) of a 12.5 day mouse embryo (Theiler stage 21) (x20). a, b. A parasagital section of the embryo. The positions of the liver (Li), lung (L), heart (H), tongue (T) and genital tubercule (Gt) are indicated.

This photograph gives a general impression of the expression levels in the 12.5 d.p.c. embryo. Details are provided in the text.

Figure 5.19

In situ hybridization photographs (bright field dark field sections) of a 13.5 day mouse embryo (Theiler stage 22) (x18). a, b. A parasagital section of the embryo. The positions of the liver (L), lung (l), kidney (k), heart (H), tongue (T) and genital tubercule (Gt) are indicated.

This photograph gives a general impression of the expression levels in the 13.5 d.p.c. embryo. Details are provided in the text.

The planes of section for the photographs are indicated below on diagrams of a 12.5 and 13.5 d.p.c. embryo respectively (adapted from Kaufman, 1992).

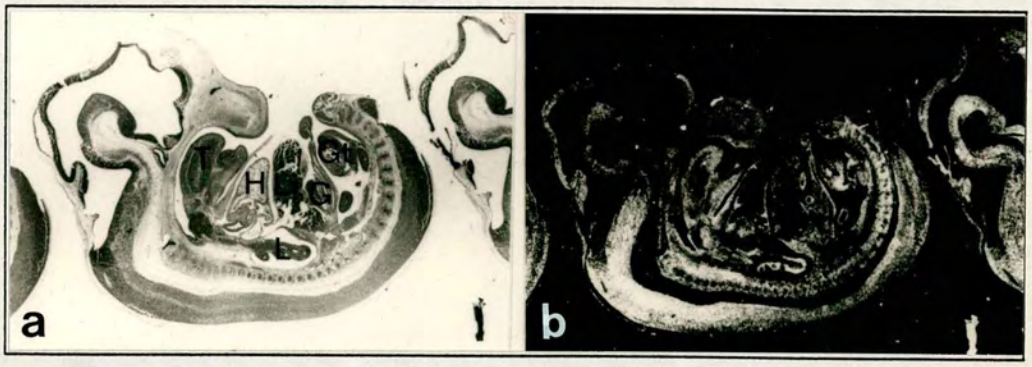


Figure 5.18

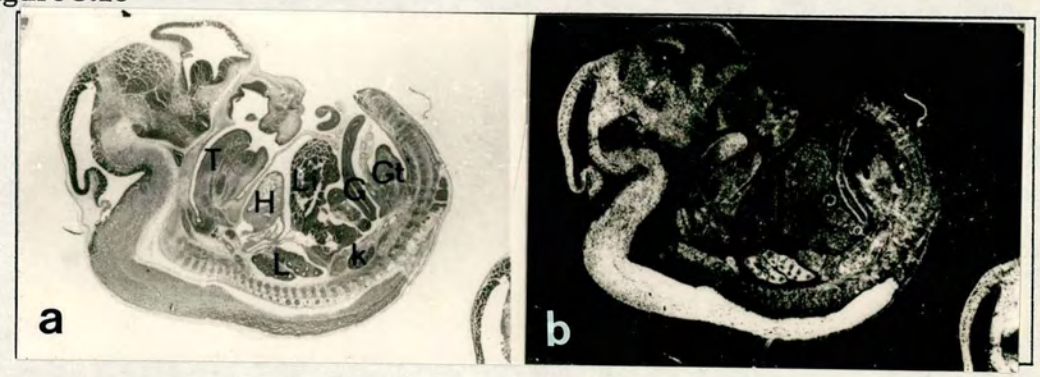
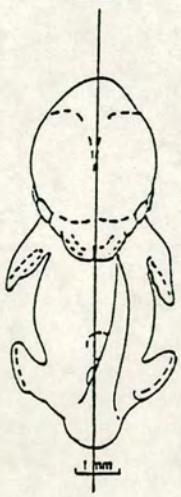


Figure 5.19

5.18



5.19

than mesenchymal expression levels (the tongue and jaw). The gene is also highly expressed in the genital tubercule (indicated by Gt in Figure 5.18).

Looking at whole-mount embryos, the hair follicles in the face are positive. More accurate localization of the signal on *in situ* hybridization sections containing hair follicles shows the expression is concentrated in the underlying mesenchyme underneath the ridges that are the basis of the whisker rudiments. This is another example of the gene being concentrated in the undifferentiated mesenchyme. There is weak expression in the eyes, and the expression in the limbs has become more restricted to interdigital regions. This may be indicative of staining of the undifferentiated mesenchymal regions of the limb. The expression is still strong in the tail and has increased to a high level in the genital area, and is also very strong in the umbilicus (Figure 5.14).

By 13.5 d.p.c. expression levels generally are comparatively lower than at 12.5 d.p.c., except in neural and some mesenchymal tissue (Figure 5.19). The level of expression in the brain, spinal cord and dorsal root ganglia is still as high if not higher than 12 d.p.c. with the roof of the third ventricle (R in Figure 5.20a) and neopallial cortex expressing the gene very highly. The gut, lung and kidney (Figure 5.20b) mesenchyme is still expressing the gene highly although the epithelial expression is almost non-detectable.

In the 15.5 d.p.c. embryo, a general drop in expression was seen in most tissues. The gut and lung mesenchyme now only show low expression where the expression had previously been high, but the mesenchyme around the cartilage appears to still express the gene at a moderate level. In epithelial tissues the expression was noticeably very low whilst the liver (marked L in Figure 5.21e and f) has moderate expression levels. In the kidney (indicated by a K in Figure 5.21e and f, expression was low in both the collecting ducts and tubules, but was still at intermediate levels in the glomeruli, stem cells, early condensates and S-shaped bodies.

Figure 5.20

Colour *in situ* hybridization photographs of a 13.5 d.p.c. mouse embryo (Theiler stage 22) showing the A) roof of the hindbrain (R) (x140), and B) the kidney [x 140]. The photograph shows that the level of expression in the roof of the third ventricle (R in figure 5.20a) is very high. There is also high expression in the brain, spinal cord and dorsal root ganglia at this stage of development.

In B) the mesenchyme (M), a comma-shaped body (c) and the stem cells (s) are indicated.

The photograph shows that expression is high in the kidney mesenchyme (M) and in the stem cells but the expression in the epithelial structures such as the comma-shaped body (C) was almost non-detectable.

The plane of section in the photograph is indicated here on a diagram of a 13.5 d.p.c. embryo (adapted from Kaufman, 1992).

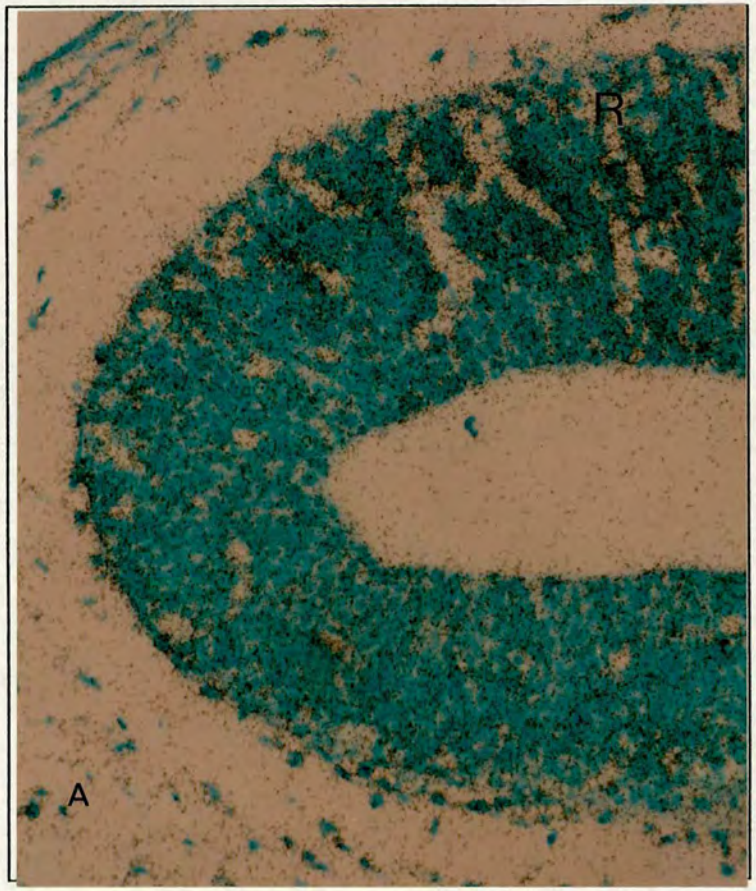


Figure 5.20A

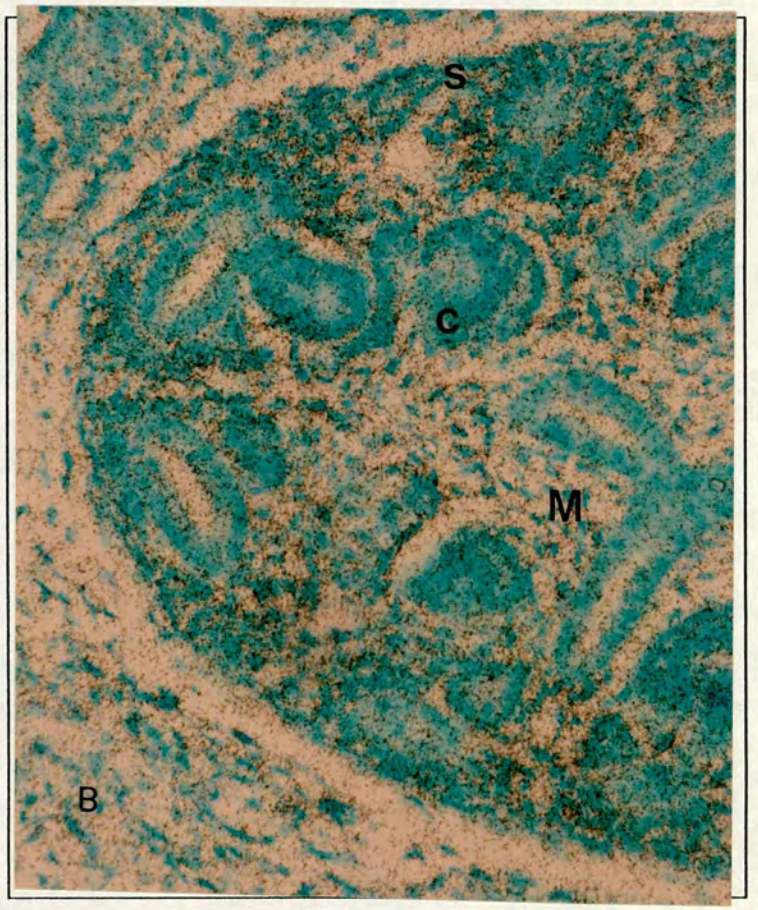


Figure 5.20B

The expression levels were still very high in the brain and dorsal root ganglia (indicated by thick arrows in Figure 5.21a, b). More specifically there is high expression in the diencephalon (indicated by D in Figure 5.21a, b, the neopathial cortex (indicated by Nc in Figure 5.21a, b) and the roof of the mid brain (indicated by a thin arrow in figure 5.21a, b).

High expression can also be seen in the presumptive molar and incisor teeth, this result is of interest as the expression pattern by whole-mount *in situ* hybridization for the 14-3-3 e gene is very similar to the pattern seen for the *Hox-7.1* gene (Figures 5.12 and 5.14) and *Hox 7.1* has also been shown to be expressed in the developing tooth (Mackenzie *et al.*, 1991a, b). The areas of expression of the mouse 14-3-3 ε gene detected by whole-mount *in situ* hybridization were similar to the areas of expression of *Hox-7.1* expression, but slightly less restricted to defined areas. (This effect was noticed as *Hox-7.1* is a control reaction in the laboratory performed alongside all reactions to ensure that the protocol is working).

In summary, the 14-3-3 ε isoform gene is expressed ubiquituously at low levels in the early embryo. Later, expression levels become higher in undifferentiated mesenchyme but the level decreases as the mesenchyme differentiates. Conversely expression rises in the nervous system as it differentiates.

The mouse ε cDNA was isolated three times by the subtractive hybridization selection criteria. Reviewing the expression data, both northern blot and *in situ* hybridization analysis, (both whole-mount and sections), an expression profile for this cDNA can be inferred. Consideration of this expression profile and the functional data currently available about the 14-3-3 gene family helps to understand the possible function that this gene may play in development.

The expression of the gene in the nephric system has been studied during development. In the early mesonephros, expression is low in the epithelial structures in comparison to the higher expression levels in the surrounding mesenchyme

(Figure 5.23a). In 11 d.p.c. embryo the metanephros has started to form but is still uninduced. At this stage a moderate expression level can be seen in the uninduced kidney mesenchyme and a very low level in the ureteric bud (data not shown). By day 12.5 d.p.c., a day or so later, expression is high in the blastema and mesenchymal cells, while the epithelial collecting ducts and nephrons (which are originally derived from mesenchyme) show much lower levels of expression. In common with the other epithelial tissues the ureteric bud derivatives (arrow head) and the differentiating nephrons (arrow), both of which are epithelial, now show very low expression (Figure 5.23b, c and d). A day later at 13.5 d p.c. expression is still high in the kidney mesenchyme (M) and in the stem cells but the expression in the epithelial structures such as the comma-shaped body (C) is almost non-detectable (Figure 5.23e). By day 15.5 p.c., expression is still high in the kidney, although expression is low in both the collecting ducts and tubules, but is still at intermediate levels in the glomeruli, stem cells, early condensates and **S**-shaped bodies. Analysis of sections of adult kidneys does not show specific signal despite the fact that northern analysis suggests that expression is relatively higher in the adult than the embryonic kidney. Therefore expression appears to be present at low levels in the early nephric system and rises to become highly expressed in the uninduced mesenchyme, perhaps meeting a requirement for signal transduction in the mesenchyme preparing to be induced. As the mesenchyme is induced and epithelializes the expression levels drop. A basal amount of expression is maintained throughout the adult kidney but this is not restricted to certain structures. This may reflect a basal level that is maintained to allow the kidney to be able to react to its environment through the signal transduction pathway that the ε isoform is postulated to belong to.

A similar story can be seen in the early developing limb, where the mesenchyme of the 10.5 d.p.c. embryo (Figure 5.10a) has high expression but by day 12.5 d.p.c. the expression has become restricted to the interdigital areas of the limb (Figure

5.14). This story is repeated in many tissues such as the whisker rudiments, lung mesenchyme and bone; undifferentiated mesenchyme has a high expression level and as the mesenchyme differentiates the expression falls. The expression of the whisker rudiments is a clear example of this. Looking at the 12.5 d.p.c. embryo, in whole mount and sections, high expression is found in the mesenchymal whisker rudiments, (this is seen as expression in the whiskers in the whole-mount Figure 5.14) but by 15.5 d.p.c. expression is undetectable (Figure 5.21a, b, c, d).

An interesting expression pattern was seen in the testes, where other isoforms were also detected by northern analysis (Figure 5.4). The gene is expressed in the genital ridge of early 10.5 d.p.c. embryos as detected by whole-mount *in situ* hybridization, and by 12.5 and 13.5 d.p.c. expression was very strong in the genital tubercule (Figures 5.18 and 5.19). By 15 d.p.c. very strong expression was found in the early seminiferous tubules; unfortunately we could not detect female genital structures to compare expression levels. During testes development the gene is up-regulated to give extremely high expression in the adult testes but only moderate expression levels are found in the adult ovary (as detected by northern analysis Figure 5.4). This expression difference was not selected for but is very intriguing. It is tempting to speculate that the gene is involved in regulating a kinase pathway required for spermatogenesis which would identify an important function for this gene.

The expression in the neural system is very different. In the early embryo (seen in the 8.5 to 10.5 d.p.c. embryos here) the neural tissue has very low expression and as the neural system develops the expression increases. Work by other groups has shown that the ε isoform can participate in regulating the protein kinase C-dependent activation of the tyrosine and tryptophan hydroxylase that regulates the catecholamine and seratonin biosynthetic pathways (Isobe *et al.*, 1991). The up regulation of gene expression in the neural system is probably correlated to the

activation of this pathway during embryogenesis. Of course it may also function to regulate signal transduction pathways required for neurogenesis, however study of this second possibility would be difficult to carry out due to the primary function of this gene family.

The high expression in mesenchyme may be preparing the mesenchyme for differentiation, allowing conformational changes of kinases required in important pathways. My hypothesis would then suggest the pathways are required at a lower level in differentiated cells. The next chapter explains the previous work that has suggested these hypotheses. The possible function of the 14-3-3 gene family is discussed further in Chapter 7 and in particular considers the possible function of the gene during development in relation to what is already known about the possible functions of the other gene members. Further experiments are also discussed that may provide more information on this protein family.

Therefore to fully understand the gene I have isolated, I now go onto a review of the 14-3-3 protein family and try to show where the ε isoform fits into the picture and how it may be involved in regulation of developmental processes.

Figure 5.21

In situ hybridization photographs (bright field dark field sections) of a 15.5 day mouse embryo (Theiler stage 24) (bright field a, c, d and e and dark field b, c, d and f).

a, b) A parasagital section of the head of the embryo (x18). The positions of the neopallial cortex (Nc), diencephalon (D), and roof of the hind brain (line) are indicated.

This photograph shows that expression levels are very high in the brain and dorsal root ganglia (thick arrows). More specifically there is high expression in the diencephalon (D), the neopathial cortex (Nc) and the roof of the mid brain (thin arrow).

c, d) A parasagital section of the head of the embryo (x58). The positions of the presumptive upper molar tooth (m), tongue (T), presumptive lower incisor (I) and Meckel's cartilage (M) are indicated.

These photographs show there is high expression in the presumptive molar and incisor teeth.

e, f) A parasagital section of the trunk of the embryo (x22). The positions of the liver (Li), lung (L), gut (G), kidney (K), heart (H) and ribs (R) are indicated.

This photograph shows that expression is still high in the kidney, and is present at moderate level in the liver at this stage. In the kidney (K) expression was low in both the collecting ducts and tubules, but was still at intermediate levels in the glomeruli, stem cells, early condensates and **S**-shaped bodies. The gut and lung mesenchyme now only show low expression where the expression had previously been high, but the mesenchyme around the cartilage appears to still be expressing the gene at a moderate level.

The plane of section in the photographs is indicated below on a diagram of a 15.5 d.p.c. embryo (adapted from Kaufman, 1992).

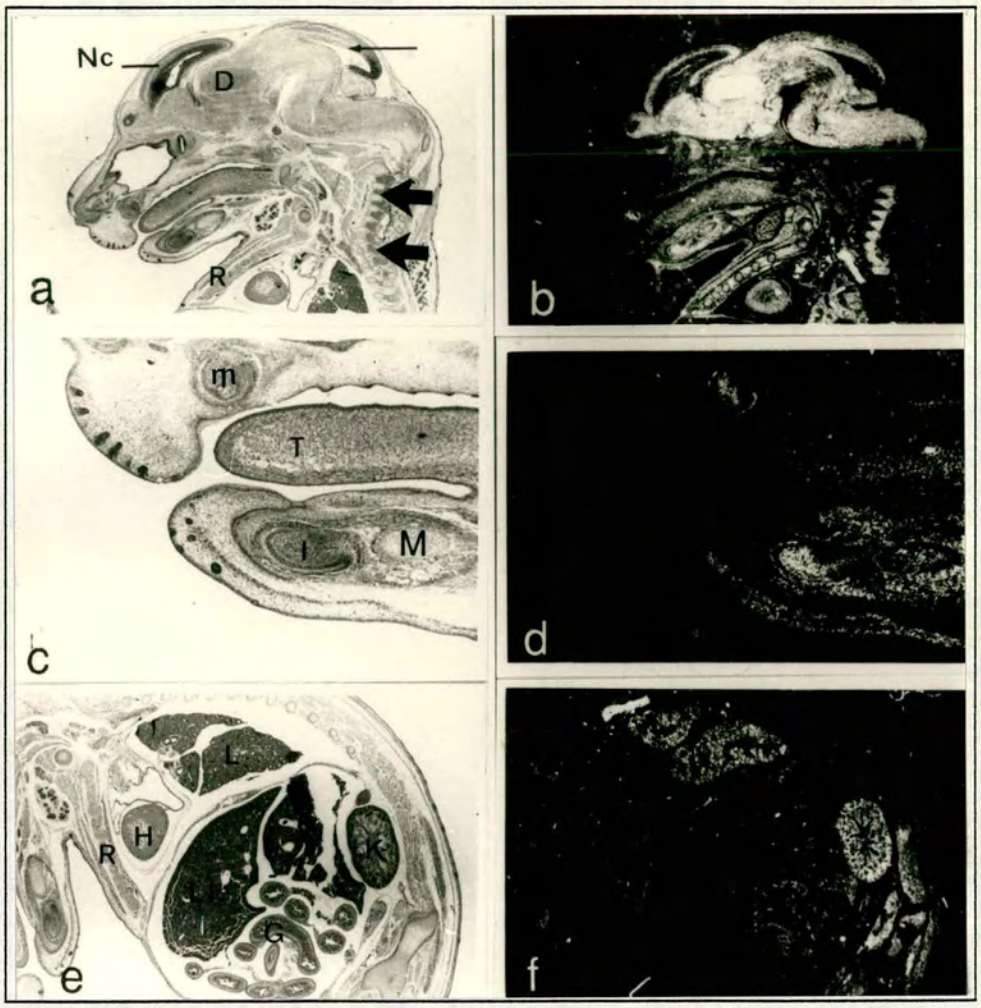


Figure 5.21

a-f

Figure 5.22

Colour *in situ* hybridization photograph of a section of 15.5 d.p.c. mouse ribs (x140). The bone (B) and mesenchyme (M) are indicated.

The photograph shows the mesenchyme around the cartilage appears to still be expressing the gene at a moderate level and the differentiated bone shows low expression levels.

The plane of section in the photograph is indicated below on a diagram of a 15.5 d.p.c. embryo (adapted from Kaufman, 1992).

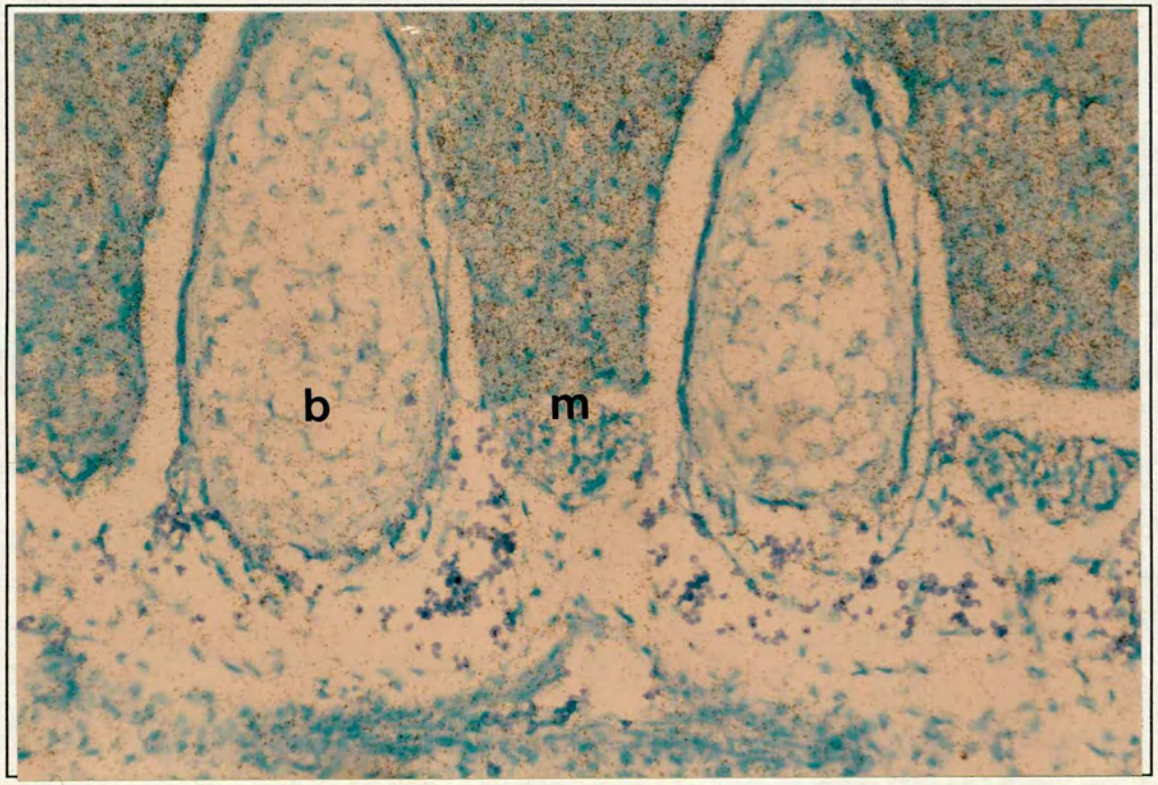


Figure 5.22

Figure 5.23

In situ hybridization of the mouse 14-3-3 ε cDNA to various stages of mouse embryos showing expression during kidney development. All section are parasagital, and correspond to the planes of section shown by Figures 5.9c, 5.15g and h, and 5.21e and f.

a) *In situ* hybridization photographs of a parasagital section of a metanephos (x100) from a 9.5 d.p.c. mouse embryo (Theiler stage 15). Left, bright field; Right, dark field.

b) *In situ* hybridization photographs of a parasagital section of a metanephros (x60) from a 12.5 d.p.c. mouse embryo (Theiler stage 21) Left, bright field; Right, dark field.

c) *In situ* hybridization photographs of a parasagital section of a metanephros (x60) from a 15.5 d.p.c. mouse embryo (Theiler stage 24) Left, bright field; Right, dark field.

Overleaf

d) Colour *in situ* hybridization photograph of a parasagital section of a metanephros from a 12.5 d.p.c. mouse embryo (Theiler stage 21), photographed at a higher magnification (x120).

e) Colour *in situ* hybridization photograph of a parasagital section of a metanephros (x 120) from a 13.5 d.p.c. mouse embryo (Theiler stage 22).

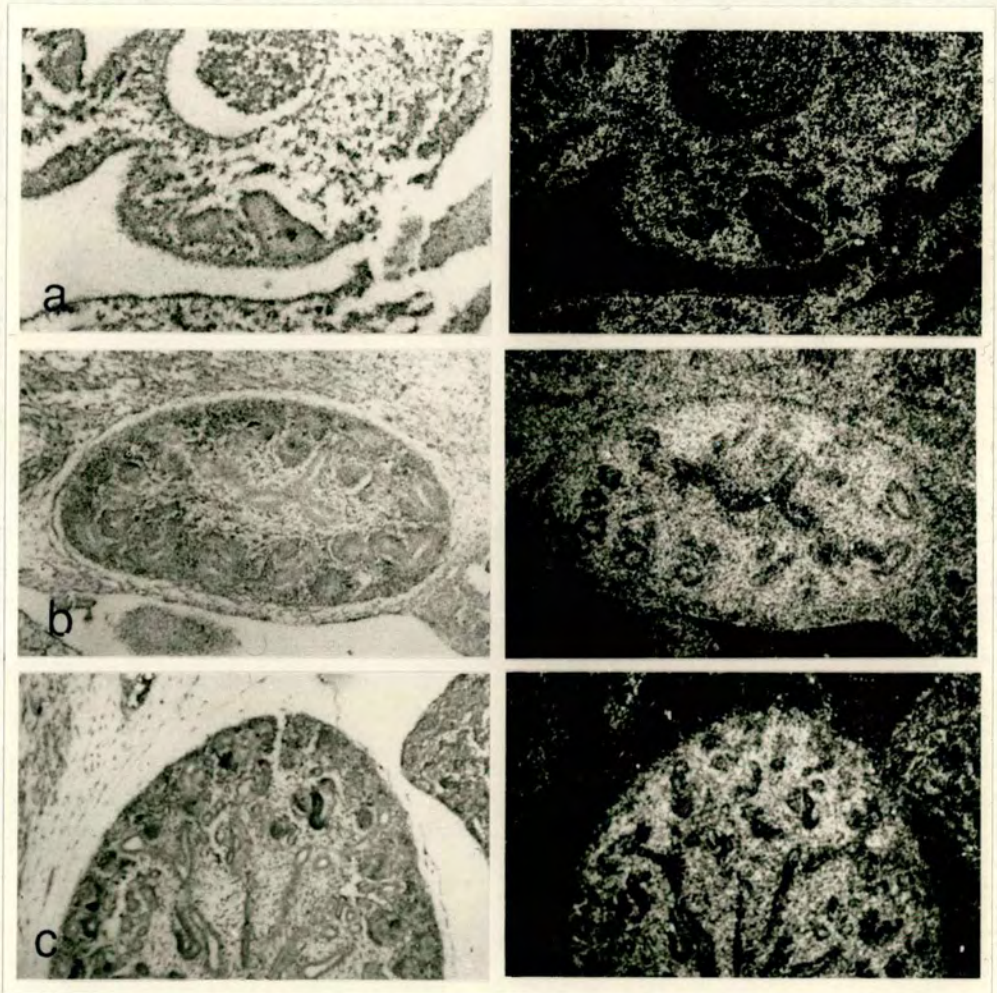
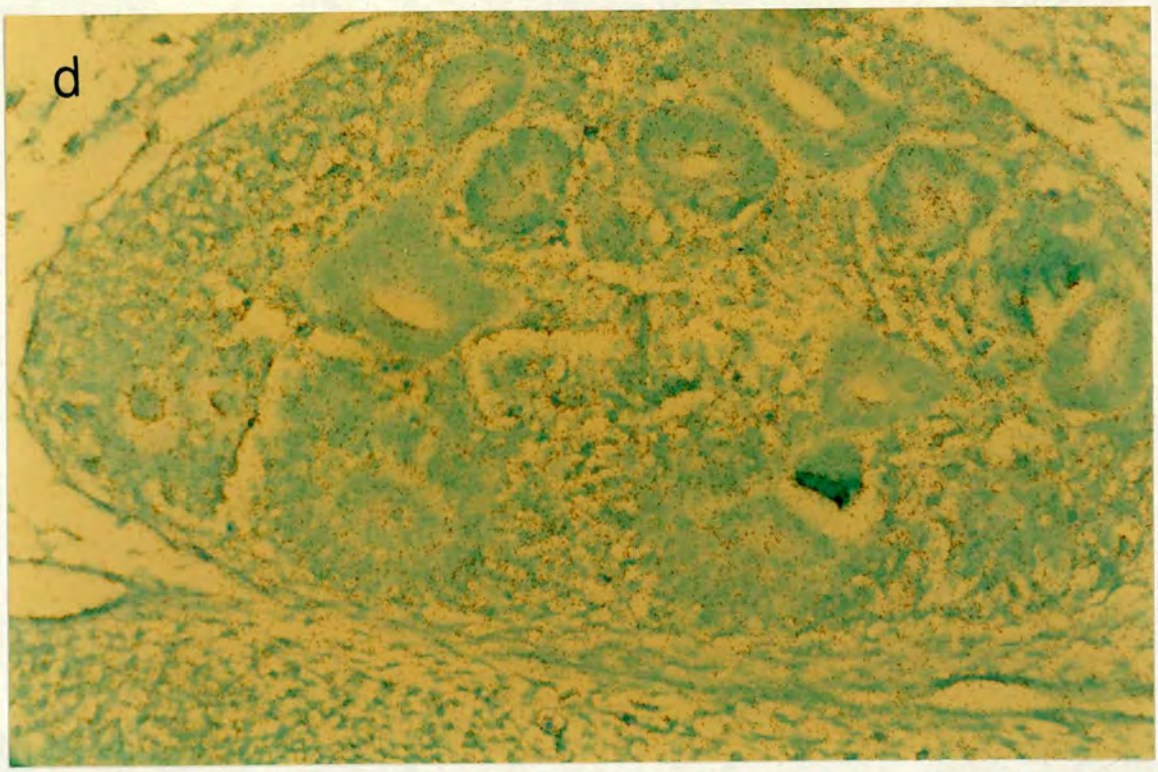


Figure 5.23



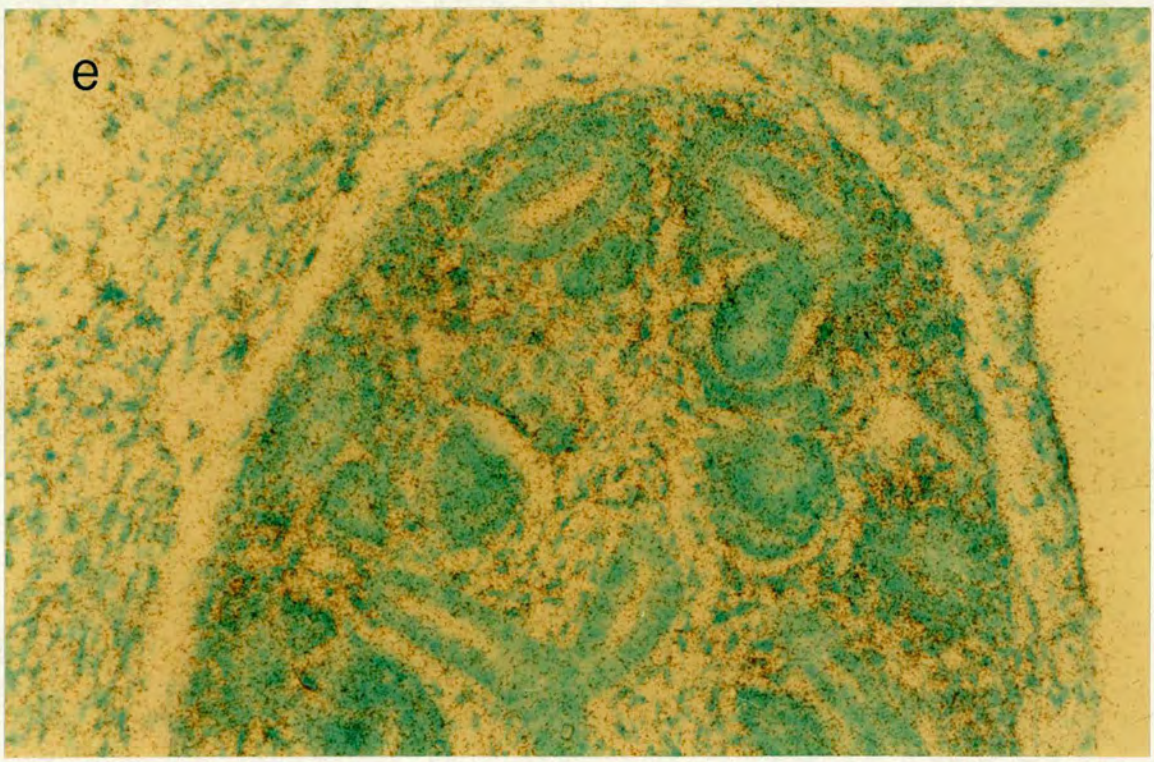


Figure 5.23

5.3 Summary of 14-3-3 ε expression during mouse embryogenesis

Tissue	8.5 d.p.c.	9.5 d.p.c.	10.5 d.p.c.	11.5 d.p.c.	12.5 d.p.c.	13.5 d.p.c.	15.5 d.p.c.
epithelia	2	2	1-1.5	1.5	1	1.5-1	1
mesenchyme	2	3	2-3	3	2-3	2-3	1.5-2
neural tube	2	2	2	3	2.5	2	
mantle					2-3		
marginal		0	0	0	0		
brain, diencephalon		2	2 3	3	2-3	2-3	2
somite			2	2	2		
lung epithelia			2	3	1.5	1.5	1
mesenchyme				3	3	3	1
cartilage					1.5	1	
bone					1.5	1.5	0.5
dorsal root ganglia				3	1.5	2	1.5
kidney: mesenchyme					2-3	2.5-3	2
stroma					-	1.5	1.5
cell duct					1.5	1.5	1.
gut: submucosa			1.5	1.5	1.5	1.5	1
muscle			2.5	3	3	3	1
serosal			1.5	2	1.5	1.5	1
heart			1-1.5	2.5	1.5	1.5	1-1.5
teeth					2	2	2-2.5
Red blood cells		1	1	1	1	0	0
liver			1.5	2	1.5	1.5	1
limb condensations				2		1.5	1
mesenchyme				3		2.5	2

Scale	0	background
	1	low
	2	medium
	3	high

Chapter 6

The 14-3-3 protein family

6.1 Introduction

The previous chapter described how three clones from the subtractive hybridization were isolated, purified and sequenced. Using the UWGCG programme TFASTA to analyse the deduced protein sequence and to search GENBANK for similar sequences showed that the sequence of the clones was very similar to the 14-3-3 protein family. To try and understand more about the cDNA that was isolated and the protein's possible function in nephrogenesis, it is helpful to review what is known about the 14-3-3 protein family. This information may help to draw conclusions about the expression data shown in the previous chapter.

6.2 Isolation of the 14-3-3 proteins

The 14-3-3 family of proteins (reviewed by Aitken *et al.,* 1992) was originally identified on two-dimensional protein gels:- ('protein maps') of brain extracts (Moore & Perez, 1967); the name arises from the systematic analysis and numerical designation of the proteins. Early studies indicated the protein was 1% of cytosolic protein in the bovine brain (Boston *et al.,* 1982a), localised exclusively in neurons (Boston *et al.,* 1982b) and axonally transported to the nerve terminals in the retinal ganglion cells (Erickson & Moore, 1980). Since then it has been shown that the protein is present in many different tissues.

It was discovered in 1987 that the family was involved in regulating kinase-dependent activation of tryptophan and tyrosine hydroxylase (Ichimura *et al.*, 1987), which are the initial and rate-limiting enzymes in the monoamine biosynthesis pathway. The biosynthesis of the neurotransmitters serotonin and the catecholamines is accelerated by increasing the activity of tryptophan or tyrosine hydroxylase, respectively, and these hydroxylases are regulated by various mechanisms including phosphorylation by protein kinases (Cooper *et al.*, 1986). The phosphorylation is coupled with a series of second messengers that include cAMP, diacylglycerol, and Ca^{2+} through three types of protein kinases, cAMP-dependent protein kinase (protein

kinase A), Ca^{2+}/phospholipid-dependent protein kinase (protein kinase C [PKC]) and Ca^{2+}/calmodulin-dependent protein kinase type II (protein kinase II) (Nairn *et al.,* 1985).

Protein kinase II requires an additional protein fraction to activate tyrosine and tryptophan hydroxylase. The protein fraction was isolated from bovine brains by Yamauchi *et al.,* (1981) and was identified as containing a number of 29-32 kDa proteins which were identified as the 14-3-3 proteins by biochemical and immunochemical means (Ichimura *et al.,* 1987). Additional work showed that the protein fraction could separated by HPLC into seven isoforms which were assigned the names α-η (Ichimura *et al.,* 1988), all of which can activate tyrosine and tryptophan hydroxylase (Isobe *et al.*, 1991). The activation of the hydroxylases has been demonstrated to be a two-step process: the hydroxylase is phosphorylated by protein kinase II, and the 14-3-3 protein then activates the hydroxylase (Yamauchi & Fujisawa, 1981), an action that has been proposed to occur by the acidic carboxyl terminal of the 14-3-3 protein binding to the basic regulatory domain of the hydroxylase (Ichimura *et al.,* 1988). It has since been conclusively proved that the 14-3-3 protein will only bind to the already phosphorylated hydroxylase (Furukawa *et al.*, 1993).

At least six of the mammalian 14-3-3 proteins, (the HME1 and τ/θ have not been tested and the ε can not,) can be phosphorylated *in vitro* by PKC but not by PK II or cAMP dependent kinase, suggesting this is a highly specific reaction (Toker *et al.*, 1992). It may be significant that a serine residue in the PKC consensus sequence is replaced by an alanine in the ε isoform or the inability to be phosphorylated might be due to conformational restraints in the tertiary structure of this isoform.

Work at the genetic level started when Ichimura and co-workers isolated and purified the η isoform of the bovine protein and sequenced a peptide fragment. With this information they cloned the cDNA for the bovine η isoform protein, the first

report of cloning of a 14-3-3 protein (Ichimura *et al.*, 1988). In retrospect it was not the first cloning of a 14-3-3 cDNA as a *Drosophila melanogaster* cDNA cloned in 1986 as an alternatively spliced exon of the *Drosophila* epidermal growth factor receptor homologue (DER) cDNA (Schejter *et al.*, 1986) was later shown to have been cloned as an artifactual cloning event (Schejter *et al.*, 1989) and when a *Drosophila* 14-3-3 protein was cloned (Swanson & Ganguly, 1992) the DER cDNA was shown to be the 5' end of the *Drosophila* 14-3-3 (D14-3-3) cDNA (McConnell & Hodges, 1993).

The proteins have been identified in many other contexts. Looking for proteins from sheep brain that inhibited PKC, Aitken and co-workers isolated two proteins, one of which, KCIP, showed three forms of 29, 30 and 32 kDa. Preliminary peptide sequencing of cyanogen bromide and trypsin fragments showed KCIP had limited homology to lipocortins and annexins and extensive homology to the 14-3-3 protein family (Toker *et al.*, 1990). When the protein was analysed by reverse phase HPLC eight isoforms were seen, the same seven isoforms as Isobe had seen but resolving the γ isoform into γ_1 and γ_2 (Toker *et al.*, 1992). The same paper reported the partial peptide sequencing of the sheep β, ε, γ_2, η and ζ isoforms and showed that all isoforms except the ε isoform could be phosphorylated by PKC. The PKC inhibitory activity is retained after the initial purification steps (DEAE-cellulose, phenyl-Sepharose and Mono Q chromatography columns) (Toker *et al.*, 1990) which purified the 14-3-3 protein to apparent homogeneity when visualised on a silver-stained protein gel. However the activity could not be recovered when the isoforms were separated by reverse-phase HPLC (Toker *et al.*, 1992). This could be due to denaturation of the proteins by the HPLC treatment or due to loss of a co-purifying molecule, such as a bound co-factor or lipid.

Another group has reported that the 14-3-3 proteins, again purified by reverse phase HPLC, actually activated PKC (Isobe *et al.*, 1992). The activation was

dependent on the presence of Ca^{2+} and phospholipid which are co-factors of PKC. The effects of Ca^{2+} and phospholipid on the 14-3-3 proteins should not be overlooked since it has been suggested that the ζ isoform is a phospholipase A_2 (Zupan *et al.*, 1992). A phospholipase A_2 was functionally isolated from sheep platelets and found as a 58 kDa complex that resolved to 30 kDa subunits, so showing that its native quaternary structure was dimeric (Loeb & Gross, 1986). These authors showed that it catalyzed the cleavage of the *sn-2* fatty acid of choline and ethanolamine glycerophospholipids through the formation of a stable acyl-enzyme intermediate and that the transesterification of the *sn-2* group was selective for arachidonic acid.

Using antibodies to this protein, the human cDNA was cloned from a λgt11 library. The sequence of the cDNA and partial peptide sequence of the sheep protein identified the proteins as the ζ isoforms of the 14-3-3 protein by comparison to the partial peptide sequence of the sheep ζ isoform sequence that had previously reported (Toker et al., 1992), although it should be said other groups have since been unable to replicate this activity (Morgan et al., 1993a, Professor R. D. Burgoyne, pers. comm.; Dr D. B. Collinge, pers. comm.) The bovine ζ isoform has also been cloned by a group purifying FAS, a eukaryotic factor required to activate Exoenzyme S, the virulence factor of Pseudomonas aeruginosa. Purification of this protein and micro-peptide sequencing identified the factor as the ζ 14-3-3 isoform (Fu et al., 1993). The protein activates ExoS allowing it to act as a ADP-ribotransferase. The ability of the ζ isoform to activate ExoS is unlikely to be reliant on the phospholipase activity as no phospholipids were required for the assay, and the protein activates ExoS in a stoichiometric manner rather than via an enzymatic pathway. This activity has also been unable to be reproduced by other groups in the 14-3-3 field (Professor R. D. Burgoyne, pers. comm.).

6.3 More functions of the 14-3-3 protein family

By identification of DNA and predicted protein homology, numerous other cDNAs have shown to be members of the 14-3-3 family. Members of the family have been now been isolated from yeast (2), *Arabidopsis* (5), rice (1), barley (3), maize (2), spinach (1), evening primrose (1), *C. elegans* (4), *Xenopus* (1), cattle (4), sheep (5), rat (6), man (5) and as this thesis reports ε is the first isoform isolated in mouse. Members of the family have in fact been found in every eukaryotic species looked at so far including chicken, turtle, frog and goldfish (Ichimura *et al.*, 1991).

The proteins have been identified by a wide range of means that suggests that the protein family may have diverse roles in organisms. The *Xenopus laevis* homologue was cloned from a pituitary cDNA library in a search for proteins co-ordinately expressed with the pro-hormone proopiomelanocortin. Although expression of the gene is higher in the brain than the liver, it is not known if the 14-3-3 gene expression is related to hormone expression (Marten *et al.*, 1992). In *Arabidopsis thaliana* (Thale cress) the ω (omega) isoform of the 14-3-3 protein was originally identified as part of a DNA-binding complex, closer examination showed that although the protein is found in the complex the 14-3-3 protein does not itself bind the DNA (Lu *et al.*, 1992). Four other *Arabidopis* 14-3-3 isoforms, χ, (chi) ϕ (psi), υ (upsilon) and ψ (phi), were recently cloned from an *Arabidopsis* cell suspension culture expression library using a monoclonal antibody to the ω isoform (Professor R. J. Ferl, pers. comm.) and several *Arabidopis* expressed sequence tags have homology to the protein family. The maize (*Zea mays*) homologue has also been found as part of a DNA-binding complex using this method (de Vetten *et al.*, 1992).

The yeast (*S. cerevisiae*) homologue, *BMH1* (brain modulosignalin homologue 1), was cloned coincidentally with a flanking gene (van Heusden *et al.*, 1992). The *BMH1* gene is not essential as disruption mutants are viable. They grow slowly when glucose is the carbon source but normally on non-fermentable sugars. Overexpression

of the gene causes the yeast to grow slowly on glucose, but very poorly on non-fermentable sugars. Both results show that *BMH1* is involved in growth regulation. A second gene *BMH2* has recently been cloned and entered into the GENBANK database.

The gene is also common in the plant world: the 14-3-3 gene cDNAs from spinach (*Spinacerea oleracea*) and evening primrose (*Oenothera hookeri)* were cloned by homology (Hirsch *et al.*, 1992), while the rice (*Oryza sativa*) 14-3-3 gene was found in a random cDNA sequencing project (Uchimiya *et al.*, 1992). Rice 14-3-3 gene expression has been shown to increase when the plant is subjected to external stresses such as low temperatures and saline (Kidou *et al.*, 1993). In barley (*Hordeum vulgare*), the gene was found to be a pathogen-induced when the plant is stressed in response to infection attempts by the powdery mildew fungus, *Erysiphe graminis* (Brandt *et al.,* 1992). When barley is infected by *Erysiphe graminis* the transcription of defence-related genes such as the 14-3-3 gene is up-regulated and this is correlated to the development of papillae. Cytoplasmic activity increases adjacent to the nearby papillae and this is thought to relate to enhanced exocytosis. Another two barley cDNAs have been cloned, and experiments have shown that the barley cDNAs can functionally complement the yeast *BMH1* mutant disrupted in its own 14-3-3 homologue gene (Dr D. B. Collinge, pers. comm.).

6.3.1 Exo1, a 14-3-3 protein

When human adrenal chromaffin cells are permeabilised and the permeable protein components from the cells allowed to leak out, the loss of proteins is accompanied by run-down of the calcium-dependent exocytotic response (Ali *et al.*, 1989). Exo1 and Exo2 proteins were found as cytosolic protein fractions that reactivated calcium-dependent exocytosis potentiated by protein kinase C activation (Morgan & Burgoyne 1992). When the Exo1 protein was purified and peptide sequenced, the peptide sequence showed homology to the 14-3-3 proteins the

protein's size, 30 kDa, also suggested it may be a member of the 14-3-3 family. Exo1 has now been purified from sheep brain (Morgan *et al.*, 1993) and shown to contain all the brain 14-3-3 isoforms, while the original adrenal chromaffin cell Exo1 extract has been found to contain multiple 14-3-3 isoforms (β, γ, ζ and low levels of η and ε). (Burgoyne pers. comm.). Recently Exo2 has been purified and shown to be the catalytic subunit of protein kinase A (Morgan *et al.*, 1993b). Therefore, if the action of Exo1 is similar to the action of Exo2, the 14-3-3 proteins may be acting on exocytosis by a mechanism similar to their kinase II dependent activation of tyrosine and tryptophan hydroxylase.

Four other human 14-3-3 cDNAs have been found. The first cDNA for a human 14-3-3 isoform was found fortuitously in a T-cell cDNA library when screening for a previously uncharacterized nuclear protein (Nielsen, 1991), the rat homologue has since been cloned and designated the θ isoform (Watanabe *et al.*, 1993b). Another isoform was found as a human epithelial cell marker protein 1 (HME1) that is down regulated as cell lines become transformed, acting as a cellular differentiation marker (Prasad *et al.*, 1992), and the human η isoform was cloned by virtue of homology to the bovine η (Ichimura-Ohshima *et al.*, 1992). Four 'new' human cDNAs were cloned as transformation-sensitive epithelial marker proteins, one is the θ isoform cloned by Nielsen, one is the HME1 protein, one is the η isoform reported by Ichimura-Ohshima *et al.*, (1992), the other cDNA is the human β isoform (Leffers *et al.*, 1993). There have also been reports of novel isoforms expressed in the spleen, one isoform specific to skin, ear and tongue while two isoforms have been identified as Golgi proteins (Celis *et al.*, 1990a; Celis *et al.*, 1990b; Aitken *et al.*, 1992).

6.4 Tissue localisation of the 14-3-3 proteins

Isobe *et al.* (1991) have cloned the β, γ and η isoforms in cattle and found that all three isoforms are highly expressed in the brain, but that the isoforms seem to have different tissue specificities. The γ isoform seemed to be brain-specific although it

has been claimed that the human γ isoform is a Golgi-specific protein found in keratinocytes, while β and η isoform expression was detected in other tissues. The β, γ and η isoform have now been isolated in rat (Watanabe *et al.*, 1991, Watanabe *et al.*, 1993a) and the expression in the embryonic and early post-natal rat brain of all three isoforms examined closely by *in situ* hybridization. The resulting data showed gene expression in the catecholamine and indolamine neurons, as previous biochemical characterization had suggested, but at least some level of expression was seen in all neurons, and neurons with large somata (*i.e.* Purkinje cells in the cerebellum) showed high levels of expression (Watanabe *et al.*, 1991, Watanabe *et al.*, 1993a, c). The rat ε (Roseboom *et al.*, 1993), θ and ζ (Watanabe *et al.*, 1993b) have also recently been cloned and entered into the GENBANK database. The rat ε isoform was found as a major contaminant of a highly purified preparation of rat pineal N-acetyltransferase. Western blot analysis of this protein found it was widely distributed in the cortex, cerebellum, hypothalamus, pineal gland, retina, testes, heart, lung and kidney in varying amounts (Roseboom *et al.*, 1992). and the η isoform in human (Ichimura-Ohshima *et al.*, 1992). The human isoforms (HME1, τ, and η) have only been examined in cell lines and they seem to be expresses in most cell lines examined although the level varies (Nielsen, 1991; Prasad *et al.*, 1992, Ichimura-Ohshima *et al.*, 1992).

The *Drosophila* homologue of the 14-3-3 protein, D14-3-3, was found in a screen for head-specific cDNAs. The cDNA detects three transcripts, 2.9 kb, 1.9 kb and 1.0 kb: the 2.9 kb transcript is only expressed in the head of the *Drosophila*, while the other two transcripts are found in the both head and the body. The northern data suggests that the D14-3-3 cDNA detects three isoform transcripts and is not specific enough to allow extrapolation on regions of expression from *in situ* hybridization data. The D14-3-3 gene is expressed at the highest level in the neural tissues of the head corresponding very well to the expression patterns of the η isoform that have

been seen in the rat, whereas at the sequence level the cDNA shows highest homology to the ζ isoforms in bovine, rat and human. The ζ isoform can also be seen to be very homologous to the β and θ isoforms and the cDNA has probably hybridized to the β and θ transcripts due to the high sequence conservation between the three isoforms.

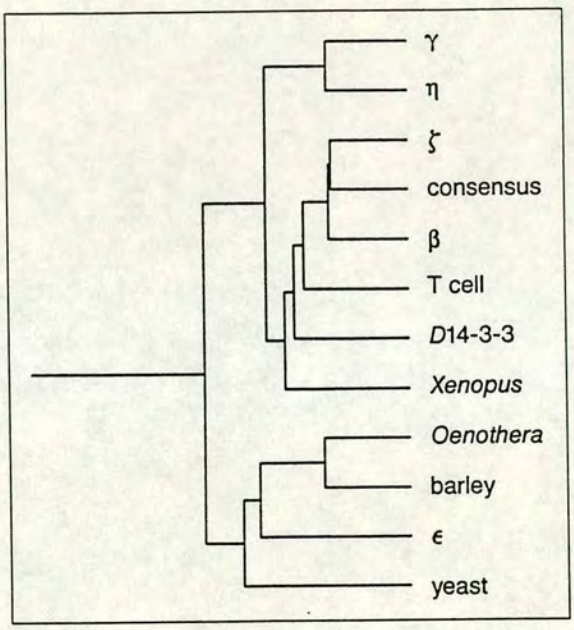
Only the ε isoform of the 14-3-3 protein has had a detailed *in situ* hybridization study done to study the pattern of expression of the protein during development. As discussed in the previous chapter, expression is high generally in the early developing mouse embryo (8.5 and 9.5 d.p.c.) with the expression becoming restricted to undifferentiated mesenchyme and the developing nervous system. Finally the expression drops in most tissues staying at a higher level in the nervous system. Other studies have used western blots to study the protein's distribution in the organism of interest but this does not tell you where in a tissue the gene is expressed. The only comparative study of the expression of several rat isoforms in fetal and early post-natal brain is similar to the results that were found for the mouse ε isoform.

6.5 The 14-3-3 protein family evolution, similarity and identity.

The level of similarity of the different members of this protein family is remarkably high when all sequences of a significant length are compared (≥100 nucleotides) (Figure 6.1, parts 1-4). Immunologically reactive proteins can be found in all the species looked at so far (Ichimura *et al.*, 1991) and sequence has been obtained from a wide range of organisms from yeast, plants, invertebrates and mammals. Currently the sequence of seven isoforms is known in mammals and invertebrates. The β, ε, γ, η, θ and ζ isoforms have been found in several species and the conservation of amino acid sequence is very high from mammals to invertebrates. The same isoforms in different mammals are 97-100% identical, while the *Drosophila* 14-3-3 protein is 81.5% identical to the ζ isoform and the *Xenopus*

protein is 86-87% identical to the mammalian β isoform. The level of similarity between isoforms in the same species is uniformly high, β and ζ are 85-87% identical, γ and η isoforms are 86-87% identical and θ and β are 81% identical. The HME1 isoform is 70.2% identical to the θ and ζ isoforms. The ε isoform shows a much lower homology to the mammalian isoforms, 63-67% identical, but it has higher identity to the plant isoforms, 75.5% identical to the *Arabidopsis* ψ isoform (Figure 6.1 and 6.4), when the amino acid sequences are compared using the UWGCG programme PILEUP, the ε isoform shows greater likeness to the isoforms expressed in yeast and plant than to the other mammalian isoforms (Figure 6.5).

Figure 6.5



Sequence comparison generated by the UWGCG programme PILEUP from Aitken *et al.*, 1992.

The mammalian isoforms can be aligned, and three blocks of high similarity and regions of more interspersed lower conservation can be seen (Figure 6.2). The hydrophilicities, calculated by the Kyte-Doolittle method (Kyte & Doolittle, 1982), of each of the seven isoforms can also be compared (Figure 6.3). The resulting graphs allow evaluation of the different isoforms protein structure and as the graphs are very

similar suggest that the proteins will also be very similar reflecting the close relationship between the proteins.

In plants only *Arabidopsis*, barley and yeast have had more than one isoform isolated currently, but the number of 14-3-3 clones reported in the GENBANK and EMBL databases is currently rising exponentially, and five distinct cDNAs have been found in *Arabidopsis* that do not seem to correspond to the range of the mammalian isoforms. When the predicted plant protein sequences are compared to the mammalian isoforms, they are all most homologous to the mammalian ε isoform. The yeast sequences are 72-74% identical and the plant isoforms are 75.5-70% identical to the ε isoform. The plant sequences appear to have diverged early in evolutionary history perhaps when only one 14-3-3 isoform (similar to the ε isoform) existed and followed a separate evolutionary pathway to the mammalian isoforms. All the plant proteins show high homology to each other (at least 80% identical) (Figure 6.1 and 6.4). Looking at the hydrophilicities of the plants, it can also be seen that the protein structure has been very highly conserved (Figure6.5).

Comparisons of the mammalian isoforms (Figure 6.2) and the plant isoforms including the yeast sequence (Figure 6.4) allow the regions of highest identity to be identified. These regions of the protein are probably under evolutionary restraint that maintains the sequences to maintain the function of the protein.

Among the isoforms, three regions of high conservation can be seen, blocks 1, 2 and 3. The constant domains are interspersed with more variable regions, block A (Figures 6.2 and 6.4). When the level of conservation in all the isoforms is considered (Figures 6.1-6.5) it becomes obvious the proteins are conserved at a level that is normally seen for essential genes such as the actin gene family. This suggests that the role this protein family plays is important and under severe evolutionary constraints.

The 14-3-3 family of proteins have homology to the annexins and particularly to annexin II (calpactin), a protein already implicated in calcium-dependent exocytosis (reviewed by Creutz, 1992). Annexins are a group of homologous proteins that bind phospholipids in the presence of calcium. They may provide a major pathway for communication between cellular membranes and their cytoplasmic environment (Creutz, 1992). A key domain, amino acid sequence KGDYQKALLYLCGGDD in block 2, was identified both in annexin II and the 14-3-3 proteins (Aitken *et al.*, 1990) as responsible for stimulating calcium-dependent exocytosis. This sequence is found as a conserved internal domain (residues 127-142 of the η form [the consensus sequence in the 14-3-3- mammalian isoform is KGDYYRYLAEVA-GD]) in the 14-3-3 family that has homology to the C-terminus of the annexins. The possession of this domain by both classes of protein suggests that is involved in protein-protein interactions in exocytosis. A synthetic peptide corresponding to this domain, KGDYQKALLYLCGGDD, inhibited PKC binding to Receptors for Activated Protein C Kinase (RACKs) in permeabilised adrenal chromaffin cells and inhibited Ca^{2+}-dependent secretion specifically (Roth *et al.*, 1993).

6.6 Conclusions

While the exact biochemical functions of these proteins are unclear, all the activities described predict a role in signal transduction. The proteins have been shown to bind the phosphorylated forms of tyrosine and tryptophan hydroxylase activating their enzyme activity. The proteins are present in plant DNA binding complexes, they can compete with activated PKC for binding to receptors, they can directly activate or inactivate PKC, they can cleave phospholipids and, in some manner, regulate exocytosis. They are also thought to be involved in pathogen resistance and response to stress in plants and growth regulation in yeast. This diverse range of roles though extensive, has not fully explained the role and action of this family of proteins and much work is required to elucidate their role. This review

of the multiple biological roles so far attributed to this family of proteins allows us to speculate on the role that the ε isoform may play in the developing kidney and the developing embryo. The possible role is speculated on and further discussed in the next chapter, the summary and discussion.

Figure 6.1

Comparison of 14-3-3 isoforms using the UWGCG programme BESTFIT to compare the deduced amino acid sequences. The top figure is the percentage similarity and the lower figure is the percentage identity. The figures are rounded to the nearest one decimal place.

Abbreviations used: Arab, *Arabidopsis thaliania*; bov, bovine; C. eleg/ c.e. *Caenorhabditis elegans*; Dros, *Drosophila melanogaster*; hum, human; pla phospholipase A_2; hvulg, barley (*Hordeum vulgare)*; mouse, *Mus musculus*; oe, evening primrose (*Oenothera hookeri);* sp, spinach (*Spinacerea oleracea*); rice, (*Oryza sativa*); Xen, *Xenopus laevis*; Yeast, *Saccharomyces cerevisiae*; *bmh,* brain modulosignalin homologue; zea/zea2, maize (*Zea mays).*

14-3-3 Protein homology table part 1

Species														
arab-ω GF14	arabω GF-14													
arabυ	88.6 81.1	arab υ												
arabψ	88.2 88.3	**91.0** **86.2**	arab ψ											
arabϕ	**91.1** **88.0**	84.4 77.7	85.8 80.3	arab ϕ										
arabχ	**94.2** **90.7**	87.5 79.2	87.4 80.0	**91.1** **88.0**	arab χ									
bovβ	74.8 64.0	72.8 60.6	74.0 63.2	70.7 57.8	74.4 62.8	bovβ								
bovγ	73.6 62.0	71.5 60.3	72.3 62.4	71.5 59.1	72.3 59.5	85.0 76.2	bov γ							
bovη	73.3 61.3	70.4 58.8	72.1 61.9	71.6 58.4	72.8 59.3	84.9 75.7	**91.4** **86.5**	bov η						
bovζ	79.1 64.3	76.6 61.5	79.4 65.4	75.8 61.1	78.6 63.0	**92.2** **86.8**	85.5 75.5	85.0 75.0	bov ζ					
C. eleg 1	85.2 75.0	81.8 71.6	85.2 77.3	85.2 77.3	84.1 73.9	85.2 76.1	86.4 77.3	87.5 80.7	84.1 77.1	C.e. 1				
C. eleg 2	69.9 62.7	71.1 60.2	73.5 62.7	65.1 56.6	69.9 60.2	88.8 86.2	85.4 79.3	82.9 76.8	**93.8** **88.9**	n/a	C.e. 2			
C. eleg 3	77.4 56.5	75.4 60.7	80.6 58.1	69.4 48.4	79.0 54.8	80.6 72.6	79.0 67.7	74.2 67.7	87.1 80.6	n/a	**95.5** **100**	C.e. 3		
C. eleg 4	63.6 54.5	60.2 53.4	63.6 54.5	62.5 52.3	64.8 53.8	71.6 62.5	77.3 68.2	73.9 67.0	77.3 65.9	85.7 80.0	n/a	85.3 85.3	C.e. 4	
Dros ζ	75.2 63.8	73.3 60.0	76.7 64.9	73.3 61.5	74.7 62.4	87.7 78.3	85.9 74.7	84.2 73.0	**91.0** **81.2**	89.8 84.1	**91.4** **77.4**	**91.9** **77.4**	79.9 68.2	Dros ζ

Table of 14-3-3 homologies part2

Species	arab ω	arab υ	arab ψ	arab ϕ	arab χ	bov β	bov γ	bov η	bov ζ	ce 1	ce 2	ce 3	ce4	Dros ζ						
hum	74.9	71.7	72.5	72.4	75.2	81.0	79.3	77.6	82.0	78.4	77.1	75.8	73.6	78.3	hum					
HME	59.1	56.7	57.9	55.7	58.1	67.8	65.7	64.7	70.2	68.2	77.1	67.7	63.2	66.4	hme					
humβ	74.8	73.9	74.9	72.0	74.8	**100**	85.1	85.0	**92.2**	85.2	88.9	82.3	71.6	88.6	80.2	hum				
	63.0	61.2	62.1	59.3	62.2	**99.6**	76.3	75.8	**86.9**	76.1	86.4	72.6	62.5	78.8	67.9	β				
humθ	74.2	71.7	73.8	73.8	73.3	88.9	82.6	83.8	89.0	86.4	88.9	82.3	73.9	86.1	80.0	88.9	hum			
	61.9	60.2	61.9	62.3	61.3	80.7	70.5	71.7	79.6	80.7	82.7	69.4	63.6	77.1	70.2	81.1	θ			
hum η	73.7	70.8	72.5	72.0	73.3	85.4	**91.8**	**99.6**	85.8	87.5	82.9	74.2	73.9	84.2	78.0	85.4	84.2	hum		
	61.3	58.8	61.9	58.4	59.3	75.7	**86.5**	**99.2**	75.4	80.7	76.8	67.7	67.0	73.0	64.7	75.8	71.7	η		
hum	79.1	76.6	79.4	75.8	78.6	**92.2**	85.5	85.0	**100**	84.1	**93.8**	87.1	77.3	**91.0**	82.0	**92.2**	89.0	85.8	hum	
pla/ζ	64.3	61.5	65.4	61.1	63.0	**86.8**	75.5	75.0	**100**	77.3	**88.9**	80.6	65.9	**81.2**	70.2	**86.9**	79.6	75.4	ζ	
hvulg	88.0	88.0	**91.7**	83.5	87.2	73.9	71.2	70.8	76.6	83.0	72.3	74.2	62.5	74.2	72.4	73.9	73.4	71.2	76.6	h
	81.0	81.0	**86.2**	76.5	80.2	64.1	59.7	59.3	61.9	73.9	62.7	61.3	56.8	62.1	59.8	63.3	61.5	59.3	61.9	vulg
mouse	84.3	81.7	84.2	81.5	82.7	78.1	74.8	74.6	80.0	86.4	80.7	74.2	62.5	78.9	74.5	77.4	77.5	75.0	80.0	81.9
ε	75.2	69.8	75.5	72.0	73.6	67.4	63.6	62.7	67.6	77.3	67.5	61.3	55.7	67.8	59.1	66.7	63.5	62.7	67.6	73.2
οε	**94.9**	87.1	89.7	89.6	**92.5**	72.2	72.7	71.6	77.3	85.2	71.1	74.2	62.5	74.4	74.4	74.4	72.8	72.0	77.3	89.6
	89.8	80.1	83.0	85.0	**87.1**	62.0	62.8	59.7	62.8	72.7	59.0	54.8	54.6	61.4	59.8	62.4	61.3	59.7	62.8	81.9
ratβ	74.4	73.1	74.1	72.0	74.4	**99.2**	85.1	84.6	**91.4**	85.2	88.9	80.6	72.7	88.6	79.4	**99.2**	88.9	85.0	**91.4**	73.1
	62.6	61.2	62.1	58.8	61.8	**97.9**	75.9	75.0	**86.1**	76.1	86.4	72.6	63.6	78.0	68.3	**98.4**	80.7	75.0	**86.1**	63.3
ratε	84.3	81.7	84.2	81.5	82.7	78.1	74.8	74.6	80.0	86.4	80.7	74.2	62.5	78.9	74.5	77.4	77.5	75.0	80.0	81.9
	75.2	69.8	75.5	72.0	73.6	67.4	63.6	62.7	67.6	77.3	67.5	61.3	55.7	67.8	59.1	66.7	63.5	62.7	67.6	73.2
ratγ	73.7	71.6	72.2	71.6	72.4	84.6	**99.6**	**91.9**	85.1	86.4	85.4	79.0	77.3	85.5	78.7	84.7	82.2	**92.3**	85.1	71.3
	61.7	60.1	62.0	58.8	59.3	75.9	**99.6**	**87.0**	75.2	77.3	79.3	67.7	68.2	74.4	65.2	76.0	69.8	**87.0**	75.2	59.4
ratη	73.3	70.4	72.1	71.6	72.8	84.9	**91.4**	**100**	85.0	87.5	82.9	74.2	73.9	84.2	77.6	85.0	83.8	**99.6**	85.0	70.8
	61.3	58.8	61.9	58.4	59.3	75.7	**86.5**	**100**	75.0	80.7	76.8	67.7	67.0	73.0	64.7	75.8	71.7	**99.2**	75.0	59.3
ratθ	74.2	71.7	73.8	73.8	73.3	88.9	82.6	83.8	89.0	86.4	88.9	82.3	73.9	86.1	80.0	88.9	**100**	84.2	89.0	73.4
	62.3	60.7	62.3	62.7	61.7	80.7	71.0	72.1	79.2	80.7	82.7	69.4	64.8	76.7	69.8	81.1	**99.6**	72.1	79.2	61.9
ratζ	78.7	76.6	79.4	75.4	78.2	**92.2**	85.5	85.0	**100**	84.1	85.4	87.1	77.3	**91.0**	82.0	**92.2**	89.0	85.8	**100**	76.6
	64.3	61.1	65.0	61.1	63.6	**87.2**	75.5	75.0	**99.6**	77.3	79.3	79.0	64.8	**81.6**	69.8	**87.3**	80.0	75.4	**99.6**	61.5

14-3-3 homology table part 3

Species	arab ω	arab υ	arab ψ	arab ψ	arab χ	bov β	bov γ	bov η	bov ζ	ce 1	ce 2	ce 3	ce 4	Dros ζ	hum hme	hum β	hum θ	hum η	hum ζ	h vulg
rice	88.7	88.7	**91.0**	85.0	88.3	74.3	72.3	72.7	77.0	84.1	72.3	72.6	60.2	74.2	71.7	74.3	73.8	73.1	77.0	**97.3**
	81.2	80.5	**84.6**	77.0	80.9	64.1	61.6	60.3	62.7	73.9	62.7	59.7	53.4	63.7	57.5	63.3	60.2	60.3	62.7	**96.2**
sheep ε	84.8	82.7	85.2	82.2	83.1	76.8	73.8	73.3	79.3	86.4	78.7	74.2	62.5	78.7	74.3	76.0	77.1	73.8	79.3	82.3
	76.1	70.6	76.5	73.0	74.5	66.1	62.7	62.2	67.4	77.3	64.0	61.3	55.7	67.6	58.7	65.3	63.9	62.2	67.4	74.0
sheep η	88.9	86.7	86.7	86.7	86.7	**95.6**	**100**	**100**	**93.3**	86.0	n/a	n/a	n/a	**93.3**	88.6	**95.6**	**91.1**	**100**	**93.3**	86.7
	80.0	73.3	80.0	77.8	75.6	**77.8**	**93.3**	**100**	**80.0**	81.4				**84.4**	75.0	**77.8**	**80.0**	**100**	**80.0**	73.3
Spinach	**90.2**	87.7	89.6	88.4	**91.2**	74.9	76.1	74.1	79.0	84.1	77.4	77.4	62.5	77.6	74.0	75.5	73.7	74.6	79.0	86.0
	85.1	79.2	82.5	83.3	**85.3**	64.5	66.2	66.2	63.9	76.1	64.2	53.2	53.4	66.8	59.1	63.2	63.9	66.2	63.9	79.4
Xenβ	75.2	75.2	77.0	71.8	76.1	**92.7**	83.6	83.6	**90.2**	84.1	**92.6**	85.5	72.7	86.8	78.6	**92.7**	86.3	84.1	**90.2**	75.6
	65.0	62.4	63.2	61.5	65.4	**87.1**	73.3	74.6	**84.2**	76.1	**90.1**	74.2	63.6	77.4	67.9	**87.2**	77.8	74.6	**84.2**	63.7
Yeast bmh1	79,8	80.6	81.0	76.7	80.0	75.5	73.3	72.8	76.6	86.4	73.5	67.7	64.4	75.7	71.1	74.7	73.6	73.7	76.6	79.3
	69.6	68.7	72.7	67.7	69.5	63.9	62.1	60.9	64.8	77.3	62.7	54.8	56.3	64.0	58.1	63.7	61.2	61.3	64.8	70.3
Yeast2 bmh2	81.3	80.8	82.9	77.4	79.5	76.4	74.4	73.3	76.2	85.2	74.7	69.4	62.5	74.4	72.8	76.0	73.4	74.1	76.2	80.4
	70.4	68.8	72.5	66.9	69.1	64.5	63.2	60.9	65.2	76.1	62.7	54.8	54.6	63.4	58.9	64.2	60.2	61.3	65.2	69.8
zea	**90.3**	**90.6**	**91.1**	87.5	88.3	76.5	75.1	74.2	77.6	85.2	78.0	75.4	62.5	75.5	75.8	76.2	75.1	74.7	77.6	**90.7**
	81.0	**81.7**	**82.9**	79.4	80.2	64.5	62.7	62.2	65.4	73.9	63.4	57.4	54.6	63.3	60.8	64.3	62.9	62.2	65.4	**83.1**
zea2	**91.8**	**90.2**	**90.2**	**90.2**	88.5	84.7	80.3	80.3	88.1	79.7	n/a	n/a	n/a	88.1	75.4	84.7	77.0	80.3	88.1	**93.4**
	77.3	**82.0**	**83.6**	**83.6**	83.6	76.3	68.9	67.2	76.3	67.8				76.3	60.7	76.3	68.9	67.2	76.3	**91.8**

Table of 14-3-3 homologies part4

Species																	
mouse	mouse ε																
oe	83.3 74.6	oe															
ratβ	77.8 66.3	74.5 61.7	rat β														
rat ε	**100** **100**	83.3 74.6	77.8 66.3	rat ε													
rat γ	74.7 63.3	72.4 62.1	84.7 75.6	74.7 63.3	rat γ												
rat η	74.6 62.7	71.6 59.7	84.6 75.0	74.6 62.7	**91.9** **87.0**	rat η											
rat θ	77.5 63.9	72.8 61.7	88.9 80.7	77.5 63.9	82.2 70.2	83.8 72.1	rat θ										
rat ζ	79.5 67.6	76.5 62.6	**91.4** **85.7**	79.5 67.6	85.1 74.8	85.0 75.0	89.0 79.6	rat ζ									
rice	83.0 73.9	87.7 80.4	73.5 63.3	83.0 73.9	72.4 61.3	72.7 60.3	73.8 60.7	77.0 62.3	rice								
sheep ε	**100** **100**	83.6 75.8	76.4 64.9	**100** **100**	73.9 62.4	73.3 62.2	77.1 64.3	78.9 67.4	83.5 75.7	sheepε							
sheep η	88.9 75.6	**93.3** **77.8**	**95.6** **77.8**	88.9 75.6	**100** **100**	**100** **100**	**91.1** **80.0**	**93.3** **80.0**	88.9 75.6	88.9 75.6	sheep η						
spinach	83.2 72.9	**93.4** **85.8**	75.0 62.7	83.2 72.9	75.5 65.2	74.1 66.2	73.7 64.4	78.5 63.9	86.8 78.8	83.2 73.6	**92.9** **88.1**	sp					
Xen β	78.6 67.1	75.6 64.1	**92.3** **86.3**	78.6 67.1	83.3 73.0	83.6 74.6	86.3 77.8	**90.2** **84.6**	76.1 62.8	76.4 65.3	**93.3** **77.8**	76.8 66.0	Xen β				
Yeast bmh1	81.9 72.4	79.8 70.4	74.3 63.3	81.9 72.4	73.0 61.9	72.8 60.9	73.6 61.2	76.2 64.8	79.5 70.9	83.0 73.9	86.7 75.6	78.9 66.5	76.5 63.7	Yeast bmh1			
Yeast2 bmh2	83.5 73.6	81.4 70.8	75.6 63.8	83.5 73.6	74.1 61.9	73.3 60.9	73.4 60.2	75.8 65.2	80.7 70.5	83.9 74.8	84.4 73.3	79.0 66.7	76.9 65.0	**92.7** **90.5**	Yeast bmh2		
zea	82.9 73.5	**90.5** **82.3**	75.3 63.4	82.9 73.5	74.6 61.9	74.2 62.2	75.1 63.3	77.6 65.8	**92.3** **83.4**	83.5 73.5	88.4 74.4	86.4 76.2	76.9 64.5	81.3 70.7	81.7 71.1	zea	
zea2	86.9 77.0	**90.2** **85.2**	84.7 76.3	86.9 77.0	80.3 68.9	80.3 67.2	77.0 68.9	88.1 76.3	**91.8** **90.2**	86.9 77.0	82.2 66.7	88.5 80.3	85.2 70.5	80.3 73.8	83.6 73.8	**90.2** **83.6**	zea2

Figure 6.2

Alignment of the seven mammalian isoforms' deduced amino acid sequences using the UWGCG programme LINEUP. Three constant regions are shown, block 1 (shown in yellow), block 2 (shown in orange), and block 3 (shown in pink), with regions of lower homology shown (coloured green). hhme is the human HME1 protein.

Comparison of the mammalian isoforms

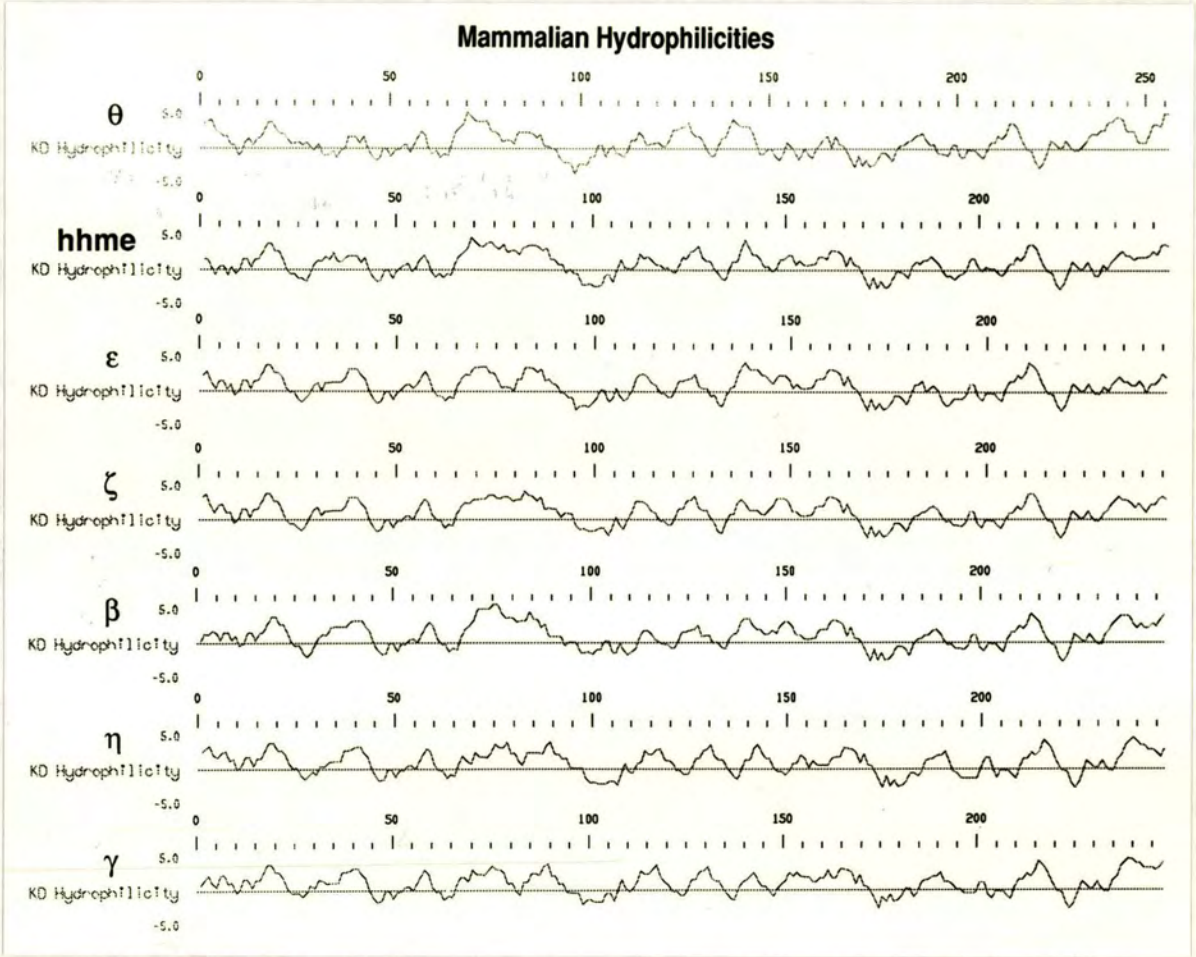


Figure 6.3

Alignment of all the mammalian proteins' deduced hydrophilicities calculated using the UWGCG programme PLOTSTRUCTURE and PEPTIDESTRUCTURE. hhme is the human HME1 protein.

Figure 6.4

Alignment of all the plant isoforms' deduced amino acid sequences plus the yeast and the mouse epsilon sequences, aligned using the UWGCG programme LINEUP. Three constant regions are shown, block 1 (shown in yellow), block 2 (shown in orange), and block 3 (shown in pink), with regions of lower homology shown (shown in green).

Abbreviations used are:,

mouse-mouse 14-3-3 epsilon isoform

bmh1/bmh2 Yeast-(*S. cerevisiae*) brain modulosignalin homologue

arab- thale cress (*Arabidopsis thaliania)*

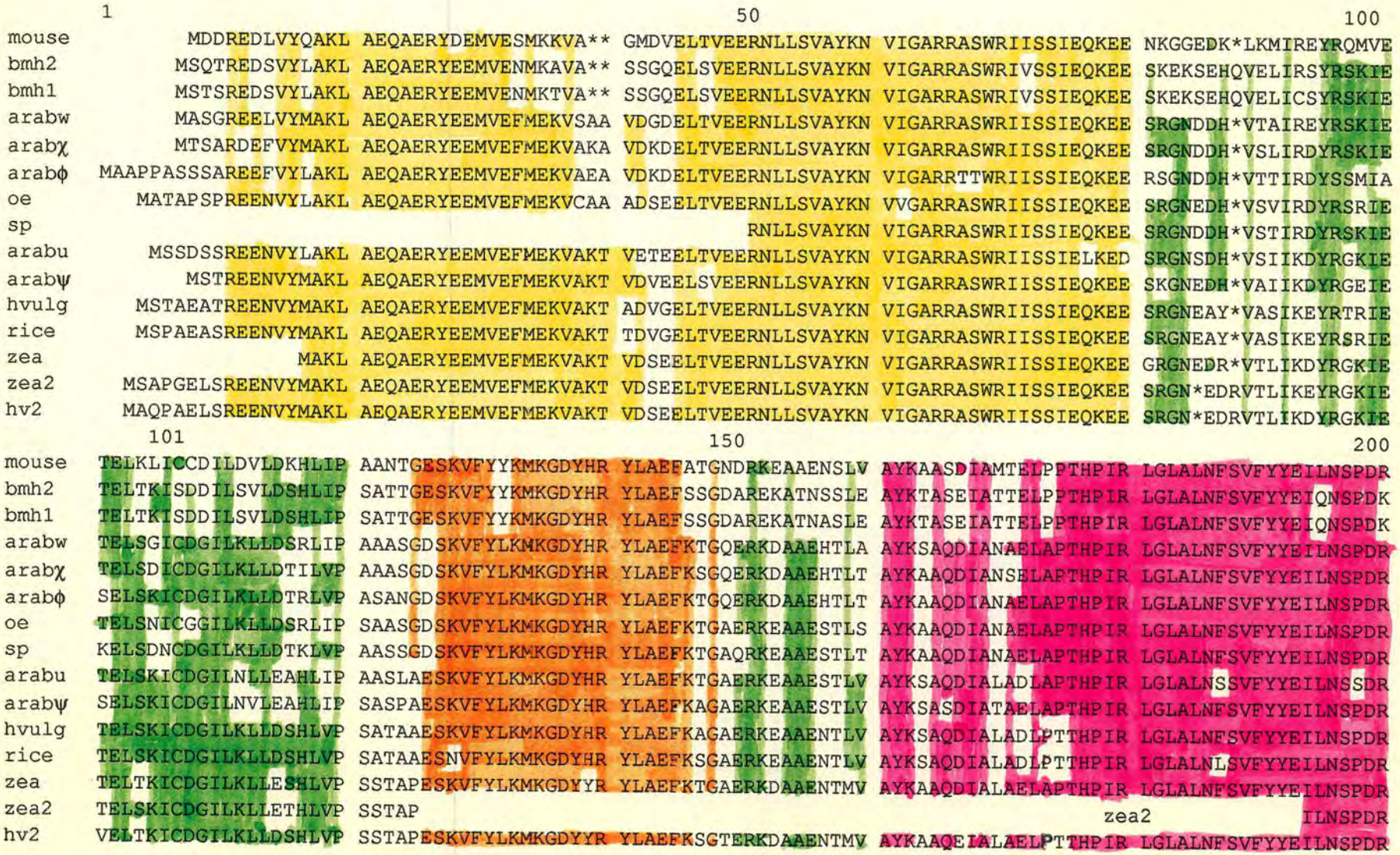
oe- evening primrose (*Oenothera hookeri)*

sp-spinach (*Spinacerea oleracea*)

rice- (*Oryza sativa*)

zea/zea2-maize (*Zea mays)*

hvulg/hv2-barley (*Hordeum vulgare)*


1
50
100
mouse MDDREDLVYQAKL AEQAERYDEMVESMKKVA** GMDVELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE NKGGEDK*LKMIREYRQMVE
bmh2 MSQTREDSVYLAKL AEQAERYEEMVENMKAVA** SSGQELSVEERNLLSVAYKN VIGARRASWRIVSSIEQKEE SKEKSEHQVELIRSYRSKIE
bmh1 MSTSREDSVYLAKL AEQAERYEEMVENMKTVA** SSGQELSVEERNLLSVAYKN VIGARRASWRIVSSIEQKEE SKEKSEHQVELICSYRSKIE
arabw MASGREELVYMAKL AEQAERYEEMVEFMEKVSAA VDGDELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGNDDH*VTAIREYRSKIE
arabχ MTSARDEFVYMAKL AEQAERYEEMVEFMEKVAKA VDKDELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGNDDH*VSLIRDYRSKIE
arabϕ MAAPPASSSAREEFVYLAKL AEQAERYEEMVEFMEKVAEA VDKDELTVEERNLLSVAYKN VIGARRTTWRIISSIEQKEE RSGNDDH*VTTIRDYSSMIA
oe MATAPSPREENVYLAKL AEQAERYEEMVEFMEKVCAA ADSEELTVEERNLLSVAYKN VVGARRASWRIISSIEQKEE SRGNEDH*VSVIRDYRSRIE
sp RNLLSVAYKN VIGARRASWRIISSIEQKEE SRGNDDH*VSTIRDYRSKIE
arabu MSSDSSREENVYLAKL AEQAERYEEMVEFMEKVAKT VETEELTVEERNLLSVAYKN VIGARRASWRIISSIELKED SRGNSDH*VSIIKDYRGKIE
arabψ MSTREENVYMAKL AEQAERYEEMVEFMEKVAKT VDVEELSVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SKGNEDH*VAIIKDYRGEIE
hvulg MSTAEATREENVYMAKL AEQAERYEEMVEFMEKVAKT ADVGELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGNEAY*VASIKEYRTRIE
rice MSPAEASREENVYMAKL AEQAERYEEMVEFMEKVAKT TDVGELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGNEAY*VASIKEYRSRIE
zea MAKL AEQAERYEEMVEFMEKVAKT VDSEELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE GRGNEDR*VTLIKDYRGKIE
zea2 MSAPGELSREENVYMAKL AEQAERYEEMVEFMEKVAKT VDSEELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGN*EDRVTLIKEYRGKIE
hv2 MAQPAELSREENVYMAKL AEQAERYEEMVEFMEKVAKT VDSEELTVEERNLLSVAYKN VIGARRASWRIISSIEQKEE SRGN*EDRVTLIKDYRGKIE
101
150
200
mouse TELKLICCDILDVLDKHLIP AANTGESKVFYYKMKGDYHR YLAEFATGNDRKEAAENSLV AYKAASDIAMTELPPTHPIR LGLALNFSVFYYEILNSPDR
bmh2 TELTKISDDILSVLDSHLIP SATTGESKVFYYKMKGDYHR YLAEFSSGDAREKATNSSLE AYKTASEIATTELPPTHPIR LGLALNFSVFYYEIQNSPDK
bmh1 TELTKISDDILSVLDSHLIP SATTGESKVFYYKMKGDYHR YLAEFSSGDAREKATNASLE AYKTASEIATTELPPTHPIR LGLALNFSVFYYEIQNSPDK
arabw TELSGICDGILKLLDSRLIP AAASGDSKVFYLKMKGDYHR YLAEFKTGQERKDAAEHTLA AYKSAQDIANAELAPTHPIR LGLALNFSVFYYEILNSPDR
arabχ TELSDICDGILKLLDTILVP AAASGDSKVFYLKMKGDYHR YLAEFKSGQERKDAAEHTLT AYKAAQDIANSELAPTHPIR LGLALNFSVFYYEILNSPDR
arabϕ SELSKICDGILKLLDTRLVP ASANGDSKVFYLKMKGDYHR YLAEFKTGQERKDAAEHTLT AYKAAQDIANAELAPTHPIR LGLALNFSVFYYEILNSPDR
oe TELSNICGGILKLLDSRLIP SAASGDSKVFYLKMKGDYHR YLAEFKTGAERKEAAESTLS AYKAAQDIANAELAPTHPIR LGLALNFSVFYYEILNSPDR
sp KELSDNCDGILKLLDTKLVP AASSGDSKVFYLKMKGDYHR YLAEFKTGAQRKEAAESTLT AYKAAQDIANAELAPTHPIR LGLALNFSVFYYEILNSPDR
arabu TELSKICDGILNLLEAHLIP AASLAESKVFYLKMKGDYHR YLAEFKTGAERKEAAESTLV AYKSAQDIALADLAPTHPIR LGLALNSSVFYYEILNSSDR
arabψ SELSKICDGILNVLEAHLIP SASPAESKVFYLKMKGDYHR YLAEFKAGAERKEAAESTLV AYKSASDIATAELAPTHPIR LGLALNFSVFYYEILNSPDR
hvulg TELSKICDGILKLLDSHLVP SATAAESKVFYLKMKGDYHR YLAEFKAGAERKEAAENTLV AYKSAQDIALADLPTTHPIR LGLALNFSVFYYEILNSPDR
rice TELSKICDGILKLLDSHLVP SATAAESNVFYLKMKGDYHR YLAEFKSGAERKEAAENTLV AYKSAQDIALADLPTTHPIR LGLALNLSVFYYEILNSPDR
zea TELTKICDGILKLLESHLVP SSTAPESKVFYLKMKGDYYR YLAEFKTGAERKDAAENTMV AYKAAQDIALAELAPTHPIR LGLALNFSVFYYEILNSPDR
zea2 TELSKICDGILKLLETHLVP SSTAP
zea2
ILNSPDR
hv2 VELTKICDGILKLLDSHLVP SSTAPESKVFYLKMKGDYYR YLAEFKSGTERKDAAENTMV AYKAAQEIALAELPTTHPIR LGLALNFSVFYYEILNSPDR

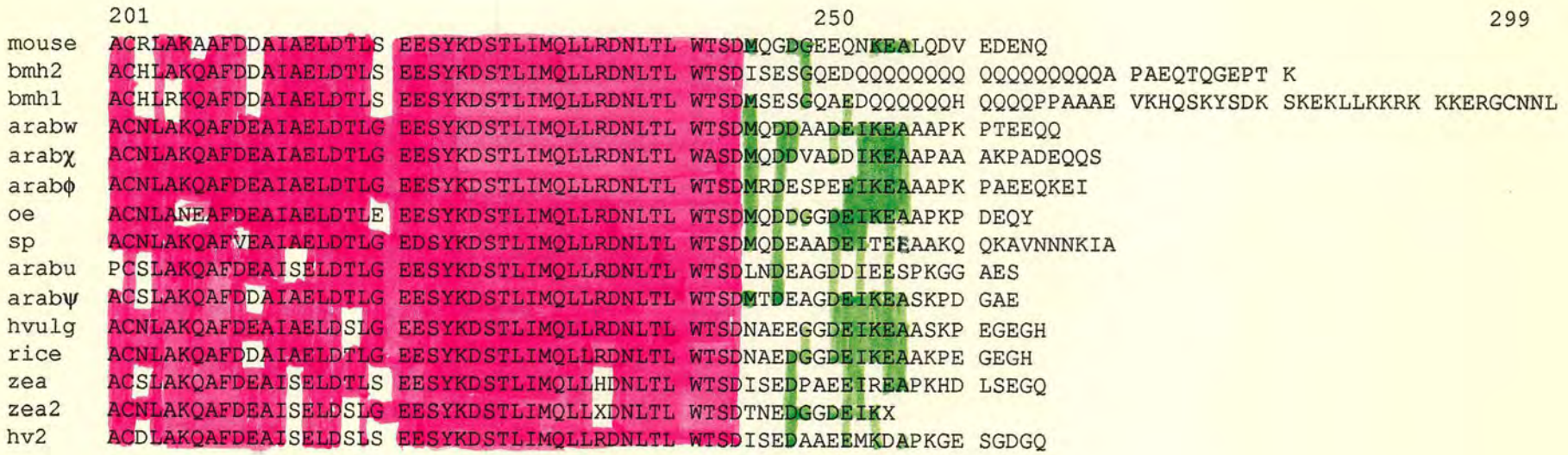

201
250
299
mouse ACRLAKAAFDDAIAELDTLS EESYKDSTLIMQLLRDNLTL WTSDMQGDGEEQNKEALQDV EDENQ
bmh2 ACHLAKQAFDDAIAELDTLS EESYKDSTLIMQLLRDNLTL WTSDISESGQEDQQQQQQQQ QQQQQQQQQA PAEQTQGEPT K
bmh1 ACHLRKQAFDDAIAELDTLS EESYKDSTLIMQLLRDNLTL WTSDMSESGQAEDQQQQQQH QQQQPPAAAE VKHQSKYSDK SKEKLLKKRK KKERGCNNL
arabw ACNLAKQAFDEAIAELDTLG EESYKDSTLIMQLLRDNLTL WTSDMQDDAADEIKEAAAPK PTEEQQ
arabχ ACNLAKQAFDEAIAELDTLG EESYKDSTLIMQLLRDNLTL WASDMQDDVADDIKEAAPAA AKPADEQQS
arabϕ ACNLAKQAFDEAIAELDTLG EESYKDSTLIMQLLRDNLTL WTSDMRDESPEEIKEAAAPK PAEEQKEI
oe ACNLANEAFDEAIAELDTLE EESYKDSTLIMQLLRDNLTL WTSDMQDDGGDEIKEAAPKP DEQY
sp ACNLAKQAFVEAIAELDTLG EDSYKDSTLIMQLLRDNLTL WTSDMQDEAADEITEEAAKQ QKAVNNNKIA
arabu PCSLAKQAFDEAISELDTLG EESYKDSTLIMQLLRDNLTL WTSDLNDEAGDDIEESPKGG AES
arabψ ACSLAKQAFDDAIAELDTLG EESYKDSTLIMQLLRDNLTL WTSDMTDEAGDEIKEASKPD GAE
hvulg ACNLAKQAFDEAIAELDSLG EESYKDSTLIMQLLRDNLTL WTSDNAEEGGDEIKEAASKP EGEGH
rice ACNLAKQAFDDAIAELDTLG EESYKDSTLIMQLLRDNLTL WTSDNAEDGGDEIKEAAKPE GEGH
zea ACSLAKQAFDEAISELDTLS EESYKDSTLIMQLLHDNLTL WTSDISEDPAEEIREAPKHD LSEGQ
zea2 ACNLAKQAFDEAISELDSLG EESYKDSTLIMQLLXDNLTL WTSDTNEDGGDEIKX
hv2 ACDLAKQAFDEAISELDSLS EESYKDSTLIMQLLRDNLTL WTSDISEDAAEEMKDAPKGE SGDGQ

Figure 6.5

Alignment of all the plant, yeast and mouse epsilon proteins deduced hydrophilicities calculated using the UWGCG programme PLOTSTRUCTURE and PEPTIDESTRUCTURE.

Abbreviations used are:-

mouse-mouse 14-3-3 epsilon isoform

bmh1/bmh2 Yeast-(*S. cerevisiae*) brain modulosignalin homologue

arab- thale cress (*Arabidopsis thaliania)*

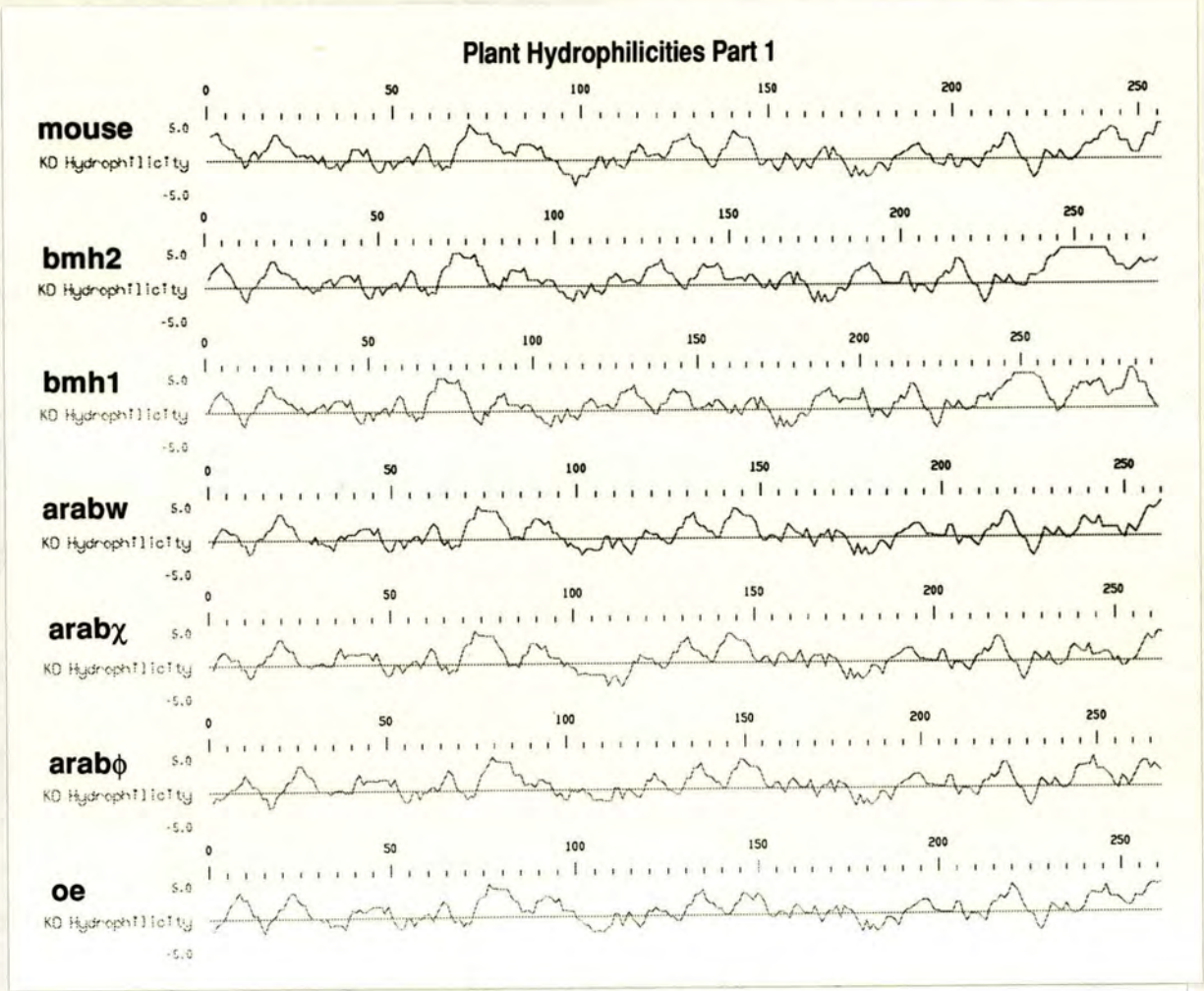
oe- evening primrose (*Oenothera hookeri)*

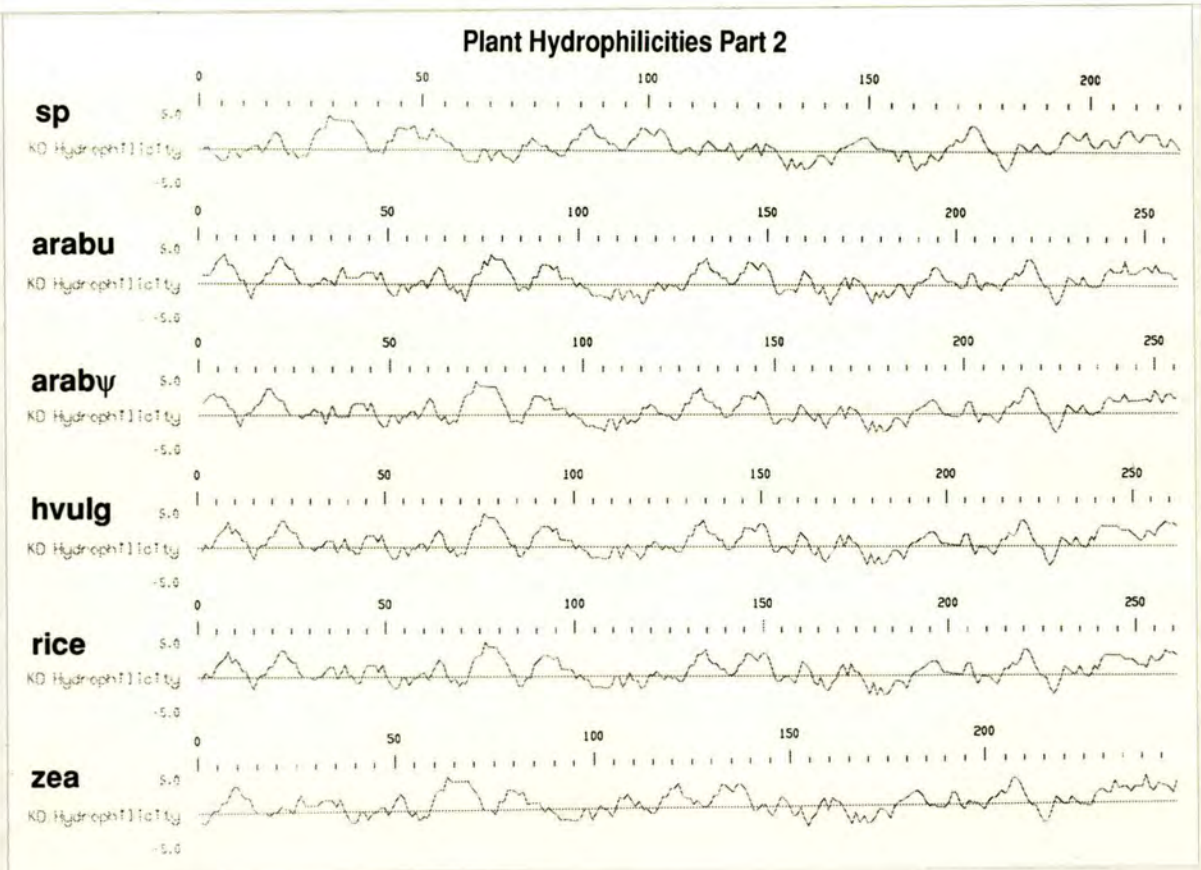
sp-spinach (*Spinacerea oleracea*)

rice- (*Oryza sativa*)

zea/zea2-maize (*Zea mays)*

hvulg/hv2-barley (*Hordeum vulgare)*


Plant Hydrophilicities Part 1
mouse
bmh2
bmh1
arabw
arabχ
arabϕ
oe
KD Hydrophilicity
5.0
-5.0
0
50
100
150
200
250


Plant Hydrophilicities Part 2
sp
arabu
arabψ
hvulg
rice
zea
KD Hydrophilicity
5.0
-5.0
0
50
100
150
200
250

Chapter 7

Discussion and Conclusions

The previous chapters have presented the results and made at least preliminary conclusions. Now all the results are considered as a whole allowing more rounded conclusions to be made. Speculation can also be made on the future work that can be done using the library and the subtraction protocol. Perhaps, more importantly the function and role of the mouse 14-3-3 ε isoform in development can be speculated upon and discussed.

7.1 Evaluation of the cDNA library

The aim of the project was to isolate cDNAs for genes involved in nephrogenesis and a major accomplishment in this project was the creation of a 14.5 d.p.c. fetal kidney library. The choice of 14.5 d.p.c. kidneys for the cDNA for the library and subtracted cDNA was chosen as this stage of kidney development should encompasses all the possible early inductive stages of nephrogenesis. Therefore, a 14.5 d.p.c. fetal kidney λZAPII cDNA library was made that has a complexity of 1.15 x 10^6 original clones. Further analysis of the library has shown that it contains representative amounts of a metabolic pathway cDNA (glyceraldehyde 3 phosphate dehydrogenase), showing that the library should contain most genes expressed in the fetal kidney. The clones that have been isolated from this library have included three full length cDNAs (B2, C3 and D1) and other cDNAs of a large size (D3 and data not shown). Therefore the reverse transcription and second strand synthesis must have been efficient. This project has created a valuable resource of a cDNA library which contains most genes expressed in the kidney and at least a proportion that provide large full length cDNAs. This library can now be used for cloning material identified by the subtracted probe discussed here or other subtracted probes and also for many other purposes.

7.1.1 Future work

Using the library it will be possible to screen for clones that may be interesting in kidney development using probes from other species or homology to conserved motifs (again as discussed in chapter 3). Although this can have drawbacks, it would be very

informative. For instance, *c-ros* expression is coincident with the proliferation of epithelium. It is expressed at the terminal ends of the branching ureteric tubules and expression disappears as the primary induced epithelium differentiates into the comma-shaped body. It is involved in the obligate cell-cell interactions of the induction and proliferation of the epithelial cells of the kidney and work on the *Drosophila* homologues suggests it is involved in reciprocal signalling between two cell types. *c-ros* is thought to be the mammalian homologue of *sevenless* suggesting that other members of the *sevenless* pathway may have homologues that act in the fetal kidney development. As the RTK-mediated signalling pathways are turning out to be key mediators of developmental pathways; mouse homologues of *boss* (the sevenless ligand) or *drk* (acting downstream of sevenless) might be ideal candidates for kidney regulator genes. These clones could be isolated by either screening the library at low stringency with the *Drosophila* cDNAs or designing degenerate primers for conserved regions and using RNA-PCR amplification. Such work could isolate genes of immense interest in the field of kidney research.

7.2 The subtractive hybridization

The project's initial aim was to devise a novel, simple and technically undemanding subtractive hybridization technique and with this technique to find genes important in kidney development. Subtractive hybridization is a very useful technique for isolating genes that are expressed in a tissue or developmental stage of interest as discussed in chapter 3. This thesis has used a strategy that incorporates several advantages from different strategies. The use of photoactivatable biotin to label the driver messenger mRNA and phenol to remove a streptavidin:biotin complex was taken from Sive & St John (1988). However the strategy that I have used improves upon this method by using tagged primers to allow PCR amplification of the subtracted cDNA. This part of the protocol is a modification of the strategies used by Timblin *et al.* (1990), Duguid and Dinaeur (1990), and Hla and Maciag (1990). The use of PCR to amplify the subtracted cDNA means that the subtraction need only be performed once to get a

renewable source of the cDNA. Other strategies have cloned the material resulting from subtracting and PCR amplifying the cDNA (Timblin *et al.*, 1990; Duguid & Dinaeur, 1990). cDNA that has been manipulated extensively then PCR amplified may give rise to PCR amplification artefacts due to misincorporation of nucleotides (Tindall & Kunkel, 1988). This reduces the probability of isolating full length clones. In contrast, using the amplified cDNA as a probe to screen non-manipulated cDNA libraries enhances the possibility of getting full length cDNAs with no sequence artefacts.

In retrospect, subtractive hybridization can be very useful but the tissues to subtract with and against must be chosen with extreme care. The most important decision that was made in this project was the tissue against which to subtract the fetal kidney cDNA. The primary decision was to use adult liver mRNA to subtract against. Looking at the clones that were ultimately analysed this proved to cause minor problems as the resulting subtracted cDNA included a large number of species that may have been isolated due to the inclusion of B1 repeats in their sequence. B1 repeats are the most prevalent transcript in the embryo and the use of an adult driver led to the isolation of cDNAs with B1 repeats in them. Northern and Southern blots using clones with B1 repeats in their UTR have a high background hybridization to the many (40-100,000) copies of B1 repeats in the genome and the large number of transcripts that contain B1 repeats, which will prevent the detection of an unique signal. As northern blot analysis was one criterion for evaluating the potential value of clones then at least one avenue of detection of potentially valuable clones is being lost.

It has been proven that the subtracted cDNA can be used to identify cDNAs that are not expressed in the adult liver but that are highly expressed in the kidney. The isolation of the 14-3-3 protein shows another value of a subtractive approach, that cDNAs that otherwise may not have been studied in kidney development can be studied. An important question is of course why the 14-3-3 ε cDNA was cloned three times in this screen. Was this just chance or is the ε isoform really expressed at an

abnormally high level in the fetal kidney? The project screened a mere 5000 clones from the library suggesting that the cDNA is present at 0.06% of the population. This level of representation would be comparable to levels of extremely highly expressed 'house-keeping' genes like actin. A further screening of the library with a different subtracted probe (data not shown) isolated some 40 clones. Very preliminary sequence has been obtained for 15 and the ε isoform has only been found once. The insert size of another 10 (too large) suggest they also are not ε isoform cDNAs. This suggests that the gene is highly represented in the fetal kidney library but not abnormally so. The analysis of the *in situ* hybridization data has shown that the gene is highly expressed in undifferentiated mesenchyme, and hence highly in the 14.5 d.p.c. developing kidney from which the library was made. It has not been detected in the mouse adult liver mRNA by northern analysis or in mid-gestation embryo liver by *in situ* hybridization analysis. Therefore it will be represented in the subtracted cDNA but should be proportionately so. It may be possible the mouse ε isoform cDNA was preferentially amplified in the PCR step of the protocol. As the ε cDNA is of a median size then it should be proportionally amplified during PCR; there is no obvious reason why the cDNA should be over-represented because of an abnormally small size. There is however the possibility that the primer used for the lone-linker PCR amplification may have primed within the ε cDNA sequence. Aligning the primer with the cDNA, a region of moderate homology can be seen with the strongest homology at the 3' priming end. This might cause internal priming of the cDNA, truncating the PCR product and helping to over-amplify the cDNA.

Figure 7.1

```
604 ctctcaacttttccgtattct 624
      || ||  | ||| |||||
  1 GGGTCGACGGTACCGAATTCT 21
```

The above figure aligns the mouse ε cDNA (in lower case) with the primer used to amplify the subtracted cDNA (capitals).

7.3 Evaluation of the cDNA library and subtraction

The fact that the library has proved so valuable and that the subtraction has been very successful shows that the project has been a success. The success of the subtraction has been further analyzed by the screening of fetal kidney and adult liver mRNA (Figure 3.3) with the cDNA showing that the cDNA has been greatly enriched for mRNAs that are expressed in the fetal kidney but not the liver. From this experiment it can be estimated that the subtraction has been 95-99% efficient at removing cDNAs expressed in the mouse adult liver. One faint band is barely detectable and it is interesting that a similar species is detected in the fetal kidney at a very high level. The persistence of this cDNA may be caused by the amount of the mRNA in the starting sample being disproportionally high. Clones representing this band have not been found in our initial screen so the residual co-expressed clones not subtracted are not an insurmountable obstacle. No subtractive screens can remove all unwanted cDNAs, realistically the most successful subtraction will aim to achieve 95-99% enrichment. Therefore a successful strategy has been devised that can be repeated using different tissues for more selective isolation of clones, a reliable library is available to allow isolation of important mRNAs.

7.3.1 Future work

The subtracted cDNA created has only started to be used effectively, only 5000 phage clones were screened. To fully utilise this resource many thousand phage would have to be screened to isolate all the cDNAs corresponding to the subtracted pool.

Secondly, as I have a greater perspective of the appropriate tissues to subtract against, the subtraction could be repeated with more effective tissues. The use of other tissue RNA as driver was considered after starting to analyse the results. Adult kidney and fetal liver mRNA were considered both individually and as joint driver mRNA sources. These possibilities were explored and the initial stages of the experiments were done but time limitations prevented the completion of these experiments.

7.4 The mouse 14-3-3 ε isoform

This thesis has discussed the isolation of the mouse 14-3-3 ε isoform while the protein family and the inter-relationships between the many different isoforms has been described in the previous chapter. Comparison of the sequences shows that the ε isoform is perhaps one of the most interesting of the mammalian isoforms as it appears to be closest to the ancestral isoform from which the family members diverged. I propose this because, when the amino acid sequences are compared using either the UWGCG programme BESTFIT or PILEUP, the ε isoform shows greater likeness to the isoforms expressed in yeast and plant than to the other mammalian isoforms. The ε isoform is also the only mammalian 14-3-3 protein that can not be phosphorylated (Toker *et al.*, 1992) suggesting it may even have a different mode of action from the others.

The chromosomal localisation of the mouse ε isoform to chromosome 11 would suggest that the different isoforms are not clustered on one chromosome, as the only two isoforms for which the chromosomal locations are known do not appear to be on syntenic chromosomes. The mouse epsilon isoform has been shown to localise to chromosome 11 (this thesis), and the human η isoform has been localised to chromosome 22, these two chromosomes are not syntenic in mouse and man. This result will need to be confirmed directly with mapping of all isoforms in both species.

The ε isoform has been isolated in a number of different contexts. The bovine brain protein fraction that was first shown to activate phosphorylated tyrosine and tryptophan hydroxylase in the catecholamine pathway (Isobe *et al.*, 1988) was shown to contain seven mammalian 14-3-3 isoforms including ε. This is the first published isolation of the ε isoform. A few years later Aitken and co-workers found the protein kinase C inhibitor protein, KCIP, which was shown to contain seven sheep 14-3-3 isoforms including ε (Toker *et al.*, 1990, 1992). More recently a protein fraction, Exo1 was shown to reactivate exocytosis from adrenal chromaffin cells (Morgan & Burgoyne, 1992) and this fraction contains all the 14-3-3 isoforms (from either human

adrenal chromaffin cells or sheep brain) including the ε isoform. Indeed, using antibodies, it has been showed that the ε isoform leaks from the chromaffin cells when they are permeabilised (Morgan *et al.*, 1993a). The rat ε isoform protein has also been found as a major contaminant of a highly purified pineal N-acetyl transferase (Roseboom *et al.*, 1992). All the above studies report the isolation of the ε protein, mostly in a protein preparation that co-purified all seven isoforms. The rat and sheep ε isoform cDNAs have now been cloned and deposited in the GENBANK database. The rat sequence has been reported as a poster abstract (Roseboom *et al.*, 1992) and the sheep sequence was reported in a review of the 14-3-3 family (Aitken *et al.*, 1992). However this thesis describes the cloning of the first mouse isoform cloned and also the first comprehensive developmental study of the expression of a 14-3-3 isoform during embryogenesis.

The cloning of both the rat and the sheep ε DNA sequences allows the comparison of the three DNA sequences, including the comparison of the 5' and 3' untranslated regions of the rat and mouse cDNAs. Only the coding region of the sheep cDNA has been cloned. As previously shown in table 5.1, the conservation between the 5' and 3' regions of mouse and rat is remarkably high (98.6 and 97.6% identical respectively). This figure is higher than would be expected for the evolutionary distance between mouse and rat (the homology between the rat and mouse β-actin 3' UTRs is 85%) suggesting there may be a conserved function of the ε isoform's 3' UTR such as a regulatory function (Yaffe *et al.*, 1985).

7.4.1 The role of the 14-3-3 ε isoform in the kidney

The 14-3-3 family was first identified as proteins that bound to phosphorylated tyrosine and tryptophan hydroxylases and induced a conformational change. The Exo1 activity may well be similar, given that Exo2, the other protein found to activate exocytosis, has been shown to be the catalytic subunit of protein kinase A (Morgan *et al.*, 1993b). This implies that a likely cause of exocytotic rundown is a lack of a particular kinase or misbalance between a protein kinase and a protein phosphatase, or

the absence of an activator of the kinase or the kinase's substrate. It is probable that the 14-3-3 proteins activate a phosphorylated protein that then allows exocytosis. Therefore I suggest that the general role of the 14-3-3 proteins is to bind to specific phosphorylated proteins and induce a conformational change that activates the phosphorylated state or prolongs the active state of an intracellular signal. This hypothesis would suggest that the 14-3-3 ε is expressed at a high level to allow a signal transduction pathway (or multiple pathways) to act at a high level both in the developing neural system and in the kidney. The pathway (not identified in this study) is presumed to be active in the adult mouse but at a far lower level than during embryogenesis when cells have to react to many signals from their surroundings. The kidney mesenchyme is a particularly good example of this, as many cells are being programmed to undergo major identity changes and this will involve many signal transduction pathways. I also suggest that the pathways are not shutdown in the adult and the pathway that the 14-3-3 proteins are involved in are maintained at a basal level to allow for cellular replacements throughout the mouse's lifetime. The signal transduction pathways may not only used during nephrogenesis but also in the adult kidney to modulate renal function.

7.4.2 Future work

The above discussion of the ε isoform and speculation on its function in the kidney of course ask more questions than this thesis can answer. It is therefore instructive to speculate on the next steps that are suggested by the information that is known.

If the developmental expression of this protein family were to be studied, then a comparative analysis of the expression profiles during embryogenesis would first be needed. This would require isolating mouse isoform-specific cDNA probes for all the known isoforms for *in situ* hybridization analysis. A first approach to this would be to screen the fetal kidney library at low stringency with the ε isoform. As the 14-3-3 family are a discrete family with no other protein families that are similar, this approach should be successful and not prone to the cloning of non-14-3-3 proteins.

Several of the other isoforms cloned in different species are expressed in the adult kidney as shown by northern blotting (β and η [Watanabe *et al.*, 1991, 1993a]; θ [Nielsen, 1991]) and hence are likely to be expressed in fetal kidney. A quick method to identify each isoform would be to sequence small regions of the cDNA that correspond to the variable regions of the protein. RNA-PCR amplification of RNA samples from different tissues could then be used to clone any isoforms missed in this screen. With the full range of mouse 14-3-3 cDNAs, *in situ* hybridization analysis could be performed on embryos of varying ages to show how the expression of each isoform fluctuates during development.

Previously, antibodies to individual isoforms have been shown to cross react between different mammals (Ichimura *et al.*, 1991) and now antibodies that are specific to a single isoform (excepting α and β; and δ and ζ for which there is one antibody for each pair) have been produced (Martin *et al.*, 1993). These antibodies could be used for immune histochemistry on embryo sections to show where the proteins are presents in the developing embryo. The comparison of the two patterns will allow both confirmation of expression pattern and also would detect secondary regulation at the level of translation.

A further important question raised by other workers in the 14-3-3 field concerns the cellular localisation of the 14-3-3 family. The work by Lu *et al.*, (1992) showed that the *Arabidopsis* homologue was found in a DNA-binding complex suggesting that at least in *Arabidopsis* one isoform may be nuclear. The majority of other work has suggested that the protein is cytoplasmic (reviewed in Aitken *et al.*, 1992). This could be resolved by subcellular fractionation, making nuclear extracts, cytoplasmic extracts and membrane preparations and showing relative amounts of 14-3-3 protein in each fraction.

Another major question concerns the dimerisation of the 14-3-3 proteins. It has been shown that the proteins when run on native protein gels run at a molecular weight of 58 kDa, when the normal monomeric molecular weight is 29-32 kDa suggesting that the 14-3-3 proteins naturally act as dimers (Loeb & Gross, 1986).

Therefore, the next point to address is whether the proteins form hetero- or homodimers within the cells and whether dimerisation is necessary for protein activity. The use of antibodies to the protein family may answer this question. The antibodies could be used to immunoprecipitate the proteins in tissue extracts. Examination of the purified proteins would allow the question of which isoforms (if any) can form hetero-dimers to be answered. However, such experiments are dependant on there being experimental conditions that allow both dimerisation of the proteins and immune precipitation without disturbing the protein-protein interaction.

Another way to answer this question would be to create chimeric proteins with peptides added to facilitate purification (the glutathione S-transferase gene fusion system, Smith & Johnson, 1988; the maltose fusion protein system, Di Guan *et al.*, 1988; the His tagging system, Janknecht *et al.*, 1991). Such fusion proteins would allow specific isolation of a particular isoform. The use of such tagged proteins may provide a more sensitive method of finding both which isoforms dimerise and may even allow purification of the protein(s) that the 14-3-3 protein interacts with. The tagged 14-3-3 proteins could be expressed in either the usual *E. coli* systems, in mammalian cells or in yeast which allow proper modification of non-yeast 14-3-3 proteins (Collinge *et al.*, 1993). Co-expression of a tagged and non-tagged isoform would show if the non-tagged isoform was co-purified with the tagged protein. Systematic analysis of the nine (or possibly more) isoforms would then show which isoforms can dimerise and which ones can not.

It is also important to find the proteins with which the 14-3-3 isoforms interact. I have already discussed the idea of co-immune precipitation and the fact that this would depend on the conditions of protein interaction being compatible with the immune precipitation. Another approach could be performed in the yeast two hybrid system which allows the identification of interacting proteins (Field & Song, 1989; Chien *et al.*, 1991). This would be of immense help in clarifying the speculation on the real function on the 14-3-3 proteins.

The function of the ε isoform in development should also be studied more directly. The gene could be removed by homologous recombination in ES cells to look at the effect of deficiency of this gene in the whole organism. This can also be studied specifically in the kidney by the use of anti-sense oligonucleotides to the ε isoform in the *in vitro* kidney culture system discussed in section 1.6.6. This would answer questions about the functional redundancy of the ε protein within the 14-3-3 family. The effect of over-expression would also be of interest, as many of the genes involved in the signal transduction pathway have been shown to be oncogenic when over-expressed (*c-ret*, *c-met*, *ras*). It is conceivable that this gene may also have a oncogenic effect when over-expressed in the embryo. All these experiments would allow a greater understanding of how the ε isoform and the 14-3-3 family function during development.

Chapter 7

Discussion and Conclusions

The previous chapters have presented the results and made at least preliminary conclusions. Now all the results are considered as a whole allowing more rounded conclusions to be made. Speculation can also be made on the future work that can be done using the library and the subtraction protocol. Perhaps, more importantly the function and role of the mouse 14-3-3 ε isoform in development can be speculated upon and discussed.

7.1 Evaluation of the cDNA library

The aim of the project was to isolate cDNAs for genes involved in nephrogenesis and a major accomplishment in this project was the creation of a 14.5 d.p.c. fetal kidney library. The choice of 14.5 d.p.c. kidneys for the cDNA for the library and subtracted cDNA was chosen as this stage of kidney development should encompasses all the possible early inductive stages of nephrogenesis. Therefore, a 14.5 d.p.c. fetal kidney λZAPII cDNA library was made that has a complexity of 1.15 x 10^6 original clones. Further analysis of the library has shown that it contains representative amounts of a metabolic pathway cDNA (glyceraldehyde 3 phosphate dehydrogenase), showing that the library should contain most genes expressed in the fetal kidney. The clones that have been isolated from this library have included three full length cDNAs (B2, C3 and D1) and other cDNAs of a large size (D3 and data not shown). Therefore the reverse transcription and second strand synthesis must have been efficient. This project has created a valuable resource of a cDNA library which contains most genes expressed in the kidney and at least a proportion that provide large full length cDNAs. This library can now be used for cloning material identified by the subtracted probe discussed here or other subtracted probes and also for many other purposes.

7.1.1 Future work

Using the library it will be possible to screen for clones that may be interesting in kidney development using probes from other species or homology to conserved motifs (again as discussed in chapter 3). Although this can have drawbacks, it would be very

informative. For instance, *c-ros* expression is coincident with the proliferation of epithelium. It is expressed at the terminal ends of the branching ureteric tubules and expression disappears as the primary induced epithelium differentiates into the comma-shaped body. It is involved in the obligate cell-cell interactions of the induction and proliferation of the epithelial cells of the kidney and work on the *Drosophila* homologues suggests it is involved in reciprocal signalling between two cell types. *c-ros* is thought to be the mammalian homologue of *sevenless* suggesting that other members of the *sevenless* pathway may have homologues that act in the fetal kidney development. As the RTK-mediated signalling pathways are turning out to be key mediators of developmental pathways; mouse homologues of *boss* (the sevenless ligand) or *drk* (acting downstream of sevenless) might be ideal candidates for kidney regulator genes. These clones could be isolated by either screening the library at low stringency with the *Drosophila* cDNAs or designing degenerate primers for conserved regions and using RNA-PCR amplification. Such work could isolate genes of immense interest in the field of kidney research.

7.2 The subtractive hybridization

The project's initial aim was to devise a novel, simple and technically undemanding subtractive hybridization technique and with this technique to find genes important in kidney development. Subtractive hybridization is a very useful technique for isolating genes that are expressed in a tissue or developmental stage of interest as discussed in chapter 3. This thesis has used a strategy that incorporates several advantages from different strategies. The use of photoactivatable biotin to label the driver messenger mRNA and phenol to remove a streptavidin:biotin complex was taken from Sive & St John (1988). However the strategy that I have used improves upon this method by using tagged primers to allow PCR amplification of the subtracted cDNA. This part of the protocol is a modification of the strategies used by Timblin *et al.* (1990), Duguid and Dinaeur (1990), and Hla and Maciag (1990). The use of PCR to amplify the subtracted cDNA means that the subtraction need only be performed once to get a

renewable source of the cDNA. Other strategies have cloned the material resulting from subtracting and PCR amplifying the cDNA (Timblin *et al.*, 1990; Duguid & Dinaeur, 1990). cDNA that has been manipulated extensively then PCR amplified may give rise to PCR amplification artefacts due to misincorporation of nucleotides (Tindall & Kunkel, 1988). This reduces the probability of isolating full length clones. In contrast, using the amplified cDNA as a probe to screen non-manipulated cDNA libraries enhances the possibility of getting full length cDNAs with no sequence artefacts.

In retrospect, subtractive hybridization can be very useful but the tissues to subtract with and against must be chosen with extreme care. The most important decision that was made in this project was the tissue against which to subtract the fetal kidney cDNA. The primary decision was to use adult liver mRNA to subtract against. Looking at the clones that were ultimately analysed this proved to cause minor problems as the resulting subtracted cDNA included a large number of species that may have been isolated due to the inclusion of B1 repeats in their sequence. B1 repeats are the most prevalent transcript in the embryo and the use of an adult driver led to the isolation of cDNAs with B1 repeats in them. Northern and Southern blots using clones with B1 repeats in their UTR have a high background hybridization to the many (40-100,000) copies of B1 repeats in the genome and the large number of transcripts that contain B1 repeats, which will prevent the detection of an unique signal. As northern blot analysis was one criterion for evaluating the potential value of clones then at least one avenue of detection of potentially valuable clones is being lost.

It has been proven that the subtracted cDNA can be used to identify cDNAs that are not expressed in the adult liver but that are highly expressed in the kidney. The isolation of the 14-3-3 protein shows another value of a subtractive approach, that cDNAs that otherwise may not have been studied in kidney development can be studied. An important question is of course why the 14-3-3 ε cDNA was cloned three times in this screen. Was this just chance or is the ε isoform really expressed at an

abnormally high level in the fetal kidney? The project screened a mere 5000 clones from the library suggesting that the cDNA is present at 0.06% of the population. This level of representation would be comparable to levels of extremely highly expressed 'house-keeping' genes like actin. A further screening of the library with a different subtracted probe (data not shown) isolated some 40 clones. Very preliminary sequence has been obtained for 15 and the ε isoform has only been found once. The insert size of another 10 (too large) suggest they also are not ε isoform cDNAs. This suggests that the gene is highly represented in the fetal kidney library but not abnormally so. The analysis of the *in situ* hybridization data has shown that the gene is highly expressed in undifferentiated mesenchyme, and hence highly in the 14.5 d.p.c. developing kidney from which the library was made. It has not been detected in the mouse adult liver mRNA by northern analysis or in mid-gestation embryo liver by *in situ* hybridization analysis. Therefore it will be represented in the subtracted cDNA but should be proportionately so. It may be possible the mouse ε isoform cDNA was preferentially amplified in the PCR step of the protocol. As the ε cDNA is of a median size then it should be proportionally amplified during PCR; there is no obvious reason why the cDNA should be over-represented because of an abnormally small size. There is however the possibility that the primer used for the lone-linker PCR amplification may have primed within the ε cDNA sequence. Aligning the primer with the cDNA, a region of moderate homology can be seen with the strongest homology at the 3' priming end. This might cause internal priming of the cDNA, truncating the PCR product and helping to over-amplify the cDNA.

Figure 7.1

```
                .
604 ctctcaacttttccgtattct 624
      || ||  | ||| |||||
  1 GGGTCGACGGTACCGAATTCT 21
```

The above figure aligns the mouse ε cDNA (in lower case) with the primer used to amplify the subtracted cDNA (capitals).

7.3 Evaluation of the cDNA library and subtraction

The fact that the library has proved so valuable and that the subtraction has been very successful shows that the project has been a success. The success of the subtraction has been further analyzed by the screening of fetal kidney and adult liver mRNA (Figure 3.3) with the cDNA showing that the cDNA has been greatly enriched for mRNAs that are expressed in the fetal kidney but not the liver. From this experiment it can be estimated that the subtraction has been 95-99% efficient at removing cDNAs expressed in the mouse adult liver. One faint band is barely detectable and it is interesting that a similar species is detected in the fetal kidney at a very high level. The persistence of this cDNA may be caused by the amount of the mRNA in the starting sample being disproportionally high. Clones representing this band have not been found in our initial screen so the residual co-expressed clones not subtracted are not an insurmountable obstacle. No subtractive screens can remove all unwanted cDNAs, realistically the most successful subtraction will aim to achieve 95-99% enrichment. Therefore a successful strategy has been devised that can be repeated using different tissues for more selective isolation of clones, a reliable library is available to allow isolation of important mRNAs.

7.3.1 Future work

The subtracted cDNA created has only started to be used effectively, only 5000 phage clones were screened. To fully utilise this resource many thousand phage would have to be screened to isolate all the cDNAs corresponding to the subtracted pool.

Secondly, as I have a greater perspective of the appropriate tissues to subtract against, the subtraction could be repeated with more effective tissues. The use of other tissue RNA as driver was considered after starting to analyse the results. Adult kidney and fetal liver mRNA were considered both individually and as joint driver mRNA sources. These possibilities were explored and the initial stages of the experiments were done but time limitations prevented the completion of these experiments.

7.4 The mouse 14-3-3 ε isoform

This thesis has discussed the isolation of the mouse 14-3-3 ε isoform while the protein family and the inter-relationships between the many different isoforms has been described in the previous chapter. Comparison of the sequences shows that the ε isoform is perhaps one of the most interesting of the mammalian isoforms as it appears to be closest to the ancestral isoform from which the family members diverged. I propose this because, when the amino acid sequences are compared using either the UWGCG programme BESTFIT or PILEUP, the ε isoform shows greater likeness to the isoforms expressed in yeast and plant than to the other mammalian isoforms. The ε isoform is also the only mammalian 14-3-3 protein that can not be phosphorylated (Toker *et al.*, 1992) suggesting it may even have a different mode of action from the others.

The chromosomal localisation of the mouse ε isoform to chromosome 11 would suggest that the different isoforms are not clustered on one chromosome, as the only two isoforms for which the chromosomal locations are known do not appear to be on syntenic chromosomes. The mouse epsilon isoform has been shown to localise to chromosome 11 (this thesis), and the human η isoform has been localised to chromosome 22, these two chromosomes are not syntenic in mouse and man. This result will need to be confirmed directly with mapping of all isoforms in both species.

The ε isoform has been isolated in a number of different contexts. The bovine brain protein fraction that was first shown to activate phosphorylated tyrosine and tryptophan hydroxylase in the catecholamine pathway (Isobe *et al.*, 1988) was shown to contain seven mammalian 14-3-3 isoforms including ε. This is the first published isolation of the ε isoform. A few years later Aitken and co-workers found the protein kinase C inhibitor protein, KCIP, which was shown to contain seven sheep 14-3-3 isoforms including ε (Toker *et al.*, 1990, 1992). More recently a protein fraction, Exo1 was shown to reactivate exocytosis from adrenal chromaffin cells (Morgan & Burgoyne, 1992) and this fraction contains all the 14-3-3 isoforms (from either human

adrenal chromaffin cells or sheep brain) including the ε isoform. Indeed, using antibodies, it has been showed that the ε isoform leaks from the chromaffin cells when they are permeabilised (Morgan *et al.*, 1993a). The rat ε isoform protein has also been found as a major contaminant of a highly purified pineal N-acetyl transferase (Roseboom *et al.*, 1992). All the above studies report the isolation of the ε protein, mostly in a protein preparation that co-purified all seven isoforms. The rat and sheep ε isoform cDNAs have now been cloned and deposited in the GENBANK database. The rat sequence has been reported as a poster abstract (Roseboom *et al.*, 1992) and the sheep sequence was reported in a review of the 14-3-3 family (Aitken *et al.*, 1992). However this thesis describes the cloning of the first mouse isoform cloned and also the first comprehensive developmental study of the expression of a 14-3-3 isoform during embryogenesis.

The cloning of both the rat and the sheep ε DNA sequences allows the comparison of the three DNA sequences, including the comparison of the 5' and 3' untranslated regions of the rat and mouse cDNAs. Only the coding region of the sheep cDNA has been cloned. As previously shown in table 5.1, the conservation between the 5' and 3' regions of mouse and rat is remarkably high (98.6 and 97.6% identical respectively). This figure is higher than would be expected for the evolutionary distance between mouse and rat (the homology between the rat and mouse β-actin 3' UTRs is 85%) suggesting there may be a conserved function of the ε isoform's 3' UTR such as a regulatory function (Yaffe *et al.*, 1985).

7.4.1 The role of the 14-3-3 ε isoform in the kidney

The 14-3-3 family was first identified as proteins that bound to phosphorylated tyrosine and tryptophan hydroxylases and induced a conformational change. The Exo1 activity may well be similar, given that Exo2, the other protein found to activate exocytosis, has been shown to be the catalytic subunit of protein kinase A (Morgan *et al.*, 1993b). This implies that a likely cause of exocytotic rundown is a lack of a particular kinase or misbalance between a protein kinase and a protein phosphatase, or

the absence of an activator of the kinase or the kinase's substrate. It is probable that the 14-3-3 proteins activate a phosphorylated protein that then allows exocytosis. Therefore I suggest that the general role of the 14-3-3 proteins is to bind to specific phosphorylated proteins and induce a conformational change that activates the phosphorylated state or prolongs the active state of an intracellular signal. This hypothesis would suggest that the 14-3-3 ε is expressed at a high level to allow a signal transduction pathway (or multiple pathways) to act at a high level both in the developing neural system and in the kidney. The pathway (not identified in this study) is presumed to be active in the adult mouse but at a far lower level than during embryogenesis when cells have to react to many signals from their surroundings. The kidney mesenchyme is a particularly good example of this, as many cells are being programmed to undergo major identity changes and this will involve many signal transduction pathways. I also suggest that the pathways are not shutdown in the adult and the pathway that the 14-3-3 proteins are involved in are maintained at a basal level to allow for cellular replacements throughout the mouse's lifetime. The signal transduction pathways may not only used during nephrogenesis but also in the adult kidney to modulate renal function.

7.4.2 Future work

The above discussion of the ε isoform and speculation on its function in the kidney of course ask more questions than this thesis can answer. It is therefore instructive to speculate on the next steps that are suggested by the information that is known.

If the developmental expression of this protein family were to be studied, then a comparative analysis of the expression profiles during embryogenesis would first be needed. This would require isolating mouse isoform-specific cDNA probes for all the known isoforms for *in situ* hybridization analysis. A first approach to this would be to screen the fetal kidney library at low stringency with the ε isoform. As the 14-3-3 family are a discrete family with no other protein families that are similar, this approach should be successful and not prone to the cloning of non-14-3-3 proteins.

Several of the other isoforms cloned in different species are expressed in the adult kidney as shown by northern blotting (β and η [Watanabe *et al.*, 1991, 1993a]; θ [Nielsen, 1991]) and hence are likely to be expressed in fetal kidney. A quick method to identify each isoform would be to sequence small regions of the cDNA that correspond to the variable regions of the protein. RNA-PCR amplification of RNA samples from different tissues could then be used to clone any isoforms missed in this screen. With the full range of mouse 14-3-3 cDNAs, *in situ* hybridization analysis could be performed on embryos of varying ages to show how the expression of each isoform fluctuates during development.

Previously, antibodies to individual isoforms have been shown to cross react between different mammals (Ichimura *et al.*, 1991) and now antibodies that are specific to a single isoform (excepting α and β; and δ and ζ for which there is one antibody for each pair) have been produced (Martin *et al.*, 1993). These antibodies could be used for immune histochemistry on embryo sections to show where the proteins are presents in the developing embryo. The comparison of the two patterns will allow both confirmation of expression pattern and also would detect secondary regulation at the level of translation.

A further important question raised by other workers in the 14-3-3 field concerns the cellular localisation of the 14-3-3 family. The work by Lu *et al.*, (1992) showed that the *Arabidopsis* homologue was found in a DNA-binding complex suggesting that at least in *Arabidopsis* one isoform may be nuclear. The majority of other work has suggested that the protein is cytoplasmic (reviewed in Aitken *et al.*, 1992). This could be resolved by subcellular fractionation, making nuclear extracts, cytoplasmic extracts and membrane preparations and showing relative amounts of 14-3-3 protein in each fraction.

Another major question concerns the dimerisation of the 14-3-3 proteins. It has been shown that the proteins when run on native protein gels run at a molecular weight of 58 kDa, when the normal monomeric molecular weight is 29-32 kDa suggesting that the 14-3-3 proteins naturally act as dimers (Loeb & Gross, 1986).

Therefore, the next point to address is whether the proteins form hetero- or homodimers within the cells and whether dimerisation is necessary for protein activity. The use of antibodies to the protein family may answer this question. The antibodies could be used to immunoprecipitate the proteins in tissue extracts. Examination of the purified proteins would allow the question of which isoforms (if any) can form hetero-dimers to be answered. However, such experiments are dependant on there being experimental conditions that allow both dimerisation of the proteins and immune precipitation without disturbing the protein-protein interaction.

Another way to answer this question would be to create chimeric proteins with peptides added to facilitate purification (the glutathione S-transferase gene fusion system, Smith & Johnson, 1988; the maltose fusion protein system, Di Guan *et al.*, 1988; the His tagging system, Janknecht *et al.*, 1991). Such fusion proteins would allow specific isolation of a particular isoform. The use of such tagged proteins may provide a more sensitive method of finding both which isoforms dimerise and may even allow purification of the protein(s) that the 14-3-3 protein interacts with. The tagged 14-3-3 proteins could be expressed in either the usual *E. coli* systems, in mammalian cells or in yeast which allow proper modification of non-yeast 14-3-3 proteins (Collinge *et al.*, 1993). Co-expression of a tagged and non-tagged isoform would show if the non-tagged isoform was co-purified with the tagged protein. Systematic analysis of the nine (or possibly more) isoforms would then show which isoforms can dimerise and which ones can not.

It is also important to find the proteins with which the 14-3-3 isoforms interact. I have already discussed the idea of co-immune precipitation and the fact that this would depend on the conditions of protein interaction being compatible with the immune precipitation. Another approach could be performed in the yeast two hybrid system which allows the identification of interacting proteins (Field & Song, 1989; Chien *et al.*, 1991). This would be of immense help in clarifying the speculation on the real function on the 14-3-3 proteins.

The function of the ε isoform in development should also be studied more directly. The gene could be removed by homologous recombination in ES cells to look at the effect of deficiency of this gene in the whole organism. This can also be studied specifically in the kidney by the use of anti-sense oligonucleotides to the ε isoform in the *in vitro* kidney culture system discussed in section 1.6.6. This would answer questions about the functional redundancy of the ε protein within the 14-3-3 family. The effect of over-expression would also be of interest, as many of the genes involved in the signal transduction pathway have been shown to be oncogenic when over-expressed (*c-ret*, *c-met*, *ras*). It is conceivable that this gene may also have a oncogenic effect when over-expressed in the embryo. All these experiments would allow a greater understanding of how the ε isoform and the 14-3-3 family function during development.

REFERENCES

Abrahamson, D. R. (1985). Origin of the glomerular basement membrane visualized after *in vivo* labeling of laminin in new-born rat kidneys. *J. Cell Biol.* **100**, 1988-2000.

Adeniyi-Jones, S. & Zasloff, M. (1985). Transcription, processing and nuclear transport of a B1 *Alu* RNA species complementary to an intron of the murine alpha-fetoprotein gene. *Nature* **317**, 81-84.

Aitken, A., Collinge, D. B., van Heusden, B. P. H., Isobe, T., Roseboom, P. H., Rosenfeld, G. & Soll, S. (1992). 14-3-3 proteins: a highly conserved, widespread family of eukaryotic proteins. *Trends in Biochem. Sci.* **204,** 498-501.

Aitken, A., Ellis, C. A., Harris, A., Sellers, L. A. & Toker, A. (1990). Kinase and neurotransmitters. *Nature* **344**, 594.

Ali, S. M., Geisow, M. J. & Burgoyne, R. D. (1989). A role for calpactin in calcium-dependent exocytosis in adrenal chromaffin cells. *Nature* **340**, 313-315.

Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1-22.

Armstrong, J. F. (1992). The Wilms' tumour gene, *WT1*, in the development of the mouse embryo and its kidney. Ph.D. Thesis.University of Edinburgh.

Armstrong, J. F., Pritchard-Jones. K., Bickmore, W. A., Hastie, N. D & Bard, J. B. L. (1992). The expression of the Wilms' tumour gene, WT1, in the developing mammalian embryo. *Mech. Devel.* **40**, 85-97.

Aufderheide, E., Chiquet-Ehrismann, R., & Ekblom, P. (1987). Epithelial-mesenchymal interactions in the developing kidney lead to expression of tenascin in the mesenchyme. *J. Cell Biol.* **105**, 599-608.

Aviv, H. & Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA* **69,** 1408-1412.

Balling, R., Deutsch U. & Gruss, P. (1988). *undulated*, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of *Pax 1*. *Cell* **55**, 531-535.

Bastian, H. & Gruss. P. (1990). A murine even-skipped homologue, *Evx 1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9**,1839-1852.

Batzer, M. A., Kilroy, G. E., Richard, P. E., Shaikh, T. H., Desselle, T. D., Hoppens, C. L. & Deininger, P. L. (1990). Structure and variability of recently inserted *Alu* family members. *Nucleic Acids Res.* **18**, 6793-6798.

Bennett, K. L. Hill, R. E., Pietras, D. F., Woodworth-Gutai, M., Kane-Haas, C., Houston, J.M., Heath, J. K. & Hastie, N. D. (1984). Most highly dispersed DNA families in the mouse genome. *Mol. Cell. Biol.* **4**, 1561-1571.

Berton, J. P. (1965). Anatomie vasculaire de mésonéphros chez certains mammifères. I. Les mésonéphros de l'embryon de porc. *C. R. Assoc. Anat.* **124**, 272-290.

Bickmore, W. A., Oghene, K., Little, M. H., Seawright, A., van Heyningen, V. & Hastie, N. D. (1992). Modulation of DNA binding specificity by alternative splicing of the Wilms' tumor *wt1* gene transcript. *Science* **257**, 235-237.

Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.

Bopp, D., Burn, M., Baumgartner, S., Frigerio, G. & Noll, M. (1986). Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*. *Cell* **47** 1033-1040.

Boston, P. F., Jackson, P. & Thompson, R. J. (1982a). Human 14-3-3 protein radioimmunoassay, tissue distribution, and cerebrospinal fluid levels in patients with neurological disorders. *J. Neurochem.* **38**, 1466-1474.

Boston, P. F., Jackson, P., Kynoch, P. A. M. & Thompson, R. J. (1982b). Purification, properties, and immunohistochemical localisation of human brain 14-3-3 protein. *J. Neurochem.* **38**, 1475-1482.

Bottaro, D. P., Rubin, J. S., Faletto, D. L., Chan, A. M., Kmiecik, T. E. Vande Woude, G. F. & Aaronson, S. A. (1991). Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product. *Science* **251**, 802-804.

Bowes, C., Danciger, M., Kozak, C. A. & Farber, D. B. (1989). Isolation of a candidate cDNA for the gene causing retinal degeneration in the *rd* mouse. *Proc. Natl. Acad. Sci. USA* **86**, 9722-9726.

Brandt, J., Thordal-Christensen, H., Vad, K., Gregerson, P. L. & Collinge, D. B. (1992). A pathogen-induced gene of barley encodes a protein showing high similarity to a protein kinase regulator. *Plant J.* **2**, 815-820.

Burgoyne, R. D. & Morgan, A. (1993). Regulated exocytosis. *Biochem. J.* **293**, 305-316.

Burns, R. K. (1955). Urogenital system. In *Analysis of Development*, ed. B. J. Willier, P. Weiss & V. Hamburger, pp. 462-491. W. B. Saunders, Philadelphia.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C. & Housman, D. E. (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* **60**, 509-520.

Celis, J. E *et al.* (1990a). Comprehensive two-dimensional gel protein databases offer a global approach to the analysis of human cells: The transformed amnion cells (AMA) master database and its link to genome DNA sequence data. *Electrophoresis* **11**, 989-1071.

Celis, J. E. *et al.* (1990b). The MRC-5 human embryonal lung fibroblast two-dimensional gel cellular protein database: Quantitative identification of polypeptides whose relative abundance differs between quiescent, proliferating and SV40 transformed cells. *Electrophoresis* **11**, 1072-1113.

Chan, A., King, H., Deakin, E. A., Tempest, P. R., Hilkens, J., Kroezens, V., Edwards, D. R., Wills, A. J., Brookes, P. & Cooper, C. S. (1988). Characterization of the mouse met proto-oncogene. *Oncogene* **2**, 593-599.

Chen, Z., Friedrich, G. & Soriano, P (1992). Characterisation of mouse promoter trap mutants affecting embryo development. Abstract presented at the Mouse Molecular Genetics Meeting, Cold Spring Harbor.

Chien, C. T., Bartel, P. L., Sternglanz, R. & Field, S. (1991). The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. USA* **88**, 9578-9582.

Chiquet-Ehrismann, R., Maccie, E. J., Pearon, C. A., & Sakakura, T. (1986). Tenascin: an extracellular matrix protein involved in tissue interactions during fetal development and oncogenesis. *Cell* **47**, 131-139.

Chomczynski, P. & Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159.

Church, G. M. & Gilbert, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1991-1995.

Cole, M. D. (1986). The myc oncogene: its role in transformation and differentiation. *Annu. Rev. Genet.* **20**, 361-384.

Collinge, D. B., Brandt, J., Christiansen, J., Palmgren, M. G. & Thordal-Christensen, H. (1993). The 14-3-3 proteins of barley: function, molecular cloning and expression of their genes. Personal communication of poster abstract.

Cooper, J. R., Bloom, F. E. & Roth, R. H. (1986). *The biochemical basis of neuropharmacology*, 5th ed., pp 203-339. Oxford University Press, New York.

Courtois, G., Morgan, J. G., Campbell, L. A., Fourel, G. & Crabtree, G. R. (1987). Interaction of a liver-specific nuclear factor with the fibrinogen and α1-antitrypsin promoters. *Science* **238**, 688-692.

Creutz, C. E. (1992). The annexins and exocytosis. *Science* **258**, 924-931.

Crocker, J. F. S. (1973). Human embryonic kidneys in organ culture: abnormalities of development induced by decreased potassium. *Science* **181**, 1178-1179.

Croisille, Y. (1976). On some recent contributions to the study of kidney tubulogenesis in mammals and birds. In *Tests of teratogenecity* in vitro, ed. J. B. Ebert & M. Marois, pp. 149-170. North-Holland Publ. Co., Amsterdam.

Davidson, D., Graham, E., Sime, C. & Hill, R. (1988). A gene with sequence similarity to Drosophila *engrailed* is expressed during the development of the neural tube and vertebrae in the mouse. *Development* **104,** 305-316.

Davis, R. L., Weintraub, H. & Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000.

De Martino, C. & Zamboni, L. (1966). A morphological study of the mesonephros of the human embryo. *J. Ultrastr. Res.* **16**, 399-427.

Deschamps, J., de Laaf, R., Verrijzer, P., de Gouw, M., Destrée, O. & Meijlink, F. (1987). The mouse *Hox2.3* homeobox-containing gene: regulation and expression pattern in embryos. *Differentiation* **35**, 21-30.

De Simone, V., De Magistris, L., Lazzaro, D., Gerstener, J., Monaci, P., Nicosia, A. & Cortese, R. (1991) LFB3, a heterodimer-forming homeoprotein of the LFB1 family is expressed in specialised epithelia. *EMBO J.* **10**, 1435-1443.

Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387-190.

de Vetten, N. C., Lu, G. & Ferl, R. J. (1992). A maize protein associated with the G-Box binding complex has homology to brain regulatory proteins. *The Plant Cell* **4**, 1295-1307.

Di Guan, C., Li, P., Riggs, P.D. & Inouye, H. (1987). Vectors that facilitate the expression of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **67**, 21-30.

Dørup, J. & Maunsbach, A. B. (1982). The ultrastructural development of distal nephron segments in the human fetal kidney. *Anat. Embryol.* **164**, 19-41.

Dressler, G. R., Erby Wilkinson, J., Rothenpieler, U. W., Patterson, L. T., Williams-Simons, L. & Westphal, H. (1993). Deregulation of Pax-2 expression in transgenic mice generates severe kidney abnormalities. *Nature* **362**, 65-67.

Dressler, G. R. & Douglass, E. C. (1992). Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms' tumor. *Proc. Natl. Acad. Sci. USA* **89**, 1179-1183.

Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O. & Gruss, P. (1990). *Pax 2,* a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109**, 787-795.

Dressler, G. R., Steele-Perkins, V., Rothenpieler, U. W., Morris, J. F., Madden, S. L. & Rauscher, F. J. (1992). *Pax-2* is a candidate target gene for WT1 repression during normal kidney development. Abstract presented at the Mouse Molecular Genetics Meeting, Cold Spring Harbor.

Drummond, I. A., Madden, S. L., Bell, G. I., Rauscher, F. J. & Sukhatme, V. P. (1991). The *IGF2* gene is a target for repression by the Wilms' tumour gene product. *J. Cell Biol.* **115**, 222(A).

Drummond, I. A., Madden, S. L., Rohwer-Nutter, P., Bell, G. I., Sukhatme, V. P. & Rauscher, F. J. (1992). Repression of the insulin-like growth factor II gene by the Wilms' tumor suppressor WT1. *Science* **257**, 674-678.

Du Bois, A. M. (1969). The embryonic kidney. In *The Kidney*, ed. C. Rouiller & A. F. Muller, vol. I, pp. 1-59. Academic Press, New York.

Duboule, D. & Dolle, P. (1989). The structural and functional organisation of the murine *Hox* gene family resembles that of the *Drosophila* homeotic genes. *EMBO J.* **5** 1497-1505.

Duguid, J. R. & Dinauer, M. C. (1990). Library subtraction of "*in vitro*" cDNA libraries to identify differentially expressed genes in scrapie infection. *Nucleic Acids Res.* **18**, 2789-2792.

Dunn, L., Glueckson-Schoenheimer, S. & Bryson, V (1940). A new mutation in the mouse. *J. Hered.* **31**, 343-348.

Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J. L. & Gruss, P. (1988). A mouse gene homologous to the *Drosophila* gene *caudal* is expressed in epithelial cells from the embryonic intestine. *Genes & Dev.* **2,** 1647-1654.

Dworkin, M. B., Dawid, I. B. (1980a). Construction of a cloned library of expressed embryonic gene sequences from *Xenopus laevis*. *Dev. Biol.* **76**, 435-448.

Dworkin, M. B. & Dawid, I. B. (1980b). Use of a cloned library for the study of abundant poly $(A)^+$ RNA during *Xenopus laevis* development. *Dev. Biol.* **76**, 449-464.

Eccles, M. R., Wallis, L. J., Fidler, A. E., Spurr, N. K., Goodfellow, P. J. & Reeve, A. E. (1992). Expression of the *PAX2* gene in human fetal kidney and Wilms' tumor. *Cell Growth & Differen.* **3**, 279-289.

Egan, S. E. & Weinberg, R. A. (1993). The pathway to signal achievement. *Nature* **365**, 781-783.

Ekblom, M., Klein, G., Mugrauer, G., Fecker, L., Deutzmann, R., Timpl, R. & Ekblom, P. (1990). Transient and locally restricted expression of laminin A chain mRNA by developing epithelial cells during kidney organogenesis. *Cell* **60**, 337-346.

Ekblom, P. (1981). Formation of basement membranes in the embryonic kidney: An immunohistological study. *J. Cell Biol.* **91**, 1-10.

Ekblom, P. (1984). Basement membrane proteins and growth factors in kidney differentiation. In *Role of Extracellular Matrix in Development*, ed. R. L. Trelstad, pp.173-206. Alan R. Liss, New York.

Ekblom, P. (1992). Renal development. In *The Kidney Physiology and Pathophysiology*, ed. D. W. Seldin & G. Giebisch, pp. 475-501. Raven Press, New York.

Ekblom, P., Alitalo, K., Vaheri, A., Timpl, R. & Saxén, L. (1980). Induction of a basement membrane glycoprotein in embryonic kidney: possible role of laminin in morphogenesis. *Proc. Natl. Acad. Sci. USA.* **77**, 485-489.

Ekblom, P., Lehtonen, E., Saxén, L. & Timpl, R. (1981a). Shift in collagen type as an early response to induction of the metanephric mesenchyme. *J. Cell Biol.* **89**, 276-283.

Ekblom, P., Miettinen, A., Virtanen, I., Wahlstrom, T., Dawnay, A. & Saxén, L. (1981b). *In vitro* segregation of the metanephric nephron. *Dev. Biol.* **84**, 88-95.

Ekblom, P., Sariola, H., Karkinen-Jääskeläinen, M. & Saxén, L. (1982). The origin of the glomerular endothelium. *Cell Diff.* **11**, 35-39.

Epstein, D. J., Vogan, K. J., Trasler, D. G. & Gruss, P. (1993). A mutation within intron 3 of the Pax-3 gene produces aberrantly spliced mRNA transcripts in the *splotch* (Sp) mouse mutant. *Proc. Natl. Acad. Sci. USA* **90**, 532-536.

Erickson, P. F. & Moore, B. W. (1980). Investigation of the axonal transport of three acidic soluble proteins (14-3-2, 14-3-3, and S-100) in the rabbit visual system. *J. Neurochem.* **35**, 232-241.

Erselius, J. R., Goulding, M. D. & Gruss, P. (1990). Structure and expression pattern of the murine *Hox-3.2* gene. *Development* **110**, 629-642.

Feinberg, A. P. & Vogelstein, B. (1984). Addenum. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**, 266-267.

Field, S. & Song, O. K. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.

Foley, K. P., Leonard, M. W. & Engel, J. D. (1993) Quantification of RNA using the polymerase chain reaction. *Trends Genet.* **9**, 380-385.

Fox, H. (1956). Compensation in the remaining pronephros of *Triturus* after unilateral pronephroctomy. *J. Embryol. Exp. Morphol.* **4**, 139-151.

Freeman, R. M., Plutzky J. & Neel, B. G. (1992). Identification of a human src homology 2-containing protein-tyrosine-phosphatase: a putative homolog of *Drosophila corkscrew*. *Proc. Natl. Acad. Sci. USA* **89**, 11239-11243.

Friedrich, G., McGregor, G., Overbeek, P. & Soriano, P. (1992). New mutations affecting early mouse development and male germ cell morphology. Abstract presented at the Mouse Molecular Genetics Meeting, Cold Spring Harbor.

Friedrich, G. & Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes & Dev.* **5**, 1513-1523.

Fu, H., Coburn, J. & Collier, R. J. (1993). The eukaryotic host factor that activates exoenzyme S from *Pseudomonas aeruginosa*, is a member of the 14-3-3 protein family. *Proc. Natl. Acad. Sci. USA.* **90**, 2320-2324.

Furukawa, Y., Ikuta, N., Omata, S., Yaumauchi, T., Isobe, T. & Ichimura, T. (1993). Demonstration of the phosphorylation-dependent interaction of tryptophan hydroxylase with the 14-3-3 protein. *Biochem. Biophys. Res. Comm.* **194**, 144-149.

Gashler, A. L., Bonthron, D. T., Madden, S. L., Rauscher, F. J., III, Collins, T. & Sukhatme, V. P. (1992). Human platelet-derived growth factor A chain is transcriptionally repressed by the Wilms' tumor suppressor WT1. *Proc. Natl. Acad. Sci. USA* **89**, 10984-10988.

Gilardi-Hebenstreit, P., Nieto, M. A., Frain, M., Mattei, D. G., Chestier, A., Wilkinson, D. G. & Charnay, P. (1992). An *eph*-related receptor protein tyrosine kinase segmentally expressed in the developing mouse hindbrain. *Oncogene* **7**, 2499-2506.

Glueckson-Schoenheimer, S (1945). The embryonic development of mutants of the *Sd* strain in mice. *Genetics* **30**, 29-38.

Goulding, M., Sterrer, S., Fleming, J., Balling, R., Nadeau, J., Moore, K. J., Brown, S. D. M., Steel, K. P. & Gruss, P. (1993). Analysis of the *Pax-3* gene in the mouse mutant *splotch*. *Genomics* **17**, 355-363.

Grobstein, C. (1953a). Morphogenetic interaction between embryonic mouse tissues seperated by a membrane filter. *Nature* **172**, 869-871.

Grobstein, C. (1953b). Inductive epithelio-mesenchymal interaction in cultured organ rudiments of the mouse. *Science* **118**, 52-55.

Grobstein, C. (1955). Inductive interaction in the development of the mouse metanephros. *J. Exp. Zool.* **130**, 319-340.

Grobstein, C. (1956). Trans-filter induction of tubules in mouse metanephrogenic mesenchyme. *Exp. Cell Res.* **10**, 424-440.

Gubler, U. & Hoffman, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.

Gurdon, J. B. (1987). Embryonic induction - molecular prospects. *Development* **99**, 285-306.

Haber, D. A., Buckler, A. J., Glaser, T., Call, K. M., Pelletier, J., Sohn, R. L., Douglass, E. C. & Housman, D. E. (1990). An internal deletion within an 11p13 Zinc finger gene contributes to the development of Wilms' tumor. *Cell* **61**, 1257-1269.

Haber, D. A., Sohn, R. L., Buckler, A. J., Pelletier, J., Call, K. M., & Housman, D. E. (1991). Alternative splicing and genomic structure of the Wilms tumor gene, *WT1*. *Proc. Natl. Acad. Sci.* **88**, 9618-9622.

Hafen, E., Basler, K., Edstroem, J. E. & Rubin, G. M. (1987). *Sevenless*, a cell-specific homeotic gene of *Drosophila*, encodes a putative transmembrane receptor with a tyrosine kinase domain. *Science* **236**, 55-63.

Hammermann, M. R., O'Shea, M. & Miller, S. B. (1993). Roleof growth factors in the regulation of renal growth. *Annu. Rev. Physiol.* **55**, 305-321.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557-580.

Harding, K., Wedeen, C., McGinnis, W. & Levine, M. (1985). Spatially regulated expression of homeotic genes in *Drosophila*. *Science* **229**, 1236-1242.

Hart, A. C., Krämer, H., Van Vactor, D. L., Jr., Paidhungat, M. & Zipursky, S. L. (1990). Induction of cell fate in the *Drosophila* retina: the bride of sevenless protein is predicted to contain a large extracellular domain and seven transmembrane segments. *Genes & Dev.* **4**, 1835-1847.

Hastie, N. D. & Bishop, J. O. (1976). The expression of three abundant classes of messenger RNA in mouse tissues. *Cell* **9**, 761-774.

Hay E. D. (1984). Cell-matrix interaction in the embryo: cell shape, cell surface, cell skeletons, and their role in differentiation. In *Role of Extracellular Matrix in Development*, ed. R. L. Trelstad, pp. 173-206. Alan R. Liss, New York.

Hill, R. E., Jones, P. F., Sime, C. M., Justice, M. J., Copeland, N. G., Jenkins, N. A., Graham, E. & Davidson, D. R. (1989). A new family of mouse homeo-box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox 7.1. *Genes and Dev.* **3**, 26-37.

Hill. R. E., Favor, J., Hogan, B. L. M., Ton, C. C. T., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan T., Hastie N. D. & Van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**, 522-525.

Hirsch, S., Aitken, A., Bertsch, U. & Soll, J. (1992). A plant homologue to mammalian brain 14-3-3 protein and protein kinase C inhibitor. *FEBS Lett.* **302**, 222-224.

Hirvonen, H., Sandberg, M., Kalimo, H., Hukkanen, V., Vuorio, E., Salmi, T. T. & Alitalo, K. (1989). The *N-myc* proto-oncogene and IGF-II growth factor mRNAs are expressed by distinct cells in human fetal kidney and brain. *J. Cell Biol.* **105**, 1093-1104.

Hla, T. & Maciag, T. (1990). Isolation of immediate-early differentiation mRNAs by enzymatic amplification of subtracted cDNA from human endothelial cells. *Biochem. Biophys. Res. Comm.* **167**, 637-643.

Hoey, T. & Levine, M. (1988). Homeobox proteins as sequence-specific transcription factors. *Nature* **322**, 858-861.

Holland, P. W. H. & Hogan, B. L. M. (1988a). Expression of homeo box genes during mouse development: a review. *Genes & Dev.* **2**, 773-782.

Holland, P. W. H. & Hogan, B. L. M. (1988b). Spatially restricted patterns of expression of the homeobox-containing gene *Hox 2.1* during mouse embryogenesis. *Development* **102**, 159-174.

Holtfreter, J. (1944). Experimental studies on the development of the pronephros. *Rev. Can. Biol.* **3,** 220-250.

Hromas, R., Moore, J., Johnston, T., Socha, C. & Klemsz, M. (1993). *Drosophila forkhead* homologues are expressed in a lineage-restricted manner in human hematopoietic cells. *Blood* **81**, 2854-2859.

Ichimura, T., Isobe, T., Okuyama, T., Takahashi, N., Araki, K., Kuwano, R., Takahashi, Y. (1988). Molecular cloning of cDNA coding for brain-specific 14-3-3 protein, a protein kinase-dependent activator of tyrosine and tryptophan hydroxylases. *Proc. Natl. Acad. Sci. USA* **85**, 7084-7088.

Ichimura, T., Isobe, T., Okuyama, T., Yamauchi, T. & Fujisawa, H. (1987). Brain 14-3-3 protein is an activator protein that activates tryptophan 5-monooxygenase and tyrosine 3-monooxygenase in the presence of Ca^{2+}, calmodulin-dependent protein kinase II. *FEBS Lett.* **219**, 79-82.

Ichimura, T., Sugano, H., Kuwano, R., Sunaya, T., Okuyama, T. & Isobe, T. (1991). Widespread distribution of the 14-3-3 protein in vertebrate brains and bovine tissues: Correlation with the distribution of calcium-dependent protein kinases. *J. Neurochem.* **56**, 1449-1451.

Ichimura-Ohshima, Y., Morii, K., Ichimura, T., Araki, K., Takahashi, Y., Isobe, T., Minoshima, S., Fukuyama, R., Shimizu, N. & Kuwano, R. (1992). cDNA cloning and chromosome assignment of the gene for human brain 14-3-3 protein η chain. *J. Neurosci. Res.* **31**, 600-605.

Isobe, T., Hiyane, Y., Ichimura, T., Okuyama, T., Takahashi, N., Nakajo, S. & Nakaya, K. (1992). Activation of protein kinase C by the 14-3-3 proteins homologous with Exo1 protein stimulates calcium-dependent exocytosis. *FEBS Lett.* **308**, 121-124.

Isobe, T., Ichimura, T., Sunaya, T., Okuyama, T., Takahashi, N., Kuwano, R. & Takahashi, Y. (1991). Distinct forms of the protein kinase-dependent activator of tyrosine and tryptophan hydroxylases. *J. Mol. Biol.* **217**, 125-132.

Jackson, I. J., Scofield, P. & Hogan, B. L. M. (1985). A mouse homeo box gene is expressed during embryogenesis and in adult kidney. *Nature* **317**, 745-748.

Janknecht, R., De Martynoff,G., Lou, J., Hipskind, R. A., Nordheim, A., Stunnenberg,H. G. (1991). Rapid and efficient purification of native histidine-tagged protein expressed by recombinent vaccinia. *Proc. Natl. Acad. Sci. USA* **88**, 8972-8976.

Jokelainen, P. (1963). An electron microscopic study of the development of the rat metanephric nephron. *Acta Anat. Suppl.* **47**, 1-73.

Jones, K. W., Shapero, M. H., Chevrette, M. & Fourier, R. E. K. (1991). Subtractive hybridization cloning of a tissue-specific extinguisher: TSE1 encodes a regulatory subunit of protein kinase A. *Cell*, **66**, 861-872.

Joyner, A. & Martin G. (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Genes & Dev.* **1**, 29-38.

Katayama, C. (1990). Single-strand DNA sequencing: A simplified protocol for single-stranded DNA rescue. *Strategies in Mol. Biol.* **4,** 56-58.

Kaufman, M. H. (1992). *The Atlas of Mouse Development.* Academic press, London.

Kawasaki, E. S. (1990). Amplification of RNA. In *PCR protocols: A guide to methods and applications*. Academic press, New York.

Kidou, S. I., Umedi, M., Kato, A. & Uchimiya, H. (1993). Isolation and characterization of a rice cDNA similar to the bovine brain-specific 14-3-3 protein gene. *Plant Mol. Biol.* **21**, 191-194.

Klar, A., Baldassare, M. & Jessell, T. M. (1992). F-Spondin a gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell-adhesion and neurite extension. *Cell* **69**, 95-110.

Klein, G., Lanlegger, M., Goridas, C. & Ekblom, P. (1988a). Neural cell adhesion molecules during embryonic induction and development of the kidney. *Development* **102**, 749-761.

Klein, G., Lanlegger, M., Timpl, R. & Ekblom, P. (1988b).Role of laminin A chain in the development of epithelial cell polarity. *Cell* **55**, 331-341.

Ko, M. S. H., Ko, S. B. H., Takahashi, N., Nishiguchi, K. & Abe, K. (1990). Unbiased amplification of a highly complex mixture of DNA fragments by 'lone linker'-tagged PCR. *Nucleic Acids Res.* **18**, 4293-4294.

Korn, R., Schoor, M., Neuhaus, H., Henseling, U., Soininen, R., Zachgo, J. & Gossler, A. (1992). Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaeric embryos with a high frequency and detect endogenous genes. *Mech. Dev.* **39**, 95-109.

Koseki, H., Zachgo, J., Mizutani, Y., Simon-Chazottes, D., Guenet, J. L., Balling, R. & Gossler, A (1993). Fine genetic mapping of the proximal part of mouse chromosome 2 excludes Pax 8 as a candidate gene for Danforth's short tail (Sd). *Mamm. Genome* **4**, 324-327.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNA. *Nucleic Acids Res.*, **12**, 857-872.

Kozeki, C., Herzlinger, D. & Al-Awqati, Q. (1992). Apoptosis in metanephric development. *J. Cell Biol.* **5**, 1327-1333.

Kramerov, D. A., Lekakh, I. V., Samarina, O. P. & Ryskov, A. P. (1982). The sequences homologous to major interspersed repeats B1 and B2 of mouse genome are present in mRNA and small cytoplasmic poly(A)$^+$ RNA. *Nucleic Acids Res.* **10**, 7477-7491.

Kramerov, D. A., Ryskov, A. P. & Georgiev, G. P. (1977). The structural organization of nuclear pre-mRNA. II Very long double-stranded structures in nuclear pre-mRNA. *Biochim. Biophys. Acta.* **475**, 461-475.

Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A. & Georgiev, G. P. (1980). The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. *Nucleic Acids Res.* **6**, 1201-1215.

Kress, C., Vogels, R., de Graaf, W., Bonnerot, C., Meijlink, F., Nicholas J. F. & Deschamps, J. (1990). Hox 2.3 upstream sequences mediate *lacZ* expression in intermediate mesoderm derivatives of transgenic mice. *Development* **109**, 775-786.

Kriedberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D. & Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell* **74**, 679-691.

Kuziora, M. A. & McGinnis, W. (1989). A homeodomain substitution changes the regulatory specificity of the *deformed* protein in *Drosophila* embryos. *Cell* **59**, 563-571.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.

Lazzaro, D., de Simone, V., de Magistris, L., Lehtonen, E. & Cortese, R. (1992). LFB1 and LFB3 are sequentially expressed during kidney development. *Development* **114**, 469-479.

Le Douarin, N. (1973). A biological cell labelling technique and its use in experimental embryology. *Dev. Biol.* **30**, 217-222.

Leffers, H., Madsen, P., Resmussen, H. H., Honoré, B., Andersen, A. H., Walbum, E., Vandekerckhove, J. & Celis, J. E. (1993). Molecular cloning and expresssion of the transformation sensitive epithelial marker Stratifin. *J. Mol. Biol.* **231**, 982-998.

Lehtonen, E., Virtanen, I. & Saxén, L. (1985). Reorganization of intermediate filament cytoskeleton in induced metanephric mesenchyme cells is independent of tubule morphogenesis. *Dev. Biol.* **108**, 481-490.

Lewis E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570.

Loeb, L. A. & Gross, R. W. (1986). Identification and purification of sheep platelet phospholipase A_2 isoforms. *J. Biol. Chem.* **261**, 10467-10470.

Loros, J. J.; Denome, S. A. & Dunlap, J. C. (1989). Molecular cloning of genes under control of the circadian clock in *Neurospora*. *Science* **243**, 385-388.

Lu, G., De Lisle, A. J., De Vetten, N. C. & Ferl, R. J. (1992). Brain proteins in plants: An *Arabidopsis* homolog to neurotransmitter pathway activators is part of a DNA binding complex. *Proc. Natl. Acad. Sci. USA* **89**, 11490-11494.

MacKenzie, A., Ferguson, M. W. & Sharpe, P. T. (1991a) Hox-7 expression during murine craniofacial development. *Development* **113**, 601-611.

Mackenzie, A., Leeming, G. L., Jowett, A. K., Ferguson, M. W. & Sharpe, P. T. (1991b) The homeobox gene *Hox 7.1* has specific regional and temporal expression patterns during early murine craniofacial embryogenesis, especially tooth development *in vivo* and *in vitro*. *Development* **111**, 269-85.

Madden, S. L., Cook, D. M., Morris, J. F., Gashler, A., Sukhatme, V. P. & Rauscher, F. J. (1991). Transcriptional repression mediated by the WT1 Wilms' tumor gene product. *Science* **253**, 1550-1553.

Maheswaran, S., Park, S., Bernard, A., Morris, J. F., Rauscher, F. J., III, Hill, D. E. & Haber, D. A. (1993). Physical and functional interaction between WT1 and p53 proteins. *Proc. Natl. Acad. Sci. USA* **90**, 5100-5104.

Maraia, R. J. (1991). The subset of mouse B1 (*Alu*-equivalent) sequences expressed as small processed cytoplasmic transcripts. *Nucleic Acids Res.* **19**, 5695-5702.

Marin-Padilla, M. (1964). The mesonephric-testicular connection man and some mammals. *Anat. Rec.* **148**, 1-14.

Marten, G. J. M., Piosik, P. A. & Danen, E. H. J. (1992). Evolutionary conservation of the 14-3-3 protein. *Biochem. Biophys. Res. Comm.* **184**, 1456-1459.

Martin, H., Patel, Y., Jones, D., Howell, S., Robinson, K. & Aitken, A. (1993). Antibodies against the major brain isoforms of 14-3-3 protein. an antibody specific for the N-acetylated amino-terminus of a protein. *FEBS Letts.* **331**, 296-303.

McConnell, J. E. & Hodges, P. E. (1993). The alternative 5'-end of the *Drosophila melanogaster* epidermal growth factor receptor cDNA (*DER)* is part of the *D14-3-3* cDNA. *Gene*, **126,** 293-294.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* **308**, 428-433.

Mendel, D. B., Hansen, L. P., Graves, M. K., Conley, P. B. & Crabtree, G. R. (1991). HNF-1α and HNF-1β (vHNF-1) share dimerization and homeo domains, but not activation domains, and form heterodimers *in vitro*. *Genes & Dev.* **5**, 1042-1056.

Miller, J., McLachlan, A. D. & Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIa from *Xenopus* oocytes. *EMBO J.* **4**, 1609-1614

Minty, A. J., Caravatti, M., Robert, B., Cohen, A., Daubas, P., Gros, F. & Buckingham, M. E. (1981). Mouse actin messenger RNAs. *J. Biol. Chem.* **256**, 1008-1014.

Moore, B. W. & Perez, V. J. (1967). Physiological and Biochemical Aspects of Nervous Integration. ed. F. D. Carlson, pp. 343-359. Prentice-Hall, Englewood Cliffs, N.J.

Morgan, A. & Burgoyne, R. D. (1992). Exo1 and Exo2 proteins stimulate calcium-dependent exocytosis in permeabilized adrenal chromaffin cells. *Nature* **355**, 833-836.

Morgan, A., Roth, D., Martin, H., Aitken, A. & Burgoyne, R.D. (1993a). Identification of cytosolic protein regulators of exocytosis. *Biochem. Soc. Transactions*, **21**, 401-405.

Morgan, A., Wilkinson, M. & Burgoyne, R.D. (1993b). Identification of Exo1 as the catalytic subunit of protein kinase A reveals a role for cyclic AMP in Ca2+-dependent exocytosis in chromaffin cells. *EMBO J.* **12**, 3747-3752.

Mosquera, L., Forristall, C., Zhou, Y. & King, M. L. (1993). Messenger RNA localized to the vegetal cortex of *Xenopus* oocytes encodes a protein with a *nanos*-like zinc finger domain. *Development* **117**, 377-386.

Mugrauer, G., Alt, F. W. & Ekblom, P. (1988). *N-myc* proto-oncogene expression during organogenesis in the developing mouse as revealed by *in situ* hybridization. *J. Cell Biol.* **107**, 1325-1335.

Mugrauer, G., & Ekblom, P. (1991). Contrasting expression patterns of three members of the *myc* family of protooncogenes in the developing and adult mouse kidney. *J. Cell Biol.* **112**, 13-25.

Mulligan, L. M., Kwok, J. B., Healey, C. S., Elsdon, M. J., Eng, C., Gardner, D., Love, R., Mole, S. E., Moore, J. K., Papi, L., Ponder, M. A., Telenius, H., Tunnacliffe, A. & Ponder, B. A. J. (1993). Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* **363**, 458-60.

Murphy, D., Brickel, P. M., Latchman, D. J., Willinson, K. & Rigby, P. W. J. (1983). Transcripts regulated during normal embryonic development and oncogenic transformation share a repetitive element. *Cell* **35**, 865-871.

Nairn, A. C., Hemmings, H. C. Jr. & Greengard, P. (1985). Protein kinases in the brain. *Annu. Rev. Biochem.* **54**, 931-976.

Naldini, L., Vigna, E., Narsimhan, R., Gaudino, G., Zarnegar, R., Michalopoulos, G. & Comoglio, P. M. (1991a) Hepatocyte growth factor (HGF) stimulates the tyrosine kinase activity of the receptor encoded by the proto-oncogen *c-met*. *Oncogene* **6**, 501-504.

Naldini, L., Weidner, K. M., Vigna, E., Gaudino, G., Bardelli, A., Ponzetto, R. P., Narsimhan, R., Hartmann, G., Zarnegar, R., Michalopoulos, G., Birchmeier, R. P. & Comoglio, P. M. (1991b) Scatter factor and hepatocyte growth factor are indistinguishable ligands for the MET receptor. *EMBO J.* **10**, 2867-2878

Neiss, W. F. (1982). Morphogenesis and histogenesis of the connecting tubule in the rat kidney. *Anat. Embryol.* **165**, 81-95.

Nielsen, P.J. (1991). Primary structure of a human protein kinase regulator protein. *Biochim. Biophys. Acta* **1088**, 425-428.

Nisen, P. D., Zimmerman, K. A., Cotter, S. V., Gilbert, F., & Frederick, W. A. (1986). Enhanced expression of the *N-myc* gene in Wilms' tumors. *Cancer Res.* **46**, 6217-6222.

O'Kane, C. J. & Gehring, W. J. (1987). Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA*. **84**, 9123-9127.

Olivier, J. P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., Schlessinger, J., Hafen, E. & Pawson, T. (1993) A *Drosophila* SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of ras guanidine nucleotide exchange, Sos. *Cell* **73**, 179-191.

Osathanondh, V. & Potter, E. L. (1963). Development of human kidney as shown by microdissection III. Formation and interrelationship of collecting tubules and nephrons. *Arch. Path.* **76**, 290-302.

Ott, M.-E., Rey-Campos, J., Cereghini, S. & Yaniv, M. (1991). vHNF1 is expressed in epithelial cells of distinct embryonic origin during development and precedes HNF1 expression. *Mech. Devel.* **36**, 47-58.

Overton, J. (1959).Studies on the mode of outgrowth of the amphibian pronephric duct. *J. Embryol. Exp. Morphol.* **7**, 86-93.

Pachnis, V., Mankoo, B. & Constantini, F. (1993). Expression of the *c-ret* protooncogene during mouse embryogenesis. *Development*, 119, 1005-1017.

Pawson, T. & Bernstein, A. (1990). Receptor tyrosine kinases: genetic evidence for their role in *Drosophila* and mouse development. *Trends Genet.* **6**, 350-356.

Pelletier, J., Bruening, W., Kashtan, C. E., Mauer, S. M., Manivel, J. C., Striegel, J. E., Houghton, D. C., Junien, C., Habib, R., Fouser, L., Fine, R. N., Silverman, B. L., Haber, D. A. & Housman, D. (1991). Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* **67**, 437-447.

Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J.-L. & Gruss, P. (1990). *Pax 8*, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* **110**, 643-651.

Poole, T. J. & Steinberg, M. S. (1981). Amphibian pronephric duct morphogenesis: segregation, cell rearrangement and directed migration of the *Ambystoma* duct rudiment. *J. Embryol. Exp. Morphol.* **63**, 1-16.

Poole, T. J. & Steinberg, M. S. (1982). Evidence for the guidance of pronephric duct migration by a craniocaudally traveling adhesion gradient. *Dev. Biol.* **92**, 144-158.

Poole, T. J. & Steinberg, M. S. (1984). Different modes of pronephric duct origin among vertebrates. *Scanning Electron Microscopy* **1**, 475-482.

Potter, E. L. (1965). Development of the human glomerulus. *Arch. Path.* **80**, 241-255.

Prasad, G. L., Valverius, E. V., McDuffie, E. & Cooper, H. L. (1992). Complementary DNA cloning of a novel epithelial cell marker protein, HME1, that may be down-regulated in neoplastic mammary cells. *Cell Growth Diff.* **3**, 507-513.

Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., Housman, D., van Heyningen, V. & Hastie, N. (1990). The candidate Wilms' tumour gene is involved in genitourinary development. *Nature* **346**, 194-197.

Proudfoot, N. J. (1991). Poly (A) signals. *Cell* **64**, 671-674.

Proudfoot, N. J. & Brownlee, G. G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**, 211-214.

Quentin, Y. (1989). Successive waves of fixation of B1 variants in rodent lineage history. *J. Mol. Evol.* **28**, 299-305.

Rauscher, F. J., Morris, J. F., Tournay, O. E., Cook, D. M. & Curran, T. (1990). Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. *Science* **250**, 1259-1262.

Rey-Campos, J., Chouard, T., Yaniv, M. & Cereghini, S. (1991). vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. *EMBO J.* **10**, 1445-1457.

Risau, W. & Ekblom, P. (1986). Production of a heparin-binding angiogenesis factor by the embryonic kidney. *J. Cell Biol.* **103**, 1101-1107.

Robert, B., Sasson, D., Jacq, B., Gehring, W. & Buckingham, M. (1989). Hox-7, a mouse homeobox containing gene with a novel pattern of expression during embryogenesis. *EMBO J.* **8**, 91-100.

Rogers, S. A., Ryan, G. & Hammerman, M. R. (1991). Insulin-like growth factors I and II are produced in the metanephros and required for growth and development *in vitro*. *J. Cell Biol.* **113**, 1447-1453.

Rose, E. A., Glaser, T., Jones, C., Smith, C. L., Lewis, W. H., Call, K. M., Minden, M., Champagne, E., Bonetta, G., Yeger, H. & Housman, D. (1990). Complete physical map of the WAGR region of 11p13 localizes a candidate Wilms' tumor gene. *Cell* **60**, 495-508.

Roseboom, P. H., Weller, J. L., Namboodiri, M. A. A., Toker, A., Aitken, A. & Klein, D. C. (1992). 14-3-3 proteins: Isolation and cloning and distribution of the ε isoform. (Abstract) *FASEB J.* **6**, A1516.

Roseboom, P. H., Weller, J. L., Babila, T., Aitken, A., Sellers, L. A., Moffet, J. R., Namboodiri, M. A. A. & Klein, D. C. (1993). 14-3-3 epsilon and zeta isoforms: Cloning, tissue distribution and developmental expression. Unpublished as cited in Genbank accesion numbers M84416 and L07914.

Roth, D., Morgan, A. & Burgoyne, R. D. (1993). Identification of a key domain in annexin and 14-3-3 proteins that stimulate calcium-dependent exocytosis in permeabilized adrenal chromaffin cells. *FEBS Letts.* **320**, 207-210.

Rothenpieler, U. W. & Dressler, G. R. (1993). Pax-2 is required for mesenchyme to epithelium conversion during early kidney development. *Development* **119**, 711-720.

Rothenpieler, U. W. & Dressler, G. R. (1992). Mesenchyme to epithelium conversion inhibited by antisense oligonucleotides directed against Pax-2. Abstract presented at the Mouse Molecular Genetics Meeting, Cold Spring Harbor.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor University Press, Cold Spring Harbor.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. **74**, 5463-5467.

Sariola, H., Ekblom, P. & Henke-Fahle, S. (1989). Embryonic neurons as *in vitro* inducers of differentiation of nephrogenic mesenchyme. *Dev. Biol.* **132**, 271-281.

Sariola, H., Saarma, M., Sainio, K., Arumae, U., Palgi, J., Vaahtokari, A., Thesleff, I., & Karavanov, A. (1991). Dependence of kidney morphogenesis on the expression of nerve growth factor receptor. *Science* **254**, 571-573.

Sariola, H., Timpl, R., Vondermark, K., Mayne, R., Fitch, J. M., Lisenmayer, T. F. & Ekblom, P. (1984). Dual origin of the glomerular basement membrane. *Dev. Biol.* **101**, 86-96.

Sargent, T. D. & Dawid, I. B. (1983). Differential gene expression in the gastrula of *Xenopus laevis*. *Science* **222**, 133-139.

Saxén, L. (1970). Failure to demonstrate tubule induction in a heterologous mesenchyme. *Dev. Biol.* **23**, 511-523.

Saxén, L. (1977). Directive versus permissive induction: a working hypothesis. In *Cell and Tissue Interaction*, ed. J. W. Lash & M. M. Burger, pp. 1-9. Raven Press, New York.

Saxén, L. (1983). *In vitro* model-systems for chemical teratogenesis. In *In Vitro Toxicity Testing of Environmental Agents*, part B, ed. A. Kolber, T. K. Wong, L. D. Grant, R. S. Dewoskin & T. J. Hughes, pp. 155-163. Pontificia Academia Scientiarum, Città del Vaticano.

Saxén, L. (1987). *Organogenesis of the Kidney*, Cambridge University Press, Cambridge.

Saxén, L. & Karkinen-Jääskeläinen, M. (1975). Inductive interactions in morphogenesis. In *The Early Development of Mammals*, ed. M. Balls & A. Wild, pp. 319-334. Cambridge University Press, Cambridge.

Saxén, L., Toivonen, S., Vanio, T. & Korhonen, P. (1965). Untersuchungen über die Tubulogenese der niere. III Die analyse der frühentwicklung mit der zeitraffermethode. *Z. Naturforsch.* **206**, 340-343

Schejter, E. D., Segal, D., Glazer, L., Shilo, B. -Z. (1986). Alternative 5' exons and tissue-specific expression of the *Drosophila* EGF receptor homolog transcripts. *Cell* **46**, 1091-1101.

Schejter, E. D. & Shilo, B.-Z. (1989). The *Drosophila* EGF receptor homolog (DER) gene is allelic to faint little ball, a locus essential for embryonic development. *Cell* **56**, 1093-1104.

Schnabel, E., Anderson, J. M., & Farquhar, M. G. (1990). The tight junction protein ZO-1 is concentrated along slit diaphragms of the glomerular epithelium. *J. Cell Biol.* **111**, 1255-1263.

Schuchardt, A., D'Agati, V., Constantini, F. & Pachnis, V. (1992). A targeted mutation in the *c-ret* receptor tyrosine kinase gene interferes with normal development of the peripheral nervous system and the kidney. Abstract presented at the Mouse Molecular Genetics Meeting, Cold Spring Harbor.

Schuchardt, A., D'Agati, V., Larssonblombeerg, C., Constantini, F. & Pachnis, V. (1994) Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor ret. *Nature*, **367**, 380-383.

Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988). λZAP: A bacteriophage λ expression vector with *in vivo* excision properties. *Nucleic Acids Res.*, **16**, 7583-7600.

Silverman, H. (1969). Uber die Entwicklung der epithelplatten in den *Corpuscula renalia* der menschlichen uriniere. *Acta Anat.* **74**, 36-43.

Simeone, A., Acampora, D., Gulisano, M., Stornaiuolo, A. & Boncinelli, E. (1992a). Nested expression domains of four homeobox genes in developing rostral brain *Nature* **358**, 687-690.

Simeone, A., Acampora, D., Mallamaci, A., Stornaiuolo, A., Rosario D'Apice, M., Nigro & Boncinelli, E. (199). A vertebrate gene related to orthodenticle contains a homeodomain of the bicoid class and demarcates anterior neuroectoderm in the gastrulating mouse embryo. *EMBO J.* **12**, 2735-2747.

Simeone, A., Gulisano, M., Acampora, D., Stornaiuolo, A., Rambaldi, M. & Boncinelli, E. (1992b). Two vertebrate homeobox genes related to the *Drosophila empty spiracles* gene are expressed in the embryonic cerebral cortex. *EMBO J.* **11**, 2541-2550.

Simon, M. A., Dodson, G. S. & Rubin, G. M. (1993). An SH2-SH3 protein is required for $p21^{Ras1}$ activation and binds to *sevenless* and *Sos* proteins in vitro. *Cell* **73**, 169-177.

Sive, H. & St John, T. (1988). A simple subtractive hybridization technique employing photoactivatable biotin and phenol extraction. *Nucleic Acids Res.*, **16**. 10937.

Sive, H., Hattori, K. & Weintraub, H. (1989). Progressive determination during formation of the anteroposterior axis in *Xenopus laevis*. *Cell* **58**, 171-180.

Smith, D. B. & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in Escherichia Coli as fusions with glutathione S-transferase. *Gene* **67**, 31-40.

Sonnenberg, E., Meyer, D., Weidner, K. M. & Birchmeier, C. (1993) Scatter factor/hepatocyte growth factor and its receptor, the *c-met* tyrosine kinase, can mediate a signal exchange between mesenchyme and epithelia during mouse development. *J. Cell Biol.* **123**, 223-235.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electophoresis. *J. Mol. Biol.* **98,** 503-517.

Swanson, K. D. & Ganguly, R. (1992). Characterization of a *Drosophila melanogaster* gene similar to the mammalian genes encoding the tyrosine/tryptophan hydroxylase activator and protein kinase C inhibitor proteins. *Gene* **113**, 183-190.

Tahira, T., Ishizaka, Sugimura, T. & Nagao, M. (1988). Expression of proto-*ret* mRNA in embryonic and adult rat tissues. *Biochem Biophys. Res. Commun.* **153** 1290-1295.

Takahashi, M., Buma, Y., Iwamoto, T., Inaguma, Y., Ikida, H. & Hiai, T. (1988). Cloning and expression of the *ret* proto-oncogene encoding a tyrosine kinase with two potential trans-membrane domains. *Oncogene* **3**, 571-578.

Tassabehji, M., Read, A.P., Newton, V. E., Harris, R., Balling, R., Gruss, P. & Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**, 635-636.

Taylor, K. D. & Pikó, L. (1987). Patterns of mRNA prevalence and expression of B1 and B2 transcripts in early mouse embryos. *Development* **101**,877-892.

Tessarollo, L., Nagarajan, L. & Parada, L.F.(1992). *c-ros*: the vertebrate homolog of the *sevenless* tyrosine kinase receptor is tightly regulated during organogenesis in mouse embryonic development. *Development* **115**, 11-20.

Tiedemann, K. (1976). The mesonephros of cat and sheep. Comparative morphological and histochemical studies. In *Advances in anatomy, embryology, and cell biology*, ed. A. Brodal, W. Hild, J. van Limborgh, R. Ortmann, T. H. Schiebler, G. Töndury & E. Wolff, pp. 1-119. Springer-Verlag, Berlin.

Tiedemann, K. & Egerer, G. (1984). Vascularization and glomerular ultrastructure in the pig mesonephros. *Cell Tissue Res.* **238**, 165-175.

Timblin, C., Battey, J & Kuehl, W. M. (1990). Application for PCR technology to subtractive cDNA cloning: identification of genes expressed specifically in murine plasmacytoma cells. *Nucleic Acids Res.* **18**, 1587-1593.

Tindall, K. R. & Kunkel, T. A. (1988). Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**, 6008-6013.

Toivonen, S. (1945). Uber Die entwicklung der Vor- und Uriniere beim kaninchen. *Ann Acad. Sci. Fenn. ser. A* **8**, 1-27.

Tokarskaya, O. N., Tchurikov, N. A., Ivanov, P. L., Kramerov, D. A. & Ryskov, A. P. (1980). Sequences hybridizing to mRNA, oligo(dT) and dsRNA from pre-mRNA are contiguous in cloned mouse DNA fragments. *Nucleic Acids Res.* **8**, 425-440.

Toker, A., Ellis, C. A., Sellers, L. A., Aitken,A. (1990). Protein kinase C inhibitor proteins. *Eur. J. Biochem.* **191**, 421-429.

Toker, A., Sellers, L. A., Amess, B., Patel, Y., Harris, A., Aitken, A. (1992). Multiple isoforms of a protein kinase C inhibitor (KCIP-1/14-3-3) from sheep brain. Amino acid sequence of phosphorylated forms. *Eur. J. Biochem.* **206**, 453-461.

Ton, C. C. T., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., Van Heyningen, V., Hastie, N. D., Meijers-Heijboer, H., Drechsler, M., Royer Pokora, B., Collins, F., Swaroop, A., Strong, L. C. & Saunders, G. F. (1991). Positional cloning and characterization of a paired box and homeobox containing gene from the aniridia region. *Cell*, **67**, 1059-1074.

Torrey, T. W. (1965). Morphogenesis of the vertebrate kidney. In *Organogenesis*, ed. R. L. Dehaan & H. Ursprung, pp. 559-579. Holt, Rinehart & Winston, New York.

Torrey, T.W. (1954). The early development of the human nephrons. *Contrib. Embryol. Carnegie Inst. Wash.* **35**, 175-197.

Uchimiya, H., Kidou, S., Shimazaki, T., Aotsuka, S., Takamatsu, S., Nishi, R., Hashimoto, H., Matsubayashi, Y., Kidou, N., Umeda, M. & Kato, A. (1992). Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice ("*Oryza sativa*" L.) *Plant J.* **2**, 1005-1009.

Vainio, S., Lehtonen, E., Jalkanen, M., Bernfield, M.& Saxén, L. (1989). Epithelial-mesenchymal interactions regulate the stage-specific expression of a cell surface proteoglycan, syndecan, in the developing kidney. *Dev. Biol.* **134**, 382-391.

van Heusden, G. P. H., Wenzel, T. J., Lagendijk, E. L., de Steensma, H. Y. & van den Berg, J. A. (1992). Characterization of the yeast BMH1 gene encoding a putative protein homologous to mammalian protein kinase II activators and protein kinase C inhibitors. *FEBS Lett.* **302**, 145-150.

Vogt, T. F., Jackson Grusby, L., Rush, J. & Leder, P. (1993). Formins: Phosphoprotein isoforms encoded by the mouse limb deformity locus. *Proc. Natl. Acad. Sci. USA*. **90**, 5554-5558.

von Melchner, H., Reddy, S. & Ruley, H. E. (1990). Isolation of cellular promoters by using a retrovirus promoter trap. *Proc. Natl. Acad. Sci. USA*. **87**, 3733-3737.

von Melchner, H., De Gregori, J. V., Rayburn, H., Reddy, S., Friedel, C. & Ruley, H. E. (1992). Selective disruption of genes expressed in totipotent embryonal stem cells. *Genes & Dev.* **6**, 919-927.

Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W. & Collins, F. S. (1991). A de novo *Alu* insertion results in neurofibromatosis type 1. *Nature* **353**, 864-866.

Wang, Z. Y., Madden, S. L., Deuel, T. F. & Rauscher, F. J., III (1992). The Wilms' tumor gene product, WT1, represses transcription of the platelet-derived growth factor A-chain gene. *J. Biol. Chem.* **267**, 21999-22002.

Wang, Z. Y., Qiu, Q. Q. & Deuel, T. F. (1993). The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains. *J. Biol. Chem.* **268**, 9172-9175.

Watanabe, M., Isobe, T., Okuyama, T., Ichimura, T., Kuwano, R., Takahashi, Y. & Kondo, H. (1991). Molecular cloning of cDNA to rat 14-3-3 η chain polypeptide and the neuronal expression of the mRNA in the central nervous system. *Mol. Brain Res.* **10**, 151-158.

Watanabe, M., Isobe, T., Ichimura, T., Kuwano, R., Takahashi, Y. & Kondo, H. (1993a). Molecular cloning of rat cDNAs for β and γ subtypes of 14-3-3 protein and developmental changes in expression of their mRNAs in the nervous system. *Mol. Brain Res.* **17**, 135-146.

Watanabe, M., Isobe, T., Ichimura, T., Kuwano, R., Takahashi, Y., Kondo, H. & Inoue, Y. H. (1993b). Molecular cloning of rat cDNAs for zeta and theta subtypes of 14-3-3 protein. Unpublished as cited in Genbank accesion numbers D17614 and D17615.

Watanabe, M., Isobe, T., Ichimura, T., Kuwano, R., Takahashi, Y. & Kondo, H. (1993b). Developmental regulation of neuronal expression for the eta subtype of the 14-3-3 protein, a putative regulatory protein for protein kinase C. *Brain Res. Dev. Brain Res.* **73**, 225-235.

Werner, H., Re, G. G., Drummond, I. A., Sukhatme, V. P., Rauscher, F. J., III, Sens, D. A., Garvin, A. J., LeRoith, D. & Roberts, C. T., Jr. (1993). Increased expression of the insulin-like growth factor I receptor gene, IGF1R, in Wilms' tumor is correlated with modulation of IGF1R promoter activity by the WT1 Wilms' tumor gene product. *Proc. Natl. Acad. Sci. USA.* **90**, 5828-3582.

Wilkinson, D. G. (1992). *In situ hybridization: a practical approach.* IRL Press, Oxford.

Wolda, S. L. & Moon, R. T. (1992). Cloning and developmental expression in *Xenopus laevis* of seven additional members of the *Wnt* family. *Oncogene* **7**: 1941-1947

Woychik, R. P., Stewart, T. A., Davis, L. G., D'Eustachio, P. D. & Leder, P. (1985). An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature* **318**, 36-40.

Yaffe, D., Nudel, U., Mayer, Y. & Neuman, S. (1985). Highly conserved sequences in the 3' untranslated region of mRNAs coding for homologous proteins in distantly related species. *Nucleic Acids Res.***13**, 3723-3737.

Yamauchi, T., Nakata, H. & Fujisawa, H. (1981). A new activator protein that activates tryptophan 5-monooxygenase and tryptophan 3-monooxygenase in the presence of Ca^{2+}-calmodulin-dependent protein kinase. Purification and characterisation. *J. Biol. Chem.* **256**, 5404-5409.

Yamauchi, T. & Fujisawa, H. (1981). Tyrose 3-monooxygenase is phosphorylated by Ca^{2+}-, calmodulin-dependent protein kinase, followed by activation by activator protein. *Biochem Biophys. Res. Commun.* **100**, 807-813.

Zappavigna, V., Renucci, A., Izpisua-Belmonte, J. C., Urier, G., Peschle, C. & Duboule, D. (1991). *HOX4* genes encode transcription factors with potential auto- and cross-regulatory capacities. *EMBO J.* **10**, 4177-4187.

Zackson, S. L. & Steinberg, M. S. (1987). Chemotaxis or adhesion gradient? Pronephric duct elongation does not depend on distant sources of guidance information. *Dev. Biol.* **124**, 418-422.

Zupan, L. A., Steffans, D. L., Berry, C. A., Landt, M. & Gross, R. W. (1992). Cloning and expression of a human 14-3-3 protein mediating phospholipolysis. *J. Biol. Chem.* **267**, 8707-8710.

Appendix a

The sequence analysis of the 14-3-3 proteins was done on the sequences available as of the 7th October, 1993. I have used the names that were used at that time for my graphs and sequence alignments. It is the nature of the study proceeding on this family that by the time I was almost ready to submit this thesis the situation had changed. There are now nine mammalian isoforms sequenced (at least at the peptide level). Martin *et al.* 1993 has shown that the α isoform previously only detected by HPLC is identical to the β isoform as far as peptide sequencing of the N-acetylated proteins can detect. Aitken speculates that one may start at the first methionine whilst the other starts at the second. The δ isoform again only previously identified as an HPLC peak has been sequenced and shown to be identical to the ζ isoform. Aitken also points out that HME1 and stratifin are identical (as I pointed out in the main text) and names this isoform σ, whilst he identifies the sheep equivalent of the isoform identified by Nielsen from a T-cell library and names it τ regardless of the fact that Watanabe has recently named this θ (pers. comm.).

Appendix b Greek alphabet

α	alpha	ν	nu
β	beta	ξ	xi
γ	gamma	o	omicron
δ	delta	π	pi
ε	epsilon	ρ	rho
ζ	zeta	σ	sigma
η	eta	τ	tau
θ	theta	υ	upsilon
ι	iota	ϕ	phi
κ	kappa	χ	chi
λ	lambda	ψ	psi
μ	mu	ω	omega

Appendix c

ene, 000 (1993) 00–00

GENE 06996

Letter to the Editor

The alternative 5′-end of the *Drosophila melanogaster* epidermal growth factor receptor cDNA (*DER*) is part of the *D14-3-3* cDNA

Recombinant DNA; bovine 14-3-3 protein; protein kinase C inhibitor; tyrosine/tryptophan hydroxylase activator protein)

Jane E. McConnell[a] and Peter E. Hodges[b]

[a]MRC Human Genetics Unit, Edinburgh, EH4 2XU, UK; and [b]Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh, EH9 3JD, UK. Tel. (44-31) 650 5359

SUMMARY

A portion of the cDNA reported as a novel cDNA encoding the *Drosophila melanogaster* homolog of the bovine 14-3-3 protein by Swanson and Ganguly [Gene 113 (1992) 183–190] already exists in the database. It was originally reported as an alternative 5′ end of the *D. melanogaster* homolog of the epidermal growth factor receptor (DER). Schejter and Shilo [Cell 56 (1989) 1093–1164] later reported that this finding was due to a cloning artifact that joined the *14-3-3* cDNA onto the *DER* cDNA.

It was with interest that we read the report of the cloning of the *D. melanogaster* homolog cDNA encoding the 14-3-3 protein (Swanson and Ganguly, 1992). The high homology of this protein to both the bovine 14-3-3 proteins and the sheep protein kinase C inhibitor protein is indicative that it is a member of the very important family of 14-3-3 proteins (J.E. McC., J.B.L. Bard and P.E. H., in preparation). We would like to also point out another homology that was not reported in the original paper. The *D14-3-3* cDNA is nearly identical to a *D. melanogaster* cDNA cloned originally as an alternative 5′ end of the *D. melanogaster DER* cDNA (Schejter et al., 1986), which was later reported to be unrelated to the *DER* gene (Schejter and Shilo, 1989; Nielsen, 1991). The original cloning of this 325-bp fragment of the gene was probably due to the artifactual joining of two cDNAs, as the joint between this cDNA and the *DER* cDNA was not at a splicing junction. When the *D14-3-3* cDNA is compared to the *DER* cDNA fragment using the BESTFIT program of GCG sequence analysis package (Devereux et al., 1984), 98% sequence identity and 99% similarity is found at the aa level (Fig. 1A) and 98.2% sequence identity at the nt level (Fig. 1B).

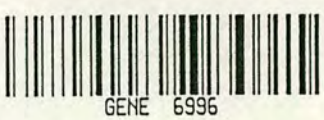
We conclude that the two cDNAs are probably from the same gene and the two nt differences may represent either population polymorphisms or sequencing errors.

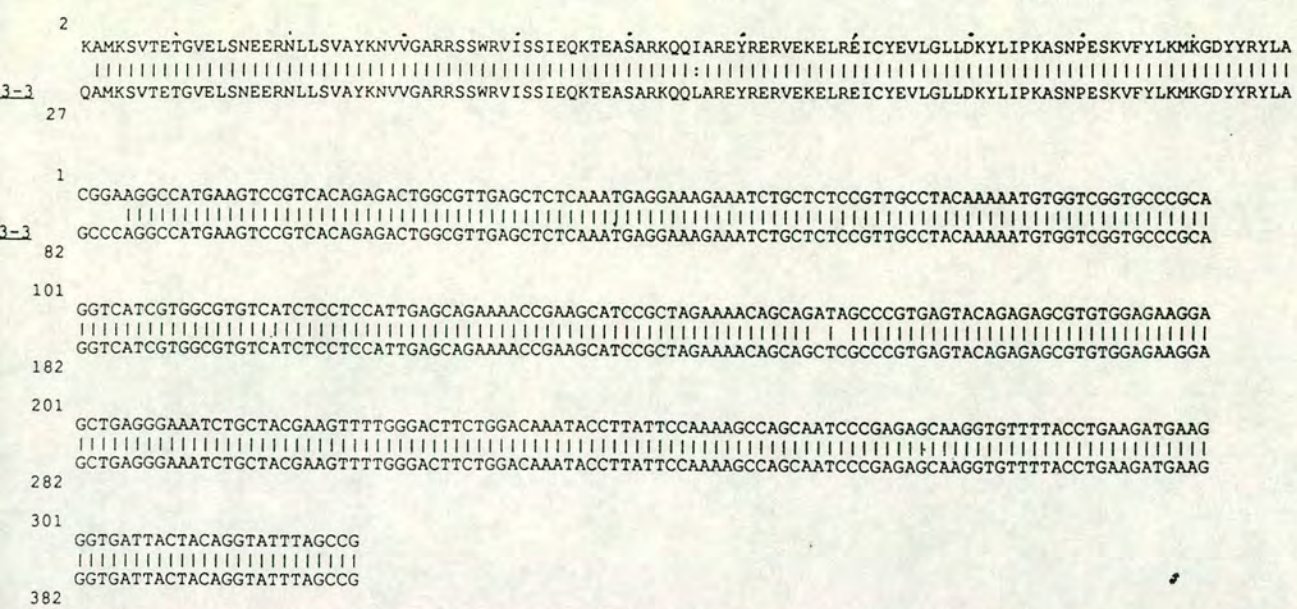
Correspondence to: Ms. J.E. McConnell, MRC Human Genetics Unit, Crewe Road, Edinburgh, EH4 2XU, UK. Tel. (44-31) 332 2471; Fax (44-31) 343 2626; e-mail: janem@uk.ac.ed.mrcvax

Abbreviations: aa, amino acid(s); bp, base pair(s); D14-3-3, putative *D. melanogaster* homolog of bovine protein 14-3-3; *D14-3-3*, gene (DNA) encoding D14-3-3; DER, *D. melanogaster* epidermal growth factor receptor; *DER*, gene (cDNA) encoding DER; GCG, Genetics Computer Group (Madison, WI, USA); nt, nucleotide(s).

ACKNOWLEDGEMENTS

J.E. McC. is currently funded by the National Kidney Research Fund, UK.





The BESTFIT aa (A) and nt (B) sequence comparison between *DER* (GenBank accession No. K03418) and *D14-3-3* (GenBank accession No. 8). Numbering of nt sequences are according to their respective GenBank entries, and aa are numbered from the initiator Met. The aa identities icated by short vertical lines, and a single conservative substitution by a colon.

RENCES

ux, J., Haeberli, P. and Smithies, O.: A comprehensive set of uence analysis programs for the VAX. Nucleic Acids Res. 12 84) 387–190.

, E., Glazer, L., Segal, D., Schlessinger, J. and Shilo, B.-Z.: The *sophila* EGF receptor gene homolog: conservation of both hor- ne binding and kinase domains. Cell 40 (1985) 599–607.

, P.J.: Primary structure of a human protein kinase regulator tein. Biochim. Biophys. Acta 1088 (1991) 425–428.

Schejter, E.D., Segal, D., Glazer, L. and Shilo, B.-Z.: Alternative 5′ exons and tissue-specific expression of the *Drosophila* EGF receptor homolog transcripts. Cell 46 (1986) 1091–1101.

Schejter, E.D. and Shilo, B.-Z.: The *Drosophila* EGF receptor homolog (DER) gene is allelic to *faint little ball*, a locus essential for embryonic development. Cell 56 (1989) 1093–1104.

Swanson, K.D. and Ganguly, R.: Characterization of a *Drosophila melanogaster* gene similar to the mammalian genes encoding the tyrosine/tryptophan hydroxylase activator and protein kinase C inhibitor proteins. Gene 113 (1992) 183–190.