HOMOLOGOUS RECOMBINATION IN EMBRYONIC STEM CELLS: TARGETING OF THE MURINE THY-1.2 AND ZFP-37 GENES

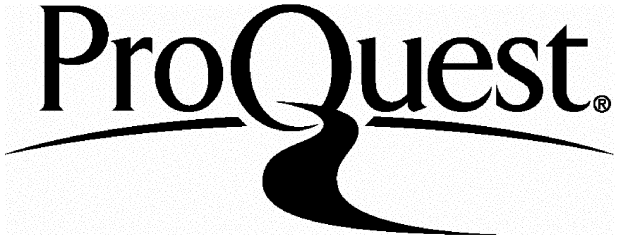
A Thesis Presented for the Degree of Doctor of Philosophy

by

David Michalovich B.Sc.

National Institute for Medical Research
The Ridgeway
Mill Hill
London

September 1995



Acknowledgements

I would like to thank Frank Grosveld for the help, encouragement and confidence he has given me in the last five years. I have greatly enjoyed the time spent in his lab, especially the Christmas pantomimes. Special thanks goes to my good friend and colleague Niels Galjart for all the help, both in and out of the lab, "Kaas Mate" and I look forward to more of those phenotype sort of days. I am also very grateful to Mike Lindenbaum for introducing me to the ES cell work and my collaborator on the Thy-1 project Roger Morris. Thanks also goes to Mirko Kuit for his photographic expertise.

I would also like to thank the members of Gene Structure and Expression and Cell Biology and Genetics (Rotterdam) for all the help, advice and friendship they have extended to me, particularly the members of labs 110, 710 and odd appendages.

Finally, I thank my parents for their constant support and Nicky for her smile. This thesis is dedicated to my Dad, who would have enjoyed its completion more than most.

Abstract

Establishment of the central nervous system follows a complex pathway of cell-to-cell interactions and sequential gene expression. Interest in our lab has focused on two genes showing neuronal expression. The *Thy-1* gene encodes an immunoglobulin-like cell surface glycoprotein which is thought to function as a receptor/adhesion molecule involved in the cessation of axonal growth. The murine *Zfp-37* gene is a member of the *Krüppel*-associated box (KRAB) Cys_2His_2 zinc-finger family of genes and is postulated to play a role in transcriptional regulation. To address the *in-vivo* functions of these genes a targeting approach in mouse embryonic stem cells (ES) has been followed to produce "loss of function" mutants. An ES cell clone carrying a targeted *Thy-1* allele was succesfully obtained, however this failed to give germline transmission. Mice carrying a mutated *Thy-1* allele became available from a another laboratory and therefore the project was pursued as a collaboration. The murine *Zfp-37* gene was disrupted by insertion of a *lacZ* reporter gene via homologous recombination in ES cells. On a three strain genetic background homozygous *Zfp-37*$^{-/-}$ mice appear healthy and fertile and are represented at close to the predicted Mendelian ratio. On a (129 x C57BL/6) genetic background three of the six homozygous animals obtained have died, however, it remains to be determined if the deaths are linked to the targeted mutation.

The targeted reporter gene expression has shown *Zfp-37* to be consistently expressed in the developing nervous system of the embryo from embryonic day 9.5 to 12.5. Other areas of the embryo also show expression such as the hind gut (e9.5) and the developing limbs (e10.5-e12.5). In the three week postnatal brain high levels of reporter gene expression are found in the hippocampus, superior colliculus, hypothalamus, olfactory tuberculum and granular and Purkinje cell layers of the cerebellum. Analysis of the neuronal reporter gene expression in *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice suggest ZFP-37 to be part of an autoregulatory mechanism regulating its own expression.

Table of Contents

List of figures

List of tables

Abbreviations

a.a	amino acid
ATP	adenosine triphosphate
bp	base pair
BRL	buffalo rat liver
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
cDNA	complementary DNA
$CHCl_3$	chloroform
Ci	Curie
cpm	counts per minute
CsCl	caesium chloride
CTP	cytosine triphosphate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytodine triphosphate
dGTP	deoxyguanosine triphosphate
ddH_2O	double distilled water
DMSO	dimethyl sulphoxide
DMEM	Dulbeco's modified Eagle's media
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease 1
dNTP	deoxyribonucleotide triphosphate
dpc	days post coitum
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
e	embryonic day
EDTA	ethylenediamine tetra-acetic acid
EGTA	ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetra-acetic acid
ES	embryonic stem
F	Farad

FCS	fetal calf serum
FIAU	1(1-2-deoxy-2-fluoro-β-Darabinofuransyl)-5-iodouracil
g	gram
GANC	gancyclovir
GPI	glycosyl phosphatidylinositol
GTP	guanosine triphosphate
HEPES	N-2 hydroxyethylpiperazine-N-2-ethanol sulphonic acid
HCl	hydrochloric acid
hprt	hypoxanthine phosphoribosyl transferase gene
HSV1	Herpes simplex virus type 1
Ig	immunoglobulin
k	kilo
kb	kilobase
KCl	potassium chloride
KDa	kilodalton
$K_3Fe(CN)_6$	potassium ferricyanide
$K_4Fe(CN)_6$	potassium ferrocyanide
K_2HPO_4	dipotassium phosphate
KH_2PO_4	monopotassium phosphate
KOAc	potassium acetate
l	litre
LiCl	lithium chloride
LTP	long term potentiation
m	milli (10^{-3})
M	molar
Mb	Mega base
$MgCl_2$	magnesium chloride
MOPS	3-(N-Morpholino)propanesulphonic acid
mRNA	messenger RNA
n	nano (10^{-9})
NaCl	sodium chloride
Na_2HPO_4	disodium phosphate

NaOAc	sodium acetate
NaOH	sodium hydroxide
neo	neomycin phosphotransferase gene
NP40	ethylphenyl(ethyleneglycolether)
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pfu	plaque forming unit
PMSF	phenylmethylsulphonyl fluoride
RNA	ribonucleic acid
RNase	ribonuclease
RT	reverse transcriptase
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SSC	sodium chloride, sodium citrate
SV40	simian virus 40
TEMED	N,N,N',N'-tetramethylethylenediamine
tk	thymidine kinase gene
Tris	Tris(hydroxymethyl)aminoethane
TTP	thymidine triphosphate
tRNA	transfer RNA
uv	ultraviolet
V	volts
vols	volumes
v/v	volume per volume
w/v	weight per volume
w/w	weight per weight
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
μ	micro (10^{-6})

In DNA sequence:

"A" adenine residue

"C" cytosine residue

"G" guanine residue

"T" thymine residue

Chapter 1

Introduction

1.1 General introduction

The generation and establishment of the mammalian central nervous system follows a complex pathway of cell-to-cell interactions and sequential gene expression. To attempt to dissect the cellular and molecular mechanisms underlying these processes, several approaches have been followed including; cell lineage marking, analysis of gene expression patterns and analogies drawn from the invertebrate nervous systems. A number of naturally occurring mouse mutants showing neuronal defects are also available for study, however, this number is far out-weighed by the number of genes that have now been implicated in nervous system development and function. The advent of gene targeting to produce defined "loss of function" mouse mutants has provided the means to assess the *in-vivo* role of these genes.

This thesis attempts to define the *in-vivo* roles of the murine *Thy-1* and *Zfp-37* genes via gene targeting in murine embryonic stem cells followed by production of "loss of function" phenotypes.

Thy-1 was first described as an allelic antigen on the surface of mouse thymocytes over thirty years ago (Reif and Allen, 1964). It has subsequently been shown to be a major cell surface glycoprotein on virtually all neurons (reviewed by Morris, 1985). The protein, carbohydrate and nucleic acid structure of Thy-1 are known. As a member of the immunoglobulin-like super gene family (Williams and Gagnon, 1982) Thy-1 is thought to behave as a recognition molecule. Based upon its structure, conserved neuronal expression and *in-vitro* inhibition of neurite outgrowth, Thy-1 is believed to function as a neuronal recognition molecule involved in the control of axonal outgrowth (reviewed by Morris, 1992). Despite the wealth of information available regarding Thy-1, its true *in-vivo* role remains elusive.

Through an interest in the transcriptional events involved in nervous system development we have focused on a putative neuronal transcription factor gene, *Zfp-37.* First described as a testis specific cDNA (Nelki *et al.*, 1990; Burke and Wolgemuth, 1992) it was later shown to be expressed in both the developing nervous system and adult brain (Mazarakis *et al.*, 1995). The gene encodes several protein isoforms, via the use of alternative promoters and splice variants (Mazarakis *et al.*, 1995), each isoform containing twelve C_2H_2 zinc-finger motifs indicative of potential nucleic acid binding.

The major protein isoform of ZFP-37 has been shown to contain a *Krüppel* associated box or KRAB domain (Mazarakis *et al.*, 1995). This places *Zfp-37* into the large KRAB containing gene family (Pieler and Bellefroid, 1994). Recently the KRAB domain has been postulated to act as a transcriptional repressor (Deuschle *et al.*, 1995; Margolin *et al.*, 1994; Witzgall *et al.*, 1994; Pengue *et al.*, 1994).

Part two of this chapter discusses the application of homologous recombination in embryonic stem cells to produce defined mouse mutants. Part three and four will concentrate on the murine *Thy-1* and *Zfp*-37 genes respectively.

1.2 Homologous recombination in embryonic stem cells

1.2.1 Introduction

Gene targeting can be defined as the homologous recombination between a defined genomic locus and exogenous DNA to produce a specific genomic alteration (Bollag *et al.*, 1989; Capecchi, 1989). Combining gene targeting with murine embryonic stem cell technology has made it possible to introduce specific alterations into the murine genome, allowing both gene inactivation and more subtle mutations to be created. The application of this technique has provided valuable insight into the function of a number of genes within the context of the whole animal and has also been used to produce animal models for specific human genetic diseases.

At present the general procedure for gene targeting in the mouse involves the introduction of a defined mutation into a cloned fragment of the particular gene of interest using normal recombinant DNA technology. This fragment, or targeting vector, is introduced into pluripotent embryonic stem (ES) cells, the mutation is then transferred via homologous recombination into the genome of the ES cell. These ES cells, now carrying the mutated allele, are microinjected into a developing mouse blastocyst or co-cultured with recipient embryos. The ES cells have the ability to contribute to all tissues of the developing mouse including the germ cells. The resulting animal, or chimera, therefore contains tissues derived from both the ES cells and host embryo. These mice can subsequently be bred to give heterozygous and hence homozygous mice for the desired mutation. This general scheme for gene targeting in the mouse is described in

Figure 1. (For reviews see Bollag *et al.*, 1989; Capecchi *et al.*, 1989; Porter *et al.*, 1989; Melton, 1990; Joyner, 1991; Rossant, 1991; Zimmer, 1992; Bronson and Smithies, 1994).

Early recombination experiments initially involved the homologous recombination between introduced exogenous DNA molecules only (Folger *et al.*, 1982; Miller and Temin, 1983). Subsequent studies using transfected cell lines carrying a mutation in an introduced selectable marker gene, demonstrated recombination between an introduced plasmid and plasmid sequences integrated into the genome of the cell (Lin *et al.*, 1985; Thomas *et al.*, 1986). The first targeting of an endogenous gene was demonstrated by Smithies *et al.* (1985), who successfully introduced a specific mutation into the human β-globin gene in human EJ carcinoma and HU11 cell lines.

Although these experiments demonstrate that mammalian cells possess the enzymatic machinery to carry out homologous recombination, the recombination is a relatively rare event. The ratio of homologous recombination compared to random integrations in the above experiments is in the order of 1×10^{-2} to 1×10^{-5}. The exogenous DNA predominantly integrates into non-homologous sites in the genome (Roth and Wilson, 1985; Roth *et al.*, 1985).

Since the recombination event is rare, introduction of specific mutations into the germline of an animal requires a large number of cells to be transfected. Furthermore, these cells must be able to contribute to the germline after being manipulated. Mouse embryonic stem (ES) cells fulfil these criteria. ES cells can be grown to large numbers and be genetically manipulated in culture and still retain their ability to populate the germline (Robertson *et al.*, 1986; Doetschman *et al.*, 1987; Thomas and Capecchi, 1987).

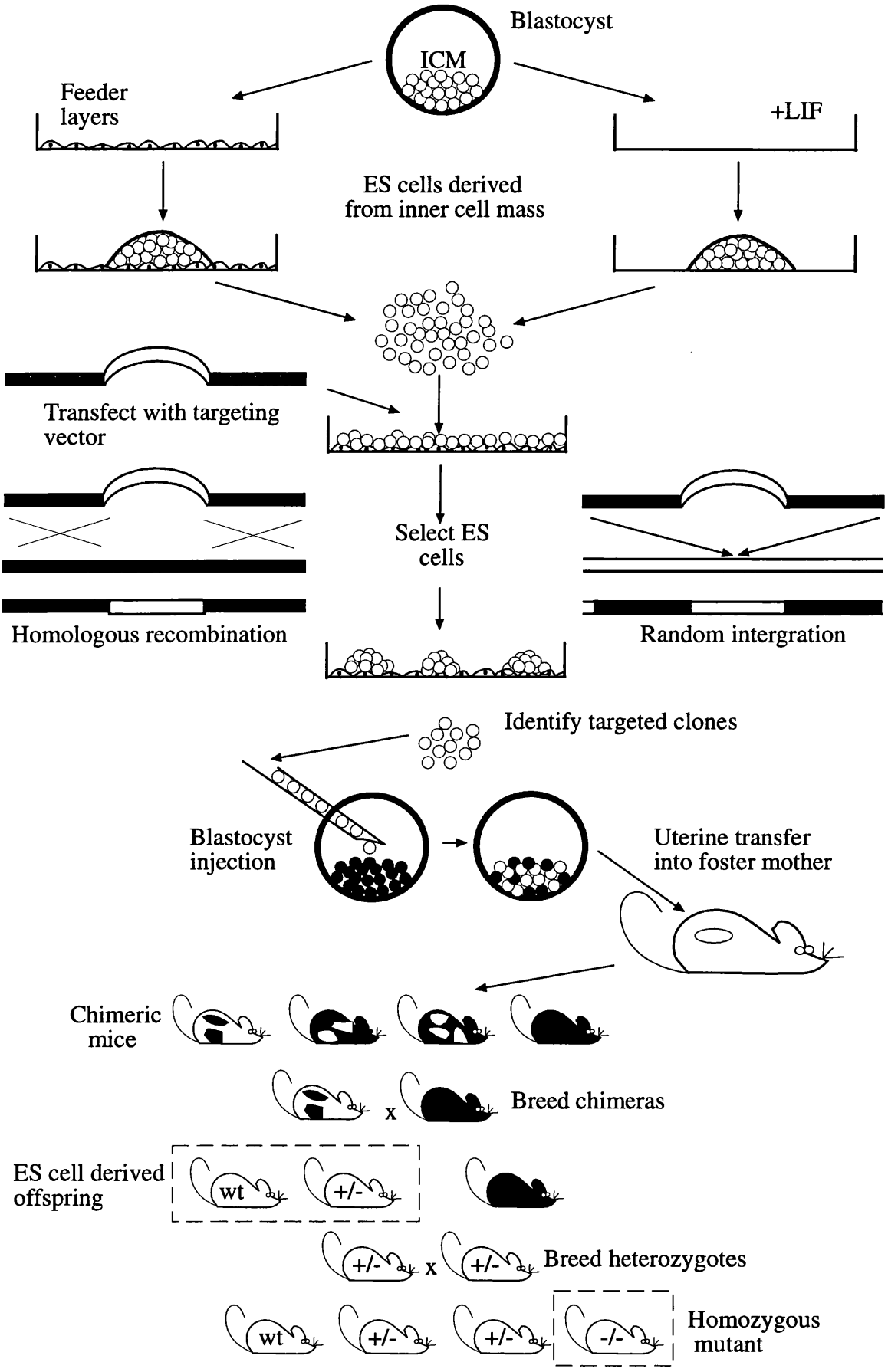
The following sections discuss the isolation and manipulation of ES cells and application of homologous recombination to introduce specific mutations into the mouse germline.

1.2.2 Murine embryonic stem cells

The pluripotent nature of cells within the pre-implantation mammalian embryo was first alluded to through the formation of composite or chimeric animals produced

Figure 1. General scheme for gene targeting in the mouse

Embryonic stem cells (ES) are derived from the inner cell mass (ICM) of wild type blastocysts. The blastocyst is disaggregated onto a monolayer of mouse fibroblasts which have been mitotically inactivated or into leukaemia inhibitory factor (LIF) conditioned media. The ES cell culture is expanded and can be maintained in an undifferentiated state. The DNA targeting vector containing the defined mutation is used to transfect the ES cells. The mutation is transferred to the genome of the ES cells via homologous recombination. ES cells containing the desired targeting event can be selected using a number of selection strategies. Once identified, the targeted ES cells are reintroduced into blastocysts isolated from mice which differ at coat colour alleles. This can be achieved by microinjection of the ES cells into the blastocyst cavity (as depicted here) or by coculture with eight cell stage embryos. After transfer of the blastocysts or eight-cell stage embryos to the uterine horns of a pseudopregnant mother, the ES cells participate in the development of the embryo and can contribute to all forms of tissue including the germline, forming a chimeric animal. Chimeras can be identified by coat colour and are mated to mice of a coat colour over which the ES cell pigmentation markers are dominant, allowing rapid screening for mice containing an ES cell derived genotype. Half of the ES cell derived offspring should possess the desired mutation. Interbreeding of heterozygotes carrying the mutant allele can give rise to mice homozygous for the desired mutation.


Blastocyst
ICM
Feeder layers
+LIF
ES cells derived from inner cell mass
Transfect with targeting vector
Select ES cells
Homologous recombination
Random intergration
Identify targeted clones
Blastocyst injection
Uterine transfer into foster mother
Chimeric mice
x
Breed chimeras
ES cell derived offspring
wt
+/-
+/-
x
+/-
Breed heterozygotes
wt
+/-
+/-
-/-
Homozygous mutant

by mechanically aggregating different pre-implantation embryos. Each embryo could be shown to contribute to tissues in the resulting chimeric animal by following genetic markers in which the embryos differed (Tarkowski, 1961; McClaren, 1976). Subsequently, the characterisation of cell lines isolated from teratocarcinomas, induced by re-implantation of 1 to 7.5 day old mouse embryos into an extra-uterine position, showed these cells to possess the ability to partly contribute to the development of an embryo in which they had been re-introduced (Brinster, 1974; Papaioannou *et al.*, 1975; Mintz and Illmensee, 1975). These cells, termed embryonal carcinoma (EC) cells, although compromised in their developmental ability, did show that a pluripotent species of cell was present in the pre-implantation embryo. The above studies finally led to the isolation and establishment in culture of mouse embryonic stem (ES) cell lines by Evans and Kaufman (1981) and independently by Martin (1981).

The isolation of ES cells was initially carried out by inducing the implantation delay of blastocysts at 2.5 days p.c and isolating the embryos 4-6 days later (Evans and Kaufman, 1981). The inner cell mass (ICM) was subsequently isolated and allowed to expand and was then disaggregated on to a mitotically inactivated feeder cell monolayer, consisting of a mitomycin-C treated mouse fibroblasts cell line (STO cell line). Colonies bearing morphological similarities to EC colonies were isolated and passaged to mass cultures (Evans and Kaufman, 1981). This method has been adapted to allow the direct isolation of ES cells from 3.5 day p.c ICMs without inducing implantation delay (Axelrod and Lader, 1983; Axelrod, 1984). Martin (1981) described a different method for isolation of ES cells, culturing isolated ICMs of 3.5 day old blastocysts in EC cell conditioned media. Again a mitotically inactivated mouse fibroblast feeder layer was used.

When re-introduced into a recipient blastocyst ES cells can contribute to all tissue types including the germ cells of the resulting chimeric animal (Bradley *et al.*, 1984). With appropriate culture conditions they can be maintained and propagated in an undifferentiated state *in vitro* and still retain their pluripotency even after 250 replications (Suda *et al.*, 1987).

1.2.3 Maintenance of ES cells

For gene targeting purposes ES cells have to remain in an undifferentiated state. Classically this has been achieved by the isolation and growth of ES cells in media containing a high percentage of fetal calf serum, 10-20%, and upon a mitotically inactivated fibroblast feeder layer (Evans and Kaufman, 1981; Martin, 1981; Robertson, 1987). Removal of the feeder cell layer leads to the differentiation of ES cells (Doetschman *et al.*, 1985). This implies that the feeder layer provides some inhibitory signal or factor which suppresses stem cell differentiation. However, differentiation can also be inhibited by the addition of conditioned media from Buffalo Rat Liver cells (BRL) showing that a soluble factor can substitute for the activity of feeder layers (Smith and Hooper, 1987). This activity was identified to be a soluble glycoprotein, termed ES cell differentiation inhibiting activity (DIA) (Smith *et al.*, 1988). Subsequent structural and functional comparisons showed this molecule to be identical to the myeloid leukaemia inhibitory factor (LIF) (Smith *et al.*, 1988) and the human interleukin differentiating activity (Moreau *et al.*, 1988; Williams *et al.*, 1988a). Further study has shown that DIA/LIF is not only produced as a soluble protein but also in an extracelluar matrix associated form. The different forms are produced by the use of alternate first exons in the gene and hence produce proteins which differ at their amino terminus (Rathjen *et al.*, 1990). The above study also showed that embryonic fibroblast feeder layers produce the matrix-associated form but not the soluble form of DIA/LIF.

A number of strategies are now used for the routine maintenance of ES cells. Mitotically inactivated feeder layers are commonly employed, their inactivation is either achieved by γ-irradiation or the use of a mitosis inhibiting agent, mitomycin-C. The feeder layers can either be produced from primary embryonic fibroblasts (Doetschman *et al.*, 1985) or by using a mouse fibroblast cell line such as the STO cell line (Ware and Axelrod, 1972). An adaptation of the latter is the use of SNL feeder cells, a STO subline which has been stably transfected with constructs expressing LIF and also the neomycin resistance gene, these cells secrete a soluble form of LIF into the media and can be used in conjunction with neomycin resistance selection (McMahon and Bradley, 1990). As mentioned above, BRL cell conditioned media containing LIF (Smith and Hooper, 1987) and also conditioned media from the human bladder carcinoma 5637 cell

line (Williams *et al.*, 1988) can been used instead of feeder layers (Handyside *et al.*, 1989). In addition, cloning of the LIF cDNA and expression in bacteria has allowed the recombinant molecule to be used directly in the culture media (Baribault and Kemler, 1989; Bradley 1990).

ES cell lines have been established directly in media supplemented with recombinant LIF, however these lines were shown to have a high degree of chromosomal abnormalities (Nichols *et al.*, 1990). This suggests that the use of LIF alone may not be sufficient to maintain the ES cells in prolonged undifferentiated culture. Furthermore, some feeder cell lines do not produce LIF to any detectable level (Bradley, 1990), suggesting that as yet unidentified factors supplied by these cells may compensate for and/or act in concert with LIF to maintain the undifferentiated state of ES cells in culture.

A number of ES cell lines with germline potential have been established and are commonly used by many research groups world wide. These include the CCE (Robertson *et al.*, 1986), D3 (Doetschman *et al.*, 1985), E14 (Handyside *et al.*, 1989), J1 (Li *et al.*, 1992), AB1 (McMahon and Bradley, 1990) and R1 (Nagy *et al.*, 1993) cell lines. These are all derived from male blastocysts from substrains of the 129 inbred mouse strain. This strain was initially chosen as a donor line for ES cells because of its use in the isolation of EC cells (Stevens, 1973; Evans and Kaufman, 1981). ES cell lines from the C57BL/6J strain have also been established (Doetschman *et al.*, 1985; Lederman and Bürki, 1991). The D3 and J1 ES cell lines were originally established on primary embryonic fibroblasts, CCE on STO cells, AB1 and R1 on the STO subline SNL and E14 cells on 60% BRL-conditioned media in combination with feeder layers.

Maintaining the potential of the ES cells to colonise the germline is crucial for a successful gene targeting project. ES cells may lose their totipotency in culture (reviewed by Wurst and Joyner, 1993). This may occur for several reasons. Firstly the cells may be triggered to enter a certain differentiation pathway which may not be detectable by morphology or growth rate (Doetschman *et al.*, 1985). Secondly karyotypic changes may occur in culture (Robertson, 1987). A 6:14 translocation has been documented in clonal isolates of the CCE line (Schwartzberg *et al.*, 1989) occurring after approximately thirty five doublings. A trisomy of chromosome 8 has also been found in a subclone of the E14 line (M. Jaegle personal communication).

Aneuploid cells may gain growth advantages in culture and rapidly dominate normal ES cells. Changes in culture conditions may give rise to these events. In general, ES cell cultures should be maintained under the same conditions as used in their primary isolation to avoid these events occurring (Wurst and Joyner, 1993).

Checking for chromosomal abnormalities, such as aneuploidy and translocations, can give an indication as to whether the ES cells are suitable for use in gene targeting experiments. Subtle mutations will be missed, however, and therefore the only absolute test of the ability of an ES cell line to contribute to the germline is the reintroduction of the cells into blastocysts and to confirm the contribution of that cell line to the germline of the resulting chimeric animal by breeding.

1.2.4 Generation of chimeric mice

As mentioned above, ES cells are capable of contributing to all tissues of a chimeric animal, including the germline, once they have been reintroduced into the early preimplantation embryo. The most commonly used method of chimera generation is direct injection of ES cells into the blastocoel cavity of 3.5 day old mouse blastocysts (Bradley *et al.*, 1984; Bradley, 1987). Other methods of producing chimeras include sandwiching of a clump of ES cells between two eight-cell stage embryos (Nagy *et al.*, 1990; Nagy and Rossant, 1993), co-culture of ES cells with tetraploid embryos (Nagy and Rossant, 1993) and simple co-culture of ES cells with eight-cell stage embryos *en masse* (Wood *et al.*, 1993).

As discussed earlier the most common substrain of mouse used for ES cell isolation is the 129 strain. C57Bl/6J mice are commonly used to provide recipient blastocysts as they have proved to be compatible hosts for 129 ES cells. The C57BL/6J strain differs from the 129 strain in both coat colour loci and glucose phosphate isomerase (GPI) isozyme. The contribution that the ES cells make in the resulting chimera can therefore be determined using these genetic markers. For example AB1 ES cells are derived from the 129/Sv/Ev substrain, which carries the coat colour alleles agouti(A)/agouti white belly (A^w) and is wild type at the brown (B) coat colour locus. The brown allele is so named because mutations at this locus lead to a brown coat colour, whereas a wild type brown allele (B) results in a black coat. These two loci,

agouti and brown, affect hair follicle formation and melanocyte pigmentation respectively. The resulting coat of a 129/Sv/Ev mouse is made up of black hairs with a subapical band of yellow pigmentation giving the overall effect of a brown coat (Lyon and Searle, 1990). The agouti phenotype is dominant over brown. C57BL/6 mice carry agouti (a) and brown (B), this combination results in a black coat. AB1-C57BL/6 chimeras can therefore be detected by the mixed coat of agouti and black.

The GPI isozyme has three variants which can be distinguished by their electrophoretic mobility. The host strain C57BL/6J possesses the GPI-1^b isozyme whereas ES cells AB1, D3 and E14 posses the GPI-1^a form. Analysis of the different GPI content of a tissue allows the contribution of the ES cell derived cells within that tissue to be determined (reviewed by Papaioannou and Johnson, 1993).

The degree of coat colour chimerism is commonly used as an indication of the contribution the ES cells have made to other tissues, such as the gonads. Therefore chimeras showing high coat colour contribution from the ES cells (> 50%) are considered likely to have an equivalent contribution of ES cell derived tissue in the gonads. However, these comparisons should be treated cautiously as mutation in the ES cells or differentiation of the ES cell to a particular cell lineage may result in high coat colour contribution but low germline contribution (van Deursen personal communication and this study). A better indication of potential germline contribution is a distortion in the sex ratio of the chimeric offspring obtained. ES cells containing XY chromosomes will cause sex conversion of a female embryo because of a threshold level of contribution of ES cells to the sex determining tissue is reached. Sex converted chimeras should therefore prove to be good germline transmitters as the gonads must carry ES cell derived cells (Bradley *et al.*, 1984). Exceptions to this can occur when the Y chromosome is lost or develops a mutation in the ES cells, which can result in transmission of the male derived ES cell genome by female chimeras (Kuehn *et al.*, 1987).

Finally, transmission of the ES cell genome from the chimera to the offspring can again be scored using the coat colour markers. Chimeras are routinely bred with mouse strains that differ at coat colour alleles from those possessed by the ES cells. A commonly used cross is between the 129 chimera and C57BL/6 mice. Heterozygous wild type agouti (A/a) is dominant over brown (B) and therefore pups which are partly

ES cell derived will appear to be agouti.

1.2.5 Homologous recombination mechanisms

Gene targeting is reliant on the process of homologous recombination to direct pre-determined mutations into the genome. The precise mechanisms of homologous recombination in mammalian cells are not well defined, however extrapolating from work carried out in yeast and bacteria it is suggested that recombination occurs via Holliday structure intermediates with a mechanism analogous to double strand break repair.

Several models have been proposed to account for meiotic recombination in lower fungi and yeast (Holliday, 1964; Meselson and Radding; 1975; Szostak *et al.*, 1983). As shown in Figure 2, recombination is thought to initiate from a double strand break, after which exonuclease activity leaves a gap flanked by 3' single strand extensions (Cao *et al.*, 1990; Sun *et al.*, 1991). One of the 3' extensions invades the other duplex in a region of homology, forming a D-loop. The D-loop is enlarged by repair synthesis primed from the 3' invading strand until the other 3' extension is reached. Repair synthesis can now occur from this end so that the gap is repaired. Branch migration results in the formation of two Holliday junctions, resolution of which results in either the crossover or non-crossover recombination products (Szostak *et al.*, 1983). The earlier models proposed by Holliday (1964) and Meselson and Radding (1975) have similar features involving strand invasion and Holliday junction formation but initiating from nicks in the DNA rather than double strand breaks.

Enzymes which can carry out strand exchange, formation of Holliday cross-overs, branch migration and resolution have been identified in *E. coli*. RecA protein can form a helical filament along single stranded DNA, can search for target sequences, mediate strand invasion *in-vitro* and Holliday junction formation (DasGupta *et al.*, 1981; Dunderdale *et al.*, 1991; Ferrin and Camerini-Otero, 1991, reviewed by Kowalczykowski 1991; West, 1992). The RuvA and RuvB proteins are involved in Holliday junction branch migration (Parsons *et al.*, 1992; Shiba *et al.*, 1991; Tsaneva *et al.*, 1992). RuvC protein acts to resolve the Holliday junction (Dunderdale *et al.*, 1991; Bennett *et al.*, 1993). Furthermore, RecA homologues have been found in yeast

Figure 2. Double-strand-break repair model for homologous recombination

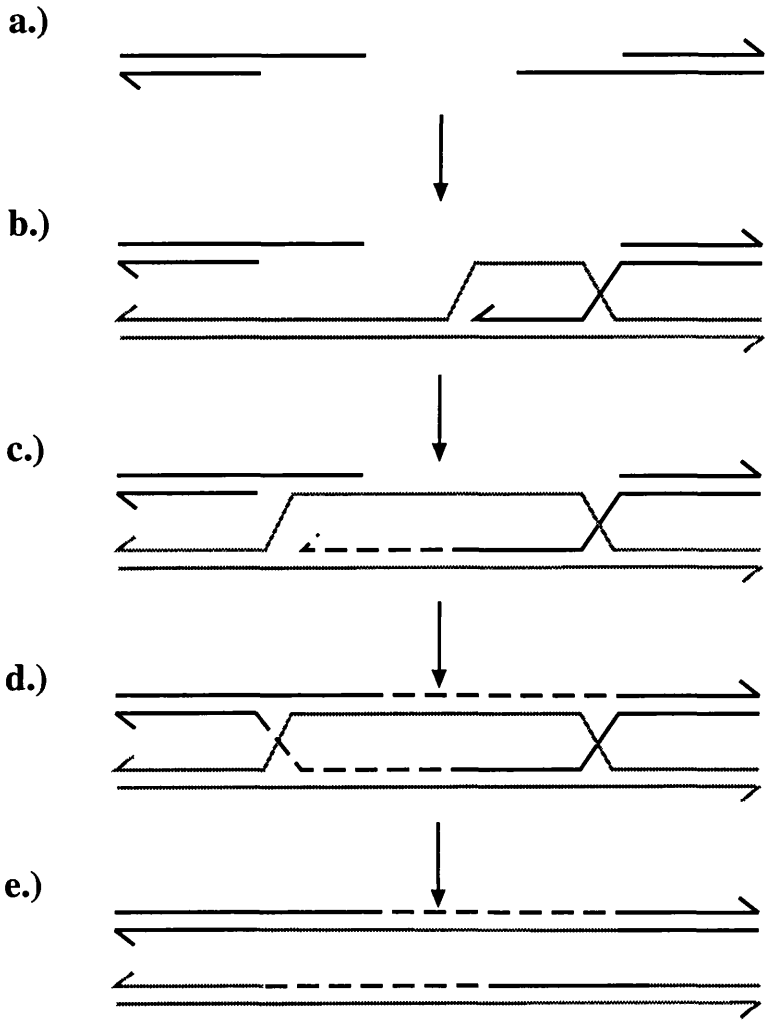
a.) A double strand break is made in one duplex, and a gap flanked by 3' single strand extensions is formed by the action of exonucleases.

b.) One of the 3' extensions invades a homologous duplex, forming a D-loop.

c.) The D-loop is enlarged by repair synthesis until the other 3' end can anneal to complementary single-stranded sequences.

d.) Repair synthesis primed from the second 3' end completes the process of gap repair, resulting in the formation of two Holliday junctions. Resolution of two junctions by cutting either inner or outer strands leads to the noncrossover products shown in (e.) or the crossover products (f.) to be formed.

Adapated from Szostak *et al.*, 1983.


a.)
b.)
c.)
d.)
e.)
non-crossover
f.)
crossover

(Shinohara *et al.*, 1992) and recombinase activity has been partially purified from human cells (Hseih *et al.*, 1986) suggesting homologous members of the recombination machinery exist in eukaryotes.

1.2.6 Targeting vectors

Introduction of a specific mutation into the genome of mammalian cell requires the use of targeting vectors. These consist of cloned fragments of genomic DNA homologous to the target gene, into which the specific mutation is designed. Based on recombination into yeast chromosomal sites (Hinnen *et al.*, 1978; Orr-Weaver *et al.*, 1981) two classes of targeting vectors have been designed, replacement or "Ω" type and insertion or "0" type vectors (Thomas and Capecchi, 1987) (Figure 3). The precise recombination mechanisms employed by these vectors are not well understood as the host enzymes involved have not been identified fully and homologous recombination has not been reconstituted *in-vitro*. However, a large amount of information can be gained from altering the DNA substrates presented to the recombination machinery and studying the reaction products.

Replacement vectors are designed to integrate foreign DNA sequences into a defined site, adding to, shifting or deleting sequences within the target locus, based on the particular vector design (Figure 3a). The recombination is thought to occur via the formation of cross overs at either side of the disruption, possibly initiated by nicks in the DNA (Hasty *et al,*. 1992). There is evidence for these crossovers to occur towards the ends of the homologous vector sequences (Deng *et al.*, 1993). Resolution of the resulting Holliday structures flanking the disruption can lead to insertion of the disrupting sequence into the desired location.

Insertion vectors insert the entire vector into the target site. As the vector is linearised in the region of homology to the target it is thought that the insertion occurs via the double strand break repair model discussed above (Szostak *et al.*, 1983; Hasty *et al.*, 1992). Correct recombination of an insertion type vector results in duplication of the target locus (Figure 3b).

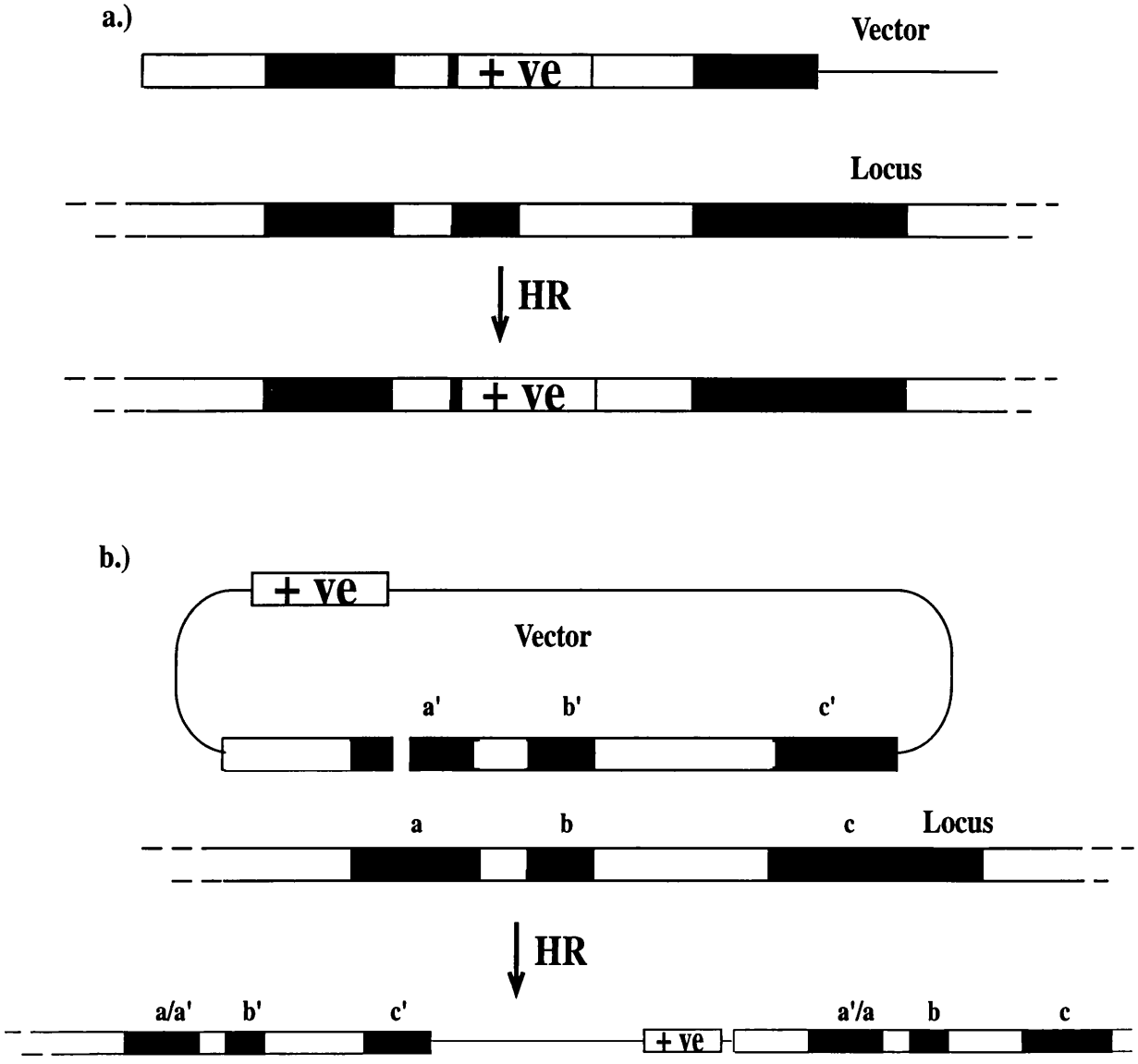
Detailed analysis of the recombination products using both these vector types has shown that a variety of events including single crossovers at one homologous arm only,

Figure 3. Replacement and insertion recombination vectors

a.) Replacement type targeting vector and recombination product. Recombination between targeting vector and locus results in sequence replacement, in this case a positive selectable marker gene, such as a bacterial gene for neomycin resistance (*neo)* or hygromycin resistance (*hph)* is recombined into the gene locus.

b.) Insertion type targeting vector and recombination product. The insertion type targeting vector is linearised within a region of homology to the target gene locus ie in exon (a). In this case the positive selectable marker is located in the vector backbone. Recombination with the target locus results in a duplication of the locus and the inclusion of the vector backbone.

Open boxes represent introns, closed boxes represent exons and single lines represent plasmid sequences. HR, Homologous Recombination. + ve, positive selection cassette.


a.)
Vector
+ ve
Locus
HR
+ ve
b.)
+ ve
Vector
a'
b'
c'
a
b
c
Locus
HR
a/a'
b'
c'
+ ve
a'/a
b
c

gene conversion events and combinations of the above can occur. The analysis can be further complicated by inter-molecular recombination and ligation of the vector before it recombines with the chromosomal site (Hasty *et al.*, 1991b; Hasty *et al.*, 1992).

1.2.6i Factors affecting homologous recombination efficiency

Although a number of parameters affecting homologous recombination efficiency had been explored during experiments in somatic mammalian cells, it has been the application of gene targeting in ES cells that has prompted further investigation. In these experiments absolute targeting efficiency can be defined as the number of cells identified to contain the recombined locus per number of cells surviving transfection.

The length of homologous sequence used in the targeting vector greatly influences the frequency of gene targeting. In the case of the murine *hprt* gene, a stepwise increase of homology from 2 to 12kb resulted in a linear increase in absolute targeting frequency from $1x10^{-6}$ to $1x10^{-4}$, a 100 fold increase. The system appeared to saturate at approximately 14 kb of homology. In this study both replacement and insertion vectors were used (Deng and Capecchi, 1992). A similar relationship was found by others (Hasty *et al.*, 1991a).

Base sequence divergence between vector and chromosomal target has also been shown to be an important factor in targeting efficiency. The use of so called isogenic DNA, DNA from the same strain of mouse as the ES cell, has greatly improved targeting efficiency for several loci including the retinoblastoma susceptibility gene (Rb) (te Riele *et al.*, 1992), the creatine kinase M gene (van Deursen and Wieringa, 1992) and *hprt* gene (Deng and Capecchi, 1992). The increased efficiencies were in the order of 20-fold, 25-fold and 4 to 5-fold respectively. The importance of this factor is obviously dependant on the degree of polymorphism which will vary between loci.

Comparison of the recombination efficiency of insertion type vectors against replacement vectors has been made by two groups using the *hprt* gene as the chromosomal target. Hasty *et al.* (1991b) found that insertion vectors recombine up to 9-fold more efficiently than replacement vectors with the same homology. A second study by Deng and Capecchi (1992) initially confirmed this result. However, when the polymorphism between vector and target site were taken into account, i.e. when isogenic

DNA was used for making the targeting constructs, the two types of vector recombined with comparable efficiency. They concluded from this that insertion vectors are less sensitive to base pair heterology than replacement vectors.

For gene replacement vectors the distribution of the homology about the disruption is also an important consideration. 472 bp of homology in one arm has shown to be sufficient to support formation and resolution of crossover junctions (Hasty *et al.*, 1991a), though generally if the extent of homology contained by one arm is below 1kb then the fidelity of recombination may be affected (Berinstein *et al.*, 1992; Thomas *et al.*, 1992).

The frequency of gene targeting using replacement vectors is not affected by the length of nonhomologous disrupting DNA. In the context of the *hprt* gene, disruptions of up to 12 kb have been shown not to affect targeting efficiency (Mansour *et al.*, 1990). It has therefore been possible to include an *E.coli lacZ* gene encoding β-galactosidase as a reporter gene together with the neomycin resistance gene, total 5.4 kb, into the *int-2* locus. This has also been achieved in a different locus by Le Mouellic *et al.* (1990 and 1992), in this study the entire plasmid sequence was also included within the disrupting DNA.

Replacement vectors where the homologous arms are directed against sequences which are a distance apart in the chromosomal target, so called deletion vectors, may also target efficiently. This was shown in experiments where a 15 kb region of the T-cell receptor β-subunit locus was replaced with a 2 kb neomycin resistance cassette (Mombaerts *et al.*, 1991). A more detailed study of these deletion type replacement vectors, showed a deletion of up to 19.2 kb could be made with an efficiency comparable to that of a 3 kb deletion or when a 1.7 kb selection cassette was inserted into the locus (Zhang *et al.*, 1994).

Topological constraints such as linearisation and inclusion of DNA nicks in the targeting vector have also been studied. Linearisation of targeting vectors has been shown to be an important factor in increasing targeting efficiency (Folger *et al.*, 1982, Kucherlapati *et al.*, 1984; Lin *et al.*, 1984; Thomas *et al.*, 1986; Doetschman *et al.*, 1987; Hasty *et al.*, 1992). Targeting efficiency of a nicked circular vector has also been compared against the targeting efficiency of an insertion type vector where a double strand break is included in the region of homology. The targeting efficiency was 10 fold

lower with nicked DNA whereas supercoiled DNA targets with a 34 fold lower efficiency (Hasty *et al.*, 1992).

It is worth noting that many of these results are based upon the targeting of the mouse *hprt* gene encoding hypoxanthine-guanine phosphoribosyltransferase. Since *hprt* is X-linked, in XY ES cell lines there is only one *hprt* allele and targeted clones can be directly selected for the lack of *hprt* function in medium containing 6-thioguanine (Melton *et al.*, 1984). This allows the parameters affecting recombination to be assessed with ease. Although the findings from targeting the *hprt* gene are applicable to other loci, absolute efficiencies are rarely comparable.

A parameter not connected to the vector design which can affect the efficiency of the experiment is the means in which the vector is introduced into the ES cell. Two modes of transfection have been employed; nuclear microinjection and electroporation. Direct injection of DNA into the nucleus of cells is a very efficient method of transfection, with up to 20% of the cells integrating the injected DNA (Capecchi, 1980). This method of introducing DNA has been employed in the targeting of the *Hox-a7 (Hox 1.1)* allele (Zimmer and Gruss, 1989). A recombination efficiency of 1 in 150 injected cells was obtained when injecting a modified 1.5 kb *Hox-a7* fragment into ES cells. This technique is advantageous in that transfection efficiency is high and the ES cells do not have to be placed in drug selection. However, technical constraints such as requirement of injection expertise and costly injection equipment, plus the fact that injection of large numbers of cells is highly laborious and time consuming mean this is not a favoured technique for transfection.

Electroporation of ES cells is easy to carry out and has the advantage that a large number of cells can be handled with reproducible results and hence this has become the chosen means of transfection. The disadvantage of the system is that transfection efficiency is relatively low as compared to microinjection, ranging from 1×10^{-3} to 1×10^{-4} (stably transfected cells/cells surving transfection) and lower. The DNA is transfected into the cell by application of a high voltage pulse which introduces pores in the cellular membrane (Shigekawa and Dower, 1988). Even under optimal conditions this can result in approximately 50% cell death. Parameters which can influence the success of electroporation as a means of transfection include ion concentration, DNA and cell concentration.

The concentration of the targeting vector has also been tested as a parameter affecting recombination efficiency. However, an increase of targeting vector concentration from 5-100 molecules per cell does not increase efficiency (Thomas *et al.*, 1986). This would suggest that locating the homologous chromosomal site is not the limiting step of the reaction.

1.2.6ii Enrichment and screening strategies for homologous recombination events

Several other factors apart from targeting efficiency should also be considered in vector design. As discussed above homologous recombination frequency is generally low in mammalian cells as compared to random integrations. A means of selecting for transformants, enriching for the rare targeting event and screening for its presence are therefore required.

Generally transformants are selected for using dominant positive selectable markers genes such as the gene encoding neomycin phosphotransferase from the bacterial transposon Tn5, *neo,* (Southern and Berg, 1982) and the *hph* gene encoding hygromycin B phosphotransferase (Santerre *et al.*, 1984). These confer resistance to toxins G418 (neomycin) and hygromycin B, respectively. For the selectable marker genes to be effective, they are required to be expressed and translated in ES cells. Promoter-enhancer combinations which are active in ES cells have been defined and include the phosphoglycerate kinase I (PGK-I) promoter, RNA polymerase II (pol II) promoter and β-actin promoter (Hasty and Bradley, 1993; Soriano *et al.*, 1991). The mutated polyoma virus enhancer-HSV1 *tk* promoter combination has also been used to express selectable marker genes in ES cells (Thomas and Capecchi, 1987), although the expression with this enhancer-promoter combination is more variable. The transcriptional status of the target locus in the ES cell is an important consideration when choosing which enhancer-promoter combination to use, as can be gathered from the following example. Recombination of a weakly transcribed selectable marker gene into a transcriptionally silent locus may suppress the expression of the marker. Cells in which the marker integrates randomly into transcriptionally active regions of the genome will survive, however recombination into the desired locus will not be detected.

Transfection followed by only positive selection will result in both random

insertion events and recombination events being selected. As insertion events occur more frequently in mammalian cells, it is likely insertional clones will predominate over clones in which homologous recombination has occured and hence a large number of clones must be screened to identify the legitimate recombinations. However, recent improvements in clone handling (Wurst and Joyner, 1993) and DNA isolation (Laird *et al.*, 1991) make the screening of large numbers of ES clones by Southern blot more practical. The improvement of recombination efficiencies using larger isogenic vectors has also made the use of positive selection alone more practical.

An adaptation of positive selection is to make it conditional on insertion into the correct locus, therefore enriching for homologous recombination. The principle of this method is to place the positive selectable marker in the vector so that its expression is dependent on the use of either an endogenous promoter or polyadenylation site, this only being possible when the vector correctly recombines (Joyner *et al.*, 1989; Zijlstra *et al.*, 1989; Schwartzberg *et al.* 1990; Donehower *et al.*, 1992). Enrichment with this technique can be in the order of 100-fold. Disadvantages include the requirement of the target gene to be expressed in the ES cells. There is also the possibility that the vector may insert into another active gene and, depending on the precise vector design, may allow expression of the selectable marker resulting in false positives. Recent modifications of this type of vector make use of viral internal ribosome entry sites (IRES) (Pelletier and Sonenberg, 1988). This has been found to improve the translation of the selectable marker in poorly transcribed loci by increasing the translation efficiency of the marker gene mRNA (Mountford *et al.*, 1994). This type of targeting construct is known as a conditional targeting vector.

A second enrichment strategy takes advantage of the fact that non-homologous sequences outside the homologous arms are lost upon a correct recombination. Placing a negative selectable marker in this position will allow for random insertions, which include this marker, to be selected against, whereas clones which have undergone legitimate recombination will not be affected. This strategy is known as positive-negative selection (PNS) and is shown in Figure 4 (Mansour *et al.*, 1988). The degree of enrichment with this type of vector, ranges from around 2 to 20 fold although up to 200 fold enrichment has been described (Mansour *et al.*, 1988). Enrichment will vary according to the efficiency with which the vector legitimately recombines and is

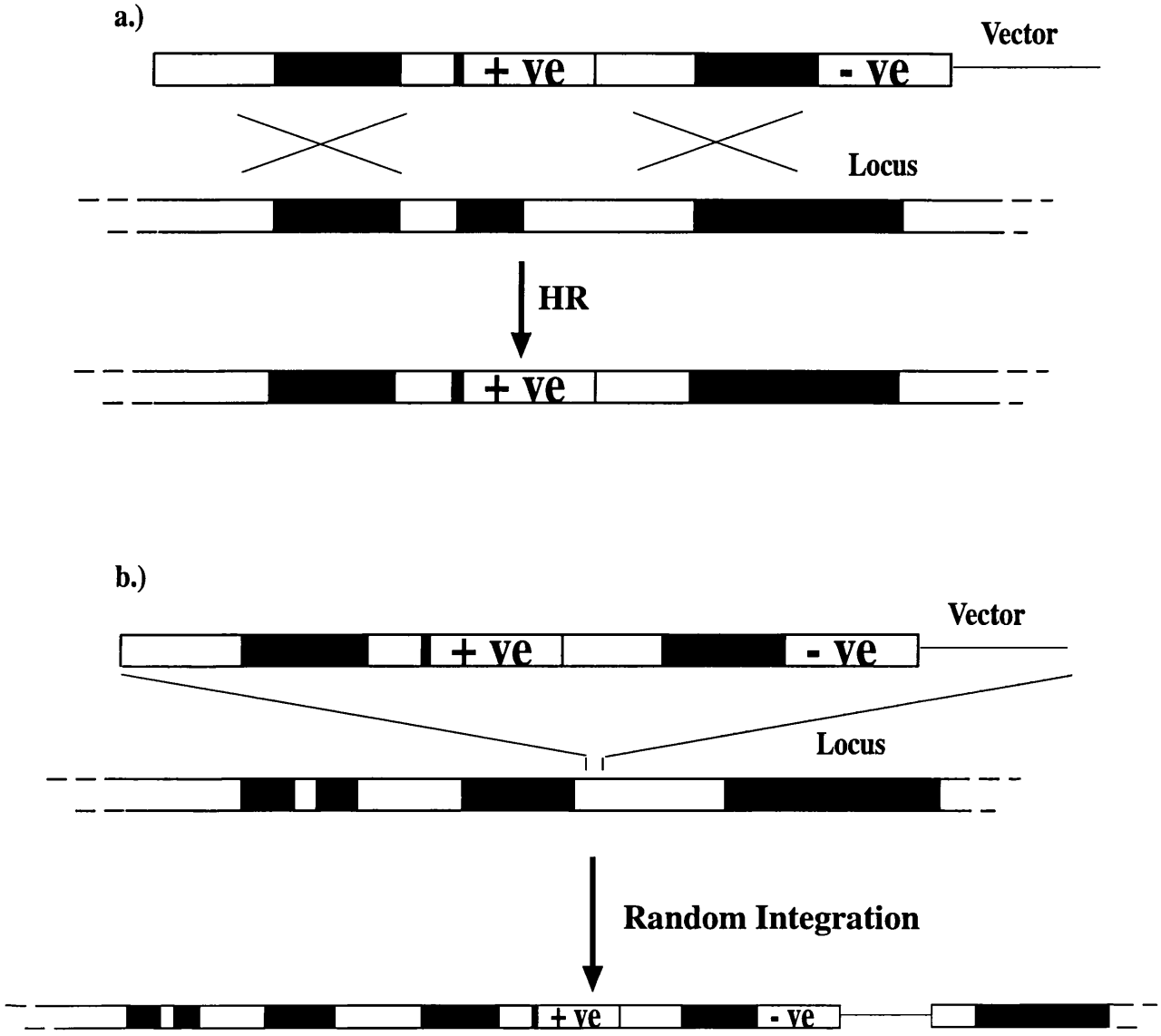
Figure 4. Positive/negative selection strategy for homologous recombinants

a.) Homologous recombination of the targeting vector with the target locus. The positive selection marker is recombined into the target locus. The negative selection cassette is lost upon correct recombination. The loss of negative selection allows survival of the cells in which the homologous recombination event has occurred.

b.) Integration of the targeting vector into a nonhomologous site. Random integration results in the vector inserting into a random site in the genome. The negative selection cassette is included in the vector integration and therefore cells in which this event has occurred will be selected against.

Open boxes represent introns, filled boxes represent exons and single lines represent plasmid sequences. HR, Homologous recombination, +ve, positive selection cassette such as the bacterial neomycin resistance gene. -ve, negative selection cassette such as HSV1 *tk* or diptheria toxin A-fragment gene.

Adapted from Mansour *et al.*, 1988.


a.)
Vector
+ ve
- ve
Locus
HR
+ ve
b.)
Vector
+ ve
- ve
Locus
Random Integration
+ ve
- ve

dependent on other factors such as length of homology. However, false positives can occur when the negative selectable marker is inactivated by point mutation or partial deletion before intergrating into the genome.

The herpes simplex virus type 1 thymidine kinase gene (HSV1*tk*) is commonly used as a negative selection marker (Mansour *et al.*, 1988). Expression of this gene confers sensitivity to the nucleoside analogue chain terminators gancyclovir (GANC) and 1(1-2-deoxy-2-fluoro-β-darabinofuransyl)-5-iodouracil (FIAU) (Mansour *et al.*, 1988; McMahon and Bradley, 1990). The diphtheria toxin A-fragment (DT-A) has also been used for negative selection (Yagi *et al.*, 1990).

Other modifications to the vector include addition of hairpin loops to the ends of the vector, described as "splinkers" (Bernet-Grandaud *et al.*, 1992). These are thought to aid the enrichment of homologous recombination by preventing extrachromosomal recombination and exonuclease digestion of PNS vectors prior to recombination.

Screening of the targeted clones can be carried out by Southern blot analysis solely or in combination with the polymerase chain reaction (PCR) (Mullis and Faloona, 1987). When using PCR the primers are designed to amplify sequences over unique vector-target junctions (Kim and Smithies, 1988). The ability of the PCR technique to detect very low concentrations of DNA allows pools of ES cell clone DNA to be made and screened. Once a positive pool is identified the clones from that pool can then be expanded to isolate the correctly recombined clones (Joyner *et al.*, 1989; Zijlstra *et al.*, 1989).

1.2.7 Strategies for gene modification

Depending on the design of the vector a number of different chromosomal modifications can be made including gene inactivation, locus deletions or introduction of single base pair mutations (reviewed by Bronson and Smithies, 1994; Hasty and Bradley, 1993). How these modifications are achieved is discussed below.

1.2.7i Gene inactivation

To date gene inactivation has been the main application of gene targeting, using

mainly replacement type vectors to create a null mutation in the gene of interest. In this type of vector the disrupting DNA, normally the selection marker, is placed within the vector so as to prevent the normal protein product of the gene being formed. Modifications of this approach have been to include a reporter gene, namely the bacterial *lacZ* gene, in frame with the coding region of the gene of interest (Mansour *et al.*, 1990; Le Mouellic *et al.*, 1990 and 1992; Schneider-Manoury *et al.*, 1993). This allows the cells that would normally express the target gene to be visualised by staining for the product of *lacZ*, β-galactosidase, using a chromogenic substrate such as 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). In the context of a conditional selection vector, a *lacZ-neo* fusion gene can be employed (Fredrich and Soriano, 1991; Mountford *et al.*, 1994) Apart from making null mutations with replacement vectors, larger genomic disruptions can be made by employing deletion vectors (Zhang *et al.*, 1994) (discussed in section 1.2.6ii).

A further modification of replacement vector targeting is the technique described as "plug and socket" (Detlott *et al.*, 1994). The strategy allows repeated targeting of the same allele with high enrichment. The first targeting event is performed as a normal replacement experiment. However, the vector not only carries a normal selection marker e.g. *neo* but also two thirds of an incomplete second positive selectable marker. Once such a construct has been targeted into the desired location a second targeting experiment can be carried out. One recombination arm of the incoming vector is homologous to part of the disabled marker already present in the target site and contains the remaining third, allowing the complete sequence of the second selectable marker to restored upon correct recombination. The other recombination arm of the second vector can be homologous to sequences much further from the original targeting event thus facilitating a larger deletion. This technique allows the repeated targeting of a single locus with a high degree of enrichment once the first targeting has been achieved.

1.2.7ii Subtle mutations

As discussed earlier both gross and subtle mutations can be made into the genome. In many cases it is more interesting to subtly modify the action of a gene product rather than to completely ablate it. One obvious application is to mimic human

genetic diseases where often the mutation is more subtle than complete gene inactivation, as in the case of cystic fibrosis (Kerem *et al.*, 1989).

The simplest way in which to produce a more subtle modification is to use a replacement type vector where the selectable marker gene is placed either upstream of the promoter or downstream of the polyadenylation site of the target gene. A point mutation is made in the region of homology to the gene. This method has been successfully used in culture to change the human sickle cell β^Sglobin to a normal β^A (Shesely *et al.*, 1991) and to create a single amino acid change in the proopiomelanocortin gene in ES cells (Rubinstein *et al.*, 1993). One obvious disadvantage with this strategy is that the selectable marker gene is still left in the targeted locus. This may either disrupt cis-acting regulatory sequences, disrupt another gene in the same locus or affect expression of the target gene itself via the enhancer elements of the selectable marker gene.

One way around this is to employ a double or repeated targeting strategy as described by Askew *et al.* (1993) in their "tag and exchange" protocol and by Stacey *et al.* (1994) in their "double replacement" method. In both cases the basis of the protocol is the same. The first targeting event uses a replacement type vector which contains both positive and negative selection markers, but only positive selection is employed. In the second targeting event a vector without selectable markers but with a mutation in the homologous region is used. Upon correct recombination both selectable markers will be replaced in the recombination and this event can be selected for by the loss of the negative marker e.g. HSV1tk (Askew *et al.*, 1993). Stacey *et al.* (1994) used the same strategy but instead of two selectable marker genes a *hprt* minigene was used in conjuction with the $HPRT^-$ ES cell line E14TG2a (Hooper *et al.*, 1987). HPRT can be positively selected for using hypoxanthine-aminopterin-thymidine and negatively selected for using 6-thioguanine.

A third strategy named "hit and run" (Hasty *et al.*, 1991c), or "in and out" (Valancius and Smithies, 1991), uses insertion type vectors. As described above, these vectors produce a duplication of the target locus. In the case of "hit and run" the vector is designed with a desired mutation in the region of homology and positive and negative marker genes in the vector backbone. The initial recombination event can be positively selected for, resulting in a duplication of the locus in which one of the duplicated

regions will carry the mutation. A second rare intrachromosomal event between the duplicated regions on the same chromosome or between sister chromatids, which removes the one copy of the duplicated region and markers, can be selected for by loss of the negative marker. This recombination will result in either restoring the locus to wild type or inclusion of the mutated sequence. This technique has successfully been used to generate subtle modifications in the Hoxb-4 gene (Ramirez-Solis *et al.*, 1993), IgH locus (Serwe and Sablitzky, 1993) and the cystic fibrosis transmembrane conductance regulator gene (van Doorninck *et al.*, 1995).

A number of potential disadvantages accompany these strategies. Firstly, the negative selectable marker may be inactivated due to mutation rather than recombination. Secondly, the techniques require prolonged culture of the ES cells increasing the chance that germline potential is lost. Germline transmission can be tested for after the first targeting. However, in the case of the "tag and exchange" and "hit and run" protocols, expression of the negative marker HSV1tk in the testes can result in males with reduced fertility (Braun *et al.*, 1990). Therefore testing the efficacy of the ES cells after the first targeting event is, in these cases not meaningful.

1.2.7iii Conditional mutations

In some cases full analysis of a targeted mutation may not be possible as the disruption leads to a lethal phenotype early in development. Although of interest, it obviously prevents analysis of the mutation in the adult animal. It would therefore be desirable to inactivate a gene in a particular tissue type or after a particular developmental time point. The recent application of site specific recombinase technology in combination with gene targeting and transgenic techniques, should facilitate the production of spatial or temporal conditional mutants (reviewed by Kilby *et al.*, 1993). Site-specific recombinases cleave specific DNA target sequences and then either religate the ends or ligate them to the cleaved DNA of a second site. This mechanism can be used to excise sequences lying between two recombinase recognition sites or to introduce sequences into a recombinase site. Both the yeast FLP-FRT system (Jung *et al.*, 1993; Fiering *et al.*, 1993) and P1 bacteriophage Cre-loxP (Gu *et al.*, 1993) system can function in ES cells. In both these cases the recombinases Cre and FLP

recognise target sequences of 34bp which are made up of two inverted repeats. The Cre-loxP system has also been shown to function in transgenic mice (Lakso *et al.*, 1992; Orban *et al.*, 1992). Several applications for these systems have been suggested. These include removal of a selectable marker gene, leaving only the relatively small recombinase sequence behind (Jung *et al.*, 1993), and spatial/temporal gene inactivation or activation. The latter can be achieved by making the expression of the recombinase dependent on a specific promoter/enhancer in a transgenic animal. Crossing this animal with one that contains targeted recombinase sites placed to either excise the gene of interest or remove disrupting sequences will allow the production of conditional loss or gain of function mutants. It may also be possible to direct interchromosomal recombination with these systems allowing human diseases where translocations are involved to be mimicked in transgenic mice.

1.2.8 Alternative approach to gene targeting in the mouse

Although gene targeting in the mouse using ES cells has now become a standard procedure, the transmission of the mutation into the germline can still be a major hurdle. It would be highly desirable if the mutation could be made directly in the oocyte using transgenic procedures. This was attempted by Brinster *et al.* (1989). The authors attempted to correct a mutation in the major histocompatibility class II E gene. Four fragments ranging in size from 2.5 kb to 8.9 kb were used. From 10,602 oocytes injected 1,841 animals were born of which 506 were transgenic. Of these animals only one was shown have undergone homologous recombination at the targeted loci. This study shows that this type of experiment is possible although very arduous. The absolute recombination efficiency can not be calculated since only one homologous recombination event was found. However, it is worth pointing out that a number of polymorphisms existed between the target locus and fragment, and according to studies discussed earlier this will affect recombination efficiency. It would be interesting to see if a gene which targets at high efficiency in ES cells could also be targeted in oocytes. A greater understanding of the recombination machinery in mammalian cells may make a gene targeting in approach in oocytes more feasible in the future.

1.2.9 Phenotypes of mice with targeted alleles-general considerations

To date a large number of targeted mouse mutants have been created. These have provided further insight into a diverse range of fields, from immunology to developmental biology and neurobiology. In many cases the mutant mice exhibit phenotypes which are in accordance with the predicted role of the gene product. However, targeted mutations have resulted in less servere phenotypes than expected, unexpected phenotypes or no discernable phenotypes. Examples are given which illustrate these points.

With regards to the first criterium, many genes when disrupted gave an expected phenotype in mutant mice. For instance disruption of the RAG-1 and -2 recombinase genes, which are involved in the rearrangement of the V(D)J gene segments of antibodies and T-cell receptors, result in mice which do not have mature T or B cells (Mombaerts *et al.*, 1992; Shinkai *et al.*, 1992). Mice deficient in the lymphoid specific $p56^{lck}$ tyrosine kinase have a block in T-cell development, consistent with its proposed role as a signal transduction molecule involved in T-cell maturation (Molina *et al.*, 1992) and targeting of the immunoglobulin μ chain resulted in B-cell deficient mice (Kitamura *et al.*, 1991). As expected targeted mutations in the mouse *Hox* homeobox genes result in various homeotic transformations consistent with their proposed role in anterior-posterior positional identity (reviewed by Krumlauf, 1994). Mice deficient in myogenin show a severe reduction of all skeletal muscle which is consistent with its presumed function as a muscle-specific transcription factor that can induce myogenesis (Hasty *et al.*, 1993; Nabeshima *et al.*, 1993).

However, there are several examples where the loss of function either leads to no phenotype or to a less severe phenotype than expected. Examples of this include β_2-microglobulin (Zijlstra *et al.*, 1989), interferon-γ (Huang *et al.*, 1993a) and the T-cell surface glycoprotein CD2 (Killeen *et al.*, 1992). These experiments show that either the proposed function of a gene based on its expression pattern and protein structure may not always be true or functional homologs may be present suggesting the existance of redundancy. An example of the latter is shown from the targted disruptions of MyoD (Rudnicki *et al.*, 1992) and Myf-5 (Braun *et al.*, 1992) genes. Both of these genes are thought to act as transcriptional regulators during myogenesis, but when they are

disrupted no obvious defects in skeletal muscle formation is found. However, upon crossing the two targeted mouse lines, the double disruption of MyoD and Myf-5 leads to no muscle formation, suggesting that MyoD and Myf-5 share an overlapping function (Rudnicki *et al.*, 1993).

The genetic background in which the mutation is bred can sometimes affect the severity of the disruption. A dramatic example is seen in mice homozygous for a mutation in the keratin-8 gene. When on a 129 x C57BL/6 background 94% of the homozygous mutant animals die between embryonic day 12 and 13, due to internal bleeding and liver defects. However 6% of the animals survive to adulthood (Baribault *et al.*, 1993). When the mutation is made congenic for the FVB/N strain, the majority of animals survive to adulthood but suffer from colorectal hyperplasia and inflammation (Baribault *et al.*, 1994). The authors conclude that genetic modifiers exist which are differentially active or present in the different strains and act to rescue the mutation.

Targeted disruption of the *Hoxc8 (Hox-3.1)* gene (Le Mouellic *et al.*, 1992) leads to death of the animal within the first few days after birth, however a number of homozygous animals for this disruption survive into adult hood. These animals are smaller than their normal litter mates and also have abnormal reflexes and spasms. The authors suggest that the survival may be due to the complex genetic background (129 x C57BL/6 x CBA) in which the homozygous animals are bred. The effect of the genetic background on a phenotype is an important consideration when studying targeted mutations. Analysis of phenotype differences between strains can potentially allow the identification of genes encoding functional homologues of the targted gene or of genes which are involved in compensatory mechanisms.

1.2.10 Gene targeting in the nervous system-nervous system development

The development of the nervous system can be roughly divided into four stages; neural induction, regionalisation of the neuroepithelia and patterning, proliferation and differentiation of neuronal precursors and survival and terminal differentiation of neuronal cell types (for reviews see Altaba, 1993; Krumlauf *et al.*, 1993; Goodman and Shatz, 1993). In the mouse neurogenesis begins around embryonic day 7.5 and carries on into postnatal development with the continuing formation and refinement of synaptic

connections. A number of targeted mutants which affect the formation and function of the nervous system have been made (reviewed by Joyner and Guillemot, 1994; Grant, 1994; Grant and Silva, 1994) examples of which are given below.

Null mutations for the *Wnt-1*(Thomas and Capecchi, 1990; McMahon and Bradley, 1990; McMahon *et al.*, 1992), *En* (Joyner *et al.*, 1991; Millen *et al.*, 1994; Wurst *et al.*, 1994), *Hoxa-1* (Chisaka *et al.*, 1992; Lufkin *et al.*, 1992; Carpenter *et al.*, 1993; Dolle *et al.*, 1993; Mark *et al.*, 1993) and *Krox-20* genes (Schneider-Manoury *et al.*, 1993; Swiatek and Gridley, 1993) affect regionalisation and patterning of the neuroepithelia.

The *Wnt-1(int-1)* gene encodes a putative signalling molecule expressed in the developing midbrain at e8.5 day (Wilkinson *et al.*, 1987) and shows a similar distribution to that of the mouse homeobox-containing *En (engrailed)* genes, *En-1* and *En-2* (Davis and Joyner, 1988; Davis *et al.*, 1988). Both *Wnt-1* and *En* genes have *Drosophila* homologues, namely *wingless* and *engrailed.* In *Drosophila* it has been suggested that *engrailed* is expressed as a result of *wingless* signalling (Bejsovec and Martinez-Arias, 1991; Heemskerk *et al.*, 1991). Disruption of *Wnt-1*(*int-1*) results in a dramatic disruption of midbrain formation and absence of the cerebellum (Thomas and Capecchi, 1990; McMahon and Bradley, 1990) which could first be detected as early as e8.5 based upon a decreased domain of *En-1* expression and complete loss of both *En-1* and *En-2* expression by e9.5 (McMahon *et al.*, 1992). In *En-1* mutants a smaller deletion of mid-hindbrain tissues occur as compared to *Wnt-1*$^{-/-}$ embryos, this correlates with a decrease but not complete loss of *En-2* expressing tissue (Wurst *et al.*, 1994). A less servere phenotype is described for mutations in the *En-2* homeobox gene, disruption of which results in abnormal formation of the cerebellum and colliculi, affecting foliation patterning of the cerebellum (Joyner *et al.*, 1991; Millen *et al.*, 1994). These results have suggested that *En-1* and *En-2* are required for the formation of overlapping regions of the mid-hindbrain border which are dependent on *Wnt-1* expression (McMahon *et al.*, 1992; Wurst *et al.*, 1994). However, evidence has yet to be provided that shows *En* expression to be a direct consequence of *Wnt-1* signalling in the mouse.

Other homologs of *Drosophila* pattern formation genes which have been disrupted include the homeobox gene *Hoxa-1* and the zinc finger transcription factor gene *Krox20*, both mutations affect the neuromeric organisation of the hind brain. *Hoxa-*

1 is normally expressed with an anterior boundary at the junction of rhombomeres 3 and 4. The disruption of *Hoxa-1* affects the cellular identity of rhombomeres 4-7 as shown by abnormal expression pattern of other hindbrain markers such as *Krox-20* and *int-2*. The positioning of cranial ganglia is also affected (Chisaka *et al.*, 1992; Lufkin *et al.*, 1992; Carpenter *et al.*, 1993; Dolle *et al.*, 1993; Mark *et al.*, 1993). The *Krox20* disruption results in apparent fusion of rhombomere 3 and 5, the normal site for *Krox20* expression, and loss of rhombomere 4. Again the cranial ganglia are affected and mice homozygous for the mutation die within the first two weeks of birth (Schneider-Manoury *et al.*, 1993; Swiatek and Gridley, 1993). In the above examples, the targeted mutations led to defined regions of the developing neuroepithelia being perturbed. This is concurrent with the postulated functions of the genes described above, i.e. that their action defines specific regions of the developing brain.

Mutations which appear to have a more general effect on proliferation and differentiation of neural precursors, rather than affecting defined regions include mutations in the murine homolog of the *Drosophila* basic-helix-loop-helix transcription factor *achaete-scute* (*Mash-1*) (Guillemot *et al.*, 1993), the retinoblastoma susceptibility gene, *Rb* (Lee *et al.*, 1992; Jacks *et al.*, 1992; Clarke *et al.*, 1992), and the proto-oncogene *c-ret* tyrosine kinase receptor (Schuchardt *et al.*, 1994). Affected neuronal systems include PNS defects ie sympathetic ganglia in the case of *Mash-1*$^{-/-}$ mice and the enteric nervous system in the case of *c-ret*$^{-/-}$ mice. Mutation of the *Rb* gene, as well as affecting the developing haematopoietic system, results in cell death in various regions of the CNS and PNS. Ectopic cell division is seen in the neural tube where differentiating rather than dividing neurons are normally located.

A series of targeting experiments have been carried out which disrupt the neurotrophin genes, neuronal growth factor (NGF) (Crowley *et al.*, 1994), brain derived neurotrophic factor (BDNF)(Ernfors *et al.*, 1994a; Jones *et al.*, 1994), neurotrophic factor 3 (NT-3)(Ernfors *et al.* 1994b) and their high affinity tyrosine kinase type receptors TrkA (Smeyne *et al.*, 1994), TrkB (Klein *et al.*, 1993) and TrkC (Klein *et al.*, 1994) and also the low affinity NGF receptor p75 (Davies *et al.*, 1993). The results obtained from the above gene disruptions are consistent with the proposed role of these neurotrophins and their receptors in neuronal survival. In many cases the observed neuronal loss is in accordance with the distribution of receptor and growth factor.

However, in some areas a less servere phenotype is observed than would be predicted, it has been suggested that this is due to the potential cooperative functions of the growth factors (reviewed by Snider, 1994). Targeting of the various neurotrophins and their receptors has highlighted the role these molecules play in survival of neuronal cell types.

1.2.11 Gene targeting in the nervous system-learning

One particularly interesting application of gene targeting in the mouse is to attempt the molecular dissection of higher neuronal functions such as behaviour and learning (reviewed by Grant, 1994; Grant and Silva, 1994). The present theory on the cellular basis of learning suggests that compounded stimuli can affect the efficiency of synaptic transmission, an activity dependent synaptic plasticity (reviewed by Bliss and Collingridge, 1993). The most studied example of this is the phenomena of long-term potentiation (LTP) in the hippocampus, a cortical structure associated with learning. Upon a brief train of high-frequency stimuli to presynaptic fibres a sustained enhancement in synaptic transmission results which can last from hours to weeks, this phenomenon is termed LTP. The tetani which induce LTP cause the release of the excitatory neurotransmitter glutamate which can activate the N-methyl-D-aspartate (NMDA) receptor. The associated depolarisation of the membrane following the high frequency stimulation and the associated stimulation of the neighbouring synapse result in the removal of blocking Mg^{2+} ions at the NMDA receptor. The coupled activation of the NMDA receptor and removal of Mg^{2+}causes an influx of Ca^{2+} ions, triggering second messenger cascades and signalling events required for the enhanced synaptic transmission.

Proteins involved in the induction of LTP have been the subject of gene targeting experiments. Targeting of the NMDA receptor unfortunately resulted in mice which died at postnatal day 2 and it was therefore not possible to study the effect of the disruption on LTP (Li *et al.*, 1994a). However the homozygous mutant mice did show defects in the refinement of their synaptic connections. Disruption of protein kinases *c-fyn* (Grant *et al.*, 1992), α-calcium calmodulin kinase II (αCamKII) (Silva *et al.*, 1992) and protein kinase Cγ (PKCγ) (Abeliovich *et al.*, 1993) have all resulted in an altered threshold

levels for hippocampal LTP. The mice, homozygous for the above mutations, also exhibit spatial learning deficits. Targeting of the α and δ isoforms of the cAMP-responsive element binding protein (CREB) (Bourtchuladze *et al.*, 1994) has also been shown to affect LTP response. This is of particular interest as this result implicates cAMP-dependent transcription in the process of mammalian long term memory.

The above disruptions clearly show that learning can be affected by the disruption of particular genes, however it is unclear whether the disrupted LTP is a result of a developmental defect in these mice or direct perturbation of the cellular mechanisms involved in LTP. The application of conditional gene inactivation would allow this distinction to be made.

1.2.12 Gene targeting in the nervous system-neuronal cell adhesion molecules

Pertinent to this thesis is the inactivation of neuronal cell adhesion molecules by gene targeting. Cells of the mammalian nervous system migrate large distances and form highly specific connections. These processes are dependent on the interaction of the cell with its enviroment (reviewed in Goodman and Shatz, 1993; Goodman, 1994). One group of molecules involved in cell-cell interactions are the cell adhesion molecules (CAMs). CAMs are thought to be involved in axonal guidance, fasciculation and interactions of neurons with glia. The CAMs can be subdivided according to their structural protein domains into immunoglobulin like molecules, integrins and cadherins (reviewed in Muller and Kypta, 1995). Applicable to this study are the targeting of P_O, neural cell adehesion molecule (NCAM) and myelin associated glycoprotein (MAG) which, like Thy-1, all possess immunoglobulin-like (Ig-like) domains and are thought to behave as recognition and adhesion molecules in the nervous system.

P_O is a major glycoprotein of PNS myelin, accounting for 60% of protein in the peripheral myelin sheath (Lemke and Axel, 1985). It possesses one Ig-variable domain, but in contrast to Thy-1 it is linked to the cell surface via a transmembrane domain. Homophilic interactions between P_O proteins are likely to be involved in myelin compaction. Heterophilic interactions are postulated to play a role in axonal ensheathment by Schwann cells (Schneider-Schaulies *et al.*, 1990). Targeted disruption of the P_O gene results in hypomeylination and axonal atrophy (Giese *et al.*, 1992). The

mice homozygous for the P_0 mutation are deficient in normal motor coordination and suffer tremors and convulsions. Disruption of the P_0 protein has shown the important role this adhesion molecule plays in myelination and hence nervous system function.

NCAM contains five Ig-like domains and also two fibronectin type III domains. Through alternative splicing three predominant isoforms are generated with molecular weights of 120 kDa, 140 kDa and 180 kDa. The isoforms have different cytoplasmic domains and a glycosyl phosphatidylinositol (GPI) membrane anchored form also exists. Two targeted mutations have been made in the NCAM gene. One inactivates NCAM-180 (Tomasiewicz *et al.*, 1993), the other produces a complete null allele (Cremer *et al.*, 1994). Although NCAM is widely distributed in the CNS and has been implicated in a variety of functions including axonal growth, guidance and fasciculation, the phenotypes of both disruptions are surprisingly mild. In both mutations the olfactory bulbs are strikingly smaller as compared to wildtype animals due to a reduction in granule cells. This reduction has been shown to result from a migratory deficit of the granule cell precursors (Tomasiewicz *et al.*, 1993). The hippocampus is also affected and the mice display spatial learning problems, a function thought to be associated with the hippocampus. Interestingly, the deficits in these mice are seen in the regions of the brain where the predominant form of NCAM is posttranslationally modified with polysialic acid (PSA). This highly charged sugar moiety is predominantly associated with NCAM in the brain. Enzymatic removal of PSA results in a phenocopy of the targeted mutations (Ono *et al.*, 1994). This suggests that it is the lack of PSA modified NCAM that results in the observed phenotype in NCAM deficient mice.

MAG is linked to the cell surface via a transmembrane domain and comprises five Ig-constant domains. Like P_0 , MAG is associated with the myelin sheath, particularly the non-compacted regions around the axonal membrane. Disruption of MAG leads to a more subtle phenotype than in the case of P_0 disruption, the mice displaying mild trunk tremors, although overall motor coordination appears unaffected. At the ultrastructural level there are some abnormalities including abnormal formation of the axonal-cytoplasmic myelin collar and excess myelination (Li *et al.*, 1994b; Montag *et al.*, 1994). This subtle phenotype may be the result of other adhesion molecules substituting for MAG function. In support of this, there is an increase in NCAM at sites normally expressing MAG (Montag *et al.*, 1994). Intercrossing mice

homozygous for the targeted CAMs could therefore result in more severe phenotypes.

Application of gene targeting to address the role of Ig-like CAMs has therefore gone some way in identifying their function. However, the targeted disruptions have also highlighted the possibility of overlapping function of CAMs and the need to consider other properties of the molecules, such as their glycosylation pattern.

1.2.13 Summary

Gene targeting has proved to be a powerful technique in dissecting gene function. The emergence of technical refinements in vector design in combination with established transgenic technology should result in more specific mutations, whether subtle mutations altering single nucleotides or directed mutations inactivating genes in a particular spatial or temporal context. Results from targeted genes have highlighted the existence of strain specific effects and unexpected phenotypes. Results from targeting experiments have also shown that in many cases functional homologs are present which can fully or partially alleviate the effect of the disruption. It is therefore important to realise that a lack of phenotype from targeting a mutation in a particular gene does not mean that the gene has no function. It is more likely that the function can be compensated by functional homologs or by compensatory mechanisms in the pathway in which the gene features. Identification and targeting of other members of the pathway and functional homologs will therefore allow complex systems such as LTP to be unravelled at the molecular level.

1.3 Murine Thy-1.2

1.3.1 Discovery of Thy-1

Thy-1 was first described by Reif and Allen (1964) as an allelic antigen on the surface of mouse thymocytes. In mice it is found predominantly on the surface of T-lymphocytes and mature neurons (Reif and Allen, 1964; Reif and Allen, 1966a and b). The existence of the molecule was first implied through the production of strain specific antisera produced when immunising one mouse strain with thymocytes from another.

Immunisation of the C3H strain of mice with thymocytes from AKR mice resulted in antisera which reacted strongly against AKR thymocytes but not against C3H thymocytes (or those of other inbred strains). Therefore the implied antigen, which was named theta-AKR and theta-C3H, existed in two allelic forms. The theta antigen was later renamed as Thy-1 and sub divided into Thy-1.1 which is the allelic marker found on thymocytes of AKR/J, RFM/Un and RIII/Jem mice and Thy-1.2, the allelic marker found on thymocytes from C3H, RF, AKR/Cum and all other mouse strains tested (Reif and Allen, 1966a; Acton *et al.*, 1973; Schlesinger and Hurvitz, 1969; Morris, 1985).

Thy-1 was subsequently shown to exist at high levels in the adult brain but at only 1% of this level in the developing neonatal brain. The dramatic rise in the level of the antigen in the CNS was seen during the second and third week of postnatal development (Reif and Allen, 1964; Reif and Allen, 1966b). Thy-1 homologues have been described for a diverse variety of species including rat (Douglas, 1973), hamster (White and Streilein, 1983), man (Cotmore *et al.*, 1981; Dalchau and Fabre, 1979; Hamann *et al.*, 1980), dog (Dalchau and Fabre, 1979; McKenzie and Fabre, 1981a), sheep (Sheldrake *et al.*, 1985), chicken (Rostas *et al.*, 1983), frog (Mansour and Cooper, 1984a), squid (Williams and Gagnon, 1982) and tunicate (Mansour and Cooper, 1984b).

1.3.2 Tissue distribution of Thy-1

Thy-1 is a major component of the surface of mature nervous tissue, a feature which is conserved for all species that it has been described in (for reviews see Morris, 1985; Morris and Grosveld, 1989). In the rat and mouse Thy-1 is also expressed at high levels in the lymphoid system, however the cell type and maturation stage of the Thy-1 bearing cells differs between the two rodents. The molecule can also be found at lower levels on immature muscle (Lesley and Lennon, 1977; Walsh and Ritter, 1981), rat mammary gland (Monaghan *et al.*, 1983), kidney epithelia in man (McKenzie and Fabre, 1981b), kidney glomeruli in rat (Paul *et al.*, 1984), connective tissue in rat (Morris and Beech, 1984), mouse and rat fibroblasts (Stern, 1973) and mouse epidermal cells (Scheid *et al.*, 1972). From this it appears that excluding nervous system expression, Thy-1 shows a high degree of variation in tissue distribution between species (Morris, 1985; Morris and Grosveld, 1989; Morris 1992).

The spatial and temporal expression of Thy-1 in the nervous system has been most extensively studied in the rat and mouse (Xue *et al.*, 1990; Xue *et al.*, 1991; Xue and Morris, 1992; reviewed by Morris, 1985; Morris and Grosveld, 1989; Morris, 1992). Thy-1 expression in nervous tissue has also been described for chicken (French and Jefferey, 1986), dog (Dalchau and Fabre, 1979), human (Mckenzie and Fabre, 1981b; Kemshead *et al.*, 1982) and squid (Williams and Gagnon, 1982).

The appearance of Thy-1 mRNA and protein was documented for several neuronal cell types and their axons during rodent neuronal development (Xue *et al.*, 1990; Xue *et al.*, 1991; Xue and Morris, 1992). These include projections of CNS neurons such as the hippocampal granular and pyramidal neurons, Purkinje cells of the cerebellum (Xue and Morris, 1992) and mitral neurons of the olfactory bulb (Xue *et al.*, 1990). In the PNS, projections of pontine and vestibular ganglion neurons to the cerebellum have been closely examined (Xue *et al.*, 1991). Thy-1 mRNA can be detected as early as mouse embryonic day 13 (e13) for vestibular ganglion neurons, e15 for mitral cells of the olfactory bulb, e16-17 for hippocampal pyramidal neurons, e17 for the Purkinje cells, postnatal day 0 (p0) for pontine neurons and p5 for hippocampal granular cells. In each case the mRNA expression coincides with the completion of neuronal cell migration and beginning of dendritic growth (Xue *et al.*, 1991). However, detection of Thy-1 protein on the surface of these neuronal cell types does not immediately follow the appearance of the mRNA. A delay between detection of mRNA and protein varying from 2 days (mitral cells) to 12 days (vestibular ganglion neurons) is found. The expression of the protein on the cell body and dendrites follows the end of initial axon elongation. Thy-1 is therefore under tight postranscriptional control, its appearance coinciding with the end of initial axonal elongation (Xue *et al.*, 1991). Furthermore, Thy-1 is restricted from the membrane of the growing axon until it has reached its terminal field, but can be found on the cell body and dendrites. In the case of the vestibular ganglion neurons, whose axons sprout into multiple fields over an extended time period, Thy-1 can be found on the non-growing proximal region of the axon but not on the growing region (Xue *et al.*, 1991). How this fine regulation of Thy-1 on the axonal surface is achieved is not known but the end result is to exclude Thy-1 from the growing region of axons. Therefore during neuronal development the appearance of Thy-1 is tightly regulated at the levels of transcription and translation.

The localisation of Thy-1 on the axonal surface is also controlled in a temporal manner.

The only neurons in the adult animal that do not express Thy-1 protein on their axonal surface are the primary olfactory neurons (Barclay and Hyden, 1978; Morris, 1985; Morris and Grosveld, 1989). These cells, located in the olfactory epithelia, are of particular interest since their axons are able to grow into the adult nervous system (Graziadei and Graziadei, 1978; Morris and Barber, 1983). Again this shows Thy-1 to be excluded from growing axons.

The density of Thy-1 on the axonal surface has been determined for CNS axons of the optic nerve (500 molecules/μm^2), for PNS axons of the hypoglossal nerve (1000 molecules/μm^2) and for axons of the preganglionic sympathetic chain (1500 molecules/μm^2) (Beech *et al.*, 1983). Thy-1 is therefore estimated to constitute between 2.5-7.5% of the surface protein of these cells (Beech *et al.*, 1983), showing it to be a major component of the axonal surface in the adult nervous system. Thy-1 expression in the nervous system is restricted to neurons. The glial cells of the nervous system appear negative for Thy-1 throughout their life (Morris and Beech, 1987).

The most extensively studied area of Thy-1 expression outside nervous tissue is that of the rat and mouse lymphoid system. In the mouse Thy-1 is expressed on all stages of T-lymphocytes and at low levels on very early B-cells. However, whether the latter are committed B-cell precursors or haematopoietic stem cells (HSCs) is not well defined (Ritter *et al.*, 1980; Goodwin *et al.*, 1986; Muller-Sieburg *et al.*, 1986). Murine HSCs have been described as Thy-1^{low} (Basch and Berman, 1982; Muller-Sieburg *et al.*, 1986), although recent results suggest that Thy-1 expression is not an invariant characteristic of all murine HSCs. In Thy-1.1 mice bone marrow cells that can repopulate irradiated animals are Thy-1^{low}Ly6A/E^+ whereas Thy-1.2 mice have a population of Thy-1^{neg} bone marrow cells which still have the same potential (Spangrude and Brooks, 1992). This suggests that the distribution of Thy-1 on these cells varies for the different alleles.

Expression of Thy-1 in the rat lymphoid system shares some similarities with the mouse. Both HSCs and lymphoid progenitor cells are Thy-1 positive (Ritter *et al.*, 1978). However, in the rat extended expression into the B-cell lineage is seen as compared to the mouse and also circulating T-lymphocytes appear negative (Acton *et al.*, 1974).

In man Thy-1 is found on Pre-B and T-lymphocytes but not on mature cells (Ritter *et al.*, 1983). However, thymic epithelial cells also appear Thy-1 positive (Ritter *et al.*, 1981).

The conserved distribution of Thy-1 in the nervous system of a number of species suggests a conserved function for the molecule on neurons. The exclusion of Thy-1 from the growing axon and its appearance upon the cessation of growth suggests that Thy-1 plays a role in axonal outgrowth. It remains unclear whether the appearance of Thy-1 causes the cessation of growth or whether completion of growth then allows Thy-1 to appear on the axonal surface. The role of Thy-1 on other tissue types remains more elusive, although Thy-1 has been shown to contribute to the adhesion of mouse thymocytes to the thymic epithelium (He *et al.*, 1991).

1.3.3 Structure of the murine *Thy-1* gene

Thy-1 mRNA was first isolated from rat thymocytes (Moriuchi *et al.*, 1983) and subsequently genomic clones from mouse (Chang *et al.*, 1985; Giguere *et al.*, 1985), rat (Seki *et al.*, 1985a) and human (Seki *et al.*, 1985b) were isolated and sequenced. The gene has been localised to chromosome 9 in the mouse (Itakura *et al.*, 1972) and the long arm of chromosome 11 in man (van Rijs *et al.*, 1985; Seki *et al.*, 1985b).

The exon/intron structure of the mouse *Thy-1* gene has been determined (Giguere *et al.*, 1985). The *Thy-1* gene possess four exons; exon I is non-coding and exists in two alternate forms, Ia and Ib, allowing two alternate transcripts to be produced (Ingraham and Evans, 1986; Ingraham *et al.*, 1986), However, Ib is present in less than 1% of the total message (Spanopoulou *et al.*, 1988). The translation initiation codon is located in exon II, exon II encodes twelve of the 19 amino acids that make up the signal peptide. Exon III encodes the remaining 7 amino acids of the 19 amino acid signal peptide and also 106 amino acids of the extracellular domain. Exon IV encodes the remaining six amino acids of the extracellular domain, a putative transmembrane domain of 31 amino acids and contains the 1.1 kb 3'UTR (Giguere *et al.*, 1985). The different functional domains of the Thy-1 protein are essentially restricted to seperate exons, with the exception of leader sequence, this is a feature shared by other members of the Ig supergene family (Hood *et al.*, 1985). The sizes of the three introns are 1500, 590 and

386 bp respectively. The mRNA is 1850 nucleotides in length as determined by Northern blot and sequence analysis (Giguere *et al.*, 1985). The structure of the murine *Thy-1.2* gene is shown in Figure 5.

Sequence comparison between *Thy-1.1*(AKR/j) and *Thy-1.2* (Balb/c) genomic clones showed a base substitution (G →A) giving rise to the Gln →Arg change at residue 89 accounting for the allelic difference detected by antibodies (Giguere *et al.*, 1985). Further comparison, including RFLP analysis, showed a number of other polymorphisms between *Thy-1.1* (AKR/j), *Thy-1.2* (Balb/c) and *Thy-1.2* (C57BL/6) mouse strains, these included insertions and deletions, both in introns and 5' to the cap site (Giguere *et al.*, 1985; Chang *et al.*, 1985).

The *Thy-1* gene has a number of interesting features associated with its promoter. It lacks a canonical TATA box (Giguere *et al.*, 1985) and is located within a methylation-free CpG-rich island (Bird, 1986; Spanopoulou *et al.*, 1988). This methylation-free island spans a 1.5 kb region, covering exons Ia, Ib and part of intron 1. Transcription initiation can occur from multiple sites, which differ between murine brain and thymus (Spanopoulou *et al.*, 1988).

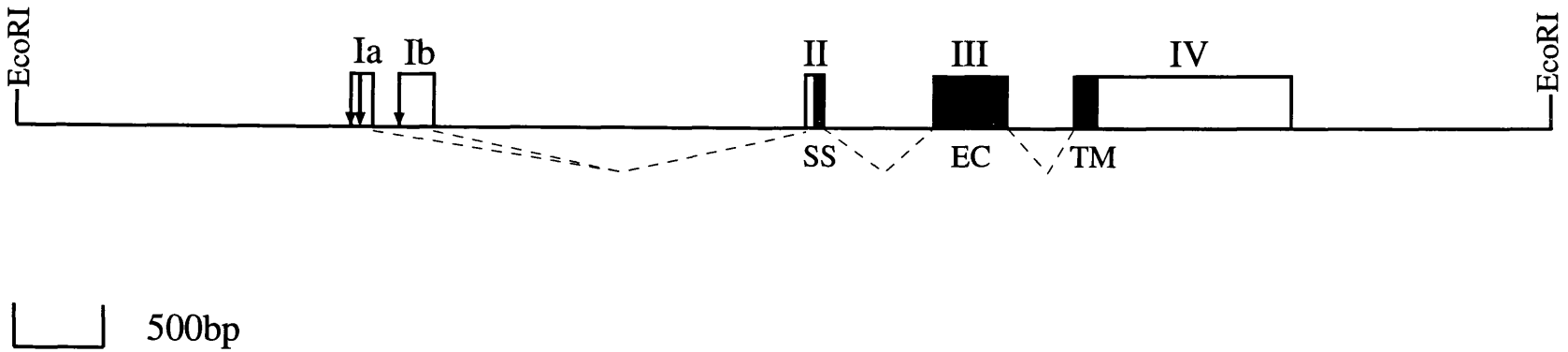
Initial experiments designed to identify cis-acting control elements within the *Thy-1* gene were carried out both in cell lines (Evans *et al.*, 1984) and in transgenic mice (Kollias *et al.*, 1987). These showed that an 8.2 kb cloned fragment of the murine *Thy-1* gene was sufficient to confer expression of the gene in the correct tissue specific and developmental manner, in both brain and thymus (Kollias *et al.*, 1987) but that it lacked sequences required for expression in peripheral T-cells in transgenic mice.

Further analysis of the *Thy-1* cis-acting regulatory elements using deletion constructs and mouse-human hybrid constructs isolated two regions of the gene responsible for this expression pattern (Vidal *et al.*, 1990). The region which directs thymic expression was found in the third intron. This element can act as a classical enhancer, functioning in a position and orientation independent manner and it can also act via a heterologous promoter (ie human β-globin promoter). The second region, located in the 3' half of intron 1, directs expression in brain. This element can function through heterologous promoters (ie that of mouse neurofilament light chain promoter and MHC class I H2-K gene promoter) however this element appears to be position and orientation dependent. This was shown by moving the element upstream of the promoter

Figure 5. Organisation of the murine *Thy-1* gene

Organisation of the murine *Thy-1* gene, 8.2 kb EcoRI fragment. Exons are depicted as boxes, introns and flanking regions are shown as lines. Filled boxes show translated regions. Dotted lines indicate splicing pattern. Vertical arrows show transcriptional start sites. SS signal peptide encoding sequence. EC extracellular region encoding sequence. TM putative transmembrane domain encoding sequence.

Adapted from Giguere *et al.*, 1985 and Spanopoulou *et al.*, 1988.


EcoRI
Ia
Ib
II
III
IV
EcoRI
SS
EC
TM
500bp

or increasing the distance between the element and the promoter by inclusion of a 4.8 kb c-myc fragment (Vidal *et al.*, 1990). Therefore, the tissue specific expression of Thy-1 appears, in the brain and thymus, to be directed by discrete *cis*-acting elements located 3' of the *Thy-1* promoter.

1.3.4 Protein and carbohydrate structure of Thy-1

Thy-1 protein was first successfully isolated and characterised from the membranes of rat thymocytes and brain (Barclay *et al.*, 1975; Barclay *et al.*, 1976; Campbell *et al.*, 1981). The molecule was shown to be a glycoprotein of MW 17,500 in brain and 18,700 on thymocytes (Kuchel *et al.*, 1978). The different molecular weights are due to differing carbohydrate compositions of the molecule in the various tissue.

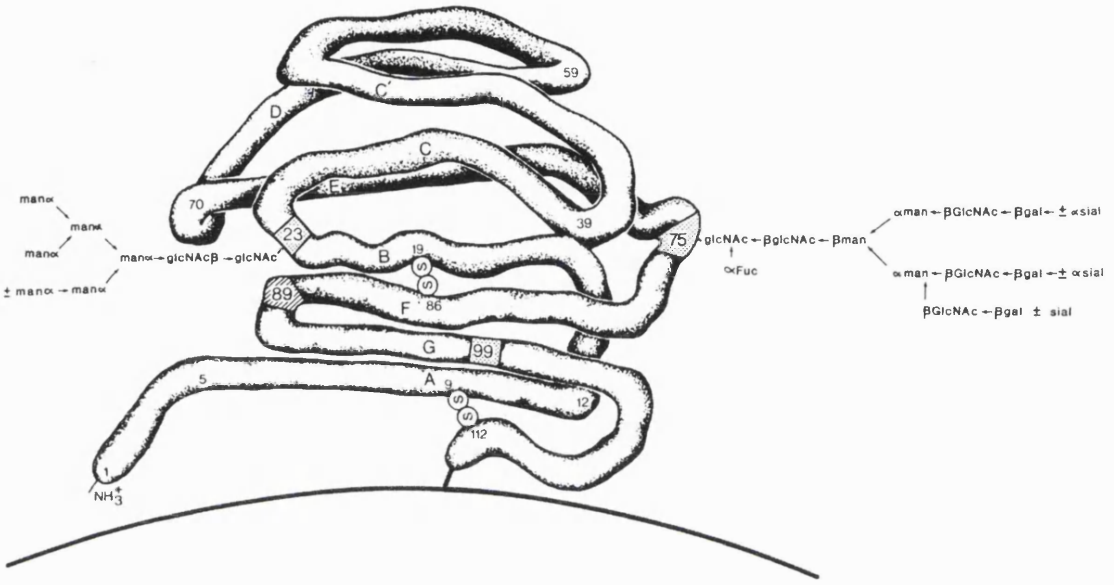
The polypeptide chain of mature Thy-1 consists of 112 amino acids in the mouse (Williams and Gagnon, 1982) and 111 amino acids in the rat (Campbell *et al.*, 1981). As mentioned above, the Thy-1.1 and Thy-1.2 allelic difference in mice is determined by residue 89, an arginine residue in Thy-1.1 mice and glutamine in Thy-1.2. The rat Thy-1 molecule shares the Thy-1.1 antigenic determinant Arg 89 (Campbell *et al.*, 1981), consistent with its recognition by anti-Thy-1.1 antibody (Douglas, 1973). The Thy-1 molecule possesses two disulphide bonds between cysteine residues 9 and 112 and 19 and 86 (numbering as for mouse Thy-1). Three N-glycosylation sites are present at asparagine residues 23, 75 and 99 (23, 74 and 98 in rat Thy-1) (Campbell *et al.*, 1981; Williams and Gagnon, 1982). The above features are described in Figure 6.

The carbohydrate composition has been determined for both rat and mouse Thy-1 showing a high degree of tissue specificity. This was first alluded to by the differing electrophoretic mobilities and lentil lectin binding affinity of rat Thy-1 isolated from brain and thymus (Barclay *et al.*, 1975; Letarte-Muirhead, 1975; Barclay *et al.*, 1976). The differing carbohydrate structure was later shown to be due to tissue specific glycosylation of Thy-1 (Carlsson and Stigbrand, 1984; Carlsson, 1985; Parekh *et al.*, 1987). At residue Asn 23 oligomannose structures are the predominant carbohydrate group found, with a low percentage of more complex structures found in thymus but not in brain. At the other two sites, Asn 75 and 99, Thy-1 shows a higher degree of

Figure 6. Structure of mouse Thy-1 protein

Schematic drawing illustrating structural features of mature cell surface Thy-1. The polypeptide chain is shown folded into the Ig-variable domain configuration of β-pleated sheets. Individual peptide strands of the β-sheet are designated by letter. The molecule is shown attached to the membrane at Cys_{112}. The disulphide bonds between Cys_9 and Cys_{112}, and between Cys_{19} and Cys_{86} are shown. Residue 89 specifies the Thy-1.1/Thy-1.2 allelic variation (highlighted with diagonal stripes). Asn residues 23, 75 and 99 which the carbohydrate chains are attached are shown (highlighted with spots). Proposed structure of two carbohydrate chains associated with Thy-1 are shown attached to residues 23 and 75 (Carlsson and Stigbrand, 1984; Carlsson, 1985).

Reproduced from Morris, 1985.


C'
D
C
E
59
70
39
23
B
19
S
S
75
89
F
86
G
99
A
9
S
S
5
12
112
1
NH3+
manα
manα
manα
manα→glcNAcβ→glcNAc
± manα→manα
glcNAc←βglcNAc←βman
αFuc
αman←βGlcNAc←βgal←±αsial
αman←βGlcNAc←βgal←±αsial
βGlcNAc←βgal ± sial

variability in its carbohydrate groups between tissues (Parekh *et al.*, 1987). Interestingly charged oligosaccharides in thymus Thy-1 are found at the site adjacent to the lipid bilayer whereas at the same site 98 in the brain only neutral mannose chains are found. One may speculate that this reflects different interactions with other membrane molecules in the two tissue types. In the brain it is only at site 74 that charged groups are found. In total Parekh *et al.* (1987) describe twelve "glycoforms" of Thy-1, four in brain and eight in thymus, non of which are shared between the two. Furthermore positions 75 and 99 in the mouse were shown to have differing carbohydrate moieties depending on differentiation status of the T-lymphocyte (Carlsson, 1985). These different glycoforms may therefore be a means of achieving functional diversity of the molecule.

More recently, analysis of the N-glycosylation patterns of Thy-1 in mouse, rat and man were shown to be conserved at corresponding sites in the three species (Williams *et al.*, 1993). This is particularly evident for sites at Asn 74/75 in rat/mouse and the equivalent site 60 in human Thy-1 and it is also seen at sites Asn 98/99 and 100 respectively. At Asn 23 mouse and rat were almost identical but differed significantly with human. So not only does Thy-1 exhibit a tissue specific glycosylation pattern but this pattern of glycosylation appears conserved in a number of species. The authors conclude that the role of Thy-1 may be to function as a carrier for specific carbohydrates.

A striking feature elucidated from the sequencing of the Thy-1 molecule was its homology to immunoglobulin domains (Campbell *et al.*, 1979; Cohen *et al.*, 1981; Williams and Gagnon, 1982). Immunoglobulin (Ig) heavy and light chains are made up of structural domains of around 110 amino acids (Edelman, 1970). The chains are composed of two functional types of domain, constant (C) and variable (V). The constant domain is involved in effector functions such as cell attachment, signal transduction and complement binding, the variable region is involved in antigen recognition. Immunoglobulin domains are characterised by two beta-sheets of four and three beta-strands each. The amino acid sequence is such that each sheet presents a hydrophobic surface internally and a hydrophilic surface externally. The Ig structure is then stabilised by two disulphide bonds. The above features define the Ig-fold (reviewed in Amzel and Poljak, 1979). The modular facet of immunoglobulins and

immunoglobulin related molecules has suggested that they arose through the duplication and rearrangement of a primordial Ig domain or domains (Hill *et al.*, 1966; Hood *et al.*, 1985). This theory is supported by analysis of Ig gene structure showing each V and C domain to have a modular organisation (Sakano *et al.*, 1979), a property that Thy-1 shares (Seki *et al.*, 1985b; Giguere *et al.* 1985)

The circular dichroism spectrum of Thy-1 has shown it to be devoid of alpha helices (Campbell *et al.*, 1979). Further structural evidence based on position of the disulphide bonds, secondary and tertiary structural predictions and sequence homology have shown Thy-1 to be most homologous to Ig variable domains, though homology is also shared with C domains (Cohen *et al.*, 1981; Williams and Gagnon, 1982). Furthermore, Williams and Gagnon (1982) isolated and partially sequenced a Thy-1 like protein from squid nervous tissue. Within a sequenced block of 25 residues, around the C-terminal Cys 86, a block of five identities could be aligned with rodent Thy-1 but higher degree of homology was seen between the squid sequence and the mouse V_λ domain.

Together the above data led Williams and Gagnon (1982) to suggest that Thy-1 may represent the primordial Ig domain. The further identification of a host of molecules possessing high homology to Ig domains has given rise to the Ig supergene family (reviewed in Hood *et al.*, 1985; Williams and Barclay, 1988) of which Thy-1 appears to be one of the simplest members.

1.3.5 Attachment of Thy-1 to the cell membrane

Initially several lines of evidence suggested that Thy-1 was not attached to the cell membrane via a conventional transmembrane domain. Purification procedures using deoxycholate showed Thy-1 to bind large amounts of detergent, furthermore removal of deoxycholate resulted in formation of roughly spherical Thy-1 oligomers, properties found in lipid binding membrane molecules (Kuchel *et al.*, 1978). Upon sequencing of the rat Thy-1 C-terminus no hydrophobic amino acid residues were found, however, C-terminal peptides were associated with hydrophobic properties not due to the amino acid sequence. The unknown material contained galactosamine and glucosamine and was possibly a lipid (Campbell *et al.*, 1981).

Contradictory results were obtained upon the sequencing of both cDNA and genomic clones of rat Thy-1 (Seki *et al.*, 1985a and c). Sequence of the coding region predicted a mature protein of 142 amino acids in contrast to the 111 amino acid protein purified. The extra 31 amino acids also contained a hydrophobic region of 20 amino acids, enough to function as a transmembrane domain. The predicted transmembrane domain was therefore postulated as the means of membrane attachment for Thy-1 (Chang *et al.*, 1985). Subsequently, mature cell surface Thy-1 was shown to possess a glycosyl phosphatidylinositol (GPI) tail attached to its carboxyl terminus (Tse *et al.*, 1985). Moreover, the molecule could be cleaved from the cell surface by phospholipase C treatment (Low and Kincade, 1985) confirming the use of the GPI tail for membrane attachment. Low and Kincade (1985) suggest that the discrepancy between the predicted sequence and the isolated protein is due to the postranslational removal of the transmembrane sequence prior to addition of the GPI tail. The structure of the GPI anchor was subsequently determined for rat brain Thy-1 (Homans *et al.*, 1988), and is shown in Figure 7. Tissue specific differences in the GPI-anchor are known to be present between brain and thymocyte Thy-1 (Tse *et al.*, 1985; Parekh *et al.*, 1987; Homans *et al.*, 1988) although the functional importance of this is not known.

The biosynthetic pathways involved in attaching the GPI anchor to Thy-1 are not well characterised. It is postulated that the C-terminus of the 142 amino acid form of Thy-1 is required to hold Thy-1 at the luminal surface of the endoplasmic reticulum. It is here that the C-terminal 31 amino acid is thought to be cleaved off, promoting the addition of the amide group of the ethanolamine to the free carboxy group of Cys 111 (Low, 1989). The GPI tail is added *en bloque* to Thy-1 soon after translation (Conzelman *et al.*, 1987).

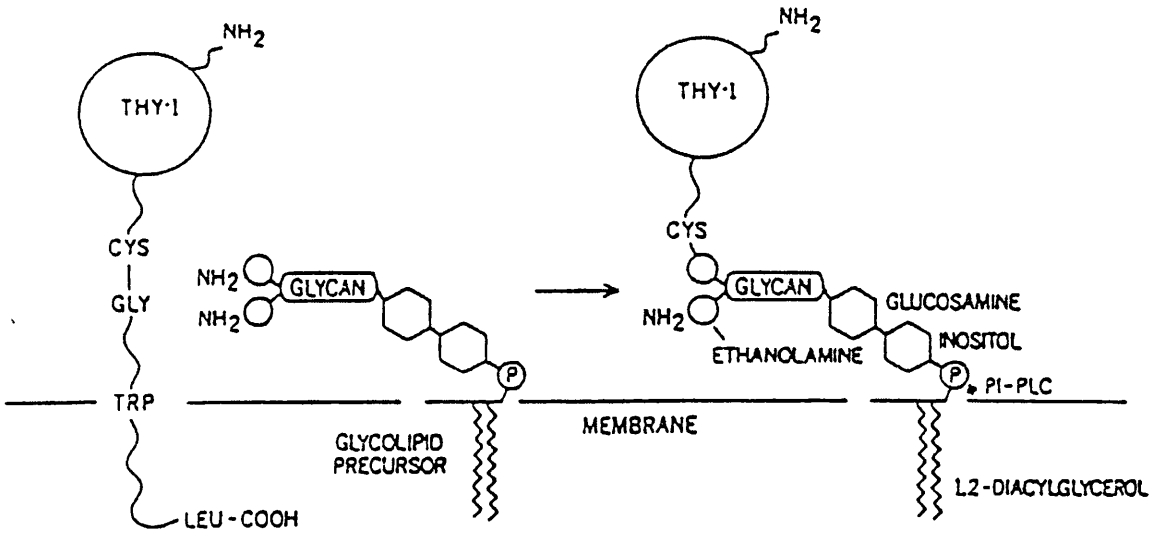
1.3.6 Signal transduction by Thy-1

A large group of eukaryotic cell-surface proteins are now known to employ a GPI anchor as means of membrane attachment. These proteins belong to diverse functional groups including enzymes, adhesion molecules of the Ig and fibronectin type III supergene families, and receptors (reviewed by Ferguson, 1992). Recent findings suggest that GPI anchors not only serve as a membrane tether but are involved in

Figure 7. Structure of Thy-1 membrane anchor.

Attachment of Thy-1 protein to the cell surface. Thy-1 is covalently linked to the phosphatidylinositol anchor via the α-carboxyl group of Cys_{112} of Thy-1 to the amino group of ethanolamine in the glycolipid anchor. In turn the acyl chain of phosphatidylinositol is thought to make hydrophobic interactions with lipid bilayer, anchoring the protein to the membrane.

Reproduced from Low, 1989b.


NH2
THY·1
CYS
GLY
TRP
LEU-COOH
NH2
NH2
GLYCAN
P
GLYCOLIPID
PRECURSOR
MEMBRANE
NH2
THY·1
CYS
GLYCAN
NH2
ETHANOLAMINE
GLUCOSAMINE
INOSITOL
P
PI-PLC
1,2-DIACYLGLYCEROL

transmembrane communication.

Crosslinking of GPI-anchored proteins on the surface of lymphocytes with antibodies, can result in the rapid influx of Ca^{2+}, increased kinase activity, cytokine release and finally proliferation (reviewed by Robinson, 1991). This *in-vitro* lymphocyte activation has been shown to be mediated via the GPI-anchor. Since exchanging the GPI anchor for a transmembrane domain abolishes the activation of the lymphocytes. The reciprocal experiment, substituting a GPI anchor for a transmembrane domain on a normally non GPI-linked protein, now allows activation by crosslinking of antibodies to this protein (Robinson *et al.*, 1989; Su *et al.*, 1991; Keller *et al.*, 1992).

From immunoprecipitation experiments of GPI-linked cell surface proteins from non-denaturing detergent extracts of lymphocytes, GPI-linked proteins have been found to be associated with protein kinases and phosphatases. Human GPI molecules CD14, CD24, CD48, CD55(DAF) and CD59 as well as Thy-1 and Ly-6 have, through these experiments, been found associated with the protein kinase $p56^{lck}$ (Stefanova *et al.*, 1991). Thy-1 has also been found associated with the kinase $p60^{fyn}$ (Thomas and Samelson, 1992; Garnett *et al.*, 1993) and with CD45 tyrosine phosphatase (Volarevic *et al.*, 1990). Again this co-precipitation of kinase activity is abolished when GPI-anchor is exchanged for a transmembrane domain (Shenoy-Scaria *et al.*, 1992). The *in-vivo* importance of the molecular association between the GPI-anchored proteins and kinases has yet to be shown, although it suggests GPI-linked molecules to be involved in a signal transduction pathway.

A number of GPI-anchored proteins have been localised to so called caveolae, which are non-clathrin-coated pits on the cell surface (reviewed by Hooper, 1992). Clustering at these domains appears to be dependent on the GPI-anchor as was shown when the transmembrane domain of CD4 was exchanged for a GPI-anchor (Keller *et al.*, 1992). Since calcium channels and kinase activity have been found associated with these membrane structures, caveolae may be microdomains which represent signal transduction centres on the cell surface (Fujimoto *et al.*, 1992; Fujimoto, 1993; Sargiacomo *et al.*, 1993). The association of GPI-anchored proteins with specialised microdomains on the cell surface again suggests that these proteins are involved in signal transduction. Interestingly, Thy-1 has been located at these structures (Ying *et al.*, 1992).

Cross linking of cell surface Thy-1 with antibody induces a mitogenic response in lymphocytes (Kroczek *et al.*, 1986), the antibody triggered activation is associted with an influx of Ca^{2+} (Barboni *et al.*, 1991). Anti-Thy-1 antibodies immobilised on a substratum (Leiffer *et al.*, 1984; Lipton *et al.*, 1992) or present in the growth media (Mahanthappa and Patterson, 1992; Doherty *et al.*, 1993) can trigger neurite outgrowth in culture from primary neurons and/or the PC12 neuronal cell line. Thy-1 antibody triggered neurite outgrowth, like the Thy-1 antibody activation of lymphocytes, is associated with an influx of Ca^{2+} (Doherty *et al.*, 1993). The above experiments show that Thy-1 can act as a signal transduction molecule in both lymphocytes and neuronal cells. However the *in-vivo* relevance of anti-Thy-1 antibody activation of lymphocytes or promotion of neurite outgrowth is yet to be established as no ligand has been found which can mimic this result.

Thy-1 has also been shown to inhibit neurite outgrowth of a transfected neuronal cell line on mature astrocytes *in-vitro*, the inhibition being dependent on the nature of the membrane anchor present on the transfected Thy-1 molecule (Tiveron *et al.*, 1992; Tiveron *et al.*, 1994). The neuronal cell line NG115-40L, which does not express Thy-1 normally, will extend neurites over astrocytes in co-culture. Expressing GPI-linked Thy-1 on the surface of this cell line inhibits this outgrowth when the neurite contacts the astrocyte. The addition of soluble Thy-1 to the media again allows neurite extension, showing that Thy-1 is interacting with a ligand on the surface of the astrocyte and that the inhibitory signal is transduced via the Thy-1 molecule (Tiveron *et al.*, 1992). Neurite outgrowth is not inhibited by fibroblasts, epithelial cells, Schwann cells or embryonic glia showing the inhibition to be specific to astrocytes. Exchanging the GPI anchor of transfected Thy-1 for the transmembrane and intracellular domain of NCAM abolishes the inhibitive effect on neurite outgrowth (Tiveron *et al.*, 1994). However, when only the transmembrane spanning region from NCAM is swapped for the GPI anchor, inhibition of outgrowth can be achieved albeit inefficiently. The normal GPI-anchored form of Thy-1, when transfected into the NG115-40L cell line, was shown to be localised in discrete microdomains on the cell surface (Tiveron *et al.*, 1994). Interestingly the Thy-1 protein with only the transmembrane domain from NCAM showed a similar membrane localisation as the GPI-linked Thy-1. It may therefore be the that the GPI anchor is a signal for targeting Thy-1 to the caveolae or microdomain

on the cell surface and that it is not involved in signal transduction per se (Tiveron *et al.*, 1994).

1.3.7 A model for Thy-1 function

Structurally Thy-1 shares a high degree of homology with the Ig variable domain (Campbell *et al.*, 1979; Cohen *et al.*, 1981; Williams and Gagnon, 1982) and is regarded as the smallest member of the Ig super gene family. The Ig domain is shared by a large variety of molecules which are known to act as receptors, adhesion and recognition molecules (Figure 8). As a member of this group of genes Thy-1 is believed to function in some guise as a recognition molecule.

As discussed in section 1.3.6, Thy-1 can be shown to act as a signal transduction molecule and this function is dependent on the membrane localisation which is confered by its GPI-anchor. Furthermore, the experiments described above show, *in-vitro*, that Thy-1 can inhibit neurite outgrowth in conjunction with an unknown ligand present on astrocytes (Tiveron *et al.*, 1992).

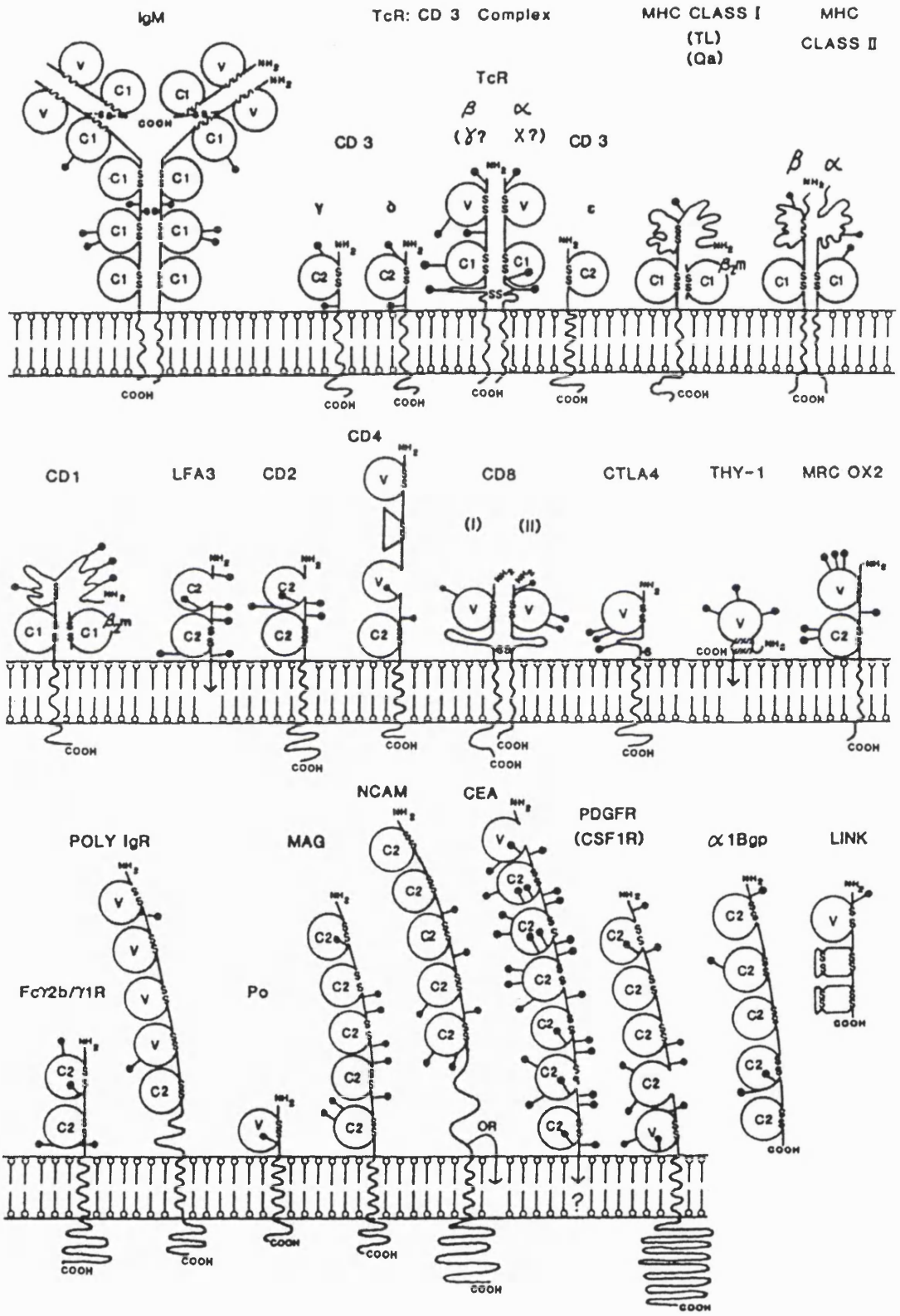
The *in-vitro* inhibition of neurite outgrowth by Thy-1 is particularly interesting when considered in conjunction with the temporal and spatial appearance of Thy-1 in the developing nervous system. As discussed in section 1.3.3, the appearance of Thy-1 in the nervous system correlates with the cessation of axonal growth, suggesting that the molecule participates in processess occuring at this time point, such as growth inhibition and/or stabilising axonal interactions possibly with other axons or glia cells (Morris, 1992; Xue and Morris, 1992). One may therefore speculate that until the axon has reached its terminal field Thy-1 is excluded from its surface, however once this has happened Thy-1 appears. This allows stable interactions with the astrocytes to be made and/or produces an inhibitory signal, hence preventing further growth to occur. The signal that stimulates the final appearance of Thy-1 on the surface of the axon is not known. It clearly does not act at the levels of trancriptional or translational regulation as Thy-1, at this point, is already found on the cell body and/or dendrites (Xue *et al.*, 1991; Xue and Morris, 1992). It may be the case that the nature of the axonal membrane or axonal transport changes upon the axon reaching its terminal field. This then allows the GPI-linked Thy-1 to be present on the surface, however this is highly

Figure 8. Immunoglobulin supergene family

Diagrammatic representation of members of the immunoglobulin supergene family, showing the structural characteristics of this family. Circles represent sequence segments homologous to the Ig domain. V indicates homology to Ig variable domain, C indicates homology to constant domain. SS indicates disulphide bonds. GPI anchors are show as arrows in membrane. Glycosylation sites are shown as stalks.

Reproduced from Williams and Barclay, 1988.

THE IMMUNOGLOBULIN SUPERFAMILY



speculative.

Interestingly, over expression of Thy-1 by 7 to 16 fold in transgenic animals does not lead to an abnormal phenotype in the tissues in which it is normally found (Kollias *et al.*, 1987). Thy-1 is already expressed at high levels on its host tissue, this could suggest that its function is dependent on concentration of a particular ligand and not upon its own concentration. Therefore increasing Thy-1 expression would not cause any effect unless the expression of the ligand was altered as well. Ectopic expession of Thy-1 in the B-cells of transgenic mice does however lead to lymphoid hyperplasia in bone marrow and lymph nodes (Chen *et al.*, 1987). This is consistent with the proliferative signal transduction found when cross-linking Thy-1 with antibodies on lymphocytes.

The glycosylation pattern of Thy-1 may also provide a clue to its function. The glycosylation pattern of neural Thy-1 is conserved between rat, mouse and human (Williams *et al.*, 1993). Thy-1 also exhibits conserved neuronal expression, suggesting that the role or one of the roles of Thy-1 is to present a particular carbohydrate structure on the neuronal surface. Interestingly NCAM, another Ig-related molecule expressed in the nervous system, also displays cell type specific glycosylation allowing the molecule to serve different functions. A highly α-(2,8)-polysialylated (PSA) form is expressed in the granular cells of the olfactory bulb but a less sialylated form is expressed in the glomeruli (Zuber *et al.*, 1992). This PSA form of NCAM promotes cell and axon migration whereas the non-PSA form of N-CAM is involved homophilic interactions and performs an adhesive role in axonal fasciculation (Rutishauser *et al.*, 1988). Therefore, one may speculate that like NCAM, Thy-1 can achieve functional diversity by displaying different glycoforms.

Together the above data suggest that Thy-1 plays a role as a receptor and/or adhesion molecule. The conserved distribution and glycosylation pattern in the nervous system would suggest that this is the main target site for its function, though this does not exclude a role in signal transduction or adhesion in other tissues in which it is found, such as the lymphoid system. During neural development the appearance of Thy-1 is under tight control at various levels. The most striking outcome of this control is that Thy-1 is excluded from growing regions of the axon. Coupled with the inhibition of neurite outgrowth by Thy-1 in culture allows one to hypothesise that Thy-1 is

involved in mediating growth inhibiting signals in conjunction with an unknown ligand present on astrocytes during or after the completion of axonal outgrowth.

Understanding the cell surface interactions which mature axons make with their environment is of clinical importance. If the interaction of Thy-1 with astrocytes occurs *in-vivo*, then Thy-1 could be involved in surpressing axonal regrowth after injury in astrocyte rich areas of the adult CNS. This has obvious implications for the study of axonal regrowth following spinal injury.

The application of gene targeting to produce a null Thy-1 allele should test whether Thy-1 does play a crucial role in the above processes

1.4 The murine *Zfp-37* gene

1.4.1 Introduction

A central theme in molecular neurobiology is the study of cell type diversity in the mammalian nervous system and how it is established during neurogenesis. Ultimately, the understanding of this process will be reliant on the elucidation of the cascade of genes encoding regulatory factors and other proteins which are key to the proliferation, migration and differentiation of neuronal precursor cells. These events rely on specific spatial and temporal control of gene expression. At the genomic level this is mediated by a set of nuclear proteins, many of which are DNA binding proteins.

Based on the presence of conserved structural motifs, DNA binding proteins can be separated into different regulatory families. Members of these families have been implicated in mammalian nervous system development and have been shown to be conserved during evolution both at structural and functional levels. One such class of DNA binding protein uses the zinc-finger motif to bind DNA (reviewed by Harrison, 1991). Some examples of zinc-finger proteins include the *Drosophila* gap genes *Krüppel* and *hunchback*, which play a role in controlling the expression of other regulatory genes, including homeotic genes, which are involved in the segmentation of the *Drosophila* body plan (Rosenberg *et al.*, 1986; Tautz *et al.*, 1987; reviewed by Wilkinson and Krumlauf, 1990). Another important example is the murine *Krox20* gene, which is involved in positional organisation of the hindbrain (Wilkinson *et al.*, 1989;

see section 1.2.10).

Part of this thesis addresses the function of the zinc-finger gene *Zfp-37* in murine development and particularly in the development of the nervous system. Evidence, based on the structure and expression pattern of this gene, suggests a role in transcriptional regulation during neurogenesis. In this context, it is important to first provide a brief overview of how eukaryotic genes are regulated at the transcriptional level.

1.4.2 Eukaryotic transcriptional regulation

1.4.2i Eukaryotic promoters

Three RNA polymerases exist, RNA polymerase I, II and III. RNApolI synthesise ribosomal RNA, RNApolII synthesises messenger RNA and RNApolIII transfer RNA, 5S RNA and other small cellular RNAs and viral RNAs (reviewed by Struhl, 1994). In this study transcription of mRNA is of primary interest and therefore transcription by RNApolII will be discussed only. Initiation of transcription requires the RNA polymerase to be recruited to the transcriptional start site within the promoter region. This is achieved through the binding of cellular transcription factors to promoter elements which lie directly upstream of the site of initiation (reviewed by Saltzman and Weinmann, 1989).

The promoter region of RNApolII genes usually spans a short sequence of a few hundred base pairs, found immediately upstream of the initiation site. This sequence normally contains an AT-rich region, the TATA-box, located 25-30 nucleotides upstream of the initiation site and one or more other regulatory sequences (reviewed by Lewin, 1990). A cascade of transcription factors are required to bind to the promoter region to recruit the polymerase. This is initiated by binding of TFIID to the TATA-box, followed by other factors such as TFIIA, TFIIB, TFIIE, TFIIF(RAP30/74) (together called the preinitiation complex) and the recruitment of the RNApolII onto the DNA (Workman and Roeder, 1987; reviewed by Saltzman and Weinmann, 1989). TFIID has been found to be a multisubunit complex comprised of the TATA-binding protein (TBP) and at least seven other protein subunits, TBP-associated factors or TAFs (Dynlacht *et al.*, 1991; Tanese *et al.*, 1991). However, not all class II promoters posses a TATA-box, in this case multiple sites of initiation are found. In the absence of a

TATA-box, the initiation complex may be recruited via the binding of TFII-I to an initiator element (Roy *et al.*, 1991; Roy *et al.*, 1993).

Class II gene promoters also contain upstream elements in their promoter where other factors may bind. These factors may stimulate transcription by acting on the basic transcription complex, for instance the E2 transactivator protein stimulates transcription after binding of TFIID (Ham *et al.*, 1994). They may also help in the binding of the preinitiation complex to the promoter, for instance Sp1, in association with TAF co-factors, will cooperatively interact with TFIID to bind DNA (Pugh and Tjian, 1990; Roeder, 1991; Hoey *et al.*, 1993). It is beginning to appear that in many cases TAFs function to mediate the interactions between upstream binding factors and the basal transcriptional machinery (Chen *et al.*, 1994). Certain TAFs can be shown to specifically interact with the activation domains of other promoter/enhancer binding factors such as Sp1 and the neurogenic element binding factor 1 (NTF-1) *in-vitro* (Chen *et al.*, 1994). Identifying the link between the basal transcriptional machinery and more distant factors.

1.4.2ii Enhancer and silencer elements

Enhancers can be defined as cis-acting elements that function at relatively larger distances than promoters to increase transcription. Classically they may be positioned upstream or downstream and in either orientation to the transcription initiation site. The first enhancer was characterised in the DNA tumour virus SV40. In this case the enhancer element is a 72 bp repeated sequence motif (Banerji *et al.*, 1981), containing multiple factor binding sites. The combination of the individual elements and their positioning was shown to be important for the functioning of this enhancer element (Ondek *et al.*, 1987). In other enhancers the arrangement of the factor binding sites is not always important as is shown by the example of the heat shock response element (Bienz and Pelham, 1986). Enhancer elements have been found surrounding a large number of inducible and tissue-specific genes. A number of enhancer elements conferring tissue specific gene activation have been pinpointed via experiments in transgenic animals, for instance the thymic enhancer element located in intron 3 of the murine *Thy-1* gene (See section 1.3.3).

The mode of action which allows enhancers to act over large distances is not clearly defined. The most popular model suggests that sequence specific factors bound at the enhancer element form protein-protein contacts with the promoter bound factors, looping out the intervening DNA (reviewed by Ptashne, 1988). DNA looping has been shown for the interaction of the zinc-finger factor Sp1 and E2 by electron microscopy (Li *et al.*, 1991), adding credence to this theory.

Like enhancer elements, silencers are believed to repress transcription of linked genes in an orientation and distance independent manner. First identified in the yeast mating type locus (Brand *et al.*, 1985), silencer elements have also been mapped in the α T cell receptor gene (Winoto and Baltimore, 1989), the chicken lysozyme gene (Baniahmad *et al.* 1990), the *Drosophila decapentaplegic* gene (Huang *et al.*, 1993b) and the mouse neuron-specific gene SCG10 (Mori *et al.*, 1990).

1.4.2iii Locus control regions

Locus control regions (LCRs) are another type of DNA element identified in controlling transcription. Like enhancers they can enhance transcription from a distance either upstream or downstream and in either orientation relative to the site of transcription initiation (Talbot *et al.*, 1989). However, in gene expression experiments in cell lines or transgenic mice, LCRs can function independently of their site of integration in the genome and transcribe the target gene in a copy-number dependent manner (Grosveld *et al.*, 1987; Blom van Assendelft *et al.*, 1989). In such assays classical enhancers tend to be influenced by the activity status of the site of integration in the genome (Wilson *et al.*, 1990; Krumlauf *et al.*, 1986). The first LCR was described for the human β-globin locus, where it was found to be able to transcribe its target gene in an integration independent and copy-number dependent manner (Grosveld *et al.*, 1987; Blom van Assendelft *et al.*, 1989). This feature of the LCR suggests that it is able to overcome the local regulatory status. LCRs have now been identified for the chicken lysozyme gene (Bonifer *et al.*, 1990) and the human CD2 gene (Greaves *et al.*, 1989).

1.4.2iv Chromatin structure and eukaryotic gene expression

The DNA in the nucleus of a eukaryotic cell is compacted by association with nuclear proteins, histones, to form chromatin. The basic repeating structure of chromatin is the nucleosome. This consists of eight histone subunits, two of each H2A, H2B, H3 and H4 subunits, and 146bp of DNA (Richmond *et al.*, 1984). The DNA is coiled in a left hand super helix 1.75 times around the histone core, changing the helical periodicity from 10.5 bp per helical turn to 10bp per turn (Klug and Lutter, 1981).

Each nucleosome unit is joined by a short length of DNA. One molecule of histone H1 associates with this linker DNA as it enters and exist the nucleosome octamer. In the presence of H1 these nucleosomes compact to form the 30nm fibre. This structural organisation has been shown to exist by electron microscopy (Thoma *et al.*, 1979). Further analysis suggested that the 30nm fibre is a superhelical structure with the nucleosomes organised radially and at a tilted attitude to the axis of the helix (Widom and Klug, 1985; Felsenfeld and McGhee, 1986).

Non-histone proteins are also regarded as being involved in the organisation of the 30nm fibre in interphase and mitotic chromosomes. Evidence for this comes from electron microscopy of metaphase chromosomes which are depleted of their histones. The DNA is seen to be attached in loops to a central scaffold (Laemmli *et al.*, 1977). This has led to the suggestion that chromatin is organised into discrete and topologically defined domains and that, in the chromosome, the 30nm fibre is organised into 30-100kb loops which are attached at their base via non-histone proteins to a nuclear scaffold (Gasser and Laemmli, 1987). One of the proteins associated with the scaffold was found to be topioisomerase II (Lewis and Laemmli, 1982; Gasser and Laemmli, 1987), which may be not only acting as a structural component but may also be involved in the release of torsional stress caused during transcription or replication.

DNA elements associated with scaffold attachment, scaffold attachment regions (SARs), have been described in *Drosophila*, chicken, mouse and human (Mirkovitch *et al.*, 1984; Cockerill and Garrard, 1986; Jarman and Higgs, 1988; Phi-Van and Stratling, 1988). These regions are notable for their AT-rich content, they show a high frequency of topioisomerase II cleavage sites and more recently have been shown to account for the banding pattern of metaphase chromosomes (Saitoh and Laemmli, 1994). It has been

suggested SARs function as transcriptional "insulator" regions, isolating gene loci from surrounding chromatin (Gasser and Laemmli, 1987). The insulating function of SARs has been tested in transgenic mice to determine whether they can act to "insulate" a transgene from integration effects asserted by other loci. However, only a limited insulating effect was seen (Thompson *et al.*, 1994). SARs have also been shown to stimulate transcription of linked genes in stable integration events (Stief *et al.*, 1989; Phi-Van *et al.*, 1990). The role SARs play in transcriptional regulation is yet to be determined.

Actively transcribed genes are generally more sensitive to enzymatic cleavage, ie by DNase1, this is thought to reflect altered chromatin conformation, chromatin modifications and/or nucleosomal displacement in transcribed genes (Weintraub and Groudine, 1976; Gross and Garrard, 1988). The precise dynamics of nucleosome positioning during transcription is not known (reviewed by Felsenfeld, 1992; Lewin, 1994; Kornberg and Lorch, 1995). *In-vitro* experiments have demonstrated that nucleosomes can be translocated by advancing RNA polymerase (prokaryotic SP6) via a suggested "spooling" of the DNA around the nucleosome, i.e. the nucleosome core remains in contact with the DNA (Studisky *et al.*, 1994; reviewed by Travers, 1994). Therefore a spooling mechanism may account for how chromatin transcription proceeds.

Disruption and compaction of the nucleosomes may allow transcriptional activation and cause repression of transcription, respectively. This may be achieved through the interaction of nuclear proteins with the histone components of the nucleosome or non-specifically with the DNA (reviewed by Kornberg and Lorch, 1995; Simon 1995). The SWI/SNF genes of *Saccharomyces cerevisiae* are required for induction of transcription of yeast promoters (reviewed by Carlson and Laurent, 1994). The SWI/SNF gene products are components of a multiprotein complex (Cairns *et al.*, 1994; Peterson *et al.*, 1994) which were shown to disrupt nucleosomes *in-vitro* and facilitate access of sequence specific DNA binding proteins to the DNA template (Kwon *et al.*, 1994; Cote *et al.*, 1994; Imbalzano *et al.*, 1994). It is unclear whether this process relies upon interaction of SWI/SNF complex with histones or direct interaction with the DNA. Nucleosome disruption *in-vitro* has also been shown for the *Drosophila Trithorax*-like gene product GAGA. A pre-assembled nucleosome-DNA complex formed on the heat shock promoter, *hsp70* could be disrupted with addition of purified

GAGA as shown by increased nuclease-sensitivity (Tsukiyama *et al.*, 1994). The above experiments suggest that nucleosome displacement is required for transcriptional activation to proceed.

Members of the *Drosophila* polycomb group (PcG) of proteins act as transcriptional repressors of homeotic genes. Of the PcG proteins thirteen have been identified and are known to cause homeotic phenotypes, none of these proteins have shown site specific DNA-binding. The repression by PcG proteins is thought to be achieved by altering the chromatin structure possibly by formation of heterochromatin over the relevant gene making it inaccessible to the transcriptional machinery (reviewed by Paro, 1990; Simon, 1995). Interestingly PcG homologs have been identified in mouse, these include *bmi-1* (Haupt *et al.*, 1991; van Lohuizen *et al.*, 1991), *mel-18* (Ishida *et al.*, 1993) and *rae-28* (Nomura *et al.*, 1994). The bmi-1 protein shares a 43% homology with the fly PcG protein posterior sex combs (Psc). Interestingly, targeted mutation of *bmi-1* in the mouse leads to homeotic transformations in skeletal vertebrate, neurological and haematopoietic defects (van der Lugt *et al.*, 1994). The skeletal transformations mimic the homeotic transformations seen in mutants of the *Drosophila* PcG proteins (Simon *et al.*, 1992). The means in which the PcG proteins achieve transcriptional repression has yet to be fully established, however they may represent another level of transcriptional regulation through chromatin structure.

1.4.2v DNA Methylation and eukaryotic gene expression

The mammalian genome shows a tissue specific pattern of DNA methylation. Methylation occurs in the form of 5-methylcytosine and is restricted to the dinucleotide CpG. In the mammalian genome the bulk (70%) of CpGs are methylated, the remaining CpGs are unmethylated. These methylation free CpGs are found clustered in regions of 1-2 kb and are associated with all house keeping gene promoter regions and 40% of tissue specific gene promoter regions. The methylation free CpG clusters are termed CpG islands or HpaII Tiny Fragment (HTF) islands because of the predominance of sites for the HpaII restriction endonuclease (reviewed by Bird, 1986; Lewis and Bird, 1991; Bird, 1992; Cross and Bird, 1995). Research on the role that methylation takes in transcription has centred on the methylation status of the CpG islands associated with

the promoter regions.

Transcriptional activity appears to correlate with a lack of methylation (Cedar, 1988) and reciprocally, methylation of promoter regions have been shown to repress transcription (Ben-Hatter and Jiricny, 1988; Watt and Molloy, 1988; Iguchi-Ariga and Schaffner, 1989). The methylation of CpG islands also appears to be important for the correct repression of the inactive X chromosome (Riggs and Pfeifer, 1992). However, it is unclear whether it is the methylation which causes the genes inactivity or visa-versa. Whether silenced, genes can acquire methylation (Enver *et al.*, 1988) and activation of a gene can be followed by demethylation (Sullivan *et al.*, 1989). It has been suggested that methylation may play a role in suppressing basal transcription in inappropriate cell types (Bird *et al.*, 1992). So although methylation does not play an active part in forming transcriptional domains it may well act as another level of control.

The biological importance of DNA methylation in the mammalian genome is called in to question by the fact that the genomes of lower eukaryotes such as *Drosophila, Caenorhabditis elegans* and yeast are devoid of methylation. However, targeted disruption of the murine gene encoding DNA (cytosine-5)-methyltransferase leads to the stunted growth and death at midgestation of the homozygous animals (Li *et al.*, 1992). The homozygous mice for the mutated allele having only 30% of the wild-type levels of DNA methylation. This experiment shows that DNA methylation is a required feature of the mammalian genome.

1.4.3 Sequence specific DNA-binding proteins

The control of temporal and tissue specific gene expression is achieved to a large extent by the presence of sequence specific DNA binding proteins which finely modulate the level of gene expression. These proteins appear modular in their structure and function, the DNA-binding domains can be separated from the transactivation domains (reviewed by Frankel and Kim, 1991). The DNA-binding motif allowing specific recognition of target sequence within a regulatory element, the transactivation domain then mediating protein-protein interactions which directly or indirectly affect the activity of the basal transcriptional machinery. This modular organisation can be

demonstrated by exchanging DNA binding motifs, for example changing the DNA binding domain from the yeast transcriptional activator GAL4 with the DNA binding domain of the bacterial repressor protein LexA could stimulate transcription when bound to a promoter with a LexA DNA binding site (Brent and Ptashne, 1985). It is also possible that transcription factors can exert their effect by affecting local DNA topology which in turn could promote the formation of productive complexes and finally transcription initiation.

DNA-binding proteins can be classified by the protein motif which they use to bind to DNA or dimerise. These motifs include helix-turn-helix, leucine-zipper coiled coil motifs, basic helix-loop-helix and zinc-finger motifs. The latter are discussed in more detail in the following section.

The helix-turn-helix (HTH) motif was the first DNA-binding domain identified in a transcription factor, being found in the cro and ci repressors of the λ phage (Anderson *et al.*, 1981; Pabo and Lewis, 1982) and the *E.coli* catabolite activator protein (CAP) (McKay and Steitz, 1981). The HTH motif is defined as a 20 a.a. sequence containing two α-helixes. These helices are linked by a sharp β-turn forming a cross like structure which is angled at about 120° (Steitz *et al.*, 1982; reviewed by Harrison, 1991; Freemont *et al.*, 1991). The second of the two helices is referred to as the "recognition" helix since it lies in the major groove of the DNA and appears to direct the sequence specific binding (Wharton *et al.*, 1984). A number of eukaryotic factors have been identified which use this motif for DNA-binding including homeodomain proteins (McGinnis *et al.*, 1984; Scot and Weiner, 1984) and products of the *S.cerevisiae* mating locus MATα2 and MATa1 (Nasmyth, 1982; Shepard *et al.*, 1984). HTH motifs also occur in the POU family of proteins, named after the first three identified members, the mammalian proteins Pit-1, Oct-1 and Oct-2 and the *C.elegans* protein unc-86 (Levine and Hoey, 1988). The POU domain is a conserved region of 150-160 a.a. and is subdivided into the POU specific region or box and the POU-homeodomain within which is the HTH motif (Herr *et al.*, 1988).

The leucine-zipper coiled-coil/basic region proteins also termed as basic zipper (bZIP) and basic coiled-coil (bCC) motifs (reviewed by Harrison, 1991) were originally defined by their dimerisation domains. These proteins contain a 30 a.a. α-helix with leucine residues occurring at around every seventh position, forming a row of leucine

residues along one face of the helix, the leucine zipper (Landschultz *et al.*, 1988). This zipper can then dimerise with the leucine zipper of its partner to form an α-helical coiled coil (O'Shea *et al.*, 1989; Oas *et al.*, 1990). The DNA-binding is achieved via an α-helical basic regions linked to each zipper which, when dimerised, interact, with one half of a symmetrical 10bp recognition sequence (Vinson *et al.*, 1989).

The binding of the basic region to DNA is not dependent on the leucine zipper, this has been shown by linking the two basic regions of the yeast factor GCN4 by a disulphide bond after which DNA binding ability was still observed (Talanian *et al.*, 1990). The dimerisation does play an indirect role as is shown in the case of the proto-oncogene products FOS and JUN, which both contain bZIP regions. JUN can homodimerise to bind DNA but FOS must heterodimerise with JUN to bind DNA, suggesting that the leucine zipper is indirectly required for binding (Kouzarides and Ziff, 1989).

The basic helix-loop-helix (bHLH) motif is a basic DNA-binding domain joined to two amphiphatic α-helices separated by a variable loop region. Dimerisation is mediated by the helices, which like the leucine zipper helices, present an alignment of hydrophobic residues. This allows for the formation of both homo- and heterodimers (reviewed by Jones, 1990). The motif was first identified in the proteins E12 and E47 which bind to the immunoglobulin enhancers (Murre *et al.*, 1989). Other members of this family include for example MyoD, Myf5 and myogenin involved in muscle differentiation and several proteins involved in neurogenesis in *Drosophila,* encoded by the *achaete-scute, daughterless, hairy* and *extramachrochaete* genes (reviewed by Jones, 1990).

1.4.3i Zinc finger motifs

The zinc finger motif was first described for the Xenopus transcription factor TFIIIA (Brown *et al.*, 1985; Miller *et al.*, 1985). This motif is a 28-30 a.a. module with one zinc ion coordinated between two cysteine and two histidine residues (reviewed by Harrison, 1991). The term "finger" describes the two dimensional representation of the primary structure. The motif can be defined as an independently folded domain, using a zinc ion to stabilise the packing of α-helix with a β-sheet (Berg, 1988).

Other zinc-finger motifs have been identified and can be divided into five different classes based upon the number of cysteine and/or histidine residues which are used to coordinate the zinc ion. The Class 1 type of finger is referred to as Cys_2His_2 or TFIIIA-type, and probably accounts for the finger domain in the majority of zinc finger proteins. The class 2 proteins coordinate zinc via four cysteine residues, Cys_2Cys_2, members of which include the steroid hormone receptors (Freedman *et al.*, 1988). Class 3 proteins coordinate two zinc ions via six cysteine residues, this structure is seen in the yeast GAL4 protein (Pan and Coleman, 1990). Class 4 proteins, like class 2, coordinate zinc via Cys_2Cys_2, however nuclear magnetic resonance (NMR) spectroscopy has shown that the manner of interaction with DNA is different from class 2 (Omichinski *et al.*, 1989). The class 4 proteins are members of the GATA-binding family of proteins. Class 5 proteins use three cysteine and one histidine residue to coordinate the zinc ion. This motif has been seen in the viral gag protein and the myelin transcription factor 1, MyT1 (Summers *et al.*, 1990; Kim and Hudson, 1992). Pertinent to this thesis are class 1 Cys_2His_2 zinc finger proteins and these are described in detail below.

There are estimated to be hundreds of structurally distinct class 1 Cys_2His_2 zinc fingers in the vertebrate genome. In humans there are estimated to be 300-700 different genes (Bellefroid *et al.*, 1989), comparable estimates are also given for Xenopus (Koster *et al.*, 1988) and mouse (Chowdhury *et al.*, 1987). Sequence analysis of over 200 of these Cys_2His_2 fingers reveals the following consensus sequence:

$$O\text{-}X\text{-}Cys\text{-}X_{2\text{-}4}\text{-}Cys\text{-}X_3\text{-}O\text{-}X_5\text{-}O\text{-}X_2\text{-}His\text{-}X_{3\text{-}4}\text{-}His\text{-}X_{2\text{-}6}$$

where O is a hydrophobic residue and X is variable. The invariant Cys and His residues define this structure (reviewed by Harrison, 1991). These proteins tend to contain at least three repeats of the above sequence, (reviewed by Pieler and Bellefroid, 1994). The vast majority of these proteins also possess a common seven amino acid linker sequence, TGEKPY/F, between each finger. This was first described for the product of the *Drosophila* gene *Krüppel*, hence these proteins are described as *Krüppel*-like (Schuh *et al.*, 1986; Chavrier *et al.*, 1987; Chowdhury *et al.*, 1988).

X-ray crystallography studies of the structure of the three finger protein Zif268 bound to DNA revealed that the zinc fingers acted in a modular fashion, each folding and binding as an independent unit (Pavletich and Pabo, 1991). Each finger contacted an adjacent 3bp DNA subsites. These contacts were also found to be on one strand of

the DNA with the amino to carboxy order of the finger alignment corresponding to the 3' to 5' direction of the DNA strand. The α-helix of each finger fitted directly into the major groove of the DNA. The DNA recognition was established by only three amino acid residues located within the helix. This one-to-one interaction between amino acid and base was further investigated using the Cys_2His_2 zinc finger proteins Krox20 and Sp1 which contain similar modules recognising 3bp DNA subsites. The sequence recognised by a Krox20 finger could be altered to that of an Sp1 finger by simply mutating two amino acids, a change in DNA binding site recognition from GGG to GCG could be made (Nardelli *et al.*, 1991). Based on these observations Desjarlais and Berg (1993), designed artificial GGG, GCG and GCT binding finger modules which differed in only four "specificity determining" positions.

More recent structural analysis of two other Cys_2His_2 proteins, the two finger *Drosophila* protein tramtrack (Fairall *et al.*, 1993) and the human five finger protein GLI (Pavletich and Pabo, 1993) show that protein-DNA contacts outside the 3 bp subsite and interactions with the DNA phosphate backbone also play a role in the binding specificity. This suggests that the one-to-one interactions between amino acid residue and base, although applicable to an extent, does not explain the sequence specificity of all Cys_2His_2 zinc finger proteins. Attempts to define a code which would allow predictions of the recognition sequence from the amino acid structure of the α-helix have been made (Klevit, 1991; Choo and Klug, 1994 a and b). Using phage libraries to display zinc fingers with an α-helix containg randomised Choo and Klug (1994a) concluded that positions -1, 2, 3 and 6 in the helix were involved in recognition of DNA and that certain amino acid residues were conserved when binding to a particular sequence, however a direct relationship between amino acid residue and sequence could not be predicted precisely.

Members of class 1 Cys_2His_2 zinc fingers are not restricted to DNA binding only. TFIIIA exhibits dual function, acting as a DNA binding protein involved in the transcriptional regulation of the 5S ribosomal RNA (Pelham and Brown, 1980; Honda and Roeder, 1980) and as a RNA binding protein, functioning in 5S RNA storage and nucleocytoplasmic transport as part of a 7S ribonuclear protein complex with the 5S ribosomal RNA (Guddat *et al.*, 1990). This duality is mediated by the nine zinc fingers which show overlapping DNA and RNA binding ability, finger 3 is the only essential

finger for DNA binding and finger six exhibits RNA binding characteristics which are distinct from the other fingers (Theunissen *et al.*, 1992). The other example of a zinc-finger protein binding RNA is described for the Xenopus factor XFG5-1 which can bind specifically to RNA, though its *in-vivo* function is not known (Koster *et al.*, 1991). The mechanism of RNA binding is ill-defined and may well be determined on secondary and tertiary RNA structure (reviewed by Pieler and Bellefroid, 1994).

The class 1 zinc finger genes can be further divided into sub groups based upon common structural and functional attributes and also evolutionary conservation (Pieler and Bellefroid, 1994). The largest group is made up of the FAX and FAR families of finger proteins identified in Xenopus (Knochel *et al.*, 1989; Klocke *et al.*, 1994) and KRAB domain proteins (Bellefroid *et al.*, 1991), these large families of finger genes are characterised by multi finger cluster, common exon/intron organisation and are proposed have arisen late in evolution (Pieler and Bellefroid, 1994). The second group fall into small zinc finger subfamilies with evolutionary conserved finger clusters of three to five units which share similar DNA binding sequence specificities within each family. These include the Gli family, Krox20 family and Sp1 family of proteins the majority of these proteins show restricted expression pattern and are proposed to play a regulatory role in vertebrate embryogenesis.

1.4.3ii Transcriptional activation and repression domains

Once bound to its target sequence the transcription factor must exert its effect on the basal transcriptional machinery. The identification of transcriptional activation domains has shown that protein-protein interactions play a role in this mechanism (Ptashne, 1988; Brent and Ptashne, 1985). This interaction may occur directly with the transcriptional machinery or via an intermediary protein or coactivator (reviewed by Ptashne and Gann, 1990). Transcriptional activation domains have been divided into three major classes based on their amino acid sequence which can be the acidic, glutamine or proline rich domains (reviewed by Mitchell and Tjian, 1989).

The first identified transactivation domain was the acidic domain (reviewed by Hahn, 1993) and was described for the yeast transcription factors GAL4 and GCN4 (Ma and Ptashne, 1987). This domain was originally postulated to function via an α-helix

which provided a negative charged surface (Giniger and Ptashne, 1987). However, mutational analysis of this α-helix showed that disruption of this structure did not effect function (Cress and Triezenberg, 1991). It has subsequently been shown that the acidic activator regions of GAL4 and GCN4 may in fact form β-sheets and that the transactivation potential is not dependent on the formation of the helix or net negative charge (van Hoy *et al.*, 1993; Leuther *et al.*, 1993). It remains unclear what the functional component of the acidic domain is. However, the acidic activator domain of the viral protein VP16 has been shown to interact with components of the basal transcriptional machinery such as TFIIB and also with the $TAF_{II}40$ subunit of TFIID (Goodrich *et al.*, 1993) and to activate transcription in *in-vitro* transcription systems (Chen *et al.*, 1994).

The zinc finger factor Sp1 contains two glutamine-rich activator domains (Courey and Tjian, 1988; Courey *et al.*, 1989). Recent studies have shown that the glutamine rich domain of Sp1 can productively interact with the $TAF_{II}110$, a subunit of the TATA binding complex TFIID (Chen *et al.*, 1994; Hoey *et al.*, 1993; Gill *et al.*, 1994). Neither proline rich domains nor acidic domains functioned in this *in-vitro* transcription assay with $TAF_{II}110$, suggesting different activator domains may act on different components of the basal transcription machinery. Glutamine-rich domains have been identified in a range of factors including the *Drosophila* factors antennapaedia, bicoid, ultrabithorax and zeste (Courey *et al.*, 1989; Driever, 1989) and other mammalian factors such as Oct-1, Oct-2 (Herr *et al.*, 1988; Clerc *et al.*, 1988), JUN (Bohmann *et al.*, 1987) and Ap2 (Williams *et al.*, 1988b). Interestingly substituting the glutamine domains from antennapaedia and bicoid for Sp1 in the above experiments failed to activate transcription via interactions with the $TAF_{II}110$ subunit. Although these factors share similarities at the primary amino acid sequence level this does not guarantee that they can function via the same mechanism (Hoey *et al.*, 1993).

The third identified domain is the proline-rich domain first identified in the CTF/NF1 factor (Mermod *et al.*, 1989) and has been identified in a number of other factors including the zinc-finger proteins WT-1 (Call *et al.*, 1990), EKLF (Miller and Bieker, 1993) and NGFI-C (Crosby *et al.*, 1991). The proline rich domain has been shown to activate transcription in domain exchange experiments (Mitchell and Tjian, 1989).

Transcriptional repressor proteins have also been studied with the aim of identifying specific repressor domains. Domains have been mapped in the carboxy terminal region of the thyroid hormone receptor T_3 and also in the related retinoic acid receptor (RAR) protein (Baniahmad, 1992). Transcriptional repression activity has also been postulated for an alanine-rich domain found in four *Drosophila* proteins: Krüppel (Licht *et al.*, 1990), engrailed (Jaynes and O'Farrell, 1991), even-skipped (Han and Manley, 1993) and AEF-1 (Licht *et al.*, 1990).

Transcriptional repressor activity has been attributed to the *Krüppel* associated box (KRAB), a domain found in approximately a third of all class 1 zinc-finger proteins (Bellefroid *et al.*, 1991). More recently a non-zinc finger KRAB domain protein has been described suggesting this may be a more ubiquitous protein motif (Crew *et al.*, 1995). The KRAB domain is a 75 amino acid sequence which is subdivided into two boxes, A and B, each predicted to form an amphiphatic helix (Bellefroid *et al.*, 1991). The evidence for transcriptional repression is provided from three studies using a transient transfection assay (Margolin *et al.*, 1994; Witzgall *et al.*, 1994; Pengue *et al.*, 1994). Various recombinant fusion constructs expressing the KRAB domains from a total of seven KRAB containing zinc finger proteins linked to the DNA binding domain of GAL4 were used as the effector to transactivate chloramphenicol acetyltransferase reporter constructs. KRAB domain fusion products appeared to repress transcription and this repressing activity was found when only conserved residues in the A box helix were included (Margolin *et al.*, 1994). Repression appeared to be dependent on DNA binding by GAL4 and was not due to some "squelching" effect, however. More recently the KRAB domain of the human zinc finger protein KOX1 has been used to make a fusion protein with the DNA binding domain of the bacterial tetracycline repressor protein (TetR) (Deuschle *et al.*, 1995). The DNA binding of the TetR can be inhibited by addition of tetracycline, as is used in the VP16-TetR inducible transactivator system (Gossen and Bujard, 1992). In this experiment the authors show that the KRAB-TetR fusion protein can repress expression of a luciferase reporter gene when bound to *tet* operator sequences, coupled to either a cytomegalovirus promoter or to a HSV1 *tk* promoter. On addition of tetracycline the fusion protein can no longer bind to its DNA site and the repression is lost. Again this suggests that the KRAB domain can function as a transcriptional repressor domain and that this function is dependent upon DNA

binding. Furthermore, the authors find a 110 kDa protein coimmunoprecipitates with the KRAB-TetR fusion protein, which they suggest is involved in mediating the repressing effect of the KRAB domain (Deuschle *et al.*, 1995). The mechanism with which the KRAB domain represses transcription remains to be determined, as does the *in-vivo* relevance of these findings, however in these transfection experiments the KRAB domain can function to repress transcription.

1.4.4 Cloning and expression pattern of the murine *Zfp-37* gene

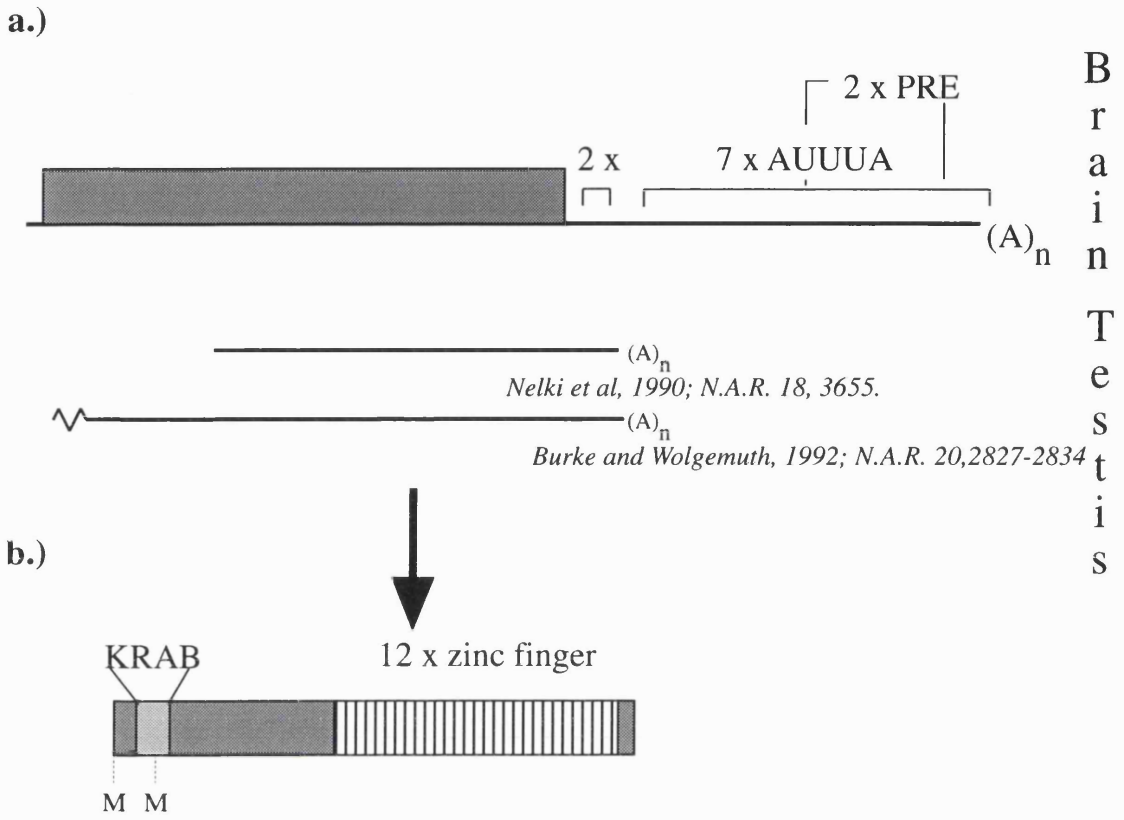
The murine *Zfp-37* gene was first identified as a 1.4 kb testis specific cDNA sequence (Nelki *et al.*, 1990). By sequence comparison the clone contained twelve contiguous *Krüppel*-like Cys_2His_2 zinc-fingers, positioned toward the carboxy terminus of the putative open reading frame. A second 1.9 kb cDNA clone was independently isolated from an adult testis cDNA library and shown to share 99.6% sequence homology with the original 1.4 kb clone (Burke and Wolgemuth, 1992) but with extended 5' sequence, containing a putative methionine translational initiation codon. In our laboratory, the 1.4 kb testis cDNA clone was used to screen Northern blots of RNA from developing mouse nervous system and under stringent conditions a message of approximately 3.7 kb was detected in the developing and adult brain (Mazarakis *et al.*, 1995). Subsequent construction and screening of a mouse P0 brain library identified a 3.4 kb cDNA clone (Figure 9). The 5' end of the transcript was cloned by rapid amplification of cDNA ends (RACE) analysis.

Sequence analysis of the products of the 5' end RACE analysis of e14.5 mouse brain revealed that a high degree of heterogeneity exists at the 5' end of the transcript. Sequence analysis of 19 5' end clones from the RACE analysis show three homology groups called I_a, I_b and II. Sequencing of the largest and most prominent group, I_a, identified a methionine initiation codon in frame with the twelve zinc fingers and upstream of the methionine described by Burke and Wolgemuth (1992). A composite message of I_a and the 3.4 kb cDNA sequence would encode a protein of 594 amino acids, approximately 67 kDa. No inframe methionine codons were described in I_b and II and it is predicted that these transcripts use the initiation codon described by Burke and Wolgemuth (1992) producing a protein of 553 amino acids (62 kDa) (Figure 9).

Figure 9. Structure of *Zfp-37* cDNAs and predicted proteins

a.) Organisation of 3.7 kb *Zfp-37* I_a cDNA as determined from sequence of cDNA clone and 5' RACE analysis. The protein coding region of the cDNA is shown as a shaded box. The locations of the nine AUUUA destabilising sequences are shown as are the PDGF responsive elements. The testis and brain polyadenylation sites are also marked. The extent of the published testis cDNA sequences are also shown.

b.) Predicted structure of ZFP-37 protein. Location of initiation codons are marked. The position of the KRAB domain and 12 zinc finger domains are also shown.


a.)
2 x PRE
2 x
7 x AUUUA
(A)n
Brain Testis
(A)n
Nelki et al, 1990; N.A.R. 18, 3655.
(A)n
Burke and Wolgemuth, 1992; N.A.R. 20,2827-2834
b.)
KRAB
12 x zinc finger
M M

All the predicted ZFP-37 proteins contain twelve Cys_2His_2 zinc-fingers located at the carboxy terminus. The fingers are connected by *Krüppel*-like amino acid linkers of the form TGEKPYE (Chavrier *et al.*, 1988; Chowdhury *et al.*, 1988). A nine amino acid tail of acidic residues follows the last finger. The amino-terminal part of the protein is rich in charged amino acid residues. Proteins of the I_a form contain an almost complete KRAbox at the amino terminus (Bellefroid *et al.*, 1991; Rosati *et al.*, 1991). Type I_b and II would only incorporate the B domain of the KRAbox. This homology classes *Zfp-37* into the KRAB containing group of class 1 zinc-finger genes (Mazarakis *et al.*, 1995).

The 3' untranslated region (UTR) of the brain derived cDNA is approximately 1.4 kb longer than the reported testis cDNA clones. A number of sequence elements reside within this region. The longer brain-derived cDNA contains 7 AUUUA sequence elements in addition to the two present in the shorter testis cDNAs. These AUUUA elements have been shown to shorten the half-life of mRNAs (Shaw and Kamen, 1986; reviewed by Decker and Parker, 1994). Two putative immediate response boxes (IRB) are also present (Freter *et al.*, 1992). IRBs were identified as TTTTGTA sequences in the 3'UTR of the immediate early gene encoding the JE cytokine. This *cis*-element acts at the transcriptional level in conjunction with an element 5' of the promoter region and is involved in the induction of the JE cytokine gene in response to serum, platelet derived growth factor (PDGF) and interleukin-1. The sequence has been identified in the 3'UTRs of a number of immediate early genes including zinc-finger transcription factors.

The tissue distribution of *Zfp-37* mRNA expression was first addressed by Burke and Wolgemuth (1992). Expression of a 4.0 kb mRNA was detected by Northern blot analysis in whole embryo RNA preparations from embryonic time points e10.5-e13.5. The same transcript was also detected in the placenta at these time points, no earlier developmental time points were included. In adult tissue samples two abundant messages of approximately 2.3 and 2.6 kb could be detected in the testis RNA as well as a less abundant message of approximately 4.0 kb. No expression of *Zfp-37* was detected in brain, heart, lung, kidney, liver, spleen, intestine and ovary. More detailed analysis during spermatogenesis showed expression to be restricted to post-meiotic spermatocytes and to be most abundant in round spermatids, this was confirmed by *in-*

situ RNA hybridisation (Burke and Wolgemuth, 1992). Similar findings were described by Hosseini *et al.* (1994).

In contrast to these published data, Northern blot RNA analysis in our laboratory has shown *Zfp-37* mRNA to be present in adult brain as a 3.7 kb message (Mazarakis *et al.*, 1995). This 3.7 kb message is also found in e10.5, e12.5, e14.5 and e16.5 head RNA samples (Mazarakis *et al.*, 1995). Furthermore expression of *Zfp-37* has been described in undifferentiated embryonic stem cells (Dr. N.Galjart unpublished results), suggesting expression may extend into earlier post-implantation and pre-implantation time points. *In-situ* RNA hybridisation data of e12.5 embryos indicate that *Zfp-37* is expressed most strongly in cells of the developing neuroepithelia. In the adult brain expression appears to be pan neuronal (Mazarakis *et al.*, 1995).

Polyclonal antibodies raised against recombinant ZFP-37 protein made in *E.coli* have shown a protein of approximately 70 kDa to be present in the testis (Hosseini *et al.*, 1994) and in brain (Dr. N.Galjart unpublished results). However, Hosseini *et al.* (1994) do not detect ZFP-37 protein in the brain using their antisera on Western blots. We have found ZFP-37 protein to be present in the testis by Western blot analysis though at a far lower concentration than the mRNA levels would predict (Dr. N.Galjart unpublished results). Hosseini *et al.* (1994) also detect a 40 kDa protein, which in purified germ cell preparations is only detectable in the elongating spermatids or residual bodies. The significance of this smaller protein is not known, the authors speculate that it may be a breakdown product of the larger 70 kDa protein. The cellular localisation of the ZFP-37 protein appears to be predominantly nuclear though some signal is also detected in the cytoplasm (Hosseini *et al.*, 1994), this is consistent with the proposed function as a transcriptional regulator.

1.4.5 Genomic organisation of the *Zfp-37* gene

A genomic cosmid clone of *Zfp-37* was isolated from a mouse genomic library from the 129 Sv/Ev mouse ES cell line, AB1 (Mazarakis *et al.*, 1995). Comparison of cDNA sequences with those of genomic subclones in combination with S1 nuclease analysis, have defined the exon-intron boundaries and have shown the *Zfp-37* gene to be located over a 20kb region. *Zfp-37* contains five exons, Ia, Ib, II, III and IV, the

genomic organisation is depicted in figure 10. Exon Ia encodes the first 14 amino acids Ib is non-coding, exon II and III encode the KRAB domain sequence, exon IV contains sequence for all twelve zinc fingers. The *Zfp-37* gene is unique in the mouse genome.

The different isoforms of ZFP-37 appear to be produced via alternate transcriptional start sites for exon Ia, Ib and exon II and alternative splicing. S1 analysis and sequence of RACE products have shown multiple start sites for exon Ia, sequence of RACE products also show that alternate splice donor sites can be used at the exon Ia/intron 1 boundary, increasing or decreasing the coding region by 9 bp. Alternate splice acceptor sites are also present at the intron 1/exon II boundary differing by 3bp. The transcription initiation site(s) of exon Ib are yet to be mapped. From the RACE analysis of exon Ib it appears that only one splice donor site is used, whereas either of the splice acceptor sites of exon II are utilised. A number of RACE products start at exon II suggesting the possibility of a promoter region within intron 1, though this has yet to be confirmed. The predicted exon usage at the 5' end of the *Zfp-37* gene is shown in Figure 10. Alternate exon use allows *Zfp-37* to encode different protein isoforms. If exon Ia is used all of exon II is coding and the KRAB A and B domains are produced. However if exon Ib is used or transcription starts at exon II then the methionine at exon II will be used and only the KRAB B domain will be encoded for (Mazarakis *et al.*, 1995). The representation of these isoforms in testis and in adult brain has yet to be described, although polyclonal antisera raised against the recombinant amino terminal region of the protein suggest that the KRAB A form (exon Ia) is the major isoform in adult brain in Western blot analysis (Dr. N. Galjart unpublished results).

Sequence comparison of genomic subclones with cDNA sequence has identified a 17 bp sequence described in the 5' end of the testis cDNA sequence of Burke and Wolgemuth (1992) which is homologous to a sequence downstream of exon Ib. This is followed by a putative splice donor site and has therefore been described tentatively as exon It and may represent a sixth exon.

The promoter region of *Zfp-37* has not been fully characterised. However, no TATA box sequences have been identified upstream of exon Ia. The region covering exon Ia is rich in CpG dinucleotides and is thought to represent a CpG island (Bird, 1986) although the methylation status of this region has not been investigated.

Through genetic linkage mapping *Zfp-37* has been localised to chromosome 4

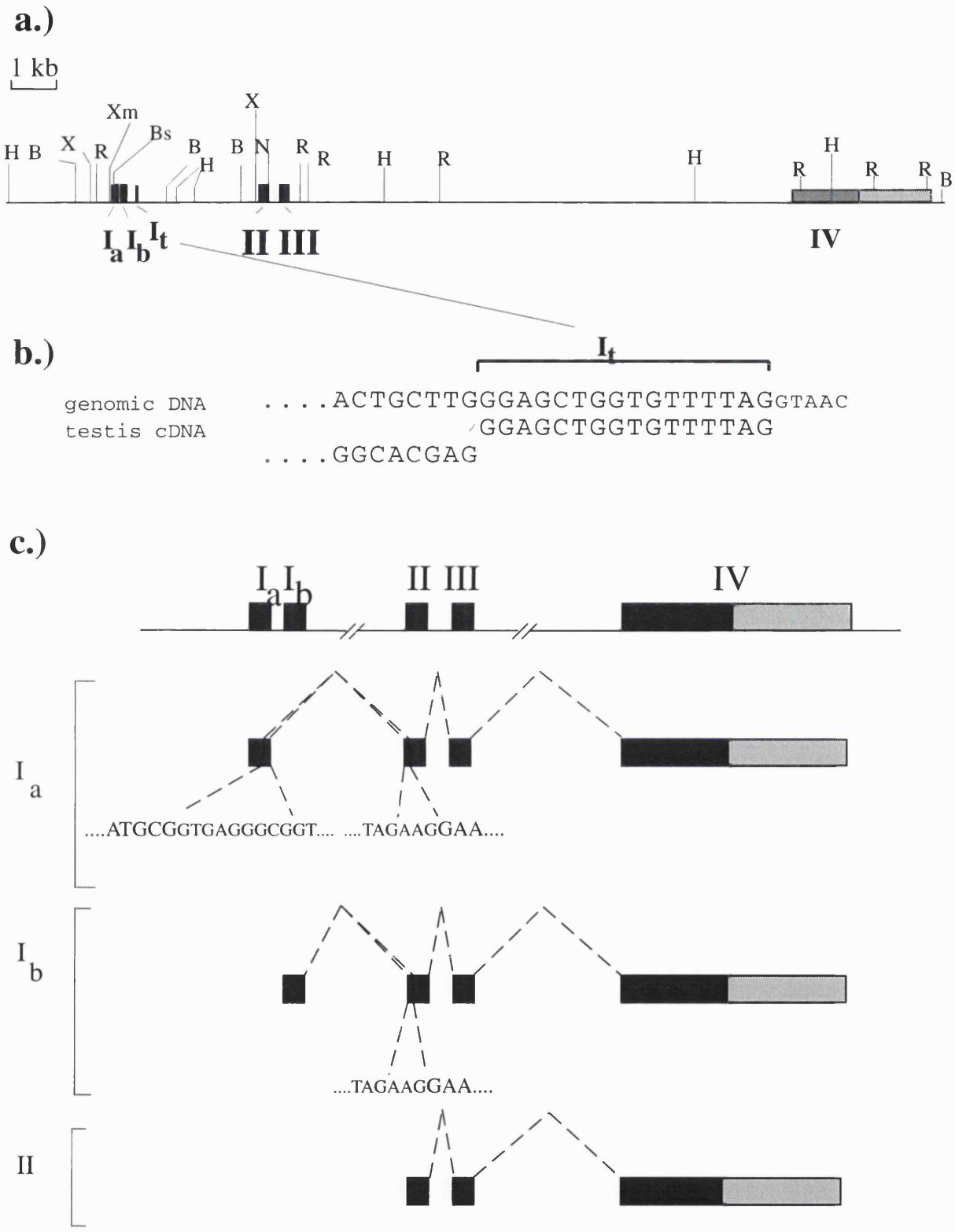
Figure 10. Genomic organisation and splicing pathways of the *Zfp-37* gene

A) Organisation of the *Zfp-37* gene. The *Zfp-37* gene spans a region of approximately 20 kb. The five exons Ia, Ib, II, and IV are shown as boxes. Exon IV contains the 3'UTR which is shown as a shaded box. Partial restriction enzyme map is also shown. H HindIII, B BamHI, R EcoRI, Xm XmaI, Bs BspEI, N NcoI.

B) Location of the putative exon I_t and homology to published cDNA sequence (Burke and Wolgemuth, 1992). The 5' end of the published testis cDNA was found to align perfectly to a genomic DNA sequence of 17 nucleotides, and was therefore termed I_t. The first seven nucleotides of the cDNA were not homologous to the genomic sequence and originate from a cloning vector.

C) Putative *Zfp-37* transcription initiation sites and modes of alternative splicing of *Zfp-37* transcripts in the brain. 5'end RACE analysis was used to establish the extent of the full length cDNA. Results from this analysis suggested that in the brain transcription initiation takes place at three different positions within the gene at exons Ia, Ib and II. Splicing of exon Ia to exon II occurs via two splice donor and acceptor sites located 9 and 3 nucleotides apart respectively, as shown by sequence given below Ia-II splicing pathway. Splicing of Ib to II makes use of the same two acceptor sites.

Adapted from Mazarakis *et al.*, 1995.


a.)
1 kb
H B X R Xm Bs B H B N X R R H R H R H R R R B
I_a I_b I_t
II III
IV
b.)
I_t
genomic DNA ....ACTGCTTGGGAGCTGGTGTTTTAGGTAAC
testis cDNA GGAGCTGGTGTTTTAG
....GGCACGAG
c.)
I_a I_b
II III
IV
I_a
....ATGCGGTGAGGGCGGT.... ....TAGAAGGAA....
I_b
....TAGAAGGAA....
II

in the mouse, mapping within 3-4 M of the major urinary protein (*Mup-1*) and *whirler* loci (Fleming *et al.*, 1994). This has been confirmed in this laboratory via fluorescent *in-situ* hybridisation of *Zfp-37* cosmid to mouse metaphase chromosomal spreads (N. Galjart unpublished).

1.4.6 The function of *Zfp-37*

At present there is little functional data for ZFP-37, clues as to its role can only be drawn from its sequence and expression pattern. *Zfp-37* has the potential to encode twelve zinc finger motifs. Sequence alignment of the fingers with the consensus sequence for Cys_2His_2 zinc finger place it in the class 1 group of finger genes. Each finger is linked by a seven amino acid sequence conforming to the consensus sequence TGEKPYE, assigning the gene to the *Krüppel*-like class of zinc fingers. Alignment of the twelve fingers with themselves shows that similar amino acid residues are present at the edges of the zinc finger domain but the residues predicted to contact DNA are different (Nardelli *et al.*, 1991; Paveltich and Pabo, 1991). We have been unable to align any of the twelve zinc fingers in ZFP-37 with any other zinc finger protein with a known DNA binding site and we are therefore unable to predict the DNA sequence which ZFP-37 could bind.

Identification of a KRAB domain in the major isoform of ZFP-37 places it within this subfamily of zinc finger genes. The exon/intron organisation of the *Zfp-37* gene also shares a number of common features with other KRAB containing zinc finger genes. Encoding of the KRAB A and B domains by two exons has been documented for the human KRAB finger gene *ZNF91* and related genes (Bellefroid *et al.*, 1993). As is the case for ZFP-37, different isoforms varying in their respective KRAB domains have been described for the human proteins ZNF43 (Bellefroid *et al.*, 1991; Lovering and Trowsdale, 1991) and ZNF2 (Rosati *et al.*, 1991). Thus, the organisation and alternate use of the 5' KRAB exons appears to be a common feature of KRAB domain zinc finger genes (Bellefroid *et al.*, 1991). If the KRAB A domain functions as a transcriptional repressor, as has been suggested (Deuschle *et al.*, 1995; Margolin *et al.*, 1994; Witzgall *et al.*, 1994; Pengue *et al.*, 1994), then the alternate use of exon Ia or Ib would allow the *Zfp-37* gene to produce functionally diverse products. A second

feature shared by KRAB zinc finger genes is the encoding of the zinc finger cluster by one large exon, again this is the case for *Zfp-37*.

The *in-vivo* role of KRAB zinc fingers proteins is not known. No specific DNA binding site has been described for any of the members of this class of protein. Bacterially expressed human ZNF85 zinc-finger protein has been mentioned to be able to bind both DNA and RNA though non-specifically (Bellefroid *et al.*, 1993). However, function has been assigned to other multi zinc finger proteins. Recently the Xenopus zinc finger protein XFG 20, containing 19 Cys_2His_2 fingers, was shown to specifically recognise a 54 bp sequence which is part of a dispersed repetitive element in the Xenopus genome, termed REM-1 (Schafer *et al.*, 1994). The authors speculate that the diversity of multi zinc finger gene families that appear in evolution may correlate with the diversity of repetitive elements found in different organisms. Further to this theory, binding of a repetitive element in the genome has also been described in *Drosophila* for the twelve zinc finger protein suppresor of hairy wing, su(Hw). This finger protein binds multiple octameric sequences located downstream of the 5' long terminal repeat of the *Gypsy* retrotransposable element (Parkhurst *et al.*, 1988). Again a repetitive element which occurs many times in the genome. Insertion of *Gypsy* can cause gene activation ~~mutations~~ over large distances from the site of insertion by affecting the transcriptional status of loci. su(Hw) functions to suppress these affects and recent studies suggest this is achieved by insulating position effects caused by *Gypsy* insertion possibly by setting up specific chromatin boundaries (Roseman *et al.*, 1993). su(Hw) has been proposed to mediate this affect by altering the DNA topology, which can be demonstrated by DNA bending *in-vitro* (Shen *et al.*, 1994). It is possible that other multi finger proteins such as ZFP-37 function in an analogous role.

Identification of IRB sequences within the 3'UTR of *Zfp-37* allows one to speculate that *Zfp-37* may be expressed as an immediate early gene in response to external growth signals (Freter *et al.*, 1992). IRB sequences have been identified in other murine zinc finger immediate early genes including Krox 20, NGFI-A, NGFI-B, Egr-1 and Zif268 (Freter *et al.*, 1992). The nine AUUUA destabilising elements found in the 3'UTR of the 3.7 kb message suggest that the brain mRNA has a higher turnover rate than the testis mRNA which contains only two of these elements. A high mRNA turnover rate would be compatible with the hypothesis that *Zfp-37* is an immediate early

response gene. Interestingly the CREM gene, which is expressed in both brain and testis, also contains nine AUUUA destabilising sequences in the longer brain 3'UTR (Foulkes *et al.*, 1993). Like *Zfp-37,* CREM also shows a high expression in the testis but lower levels of expression in the brain. This was shown to be due to the presence of the nine destabilising sequences in the 3'UTR in the brain mRNA (Foulkes *et al.*, 1993).

Burke and Wolgemuth (1992) and Hosseini *et al.* (1994) have both suggested that ZFP-37 functions as a transcriptional regulator during spermatogenesis. This is based upon the zinc finger structure and regulated expression during this process. The cellular localisation has been shown to be predominantly nuclear (Hosseini *et al.*, 1994) adding further credence to this hypothesis.

In this laboratory we have also shown expression of *Zfp-37* in neuronal tissue and hypothesise that ZFP-37 is playing a role in transcriptional regulation in this tissue. This may well be the case. However, ZFP-37 may not necessarily function in the manner of a classical site specific transcription factor, a function which can be attributed to Cys_2 His_2 zinc finger proteins of the GLI, Krox 20, Sp1 and Wt-1 families. Sequence comparison does not class ZFP-37 into these groups of proteins but into the much larger family of KRAB zinc finger genes whose function are not known.

In the context of other functional studies aimed at dissecting the function of ZFP-37 a gene targeting approach is to be taken. This approach should shed light upon the importance of the *Zfp-37* gene during development and adult life and may also identify the function of KRAB domain containing zinc finger proteins.

Chapter 2

Targeting of the murine *Thy-1.2* gene in mouse embryonic stem cells

2.1 Introduction and strategy

Thy-1 was described over thirty years ago as an allo-antigen on murine lymphocytes (Reif and Allen, 1964). It has subsequently been shown to be a major protein constituent of the membranes of mature neurons (reviewed by Morris, 1985; Morris and Grosveld, 1989; Morris, 1992). Thy-1 displays a conserved neuronal distribution in all species in which it has been described. Furthermore, the fine control of Thy-1 distribution during the cessation of axonal growth coupled with a marked inhibitory effect on neurite outgrowth *in-vitro* has suggested a role for Thy-1 in the process of axonal outgrowth regulation (Xue *et al.*, 1990; Xue *et al.*, 1991; Xue and Morris, 1992; Tiveron *et al.*, 1992).

In the mouse haematopoietic system Thy-1 again shows a strict pattern of regulation and has been shown *in-vitro* to mediate mitogenic response. The fact that Thy-1 has been found associated with a number of kinases and phosphatases is indicative of it being a part of a cell signalling pathway (Stefanova *et al.*, 1991; Shenoy-Scaria *et al.*, 1992; Thomas and Samelson, 1992; Volarevic *et al.*, 1990). It has also been shown that Thy-1 can act as an adhesion molecule involved in the interactions of thymocytes with their thymic environment (He *et al.*, 1991).

Attempts to ascertain Thy-1 function *in-vivo* have been made. Gain of function transgenic experiments increasing the concentration of Thy-1 on nerve and thymocytes by 7-16 fold have not resulted in an aberrant phenotype suggesting that Thy-1 is already in abundance and its function is dependent on a putative ligand concentration (Kollias *et al.*, 1987). Ectopic expression of Thy-1 on B-cells has resulted in a proliferative abnormality of these cells (Chen *et al.*, 1987). The exact mechanism of the hyperplasia is not known, but the proliferative effect seen in the B-cells is consistent with a normal role in T-cell proliferation, as suggested from the mitogenic repsonse induced by anti-Thy-1 antibodies on T-lymphocytes.

To asses the *in-vivo* role of Thy-1 it was proposed to create loss of function murine mutants via gene targeting technology. Thy-1 has already been extensively studied in this laboratory. The murine gene has been isolated, its exon/intron organisation elucidated and cis-acting elements which direct tissue specific expression have been located via experiments in transgenic animals (Giguere *et al.*, 1985, Kollias

et al., 1987, Spanopoulou *et al.*, 1988 and 1991; Vidal *et al.*, 1990). Therefore a great deal is already known about the expression pattern of Thy-1 and the organisation of the *Thy-1* gene and reagents are readily available. Targeting of *Thy-1* has previously been carried out in this laboratory (Watson, 1992) using the CCE embryonic stem cell line (Robertson *et al.*, 1986). ES cells which had undergone a correct recombination event at the *Thy-1* locus were isolated but failed to transmit the mutation through the germline of chimeric mice. It was subsequently shown that the subclone of the CCE line used carried a chromosomal translocation which was thought to account for the inability of the ES cell line to contribute to the germ line of the chimeric animals (Watson, 1992).

In this chapter I describe the targeting of the *Thy-1* gene in a different ES cell line with the aim of producing chimeric mice that will pass on the mutated *Thy-1* allele. The ES cell line AB1 (McMahon and Bradley, 1990) was made available to us and this line would therefore be used to repeat the targeting of *Thy-1*. As previous work in this laboratory has located the DNA elements which direct tissue specific regulation of the gene in transgenic mice to either brain or thymus (Vidal *et al.*, 1990) it could be envisaged that, depending on the severity of the Thy-$1^{-/-}$ phenotype, rescue constructs could be introduced as a transgene to complement the mutation in either brain or thymus. These transgenes could be used to assess the importance of Thy-1 in the neuronal or thymic enviroment or could be used to introduce specific Thy-1 mutants which could then be studied in a null background. Analysis of the Thy-$1^{-/-}$ mice is proposed to be carried out in collaboration with the laboratory of Dr. R. Morris, who has described the expression of Thy-1 in the rodent nervous system in detail and who has also developed an *in-vitro* system to study Thy-1 function.

2.2 Results

2.2.1 *Thy-1* targeting vector-design

The vector Thy-neo5'3' was designed previously and is a replacement type targeting vector (Watson, 1992). The vector Thy-neo5'3' was successfully used to obtain targeted clones in the CCE line (Watson, 1992) and this would therefore be used as the basis for the targeting vector used in the AB1 ES cells. In the Thy-neo5'3'

targeting vector the *Thy-1* sequence is disrupted using the MC1NeopolyA cassette (Thomas and Capecchi, 1987). This cassette is designed to give efficient transcriptional levels of the neomycin resistance gene in embryonic stem cells (Thomas and Capecchi, 1987). MC1NeopolyA contains the neomycin resistance gene from the bacterial transposon Tn5 (Beck *et al.*, 1982), under the transcriptional control of mutated polyoma virus enhancer (Fujimura *et al.*, 1981) linked to the herpes simplex virus type 1 (HSV1) thymidine kinase (tk) promoter region (McKnight, 1980). A polyadenylation sequence is provided by the SV40 polyadenylation site. The translational initiation site has also been modified with the aim of improving efficiency. In the vector Thy-neo5'3' the neomycin resistance cassette is inserted at the BstEII site located in exon three of the *Thy-1* sequence and disrupts the coding region of Thy-1 between codons 21 and 22, which are part of the extracellular domain of the protein. The transcriptional orientation of *Thy-1* and the neomycin resistance genes are identical.

The vector contains a total of 5.4 kb of homology to the *Thy-1* gene. This is distributed as 4.6 kb 5' and 813 bp 3' of the MC1NeopolyA disruption. The *Thy-1* sequences used in this construct were cloned from a non-isogenic cosmid clone of the Thy-1 gene (Giguere *et al.*, 1985).

The Thy-neo5'3' vector was successfully used in the CCE ES cell line to target *Thy-1* (Watson, 1992). In this experiment only selection for the neomycin resistance gene was employed and the screening of clones was carried out using the polymerase chain reaction according to the protocol of Zijlstra *et al.* (1989). From 1 x 10^7 CCE ES cells transfected, five independent targeting events were identified (Watson, 1992). The absolute targeting efficiency with this vector was 1 x 10^{-6} assuming 50% of the cells survived the electroporation. This is in the expected range of absolute targeting efficiency for a non-isogenic vector with 5.4 kb homology (Deng and Capecchi, 1992). Since, from the experiments with the CCE line, it was clear that a large number of clones would need to be screened, it was proposed to use the positive/negative selection protocol of Mansour *et al.*, (1988) in the targeting experiment with the AB1 ES cells. This allows an enrichment of the ES cell clones which have undergone homologous recombination to be made (Mansour *et al.*, 1988). Such a selection procedure has resulted in an enrichment of up to 2000 fold as compared to single selection when targeting the *Hprt* and *Int-2* loci (Mansour *et al.*, 1988). Positive/negative selection has

also been successfully used for targeting the *Int-2* locus in the AB1 ES cell line (McMahon and Bradley, 1990). However, when positive/negative selection was attempted for targeting of *Thy-1* it was unsuccessful in yielding targeted clones (Watson, 1992). In this case the vector used possessed only 3.3 kb of homology to the target locus and it is therefore likely that the length of homology used in this targeting vector was too short to support efficient recombination.

As discussed earlier, the HSV1 thymidine kinase gene is commonly used as a negative selection marker in combination with the nucleoside analogues GANC or FIAU (Hasty and Bradley, 1993). The thymidine kinase expression cassette, py*tk*, was constructed in this laboratory. It comprises the repeated PyF441 polyoma enhancer sequence employed in pMC1NeopolyA (Fujimura *et al.*, 1981), cloned 5' to the HSV1 thymidine kinase gene in the pBluescript KS+ cloning vector and was shown to be functional (Watson, 1992). The Thy-neo5'3' vector was modified by inclusion of the py*tk* cassette as a negative selection marker.

2.2.2 *Thy-1* targeting vector-construction

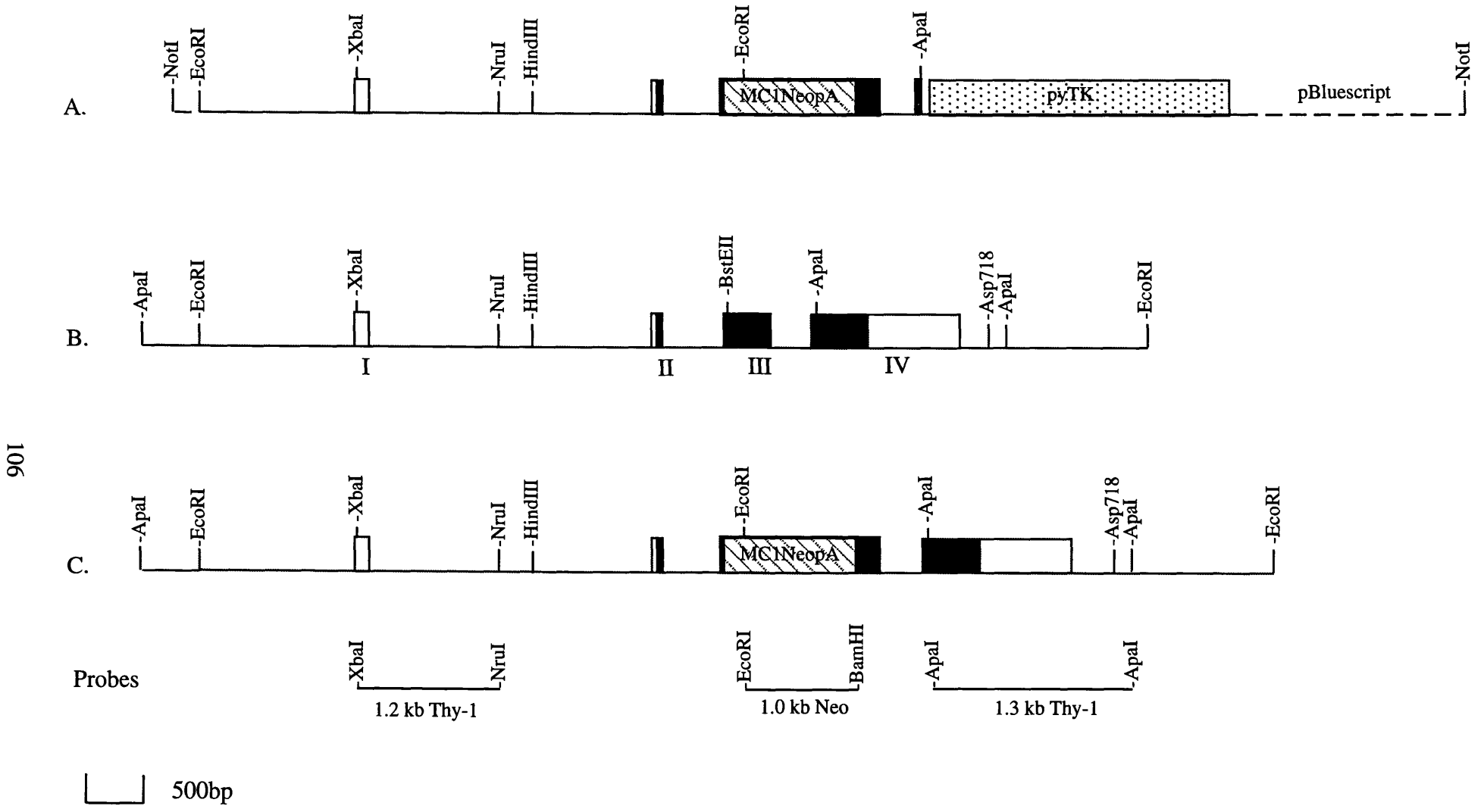
An EcoRI-BamHI subclone of the murine *Thy-1.2* gene containing the pMC1NeopolyA cassette inserted into the BstEII site in exon three (Watson, 1992) was excised from the vector pPolyIII as a NotI fragment and cloned into the vector pBluescript II SK+. The 3' ApaI fragment from this clone was deleted which removes part of exon IV allowing this region to be used to generate probes in the subsequent screeening of the ES cell clones. A blunted 2.3 kb Asp718-BamHI fragment, containing the py*tk* selection cassette, was subsequently cloned into the blunted Asp718 site located 3' of the *Thy-1* insert. The resulting vector carried the same distribution of homology as Thy-neo5'3' but in addition carried the py*tk* negative selection cassette in the same transcriptional orientation as both the *Thy-1* gene and the neomycin resistance gene. The vector was linearised at the unique NotI site located at the 5' end of homology, to create a replacement type targeting vector possessing a total of 5.4 kb of homology to the target site. Figure 11 shows the targeting vector and predicted structure of the correctly targeted *Thy-1* locus.

Figure 11. *Thy-1* targeting vector and predicted organisation of targeted *Thy-1* allele.

A.) Diagrammatic representation of the 5.4 kb *Thy-1* targeting vector. The 1.1 kb MC1NeopolyA cassette (shown as diagonally marked box) is cloned into the blunted BstEII site in exon 3 as described by Watson (1992). The vector contains 4.6 kb of homology to the *Thy-1* gene 5' of the MC1NeopolyA disruption and 813 bp 3'. The 2.2 kb py*tk* negative selection cassette (shown as dotted box) is cloned into the blunted Asp718 site located in the pBluescript vector (shown as dashed line). The vector is shown linearised at the NotI restriction enzyme site of pBluescript.

B.) Diagrammatic representation of the wild type *Thy-1* allele showing exons I-IV. White boxes represent untranslated regions, black boxes represent protein coding regions.

C.) Predicted structure of the correctly targeted *Thy-1* allele. The diagnostic probes used for the Southern blot analysis of the targeted allele are shown below. The 1.2 kb XbaI-NruI and 1.3 kb ApaI *Thy-1* probes were isolated from an 8.2 kb clone of *Thy-1.2* (Giguere *et al.*, 1985). The 0.6 kb Neo probe was isolated from the plasmid pMC1NeopolyA (Thomas and Capecchi, 1987).


A.
-NotI
-EcoRI
-XbaI
-NruI
-HindIII
-EcoRI
-ApaI
MC1NeopA
pyTK
pBluescript
-NotI
B.
-ApaI
-EcoRI
-XbaI
-NruI
-HindIII
-BstEII
-ApaI
-Asp718
-ApaI
-EcoRI
I
II
III
IV
C.
-ApaI
-EcoRI
-XbaI
-NruI
-HindIII
-EcoRI
MC1NeopA
-ApaI
-Asp718
-ApaI
-EcoRI
Probes
XbaI
NruI
1.2 kb Thy-1
EcoRI
BamHI
1.0 kb Neo
ApaI
ApaI
1.3 kb Thy-1
500bp

2.2.3 Transfection and enrichment

The AB1 ES cell line is derived from the 129/Sv/Ev mouse strain (McMahon and Bradley, 1990) and carries the agouti/agouti white belly coat colour alleles. This line has been successfully employed in the targeting of the *wnt-1(Int-2)*, *c-src*, myogenin and *Hoxb-4* genes (McMahon and Bradley, 1990; Soriano *et al.*, 1991; Hasty *et al.*, 1993; Ramirez-Solis *et al.*, 1993). The AB1 ES cell line was made available to us by Dr. Allan Bradley. ES cells were maintained and subcultured according to Robertson (1987) on SNL76/7 feeder cell layers which were mitotically inactivated with mitomycin-C. SNL76/7 is a subline of the STO mouse fibroblast cell line (Ware and Axelrad, 1972). This line has been stably transfected with a neomycin resistance cassette and a construct expressing the secreted form of the molecule LIF (McMahon and Bradley, 1990).

For transfection experiments, three 0.8 ml aliquots of 1 x 10^7 AB1 ES cells were prepared. To two of the aliquots the linearised targeting vector was added (final concentration 25 μg/ml) the third served as mock transfected control. The cells were electroporated as described by McMahon and Bradley (1990). After each electroporation the cells were diluted in normal ES media and plated. An aliquot (1/200 dilution) was taken and plated, this was used as a G418 selection control in the case of the vector transfections. The mock transfection was used as a control for G418 killing in the abscence of the targeting vector. The 1/200 aliquot was taken to asses cell death after electroporation, a "no shock" control was also plated at the same cell density to serve as a comparison.

Twenty four hours after electroporation G418 (350μg/ml) and GANC (2μM) selection was applied to the cultures. The 1/200 aliquots were put under G418 selection only to assess transfection efficiency. The survival rate after electroporation was approximately 50% as determined from cell counts of shocked versus unshocked controls twenty four hours after plating.

Mock transfected ES cells began to die in G418 selection at day 3 of drug selection. No cells could be seen by day 8. In contrast, ES cells transfected with the targeting vector gave rise to colonies which were visible from day 8 onward in both G418 and G418/GANC selection media. The number of G418 and G418/GANC

resistant colonies obtained are shown in Table 1. G418/GANC selection gave an approximate 16-fold enrichment over G418 selection alone.

Between 9 and 11 days after transfection a total of 82 $G418^r/GANC^r$ colonies were picked from the two independent transfection experiments. Each colony was plated into duplicate wells of a 24 well multi well plate and then expanded to confluency in 60mm plates in G418 selection only. One of the duplicate cultures was frozen, the other was used for DNA isolation. From the 82 colonies picked 29 were lost during this process, some of these loses were due to differentiation of the ES cell clones, unstable integration of the neomycin resistance gene may have accounted for other loses.

2.2.4 Screening for homologous recombination at the *Thy-1* allele

A Southern blot screening procedure was used to identify targeted clones. Correct recombination of the targeting vector into the *Thy-1* locus introduces a diagnostic EcoRI site into the targeted allele which is derived from the 5' end of the MC1NeopolyA cassette. Hybridisation with a 3' *Thy-1* ApaI probe, located in exon 4 (Figure 11), detects an 8.2 kb EcoRI fragment in the wild type allele, whereas in a correctly targeted allele a 4.5 kb hybridising fragment will be detected, due to the new EcoRI site. As the ApaI probe is external to the region of homology shared between locus and targeting vector, it will only detect homologous recombination events and not random integrations of the vector. This assay allows both the detection of recombination events at the *Thy-1* allele and at the same time demonstrates the integrity of the 3' recombination junction. To show the EcoRI site is introduced by the MC1NeopolyA cassette, the filter is stripped of the ApaI probe and hybridised with a 600bp Neo specific probe, which should detect the same fragment. One clone, 3.10D.2 was positive for this assay (Figure 12). In other $G418^r/GANC^r$ clones random integration events can be seen when probing for neo sequences.

To determine the integrity of the 5' recombination junction, restriction sites 5' to sequences encompassed by the targeting vector were mapped. Restriction enzymes known to cut the *Thy-1* gene 3' of the *neo* insertion were used to map convenient 5'sites. A 1.2 kb Xba-NruI probe located over exons Ia and Ib was used to assess these sites (Figure 11). The enzymes tested were ApaI, AflIII, EagI, EcoRV and SphI of

Table I. *Thy-1* targeting experiment

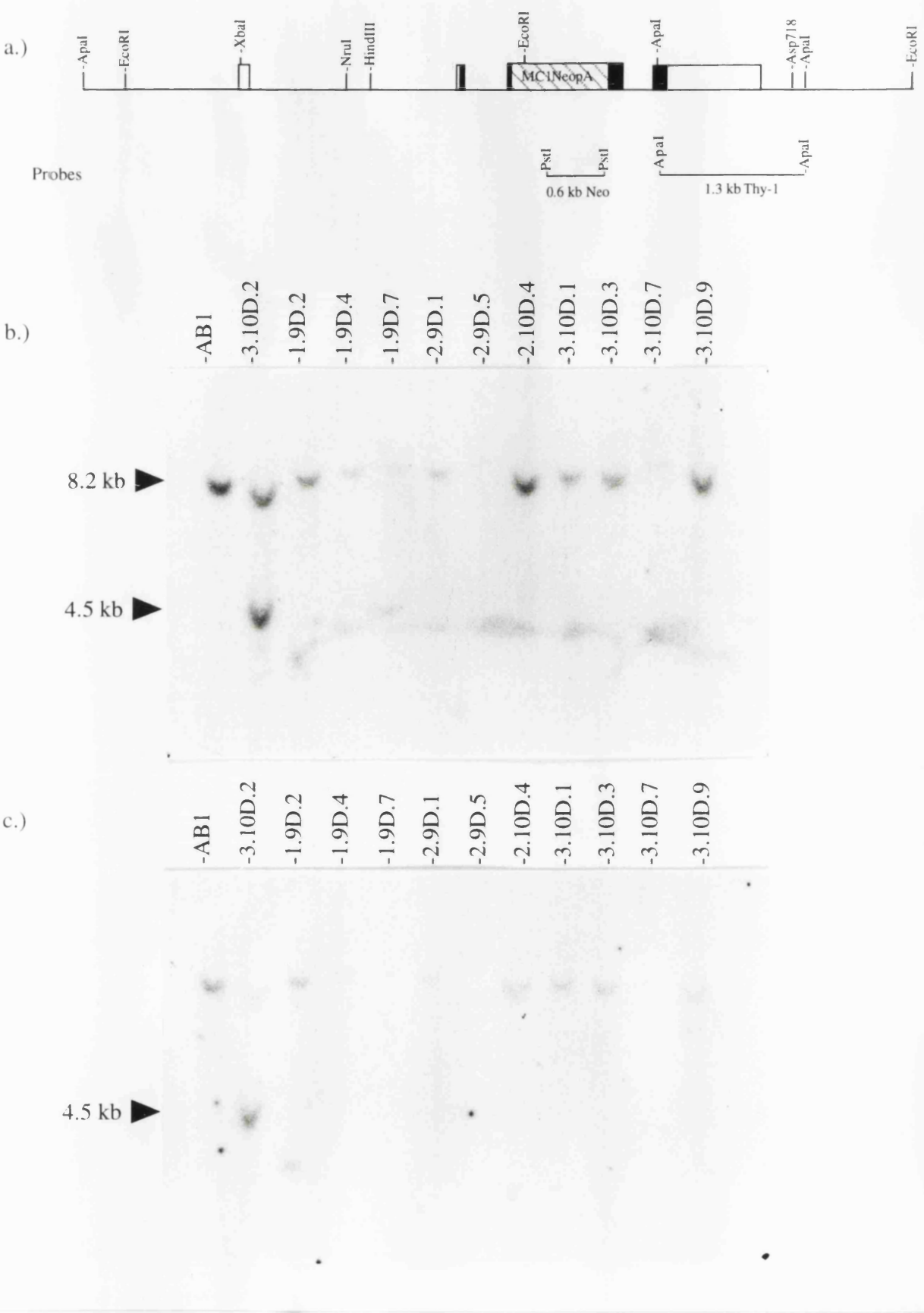
Experiment	No. of cells transfected	a) No. of G418^r colonies	b) No. of G418^r/GANC^r colonies	b/a
1	1x 10^7	1512	84	1/18
2	1x 10^7	1155	80	1/14.5

Figure 12. Southern blot analysis of G418^r/GANC^r ES cell clones from Thy-1 targeting experiment.

a.) Predicted structure of targeted *Thy-1* allele and position of diagnostic probes used in establishing the integrity of the 3' recombination junction.

b.) Southern blot analysis of eleven of the 53 G418^r/GANC^r clones obtained in the Thy-1 targeting experiment. ES cell clone DNA was digested with EcoRI restriction enzyme and the filter hybridised with the 1.3 kb ApaI 3' probe which is external to the 3' recombination junction. The wild type *Thy-1* allele is seen as an 8.2 kb hybridising band as can be seen in the AB1 lane, the targeted allele hybridises as a 4.5 kb band due to the additional EcoRI site introduced into the *Thy-1* locus via the MC1NeopolyA cassette.

c.) Filter shown in panel b. partially stripped of 1.3 kb ApaI 3' probe and rehybridised with the 0.6kb Neo probe. In sample 3.10D.2 the 4.5 kb band can be seen to hybridise with the Neo probe, showing the introduced EcoRI polymorphism to be due to the recombination of the MC1NeopolyA cassette into the *Thy-1* allele. Random integrations of the targeting vector can be seen in the other samples.


a.)
ApaI
EcoRI
XbaI
NruI
HindIII
EcoRI
MC1NeopA
ApaI
Asp718
ApaI
EcoRI
Probes
PstI
PstI
0.6 kb Neo
ApaI
ApaI
1.3 kb Thy-1
b.)
-AB1
-3.10D.2
-1.9D.2
-1.9D.4
-1.9D.7
-2.9D.1
-2.9D.5
-2.10D.4
-3.10D.1
-3.10D.3
-3.10D.7
-3.10D.9
8.2 kb
4.5 kb
c.)
-AB1
-3.10D.2
-1.9D.2
-1.9D.4
-1.9D.7
-2.9D.1
-2.9D.5
-2.10D.4
-3.10D.1
-3.10D.3
-3.10D.7
-3.10D.9
4.5 kb

which ApaI was chosen as a suitable digest. In a wild type allele a 6.2 kb band (Figure 13) is detected with the aforementioned probe. Whereas in a targeted allele this fragment will appear larger by 1.1kb because of the inclusion of the MC1NeopolyA cassette, giving a diagnostic 7.3 kb fragment. Clone 3.10D.2 was tested using the ApaI diagnostic digest and probed with the internal XbaI-NruI probe. The predicted 7.3 kb band was detected in this sample and the 3' and 5' mapping of the targeted locus is shown in Figure 14. This confirms the correct targeting of one allele of the *Thy-1* gene in the clone 3.10D.2. Since only one clone showed a correct targeting event from the 53 clones screened, a conclusion on the absolute efficiency of this vector can not be drawn precisely.

2.2.5 Generation of chimeric mice

A series of seven blastocyst injections were carried out by Dr. M. Lindenbaum using the targeted ES cell clone 3.10D.2, from which one male chimeric mouse was obtained. This chimera showed a low degree of ES cell contribution, 5-10%, as determined from the coat colour chimerism. At six weeks of age the chimera was set up to breed with C57BL/6 females. From two matings nine offspring were obtained. No transmission events were seen in these litters as determined by coat colour assay. As small litters were obtained with C57BL/6 mouse strain, the chimera was also bred with CBAxC57BL/10 F1 females. Since the chimera was being bred into an agouti genetic background, the offspring were screened for the EcoRI polymorphism at the DNA level. In total 45 offspring were screened by both methods (data not shown), no germline transmission of the mutated allele was detected.

Therefore, although clone 3.10D.2 was correctly targeted, it appeared not to be able to contribute efficiently to the formation of a chimeric mouse when injected in to blastocysts. The lack of germline transmission from this clone is unlikely to be due an inherent defect in the original ES cell line used. Altough at the time of this experiment wild type ES cells had not been tested for germline transmission, the AB1 ES cell line was later used successfully in this lab (Pandolfi *et al.*, 1995; and this study) and therefore it does retain germline contributing capability. It seems more likely that during the targeting experiment the ES cells had entered some differentiation pathway or

Figure 13. Mapping of 5' restriction enzyme sites in the murine *Thy-1* gene.

Southern blot analysis of AB1 ES cell genomic DNA digested with EcoRV, ApaI, SphI EagI and AflIII, and hybridised with the 1.2 kb XbaI-NruI probe (Figure 11).

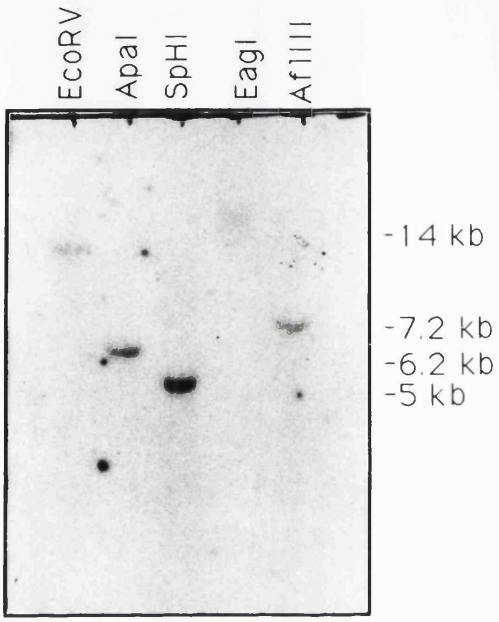

EcoRV
ApaI
SphI
EagI
AflIII
-14 kb
-7.2 kb
-6.2 kb
-5 kb

Figure 14. Southern blot analysis of both the 3' and 5' recombination junctions of the targeted *Thy-1* allele.

a.) Organisation of the targeted *Thy-1* allele. The positions of the diagnostic XbaI-NruI 1.2 kb *Thy-1* probe internal to the targeting vector and external 1.3 kb ApaI probe are shown below.

b.) Repeat Southern blot analysis of EcoRI digested ES cell DNA mapping the 3' recombination junction using the external 1.3 kb *Thy-1* probe. Non-transfected AB1 DNA and a non-targeted clone 4.10D.1 are included as well as the targeted clone 3.10D.2. The difference in hybridisation intensities between the 8.2 kb and 4.5 kb bands in sample 3.10D.2 has been ascribed to contamination by feeder cell DNA which will result in a more intense wild type signal.

c.) Southern analysis of ApaI digested ES cell DNA, mapping the 5' recombination junction using the internal 1.2 kb XbaI-NruI *Thy-1* probe. Clone 3.10D.2 shows the wild type allele band of 6.2 kb and the predicted 7.3 kb band resulting from the inclusion of the 1.1 kb MC1NeopolyA selectable marker gene in the targeted allele. In clone 4.10D.1 random integration of the targeting vector is evident has a higher molecular weight hybridising band.

a.)
-ApaI
-EcoRI
-XbaI
-NruI
-HindIII
-EcoRI
MC1NeopA
-ApaI
-Asp718
-ApaI
-EcoRI
Probes
XbaI
1.2 kb Thy-1
NruI
ApaI
1.3 kb Thy-1
ApaI
b.)
8.2 kb
4.5 kb
-4.10D.1
-3.10D.2
-AB1
c.)
-4.10D.1
-3.10D.2
-AB1
7.3 kb
6.2 kb

acquired a mutation impairing their pluripotent nature.

2.2.6 Improved *Thy-1* targeting vector

As only one targeted clone was obtained in the AB1 ES cell line and this was not able to transmit the targeted allele, a second targeting experiment was proposed because a number of factors were not considered in the design of the 5.4 kb targeting vector. The extent of homology shared between the targeting vector and locus was known to be important in targeting efficiency (Thomas and Capecchi, 1987). However, new studies appeared that showed a dramatic increase in targeting efficiency with increased length of homology (Hasty *et al.*, 1991a; Deng and Capecchi, 1992). For example, increasing the length of homology from 1.3 kb to 6.8 kb gave a 190-fold increase in targeting efficiency at the *Hprt* locus (Hasty *et al.*, 1991a). Another important factor which emerged was the increased targeting efficiency found when using targeting vectors made from DNA isogenic for the ES cells (te Riele *et al.*, 1992). A 20 fold increase in targeting efficiency was found when the retinoblastoma susceptibility gene was targeted using an isogenic targeting vector (te Riele *et al.*, 1992). This result was also described for other loci (van Deursen *et al.*, 1992; Deng and Capecchi, 1992). The *Thy-1* sequences used for construction of the 5.4 kb targeting vector were derived from the Balb/C mouse strain whereas the ES cells used were from the 129 Sv/Ev strain (Giguere *et al.*, 1985). Since polymorphisms were already known to exist between *Thy-1* genes from different mouse strains (Giguere *et al.*, 1985), it was proposed to remake the *Thy-1* targeting vector using isogenic DNA. The length of homology would also be increased to aid recombination efficiency.

It is important to note that upon completion of *Thy-1* isogenic targeting vector the existance of mice homozygous for a null Thy-1 allele was brought to our attention (Hoffman La Roche Report 1991). These mice were made available to Dr. R. Morris for study by Dr. C. Stewart of Hoffman La Roche. In view of this fact further attempts to target the *Thy-1* gene using the isogenic vector were discontinued. For the analysis of the Thy-1 mutation I have genotyped mice from Thy-$1^{+/-}$ x Thy-$1^{+/-}$ crosses for behavioural and electrophysiological analysis. The construction of the isogenic vector is described below.

The murine *Thy-1* gene is known to be located on an 8.2 kb EcoRI fragment in all mouse strains tested (Giguere *et al.*, 1985). This fragment also contains the promoter and *cis*-acting elements which direct expression of *Thy-1* in both thymus and brain (Vidal *et al.*, 1990). To obtain an isogenic clone of the *Thy-1* gene a size selected genomic DNA library was to be made from AB1 ES cell DNA. Cells were isolated from three confluent 60mm dishes by trypsinisation, (approximately $1x10^7$ cells). To avoid contamination from feeder cell DNA the AB1 ES cells were plated onto bacterial petri dishes. This allows the ES cells to be separated from the feeder cells due to a difference in adhesion properties. Feeder cells rapidly attach to the surface of the bacterial plates whereas the ES cells remain in suspension for longer periods. The suspension was taken and DNA was isolated from the ES cells, subsequently 100μg of ES cell DNA was digested with EcoRI and fragments were separated on a low melting point agarose gel. Ten micrograms of lambda DNA digested with BstEII was mixed in with the genomic DNA to provide an accurate size estimation. Sections of the gel around the 8.4 kb marker band were cut and DNA was isolated from each slice. A fraction was tested by Southern blot analysis for the presence of *Thy-1* sequence (Figure 15). Fraction 6 appeared enriched for the *Thy-1* sequence and was used to create the size selected library.

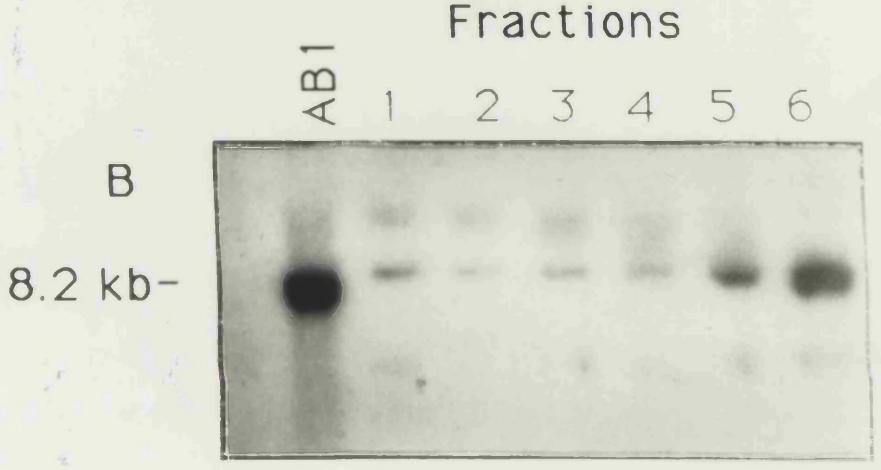
Lambda ZAP II (Stratagene) was chosen as a vector for the library as it can accommodate inserts of up to 14kb in size and inserts can be excised from the vector in the form of a pBluescript II SK- clone. This is achieved by using a helper phage which encodes a site-specific recombinase. This excision protocol avoids the subcloning step from the lambda vector. From fraction 6, 100ng was ligated into pre-digested EcoRI vector arms. The titre of the library obtained was 1.72×10^5 pfu/ml. Half of the library was plated and duplicate filters were screened for the presence of Thy-1 sequences using the 1.4 kb ApaI probe. Five duplicate positive plaques were picked, purified and the inserts excised. All five clones contained the correct 8.2 kb EcoRI insert of the *Thy-1* gene. Two clones designated Thy-129+ and Thy-129-, containing the 8.2 kb EcoRI insert in either orientation in the pBluescript multiple cloning site, were used for construction of the new targeting vector.

Figure 15. Screening for the presence *Thy-1* sequences in size selected fractions of AB1 genomic DNA.

A.) Ethidium bromide stained 0.8 % agarose gel showing the lambda BstEII digested molecular weight marker, EcoRI digested AB1 ES cell genomic DNA and size selected fractions from EcoRI digested AB1 ES cell genomic DNA, 1-6. The minor 8.4 kb and 5.6 kb bands visible in fractions 1-5, originate from the 14 kb band of the lambda BstEII marker and contain the lambda cohesive ends. The 14 kb band is separated into the 5.6 kb and 8.4 kb bands upon heating during the isolation of DNA from the low melting point agarose.

B.) Southern blot analysis of the gel show in panel A. hybridised with the 1.3 kb ApaI *Thy-1* probe. The 8.2 kb EcoRI fragment containing the *Thy-1* can be seen to hybridise in the genomic DNA sample. The size selected fractions sample 6 appears enriched for *Thy-1* sequence.


Lambda BstEII
AB1
Fractions
1 2 3 4 5 6
A


Fractions
AB1
1 2 3 4 5 6
B
8.2 kb-

2.2.7 Construction of the isogenic targeting vector

The construction of the isogenic targeting vector is shown in Figure 16. From the *Thy-1* isogenic clone, Thy-129-, a 5.2 kb HindIII-EcoRI fragment was subcloned into the polylinker of the vector pUC19. This made the BstEII site in exon III of the Thy-1 gene unique. Into the blunted SmaI-BstEII sites of this subclone a blunted 1.1 kb XhoI-BamHI fragment containing MC1NeopolyA was cloned. To this intermediate clone the remaining 5' homology was restored by including the 2.8 kb HindIII fragment from Thy-129-. Sequences 3' of the Asp718 site in exon IV of *Thy-1* were removed from the vector to allow this region to be used to generate external probes. To achieve this, a 7kb BglII-Asp718 fragment from the above subclone was placed into the equivalent sites in Thy-129+, replacing the normal *Thy-1* sequence with the MC1NeopolyA disruption and deleting sequences 3' of the Asp718 site. Finally the py*tk* vector was cloned into the unique blunted Asp718 site as a blunted 2.3 kb BamHI-Asp718 fragment.

The final vector contains MC1NeopolyA cassette placed between the SmaI site located in intron 2 and the BstEII site in exon III of the *Thy-1* gene, deleting a region of 565bp. As compared to the previous targeting vector, 2 kb of 3' homology was included. Again the py*tk* counter selection cassette is placed downstream of the 3' homology and the resulting vector was linearised at the unique NotI site located in the pBluescript polylinker, adjacent to the 5' region of homology.

2.3 Discussion

2.3.1 Summary of targeting experiment

In this chapter the targeting of the murine *Thy-1* gene by homologous recombination in ES cells is described. A replacement type targeting vector possessing 5.4 kb of homology to the target gene was used in combination with a positive/negative selection protocol (Mansour *et al.*, 1988) to obtain targeted clones. From this experiment one targeted ES cell clone was successfully isolated and the integrity of the targeted locus established.

Figure 16. Diagram showing construction of *Thy-1* isogenic targeting vector.

Cloning step 1. 5.2 kb HindIII-EcoRI fragment from pThy-129- is subcloned into the pUC19 vector.

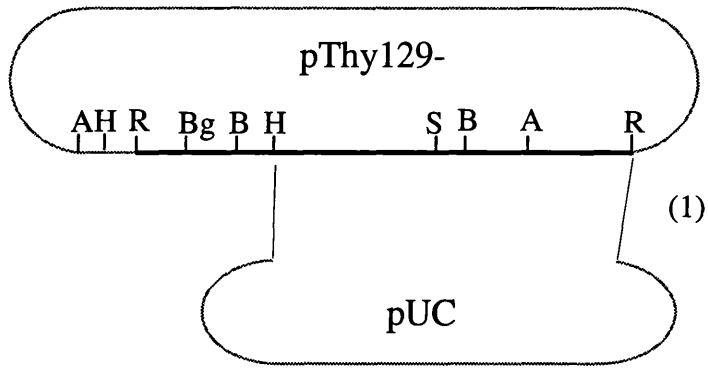
Cloning step 2. The unique BstEII site in exon III was cut and blunted. Into the blunt end SmaI-BstEII sites a 1.1 kb blunted XhoI-BamHI fragment from pMC1NeoployA (Thomas and Capecchi, 1987) was cloned.

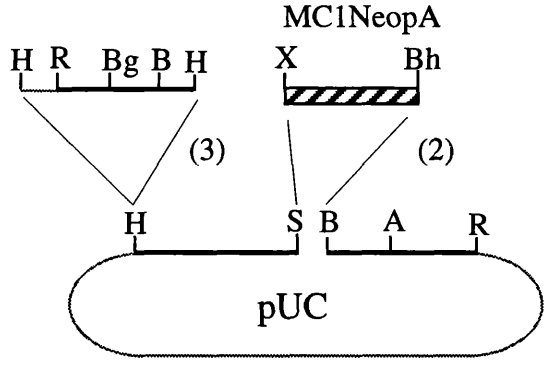
Cloning step 3. Further 5' *Thy-1* sequence was added as a 2.8 kb HindIII fragment.

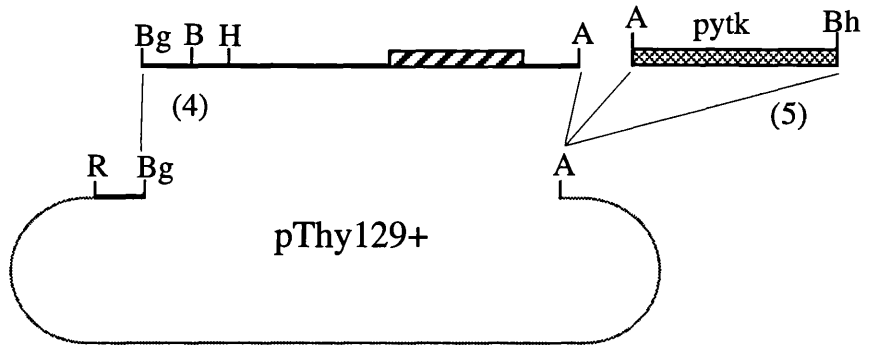
Cloning step 4. The BglII-Asp718 fragment from the above intermediate clone was cloned into the BglII-Asp718 sites of pThy-129+.

Cloning step 5. A blunted 2.3 kb BamHI-Asp718 fragment from py*tk* (Watson, 1992) was cloned into the blunt Asp718 site.

Restriction enzymes sites are as follows: A, Asp718; B, BstEII; Bg, BglII; Bh, BamHI; H, HindIII; R, EcoRI; S, SmaI; X, XhoI. Fragment lengths are not to scale.


pThy129-
AH R Bg B H
S B A R
(1)
pUC


MC1NeopA
H R Bg B H
X Bh
(3)
(2)
H
S B A R
pUC


Bg B H
A
A pytk Bh
(4)
(5)
R Bg
A
pThy129+

The transfection efficiency of the vector, as defined by the number of $G418^r$ colonies per number of cells electroporated, was in the order of 1×10^{-4} (Table I). This is 10 fold lower than the 1×10^{-3} transfection efficiency described previously for pMC1NeopolyA cassette alone (Thomas and Capecchi, 1987), however, both the ES cell type and electroporation conditions were different which could account for the differences. The application of the positive/negative protocol for clone selection yielded an approximate 16-fold enrichment of $G418^r/GANC^r$ colonies compared to $G418^r$ colonies (Table I). Although this is 100-fold lower than documented by Mansour *et al.* (1988) when targeting the *Hprt* and *Int-2* locus, it is concurrent with levels of enrichment described by other groups, the range of enrichment documented in the literature is in the order of 2 to 20 fold (reviewed by Hasty and Bradley, 1993). The degree of enrichment is dependent on a number of factors integral to the recombination efficiency of the vector, such as the length of homology and number of polymorphisms between vector and locus. The extent of enrichment will therefore vary from one gene to another.

The absolute targeting efficiency of the vector used can not be judged, as only one correctly targeted clone was obtained and only 53 of the 82 double resistant colonies that arose were finally analysed. The targeting vector contains the same degree of homology as the Thy-neo5'3' vector used in the CCE line (Watson 1992), but in this experiment a PCR/pool selection procedure (Zijlstra *et al.*, 1989) was used rather than positive negative selection. In the latter experiment five independent targeting events were detected from the 1×10^7 electroporated cells, but the transfection efficiency was not calculated and not all the pools of ES cells analysed. Overall, the experiment in the CCE line did yield more targeted clones, however it can not be judged if this is due to a more efficient transfection protocol or to the screening procedure used.

The clone handling procedure used in the AB1 targeting experiment was probably a contributing factor to the loss of the 29 clones from the 82 picked. Expanding ES cell colonies from a 24 multi well plate to a 60mm plate leads to extended time in culture which increases the likelihood of clone loss due to differentiation. New clone handling protocols employ improved freezing techniques which allow the clones to be frozen in the multi well plates. This reduces the passage number of the clones and eases the manipulation of large numbers of samples (Wurst

and Joyner, 1993). Isolation of DNA can also be carried out in multi well plates (Laird *et al.*, 1991), again increasing experimental efficiency.

The failure of the targeted ES cells to produce germline chimeras may have been due to a number of factors. First, if the ES cells entered some differentiation pathway, this in itself may have been detrimental to the development of the injected blastocyst. This appears to be the case from the lack of chimeras obtained. However, the AB1 ES cell line used in this experiment was later shown to be capable of germline transmission in our hands, although chromosomal counts revealed that 20% of the original culture was aneuploid (Anne Bygrave, personal communication). It is possible that the *Thy-1* targeted ES cells isolated and injected were from this population of aneuploid cells. It is also feasible that chromosome abnormalities were acquired during the targeting experiment. As no chromosomal analyses were performed on the targeted ES cells, these factors can not be discounted. This highlights the need to karyotype the ES cell line and targeted clones prior to injection.

As germline transmission was not obtained in the above experiment a second targeting experiment was embarked upon. Though the original 5.4 kb targeting vector was successful in establishing ES cells targeted at the *Thy-1* allele, improvements to the vector design were made with the aim of improving recombination efficiency. The new targeting vector contained 2.0 kb of 3' homology as compared to the 800 bp in the 5.4 kb targeting vector and was constructed from a 129/AB1 library. As discussed in detail earlier, both increasing homology and use of isogenic DNA are known to improve relative targeting efficiencies. This new vector should therefore have targeted with improved efficiency. However, the isogenic targeting vector constructed was not tested since mice carrying a targeted *Thy-1* allele became available from another lab. No comparison between the targeting efficiency of the new and old vectors were therefore made.

2.3.2 Fate of *Thy-1*$^{-/-}$ mice.

It has been suggested that Thy-1 acts as a signal transduction and/or adhesion molecule in the mouse nervous and lymphoid systems (Morris, 1992; He *et al.*, 1991). In the rodent nervous system Thy-1 is thought to play an inhibitory role in the process

of axonal outgrowth, a function which is believed to be conserved between species (Morris, 1992). Evidence from studying the nervous system of *Thy-1*$^{-/-}$ mice suggest that this may be the case.

The existence of mice homozygous for a targeted *Thy-1* allele was first described in the Hoffman La Roche Research Report 1991. The animals were described as "overtly normal adult mice". No aberrations in the nervous or lymphoid system were described for these mice in this report, though the degree to which they were examined was not alluded to. Subsequent analysis has revealed a number of defects in these mice.

It has been previously shown that the presence of Thy-1 on the surface of a neuronal cell line can cause an inhibitory effect on neurite outgrowth on astrocytes (Tiveron *et al.*, 1992). In this *in-vitro* assay for Thy-1 function, the rate of axonal outgrowth of sensory neurons from *Thy-1*$^{-/-}$ mice shows a marked difference to that of neurons from normal mice. The axons of *Thy-1*$^{-/-}$ neurons grow 50% faster over astrocytes than those of wild type animals. It is clear, in this cell culture assay, that Thy-1 is involved in inhibition of neurite outgrowth.

Electrophysiological analysis of LTP in the hippocampus has also revealed differences between *Thy-1*$^{-/-}$and wild type mice. LTP has been described for three linked synaptic systems within the hippocampus. The first is the projection of entorhinal axons onto the granule cells of the dentate gyrus, the perforant path. The granule neurons in turn send out processes known as mossy fibres which synapse onto the dendrites of pyramidal neurons of the CA3 area of the hippocampus, the second synaptic connection. Thirdly, axons of the CA3 region, known as Schaffer collateral axons, project onto the dendrites of the CA1 pyramidal neurons (Figure 17). Inhibition of LTP using N-methyl-D-aspartate receptor antagonists has shown LTP in the first and third synaptic relays to be dependent on this receptor (Morris *et al.*, 1986). On examination, LTP in the dentate gyrus of *Thy-1*$^{-/-}$ mice appears severely suppressed. In the other NMDA receptor dependent system, projection of the Schaffer collaterals onto the CA1 neurons, the mice appear normal. How the LTP in the *Thy-1*$^{-/-}$ mice is suppressed in the dentate gyrus is not known. As Thy-1 is suggested to play a role in the inhibition of axonal outgrowth one could hypothesise that abnormal numbers of afferent projections into the dentate gyrus may cause this result. Why then only this one area is affected is difficult to speculate upon. Certainly a closer inspection of the dentate gyrus should prove

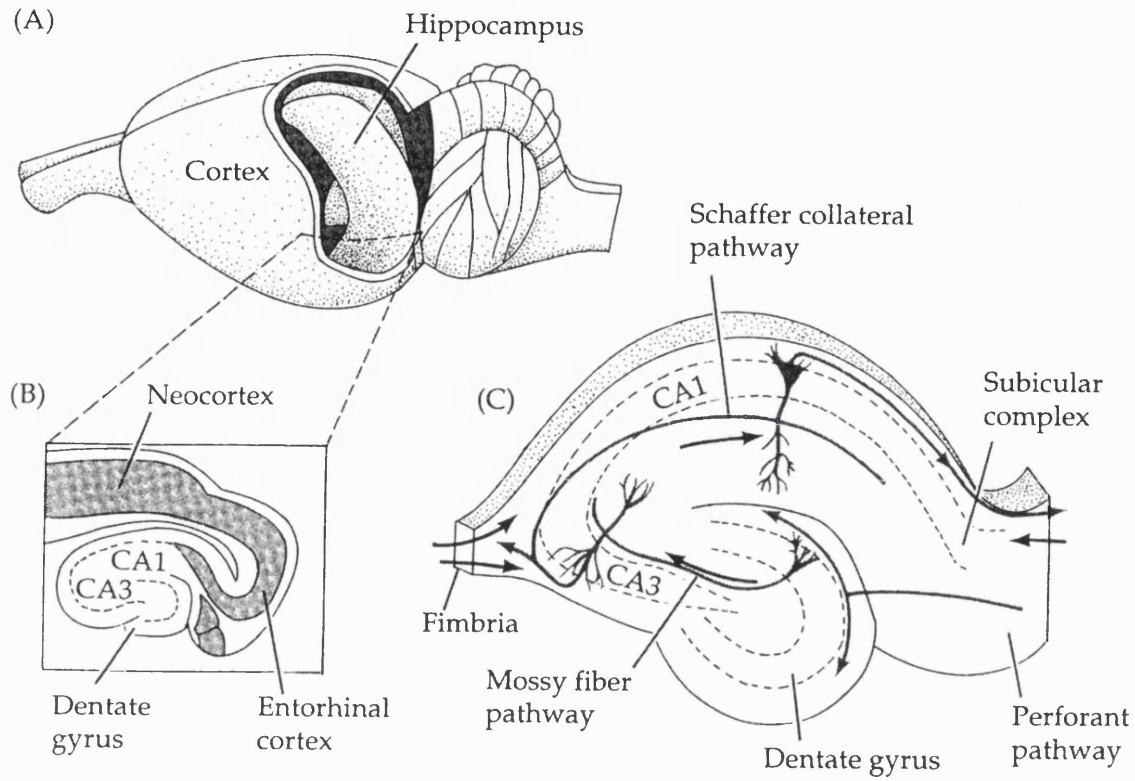
Figure 17. Organisation of the rodent hippocampus and synaptic pathways involved in hippocampal LTP

A.) Diagrammatic representation of the rodent brain showing the position of the hippocampus.

B.) Horizontal cross section of the hippocampus and surrounding neocortex. Dashed lines represent the dense neuronal cell layers of the dentate gyrus, and CA1 and CA3 (cornu Ammonis) pyramidal neuron fields.

C.) Detailed view of synaptic circuit in hippocampal section. Dashed lines depict the tightly packed layers of neuronal cell bodies. Areas outside the dashed lines contain axons, dendrites and synaptic contacts. The positions of the perforant, mossy fibre and Schaffer collateral axon pathways are indicated.

Reproduced from Kennedy and Marder, 1992.


(A)
Hippocampus
Cortex
(B)
Neocortex
CA1
CA3
Dentate gyrus
Entorhinal cortex
(C)
Schaffer collateral pathway
CA1
Subicular complex
CA3
Fimbria
Mossy fiber pathway
Dentate gyrus
Perforant pathway

interesting.

Despite this abnormal electrophysiology, the *Thy-1*$^{-/-}$ mice appear to function normally in the Morris watermaze test (Dr. R. Morris personal communication). The Morris watermaze test is a test designed to determine spatial learning capability in mice (Morris *et al.*, 1981). The normal learning behaviour of the *Thy-1*$^{-/-}$ mice in this test is an interesting finding, since LTP is widely thought to correlate with spatial learning (Morris *et al.*, 1982; Sutherland *et al.*, 1982; reviewed by Brandeis *et al.*, 1989). It should however be mentioned that discrepancies between LTP and hippocampal learning tasks have been previously documented (Collinge *et al.*, 1994; Brace *et al.*, 1985).

Interestingly, a similar phenotype to the *Thy-1*$^{-/-}$ mice is seen in the targeted mutation of the murine PrP^{c} protein (Bueler *et al.*, 1992; Bueler *et al.*, 1993; Collinge *et al.*, 1994). PrP^{c} is the cellular or normal form of the prion protein. The prion protein is regarded to be the causative agent in transmissible spongiform encephalopathies such as scrapie in sheep and human diseases such as Creutzfeld Jakob disease and kuru (reviewed by Pruisner 1994). Like Thy-1, PrP^{c} is a cell surface glycoprotein which is anchored to the cellular membrane via a GPI tail (Stahl *et al.*, 1987) and has also been found associated with caveolae (Ying *et al.*, 1992). Unlike Thy-1, PrP^{c} is not a member of the Ig super gene family (Stahl *et al.*, 1993). The developmental profile of *PrP^{c}* gene expression is different from Thy-1, though the main site of expression is in the nervous system (Oesch *et al.*, 1985; Kretzschmar *et al.*, 1986; Manson *et al.*, 1992). Unlike Thy-1 *PrP^{c}* mRNA is also detected in glial cells (Moser *et al.*, 1995). Mice homozygous for a disrupted *PrP^{c}* allele were first described as being normal both in development and behaviour (Bueler *et al.*, 1992). Subsequent electrophysiological analysis of LTP in the CA1 region, the other synaptic system of the hippocampus in which LTP is NMDA receptor dependent, has shown that these mice have a reduced LTP response (Collinge *et al.*, 1994). LTP in the dentate gyrus is not described in this study.

Therefore, like the *Thy-1* disruption, *PrP$^{c-/-}$* mice show normal spatial learning behaviour but display abnormal NMDA receptor dependent LTP in the hippocampus, albeit in a different synaptic system. The similar phenotypes of the *Thy-1* and *PrP^{c}* disrupted mice and structural similarities between the two proteins, such as GPI anchor and glycosylation, suggest that PrP^{c} and Thy-1 may have similar and perhaps overlapping functions. The conserved glycosylation pattern of Thy-1 in three different

species has lead to the suggestion that the role of Thy-1 is to present specific surface carbohydrates (Williams *et al.*, 1993). It would be interesting to determine whether these N-glycans are shared between Thy-1 and PrP^c. PrP^c is known to possess complex-type N-linked oligosaccharides, though their precise structure is not known (Haraguchi *et al.*, 1989). As it is now feasible to produce *Thy-1*$^{-/-}$ *PrP*$^{c-/-}$ mice by a simple cross, it will be interesting to see whether combining the two mutations will lead to a more severe phenotype. If PrP^c and Thy-1 do overlap in function, but not in expression and/or glycosylation pattern, then this could account for the restricted phenotypes seen in the *Thy-1*$^{-/-}$ and *PrP*$^{c-/-}$ mice.

Although less analogous to the Thy-1 disruption, inactivation of the murine *NCAM* gene, which encodes another Ig like molecule, leads to a milder phenotype than the distribution of NCAM would predict (discussed in section 1.2.12 and 1.3.7). A phenotype is only seen in systems where a particular glycoform of NCAM is predominant. NCAM is one of the major carriers of polysialic acid moiety (Rutishauser, 1992) and enzymatic removal of the polysialic acid results in the phenotype of the NCAM disruption being mimicked (Ono *et al.*, 1994). Different glycoforms of Thy-1 are known to exist in the brain (Parekh *et al.*, 1987; Williams *et al.*, 1993) although the distribution of these glycoforms is not known. It is possible, that like the NCAM disruption, what appears to be highly restricted phenotype in the *Thy-1*$^{-/-}$ mice is due to the loss of a particular N-glycan in an area where Thy-1 is the major carrier.

The above theories have yet to be tested, however establishment of *Thy-1*$^{-/-}$ mice now provides a "null" genetic background in which different functional aspects of the Thy-1 molecule can be tested *in-vivo* by the introduction of *Thy-1* transgenes. Since the *Thy-1*$^{-/-}$ mice show no abnormality in their ability to breed, any transgene can be directly introduced into *Thy-1*$^{-/-}$ oocytes. As discussed above, the glycosylation of Thy-1 may be important to its function. This theory could be tested by introducing transgenes into the Thy-1$^{-/-}$ mice which carry mutations at the asparagine glycosylation sites. These transgenic mice could then be tested to see if the LTP deficit in the dentate gyrus could be rescued or not. Furthermore, *in-vitro* experiments have shown the membrane anchor of Thy-1 to be important for its inhibition of neurite outgrowth upon astrocytes (Tiveron *et al.*, 1994). These membrane anchor mutants could also be tested *in-vivo*. These type of experiments are in progress at the moment and should provide valuable information

as to how Thy-1 functions and what aspects of the molecule are important.

Chapter 3

Targeting of the murine *Zfp-37* gene in mouse embryonic stem cells

3.1 Introduction and strategy

Research on the murine *Zfp-37* gene was initiated in this laboratory through an interest in identifying transcription factors that are involved in nervous system development. Zinc finger transcription factors have been identified to play important roles in cell type and regional specification in the nervous system of both *Drosophila* and mouse. *Zfp-37* appeared to be an interesting candidate as it could encode 12 *Krüppel*-like Cys_2His_2 zinc finger motifs indicative of potential DNA binding and hence a role as a transcription factor (Nelki *et al.*, 1990). Northern blot analysis of RNA expression and *in-situ* RNA hybridisation data showed *Zfp-37* gene to be expressed in testis and in the developing and adult mouse brain from at least e10.5 (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994; Mazarakis *et al.*, 1995). Taken together, these data suggested that ZFP-37 protein could play a role as a transcriptional regulator in the testis and the developing and adult brain. In this context further investigation of ZFP-37 was initiated. To complement biochemical analysis of ZFP-37 function, disruption of the *Zfp-37* gene in the mouse via homologous recombination in ES cells was undertaken. This approach should help address the functional role of ZFP-37 in the mouse.

As little information was available regarding the expression pattern of *Zfp-37* during development, a bacterial reporter gene, *lacZ*, was included in the disruption. Placing the β-galactosidase gene, *lacZ*, in-frame with the ZFP37 open reading frame and recombining it into the genomic context of the *Zfp-37* gene was carried out for the following reasons. Firstly, expression of the *Zfp-37* gene could be easily identified by staining tissue samples with the β-galactosidase chromogenic substrate X-gal. As the reporter gene is placed within the chromosomal context it was expected to mimic the normal expression pattern of the *Zfp-37* gene and not be subject to position effects that can affect randomly integrated transgenic constructs (Krumlauf *et al.*, 1986; reviewed by Wilson *et al.*, 1990). Secondly, in mice homozygous for the disrupted allele, the fate of cells that would normally express *Zfp-37* but now express *lacZ* could be visualised. This strategy has successfully been applied to the targeting of the *int-2* (Mansour *et al.*, 1990), *Hoxc8* (*Hox-3.1*) (Le Mouellic *et al.*, 1990 and 1992) and *Krox20* (Schneider-Maunoury *et al.*, 1993) genes. Although inclusion of the *lacZ* gene in the disruption

increases the amount of non-homology used as the disrupting sequence, this has been shown not to affect targeting efficiency (Mansour *et al.*, 1990; Le Mouellic *et al.*, 1990). The strategy outlined above should therefore help elucidate the role of ZFP-37 and should provide further information as to the expression pattern of the *Zfp-37* gene.

The D3 ES cell line (Doetschman *et al.*, 1985) had been made available to us and this line would be tested for its ability to contribute to the germline by injection of unmodified cells into blastocysts. Resulting chimeric mice will be bred and assessed for germline transmission by the coat colour of the offspring. The AB1 ES cell line has also been shown to be capable of germline transmission in our laboratory (Pandolfi *et al.*, 1995). Depending on the results from the wild type D3 injections these may also be used.

Once the ES cell line has been chosen, targeting of the *Zfp-37* gene will be attempted. Targeted clones will be characterised by Southern blot analysis. Positive clones will be karyotyped and clones showing a normal chromosome number will be injected into blastocyst. The resulting chimeric mice will be bred to establish mice heterozygous for the targeted allele. These mice will be intercrossed to establish the fate of animals homozygous for the targeted *Zfp-37* allele.

3.2 Results

3.2.1 Design of the *Zfp-37* targeting vector

The design of the *Zfp-37* targeting vector was required to both disrupt ZFP-37 function but also place a reporter gene in frame with the ZFP-37 open reading frame. When this project was initiated, the initiation codon of ZFP37 had been identified as being located within exon II, one base from the exon/intron junction (Burke and Wolgemuth, 1992). An in frame initiation codon was subsequently identified in exon Ia, although it is estimated that 90% of ZFP-37 proteins initiate at this codon, placing the *lacZ* gene in frame here would not have targeted all isoforms of the protein. The ATG in exon II is contained in all ZFP-37 isoforms and was therefore chosen as the site to insert the disrupting sequence (Mazarakis *et al.*, 1995). Prior and during construction of the targeting vector I have been involved in mapping the *Zfp-37* cosmid clone, the

results of which are discussed in section 1.4.5 of Chapter 1 and Mazarakis *et al.* (1995).

Two types of *lacZneo* cassettes could be used for this targeting. Either a *E.coli lacZ* gene followed by neomycin resistance cassette under control of its own promoter enhancer sequences could be engineered, as was used by Mansour *et al.* (1990), or a *lacZneo* fusion cassette could be employed (Fredrich and Soriano, 1991; Mountford *et al.*, 1994). In the latter design *lacZ* and *neo* are expressed as a fusion protein, however for this cassette to work for selection purposes the gene to be targeted must be expressed in the ES cells. This type of construct functions as a promoterless selection cassette and would only allow survival of a targeted clone if the fusion protein had recombined into the correct chromosomal position, conditional selection. At the time *Zfp-37* was not known to be expressed in ES cells and therefore a conditional selection strategy may not have worked. Secondly a putative promoter sequence had been identified upstream of exon II. Inclusion of this promoter region in a conditional targeting vector may allow the selectable marker to be expressed when the construct integrates into a random position. Excluding this putative promoter region would limit the extent of 5' homology that could be employed to less than 200bp if the *lacZneo* cassette is to be placed at the ATG in exon II. Limiting the 5' homology to 200bp is predicted to severely affect the recombination efficiency of the vector (Berinstein *et al.*, 1992; Thomas *et al.*, 1992). For the above reasons the first vector design was chosen.

As neurons are one of the cell types predicted to expressing *Zfp-37* it was desirable to confine the β-galactosidase reporter protein to the nucleus. This should give better cellular resolution, avoiding axonal staining which could confuse interpretation of the expression pattern. The construct L7RH β-gal (Kalderon *et al.*, 1984) was available in the laboratory, this construct consists of the bacterial *lacZ* gene encoding β-galactosidase with an N-terminal 20 amino acid nuclear localisation signal from the SV40 Large T antigen (Kalderon *et al.*, 1984).

The construction strategy was as follows: Firstly a *lacZneo* cassette would be engineered based upon the L7RH β-gal construct (Kalderon *et al.*, 1984) and pMC1NeopolyA (Thomas and Capecchi, 1987). The target ATG located at the 3' end of exon II is within a NcoI restriction endonuclease site. This cassette would be compatible with the NcoI site such that the *lacZ* gene would be place in-frame with an ATG initiation codon within that sequence. Further homologous sequences would be

added to the vector and the py*tk* cassette (Watson, 1992) included for negative selection.

3.2.2 Construction of the *Zfp-37* targeting vector

The nlslacZneo cassette was constructed using the pMC1NeopolyA plasmid as the host vector. A 3.4 kb BamHI fragment containing L7RH β-gal was blunted and ligated into the unique blunted XhoI site upstream of the polyoma enhancer of pMC1NeopolyA. To make this vector compatible with a NcoI restriction site, two linker sequences were designed which included the NcoI compatible BspHI site. One was inserted into the BspEII site within the *lacZ* coding region such that a new ATG sequence is introduced, the second was inserted at the unique BamHI site 3' of pMC1NeopolyA. Sequence and cloning of the BspHI linkers is shown in Figure 18. The 5' linker introduces a base change of G→A as compared to the normal L7RH β-gal sequence which results in an amino acid change of aspartate(D) to asparagine(N) in the resulting protein. Since this residue is not part of the nuclear localisation signal this change should not be detrimental to this function.

A 3.5 kb BamHI-HindIII subclone of *Zfp-37* was isolated from the *Zfp-37* cosmid and cloned into pBluescript KS+. This fragment contains exons II and III, the NcoI site at the end of exon 2 is unique. Into this NcoI site the 4.5 kb BspHI nlslacZneo fragment was cloned. The orientation of the cloning was confirmed by restriction mapping and the *Zfp-37-lacZ* junction sequenced to verify that the *lacZ* gene was in frame with the *Zfp-37* coding sequence.

To this intermediate construct the py*tk* cassette was cloned into the unique Asp718 site in the pBluescript polylinker at the end of the 3' homology as a blunt end ligation. To this vector a further 3.6 kb of *Zfp-37* sequence was cloned 5' of the disruption as a XbaI fragment into the partially digested XbaI backbone of the intermediate vector. The resulting targeting vector contains a total of 7.1 kb of homology to the *Zfp-37* locus, 4.1 kb 5' and 3kb 3' of the disrupting sequence, the total size of the vector was 16.9 kb. The 7.1 kb zinc finger targeting vector (7ZTV) could be linearised at the unique NotI site located at the end of the 5' stretch of homology. Figure 19a shows the design of the completed targeting vector.

Figure 18. Construction of the NcoI compatible nlslacZneo cassette

a) Diagram represents the vector nlslacZ (Kalderon *et al.*, 1986) cloned into pMC1NeopolyA (Thomas and Capecchi, 1987). A 3.4 kb BamHI fragment containing nlslacZ was blunted and ligated into the unique XhoI site of pMC1NeopolyA. This cloning step recreates XhoI sites. Dashed lines represent the vector backbone.

b) Cloning of BspHI linker sequences. The sequence surrounding the ATG of nlslacZ is shown in detail as is the location of the BspEI restriction enzyme site. Into BspEI the 5' linker shown below was cloned. This linker contains a BspHI site located over the ATG sequence, this sequence modification results in an amino acid change from aspartate (D) to asparaginine (N), highlighted in bold. At the BamHI site located at the 3' end of the MC1NeopolyA cassette a 3' linker was cloned, this destroys the BamHI site but places a BspHI restriction site at the 3' end of MC1NeopolyA.

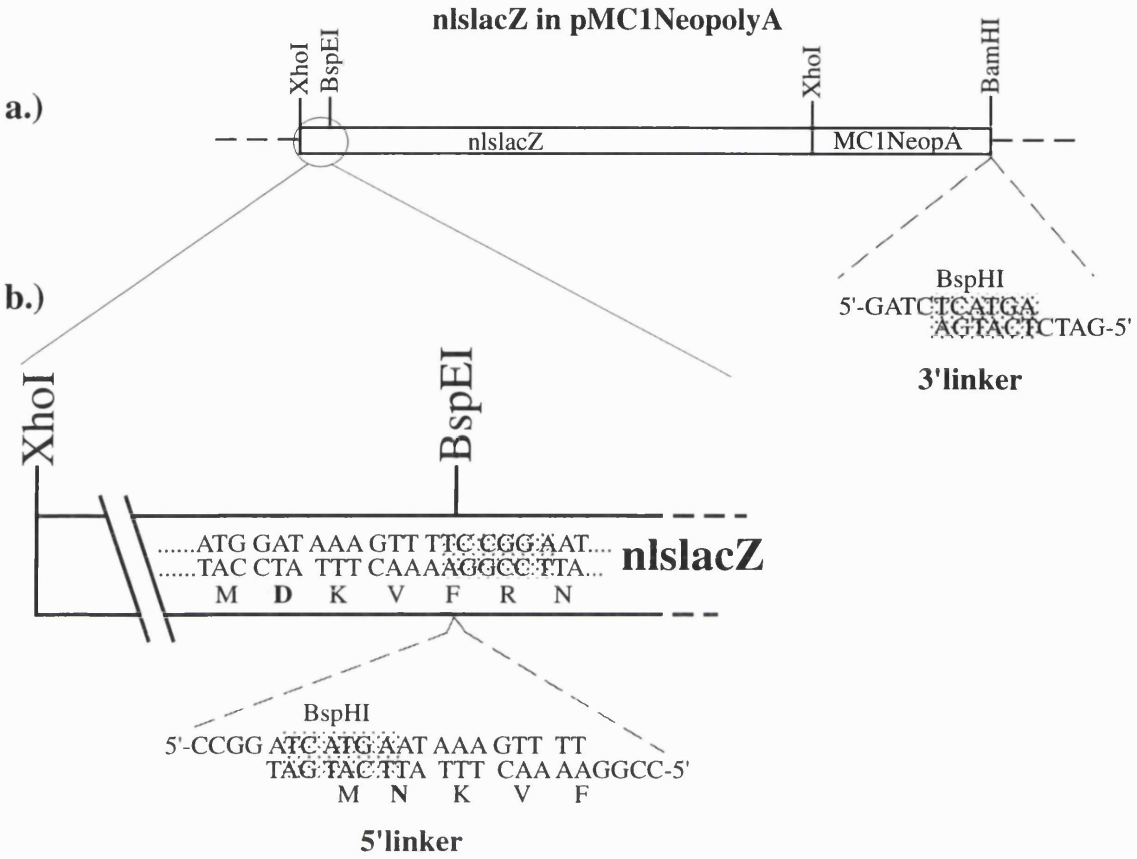

nlslacZ in pMC1NeopolyA
a.)
XhoI
BspEI
XhoI
BamHI
nlslacZ
MC1NeopA
BspHI
5'-GATCTCATGA
AGTACTCTAG-5'
3'linker
b.)
XhoI
BspEI
.....ATG GAT AAA GTT TTC CGG AAT....
.....TAC CTA TTT CAAAAGGCC TTA...
nlslacZ
M D K V F R N
BspHI
5'-CCGG ATC ATG AAT AAA GTT TT
TAG TAC TTA TTT CAA AAGGCC-5'
M N K V F
5'linker

Figure 19. *Zfp-37* targeting vector and predicted structure of the targeted *Zfp-37* allele

a) Organisation of the *Zfp-37* 7ZTV targeting vector. The 4.6 kb NcoI compatible nlslacZneo cassette is inserted into the NcoI restriction enzyme site located at the ATG in *Zfp-37* exon II (shown in b.). The py*tk* cassette is inserted in the polylinker sequence of pBluescript. Further 5' *Zfp-37* sequences is added as a 3.6 kb XbaI fragment. The resulting vector is linearised at the NotI site and possess 4.1 kb of homology to *Zfp-37* 5' of the nlslacZneo disruption and 3.0 kb 3'. Cloning vector sequences are depicted as a dotted line.

b) Organisation of the murine *Zfp-37* allele. Location of the *Zfp-37* exons Ia, Ib, II, III and IV are shown as boxes, coding region of exons are shown as filled boxes. Intervening sequences are shown as a single line. A partial restriction map of the *Zfp-37* gene is shown.

c) Predicted organisation of the targeted *Zfp-37* allele. Restriction enzymes used in Southern blot analysis of the targeted allele are shown, these are NcoI, PstI, SphI and StuI.

d) Location of probes used to screen for homologous recombination at the *Zfp-37* allele. 0.9 kb SphI-XbaI Zfp37 5' probe was isolated from an XbaI subclone (data not shown) of the *Zfp-37* cosmid (Mazarkis *et al.*, 1995). The Zfp37 cDNA probe is a 0.4 kb EcoRI -HindIII fragment from the 3.4 kb *Zfp-37* cDNA (Mazarakis *et al.*, 1995). The Neo probe is a 0.6kb PstI fragment from the vector pMC1NeopolyA (Thomas and Capecchi, 1987).

a.)
-NotI
-XbaI
StuI
-NcoI
-XbaI
-PstI
ATG
nlsLacZ
-StuI
-PstI
Neo
-NcoI
SphI
SphI
-XbaI
-HindIII
pγTk
pBluescript
-NotI
b.)
SphI
-XbaI
-StuI
a b
-NcoI
-XbaI
-PstI
-NcoI
II III
SphI
-XbaI
-HindIII
-PstI
-NcoI
-NcoI
-StuI
-HindIII
-HindIII
-XbaI
IV
c.)
SphI
-XbaI
-StuI
-NcoI
-XbaI
-PstI
nlsLacZ
-StuI
-PstI
Neo
-NcoI
SphI
SphI
-XbaI
-HindIII
-PstI
-NcoI
-NcoI
-StuI
-HindIII
-HindIII
IV
d.)
Zfp37-5'
SphI
-XbaI
Neo
-PstI
-PstI
Zfp37 cDNA

3.2.3 Test injection of the D3 ES cell line

The D3 ES cell line (Doetschman *et al.*, 1985) had been made available to us by Dr. R. Kemler. This ES cell line has been successfully used in the targeting of a number of genes including the β2-microglobulin gene (Zijlstra *et al.*, 1989), *En-2* gene (Joyner *et al.*, 1989) and the retinoblastoma gene (Jacks *et al.*, 1992). Blastocyst injection of the parent D3 line served two purposes. The line could be tested for its ability to contribute to the germline tissue of the animal prior to its use in gene targeting experiments and secondly, its testing allowed blastocyst injection expertise to be developed. The D3 line is normally cultured on primary embryonic fibroblasts (Doetschman *et al.*, 1985). However, the passage available to us had been grown on SNL feeder cells, a culture method used for the AB1 ES cell line (McMahon and Bradley, 1990). Injection of the D3 ES line would also test whether this change in culture conditions affected the pluripotency.

Two sets of injections were performed. Ten to fifteen ES cells were injected into the blastocoel cavity of 3.5 day mouse blastocysts. The result of the injections are summarised in Table II. A total of 5 male chimeras were produced from injection of 52 blastocysts, which had given rise to 13 live births. The contribution of the ES cells was approximately 50% as deemed by coat colour chimerism (Figure 20). When six weeks old, the male chimeras were crossed with C57BL/6 females and the resulting offspring were scored for transmission by presence of agouti coat colour. Two of the chimeras from the first set of injections transmitted the agouti coat colour allele from the ES cells. In one case transmission appeared to be 100% as all pups in the litter were agouti. From this experiment the unmodified D3 line grown under the new culture conditions was shown to be capable of contributing to tissue in a chimeric animal and efficiently contributing to the germline.

Based upon these results the D3 line was used for targeting experiments in the laboratory. However, upon transfection and selection of these cells using variable culturing conditions no viable clones were obtained with a number of different targeting vectors (E. Cook and A. Karis personal communication). The ES cells appeared not to tolerate transfection conditions. Upon these observations it was decided to continue with the AB1 ES cell line as this had been shown to give germline transmission of a targeted

Table II-Blastocyst injections with the D3 ES cell line

Experiment	Blastocysts injected	Births	Chimeras	Male	Female
1	26	9	3	3	-
2	26	4	2	2	-
Total	52	13	5	5	-

Figure 20. Chimeric mouse from blastocyst injections of the D3 ES cell line.

Picture shows a male chimeric mouse from blastocyst injection of the D3 ES cell line into a C57BL/6 host blastocyst. The contribution of the ES cell line in the case is approximately 50% as shown by the agouti coat colour.

allele in our hands (Pandolfi *et al.*, 1995).

3.2.4 Isolation of targeted *Zfp-37* ES cell clones

The transfection conditions used to electroporate the AB1 ES cells with the *Zfp-37* targeting vector were as those described for the targeting of Thy-1. As the targeting vector was larger as compared to the *Thy-1* targeting vector, the DNA concentration was increased to 50μg per shock. For each transfection 8 x 10^6 cells were used. As for the Thy-1 targeting, control plates to assess positive selection were also plated. The results from the electroporation are summarised in Table III. A total of 39 $G418^r/GANC^r$ were obtained. The enrichment using negative selection was 1/35 to 1/38 $G418^r/GANC^r$ clones per $G418^r$ clones.

Clones were picked into 24 well plates. Each clone was allowed to grow until confluent and at this point the contents of the well was harvested and the culture divided in two. Half of the sample was allowed to grow to confluence for isolation of DNA. The other half was grown to confluence trypsinised from the well and frozen.

Clone DNA was screened for recombination events by Southern blot analysis. As no probe external to the targeting vector homology had been isolated at this time, clones were initially assessed for a specific *neo* integration event which is predicted in a correct recombination. StuI restriction endonuclease sites, which cut both within the MC1NeoployA sequence and outside of the region of homology between vector and locus, were used. Southern blots were probed with a *Neo* specific 685 bp probe. In the case of homologous recombination a 6.2 kb fragment containing the neomycin resistance gene should be detected, the predicted structure of the complete targeted locus is shown in Figure 19c. However, this strategy will also detect random insertions of the targeting vector, which could give a false positive result if the 3' end of the vector inserts 2 kb from a StuI site in the genome. Therefore this is not a definitive screen but identifies candidate clones to be further analysed. Based upon this screening strategy three potentially targeted clones were originally identified from the 39 screened (Figure 21).

From the three clones identified, only one, ZT11, survived the freeze-thaw. Further DNA was isolated from this sample. To confirm this clone had undergone the

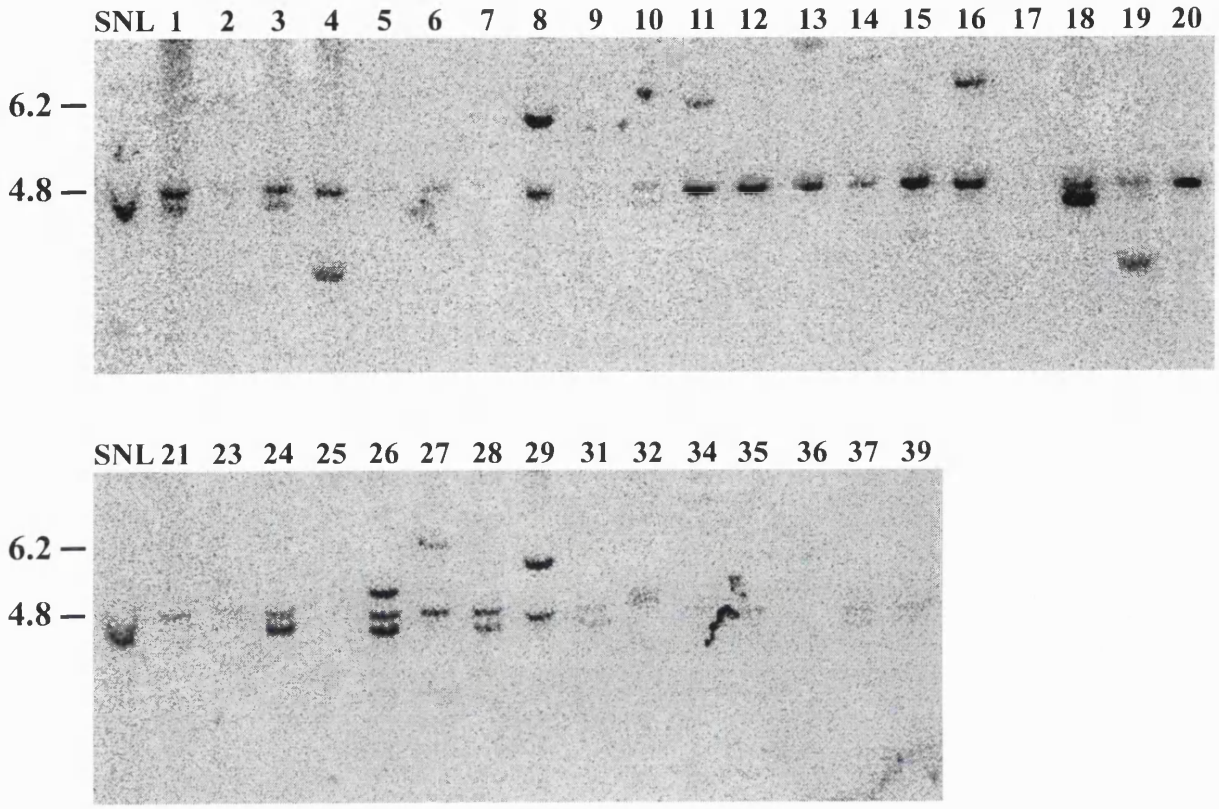
Table III- *Zfp-37* targeting experiment in AB1 ES cells

Experiment	No. of cells transfected	a) No. of G418^r colonies	b) No. of G418^r/GANC^r colonies	b/a
1	8x 10^6	680	19	1/35
2	8x 10^6	760	20	1/38

Figure 21. Southern blot analysis of G418^r/GANC^r AB1 ES cell clones from *Zfp-37* targeting experiment.

Genomic DNA from the 39 G418^r/GANC^r clones obtained digested with StuI restriction enzyme, Southern blotted and probed with the Neo probe shown in Figure 19d. A 4.8 kb hybridising fragment is seen in most lanes this originates from the *neo* construct integrated in the SNL feeder cell genome. In sample 2, 11 and 27 a 6.2 kb fragment hybridises with the Neo probe suggesting these clones to have undergone homologous recombination at the *Zfp-37* allele. In other samples random integrations of the targeting can be detected. Clones for which no or little DNA was available were regrown and checked for recombination, no other targeted clones were detected with this assay.

The hybridised Southern blot was visualised on a Phosphorimager.


SNL 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
6.2
4.8
SNL 21 23 24 25 26 27 28 29 31 32 34 35 36 37 39
6.2
4.8

correct recombination at the *Zfp-37* locus a further three restriction sites were chosen which cut both within the *Neo* gene and outside of the region of homology shared between the *Zfp-37* gene and targeting vector. The NcoI, SphI and PstI digest were used in conjunction with the *neo* specific probe. The location of these sites are shown in Figure 19. With a NcoI digest, a correct recombination of the targeting vector with the *Zfp-37* locus results in two fragments, since one NcoI site is within the probe region. The predicted fragments seen are an external 3' fragment of 3.7 kb and internal fragment of 5kb fragment. For a single integration of the vector into the locus, the 5 kb internal band should be of the same intensity as the 3.7 kb external band, this appears to be the case. The SphI digest cuts 5' of the targeting vector and within *neo*, showing the predicted band of 9 kb and an internal fragment of 1 kb. A PstI digest gives the predicted fragment size of 3.9 kb, a PstI site is located 200 bp from the end of the 3' homologous arm and within the *neo* sequence. The Southern analysis using these three digests is shown in Figure 22b. Based on the result from this analysis the clone ZT 11 was confirmed as possessing a targeted *Zfp-37* allele.

Further confirmation of the structure of the targeted allele was provided by hybridising a 0.4 kb cDNA probe to the NcoI digested clone DNA (Figure 22c). A 5' external SphI-XbaI probe was subsequently isolated and shown on a SphI digest to hybridise to the predicted 9 kb SphI fragment of the disrupted allele and 5.5 kb fragment of the unmodified allele of *Zfp-37* (Figure 22d).

As only one targeted clone was used in the production of *Zfp-37*$^{+/-}$ mice a second targeting experiment was carried out in an independent ES cell line. A subculture of E14 ES cell line derived by Handyside *et al.* (1989) was available to the laboratory via Dr. J. van Deursen. This subclone had been passaged on SNL76/7 feeder layers. The E14 ES cell line was cultured in larger numbers than the AB1 ES cells. Thereby reducing the length of time the ES cells spend in culture prior to transfection (see Materials and Methods). One confluent T75 tissue culture flask provides approximately 3x 10^7 cells for transfection purposes. The 1 x 10^7 cells were electroporated with the linearised 7.1 ZTV targeting vector. Drug selection (G418 350μg/ml, FIAU 0.2μM) was applied 24 hours post electroporation according to McMahon and Bradley (1990). The pyrimidine analogue FIAU was used instead of GANC because it was reported to be less cytotoxic to normal cells and more potent in killing expressing the HSV1 *tk* gene

Figure 22. Further southern blot analysis of clone ZT11

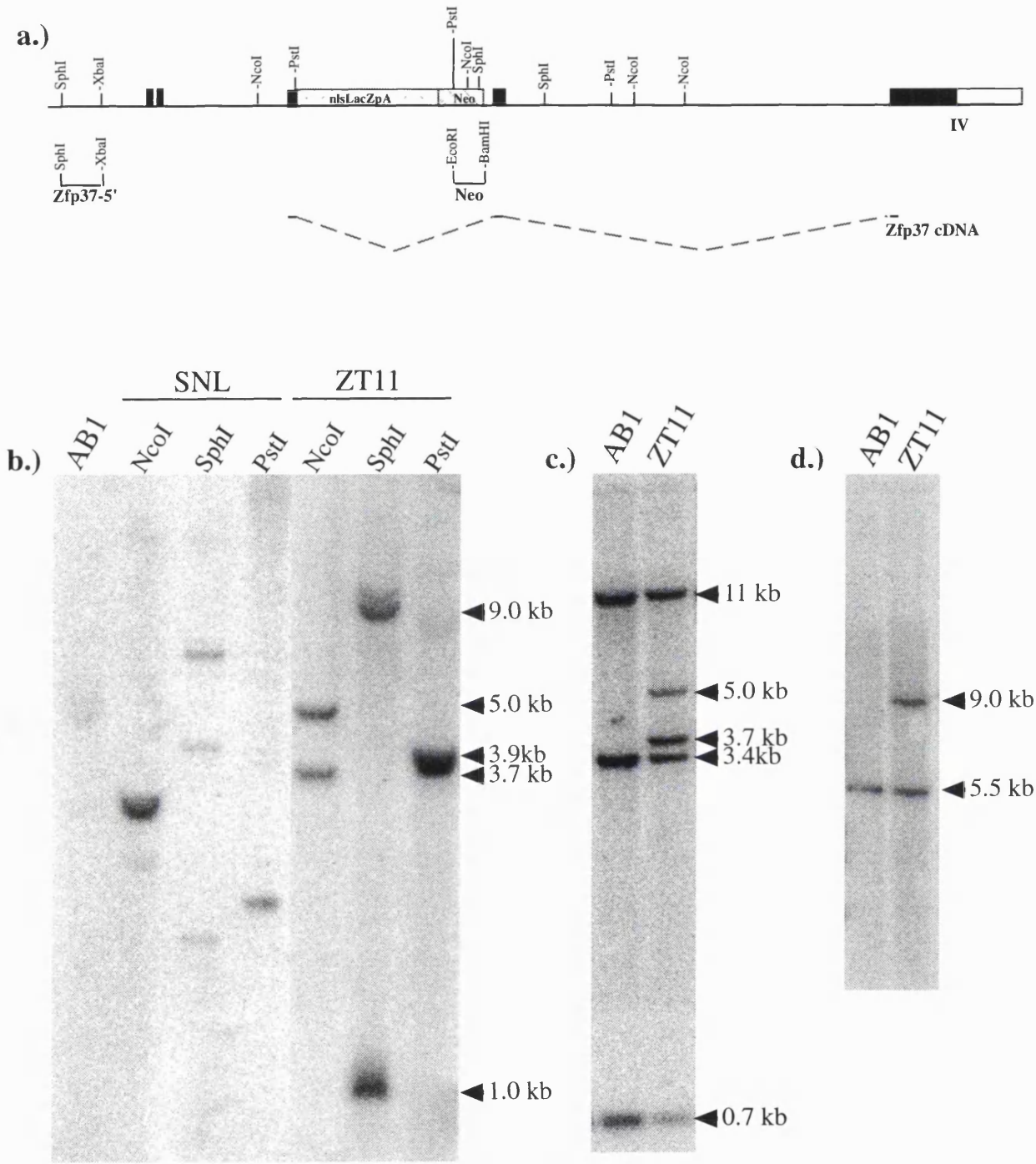
a) Predicted organisation of the targeted *Zfp-37* allele. Location of restriction enzyme sites used for Southern analysis and position of probes used are shown below.

b) Southern blot of ZT11 DNA digested with NcoI, SphI and PstI and probed with the Neo specific probe. SNL DNA digested with the above enzymes and hybridised with the Neo probe is included to show the Neo hybridising bands in this cell line. Wild type AB1 ES cell DNA cut with NcoI is included as a negative control. In the ZT11 DNA samples the NcoI digest shows the Neo hybridising fragments of 3.7 kb and 5 kb, these are predicted to occur from homologous recombination of the targeting vector into the *Zfp-37* allele. SphI digest shows the predicted fragments of 9 kb and 1kb to hybridise. The PstI digest shows the predicted 3.9 kb fragment from a correct recombination event.

c) Southern blot of AB1 and ZT11 DNA digested with NcoI restriction enzyme and hybridised with the 0.4 kb *Zfp-37*cDNA probe. In the AB1 sample hybridising fragments from the wild type allele of 700bp, 3.4 kb and 11 kb can be seen. In the ZT11 targeted clone DNA sample hybridising fragments from the targeted allele of 3.7 kb and 5 kb and 11 kb can be seen as well as the 700bp, 3.4 kb and 11 kb fragments from the wild type allele. Introduction of the nlslacZneo cassette into the *Zfp-37* allele destroys the NcoI site in exon II and therefore in the targeted allele the 700 bp fragment will be lost. This can be seen by the reduced intensity of this band as compared to the signal from the two wild type alleles.

d) Southern blot of AB1 and ZT11 DNA digested with SphI and hybridised with the SphI-XbaI *Zfp-37* 5' genomic probe. The wild type allele shows a 5.5 kb hybridising fragment. In the ZT 11 sample the 5.5 kb wild type fragment and 9 kb targeted fragment from introduction of the 4.5 kb nlslacZneo cassette into the *Zfp-37* allele are seen to hybridise.

The hybridised Southern blots were visualised on a Phosphorimager.


a.)
SphI
XbaI
NcoI
PstI
nlsLacZpA
Neo
PstI
NcoI
SphI
SphI
PstI
NcoI
NcoI
IV
SphI
XbaI
Zfp37-5'
EcoRI
BamHI
Neo
Zfp37 cDNA
b.)
AB1
SNL
ZT11
NcoI
SphI
PstI
NcoI
SphI
PstI
9.0 kb
5.0 kb
3.9kb
3.7 kb
1.0 kb
c.)
AB1
ZT11
11 kb
5.0 kb
3.7 kb
3.4kb
0.7 kb
d.)
AB1
ZT11
9.0 kb
5.5 kb

(McMahon and Bradley 1990). The efficiency of the counter enrichment was calculated to be in the order of 1 in 22. In total 165 $G418^r/FIAU^r$ clones were obtained, however some of these appeared earlier than 8 days of selection and were of increased size, these were discounted. From the remainder a total of 96 clones were picked into a 96 well feeder plate. At confluence each clone was split into three, one sample was grown to confluence in a well of a 96 well plate and frozen, the other two samples were grown to confluence and used for DNA isolation.

From the 96 clones picked, 75 clones grew well and yielded enough cells for freezing and DNA analysis. From one super confluent well of a 96 well plate enough DNA for one Southern blot analysis was isolated, approximately 5μg. The 5' external SphI-XbaI probe was used in conjunction with a SphI digest to screen for correct recombination. From the 75 clones screened two clones appeared to have undergone homologous recombination by this assay (Figure 23).

Karyotyping of the subclone initially used for the above targeting experiment has shown that 80% of the cells carry 41 chromosomes, a trisomy 8 (M. Jaegle personal communication). The clones obtained have not yet been tested by karotyping or injection into blastocysts but the karyotype of the starting subclone predicts that targeted clones obtained will not be suitable for generation of chimeric mice.

3.2.5 Generation of *Zfp-37*$^{+/-}$ mice

The karyotype of the targeted AB1 ES cell clone ZT11 was assessed by chromosomal counts from metaphase spreads. Forty spreads were counted, 77.5% of the spreads showed a normal number of 40 chromosomes, 7.5% contained 36 chromosomes, 7.5% 41 chromosomes, 5% 39 chromosomes and 2.5 % 38 chromosomes. Since this ES cell clone had a predominantly normal chromosomal count blastocyst injection was pursued.

Four sets of blastocyst injections were carried out, the results of which are summarised in Table IV. From the 49 live births 37 animals were chimeric and 80% showed coat colour chimerism of between 80-100%, with little or no non-agouti hair detectable in some cases. Of the chimeras born 65% were male, suggesting that sex conversion may have taken place in some cases. Abnormal genitalia were observed in

Figure 23. Southern blot analysis of G418^r/FIAU^r E14 ES cell clone from second *Zfp-37* targeting experiment

Southern blot of E14 ES cell clone DNA digested with SphI and hybridised with SphI-XbaI *Zfp-37* 5' external genomic DNA probe. E14 and ZT11 targeted clone DNA are included as negative and positive control lanes. Samples showing the predicted recombination fragment of 9 kb are highlighted with an asterisk.

The hybridised Southern blot was visualised on a Phosphorimager.

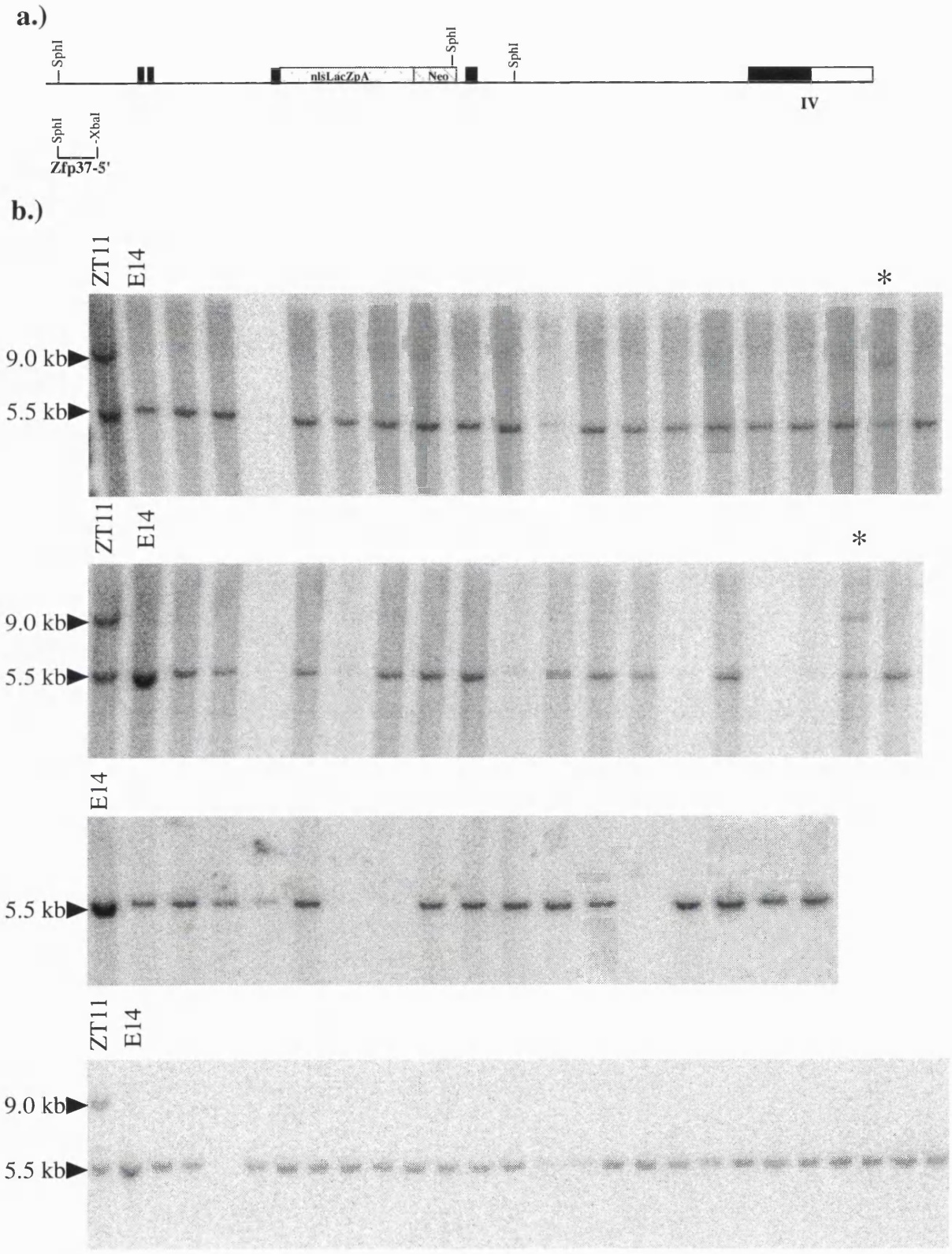

a.)
SphI
SphI
SphI
nlsLacZpA
Neo
IV
SphI
XbaI
Zfp37-5'
b.)
ZT11
E14
*
9.0 kb
5.5 kb
ZT11
E14
*
9.0 kb
5.5 kb
E14
5.5 kb
ZT11
E14
9.0 kb
5.5 kb

Table IV-Blastocyst injection of *Zfp-37*$^{+/-}$ AB1 ES cells

Experiment	Blastocysts injected	Births	Chimeras	Male	Female
1	21	12	11	8	3
2	30	14	11	6	5
3	30	6	2	2	-
4	36	17	13	8	5
Total	117	49	37	24	13

some of the females, this was manifest as an extended distance of the vagina from the anus. A number of the females also showed high coat colour chimerism suggesting that the ES cells were more efficient at contributing to somatic tissue than their contribution to the germline.

The twenty four male chimeras were bred with C57BL/6 females and germline transmission from ES cell derived germinal tissue was scored by coat colour. DNA preparations from tail biopsies of agouti pups were analysed by Southern blot for the presence of the targeted allele. Seven of the chimeras appeared infertile as they failed to sire any litters. Approximately 450 offspring were analysed from the other 17 chimeras. In total eleven sporadic germline transmission events were seen from 5 of the 17 male chimeras as determined by the presence of the agouti coat colour in the offspring. The frequencies of transmission ranged from 1 in 3 to 1 in 44 agouti per non-agouti pups. Southern blot analysis of the eleven agouti pups showed four to contain a disrupted *Zfp-37* allele (Figure 24a). This confirms the successful transmission of the targeted *Zfp-37* allele into the mouse germline.

At six weeks of age the mice heterozygous for the disrupted *Zfp-37* allele were intercrossed to establish the fate of homozygous animals. The males were also out-crossed to establish a colony of *Zfp-37*$^{+/-}$ mice. As the C57BL/6 females used proved to be poor mothers and to produce small litters, FVB/N females were used which consistently gave litters of 8-10 pups.

3.2.6 Results from crossing *Zfp-37*$^{+/-}$ mice

Mice heterozygous for the *Zfp-37* targeted disruption (*Zfp-37*$^{+/-}$) were established on a (129 x C57BL/6) genetic background and also on a (129 x C57BL/6 x FVB) background. Animals from both backgrounds were interbred to establish the fate of mice homozygous for the disrupted *Zfp-37* allele. Offspring from the crosses were genotyped by Southern blot at 8-10 days of age for the disrupted allele. This was carried out using the 0.4 kb cDNA probe on a NcoI digest (Figure 24b). Crossing heterozygous mice should yield wild type, heterozygous and homozygous animals in the Mendelian ratio 1:2:1. Numbers and fate of animals are shown in Table V.

Initial results from the first three litters on the (129 x C57BL/6) background

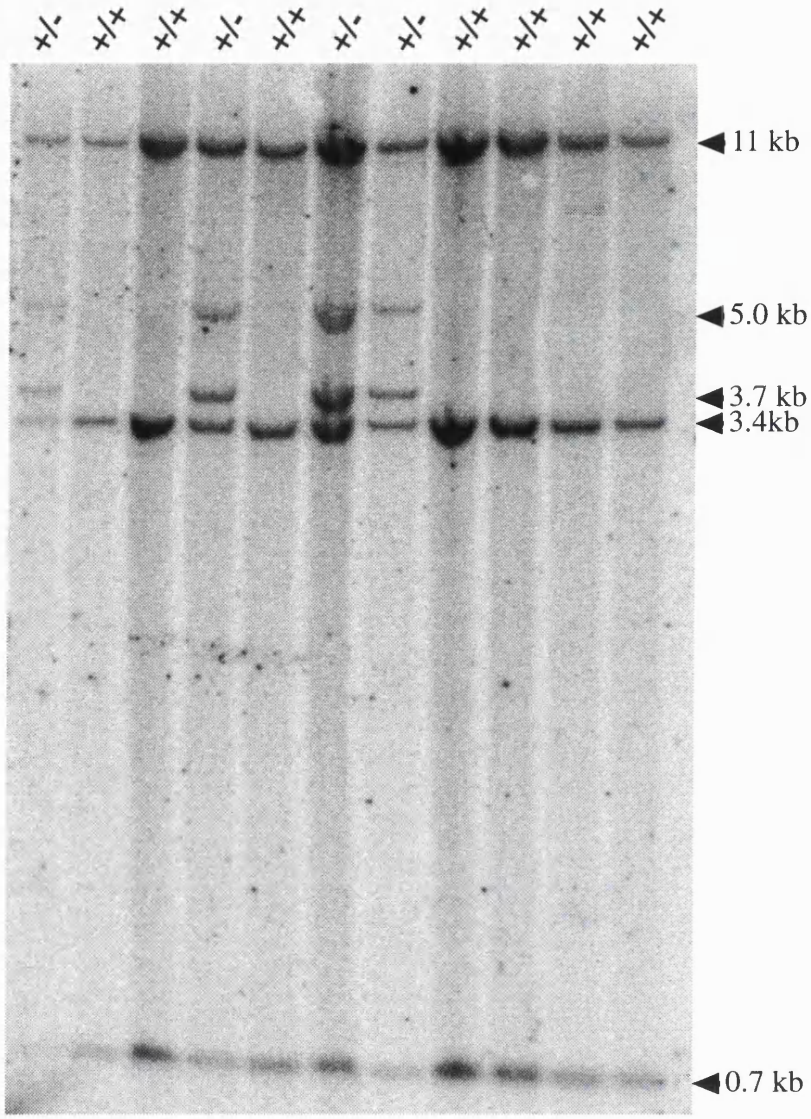
Figure 24. Southern blot screening for germline transmission of the targeted *Zfp-37* allele

a) Genomic DNA isolated from tail biopsies from 11 agouti pups digested with NcoI and hybridised with 0.4 kb *Zfp-37* cDNA probe. Transmission of the targeted allele can be seen in four of the eleven samples as shown by the hybridising NcoI restriction fragments of 3.7 kb and 5.0 kb.

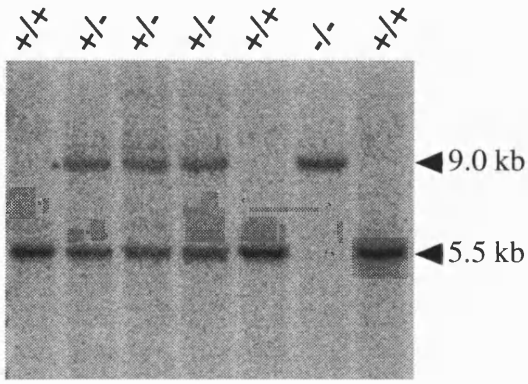
b) Genomic DNA isolated from tail biopsies from a typical litter of a $Zfp\text{-}37^{+/-}$ cross digested with SphI, Southern blotted and hybridised to the 0.9 kb SphI-XbaI *Zfp-37* genomic probe. The wild type allele and targeted alleles can be distinguished by the hybridising restriction fragments of 5.5 kb and 9.0 kb respectively. This litter contains a homozygous $Zfp\text{-}37^{-/-}$ mouse.

The hybridised Southern blots were visualised on a Phosphorimager.

a.)



b.)



showed a lower percentage of homozygous animals born than would be predicted (Table V). From the 26 animals born the numbers of wild type, heterozygous and homozygous animals obtained were 10, 12 and 4 respectively. Of the four homozygous animals two were found dead at postnatal day 2, the third died at postnatal day 20. It is also important to note that from the 26 animals analysed two heterozygous animals and one wild type animal have also been found dead at postnatal day two. The 20 day old homozygous animal that died showed a progressively retarded development, being visually identifiable by its reduced size and movement as compared to its litter mates. After death the brain was removed, fixed, stained with X-gal and sectioned. Comparison of the sections with those of a heterozygous litter mate showed no obvious structural abnormalities. From the above data it is unclear whether the deaths seen were due to the targeted mutation. To further examine the effect seen in the (129 x C57BL/6) background, heterozygous animals have now been bred with wild type C57BL/6 to generate a F_2 generation of heterozygous animals in this background. These animals are being interbred, and one more litter has been born. From the 8 animals born 2 are *Zfp-37*$^{-/-}$, 3 are *Zfp-37*$^{+/-}$ and 3 are *Zfp-37*^{wt}, bringing the total numbers of animals analysed in the 129 x C57BL/6 background to 34 of which 6 have been *Zfp-37*$^{-/-}$, no further animals have died (Table V). There appears to be a lower number of homozygous animals than would be predicted, however a larger number of animals must be screened on this background before any conclusions can be drawn with any validity.

In the 129 x C57BL/6 x FVB/N background 112 animals have been genotyped to date. The distribution of wild type, heterozygous and homozygous animals is 29, 54 and 29 respectively, this is close to the predicted ratio of 1:2:1. From these animals one wild type and one heterozygous have been found dead in the litter within the first two postnatal days. These were in litters of nine and eight pups respectively. Three homozygous animals died, either at two days of age or in one case at three weeks. An autopsy was carried out on this animal, although autolysis had begun to occur in some tissues, particularly the spleen and liver, the only abnormality observed was seen in the lungs. The lungs appeared inflamed and infiltrated with lymphocytes, which would be concurrent with a lung infection. The other homozygous animals appear normal, showing no obvious behavioural problems although more stringent tests have not as yet been carried out. Homozygous males and females have been intercrossed and produce

Table V-Genotype distribution in offspring from *Zfp-37*$^{+/-}$*xZfp-37*$^{+/-}$cross

Background*	No.	+/+	+/-	-/-	Dead +/+	Dead +/-	Dead -/-
129xC57BL6xFVB	112	29	54	29	1	1	3
129xC57BL6	34	13	15	6	1	2	3

*F_1 and F_2 generations

a viable litters of 8-10 pups. This suggests that the disruption does not affect spermatogenesis and implantation of the embryo, ZFP-37 mRNA is also found expressed in placenta and yolk sac.

3.3 Discussion

The murine *Zfp-37* gene is a recently isolated Cys_2His_2 zinc finger gene which is postulated to play a role in transcriptional regulation in the developing and adult nervous system and also during spermatogenesis (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994; Mazarakis *et al.*, 1995). More recently ZFP-37 has been shown to contain a KRAB domain, placing it into this large group of Cys_2His_2 zinc finger genes (Mazarakis *et al.*, 1995). This chapter describes the successful targeting of the murine *Zfp-37* gene in embryonic stem cells and the subsequent introduction of the mutated allele into the mouse germline.

3.3.1 Summary of targeting experiments and generation of chimeric mice

Comparison of the *Zfp-37* targeting experiments carried out in the AB1 and E14 cell lines has highlighted a number of differences in transfection efficiency and number of targeted clones obtained using the different protocols. The targeting experiment carried out in the AB1 ES cell line yielded a low number of $G418^r/GANC^r$ clones. The number of $G418^r$ colonies obtained showed a transfection efficiency of $9x10^{-5}$ $G418^r$/total cells transfected. This is 2 fold lower than the transfection efficiency of $1.8x10^{-4}$ $G418^r$/total cells transfected, obtained when the E14 ES cell line was used. This difference is likely to be due to the electroporation conditions used in the two experiments. The enrichment provided by the counter selection in the AB1 and E14 experiments was approximately 36- and 12-fold, respectively. This difference in enrichment could be accounted for by the efficacy of the different HSV1 thymidine kinase substrates, GANC and FIAU. However FIAU is reported be more efficient for counter selection than GANC (McMahon and Bradley, 1990), it is possible that the FIAU concentration was lower than assumed.

The targeting efficiency of the 7.1 ZTV vector was approximately 1 in 340

$G418^r$ clones in the AB1 experiment, assuming that the diagnostic 6.2 kb Neo integration pattern represents a legitimate recombination event. In the E14 targeting experiment the targeting efficiency was approximately 1 in 450 $G418^r$ clones. This variation in efficiency is relatively minor and may reflect to variable efficiency of the batches of G418 used for the selection procedure or discrepancies in cell numbers transfected.

The AB1 targeted clone ZT11, once injected into blastocysts, was able to contribute to the somatic tissue of the resulting chimeric animal. This was determined by coat colour contribution, with 75% of the mice born being chimeric. Many of the chimeras showed high coat colour contribution of up to 80-100%. A distortion in the predicted male/female ratio was seen with 24/37 animals being male, however this distortion was not seen in all litters and is not has high as described in other studies (Robertson *et al.*, 1986). As the ES cells injected are XY they can convert a female blastocyst into a phenotypic male chimera (Bradley *et al.*, 1984). If coat colour reflects contribution to the germ cells then all the high coat colour chimeric animals should be male. This was not the case in this study, suggesting the ES cell clone used was not as efficient in contributing to the germline as the coat colour chimerism would lead one to believe. In the eight high contribution females (70-100%), abnormalities in the urogenital system were seen. This was manifest as increased distance of the female genitalia from the anus, appearing more male than female. It is possible that some of these animals were hermaphrodite and are a result of incomplete sex conversion, though this was not examined further.

The above results predict a low contribution of the ES cells to the germline of the chimeric mice. This was reflected in the sporadic transmission of the ES cell coat colour allele and would suggest that a percentage of the ES cells used may have entered a differentiation pathway which allowed them to contribute to the somatic tissue but not to the germ cells. Extensive breeding was required before transmission was seen, some males having to sire three or four litters before agouti pups were seen. The reason for this delayed transmission is not known. From the eleven transmission events seen, four animals were found to carry the targeted allele.

Despite the low transmission frequency, animals heterozygous for the targeted allele were successfully obtained. As ZFP-37 is expressed in mature post meiotic

spermatocytes both at the level of mRNA and protein (Burke and Wolgemuth, 1992, Hosseini *et al.*, 1994) the protein may be required for the proper function in the haploid sperm. If this is the case, then germline transmission may be affected by the sperm carrying a *Zfp-37* targeted allele. However, as ES cell derived offspring not carrying the targeted allele were also represented at a low level, this explanation is less plausible. Since transmission of the targeted allele was obtained and heterozygous animals appear healthy the disruption does not have an autosomal dominant effect on murine development.

3.3.2 The fate of *Zfp-37*$^{-/-}$ mice

To determine the effect of the introduced disruption, mice heterozygous for the targeted *Zfp-37* allele were intercrossed. This was carried out on outcrossed genetic backgrounds of two and three strains. Within the 129 x C57BL/6 background twenty six animals were genotyped from three litters. Although the numbers of mice born which were homozygous for the disruption were slightly lower than expected, the numbers analysed were too low to be significant. However, of these four homozygous animals two died within the first 48 hrs and a third died at three weeks. The cause of death in these animals is not known. In large litters the loss of one pup is not unusual, the early postnatal losses may be due to this fact. Furthermore of the 26 animals born from this cross a total of six have died, three homozygous, two heterozygous and one wild type animal. The heterozygous founders used for the breeding are from the original germline transmission and therefore will have the highest content of the ES cell genome. If the ES cells carried another mutation obtained during culture or during the targeting experiment then this would be most manifest in this cross. As approximately a quarter of the animals obtained died in this cross, this could be due to these animals being homozygous for an unknown lethal mutation. To establish whether in the (129 x C57BL/6) genetic background the observed mortality segregates with the targeted *Zfp-37* allele we have bred an F_2 generation of heterozygous mice on this background. From the interbreeding of these heterozygous mice a further two mice homozygous for the disruption were obtained, these appear to be healthy.

On the three strain background homozygous animals appear healthy and fertile

with a close to normal mendelian representation. There have been three exceptions to this, two animals dying by postnatal day 2 and a third at three weeks. Postmortem examination of the third mouse suggested the most likely reason for death to be due to a lung infection as there were clear signs of immune response in this tissue. *Zfp-37* transcripts have been detected in lung but only at very low levels. It seems unlikely that this death is directly due to the *Zfp-37* mutation. Interestingly, mice homozygous for the targeted *Zfp-37* allele can successfully mate and produce young which appear normal within this outbred background. This allows one to conclude that targeted mutation made in the *Zfp-37* gene does not prevent spermatogenesis or placenta formation in the three strain background, both sites for *Zfp-37* expression (Burke and Wolgemuth, 1992).

The different effect of the mutation in the (129 x C57BL/6) compared to the (129 x C57BL/6 x FVB) can be accounted for in a number of ways. Firstly, as discussed above the 129 x C57BL/6 mice from the germline transmission may contain an unknown mutation which manifests itself when made homozygous in this cross. Outbreeding to FVB dilutes the effect of this unfortuitous mutation and it no longer appears associated with the *Zfp-37* disruption. Secondly, we may be observing a strain specific effect. The FVB strain may carry some modifier which rescues the *Zfp-37* mutation in this cross. Such an effect is well documented for the Keratin 8 disruption (Baribault *et al.*, 1993 and 1994). This is being investigated further.

The expression pattern of the *Zfp-37* gene and the presumed structure of its protein have lead us to speculate that ZFP-37 plays a role in transcriptional regulation in the nervous system during development and adult life (Mazarakis *et al.*, 1995). Though this may be the case, the survival of the mice homozygous for the targeted *Zfp-37* mutation in the three strain background would suggest that the mutation made in the *Zfp-37* genes does not lead to a life threatening defect in these system. It has not yet been determined whether the mutation leads to a behavioural phenotype. Further examination of the expression of the *Zfp-37* gene utilising the targeted *lacZ* reporter gene may provide further insights into the neuronal systems in which ZFP-37 protein functions. Furthermore, I have shown that, at the DNA level, the *Zfp-37* gene has been disrupted by the insertion of the *lacZneo* cassette, however it must be determined whether the targeted mutation leads to a null allele before conclusions can be drawn as to ZFP-37 function. The following chapters describe the analysis of *Zfp-37* expression

pattern via the targeted reporter gene and the effect of the targeted mutation upon *Zfp-37* expression.

Chapter 4

Analysis of *Zfp-37* expression pattern via the targeted *lacZ* reporter gene

4.1 Introduction

Previous analysis of the expression pattern of *Zfp-37* have revealed the major sites of expression to be the developing and adult nervous system and the testis (Burke and Wolgemuth, 1992; Hosseini *et al.*, 1994; Mazarakis *et al.*, 1995). Other sites of expression include placenta and yolk sac (Burke and Wolgemuth, 1992; Mazarakis *et al.*, 1995). *In-situ* RNA hybridisation data from e12.5 mouse embryos has also shown the gene to be expressed in the developing brain and to a lesser extent in the spinal chord. Expression in the adult brain appears to be pan neuronal (Mazarakis *et al.*, 1995), though certain neuronal cell types show higher levels of expression than others. For example, in the cerebellum the Purkinje cells show a stronger hybridisation signal than the granular neurons (Mazarakis *et al.*, 1995).

As described in chapter 3, the targeting vector used to disrupt *Zfp-37* includes a bacterial *lacZ* gene in-frame with the *Zfp-37* coding region. This allows the transcription of *Zfp-37* to be followed by staining for β-galactosidase activity using the chromogenic substrate X-gal. The inclusion of this reporter gene has a number of advantages. Firstly it allows further analysis of *Zfp-37* expression to be carried out with relative ease. Secondly, in mice homozygous for the disruption, it provides a means to mark cells that would have been expressing *Zfp-37.* If any structural abnormalities result from the targeted disruption, the fate of cells that would have expressed *Zfp-37* can be followed.

This chapter describes the analysis of *Zfp-37* expression pattern via the targeted *lacZ* reporter gene, and mainly concentrates on the neuronal expression. The work described in this chapter, although not yet a full developmental study of the expression pattern, may provide further clues as to the processes in which ZFP-37 plays a functional role.

4.2 Results

4.2.1 Expression of the targeted *lacZ* reporter gene in embryos from dpc 9.5-12.5

Embryonic expression of *Zfp-37* has so far been addressed by Northern blot

analysis from e10.5 onward and by *in-situ* RNA hybridisation of e12.5 embryo sections (Mazarakis *et al.*, 1995). In this study the targeted *lacZ* reporter gene was used to examine *Zfp-37* expression. Mouse embryos were isolated 9.5 days post coitum (dpc) from a *Zfp-37*$^{+/-}$ heterozygous cross. The embryos were mildly fixed and placed in X-gal staining solution. Yolk sac was taken for genotype analysis. No staining was observed in wild type embryos after incubation overnight in staining solution. After whole mount staining sagittal sections of the homozygous embryos were made which showed staining is in the anterior neural tube with the most intense region of staining being above the mesencephalic flexure (Figure 25a and b). Staining can also be seen in the developing gut as shown in transverse section (Figure 25c). Weaker but widespread staining is seen throughout the embryo, with the exception of the developing heart which appears to be negative. In the whole mount staining of homozygous (*Zfp-37*$^{-/-}$) embryos staining can be seen in the in the ventral neuroepithelia, most predominantly around the mesencephalic flexure (Figure 25d). Staining is also detectable in the body of the embryo. In heterozygous (*Zfp-37*$^{+/-}$) embryos, staining appears to be very weak in both whole mount stained embryos as compared to the homozygous embryos, although it is visible in the neuroepithelia surrounding the mesencephalic flexure (data not shown), this staining is too weak to be detected in sections.

A similar pattern of expression to the e9.5 embryos is seen in the e10.5 embryos. β-galactosidase expression is seen around the cephalic flexure, there appears to be an increase in staining in the hindbrain and the ventral region of the neural tube. Staining is also apparent around the developing limb buds (Figure 26a and b). Homozygous and heterozygous animals show an identical expression pattern but the staining in the homozygous embryos again appears much stronger. Again no staining is observed in the wild type embryos.

At e12.5 staining is more pronounced. Expression of the *lacZ* reporter gene is seen throughout the neuroepithelia (Figure 27a and b). From the sagittal view of the spine this appears to be restricted to the ventral half of the neural tube rather than whole neural tube (Figure 27a). From the whole mount stained embryos it is more difficult to see if this is the case for the developing brain. In the homozygous and heterozygous embryos, dorsal root ganglia also stains as do the sympathetic ganglia. This shows expression to include neural crest derived cells. *lacZ* staining can be clearly seen in the

Figure 25. Expression of the targeted *lacZ* reporter gene at 9.5 dpc

a) Sagittal section of 9.5 dpc *Zfp-37*$^{-/-}$ embryo. X-gal staining can be seen in the layer of cells surrounding the mesencephalic flexure and extends to the floor of the telencephalon. Cells in the roof of the telencephalon, diencephalon, mesencephalon and metencephalon appear negative for X-gal staining. X-gal staining can be seen throughout the body of the embryo excluding the developing heart: t, telencephalic vesicle; di, diencephalic vesicle; me, mesencephalic vesicle; met, metencephalic vesicle; mf, mesencephahilc flexure; ot, otic vesicle; ba, branchial arch; he, heart.

b) Close up of mesencepahilc flexure. Strong X-gal staining is seen in the neuroepithelia surrounding the flexure: mf, mesencephalic flexure; ne, neuroepithelia.

c) Transverse section through lower lumbar region of 9.5 dpc *Zfp-37*$^{-/-}$ embryo. X-gal staining can be seen in epithelia surrounding the hindgut.

d) Whole mount stain of 9.5 dpc *Zfp-37*$^{+/+}$ embryo. No X-gal staining detected.

e) Whole mount stain of 9.5 dpc *Zfp-37*$^{-/-}$ embryo. Staining is prominent in developing nervous tissue but is also apparent throughout the body of the embryo.

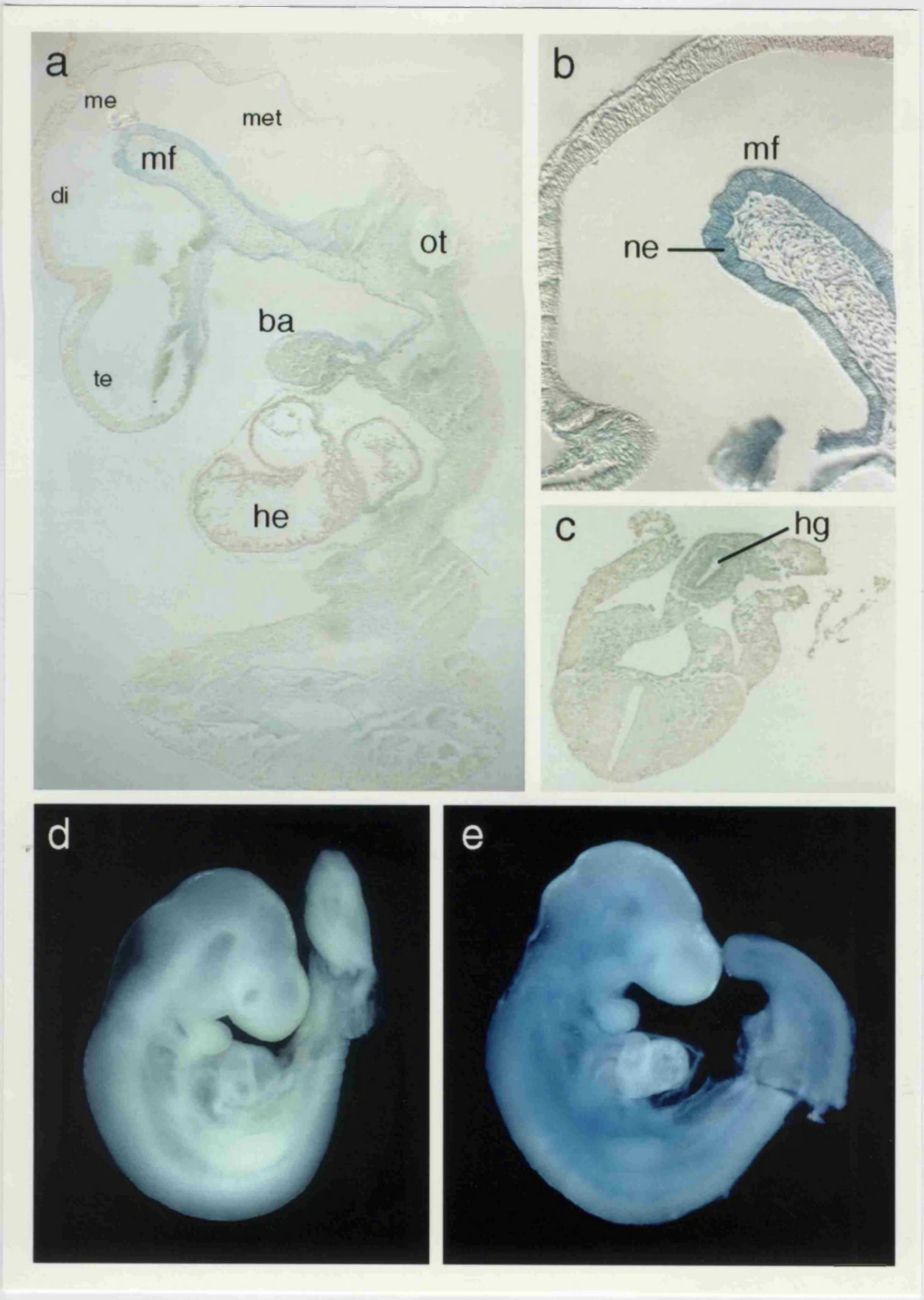

a
me
met
mf
di
ot
ba
te
he
b
mf
ne
c
hg
d
e

Figure 26. Expression of the targeted *lacZ* reporter gene at 10.5 dpc

Whole mount X-gal staining of *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$and *Zfp-37*$^{-/-}$ embryos at 10.5 dpc.

a) Sagittal view of embryos. No staining is detected in the wt embryo. Weak X-gal staining can be detected in the +/- embryo, particularly around the mesencephalic flexure. In the -/- embryo X-gal staining is seen around the mesencephalic flexure, the metencephalon and the spinal neural tube. X-gal staining appears to be reserved to the ventral half of the spinal neural tube. X-gal staining can also be detected at the base of the developing limb rudiments: mf, mesencephalic flexure; pf, pontine flexure; nt, neural tube; fl, fore limb.

b) Dorsal view of the above embryos. Weak X-gal staining is seen in the +/- embryo as compared to the -/-. In the -/- embryo X-gal staining is seen in the neural tube and developing fore limb as above.

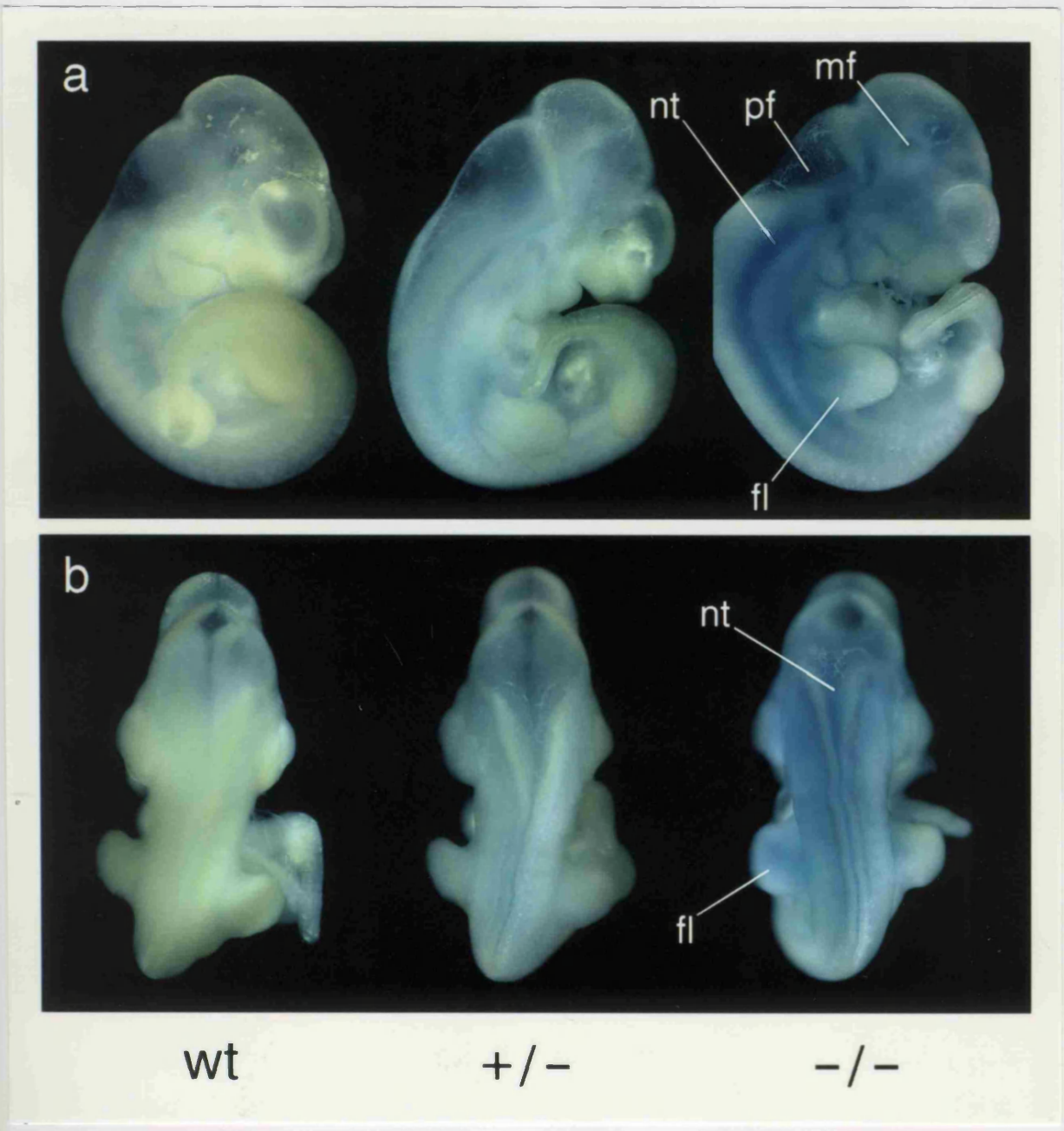
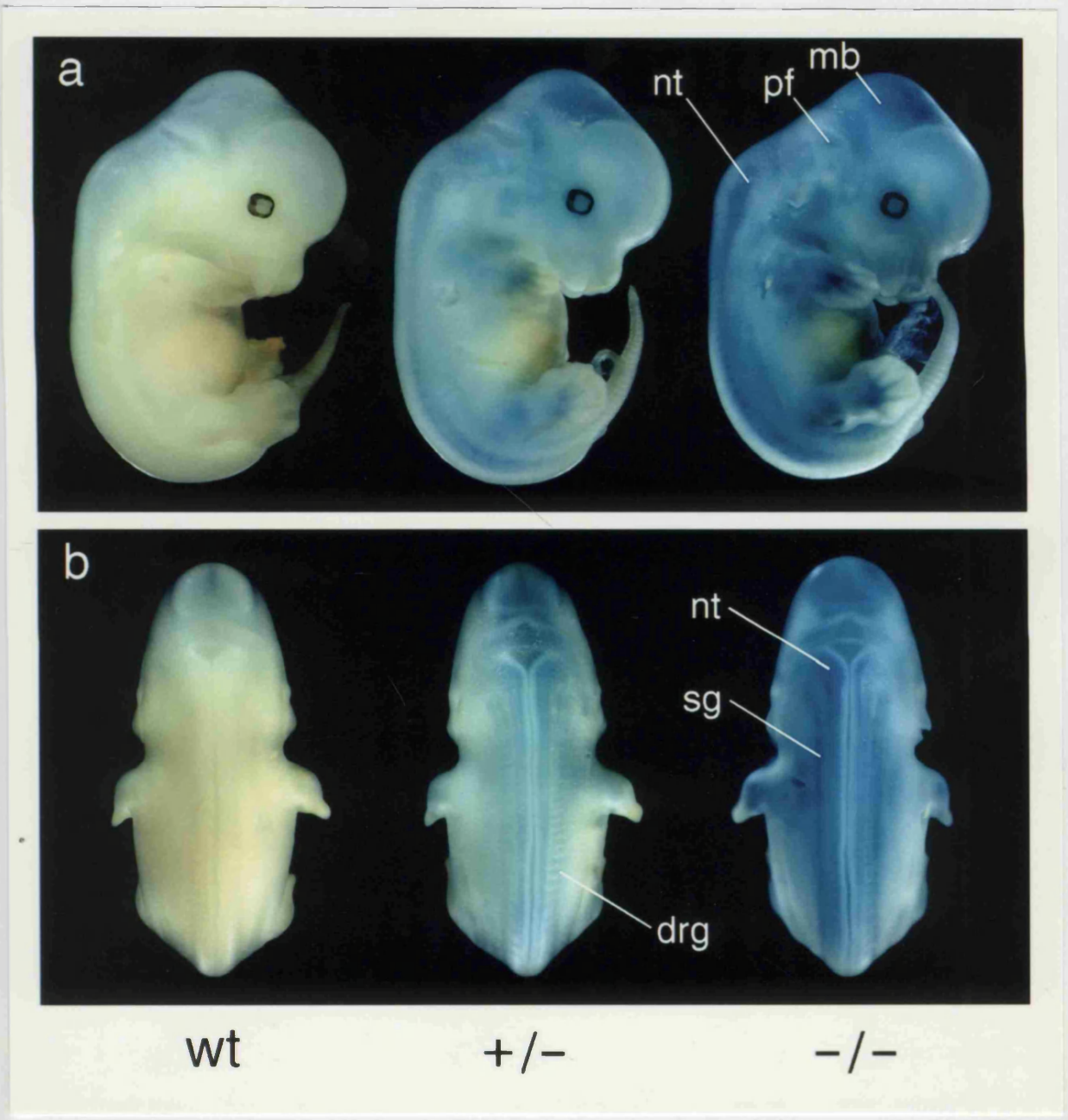

a
nt
pf
mf
fl
b
nt
fl
wt
+/−
−/−

Figure 27. Expression of the targeted *lacZ* reporter gene at 12.5 dpc

Whole mount X-gal staining of *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$and *Zfp-37*$^{-/-}$ embryos at 12.5 dpc.

a) Sagittal view of embryos. No staining is seen in the wt embryo. X-gal staining can be detected in the +/- embryo throughout the neural tube. Staining is seen in the forming digits of both the fore and hind limbs. In the -/- embryo Staining appears more intense than the +/- embryo with a similar pattern, X-gal staining appears restricted to the ventral half of the neural tube: mb, mid brain; pf, pontine flexure; nt, spinal neural tube.

b) Dorsal view of embryos shown above. Again staining is seen in the neural tube of both the +/- and -/- embryos. Dorsal root ganglia and sympathetic ganglia appear to stain in both cases: nt, neural tube; sg, sympathetic ganglia; drg, dorsal root ganglia.


a
nt
pf
mb
b
nt
sg
drg
wt
+/−
−/−

developing limbs with marked expression in the forming digits. Sectioning of both e10.5 and e12.5 embryos is being carried out to attempt to define the cell types in which the expression is found.

To address the onset of *Zfp-37* expression in development, e8.5 embryos from a *Zfp-37*$^{+/-}$ cross have also been stained. Eight embryos were stained for three days but no staining was observed. However, these embryos were not genotyped as it was difficult to obtain suitable amounts of embryo derived material for Southern blot analysis. It is unlikely that no heterozygous or homozygous embryos carrying the *lacZ* reporter gene were present in this analysis, however this has not been demonstrated and can not be discounted. This experiment needs to be repeated. Attempts to define the *lacZ* expression in e14.5 and older embryos have been made, though these have been complicated by high background from the endogenous galactosidase enzyme and poor penetration of the X-gal substrate. Staining of cryostat sections of e14.5 and older embryos are to be performed to clarify the expression pattern at these older time points.

4.2.2 Expression of the *lacZ* reporter gene in the postnatal mouse brain

A clear impression of *Zfp-37* expression in the adult brain has been provided by *in-situ* RNA hybridisation (Mazarakis *et al.*, 1995). In the study described here, brains from three week old animals were taken. The three week time point was chosen because a number of homozygous mice for the disruption died around this time. Analysis of three week old mouse brain both from heterozygous and homozygous animals are consistent with the results obtained from the *in-situ* RNA hybridisation study. However, whole mount staining has highlighted areas of high expression which were not described by the *in-situ* RNA hybridisation study.

The most striking expression of the *lacZ* reporter gene is seen in discrete neuronal structures located in the olfactory tuberculum (Figure 28). These structure are known as the Islands of Calleja and are dense clusters of granular neurons located close to the pial surface of the olfactory tuberculum (Millhouse, 1987), as shown in tissue section (Figure 28g). These structures show strong staining in both heterozygous and homozygous animals (Figure 28e and f). High expression of the *lacZ* reporter gene is also seen in the hypothalamus, particularly the suprachiasmatic nuclei and the

Figure 28. Expression pattern of the *lacZ* reporter gene in whole mount three week postnatal brains from *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice

a) Whole mount X-gal stained +/- (left) and -/- (right) brains viewed from top. Note the general increased staining in the -/- brain as compared to the +/-.

b) Above brains viewed from bottom.

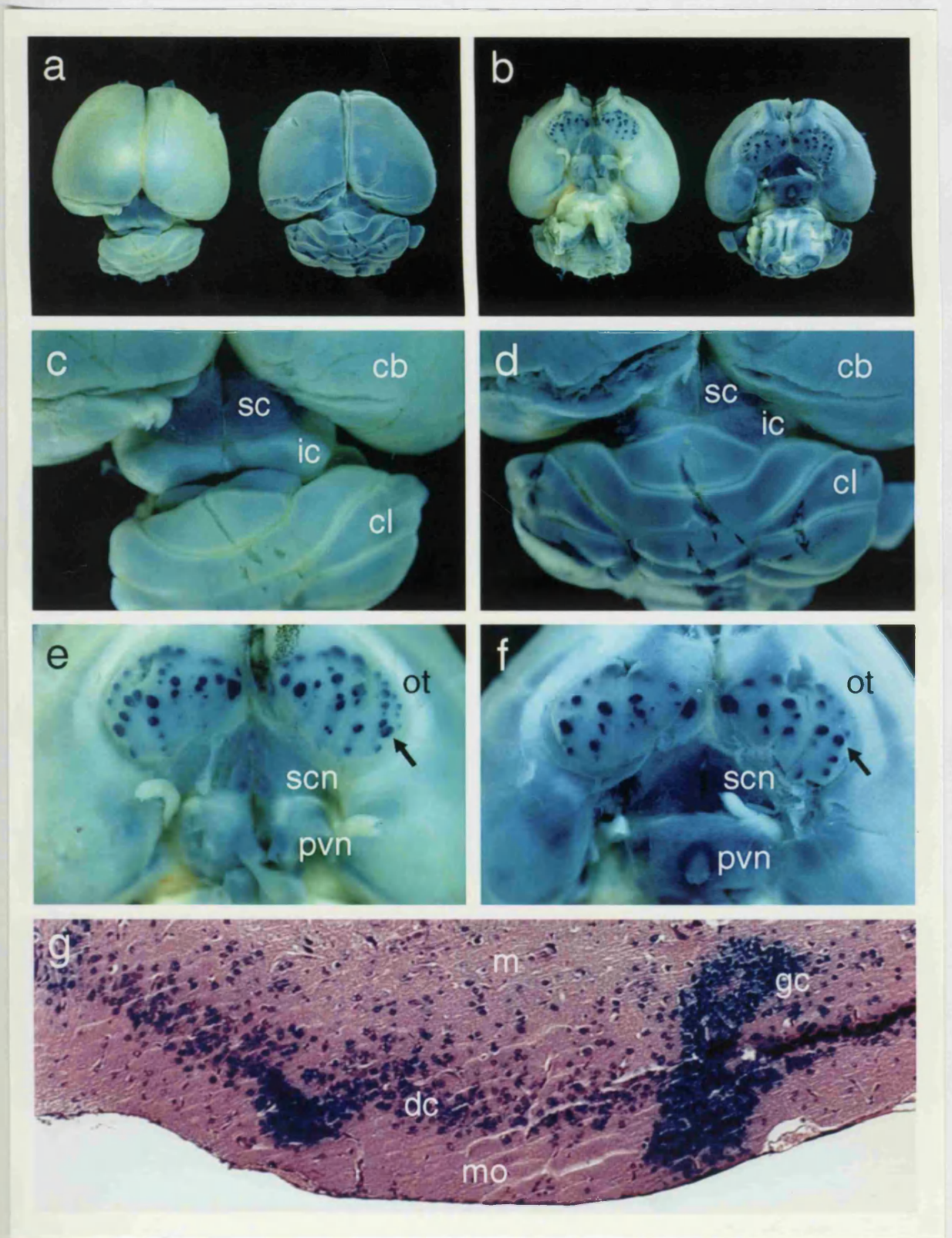
c) Close up of view of colliculi and cerebellum of +/- brain. The cerebrum shows weak staining as do the inferior colliculus and cerebellum. Note the strong X-gal staining seen in the superior colliculus as compared to the inferior colliculus; cb, cerebrum; sc, superior colliculus; ic, inferior colliculus; cl, cerebellum.

d) Close up view of colliculi and cerebellum of -/- brain. All structures shown appear to stain strongly for *lacZ* expression as compared to +/- brain. Note particularly the staining seen in the inferior colliculus as compared to the +/- sample shown in c.

e) Close up view of olfactory tuberculum and hypothalamus of +/- brain. Striking X-gal staining is seen in the islands of Calleja (denoted by arrow) as compared to surrounding structures. Staining is also seen in the suprachiasmatic nuclei and periventricular nuclei: ot, olfactory tuberculum; scn, suprachiasmatic nuclei; pvn, periventricular nuclei.

f) Close up view of olfactory tuberculum and hypothalamus of -/- brain. As for +/- brain intense staining is seen in the islands of Calleja located in the olfactory tuberculum (denoted by arrow). Stronger X-gal staining as compared to the +/- brain is seen in the suprachiasmatic nuclei and periventricular nuclei of the hypothalamus.

g) Coronal section of olfactory tuberculum of -/- brain shown in f. X-gal staining is seen in the dense cell layer particularly in the ruffled regions where cell density increases. Staining can be seen in the granular cell clusters which makes up the islands of Calleja. Weaker X-gal staining is seen in cell types located in the multiform layer: mo, molecular layer; dc dense cell layer; gc granular cell cluster; m, multiform layer.


a
b
c
cb
sc
ic
cl
d
cb
sc
ic
cl
e
ot
scn
pvn
f
ot
scn
pvn
g
m
gc
dc
mo

periventricular arcuate nuclei surrounding the infundibulum (Figure 28e and f). Looking dorsally, the superior colliculus (SC) shows expression of the *lacZ* reporter in brains from both heterozygous and homozygous (Figure 28c and d). Interestingly staining in the inferior colliculus (IC) appears to differ between heterozygous and homozygous animals. In the brain from the homozygous animal both the SC and IC stain strongly and appear to stain to the same degree. However in the heterozygous animal the IC shows a decreased staining as compared to the SC. This observation is not restricted to these two areas, in general staining appears to either be more widespread or stronger in the homozygous brain. In whole mount no X-gal staining can be seen in the wild type samples.

Upon sectioning of the stained whole mount brains only structures close to the surface showed *lacZ* expression, suggesting a problem of penetration of the X-gal substrate. To overcome this problem internal surfaces were exposed by cutting coronal sections of 2-3mm with a razor blade prior to staining. This revealed that the hippocampus shows the same expression pattern as described for the adult *in-situ* RNA hybridisation study. Weak staining of the endogenous β-galactosidase can be seen around the hippocampus in the wild type brain (Figure 29a). Strong expression is seen in the dentate gyrus and the CA1 and CA3 pyramidal fields in the heterozygous brain (Figure 29b). Staining can now also be seen in neuronal layers of the cortex. Strong staining is also seen in the habenular nuclei (Figure 29b). In the cerebellum, the Purkinje and granular cell layers show expression of the *lacZ* reporter gene (Figure 29c), no staining is seen in the white matter suggesting that glial cells are negative for *Zfp-37* expression, this has also been concluded from the *in-situ* RNA hybridisation experiments (Mazarakis *et al.*, 1995). Comparison of the brains from *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice has revealed no obvious morphological differences.

4.3 Discussion

4.3.1 General considerations of using the targeted *lacZ* reporter gene

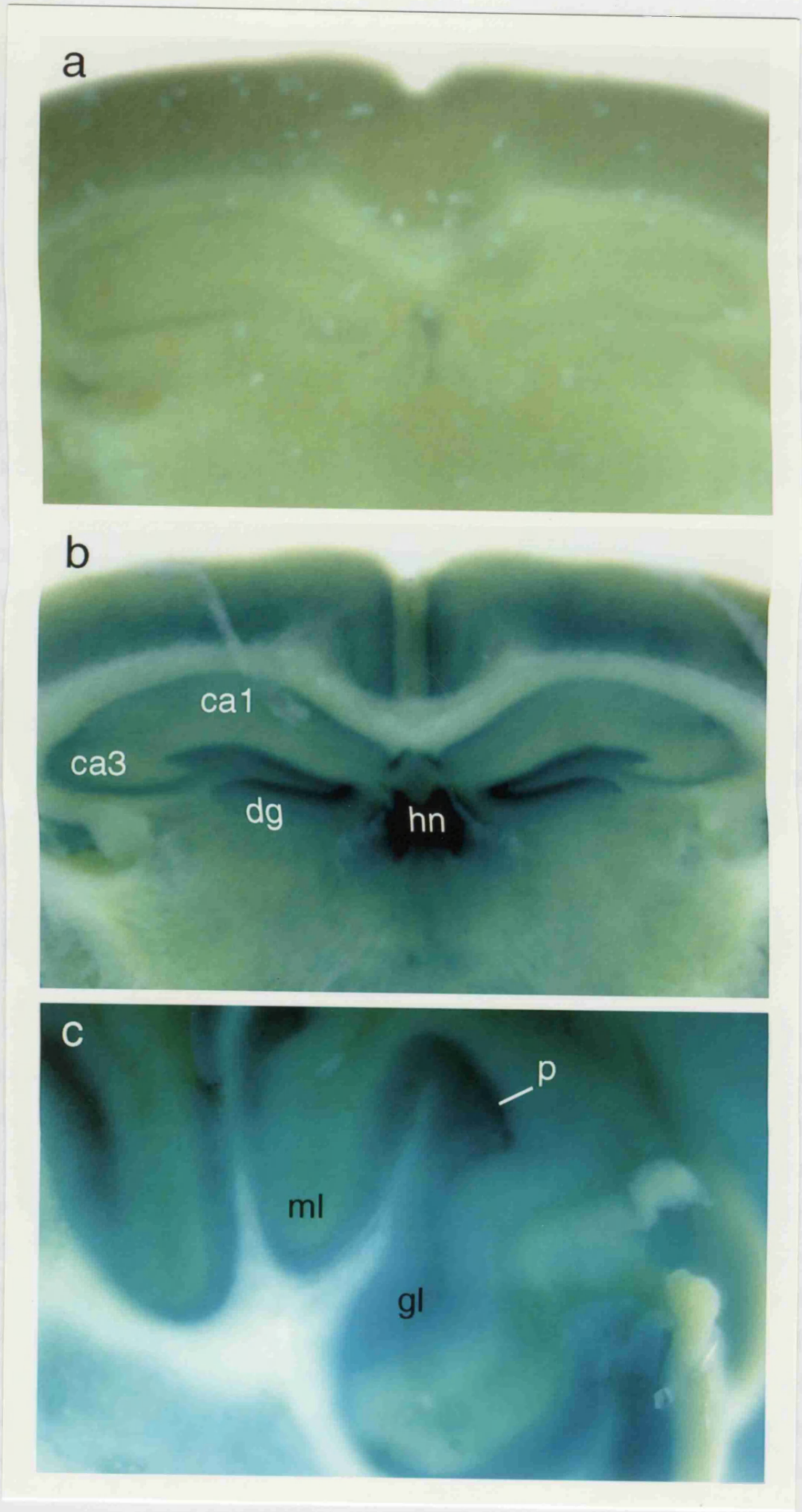
This chapter describes the preliminary analysis of the *Zfp-37* expression pattern via the targeted *lacZ* reporter gene. Although this analysis is not complete a number of

Figure 29. Expression of the targeted *lacZ* reporter gene in internal neuronal structures of *Zfp-37*$^{+/-}$ mice.

a) Coronal cross section through a wild type three week postnatal brain. Weak X-gal staining can be seen due to the endogenous β-galactosidase activity.

b) Coronal section through cerebrum of +/- brain. X-gal staining can be seen in the CA1 and CA3 pyramidal neuron layers and the dentate gyrus of the hippocampus. Staining is also seen in the habenular nuclei and various other neuronal layers of the cortex. No staining is seen in the white matter: dg, dentate gyrus; hn, habenular nuclei; ca1, ca3 pyramidal fields of the hippocampus.

c) Coronal section through cerebellum of +/- brain. X-gal staining is seen in the Purkinje cell layer of the cerebellum, appearing to stain more strongly than cells located in the granular layer. No staining is seen in the white matter: p, Purkinje cells; gl, granular cell layer; ml, molecular layer.


a
b
ca1
ca3
dg
hn
c
p
ml
gl

conclusions canbe drawn regarding the expression pattern of *Zfp-37*. Comparison of the *lacZ* reporter gene expression with that of the expression pattern of *Zfp-37* from in-situ RNA hybridisation studies show the targeted reporter gene to mimic the normal expression pattern. One concern with the *lacZNeo* cassette used in this experiment was that the enhancer-promoter region driving the neomycin resistance gene may interfere with the *Zfp-37* promoter activity, possibly causing ectopic expression of the reporter gene (Le Mouelic *et al.*, 1991). From comparison with the known expression pattern of *Zfp-37* this does not appear to be the case. Furthermore, this cassette has been included in the targeting vector used for disruption of the erythroid krüppel like factor gene (EKLF) (Nuez *et al.*, 1995), again the *lacZ* reporter gene mimics the expression pattern of the target gene EKLF. This shows that within a different genomic context, the reporter gene is expressed faithfully.

Another factor that should be considered when interpreting the expression data obtained from the reporter gene is the half life of both the reporter gene mRNA and protein. The β-galactosidase protein half life is likely to differ from that of the *Zfp-37* transcript and protein. Therefore, the staining pattern observed from the *lacZ* reporter gene expression only follows the *Zfp-37* promoter activity, reflecting the onset of expression rather than its duration.

4.3.2 Expression pattern of *Zfp-37*

In previous studies, *Zfp-37* was shown to be expressed in the embryo at e10.5 by Northern blot analysis, no earlier time points were studied (Mazarakis *et al.*, 1995). This study shows, via the targeted reporter gene, that *Zfp-37* is already expressed at day e9.5, although it is yet to be established conclusively whether this expression extends into earlier time points. At e9.5 the highest level of the reporter gene expression, in both the hetero- and homozygous embryos, is seen in the neuroepithelia of the di- and mesencephalon surrounding the mesencephalic flexure. This region of the developing neuroepithelia is known to give rise to dopaminergic neurons (Altman and Bayer, 1981; Hynes *et al.*, 1995). However, expression is also seen in the epithelia of the hindbrain and posterior neural tube, this would then argue against *Zfp-37* expression being restricted to a particular sub-population of developing neurons. Expression can also be

detected in the homozygous embryo in cells located in the epithelia of the hindgut. This shows that *Zfp-37* expression is not restricted to the developing nervous system at this stage.

At e10.5 and e12.5 staining is seen throughout the developing nervous system. In the homozygous stained embryo expression appears restricted to the ventral part of the neural tube. The significance of this is not yet known and this has yet to be seen clearly in tissue section. However, one may speculate that *Zfp-37* is expressed in response to some diffusible factor or factors originating from the notochord or neural floor plate (Yamada *et al.*, 1991; Yamada *et al.*, 1993), one possible candidate for this is the mouse homolog of the Drosophila gene hedgehog (hh) (Echelard *et al.*, 1993; reviewed by Smith, 1994). At e10.5 expression is also seen around the developing fore limb, this expression is extended in the e12.5 embryo to both the fore and hind limb. Until sections have been obtained from the developing limb it is unclear which cell types are expressing the *lacZ* reporter. In general, the expression of the *lacZ* reporter gene appears to become more widespread as the embryo develops. Expression is clearly not restricted to developing neuronal structures.

The staining pattern seen in the three postnatal brain is consistent with the *in-situ* RNA hybridisation data from the adult brain, the targeted reporter gene expression shows the *Zfp-37* promoter to be active in diverse neuronal cell types. Higher levels of *lacZ* expression are evident in particular neuronal structures these include the Islands of Calleja, regions of the hypothalamus, the hippocampus, the superior colliculi, habenular nuclei and purkinje neurons of the cerebellum. The exact role of the Islands of Calleja are not known, though it is thought they may a function as chemical receptors in the neuroendocrine system (Millhouse, 1987). The hypothalamus is involved in controlling hormone release from the anterior pituitary gland (reviewed by Kandel *et al.*, 1991). As discussed in previous chapters, the hippocampus is involved in spatial learning and memory functions. The superior colliculus is involved in coordinating ocularmotor function in response to visual and auditory stimuli (reviewed by Kandel *et al.*, 1991). Defects of the cerebellar are manifest as motor dysfunctions such as ataxia. The function of some of these neuronal processes could be tested in the mice homozygous for the targeted *Zfp-37* mutation. Neuroendocrine function could be examined by measuring various hormone levels in these mice. Defects in the superior

colliculus and cerebellum can affect coordination and result in ataxia. However, the homozygous animals appear to move normally suggesting that no gross aberrations have occurred in these areas, though more subtle effects may be present. Homozygous animals are being tested for the regulation of eye movement in response to changing environmental cues, this should test superior colliculus and cerebellum function. Hippocampal spatial learning tests such as the Morris Water maze test (Morris, 1991) will also be carried out on *Zfp-37*$^{-/-}$ mice. However, these test should be carried out on congenic strains to avoid strain specific effects, a breeding program to establish congenic lines is in progress.

One possible explanation for the widespread neuronal expression pattern of *Zfp-37*, is that *Zfp-37* expression correlates with neuronal activity. This is known to occur for immediate early genes such as *c-fos* and *Krox24* (Morgan and Curran, 1989; Morgan and Curran, 1991). We have previously postulated that *Zfp-37* could be expressed as an immediate early gene by virtue of sequence elements located in the 3'UTR of the gene (Mazarakis *et al.*, 1995). It could therefore be the case that *Zfp-37* expression is the result of some external stimulus or stimuli. This may also be the case for the regulation of *Zfp-37* expression in the embryo. Immediate early genes can be activated by neurotransmitters, for instance activation of the N-methyl-D-aspartate (NMDA) receptor results in increase of *c-fos* expression (Aronin *et al.*, 1991). Addition of kainic acid, the agonist of the kainate glutamate receptor, to transgenic *c-fos-lacZ* mice results in induced expression of the transgene in the hippocampus (Robertson *et al.*, 1995). Furthermore, antipsychotic drugs such as clozapine have been shown to increase expression of *c-fos* and the zinc finger gene *Krox24* in the Islands of Calleja (MacGibbon *et al.*, 1994). The effect of receptor agonists or antipsychotic drugs on *Zfp-37* expression in various neuronal structures could be tested by looking for increased expression of the targeted *lacZ* reporter gene in *Zfp-37*$^{+/-}$ mice after administration of the drug.

Comparison of the *lacZ* staining pattern between the brains of three week old *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice has highlighted some differences. In a number of neuronal structures the reporter gene appears to be expressed at high levels as compared to the rest of the brain, this is observed for the islands of Calleja, superior colliculus and the dentate gyrus. These areas stain strongly in both hetero and homozygous mice.

However, areas which show low or no staining in the *Zfp-37*$^{+/-}$ mice appear to stain strongly in the *Zfp-37*$^{-/-}$ brain. This is particularly evident when comparing staining in the inferior colliculi between the two genotypes. In the *Zfp-37*$^{+/-}$ mice the inferior colliculus stains weakly as compared to the superior colliculus of that animal. However, in the *Zfp-37*$^{-/-}$ mouse the superior and inferior colliculi appear to stain to the same degree (Figure 28c and d). This is a particularly intriguing observation. Similarly, homozygous embryos show increased X-gal staining as compared to the heterozygous embryos. Homozygous *Zfp-37*$^{-/-}$ animals are predicted to show an increase in X-gal staining as they possess two *lacZ* reporter genes, however, as observed by eye the X-gal staining appears to be a more than two fold. This observation could be explained in two ways. Firstly, we are observing a phenomena purely related to the X-gal staining, a threshold level of β-galactosidase activity must be reached before staining becomes visible. Staining is visible in both genotypes in areas of high *lacZ* expression, in areas where the gene is expressed at a lower level staining becomes visible in the homozygous mouse because two *lacZ* reporter genes are being expressed. The second explanation is that the expression of the *Zfp-37/lacZ* fusion allele is deregulated in the mice which are homozygous for the mutation. This would also suggest that the deregulation of the *Zfp-37/lacZ* fusion gene expression is linked to the targeted disruption made. However these observations are by eye, and are not quantitative. The following chapter examines the expression of the *Zfp-37/lacZ* targeted allele in more detail.

Chapter 5

Analysis of the disrupted *Zfp-37* allele

5.1 Introduction

In chapter 3 the introduction of the *nlslacZneo* cassette by homologous recombination into the murine *Zfp-37* gene is described. The cassette was inserted at the ATG codon in exon II. This places the nuclear localised β-galactosidase protein in frame with ZFP-37. Depending on which isoform of ZFP-37 is encoded the β-galactosidase protein will either be in frame with the first 40 amino acids of ZFP-37, if transcription initiates from exon Ia, or coding will begin at the nuclear localised β-galactosidase protein if transcription initiates from exon Ib or II. The inclusion of *nlslacZneo* cassette is predicted to disrupt the normal formation of the ZFP-37 protein and result in a null allele. The organistaion of the disrupted allele has been determined by Southern blot analysis, showing the structure of the targeted allele to be as predicted. However, before any conclusion can be drawn regarding the biological effect of the disruption, it must be confirmed that the targeted disruption of the *Zfp-37* allele does indeed prevent normal formation of ZFP-37 protein. Part of this chapter is concerned with the analysis of *Zfp-37* transcripts and ZFP-37 protein production in mice homozygous for the targeted *Zfp-37* allele.

In the previous chapter we have observed a marked difference in the intensity of X-gal staining in the homozygous mice carrying two copies of the disrupted allele as compared to the heterozygous animals. This is most evident in the three week postnatal brain. As discussed in chapter 4, areas which show no or little X-gal staining in the *Zfp-37*$^{+/-}$ mice appear to stain strongly in the homozygous *Zfp-37*$^{-/-}$ mice. The clearest example of this phenomena is seen in the inferior colliculus. This observation has lead us to speculate that in certain structures of the *Zfp-37*$^{-/-}$ brain, such as the inferior colliculus, expression of the *Zfp-37/lacZ* allele is deregulated. Part of work described in this chapter aims to investigate this phenomena further.

5.2 Results

5.2.1 RNA analysis of *Zfp-37* expression in *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice

To determine whether insertion of the *lacZneo* cassette at the ATG in exon II of

the *Zfp-37* gene results in a null allele analysis of *Zfp-37* mRNA was carried out on tissue samples from adult wild type, heterozygous and homozygous mice. RNA was isolated from testis, brain and liver from six week old litter mates of a *Zfp-37*$^{+/-}$ x *Zfp-37*$^{+/-}$ cross. Northern blots were prepared and probed with a 0.9kb HindIII-EcoRI cDNA probe directed against exon IV of *Zfp-37*, a 0.4 kb EcoRI *Zfp-37* cDNA probe which is directed against exons II, III and IV, a *lacZ* HindIII-EcoRV 1.2 kb probe and a *neo* PstI 0.6 kb probe. A GAPDH specific probe was used as a RNA loading control. The the location of the probes used and results from the Northern blot analysis are shown in Figure 30.

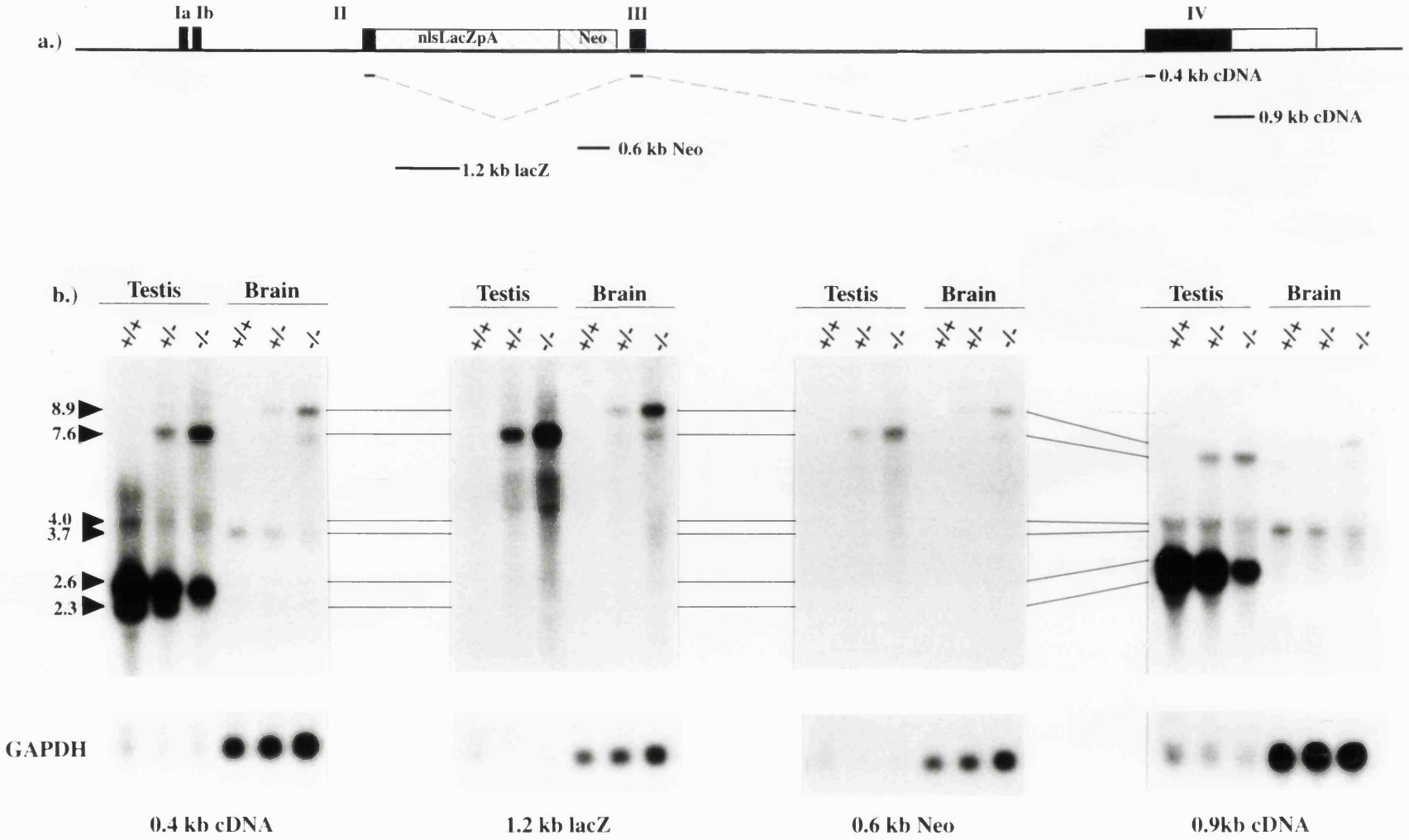
In the wild type tissue samples, the expected transcripts of 2.3, 2.6 and 4.0 kb in the testis and 3.7 kb in the brain are detected with the two *Zfp-37* specific probes. In the heterozygous and homozygous animals larger transcripts are detected. In the testis of heterozygous and homozygous animals a transcript of approximately 7.6 kb is seen whereas in the brain a major 8.9 kb and minor 7.6 kb transcript are seen. These messages also hybridise with both the *lacZ* and *neo* specific probes. This suggests that these mRNAs are fusion transcripts containing *Zfp-37* and *lacZneo* sequences. The fact that two transcripts are seen is likely to be due to the use of the two different *Zfp-37* polyadenylation sites (Mazarakis *et al.*, 1995). It appears that the polyadenylation signal at the end of MC1NeopolyA is ignored or spliced out. Though this was not predicted, these messages will not encode a functional ZFP-37 protein due to the inclusion of the *lacZ* and *neo* open reading frames.

In the homozygous samples from testis and brain probed with the 0.4 kb and 0.9 kb *Zfp-37* cDNA probes, the normal testis transcripts of 2.3 and 2.6 kb are now replaced with one transcript of approximately 2.4 kb, and in the brain, the 3.7 kb transcript appears lower in intensity and to be of lower molecular weight, approximately 3.6 kb (Figure 30). The 2.4 kb and the 3.6 kb transcripts do not hybridise with *lacZ* or *neo* probes. The difference in transcript size and lack of hybridisation with *lacZ* and *neo* probes suggested to us that these transcripts may have resulted from a splicing event around exon II, the exon in which the targeted disruption has been placed. To test this hypothesis a reverse transcriptase (RT) PCR strategy was taken to amplify sequences over exon II. Forward primers were designed which were located in sequences from exon Ia, Ib and exon II, the reverse primer was located in exon IV. The exon Ia and IV

Figure 30. Northern blot analysis of RNA from adult testis and brain of *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$and *Zfp-37*$^{-/-}$ mice.

a) Structure of the targeted *Zfp-37* allele. Probes for hybridisation to Northern blot are shown below.

b) Northern blot analysis of transcripts from the targeted *Zfp-37* allele. Each panel shows RNA samples from testis and brain of *Zfp-37*$^{+/+}$, *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$. The first panel (L to R) shows RNA samples hybridised with the 0.4 kb EcoRI cDNA probe which hybridises to exon II, III and IV sequences. In the wild type (+/+) RNA samples from testis and brain, the 4.0 kb, 2.6 kb and 2.3 kb wild type testis transcripts are shown to hybridise as does the 3.7 kb wild type brain transcript. In both heterozygous (+/-) and homozygous (-/-) samples larger transcripts of 7.6 kb, testis, and 7.6 and 8.9 kb brain hybridise. In the homozygous samples a 2.4 kb testis transcript and a 3.6 kb brain transcript are also found to hybridise with the probe. A similar hybridisation pattern to the first panel is seen when using the 0.9 kb HindIII-EcoRI cDNA probe which hybridises to exon IV sequences only (last panel R). The second panel shows samples as for first panel, hybridised with the 1.2 kb HindIII-EcoRV *lacZ* probe. This shows the 7.6 kb and 8.9 kb transcripts seen in the first panel to contain *lacZ* sequences. Third panel shows samples hybridised with the 0.6 kb PstI *neo* probe. The larger 7.6 kb and 8.9 kb transcripts hybridise with the *neo* specific probe. Lower panels show loading control for each lane. A human glyceraldehyde-3-phospate dehydrogenase (GAPDH) cDNA probe (Benham *et al.*, 1984) was used to determine RNA loading in each sample.


a.)
Ia Ib
II
III
IV
nlsLacZpA
Neo
0.4 kb cDNA
0.9 kb cDNA
0.6 kb Neo
1.2 kb lacZ
b.)
Testis
Brain
+/+
+/-
-/-
8.9
7.6
4.0
3.7
2.6
2.3
GAPDH
0.4 kb cDNA
1.2 kb lacZ
0.6 kb Neo
0.9kb cDNA

primers will amplify a fragment of 430 bp from a normal transcript which includes exon II. Exon Ib and IV will also amplify a similar sized fragment. Loss of exon two results in a smaller amplified fragment of 350 bp. The exon II and IV primers should in the normal *Zfp-37* transcript amplify a fragment of 374 bp. The positions of the primers are shown in Figure 31. The RT-PCR reaction products were visualised on a gel, Southern blotted and hybridised with the 0.4 kb *Zfp-37* cDNA probe (Figure 31). In the *Zfp-37*$^{-/-}$ samples of testis and brain a major PCR product of approximately 350 bp is seen when exon Ia and IV primers are used showing exon II to be deleted. However a minor product of 430 bp is also seen this suggests that a transcript containing exon II sequences is present in the homozygous samples. This is further confirmed by amplification of a 370 bp fragment using the exon II and IV primers.

The 350 bp amplified fragment was cloned and sequenced, confirming that in the major RT-PCR product, exon II has indeed been spliced out leading to a frame shift with the downstream exons. We conclude that this transcript will not encode a functional ZFP-37 protein. However, the minor 430 bp product amplified appears to contain exon Ia, II and III sequences from the PCR analysis. Unfortunately we have been unable to clone or sequence this product and have therefore have been unable to establish whether this can encode a functional ZFP-37 protein.

The analysis of the mRNA products from the disrupted allele have shown the large majority of the *Zfp-37* transcripts produced from the targeted allele are unable to produce a functional ZFP-37 protein. To establish whether this was the case Western blot analysis was carried out to determine if any ZFP-37 protein could be detected.

5.2.2 Western blot analysis of ZFP37 protein in *Zfp-37*$^{-/-}$ animals

Concurrent with the Northern blot analysis we have investigated the presence of ZFP-37 protein isoforms on Western blot, using nuclear extracts from *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ brain and testis samples. Polyclonal rabbit antisera had been raised against recombinant full length ZFP-37 (antiserum 1341) and also a truncated form containing the N-terminal 135 amino acids (antiserum 1343) (N.Galjart unpublished results). This has previously been shown to detect ZFP37 on Western blots of total protein extracts from transfected COS-1 cells (N.Galjart unpublished). Nuclear extracts from testis,

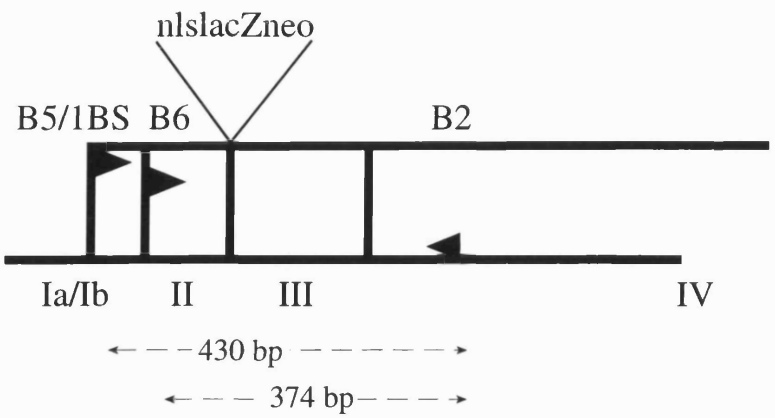
Figure 31. RT-PCR analysis of *Zfp-37* transcripts from *Zfp-37*$^{-/-}$ mice.

a) Diagrammatic representation of 5' end of *Zfp-37* Ia or Ib transcripts. Single line denotes 5'UTR sequence. Regions of transcript encoded by exons Ia/Ib, II, III and IV are shown by vertical divisions. Location of primers are shown by arrow heads. Forward primers are exon Ia (B5), exon Ib (1BS) and exon II (B6). The reverse primer is located in exon IV(B2). Sequence of primers used is given in Chapter 7 Material and methods. Below is shown the predicted PCR product size obtained using the various primer combinations. Using primer B5 or 1BS in conjunction with B2 will amplify a product of 430 bp. B6 used with B2 will amplify a product of 374 bp. For reference the disrupting nlslacZneo cassette is placed at the end of exon II.

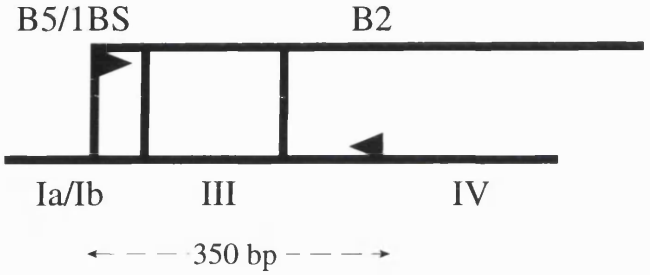
b) Organisation of 5' end of aberrant *Zfp-37* transcript which does not posses exon II sequence. The predicted size of the PCR product using primers B5/1BS and B2 if exon II is removed is 350 bp, show below.

c) RT-PCR products from brain and testis wild type (+/+) and homozygous for the *Zfp-37* disruption (-/-). Primer combinations are shown above each lane: exIb primers 1BS and B2; exIa primers B5 and B2; exII primers B6 and B2.

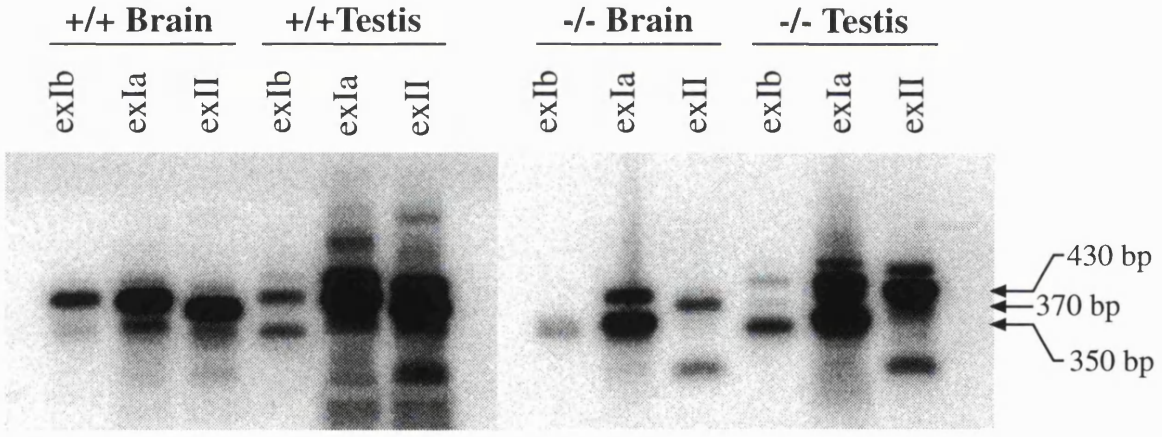
a.)


nlslacZneo
B5/1BS
B6
B2
Ia/Ib
II
III
IV
430 bp
374 bp

b.)


B5/1BS
B2
Ia/Ib
III
IV
350 bp

c.)


+/+ Brain
+/+Testis
-/- Brain
-/- Testis
exIb
exIa
exII
exIb
exIa
exII
exIb
exIa
exII
exIb
exIa
exII
430 bp
370 bp
350 bp

cerebellum and cerebrum of wild type, heterozygous and homozygous mice were tested by Western blot for the presence of ZFP-37 protein. Extracts from transfected COS-1 cells were also used as a control. Figure 32 shows the western blot analysis from cerebrum and cerebellum. The long form of ZFP-37 is approximately 70 kDA on a Western blot. A ubiquitous band appears in all tissue samples at approximately 35 KDa, this band corresponds to a highly abundant protein visible when the filters were stained for total protein. We conclude that this band does not originate from ZFP-37. A protein of approximately 70 KDa was detected in both the wild type and heterozygous samples using both antisera. On closer inspection multiple bands are detected which may reflect postranslational modifications of the ZFP-37 protein. The amount of the 70 KDa ZFP-37 protein detected by the antibodies in the heterozygous samples appears lower than the wild type and it is not detected in samples from the homozygous animals. However, it should be noted that not all samples were loaded equally. If ZFP-37 protein is produced in the mice homozygous it is below the detection level of this assay. From these data it appears that the introduced mutation severely disrupts the production of ZFP-37 protein in mice homozygous for the disrupted allele.

5.2.3 Allelic expression of the *lacZ* reporter gene in brains of *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice.

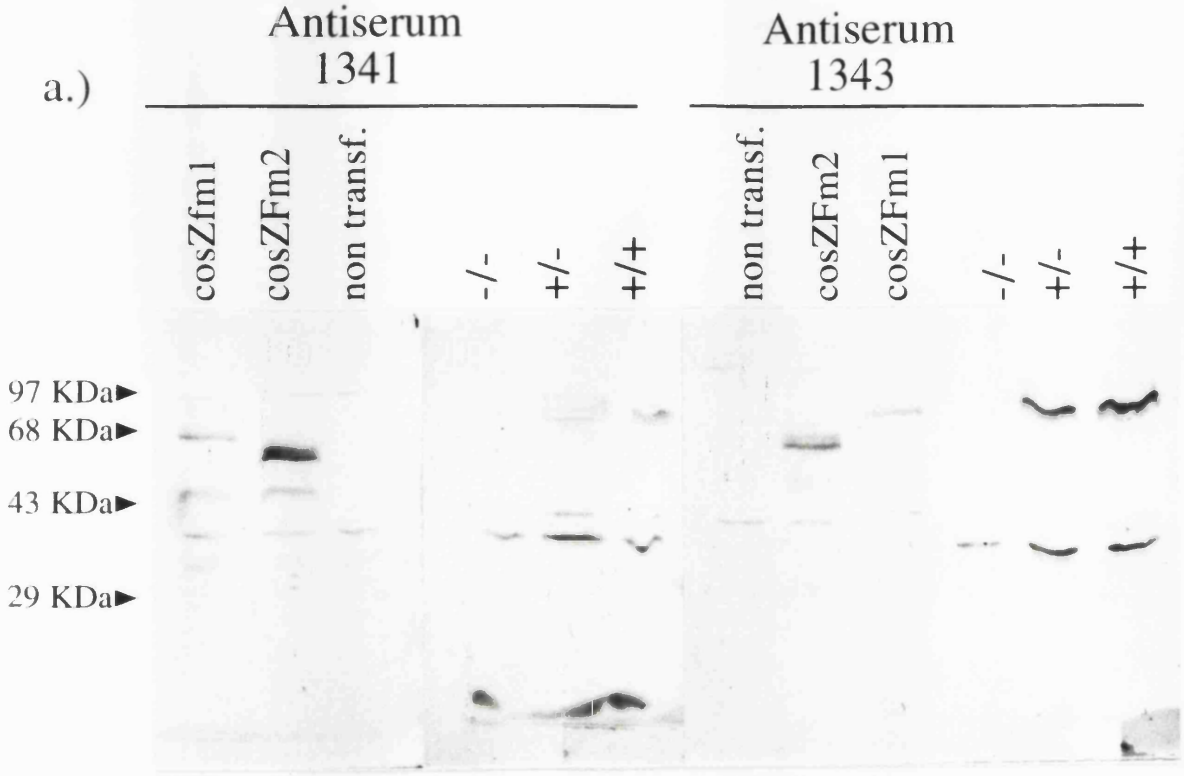
Based upon the observed difference in X-gal staining between the brains of *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice described in chapter 4, the levels of RNA expression from the *Zfp-37/lacZ* alleles in the heterozygous and homozygous mice were examined in more detail. Using the Northern blot shown in Figure 30 we quantitated the intensities of the 7.6 and 8.9 kb *Zfp-37/lacZ* fusion transcripts present in the brain This was performed using the phosphorimage of the Northern blot and analysed using Imagequant version 3.3 software. Each transcript was normalised for GAPDH expression and number of *lacZ* alleles. Expression per *Zfp-37/lacZ* allele of the 7.6 kb and 8.9 kb transcripts was calculated using the 0.4 kb *Zfp-37* cDNA probe, the 0.9 kb *Zfp-37* cDNA probe, the *lacZ* probe and the *neo* probe. The results obtained are shown in Table VI. The 8.9 kb *Zfp-37/lacZ* fusion transcript shows an increase in expression per allele in the homozygous brain samples with each probe tested. The 7.6 kb transcript also shows an

Figure 32. Western blot analysis of ZFP-37 protein

a) Cerebrum nuclear extracts from *Zfp-37*$^{+/+}$, *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice. Antiserum 1341 is directed against full length recombinant ZFP-37, Antiserum 1343 is directed against the first 135 amino acids of recombinant ZFP-37. COS-1 cell extracts are used as control. cosZfm1 extracts are from COS-1 cells transfected with full length *Zfp-37* 3.7 kb cDNA. cosZFm2 extracts are from COS-1 cells transfected with *Zfp-37* cDNA using ATG located in exon II. Non-transfected COS-1 control is included. Cerebrum nuclear extracts from -/-, +/- and +/+ mice are shown for both antisera. No ZFP-37 protein can be detected in -/- samples. A 70 KDa band can be detected in both heterozygous (+/-) and wild type samples.

b). As above but using Cerebellum nuclear extracts. Again no ZFP-37 70 Kda protein can be seen in the homozygous sample (-/-). ZFP-37 protein is detected in heterozygous and wild type nuclear extracts.

a.)


Antiserum 1341
Antiserum 1343
cosZfm1
cosZFm2
non transf.
-/-
+/-
+/+
non transf.
cosZFm2
cosZFm1
-/-
+/-
+/+
97 KDa►
68 KDa►
43 KDa►
29 KDa►

b.)

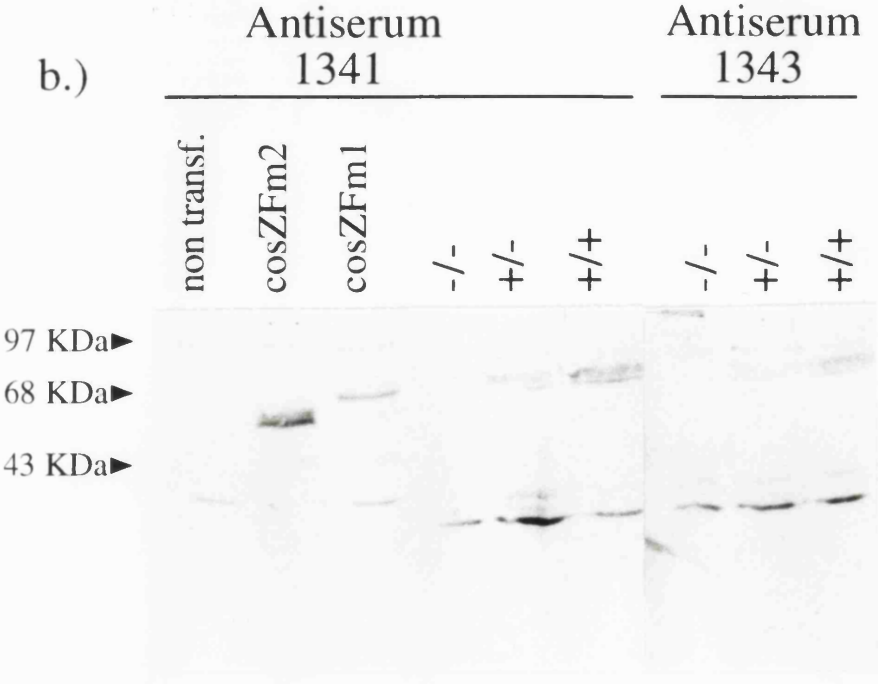

Antiserum 1341
Antiserum 1343
non transf.
cosZFm2
cosZFm1
-/-
+/-
+/+
-/-
+/-
+/+
97 KDa►
68 KDa►
43 KDa►

Table VI. Quantification of *Zfp-37/lacZ* fusion transcript in brain from *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice

Probe	Transcr.	ImageQuant values		Expression per *lacZ* allele*	
		+/-[a]	-/-[b]	+/-	-/-
0.9	**8.9 kb**	10432	28604	1	1.35
	7.6 kb	5013	10994	1	1.04
GAPDH[1]	-	1561544	1638462	-	-
0.4	**8.9**	12660	41805	1	1.16
	7.6	6039	15612	1	0.91
lacZ	**8.9**	28000	102000	1	1.28
	7.6	13000	55000	1	1.49
Neo	**8.9**	10000	30000	1	1.06
	7.6	4000	20000	1	1.76
GAPDH[2]	-	1014019	1434354	-	-

ImageQuant values obtained using ImageQuant version 3.3 software selecting "integrate volume" quantification option.
GAPDH[1] values are used to normalise results obtained from the 0.9 kb *Zfp-37* probe.
GAPDH[2] values are used to normalise results obtained with the 0.4 kb *Zfp-37* probe, *lacZ* probe and *neo* probe. In this case two measurements of the GAPDH signals were made.
*Expression per *lacZ* allele calculated as follows: For -/- samples (a/GAPDH)/(a/GAPDH). For +/- samples (b/GAPDH)/(a/GAPDH).

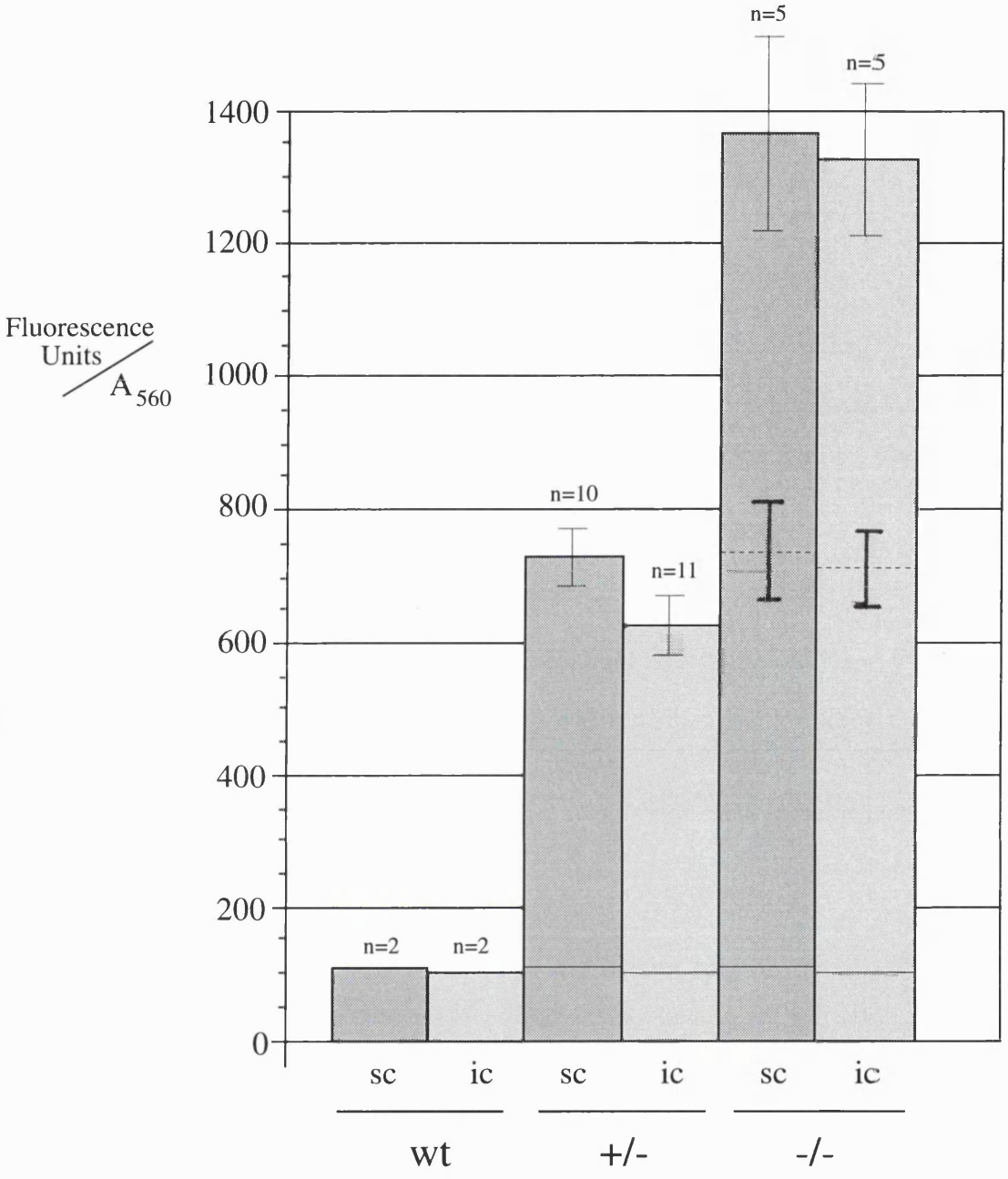
increase in expression per allele in the homozygous samples with the exception of the 0.4 kb probe. From this we conclude that the expression of the *Zfp-37/lacZ* fusion transcipt per allele is increased in the homozygous brain as compared to the heterozygous. This infers that the expression of the *Zfp-37/lacZ* allele is deregulated ie either *lacZ* expression is increased in the *Zfp-37*$^{+/-}$ brain compared to *Zfp-37*$^{+/-}$ or expression in the *Zfp-37*$^{+/-}$ brain is repressed.

From the above RNA analysis and our original observation upon the X-gal staining in the postnatal brain we attempted to address this phenomena further by quantitating the β-galactosidase enzyme activity in samples from the heterozygous and homozygous brains. This was carried out using the fluorgenic β-galactosidase substrate 4-methylumbelliferyl-β-D-galactoside (MUG). We chose to focus on the superior and inferior colliculi. As mentioned above the β-galactosidase staining in the superior colliculus (SC) of three week old *Zfp-37*$^{+/-}$ mice appears to be more intense than staining in the inferior colliculus (IC). However, in the *Zfp-37*$^{-/-}$ mice this staining appears to be of equivalent intensity. The colliculi are also simple to remove as they are defined surface structures. The original observation was made in animals of three weeks of age, at this age mice are still developing neurologically and the degree of development may vary between individual animals, though this observation is consistent for other three week old animals and is independent of the background strain (data not shown). Six week old animals were therefore chosen to provide the colliculi as their nervous system should be fully developed.

Superior and inferior colliculi were isolated by dissection from a total of nineteen six week old litter mates (2 x *Zfp-37*$^{+/+}$, 12 x *Zfp-37*$^{+/-}$ and 5 x *Zfp-37*$^{-/-}$) (129 x C57BL/6 x FVB background). The enzyme activity was determined by the formation of the reaction product 4-methylumbelliferone, which can be measured by its fluorescence at 480nm (after excitation at 355nm). The values obtained for the 19 SC and IC analysed were normalised for total protein in each sample. The results are shown in Figure 33. The mean values obtained from the colliculi of *Zfp-37*$^{-/-}$ mice are 1366 (standard deviation (S.D.) 146) and 1326 (S.D. 112) fluorescence units/mg protein for the SC and IC respectively. These values are not significantly different, as determined by the student t statistical test ($p > 0.05$), suggesting that the β-galactosidase enzyme is expressed to equivalent levels in the SC and IC in *Zfp-37*$^{-/-}$ mice, as observed by eye.

Figure 33. Histogram of β-galactosidase activity in inferior and superior colliculi from *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice.

Histogram depicting β-galactosidase activity in the superior (sc) and inferior (ic) colliculi of *Zfp-37*^{wt}, *Zfp-37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice. β-galactosidase activity is measured as fluorescence of the 4-methylbelliferyl-β-D-galactoside product 4-methylumbelliferone at 480 nm per protein absorbance at A_{560} using the BCA colometric assay (Smith *et al.*, 1985). Sample numbers are shown above each column. One sc +/- sample was lost. Standard deviation bars are shown for each column. The level of endogenous β-galactosidase activity is shown as a single line in the +/- and -/- columns as determined from the wt columns. Dashed lines in the -/- column represent the predicted activity per allele calculated as (mean (-/-) value-mean (wt) value)/2. Note the mean β-galactosidase activity in the (+/-) ic is below the predicted value, however it can not be concluded that this is significant


Fluorescence
Units
A 560
1400
1200
1000
800
600
400
200
0
n=5
n=5
n=10
n=11
n=2
n=2
sc
ic
sc
ic
sc
ic
wt
+/-
-/-

In *Zfp-37*$^{+/-}$ mice, the mean value obtained for the SC is 732 (S.D. 44) fluorescence units/mg protein, and for the IC the it is 624 (S.D. 45) fluorescence units/mg. Using the student t test the level of β-galactosidase activity in the IC is significantly lower than observed in the SC ($p< 0.01$), this observation is consistent for each individual mouse. Comparing the mean enzyme activity (subtracting the mean endogenous β-galactosidase activity found in the wild type mice) in the superior colliculi of *Zfp-37*$^{+/-}$ mice with that of the *Zfp-37*$^{-/-}$ SC shows β-galactosidase activity in the SC of *Zfp-37*$^{+/-}$ mice to be 49% of the expression found in the SC of *Zfp-37*$^{-/-}$ mice. This is the expected result if the *lacZ* reporter genes is expressed as per gene copy. In the IC of *Zfp-37*$^{+/-}$ mice enzyme activity is 42% of the activity seen in the IC of *Zfp-37*$^{-/-}$ mice. This would suggest that some form of repression of β-galactosidase activity is occurring in the IC of *Zfp-37*$^{+/-}$ mice but not in the SC. However, it is important to note that the values obtained were from one experiment only. This study needs to be repeated with a large sample group before any conclusions can be drawn with validity.

5.3 Discussion

In this chapter the expression products from the targeted *Zfp-37* allele are investigated. This was carried out with the aim of determining whether the targeted disruption has inactivated the *Zfp-37* allele and prevents production of the ZFP-37 protein. Furthermore, we have observed a difference in the X-gal staining in the postnatal heterozygous and homozygous brain which we hypothesised was due to a deregulation of the reporter gene expression and that this deregulation is dependent on the abscence or reduced level of ZFP-37. We have provided further evidence that would suggest this may be the case.

5.3.1 *Zfp-37* expression from the targeted allele

We have investigated whether the placing the nlslacZneo cassette at the ATG in exon II has lead to disruption of the normal ZFP-37 protein production. From the analysis of the RNA transcripts from the targeted allele we see two unexpected results. Firstly, the nlslacZneo cassette is expressed as a fusion transcript with the downstream

exons of *Zfp-37*. It appears that the polyadenylation signals within the disruption are ignored or are spliced around. However, this transcript is not predicted to encode a functional form of ZFP-37 due to the *lacZ* and *neo* open reading frames that will be encountered. Transcripts which are closer in molecular weight to the wild type *Zfp-37* transcripts are detected in the testis and brain of the homozygous mice, namely the 2.4 kb testis and 3.6 kb brain transcripts. We have shown by RT-PCR analysis and sequencing that the majority of these aberrant transcripts are the result of a splicing event which excludes exon II, the exon in which the disrupting cassette is placed. Removal of exon II results in a frame shift of the upstream exons with exons III and IV, these transcripts are therefore deemed to be nonfunctional. However by RT-PCR we also detect a product which contains exon II sequences and is approximately the same molecular weight as the wild type reaction product. We have been unable to sequence this product so far and are therefore unable to rule out the possibility that this may represent a functional *Zfp-37* transcript. It is unclear how such a transcript could arise although one possibility may be the use of cyptic splice donor sites within the 5' end of the nlslacZneo cassette. However, until we can sequence this product this idea remains purely speculative.

At the protein level, we have attempted to detect ZFP-37 protein using two polyclonal antisera raised against the recombinant full length ZFP-37 protein and the first 135 amino acids respectively. With both antisera, ZFP-37 protein can be detected in nuclear extracts from the wild type and heterozygous mice. We are unable to detect ZFP-37 protein with these antisera in the homozygous *Zfp-37*$^{-/-}$ samples. This shows that the targeted disruption in the *Zfp-37* allele has servely reduced the amount of ZFP-37 protein produced to below our ability to detect it. However, from the RNA analysis we can not yet rule out the possibility that a functional ZFP-37 protein is produced albeit at a very low level.

4.3.3 Regulation of *Zfp-37* expression

An observed difference in the *lacZ* staining pattern between the brains of three week old *Zfp37*$^{+/-}$ and *Zfp-37*$^{-/-}$ mice was seen in the inferior colliculus. It appeared that in the IC of *Zfp37*$^{+/-}$ mice, β-galactosidase activity was more than two fold lower than

the activity in the IC of *Zfp-37*$^{-/-}$ mice. This suggested to us that the *lacZ* gene was not being expressed as per gene copy in the IC of the *Zfp37*$^{+/-}$ mice, but was expressed per gene copy number in the IC of *Zfp-37*$^{-/-}$ mice. This would suggest that the *Zfp-37/lacZ* expression was being repressed in some manner and that this repression was alleviated when both copies of the *Zfp-37* allele were disrupted. This lead us to speculate that *Zfp-37* expression is controlled in an auto regulatory fashion and that the *Zfp-37* gene product plays a role in this regulatory mechanism. Further to this observation we have analysed RNA levels of the 7.6 kb and 8.9 kb *Zfp-37/lacZ* fusion transcripts in adult brain from mice heterozygous and homozygous for the targted disruption (Table VI). Four probes were tested, with exception of one transcript, we see an increase in allelic expression in the homozygous sample ranging from 1.04 to 1.76. This result can either be viewed as a general increase in expression per allele in the homozygous animal compared to the heterozygous, suggesting an up regulation of expression, or that expression is repressed in the heterozygous brain.

To further investigate the above observations the β-galactosidase activity in colliculi from six week old animals was measured using the β-galactosidase fluorogenic substrate MUG. In the SC *lacZ* expression as determined by the activity of β-galactosidase appears to follow the gene copy number, however in the IC, β-galactosidase activity from one *Zfp-37/lacZ* allele is 42% of the expression from two *Zfp-37/lacZ* alleles. It is attractive to conclude from this result that repression of *lacZ* expression occurs in the IC of heterozygous mice due to the prescence of ZFP-37. However, this experiment should be repeated using a larger sample population.

Assuming repression of the *Zfp-37/lacZ* allele is occurring, then one possibile explanation is that ZFP-37 could function as a transcriptional repressor, either through direct interaction with a specific DNA sequence or by affecting chromatin structure in some manner. Disruption of both copies of the *Zfp-37* gene alleviates this repression and hence increases allelic expression in the homozygous mouse. In direct evidence supporting this notion comes from *in-vitro* binding of DNA by *Zfp-37* (N.Galjart unpublished results). Furthermore, the major isoform of ZFP-37 contains a variant of the KRAB-A domain which is postulated to act as a transcriptional repressor domain (Margolin *et al.*, 1994; Pengue *et al.*, 1994; Witzgall *et al.*, 1994; Deuschle *et al.*, 1995). It has been suggested that the repressing activity of the KRAB-A domain is achieved

by interaction with the basal transcriptional machinery either by direct interaction or via some co-factor (Deuschle *et al.*, 1995). This repression could therefore be determined by the ZFP-37 isoform produced in a particular cell type, and this could theoretically account for the different picture observed for the SC. It would be interesting to determine which ZFP-37 isoform is expressed in the SC and IC and if there is a difference in the level of *Zfp-37* RNA expression in these structures. If disruption of *Zfp-37* does lead to a deregulation of expression from its own allele and ZFP-37 is involved in this regulation, then one would predict that expressing *Zfp-37* as a transgene in the disrupted background would normalise the expression of the *Zfp-37/lacZ* allele. This experiment is being carried out at present. The above observations although not conclusive have certainly highlighted an interesting phenomenon which deserves further investigation.

Chapter 6

General Discussion

6.1 Introduction

The general theme of this thesis has been the application of gene targeting technology in the mouse to study gene function in the murine nervous system. To this end the murine *Thy-1* and *Zfp-37* genes, both expressed in the nervous system were targeted. These two genes encode disparate classes of proteins which are known to function in the nervous system. Thy-1 is a cell surface glycoprotein and belongs to the immunoglobulin supergene family and as such it is postulated to act as a recognition molecule, possibly playing a role in cell adhesion and/or as a receptor (Williams and Gagnon, 1982; Morris, 1992). The murine *Zfp-37* gene by virtue of sequence can encode a protein with 12 *Krüppel*-like zinc-finger domains at its C-terminus. This classes *Zfp-37* in to the large family of zinc-finger genes, members of which are suggested to function as transcriptional regulators in nervous system development. In this chapter the results from the targeting experiments are discussed.

6.2 Targeting of Thy-1

Thy-1 was first described over thirty years ago, only now is its function beginning to be unravelled. The detailed analysis of Thy-1 expression and structure coupled with *in-vitro* experiments aimed at describing its function have lead to the present theories upon its function. The evidence points in the direction of Thy-1 playing a part in controlling axonal growth during the end of neurogenesis. To directly address the role of Thy-1 a gene targeting experiment in mouse ES cells was attempted with the aim of producing mice devoid of the Thy-1 protein. Chapter 2 describe attempts to target the *Thy-1* gene in mouse embryonic stem cells. To this end I was successful in isolating ES cells which had undergone homologous recombination at one *Thy-1* allele. However, upon blastocyst injection, the targeted ES cells were inefficient in contributing to the development of chimeric mice. Only one contribution chimera was obtained. This animal failed to transmit the mutated *Thy-1* allele through the germline. This failure of the ES cell clone to contribute to the germline may possibly have been be due to chromosomal abnormalities, which had been seen in 20% of the parent culture, or to defects acquired during the targeting experiment. The parent ES cell line was subsequently shown to be able to give germline transmission in this laboratory (Pandolfi

et al., 1995; and chapter 3 of this study). As only one targeted *Thy-1* ES cell clone was obtained in the first experiment a second targeting vector was designed and made. This vector took into account the recent finding that isogenic DNA improved targeting efficiency (te Riele *et al.*, 1992). However, this vector was never tested as mice carrying a targeted *Thy-1* allele had been produced in another laboratory. The *Thy-1* targeted mice were available to our collaborators and therefore further attempts to target the *Thy-1* gene were discontinued.

Emerging data from the study of the nervous system of Thy-$1^{-/-}$ mice has revealed some subtle defects. In culture, neurons from Thy-$1^{-/-}$ mice display abnormal growth properties. This is concurrent with the *in-vitro* inhibition of neurite outgrowth by Thy-1 (Tiveron *et al.*, 1992 and 1994). Thy-$1^{-/-}$ mice behave normally, but do show depressed long term potentiation in one synaptic relay of the hippocampus. An analogous phenotype is described for mice with a targeted disruption in the cellular form of the prion protein PrP^c mice (Collinge *et al.*, 1994). By virtue of the similar phenotypes and shared structural properties Thy-1 and PrP^c may share a common function. This hypothesis can be examined by crossing mice homozygous for both gene disruptions. As a member of an array of cell surface recognition molecules, it is certainly feasible that Thy-1 shares an overlapping function with one of these molecules and it may only be in the areas where expression and or glycosylation do not overlap that a phenotype is seen. The advent of mice carrying the targeted *Thy-1* alleles and establishment of a phenotype now provides an *in-vivo* model system in which structural attributes of the Thy-1 protein can be tested. Further study of these mice particularly in the LTP suppressed dentate gyrus should yield to a greater knowledge of Thy-1 and add to the theories upon the function of cell surface adhesion molecules in the nervous system.

6.3 Targeting of the murine *Zfp-37* gene

The murine *Zfp-37* gene is a member of the large family of KRAB domain zinc finger genes. The expression pattern and structure of *Zfp-37* has suggested to us that ZFP-37 can function as a transcriptional regulator in the developing and adult nervous system (Mazarakis *et al.*, 1995). To address the role of ZFP-37 in the nervous system

a gene targeting approach was taken. Chapter 3 describes the successful introduction of a *lacZ* reporter gene into the *Zfp-37* allele by homologous recombination and establishment of mice carrying this disrupted allele. The majority of mice homozygous for the targeted disruption, bred on a 129 x C57BL/6 x FVB outbred genetic background, appear healthy, fertile and are represented close to the predicted Mendelian ratio. In the 129 x C57BL/6 genetic background, three from the six homozygous animals have died, however mortality has also been seen in the other genotypes in the F_1 generation. It is yet to be established if these deaths were linked to the targeted mutation or are the manifestation of another unknown mutation carried by the ES clone used. Further breeding on this background should determine this. It has also to be established whether homozygous animals have subtle neurological defects, parameters such as spatial learning will be tested in these mice. In relation to the biological affect of the targeted mutation we have attempted to show whether a functional *Zfp-37* protein can be produced from the targeted allele. Analysis of RNA transcripts from the targeted allele predict the majority of these to be unable to encode a functional ZFP-37 protein, either due to inclusion of *lacZ* and *neo* open reading frames or due to a frameshift which results from the splicing out of *Zfp-37* exon II. We have, however, detected one transcript by RT-PCR analysis, which appears to contain exon II sequence as determined by Southern blotting. We have been unable to determine the composition of this PCR product and can not rule out the possibility that it may represent a functional *Zfp-37* transcript. From protein analysis on Western blot, we have been unable to detect ZFP-37 protein in homozygous samples using two antisera raised against recombinant ZFP-37 protein. We conclude that we have severely disrupted production of ZFP-37 protein but can not conclude that we have made a full null allele because of the existence of the RT-PCR product. This raises the possibility that a low level of functional ZFP-37 protein is present in the homozygous animal which could potentially alleviate a more severe phenotype.

Examining the expression pattern of *Zfp-37* via the targeted *lacZ* reporter gene has shown expression to extend into earlier embryonic time points than had previously been described. It is also clear from the embryonic expression pattern that *Zfp-37* expression is not restricted to neuronal tissues, though the developing nervous system does appear to be one of the major sites of expression. In the postnatal three week brain

expression was found to be consistent with the previous *in-situ* RNA hybridisation data (Mazarakis *et al.*, 1995). Expression is seen to be pan neuronal, although particular neuronal structures such as the hippocampus, hypothalamus, cerebellum and olfactory tuberculum show areas of higher expression as determined by the X-gal staining pattern. From the widespread neuronal expression pattern I have suggested that *Zfp-37* may be expressed as an immediate early gene. The *Zfp-37* gene sharing a number of sequence features with other immediate early expressed genes (Mazarakis *et al.*, 1995). Expression of *Zfp-37* may therefore correlate with a particular neuronal activation signal possibly provided by a particular neurotransmitter(s). Immediate early genes are known to be expressed in response to such stimuli (reviewed by Morgan and Curran, 1991). The targeted *lacZ* reporter gene is a useful *in-vivo* marker for *Zfp-37* expression and will therefore provide a means to test this hypothesis.

The effect of the targeted mutation on the neuronal function has yet to be established. The high level of the targeted reporter gene expression in the superior colliculus and cerebellum suggests that control of movement may be perturbed in the *Zfp-37*$^{-/-}$ mice. Though no gross defects are apparent upon superficial observation, homozygous mice are being tested for parameters such as control of eye movement. Expression in the hippocampus has suggested that the homozygous mice may have deficits in spatial learning, this has been found for the targeted disruption of isoforms of the cAMP-responsive element binding protein (CREB) (Bourtchuladze *et al.*, 1994), showing this immediate early response transcription factor to be required in long term potentiation. Spatial learning tests will be performed on *Zfp-37*$^{-/-}$ mice to test this function.

From the results obtained from the targeted reporter gene in chapter 4, we have observed a difference in the level of X-gal staining between neuronal regions of the heterozygous and homozygous three week postnatal brain. It appears in certain neuronal structures such as the superior colliculus that the reporter gene is expressed highly in both heterozygous and homozygous brains. In other structures such as the inferior colliculus staining dramatically increases in the homozygous sample. These observations are made by eye and are by no means quantitative, but suggested to us that regulation of the *lacZ* reporter gene may be affected in some manner in the homozygous mouse, this being due to the absence (or severe loss) of ZFP-37 protein. To further examine this

observation we analysed the levels of *lacZ/Zfp-37* fusion transcripts in hetero and homozygous mice. There appears to be an increase in expression of the *lacZ* containing transcript per targeted allele in the homozygous sample. This would suggest that any deregulation is either acting at a transcriptional or posttranscriptional level. We further examined the β-galactosidase activity in the superior and inferior colliculi and compared it to copy number of the *lacZ* allele. Although these are results from a pilot experiment, we have some evidence which suggests that the β-galactosidase activity does not follow gene copy in the inferior colliculus of the heterozygous mouse. Suggesting to us that repression of *lacZ* expression in the heterozygous animal is occurring and that this is linked to the presence of ZFP-37. Taking these three observations together, we speculate that ZFP-37 is part of an autoregulatory loop which serves to finally modulates its own expression at either a transcriptional or postranscriptional level.

Pertinent to this theory are the results that show the KRAB-A domain, which is present in the major isoform of ZFP-37, to act as a transcriptional repressor in cultured cells (Deuschle *et al.*, 1995; Margolin *et al.*, 1994; Witzgall *et al.*, 1994; Pengue *et al.*, 1994). Furthermore we have data which shows ZFP-37 to be able to bind DNA, although we have been unable to identify a specific DNA recognition sequence (Dr. N.Galajart unpublished results). It could therefore be the case that the major isoform of ZFP-37 could act as a transcriptional repressor, possibly negatively regulating its own transcription. In the homozygous disruption this repression would be removed and hence expression of the targeted reporter gene would increase. This would only be the case in cells where the Ia repressing isoform of ZFP-37 was encoded. This could account for the different expression seen in the superior and inferior colliculus. Figure 34 shows a model describing this regulation. Negative autoregulation by an immediate early transcription factor has been shown for the cAMP response element modulator (CREM) gene (Molina *et al.*, 1993; Stehle *et al.*, 1993; reviewed by Sassone-Corsi 1994). The CREM gene can encode, through the use of an alternate intron promoter, a repressing factor, inducible cAMP early repressor (ICER) which can function to repress its own transcription (Molina *et al.*, 1993). ZFP-37 may therefore control its own expression in an analogous fashion.

The above model for ZFP-37 function and regulation is highly speculative.

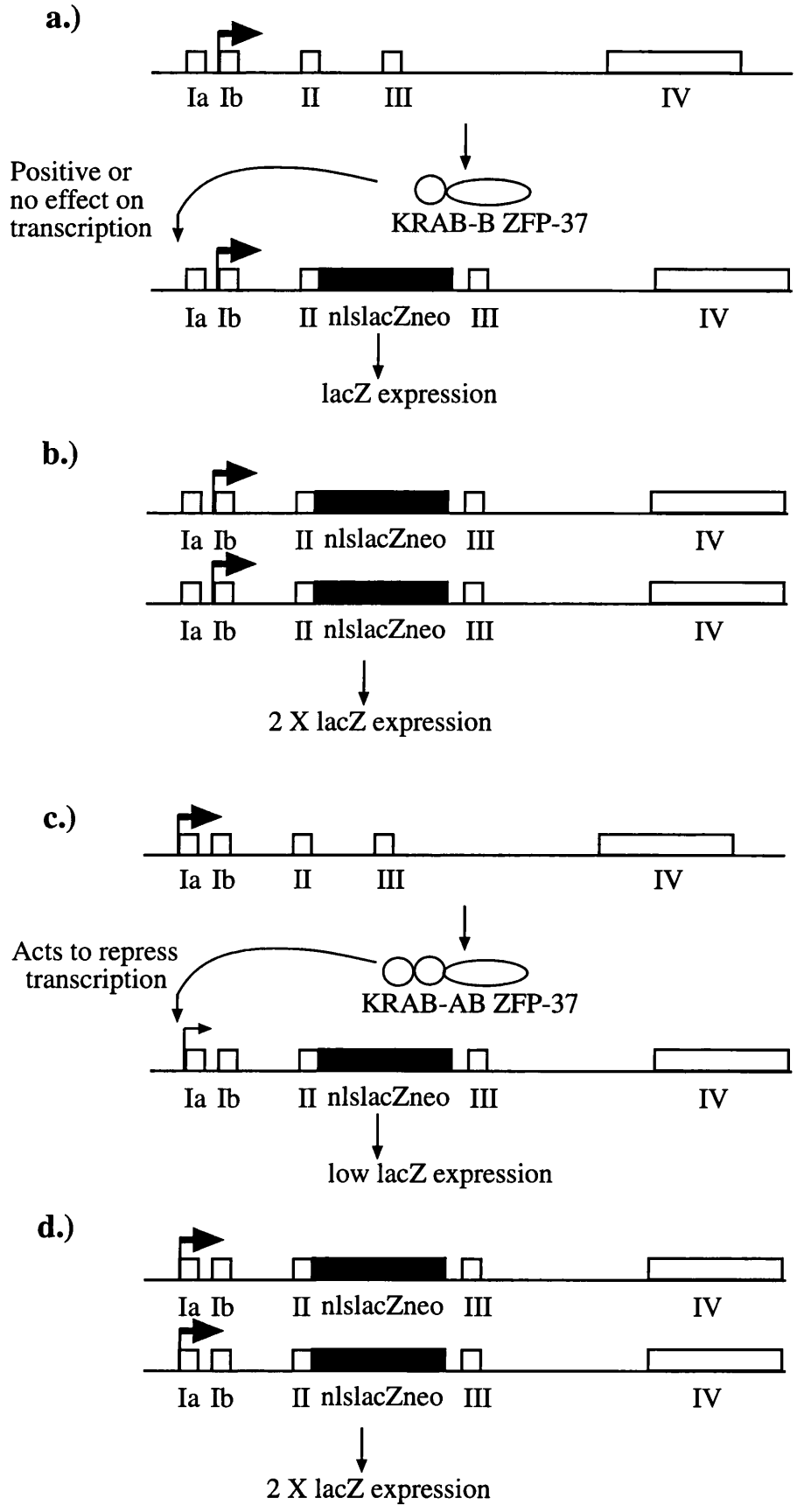
Figure 34. Model of ZFP-37 transcriptional auto regulation in the inferior and superior colliculi.

a.) Expression of *Zfp-37* and *Zfp-37* targeted allele in the superior colliculus of *Zfp-37*$^{+/-}$ mice. In this scenario transcription initiates from the wild type allele at exon Ib producing the KRAB-B 62 KDa isoform of ZFP-37. This has a positive or zero effect on the transcription of the *Zfp-37* allele and the *lacZ* reporter gene is expressed.

b.) Expression of the two targeted *Zfp-37* alleles in the superior colliculus of *Zfp-37*$^{-/-}$ mice. This results in two fold *lacZ* expression as compared to the *Zfp-37*$^{+/-}$ superior colliculus.

c.) Expression of *Zfp-37* and *Zfp-37* targeted allele in the inferior colliculus of *Zfp-37*$^{+/-}$ mice. In this case transcription initiates from exon Ia and the wildtype allele expresses the KRAB-AB 67 KDa isoform. This isoform can repress its own transcription resulting in a lower level of *lacZ* reporter gene expression from the targeted allele.

d.) Expression of the two targeted *Zfp-37* alleles in the inferior colliculi of *Zfp-37*$^{-/-}$ mice. The KRAB-AB ZFP-37 isoform is no longer produced (or at a greatly reduced level) and the reporter gene is expressed as per gene copy, greater than twice that of the *Zfp-37*$^{+/-}$ inferior colliculus.


a.)
Ia Ib
II
III
IV
Positive or
no effect on
transcription
KRAB-B ZFP-37
Ia Ib
II nlslacZneo
III
IV
lacZ expression
b.)
Ia Ib
II nlslacZneo
III
IV
Ia Ib
II nlslacZneo
III
IV
2 X lacZ expression
c.)
Ia Ib
II
III
IV
Acts to repress
transcription
KRAB-AB ZFP-37
Ia Ib
II nlslacZneo
III
IV
low lacZ expression
d.)
Ia Ib
II nlslacZneo
III
IV
Ia Ib
II nlslacZneo
III
IV
2 X lacZ expression

However, targeting of the murine *Zfp-37* gene has provided an interesting observation which warrants further investigation. In this context the targeted reporter gene provides a useful tool in which *Zfp-37* expression can be studied *in-vivo*. Mice homozygous for the targeted allele also provide a background in which the functional attributes of ZFP-37 can be tested as transgenes. If our observations hold true we will have provided *in-vivo* evidence that a KRAB domain protein can act as a transcriptional repressor protein.

Chapter 7

Materials and Methods

7.1 Autoradiography

Nylon filters from hybridisation were enclosed in heat sealed polythene bags or "saran wrap". Dried polyacrylamide sequencing gels were enclosed in saran wrap. Filters or gels were exposed to X-ray film, either Kodak "Xomat-AR" or Fuji "Nif-RX", the former being used when greater sensitivity was required. Samples and film were placed in light proof cassettes, with phosphor-intensifying screens where required, and placed at -70°C. Films were developed using a Fuji developer.

Phosphorimage analysis was carried out using Molecular Dynamics phosphorscreens and analysed using a Molecular Dynamics 425S phosphorimager and ImageQuant software version 3.3.

7.2 Bacterial cultures

Liquid bacterial cultures were grown in Luria broth (L-broth (1% w/v bacto yeast extract, 1% w/v NaCl in ddH_20) and sterilised by autoclaving prior to use. Ampicillin was added for plasmid selection when required, final concentration of 25-100μg/ml, depending on the type of vector carrying the drug resistance gene. For plate cultures, L-broth was supplemented with 1.5% w/v bactoagar, autoclaved and allowed to cool to approximately 55°C. Ampicillin was added if required and the medium poured into 90mm petri dishes and allowed to set at room temperature.

Storage of bacterial stocks: For short term storage (< two weeks) bacteria were kept at 4°C. For long term storage, 100μl of 10x Hogness buffer (36mM K_2HPO_4, 13 mM KH_2PO_4, 20mM Na_3Citrate, 10mM $MgSO_4$, 40% w/v glycerol) was added to 900μl of bacterial suspension and then stored at -70°C.

7.3 Buffers and solutions

Commonly used buffers and solutions were made as laboratory stocks. Solutions were prepared using ddH_20 where applicable and sterilised by autoclaving if possible.

10x Denhardt's Solution: 2% w/v BSA (fraction V), 2% w/v Ficoll (Type 400,

Pharmacia), 2% w/v polyvinylpyrrolidone.

10x Glucomix: 250mM Tris-HCl pH 8.0, 500mM glucose, 100 mM EDTA.

5M KOAc pH 4.8: 3M potassium, 5M acetate.

2M NaOAc pH 5: 2M NaOAc pH adjusted with glacial acetic acid.

20x SSC: 3M NaCl, 0.3M Na_3Citrate.

10x TBE: 890mM Tris, 890mM boric acid, 20mM EDTA.

TE: 10mM Tris-HCl pH 7.5, 1mM EDTA.

10x TNE: 100mM Tris-HCl pH 7.5, 1M NaCl, 10mM EDTA.

Phenol, phenol/chloroform:

25mg 8-hydroxy quinoline was added to 500mg of phenol, this was then melted at 65^{0}C. 200mls 1M Tris-HCl pH7.5 and 200 mls dd$H_2$0 were added, mixed and allowed to separate. The aqueous phase was removed by aspiration and 200 mls of TE was added. The solution was again mixed, allowed to settle and the aqueous phase removed. This was repeated a further two times, the aqueous phase being left the last time. This was then used when only phenol was required.

For phenol/chloroform, an equal volume of chloroform was added and the mixture was left to settle. The phenol/chloroform was stored away from light.

7.4 Chemicals and reagents

Restriction enzymes, T4 DNA ligase, reverse transcriptase, polynucleotide kinase, alkaline phosphatase, proteinase K, Klenow fragment DNA polymerase and BSA were purchased from Boehringer Mannheim, Bethesda Research Laboratories (BRL), New England Biolabs and Pharmacia. Nylon filters were obtained from Amersham. General chemicals were obtained from BDH, Sigma chemical company and Pharmacia unless otherwise specified. Agarose powder was purchased from Sigma and low melting point agarose was obtained from ICN Biomedicals (FMC). Tissue culture media was obtained from Flow laboratories and Gibco BRL. Fetal calf serum was purchased from Imperial laboratories.

7.5 Competent bacteria

RecA$^-$ DH5α bacteria strain were routinely used for preparation of transformation competent bacteria. The RecA$^-$ Dam$^-$ JM110 bacteria strain were used when Dam methylation sites were required to be free of Dam methylation modification. Competent cells were prepared by a modified protocol of Cohen *et al.* (1972).

A single colony from a freshly streaked plate was picked into 10mls of L-broth and grown overnight at 37°C in a shaking incubator. 0.5 mls of this culture was added to 100mls of pre-warmed L-broth and grown at 37°C in a shaking incubator to an OD_{550} of 0.7. 50 mls of this culture was added to a further 550mls of pre-warmed L-broth and grown to an OD_{550} of 0.2. The culture was chilled on ice for 15 minutes then pelleted at 3500rpm for 20 minutes at 4°C (Beckman J6B centrifuge). The supernatant was discarded and the bacterial pellet resuspended in 250 mls of ice cold 100mM $CaCl_2$. After 20 minutes on ice the bacteria were pelleted again and gently resuspended in 10mM $CaCl_2$, 50% w/v glycerol. 100µl aliquots were frozen in liquid nitrogen and stored at -70^{0}C.

7.6 Bacterial transformation

10-100ng of ligated plasmid DNA was added to 50µl of thawed competent cells and incubated on ice for 45 minutes. The cells were then heat shocked at 42°C in a waterbath for 90 seconds and returned to ice for a further 5 minutes. 0.5 ml of pre-warmed L-broth was then added and the culture incubated at 37°C for one hour. The cells were then pelleted at 1000rpm for 60 seconds, resuspended in 100µl L-broth and then plated onto pre-warmed L-agar plates containing 50µg/ml ampicillin. The plates were then incubated overnight at 37°C.

7.7 DNA preparation

7.7.1 Small scale plasmid DNA isolation

A modification of the alkali lysis method for small scale DNA preparation was used (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981).

A 5ml overnight culture from a single colony was pelleted by centrifugation at 2500rpm for 10 minutes at 4°C (Beckman J6 centrifuge). The bacterial pellet was resuspended in 200µl of 1x glucomix and transferred to a 1.5 ml reaction tube (Eppendorf). 400µl of lysis solution (0.2M NaOH, 1% SDS) was added and mixed gently to lyse the bacteria. To this 200µl of 5M KOAc was added and the reaction vortexed to precipitate the chromosomal DNA. 600µl of phenol/chloroform was added, the tube vortexed and centrifuged at 14000rpm for 5 minutes (Eppendorf centrifuge 5415C). The aqueous layer was removed, plasmid DNA and cellular RNA was precipitated by the addition of 0.6 vol isopropanol and collected by centrifugation at 14000rpm for 10 minutes. The pellet was dissolved in 100µl TE containing RNAase to a final concentration of 10µg/ml and incubated at 37°C for 30 minutes. 2-5µl was typically used in a restriction digest to identify recombinant clones.

7.7.2 Large scale plasmid preparation

Bacteria carrying plasmid DNA were grown overnight at 37°C as a one litre L-broth culture containing 50µg/ml ampicillin. DNA preparation was carried out by alkaline lysis (Birnboim and Doly, 1979) as follows. The one litre culture was pelleted at 4000rpm for 30 minutes at 4°C (Beckman J6B centrifuge) and resuspended in 40ml 1x glucomix. To this 80ml of lysis buffer was added and mixed gently, followed by 40ml 5M KOAc pH 4.8 and mixed by inversion. The precipitated chromosomal DNA and bacterial components were removed by centrifugation and the supernatant was filtered through four layers of cheesecloth. A 0.6 vol of isopropanol was added. The solution was allowed to stand at room temperature for 10 minutes and then recombinant DNA was pelleted by centrifugation at 4000rpm for 30 minutes at room temperature. The pellet was washed with 70% ethanol, air dried and resuspended in 5ml of 100mM Tris-HCl pH8.0, 1mM EDTA. The solution was then made up to a total weight of 9g with 100mM Tris-HCl pH8.0, 1mM EDTA. 10.1g of CsCl and 1ml of ethidium bromide (5mg/ml) were added before transferring the solution to a "Quick-seal" polyallomer centrifuge tube (Beckman). CsCl gradients were centrifuged at 55000rpm, 25°C for 16 hours in a vertical rotor (Beckman L8 ultracentrifuge, vTi80 rotor). Supercoiled plasmid DNA was harvested from the gradient using a hypodermic needle and syringe.

Two volumes ddH_2O were added and the DNA precipitated by the addition of two volumes absolute ethanol. The DNA was pelleted by centrifugation (8000rpm 30 minutes Sorvall RC5C centrifuge), washed with 70% ethanol and resuspended in 500µl of TE containing 100µg/ml RNAase A and incubated at 37°C for one hour. The solution was then extracted with phenol/chloroform. The aqueous phase was removed and DNA precipitated with the addition of two volumes of ethanol and one tenth volume 2M NaOAc. The DNA was pelleted by centrifugation 14000rpm for 30 minutes, washed with 70% ethanol and resuspended in 500µl of TE. The concentration was determined by OD_{260}, 1 OD_{260}unit = 50µg dsDNA/ml.

7.7.3 Preparation of genomic DNA from mammalian tissue

The tissues to be analysed were typically 5mm of mouse tail, half a placenta or yolk sac. For the harder tissue (ie tail biopsy) the sample was placed into 700µl tissue lysis buffer (50mM Tris-HCl pH 8.0, 100mM EDTA, 100mM NaCl, 1% SDS, 100µg/ml proteinase K) and incubated overnight at 55°C.

In the case of the soft tissue (placenta and yolk sac) the 1% SDS and proteinase K were excluded from the lysis solution initially. The samples were homogenised by drawing through a 0.5 x 16mm hypodermic needle attached to 1 ml syringe. SDS and proteinase K were then added to the solution and the samples prepared as for the tail biopsies.

After overnight digestion, RNAase A was added to a final concentration of 10 µg/ml and the samples incubated at 37°C for one hour. One volume of phenol/chloroform was added and the samples were mixed well by agitation. The samples were centrifuged at 14000rpm for 10 minutes and the aqueous phase recovered. To the aqueous phase, 0.6 volume isopropanol (420µl) was added and mixed well. The DNA precipitate was hooked out using a flame sealed pasteur pipette, washed in 70% ethanol, air-dried and resuspended in 100µl of TE. Concentration was then determined by measuring OD_{260}.

For isolation of DNA from cultured cells, particularly ES cells, two protocols were used depending on the number of cells available. For samples of 3-10 x 10^6 cells (ie a confluent 3.5mm petri dish ES cell culture) cells were trypsinised and harvested

by centrifugation, washed once in PBSA, pelleted by centrifugation and resuspended in tissue lysis buffer, excluding SDS and proteinase K, at a concentration of 5-10 x 10^6 cells/ml. Once the pellet was resuspended SDS and proteinase K were added and the samples prepared as for the tissue preparation above.

For smaller samples, isolation of DNA from multi well plates, the following protocol was used. For 24 well plates, 500µl, and 96 well plates, 200µl, of proteinase K solution (100mM Tris-HCl pH8.5, 5mM EDTA, 0.2%SDS, 200mM NaCl, 400µg proteinase K/ml) were added directly to the well containing a super confluent ES cell culture. The plates were wrapped in "parafilm" to reduce evaporation and incubated at 55°C overnight. The following day the contents of the well were removed, placed in a 1.5 ml reaction tube and extracted with an equal volume of phenol/chloroform. The aqueous phase containing the genomic DNA was isolated and the DNA ethanol precipitated with two volumes absolute ethanol. The DNA was pelleted by centrifugation at 14000rpm and the pellet washed in 70% ethanol. The DNA was dissolved in 20µl of TE and allowed to dissolve overnight at 4°C. For Southern blot analysis, half of the sample from a 24 well plate was used and all the sample obtained from a 96 well.

7.8 Recombinant DNA

7.8.1 Restriction digests

Restriction endonuclease digests were carried out as specified by the enzyme supplier in the recommended buffers and at the appropriate incubation temperature for the enzyme used.

Partial restriction endonuclease digestion of DNA was carried out using 0.2 Units enzyme/µg of DNA for 1-2 hours with aliquots removed at frequent intervals. The reaction was stopped by addition of 10mM Tris-HCl pH 7.5, 100mM EDTA, 1% SDS. Aliquots were checked for the extent of digestion by electrophoresis in agarose gel.

7.8.2 Blunt end formation

5' overhangs were filled in using 23 units of reverse transcriptase per 10µg of

DNA, in 50µl of 1x reverse transcriptase buffer (10x RT buffer: 500mM Tris-HCl pH 8.3, 60mM $MgCl_2$, 600mM NaCl, 100mM DTT) with 1mM dNTPs. The reaction was incubated at 37°C for one hour. The sample was phenol/chloroform extracted and ethanol precipitated.

7.8.3 Dephosporylation

DNA ends were deposphorylated by incubation at 37°C with 1 unit of calf intestinal phosphatase (CIP) per 10µg DNA for one hour. This was routinely carried out in the completed restriction digest reaction or in 1x PME buffer (10x PME: 100mM Tris-HCl pH 9.5, 10mM spermidine, 10mM EDTA. The enzyme was inactivated by heating at 75°C for 10 minutes followed by phenol/chloroform extraction and ethanol precipitation.

7.8.4 Ligation of DNA molecules

DNA fragments for ligation were generally gel purified before ligation. The vector carrying the antibiotic resistance gene was dephosphorylated prior to ligation. Vector and insert were combined at approximately equivalent molar ratios to a final concentration of 800ng in 10µl of 1x ligation buffer (10x ligation buffer: 100mM Tris-HCl pH 8.0, 75mM $MgCl_2$, 10mM ATP, 100mM DTT, 2mg/ml gelatin) containing 1 unit of T4 DNA ligase. The DNA fragments were ligated overnight at 16°C.

7.8.5 Agarose gel electrophoresis

Horizontal agarose gels were made with 1x TBE buffer containing 0.5 µg/ml ethidium bromide. 0.5-2% gels were run in 1x TBE buffer containing 0.5µg/ml ethidium bromide, at a constant voltage of 5-10V/cm.

0.2 volume of Orange G loading dye (20% Ficoll, 110 mM Tris-HCl pH 7.5, 0.25% orange G w/v) was added to each sample before loading into the gel. For standard molecular weight markers, BstEII digested lambda phage DNA was used for fragments > 1kb and HinfI digested PBR322 DNA was used for smaller DNA

fragments, < 1kb. Nucleic acids were visualised on a short wave ultraviolet transilluminator unless the gel was to be used for preparative purposes where a long wave ultraviolet illuminator was used.

7.8.6 Denaturing polyacrylamide gel electrophoresis

Gel solutions were made at 5% to 20% acrylamide w/v (acrylamide:bis-acrylamide, 19:1 w/w) in 1 x TBE buffer containing 7M urea (final concentration). The gel was polymerised by the addition of 0.1% ammonium persulphate (final concentration) and 0.1% TEMED (final concentration, N,N,N',N'-Tetramethylethylenediamine). Gels were poured between glass plates seperated by 0.4mm (analytical) or 1.5mm (preparative) spacers. Gels were run vertically in 1 x TBE buffer at 25-35V/cm. Visualisation of radioactive bands was carried using either X-ray film or Molecular Dynamics phosphorscreens. Gels were either kept "wet" and surrounded with "saran" wrap or fixed in 12% methanol, 10% acetic acid (glacial) v/v, for 15 to 30 minutes, dried under vacuum at 80°C before exposing to film or phosphorscreens.

7.8.7 Isolation of DNA from low melting point agarose gel

DNA fragments to be purified were resolved by electrophoresis through low melting point agarose of 0.5-2% dependent on the molecular weight of DNA to be isolated. After visualisation on a long-wave ultraviolet transilluminator, the required band was cut out and melted in a 1.5 ml eppendorf tube at 65°C for 20 minutes. The solution was extracted twice with an equal volume of phenol (no chloroform) and extracted once with phenol/chloroform. 2M NaOAc (pH 7.5) was added to a final concentration of 0.3 M and the DNA was precipitated with two volumes absolute ethanol, pelleted by centrifugation at 14000rpm for 30 minutes and washed with 70% ethanol. The pellet was dissolved in TE. Concentration was determined by comparison to molecular weight markers in agarose gel.

7.8.8 Preparation of oligonucleotides

Oligonucleotides were synthesised on an Applied Biosystems oligonucleotide synthesiser. The oligonucleotides were supplied heat deprotected in 35% ammonia solution in a volume of 1.5 ml. The solution was chilled to -20°C for 30 minutes. A 360µl aliquot was removed, 40µl of 3M NaOAc and 1.2 ml absolute ethanol were added to this. The solution was chilled to -20°C for 1 hour. The DNA was pelleted by centrifugation at 14000rpm for 30 minutes, washed with 70% ethanol and resuspended to a concentration of 2-5mg/ml in ddH_2O. The concentration was determined by OD_{260}, 1 OD_{260} = 30mg/ml oligonucleotide.

If further purification was required, 100µg of oligonucleotide was denatured at 85°C in the presence of 50% formamide for 5 minutes, one volume of 50%glycerol/0.01% bromophenol was added and the solution run on a 8-12% denaturing polyacrylamide gel. The oligonucleotide band was identified by UV shadowing using a hand held UV light source set to long wave UV. The band was cut out and crushed into 500µl of acrylamide elution buffer (0.5M ammonium acetate, 10mM Mg acetate, 1mM EDTA, 0.1% SDS) by passing through a 1ml syringe. After incubation overnight at 37°C, the acrylamide gel fragments were separated from the supernatant by centrifugation. The oligonucleotides were precipitated with the addition of one tenth volume 3M NaOAc and two volumes absolute ethanol. The oligonucleotides were pelleted by centrifugation at 14000rpm for 30 minutes, washed with 70% ethanol and resuspended in ddH_2O. Concentration was measured by OD_{260}.

7.8.9 Annealing of oligonucleotides

Single stranded complementary oligonucleotides were annealed for formation of linker sequences for cloning purposes. Oligonucleotides were mixed in an equimolar ratios to a total of 250ng in 100µl of 1x kinaseing/annealing buffer (100mM Tris-HCl pH 7.5, 500mM Na Cl, 100mM Mg Cl_2, 10mM EDTA, 10mM DTT). The reaction was heated to 90°C for 5 minutes and then allowed to cool to room temperature overnight.

7.8.10 Radioactive end-labelling of oligonucleotides

Radioactive end-labelling of oligonucleotide primers for sequencing purposes was carried out as follows. 50ng of single stranded oligonucleotide was end labelled in the following reaction.

50 ng single stranded oligonucleotide

2µl 10x kinasing/annealing buffer

4µl γ^{32}P-ATP

1µl 10 units/µl T4 polynucleotide kinase

ddH_2O up to 20µl

The reaction was incubated at 37°C for one hour, the reaction was halted by adding one volume of 2 x stop mix (20mM Tris-HCl pH 7.5, 200mM EDTA, 2% SDS) and the solution was phenol/chloroform extracted. Free nucleotides were removed by G-50 chromatography.

7.8.11 Sequencing of double-stranded plasmid DNA

Sequencing of double-stranded plasmid DNA was carried out using a Sequenase Kit (United States Biochemicals) according to the manufacturer's protocol with the following modifications. 10µg of plasmid DNA was denatured in the presence of radiolabelled primer (kinased to a specific activity of approximately 10µCi/µg) using 0.2M NaOH and incubated at 37°C for 30 minutes. The reaction was neutralised by the addition of 2M NaOAc to a final concentration of 0.3M and transferred to a new reaction tube. The DNA was precipitated by the addition of 2.5 volumes of absolute ethanol and pelleted by centrifugation. The DNA was resuspended in 8µl of ddH_2O and then processed as described in the Sequenase protocol. The final denatured elongation products were run on 5-8% denaturing polyacrylamide gels, dried at 80°C under vacuum and exposed to X-ray film.

7.9 Analysis of genomic DNA

7.9.1 Southern blotting

10µg of genomic DNA from either tissue or cell culture was digested with the

appropriate restriction endonuclease from 4 hours to overnight depending on enzyme. The digested genomic DNA was electrophoresed overnight through an agarose gel, 0.7% for analysis of fragments between 1-10kb, at approximately 2-4V/cm depending on separation required. Molecular weight markers were run either side of the sample lanes. The gel was subsequently visualised on an ultraviolet transilluminator and photographed with a ruler next to the molecular markers for later comparison with the autoradiograph. Blotting was carried out by a modified protocol of Southern (1975) using nylon filters under denaturing conditions. The gel was treated by gently rocking in the following solutions: 30 minutes 0.2M HCl; 45 minutes 0.5M NaOH, 1.5M NaCl. The DNA was transferred on to the 0.45 μm nylon filter using the following set up. The gel was inverted onto "saran wrap" on a level surface. The nylon filter, which had been pre-wetted in ddH_2O 1 minute and 0.5M NaOH, 1.5M NaCl 1 minute, was laid upon the gel avoiding air bubbles. Two layers of 3MM paper (Watman) soaked in 0.5M NaOH, 1.5M NaCl were placed on top followed by two dry layers of 3MM paper. On top of this was placed a wad of paper towels and a 1kg flat weight. The transfer was allowed to proceed for 4-5 hours at room temperature after which time the layers of paper were removed. The position of the loading slots were marked on the filter with a pencil. The filter was removed and briefly soaked for 10-20 seconds in 0.5M Tris-HCl pH 7.1, 3M NaCl and UV crosslinked using a "Stratlinker" (Stratagene) set to auto crosslink. The membrane was then used in hybridisation reactions with a radioactive probe.

7.9.2 Preparation of radiolabelled probes for hybridisation to filter bound DNA

Probes smaller than 400 bp were radiolabelled using an oligolabelling protocol (Feinberg and Vogelstein 1983), larger probes were labelled by nick translation (Rigby *et al.* 1977). For both procedures radioactive nucleotides had a specific activity of 3000Ci/mmol and a concentration of 10mCi/ml.

Oligolabelling was carried out using 50-100ng of double-stranded probe DNA. This was suspended in 10μl of ddH_2O and heated to 100°C for 3-5 minutes and placed on ice. To this was added:

2μl Oligolabelling buffer (OLB)

1μl 10mg/ml BSA

2.5µl (α^{32}P)dATP

5 units of Klenow fragment DNA polymerase

Oligolabelling buffer is 250mM Tris-HCl pH8.0, 25mM $MgCl_2$, 100µM dCTP, 100µM dGTP, 100µM dTTP, 1M HEPES pH7.6, 27 OD_{260}/ml hexadeoxyribonucleotides, 7mM β-mercaptoethanol.

The reaction was allowed to proceed at room temperature for 2½ hours at which time the reaction was stopped by addition of 90 µl of 10mM Tris pH 7.5, 100mM EDTA, 1% (w/v) SDS. Unincorporated nucleotides were removed by G-50 chromatography.

Nick translation was carried out using a Nick Translation Kit (Amersham) using a modification of the manufacturers protocol. The following reaction mix was set up using 100ng of probe DNA in the supplied solutions:

4µl ddH_2O

2µl dNTP mix (100µM dGTP, dCTP, dTTP in Tris-HCl pH 7.8 buffer including β- mercaptoethanol and $MgCl_2$ (as supplied by Amersham)

1µl Enzyme mix (0.5 Units/µl DNA polymerase I, 10pg/µl DNase I in Tris-HCl pH 7.5 including $MgCl_2$, glycerol and BSA (as supplied by Amersham)

3µl (α^{32}P)dATP

The reaction was incubated at 16°C for 90 minutes at which time the reaction was stopped by addition of 10mM Tris pH 7.5, 100mM EDTA, 1% (w/v) SDS. Unincorporated nucleotides were removed by G-50 chromatography.

7.9.3 Hybridisation of filter-bound DNA to radiolabelled probes

The filter was wetted in 2x SSC then transferred to 20 ml of hybridisation buffer (3x SSC, 0.1% SDS, 10x Denhardt's, 10% dextran sulphate, 100µg/ml denatured salmon sperm DNA) at 65°C in a rolling roller bottle (Hybaid) and pre hybridised for one hour in a rotary hybridisation oven (Hybaid). The radioactive probe was denatured at 100°C for 5 minutes before addition to the filter. Hybridisation was carried out in a final volume of 10 ml of hybridisation buffer and probe overnight. The filter was washed twice in 3x SSC, 0.1% SDS for 15 minutes per wash, followed by two washes in 0.3x

SSC, 0.1% SDS for 15 minutes each wash at 65°C. The filter was wrapped in "saran" wrap or placed in a heat sealed bag and exposed to either X-ray film or phosphorscreen. To remove radiolabelled probe, wet filters were stripped in 0.2 M NaOH for 20-30 minutes and washed in 2 x SSC.

7.10 Embryonic Stem cell culture

7.10.1 Media and solutions

For culture of embryonic stem cells and SNL feeder cells the following media and solutions were used.

Embryonic stem cell culture media for growth of AB1 ES cells was as follows: 20% fetal calf serum (FCS)(Imperial Laboratories), 80% Dulbeco's modified eagles medium (DMEM) (no pyruvate, high glucose in 1x concentration (Flow Labs)), antibiotics penicillin (final concentration 50 units/ml, stored as 100x stock) and streptomycin (final concentration 50μg/ml, stored as 100x stock)(Gibco BRL), 1mM L-glutamine (Gibco BRL), 0.1mM each glycine, L-alanine, L-aspartic acid, L-asparagine, L-glutamate, 0.2mM L-proline, L-serine (added as 100x stock Flow Labs Non Essential Amino Acids) and 1 μM β-mercaptoethanol (made as 100x stock in PBSA and made weekly). Fetal calf serum was batch tested for ability to support growth of ES cells and maintain pluripotency according to Robertson (1987).

E14 ES cell media was as follows: 15% FCS, 85% 1x DMEM (10x Gibco, made to 1x and buffered with 44mM $NaHCO_3$), 1mM L-glutamine, 1mM sodium pyruvate (Gibco BRL), 1μM β-mercaptoethanol, 10μg/ml gentamicin (Gibco).

For growth of the SNL feeder line, DMEM containing 10% FCS was used. Antibiotics and glutamine were as for AB1 media.

Media were maintained at 4°C and heated to 37°C prior to use. The following additional solutions were used in the culturing process:

PBS-A

Used for washing cells and rinsing cells prior to trypsinisation. One lite of PBS solution made in ddH_2O contains the following: 10g NaCl, 0.25g KCl, 1.44g Na_2HPO_4,

0.25g KH_2PO_4. Autoclaved to sterilise.

Trypsin/EDTA

A solution of 0.25% (w/v) trypsin, 0.04% EDTA was prepared as follows: For a 1 litre stock 2.5g porcine trypsin (Difco), 0.4g EDTA, 7.0g NaCl, 0.3g $Na_2HPO_4.12H_2O$, 0.24g KH_2PO_4, 0.37g KCl, 1.0 g D-glucose, 3.0g Tris, 0.01g phenol red (Difco) ddH_2O to 1 litre. The solution was adjusted to pH 7.6 with 1M HCl and sterilised through a 0.22 μM filter. The stock was dispensed into 5 ml aliquots and stored at -20°C.

Gelatin

For 1 litre of 0.1% gelatin solution, 1g swine skin gelatin (Sigma) in ddH_2O autoclaved and stored at room temperature.

β-mercaptoethanol

For a 0.1mM stock, 14μl of β-mercaptoethanol (Sigma) was dissolved in PBS and used as a 100x stock. This was stored at 4°C and made weekly.

Mitomycin-C

A 2g vial of mitomycin-C (Sigma) was rehydrated with 2 ml of sterile PBS to make a 1mg/ml 100x stock. The mitomycin-C solution was stored in the dark at 4°C and used within one week of hydration. Care should be taken in the handling and disposal of mitomycin-C solutions as it is highly toxic.

7.10.2 Routine subculture of SNL 76/7 cells

All the ES cell lines used in this study were cultured on the neomycin resistant STO subline SNL76/7 from Dr. A. Bradley (McMahon and Bradley, 1990). SNL76/7

cells were maintained in culture on 10 cm tissue culture dishes in 5% CO_2 at 37°C. SNL cells were seeded at a density of approximately 1 x 10^4 cells/cm^2, grown to confluence and then split 1:10. A 10cm culture tissue yields 8-10 x 10^6 cells. Over confluent cultures were avoided as this leads to loss of contact inhibition and produces poor feeder cells.

7.10.3 Preparation of SNL feeder layers

Two methods for preparation of feeder layers were used, for the AB1 and D3 culture mitomycin-C inactivated feeder layers were used. For growth of E14 cells γ-irradatiated cells were used.

Mitomycin-C inactivation

From confluent SNL cultures the medium was aspirated and replaced with SNL medium containing 1μg/ml mitomycin-C. The SNL cells were incubated for 3 hours. During this period the appropriate number and type of plates were gelatin coated by addition of the appropriate volume of 0.1% gelatin solution and incubated for at least one hour at room temperature. After the three hour period, the mitomycin-C media was removed from the SNL cells by aspiration and the cells were washed three times with 10ml of PBS-A. To each 10cm plate 0.5 ml of trypsin/EDTA was added and the cells briefly incubated at 37°C until they started to detach. The trypsin was quenched with SNL medium and the cells transferred to a conical centrifuge tube and pelleted at 1000rpm for 5 minutes. Cells were resuspended in SNL medium to give a single cell suspension and counted in a haemocytometer. The SNL feeder cells were diluted and plated at a density of 5 x 10^4cells/cm^2 in the gelatinised plates.

The feeder cells attach within 30 minutes and were normally used 12 hours after plating although they could be used sooner if need be. Feeder plates were usable for 10 days after they were made although fresh plates were used preferentially. Prior to plating of ES cells the SNL medium was removed.

γ-Irradiation

For γ-irradiated feeder cells, SNL cells were cultured in the tissue culture dish that the ES cells would be cultured in. Cells were plated as above at a density of 1 x 10^4 cells/cm^2. When the culture was semi-confluent, approximately 2 x 10^4 cells/cm^2 (estimated by eye) the culture was irradiated with 3000 rad of γ-radiation using a ^{137}Cs source (Gamma cell 40, Atomic energy of Canada limited). The cells then underwent a further division before mitotic arrest and were used from 3 to 13 days post irradiation.

7.10.4 Culture of embryonic stem cells

Routine subculture

Ideally ES cells were stored as frozen aliquots and only grown prior to a targeting experiment. From the original stock the cells are expanded and frozen in aliquots. ES cells are frozen at a concentration of 0.5-1 x 10^7 cells/ml. Once the cells were thawed, the media was replenished daily. The culture normally reached confluence 2-3 days after plating at which time it was divided and replated at a density of 3-3.5 x 10^5 cells/ml, normally this was a 1 in 3 dilution. Over growth was avoided as this can lead to differentiation of the ES cells.

Thawing ES cells

A frozen ampoule of ES cells was removed from liquid nitrogen and quickly thawed at 37°C in a water bath. The content of the ampoule was transferred to a clean 10ml conical centrifuge tube and 10 ml of prewarmed ES cell medium was added dropwise, carefully agitating after each drop was added. The cells were then pelleted by centrifugation at 1000rpm for 5 minutes at room temperature. The pellet was resuspended in ES cell medium to a density of 3-3.5 x 10^5 cells/ml and plated onto feeder dishes. The culture was placed into a 37°C, 5% CO_2 incubator. After 2-3 days the culture reached the desired confluence and was either frozen, passaged or used for targeting experiments.

Sub-culture of ES cell

The ES cell culture was passaged once it had reached confluence, as determined by eye, the ES cells appearing as large groups of cells on the feeder cell monolayer. Between 1-2 hours prior to passaging the medium was replenished to improve viability upon passage.

The medium was aspirated from the culture and the ES cells gently washed with the appropriate volume of PBS ie 10ml for a 90mm dish. To this 1ml of trypsin/EDTA was added for a 90mm dish, 0.5 ml for 60mm. The culture dish was returned to the incubator for 1-2 minutes by which time the cells were beginning to lift from the surface. The trypsin was quenched by adding 5-10ml (dependent on dish size) of ES medium and the cell suspension placed into a 50ml conical tube. The cells were pelleted by centrifugation at 1000rpm for 5 minutes at room temperature. The pellet was resuspended in 1-1.5 ml of ES medium. To obtain a single cell suspension the suspension was drawn through a flame polished pasteur pipette up to ten times (clumps of ES cells tend to differentiate once plated). The cell suspension was diluted to 20ml and a cell count performed using a haemocytometer, feeder cells contained in the suspension can be identified in the counting chamber as larger cells and can be easily distinguished from the ES cells. The cells were diluted to a density of 3-3.5 x 10^5 cells/ml and plated on to fresh feeder layers. As mentioned above, the ES media was replenished daily and cells routinely passaged every second or third day.

The above protocol was used for the culturing of AB1 ES cells. For E14 culture on feeder layers a number of modifications to the above protocol were used. For routine culture, E14 ES cells were cultured in T75 tissue culture flasks (Costar) containing the feeder layer of cells. The plating density was higher, 1 x 10^6 cells/ml or 2 x 10^5 cells/cm^2. Cells were passaged every two days. Prior to trypsinisation the ES cells were washed with PBS three times. To a T75 flask 1.5ml of Trypsin/EDTA was added. The cells were returned to the 37°C incubator for one minute. The morphology of the cells was then checked periodically under the phase contrast microscope. When the ES cells appeared rounded but were still attached to the feeder layer the flask was held horizontally and the side of the flask hit violently three times. The trypsin was immediately quenched and cells resuspended to the above density and replated. This

method was successful in obtaining a single cell suspension.

Freezing ES cells

After trypsinisation and cell counting the ES cells were re-pelleted by centrifugation at 1000rpm for 5 minutes and resuspended at 2 x 10^7 cells /ml in ES cell medium. To this suspension an equal volume of ES cell medium containing 20% DMSO (Sigma) was added. 1ml aliquots of the cell suspension were placed in Nunc freezing vials, the final cell density being 1 x 10^7 cells/ml. The vials were either placed into the vapour phase of a liquid nitrogen tank for 24 hours or placed in an isopropanol freezing container and placed at -70°C for 24 hours. The vials were then transferred into liquid nitrogen for permanent storage.

7.10.5 Electroporation of ES cells

Preparation of targeting vector DNA

The targeting vector DNA was purified as a large scale preparation over a CsCl gradient as described in section 7.7.2. The vector DNA was linearised by restriction enzyme digestion and extracted with phenol/chloroform. The DNA was ethanol precipitated, washed with 70% ethanol and dissolved in sterile PBS at a concentration of 1μg/μl. The last three steps were carried out in a sterile hood.

Electroporation of AB1 ES cells

AB1 ES cells were harvested by trypsinisation of confluent 60mm plates, 2 hours after replenishing the media. After centrifugation the cells were washed twice with PBS and resuspended at a concentration of 1.2 x 10^7 cells/ml. 0.8 ml of the cell suspension was transferred to a Biorad 0.4mm electroporation cuvette. 25-50 μg (depending on vector size) of targeting vector DNA was added and mixed. Electroporation was carried out at room temperature using a Biorad "gene pulser" electroporator. The cell and DNA suspension was given two pulses, one at 230V, 500μF immediately followed by another at 240V, 500μF, according to McMahon and Bradley (1990). The cells were allowed

to stand for 10 minutes after electroporation. The ES cells were then diluted 1/20 and plated onto fresh 90mm feeder plates, a 1/200 dilution was plated onto 60mm feeder plates. Twenty four hours after plating drug selection was applied. To the 90mm plates G418 (350μg/ml; Gibco BRL) and GANC (2μM; Syntex pharmaceuticals Ltd.) selection was applied. To the 60mm plates only G418 selection was applied. This served as a control to determine the extent of enrichment using GANC as negative selection. The following controls were also carried out. An electroporation as above was made excluding the targeting vector, the 90mm plate being used to test G418 killing in the absence of the targeting vector. The 60mm plate was harvested 24 hours after the electroporation to asses the cell survival after the electroporation procedure. A no-shock control was used to compare cell survival with the no-DNA electroporation.

Selection was maintained for 9-12 days after electroporation. ES cell colonies routinely appeared around day 7 post electroporation. The ES cell medium containing drug selection was replenished daily.

Electroporation of E14 ES cells

E14 ES cells were grown to confluence in a $75cm^2$ flask (T75 Costar). The cells were harvested by trypsinisation, a cell count performed and 3×10^7 cells resuspended in 1.2ml of ES cell medium. For electroporation 0.4 ml aliquots containing 1×10^7 cells were taken and placed into 2mm electroporation cuvettes (BTX) containing 10-25μg (depending on vector size) of targeting vector DNA. The cells were electroporated at 128V, 1080μF using a Hoefer PG200 Progenetor II electroporator. Cells were allowed to stand at room temperature for 10 minutes after electroporation. The cells were then diluted 1/20 and plated as for AB1 ES cells. FIAU (0.2μM, Bristol Myers obtained from Dr J. van Deursen) was used for counter selection.

7.10.6 Harvesting of ES cell colonies

ES colonies were harvested at 9-12 days post electroporation. Two techniques were used for picking ES colonies. Either flame drawn out pasteur pipettes on a mouth controlled air line or a Gilson 10μl pipette set to 2μl were used. The former technique

was used for collection of ES colonies in the Thy-1 targeting experiment.

Using the drawn out pasteur pipette, the colony was dislodged from the feeder layer and transferred to a microdrop of trypsin/EDTA. The colony was disaggregated by pulling up and down through the pipette and incubated for 2 minutes at 37°C. The disaggregated colony was placed in 2ml of ES cell medium and divided into two individual wells of a 24 multiwell feeder plate (Costar). As each well reached confluence it was passage first to 3.5 mm multi well feeder plate and then to a 60mm feeder dish. At this point the cells were frozen down as previously described and the duplicate plate was used to isolate DNA as described in section 7.7.3.

Using a Gilson 10µl pipette, the ES cell colony was dislodged from the feeder layer and drawn into the pipette tip. The colony was transferred to an individual well of a 96 U-well plate containing 25µl of trypsin/EDTA. Colonies were picked in groups of twelve. Once a row of the 96 well plate was full, the plate was transferred to the 37°C incubator and incubated for 2 minutes. The plate was then removed and the 12 colonies disaggregated using a 12 channel multi pipette. The contents of the well were then transferred to either a 24 or 96 multiwell feeder dish containing the appropriate volume of ES cell medium. The contents of the well were grown to confluence, trypsinised and split into two wells. One well was allowed to over grow, this was to be used for DNA isolation.

In the case of the 24 well plate, duplicate clones were grown until confluent and trypsinised from the well. The trypsin was quenched in 250µl of ES cell medium. To this was added 250µl of 20% DMSO in ES cell medium and the clones were frozen in 1ml Nunc freezing vials.

For the duplicate colonies passaged into wells of the 96 well plate, 25 µl of trypsin/EDTA was added after washing with PBS and the cells incubated at 37°C for 2 minutes, the cells were disaggregated using a multichannel pipette. The contents of the well were then removed and placed in a pre-prepared well containing 75µl of ES cell medium, to this 100µl of ES cell medium containing 20% DMSO was added and the plates were transferred to a polystyrene box and placed in a -70°C freezer. The clones were kept at -70°C for 1-3 months by which time positive clones had been identified.

Frozen clones were recovered as follows. Clones frozen from the 24 well plate

were thawed as for normal ES cells, resuspended in 1ml of ES cell medium and plated into a fresh 24 well feeder plate. Clones frozen in the 96 well plate were recovered by floating the dish in the 37°C water bath. The contents of the well containing the targeted clone removed, transferred to a sterilised 1.5 ml reaction tube, to which 1ml of ES cell medium was added gradually and cells were spun briefly at 1000rpm. The supernatant was removed and the cells were resuspended in to 1 ml ES cell medium and transferred to an individual well in a 24 well feeder plate.

Targeted clones were passaged up to 3.5 mm plates from which frozen stocks were made. Chromosomal counts of ES cells were made according to Robertson (1987).

7.10.7 Generation of chimeric mice

Chimeric mice were generated by blastocyst injection of ES cells as described by Bradley (1987). Both blastocyst isolation and ES cell injection were carried out using DMEM/Hepes (Sigma) supplemented with 10% fetal calf serum.

Blastocysts were obtained by flushing the uterine horns from C57BL/6 females 3.5 days post-coitus. Each uterine horn was flushed using 1 ml of the above medium into a round bottom watch glass. Blastocysts were transferred into microdrop cultures under paraffin oil and maintained in a 37°C incubator until required for injection.

ES cells were prepared for injection as follows. From a confluent 3.5 mm well ES cells were trypsinised and collected by centrifugation as above. The cell pellet was resuspended in 1 ml of DMEM/Hepes, 10% FCS and made into a single cell suspension by passing through the end of a flame polished pasteur pipette, this was confirmed under the microscope. The cell suspension was made up to 5 ml and placed on ice.

Injections were carried out using a Nikon inverted microscope and Leitz micromanipulators. A solid state cooled injection chamber was used as described by Bradley (1987). Injection needles were prepared using a Kopf vertical pipette puller and holding pipettes were prepared by hand over a small flame. A microforge (deForbrune) was used for polishing the end of the holding pipette and making bends in both the needle and holder (Bradley, 1987). Control of blastocyst pick up and ES cell pick up and injection was achieved using air lines attached to the holder and needle micromanipulators. The suction and force was provided by two 5ml plastic syringes.

This method of control proved to be more efficient than the use of oil filled control lines.

To each blastocyst, 10-15 ES cells were injected into the blastocoel cavity. The blastocysts were allowed to expand after injection for one hour. 8-10 injected blastocysts were returned to the uteri of 2.5 day post-coitus anaesthetised F1(CBA x C57BL/10) pseudo-pregnant females (Bradley, 1987). Pups were born approximately 17 days after blastocyst injection. Pups were scored for coat colour chimerism after day 7 when agouti hairs could first be seen.

Male chimeras were set up to breed with C57BL/6 females at six weeks of age. Transmission of the ES cell genome was scored by agouti coat colour on 7 day old offspring.

7.11 Construction of size selected 129 genomic library

7.11.1 Preparation of size selected AB1 ES cell genomic DNA

Genomic DNA was prepared from the AB1 ES cell line. 1.2 x 10^7 AB1 ES cells were isolated from three confluent 60mm feeder plates. After trypsinisation, centrifugation and suspension, the ES cells were plated on to 100mm bacterial plates (Sterilin). This allows the feeder cells to be separated from the suspension, feeder cells preferentially attach to the surface of the bacterial plate. The feeder cells were allowed to attach for one hour, at which time the ES cell suspension was carefully drawn off, centrifuged and the DNA prepared as described in section 7.7.3. 100µg of ES cell DNA was digested with EcoRI and size fractionated on a preparative low melting point gel, electrophoresis was carried out at 3V/cm. 10µg of lambda bacteriophage DNA digested with BstEII was added to the genomic DNA as an internal size marker. Six 2mm gel slices were isolated around the 8.4 kb marker band. The DNA was isolated from each gel slice by phenol extraction Aliquots were tested for the presence of the Thy-1 gene by Southern blot analysis, hybridising with the 3' Thy-1 ApaI probe.

7.11.2 Preparation of size selected genomic library

100 ng of the positive DNA fraction was cloned into the EcoRI digested arms

of the Lambda ZAP II vector (Stratagene). The vector DNA containing the insert were packaged using the Giga Pack Gold packaging extract (Stratagene), according to the manufacturer's protocol. Titration, plating and subsequent screening of the library was carried out according to the manufacturer's protocol. Positive plaques were screened using the 3' Thy-1 ApaI probe. Following secondary screening the positive clones were excised as pBluescript II Sk- clones according to the manufacturer's *in-vivo* ZAP excision protocol.

7.12 RNA analysis

7.12.1 Preparation of RNA from tissues

RNA from mouse tissue was prepared using a LiCl, Urea precipitation method adapted from Auffrey and Rougeon (1980). Mouse tissue which had been isolated by dissection was routinely "snap-frozen" and stored in liquid nitrogen until required for analysis. The tissue was homogenised in 3M LiCl, 6M Urea 5-10 ml/g tissue using an ultraturrax homogeniser. The samples were then sonicated for 1 minute on ice using a microtip to break up genomic DNA, the samples were kept at 4°C over night to precipitate the RNA. The following day the RNA precipitate was pelleted by centrifugation at 3500 rpm at 4°C (Beckman J6B centrifuge) and the supernatant discarded. The samples were resuspended in half the original volume of LiCl buffer (kept at 4°C) and then pelleted as above. The RNA pellet was dissolved in 0.5 ml TE containing 0.5% SDS. An equal volume of phenol/chloroform was added and the samples were shaken at room temperature for 30 minutes. The samples were then centrifuged at 14000rpm at room temperature for 10 minutes. The aqueous phase was removed to a fresh tube and the RNA precipitated with the addition of 0.15 volume of 2M NaOAc and two volumes of absolute ethanol. The RNA was pelleted by centrifugation at 14000rpm, 4°C for 30 minutes. The pellet was washed with 70% ethanol and dissolved in ddH_2O and stored at -20°C. The concentration of the RNA sample was assessed by OD_{260}, 1 OD_{260} unit = 40μg RNA/ml. The quality of RNA was checked by electrophoresis of 2μg of the sample on a 2% agarose gel.

7.12.2 Northern blot analysis of mRNA

Northern blot analysis was carried out by a modified protocol of Fourney *et al.* (1988).

Sample preparation

10-30μg of total RNA was precipitated by the addition of 2.5 volumes absolute ethanol, 0.3 M NaOAc. The RNA was pelleted by centrifugation and dissolved in 25μl 1 x MOPS buffer (10x: 0.2M MOPS, pH7.0, 50mM NaOAc, 10mM EDTA), 50% deionised formamide, 7% formaldehyde. The samples were heated at 65°C for 15 minutes and placed on ice. 1/5 volume Orange G loading buffer and 1μl 1.0 mg/ml ethidium bromide solution (1.0mg/ml in deionised RNase free H_2O) was added prior to loading on the gel.

Gel preparation and electrophoresis

A 0.8% agarose gel was prepared as follows. 2.8 g of agarose was dissolved by boiling in 340 ml (dimethyl pyrocarbonate (DEPC)-treated) autoclaved H_2O. 40 ml of 10 x MOPS buffer and 20.4 ml 37% formaldehyde were added in a fume hood to the dissolved agarose once it had cooled to 50°C. The solution was mixed and poured into an RNase free 20x20 cm gel cast. The gel was run in 1 X MOPS buffer. Sample wells were flushed prior to loading the 30μl samples. The gel was run overnight at 3V/cm at room temperature

Transfer to nylon membrane

The gel was prepared prior to blotting by soaking for two 20 minute periods in 10x SSC at room temperature with gentle shaking. The gel was blotted for 8 hours at 4°C using the following set up. Two layers of 3MM paper pre-wetted in 10x SSC were placed on a flat surface, the gel was placed on top of this. A 0.2μm nylon membrane pre-wetted in 10x SSC was placed on top of the gel. Three dry sheets of 3MM paper

followed by a wad of paper towels and 500g weight were placed on top. When the transfer was completed the filter was baked at 80°C for two hours. Prior to hybridisation the filter was pre-wetted in 2x SSC.

7.12.3 Hybridisation of filter bound RNA to radiolabelled probes

DNA probes were radiolabelled as described in section 7.9.3. Filters were prehybridised in 50% formamide, 10% dextran sulphate, 10x Denhardt's, 50mM Tris-HCl pH7.5, 100μg/ml sonicated salmon sperm DNA and 0.1% SDS for two hours at 42°C in a rolling roller bottle (Hybaid) using a Hybaid oven. The heat denatured labelled DNA probe was added and allowed to hybridise to the filter over night. The filter was washed and exposed to film or phosphorscreen as described in section 7.9.3.

7.12.4 Analysis of mRNA species by PCR

Reverse transcription followed by PCR was used to amplify mRNA species from the (targeted) *Zfp-37* alleles. PCR products were subsequently cloned and sequenced or Southern blotted. The following oligonucleotide primers were used:

B2 5'-ATC GAA TTC TTA **AAG CAG TTT TTT CTC GGA**-3' *Zfp-37* exon IV reverse primer.

B5 5'-ATC GGA TCC **ATG GCT ACA TCC GAG CCT**-3' *Zfp-37* exon Ia forward primer.

1BS 5'-ATC CAT GGA TCC **TGC TGT AAC TTT CGG AAA TC**-3' *Zfp-37* exon Ib forward primer.

B6 5'-ATC CAT GGA TCC **GAA TGG GAG CAG CTG GAA**-3' *Zfp-37* exon II forward primer.

Sequences in bold are homologous to *Zfp-37* sequence. Sequences are derived from Mazarakis *et al.* (1995).

Reverse transcriptase reaction

Total tissue RNA was isolated as above. The following reverse transcriptase reaction

was set up:

8µl total tissue RNA 100ng/µl in ddH_2O

2µl 10x RT buffer (as section 7.8.2)

8µl 3.75 mM dNTPs

2µl 10mM DTT

7.25µl RNase inhibitor RNasin (40 Units/µl, Boehringer)

0.5µl oligonucleotide primer P5 4µg/ml

0.5µl reverse transcriptase (20.5 Units/µl)

The reaction was incubated at 41°C for two hours. The reaction was made up to 500µl with ddH_2O and frozen at -20°C for storage.

Polymerase chain reaction

The following polymerase chain reaction mix was set up:

5µl cDNA from the above reverse transcriptase reaction

10µl 5x PCR buffer (250mM Tris pH 9.0, 75mM ammonium sulphate, 35 mM $MgCl_2$, 0.85 mg/ml BSA (Boehringer), 0.25% Nonidet P (Sigma))

20 µl 3.75 mM dNTPs

1µl forward primer B5,1BS or B6 (100ng/µl)

1µl reverse primer B2 (100ng/µl)

0.1µl Taq polymerase (5 Units/µl)

The reaction made up to 50µl with ddH_2O and overlaid with 30µl of paraffin oil (Sigma). Thermal cycling was carried out in a Perkin Elmer Cetus Thermal cycler, using the following heating cycles:

94°C x 1min)
42°C x 1min } X2
72°C x 3min)
followed by
92°C x 1min)
60°C x 1min } X30
72°C x 2min)

followed by
72°C x 10 min

When the cycles were completed the reaction was allowed to cool slowly to room temperature. 5-10µl of the reaction was tested on a 2% agarose gel and blotted. PCR products were cloned using the restriction enzyme sites located within the primer ends.

7.13 Protein analysis

7.13.1 Preparation of nuclear protein extracts

The following procedure is an adaptation of the method described by Lichtsteiner *et al.*, (1987). Freshly dissected tissue was placed in ice cold PBS. The tissue was minced with sharp scissors and placed in 12 ml of homogenisation buffer (2M Sucrose, 10mM Hepes-KOH pH7.6, 15mM KCl, 0.15mM spermine, 0.5mM spermidine, 2mM EDTA, 10% Glycerol, 0.5mM DTT, 20µg/ml Trasylol, 1mM Benzamidine, 1mM PMSF). The tissue was homogenised using a teflon/glass motor driven homogeniser. The homogenate was laid upon a 0.5 ml cushion of 2M sucrose homogenisation buffer in SW41 Ultra clear centrifuge tubes (Beckman) and was spun at 22000rpm for 45 minutes in a pre-chilled (0°C) SW41 rotor in an L8 ultracentrifuge. The solid disk of unbroken cells on the top of the tube was removed with a spatula. The sucrose solution was carefully removed by aspiration. The nuclei were resuspended in lysis buffer (10mM Hepes-KOH ph 7.6, 100mM KCL, 3mM $MgCl_2$, 0.1mM EDTA, 10% Glycerol, 1mM DTT, 0.1mM PMSF, 2µg/ml Trasylol) using as small a volume as possible and stored at -20°C or -80°C for longer periods

7.13.2 Western blot analysis of protein

Sample preparation

To the nuclei samples described in 7.13.1, an equal volume of 2x SDS gel-loading buffer was added (100 mM Tris-HCl (pH 7.8), 200mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol). The samples were boiled for 5 minutes, and genomic

DNA sheared by sonication for 1 minute. The samples were centrifuged at 13,000 rpm for 10 minutes and the supernatant transferred to a fresh tube.

SDS-polyacrylamide gel electrophoresis

Proteins were separated by electrophoresis on a 10% vertical SDS-polyacrylamide gel (Laemmli 1970, Sambrook *et al.* 1989). 10% acrylamide gels (bis-acrylamide:acrylamide 1:29 w/v) were prepared in 375 mM Tris (pH 8.8), 0.1% SDS and polymerised with 0.1% ammonium persulfate (final concentration), 0.1% TEMED (final concentration). Gels were run at 15V/cm in Tris-glycine electrophoresis buffer (25mM Tris, 250mM glycine (pH 8.3), 0.1% SDS) for approximately 2-3 hours.

Transfer of proteins

Proteins were blotted on to Immobilon filters using Biorad electroblotting apparatus. Immobilon filters were rinsed in methanol prior to use, washed in ddH_2O and soaked in blot buffer (19 mM glycine, 25 mM Tris (pH 8.3), 20% methanol) for 30 minutes. The blot was set up as follows: Anode (+ ve), porous pad, 3MM paper wetted in blot buffer, Immobilon filter, SDS-polyacrylamide gel, 3MM paper wetted in blot buffer, porous pad, cathode (-ve) . The gel was blotted at 300 mA for 90 minutes.

Antibody labelling

After blotting filters were rinsed in blocking buffer (50mM Tris pH8.0, 150mM NaCl, 0.02% sodium azide, 0.05% Tween 20, 3% BSA) and then incubated in blocking buffer at room temperature for at least 2 hours. Filters were then incubated for 2 hours with purified rabbit polyclonal antibodies #1341 or #1343, 1:200 dilution in 10 ml blocking buffer. The filters were washed 6 x 10 minutes with 50mM Tris pH8.0, 150mM NaCl, 0.05% Tween 20, and subsequently incubated with anti-rabbit IgG alkaline phosphatase conjugate (Sigma) at a dilution of 1:1000 in 10 ml blocking buffer. Filters were washed again 6 x 10 minutes in 50mM Tris pH8.0, 150mM NaCl, 0.05% Tween 20, and stained in 50 ml of 0.2M Tris-HCl (pH 9.0), 10 mM $MgCl_2$ containing 1mg/ml naphthol,

3mg/ml Vaciamine Blue RT (Sigma). The filter was incubated at room temperature until bands became visible. The colouring reaction was stopped by prolonged rinsing in H_20.

7.14 β-galactosidase staining

7.14.1 Staining of mouse embryos and tissue

E8.5-e14.5 mouse embryos were recovered from uteri as follows. The day in which the vaginal mating plug was seen was taken as day 0.5. The pregnant female mouse was killed by either cervical dislocation or with an overdose of CO_2. The uteri were isolated by dissection and cut between each embryo swelling using a fine pair of scissors. The uterine muscle was removed carefully using two pairs of watch makers forceps (Dumont No. 5). For e8.5-10.5 day embryos the decidua were isolated and dissected according to Hogan *et al.* (1994). Embryos were transferred using pasteur pipettes with enlarged openings, which were made by scoring the pipette close to the widening of the neck with a diamond tip pencil. The narrow shaft was broken away and the opening polished in a flame to remove sharp edges. The pipette was controlled by means of a mouth piece and tubing. Embryos were placed into individual wells of a 24 well plate containing PBS-A for further processing. Yolk sacs were also isolated and used for DNA analysis. Larger embryos, ie e12.5 and e14.5 these were dissected from the yolk sac and placenta and transferred using a curved spatula. In this case the placenta was used for DNA analysis from these larger embryos. Freshly dissected postnatal and adult tissue for whole mount staining was collected into PBS-A.

Embryos and tissues were mildly fixed in 1 % formaldehyde, 0.5% glutaraldehyde in PBS-A. Embryos up to 9.5 days old were incubated in 1 ml of fixative for 5 minutes at room temperature. Embryos between 10.5 and 12.5 days old were fixed for 15 minutes. Larger embryos and adult tissue were fixed for 30-60 minutes in ten times the sample volume. After fixation samples were washed three times in rinse buffer (100mM sodium phosphate buffer, pH7.0, 2mM $MgCl_2$, 0.02% Nonidet P40, 0.01% w/v deoxycholate). Washing times applied as for fixation times. The final wash was replaced with the X-gal staining solution (1mg/ml X-gal, 5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$, 1mM EGTA, 100mM sodium phosphate buffer, pH7.0, 2mM $MgCl_2$, 0.02% Nonidet P40, 0.01% w/v deoxycholate). X-gal was dissolved in

dimethylformamide and kept as a 50mg/ml stock at -20°C. The embryos and tissue samples were stained for 4 hours at 37°C in the dark and then transferred to 4°C for overnight staining.

7.14.2 Histological analysis of X-gal stained embryos and tissue samples

Embryos or tissue samples were paraffin embedded and sectioned using a modification of the protocol described by Gossler and Zachgo (1993).

After staining the samples were washed in rinse buffer, fixed further in 4% paraformaldehyde in PBS at 4°C for 2 hours and kept in 70% ethanol until further use. Dehydration times and embedding in paraffin were as described in Gossler and Zachgo (1993). 5-10μM sections were cut on an American Optical 820 rotary microtome. Sections were de-waxed using Paraclear and embedded using Paramount (Earthsafe industries, Inc.). Counterstaining with haemoxilin and eosin was as described in Gossler and Zachgo (1993).

7.14.3 4-methylumbelliferyl-β-D-galactoside (MUG) analysis of β-galactosidase activity

Dissected tissue was placed in liquid nitrogen. When required the sample was thawed on ice and resuspended in 500μl of rinse buffer (100mM sodium phosphate buffer, pH7.0, 2mM $MgCl_2$, 0.02% Nonidet P40, 0.01% w/v deoxycholate). The sample was further disrupted by freeze-thawing in liquid nitrogen, this was repeated twice. Protein extracts (10μl) were incubated with 20 μl of a solution of 1 mM 4-methylumbelliferyl-β-D-galactoside (MUG) in rinse buffer and the samples were incubated at 37°C for 30 minutes to 1 hour. Subsequently 500μl of 0.5M carbonate buffer, pH 10.7, was added and the release of 4-methylumbelliferone was assayed in a Fluoroskan fluorometer (excitation wavelength 355nm, emission wavelength 480nm). Protein content was measured using the BCA colometric assay (Smith *et al.*, 1985) and measured at 562nm with a BSA (bovine serum albumin) standard.

References

Abeliovich, A., Chen, C., Goda, Y., Silva, A.J., Stevens, C.F. and Tonegawa, S. 1993. Modified hippocampal long-term potentiation in PKCγ-mutant mice. Cell **75**:1253-1262.

Acton, R.T., Blankenhorn, E.P., Douglas, T.C., Owen, R.D., Hilgers, J., Hoffman, H.A. and Boyse, E.A. 1973. Variations among sublines of inbred AKR mice. Nature New Biol. **245**:8-10.

Acton, R.T., Morris, R.J. and Williams, A.F. 1974. Estimation of the amount and tissue distribution of rat Thy-1.1 antigen. Eur. J. Immunol. **4**:110-115.

Altaba, A.R. 1993. Induction of axial patterning of the neural plate: planar and vertical signals. J. Neurobiol. **24**:1276-1304.

Altman, J. and Bayer, S.A. 1981. Development of the brain stem in the rat. V. Thymidine-radiographic study of the time of origin of neurons in the mid-brain tegmentum. J. Comp. Neurol. **198**:677-716.

Amzel, L.M. and Poljak, R.J. 1979. Three dimensional structure of immunoglobulins. Ann. Rev. Biochem. **48**:961-997.

Anderson, R.G. 1993. Caveolae: where incoming and outgoing messages meet. Proc. Natl. Acad. Sci. USA **90**:10909-10913.

Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Mathews, B.W. 1981. Structure of the cro repressor from bacteriophage lambda and its interactions with DNA. Nature **390**:754-758.

Aronin, N., Chase, K., Sagar, S.M., Sharp, F.R. and Difiglia, M. 1991. N-Methyl-D-aspartate receptor activation in the neostriatum increases c-fos and fos-related antigens selectively in medium sized neurons. Neuroscience **44**:409-420.

Askew, R.G., Doetschman, T. and Lingrel, J.B. 1993. Site-drected point mutations in embryonic stem cells: a gene targeting Tag-and-Exchange strategy. Mol.Cell.Biol. **13**:4115-4124.

Auffrey, C. and Rougeon, O. 1979. Purification of mouse immunoglobulin heavy chain messenger RNAs from total myeloma tumour RNA. Eur.J.Biochem. **107**:303-314.

Axelrod, H.R. 1984. Embryonic stem cells derived from blastocysts by a simplified method. Dev. Biol. **101**:225-228.

Axelrod, H.R. and Lader, E. 1983. A simplified method for obtaining embryonic stem cells from blastocysts. In Silver, L.M., Martin, G.R. and Strickland, S. (ed.), Cold Spring Harbour Conferences on Cell Proliferation. Vol 10 Teratocarcinomas and stem cells. Cold Spring Harbour Press.

Banerji, J., Rusconi, S. and Schaffner, W. 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299-308.

Baniahmad, A., Kohne, A.C. and Renkawitz, R. 1992. A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. EMBO J. **11**:1015.

Baniahmad, A., Steiner, C., Kohne, A.C. and Renkawitz, R. 1990. Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. Cell **61**:505-514.

Barboni, E., Gormley, A.M., Rivero, B.P., Vidal, M. and Morris, R.J. 1991. Activation of T lymphocytes by cross-linking of glycophospholipid-anchored Thy-1 mobilizes seperate pools of intracellular second messengers to those induced by the antigen-receptor/CD3 complex. Immunology **72**:457-463.

Barclay, A.N. and Hyden, H. 1978. Localisation of the Thy-1 antigen in rat brain and spinal cord by immunofluorescence. J. Neurochem. **31**:1375-1391.

Barclay, A.N., Letarte-Muirhead, M. and Williams, A.F. 1975. Purification of the Thy-1 molecule from rat brain. Biochem. J. **151**:699-706.

Barclay, A.N., Letarte-Muirhead, M. and Williams, A.F. 1976. Chemical characterisation of the Thy-1 glycoproteins from the membranes of rat thymocytes and brain. Nature **263**:563-567.

Baribault, H. and Kemler, R. 1989. Embryonic stem cells culture and gene targeting in transgenic mice. Mol. Biol. Med. **6**:481-492.

Baribault, H., Penner, J., Iozzo, R. and Wilson-Heiner, M. 1994. Colorectal hyperplasia and inflammation in keratin 8-deficient FVB/N mice. Genes Dev. **8**:2964-2973.

Baribault, H., Price, J., Miyai, K. and Oshima, R.G. 1993. Mid-gestational lethality in mice lacking keratin **8**. Genes Dev. **7**:1191-1202.

Basch, R.S. and Berman, J.W. 1982. Thy-1 determinants are present on many murine haematopoietic cells other than T-cells. Eur. J. Immunol. **12**:359-364.

Beck, E., Ludwig, G., Auserwald, E.A., Reiss, B. and Schaller, H. 1982. Nucleotide sequence and exact localisation of the neomycin phosphotransferase gene from transposon Tn5. Gene **19**:327-336.

Beech, J.N., Morris, R.J. and Raisman, G. 1983. Density of Thy-1 on axonal membrane of different rat nerves. J. Neurochem. **41**:411-417.

Bejsovec, A. and Martinez-Arias, A. 1991. Roles of *wingless* in patterning the larval epidermis of *Drosophila*. Development **113**:471-485.

Bellefroid, E., Lecocq, P.J., Benhida, A., Poncelet, D.A., Belayew, A. and Martial, J.A. 1989. The human genome contains hundreds of genes coding zinc finger proteins of the *Kruppel* type. DNA **8**:377-387.

Bellefroid, E.J., Marine, J.-C., Ried, T., Lecocq, P.J., Riviere, M., Amemiya, C., Poncelet, D.A., Coulie, P.G., de Jong, P., Szpirer, C., Ward, D.C. and Martial, J.A. 1993. Clustered organisation of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. EMBO J. **12**:1363-1374.

Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O. and Martial, J.A. 1991. The evolutionary conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. USA **88**:3608-3612.

Ben-Hattar, J. and Jiricny, J. 1988. Methylation of single CpG dinucleotides with a promoter element of the Herpes simplex virus *tk* gene reduces its transcription *in vivo*. Gene **65**:219-227.

Benham, F.J., Hodgkinson, S. and Davies, K. 1984. A glyceraldehyde-3-phosphate dehydrogenase pseudogene on the short arm of the human X chromosomes defines a multigene family. EMBO J. **3**:2635-2640.

Bennett, R.J., Dunderdale, H.J. and West, S.C. 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. Cell **74**:1021-1031.

Berg, J.M. 1988. Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc.Natl.Acad.Sci.USA **99-102**:

Berinstein, N., Pennell, N., Ottaway, C. and Shulman, M.J. 1992. Gene replacement with one-sided homologous recombination. Mol. Cell. Biol. **12**:360-367.

Bernet-Grandaud, A., Ouazana, R., Morle, F. and Godet, J. 1992. A method improving the efficiency of the positive-negative selection used to isolate homologous recombinants. Nucl. Acids Res. **20**:6417-6418.

Bienz, M. and Pelham, H.R. 1986. Heat shock regulatory element function as an inducible enhancer in the Xenopus *hsp70* gene and when linked to a heterologous promoter. Cell **45**:753-760.

Bird, A. 1992. The essentials of DNA methylation. Cell **70**:5-8.

Bird, A.P. 1986. CpG-rich islands and the function of DNA-methylation. Nature **321**:209-213.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. **7**:1513-1523.

Bliss, T.V. and Collingridge, G.L. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. Nature **361**:31-39.

Blom van Assendelft, G., Hanscombe, O., Grosveld, F. and Greaves, D.R. 1989. The β-globin dominant control region activates homologous and heterologous promoters in a tissue specific manner. Cell **56**:969-977.

Bohmann, D., Bos, T.J., Adman, A., Nisshimura, T., Vogt, P.K. and Tjian, R. 1987.

Human proto-oncogene C-*jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP1. Science **238**:1386-1392.

Bollag, R.J., Waldman, A.S. and Liskay, M. 1989. Homologous recombination in mammalian cells. Annu. Rev. Genet. **23**:199-225.

Bonifer, C., Vidal, M., Grosveld, F. and Sippel, A.E. 1990. Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. **9**:2843.

Bourtchuladze, R., Frenguelli, B., Blendy, J., Cioffi, D., Schutz, G. and Silva, A. 1994. Deficient long-term memory in mice with a targeted mutation of the cAMP-Responsive element-binding protein. Cell **79**:59-68.

Brace, H.M., Jefferys, J.G. and Mellanby, J. 1985. Long-term changes in hippocampal physiology and learning ability of rats after intrahippocampal tetanus toxin. J.Physiol. **368**:343-357.

Bradley, A. 1987. Production and analysis of chimeric mice. p. 113-151. In Robertson, E.J. (ed.), Teratocarcinomas and embryonic stem cells: A practical approach. IRL Press, Oxford.

Bradley, A. 1990. Embryonic stem cells: proliferation and differentiation. Curr. Opin. Cell Biol. **2**:1013-1017.

Bradley, A., Evans, M., Kaufman, M. and Robertson, E. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature **309**:255-256.

Brand, A.H., Breeden, L., Abraham, J., Sternglanz, R. and Nasmyth, K. 1985. Characterisation of a "silencer" in yeast: a DNA sequence with properties opposite

to those of a transcriptional enhancer. Cell **41**:41-48.

Brandeis, R., Brandys, Y. and Yehuda, S. 1989. The use of the Morris water maze in the study of memory and learning. Int. J. Neurosci. **48**:26-69.

Braun, R.E., Lo, D., Pinkert, C.A., Widera, G., Flavell, R.A., Palmiter, R.D. and Brinster, R.L. 1990. Infertility in male transgenic mice: disruption of sperm development by HSV-tk expression in postmeiotic germ cells. Biol. Repro. **43**:684-693.

Braun, T., Rudnicki, M.A., Arnold, H.-H. and Jaenisch, R. 1992. Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development. Cell **71**:369-382.

Brent, R. and Ptashne, M. 1985. A eukaryotic transcriptional activator bearing the DNA specificty of a prokaryotic repressor. Cell **43**:729.

Brinster, R.L. 1974. The effect of cells transferred into the mouse blastocyst on subsequent development. J. Exp. Med. **140**:1049-1056.

Brinster, R.L., Braun, R.E., Lo, D., Avarbock, M.R. and Oram, F. 1989. Targeted correction of a major histocompatibility class II E_{α} gene by DNA microinjected into mouse eggs. Proc. Natl. Acad. Sci. USA **86**:7087-7091.

Bronson, S.K. and Smithies, O. 1994. Altering mice by homologous recombination using embryonic stem cells. J.Biol.Chem **269**:27155-27158.

Brown, R.S., Sander, C. and Argos, P. 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett. **186**:271-274.

Bueler, H., Aguzzi, A., Sailer, A., Greiner, R.-A., Autemried, P., Aguet, M. and

Weissmann, C. 1993. Mice devoid of PrP are resistant to scrapie. Cell **73**:1339-1347.

Bueler, H., Fisher, M., Lang, Y., Bluethman, H., Lipp, H.-P., DeArmond, S.J., Pruisner, S.B., Aguet, M. and Weissmann, C. 1992. The neuronal cell surface protein PrP is not essential for normal development and behaviour of the mouse. Nature **356**:577-582.

Burke, P.S. and Wolgemuth, D.J. 1992. *Zfp-37*, a new murine zinc finger encoding gene, is expressed in a developmentally regulated pattern in the male germ line. Nucl. Acids Res. **20**:2827-2834.

Cairns, B.R., Kim, Y.J., Sayre, M.H., Laurent, B.C. and Kornberg, R.D. 1994. A multi-subunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, SNF6 gene products isolated from yeast. Proc. Natl. Acad. Sci.USA **91**:1950-1954.

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. and Housman, D.E. 1990. Isolation and characterisation of a zinc-finger polypeptide gene at the human chromosome 11 Wilm's tumor locus. Cell **60**:509-520.

Campbell, D.G., Gagnon, J., Reid, K.B. and Williams, A.F. 1981. Rat brain Thy-1 glycoprotein. Biochem. J. **195**:15-30.

Campbell, D.G., Williams, A.F., Bayley, P.M. and Reid, K.M. 1979. Structural similarities between Thy-1 antigen from brain and immunoglobulin. Nature **282**:341-342.

Cao, L., Alani, E. and Nancy, K. 1990. A pathway for generation and processing of double-strand breaks during meiotic recombination in S.cerevisiae. Cell **61**:1089-1101.

Capecchi, M.R. 1980. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. Cell **22**:479-488.

Capecchi, M.R. 1989. The new mouse genetics: altering the genome by gene targeting. Trends Genet. **5**:70-76.

Carlson, M. and Laurent, B.C. 1994. The SNF/SWI family of global transcriptional activators. Curr. Opin. Cell Biol. **6**:296-402.

Carlsson, S.R. 1985. Changes in glycan branching and sialylation of the Thy-1 antigen during normal differentiation and mouse T-lymphocytes. Biochem J. **226**:519-525.

Carlsson, S.R. and Stigbrand, T. 1984. Partial characterisation of the oligosaccharides of mouse thymocytes Thy-1 glycoprotein. Biochem J. **221**:379-392.

Carpenter, E.M., Goddard, J.M., Chisaka, O., Manley, N. and Capecchi, M.R. 1993. Loss of *Hox-a1 (Hox-1.6)* function results in the reorganisation of the murine hindbrain. Development **118**:1063-1077.

Cedar, H. 1988. DNA methylation and gene activity. Cell **53**:3-4.

Chang, H.-C., Seki, T., Moriuchi, T. and Silver, J. 1985. Isolation and characterisation of mouse Thy-1 genomic clones. Proc. Natl. Acad. Sci. USA **82**:3819-3823.

Chavrier, P., Lemaire, P., Revelant, O., Bravo, R. and Charnay, P. 1987. Charcterisation of a mouse multigene family that encodes zinc finger structures. Mol. Cell. Biol. **8**:1319-1326.

Chavrier, P., Zerila, M., Lemaire, P., Almendral, J., Bravo, R. and Charnay, P. 1988. A gene encoding a protein with zinc fingers is activated during G0/G1 transition in

cultured cells. EMBO J. **7**:29-35.

Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K. and Tjian, R. 1994. Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. Cell **79**:93-105.

Chen, S., Botteri, F., van der Putten, H., Landel, C.P. and Evans, G.A. 1987. A lymphoproliferative abnormality associated with inappropriate expression of the Thy-1 antigen in transgenic mice. Cell **51**:7-19.

Chisaka, O., Musci, T. and Capecchi, M.R. 1992. Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene *Hox-1.6*. Nature **355**:516-520.

Choo, Y. and Klug, A. 1994a. Toward a code for interactions of zinc fingers with DNA: selection of randomised fingers displayed on phage. Proc. Natl. Acad. Sci. USA **91**:11163-11167.

Choo, Y. and Klug, A. 1994b. Selection of DNA binding sites for zinc fingers using rationally randomised DNA reveals coded interactions. Proc. Natl. Acad. Sci. USA **91**:11168-11172.

Chowdhury, K., Deutsch, U. and Gruss, P. 1987. A multigene family encoding several "finger" structures is present and differentially active in mammalian genomes. Cell **48**:771-778.

Chowdhury, K., Dressler, G., Breier, G., Deutsch, U. and Gruss, P. 1988. The primary structure of the murine multifinger gene *mkr2* and its specific expression in developing and adult neurons. EMBO J. **7**:1345-1353.

Clarke, A.R., Maandag, E.R., van Roon, M., van der Lugt, N.M., van der Valk, M., Hooper, M.L., Berns, A. and te Riele, H. 1992. Requirement for a functional *Rb-1* gene in murine development. Nature **359**:328-330.

Clerc, R.G., Corcoran, L.M., LeBowitz, J.H., Baltimore, D. and Sharpe, P.A. 1988. The B-cell specific Oct-2 protein contains POU-box and homeobox type domains. Genes Dev. **2**:1570-1581.

Cockerill, P.N. and Garrard, W.T. 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell **44**:273.

Cohen, F.E., Novotny, J., Sternberg, M.J., Campbell, D.G. and Williams, A.F. 1981. Analysis of structural similarities between brain Thy-1 antigen and immunoglobulin domains. Biochem J. **195**:31-40.

Cohen, S.N., Chang, A.C.Y. and Hsu, L. 1972. Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110-2114.

Collinge, J., Whittington, M.A., Sidle, K.C., Smith, C.J., Palmer, M.S., Clarke, A.R. and Jefferys, J.G. 1994. Prion protein is necessary for normal synaptic function. Nature **370**:295-297.

Conzelmann, A., Spiazzi, A. and Bron, C. 1987. Glycolipid anchors are attached to Thy-1 glycoprotein rapidly after translation. Biochem J. **246**:605-610.

Cote, J., Quin, J., Workman, J.L. and Peterson, C.L. 1994. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science **265**:53-60.

Cotmore, S.F., Crowhurst, S.A. and Waterfield, M.D. 1981. Purification of Thy-1 related glycoproteins from human brain and fibroblasts: comparison between these molecules and murine glycoproteins carrying Thy-1.1 and Thy-1.2 antigens. Eur. J. Immunol. **11**:597-603.

Courey, A.J., Holtzman, D.A., Jackson, S.P. and Tjian, R. 1989. Synergistic activation by the glutamine rich domains of human transcription factor Sp1. Cell **59**:827-836.

Courey, A.J. and Tjian, R. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55**:887-898.

Cremer, H., Lange, R., Christoph, A., Plomann, M., Vopper, G., Roes, J., Brown, R., Baldwin, S., Kraemer, P., Scheff, S., Barthels, D., Rajewsky, K. and Willie, W. 1994. Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning. Nature **367**:455-459.

Cress, W.D. and Triezenberg, S.J. 1991. Critical structural elements of the VP16 transcriptional activation domain. Science **251**:87.

Crew, A.J., Clark, J., Fisher, C., Gill, S., Grimer, R., Chand, A., Shipley, J., Gusterson, B.A. and Cooper, C.S. 1995. Fusion of *SYT* to two genes, *SSX1* and *SSX2*, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma. EMBO J. **14**:2333-2340.

Crosby, S.D., Puetz, J.J., Simburger, K.S., Fahrner, T.J. and Milbrandt, J. 1991. The early response gene NGFI-C encodes a zinc finger transcriptional activator and is a member of the GCGGGGGCG (GSG) element-binding protein family. Mol. Cell. Biol. **11**:3835-3841.

Cross, S.H. and Bird, A.P. 1995. CpG islands and genes. Curr. Opin. Genet. Dev.

5:309-314.

Crowley, C., Spencer, S.D., Nishimura, M.C., Chen, K.S., Pitts-Meek, S., Armanini, M.P., Ling, L.H., McMahon, S.B., Shelton, D.L., Levinson, A.D. and Philips, H.S. 1994. Mice lacking nerve growth factor display perinatal loss of sensory and sympathetic neurons yet develop basal forebrain cholinergic neurons. Cell **76**:1001-1012.

Dalchau, R. and Fabre, J.W. 1979. Identification and unusual tissue disribution of the canine and human homologues of Thy-1 (θ). J. Exp. Med. **149**:576-591.

DasGupta, C., Wu, A.M., Kahn, R., Cunningham, R.P. and Radding, C.M. 1981. Concerted strand exchange and formation of Holliday structures by E.coli RecA protein. Cell **25**:507-516.

Davies, A.M., Lee, K.F. and Jaenisch, R. 1993. p75-deficient trigeminal sensory neurons have an altered response to NGF but not to other neurotrophins. Neuron **11**:565-575.

Davis, C.A. and Joyner, A.L. 1988. Expression patterns of the homeo box-containing genes *En-1* and *En2* and the proto-oncogene *int-1* diverge during mouse development. Genes Dev. **2**:1736-1744.

Davis, C.A., Noble-Topham, S.E., Rossant, J. and Joyner, A.L. 1988. Expression of the homeo box-containing gene *En-2* delineates a specific region of the developing mouse brain. Genes Dev. **2**:361-371.

Decker, C.J. and Parker, R. 1994. Mechanisms of mRNA degradation in eukaryotes. Trends Biochem. Sci. **19**:336-340.

Deng, C. and Capecchi, M.R. 1992. Reexamination of gene targeting as a function of the

extent of homology between the targeting vector and the target locus. Mol. Cell. Biol. **12**:3365-3371.

Deng, C., Thomas, K.R. and Capecchi, M.R. 1993. Location of crossovers during gene targeting with insertion and replacement vectors. Mol. Cell. Biol. **13**:2134-2140.

Desjarlais, J.R. and Berg, J.M. 1993. Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. Proc. Natl. Acad. Sci.USA **90**:2256-2260.

Detlott, P.J., Lewis, J., John, S.W.M., Shehee, R.W., Langenbach, R., Maeda, N. and Smithies, O. 1994. Deletion and replacement of the mouse adult β-globin genes by a "Plug and Socket" repeated targeting strategy. Mol. Cell. Biol. **14**:6936-6943.

Deuschle, U., Meyer, W.K. and Thiesen, H.-J. 1995. Tetracycline-reversible silencing of eukaryotic promoters. Mol. Cell. Biol. **15**:1907-1914.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S. and Smithies, O. 1987. Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature **330**:576-578.

Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R. 1985. The *in-vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J. Embryol. Exp. Morph. **87**:27-45.

Doherty, P., Singh, A., Rimon, G., Bolsover, S. and Walsh, F. 1993. Thy-1 antibody-triggered neurite outgrowth requires an influx of calcium into neurons via N- and L-type calcium channels. J. Cell Biol. **122**:1818-189.

Dolle, P., Lufkin, T., Krumlauf, R., Mark, M., Duboule, D. and Chambon, P. 1993. Local

alterations of *Krox-20* and *Hox* gene expression in the hindbrain suggests lack of rhombomeres 4 and 5 in the homozygote null *Hoxa-1 (Hox-1.6)* mutant embryos. Proc. Natl. Acad. Sci. USA **90**:7666-7670.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Butel, J.S. and Bradley, A. 1992. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature **356**:215-221.

Douglas, T.C. 1973. Occurence of theta-like antigen in rats. J. Exp. Med. **136**:1054-1062.

Driever, W., Ma, J., Nusselein-Volhard, C. and Ptashne, M. 1989. Rescue of bicoid mutant *Drosophila* embryos by bicoid fusion proteins containing heterologous activating sequences. Nature **342**:149-154.

Dunderdale, H.J., Benson, F.E., Parsons, C.A., sharples, G.J., Lloyd, R.G. and West, S.C. 1991. Formation and resolution of recombination intermediates by *E.coli* RecA and RuvC proteins. Nature **354**:506-510.

Dynan, W.S. 1986. Promoters of housekeeping genes. Trends Genet. **6**:196-197.

Dynlacht, B.D., Hoey, T. and Tjian, R. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell **66**:563-576.

Echelard, Y., Epstein, D.J., St-Jaques, B., Shen, L., Mohler, J., McMahon, J.A. and McMahon, A.P. 1993. Sonic hedgehog, a member of a family of putative signalling molecules, is implicated in the regulation of CNS polarity. Cell **75**:1417-1430.

Edelman, G.M. 1970. The covalent structure of a human γ^G-immunoglobulin XI. Functional implications. Biochemistry **9**:3197-3205.

Enver, T., Zhang, J.-W., Papayannopoulou, T. and Stamatoyannopoulous, G. 1988. DNA methylation: a secondary event in globin gene switching? Genes Dev. **2**:698-706.

Ernfors, P., Lee, K.F. and Jaenisch, R. 1994a. Mice lacking brain-derived neurotrophic factor develop with sensory deficits. Nature **368**:147-150.

Ernfors, P., Lee, K.F., Kucera, J. and Jaenisch, R. 1994b. Lack of neurotrophin-3 leads to deficiencies in the peripheral nervous system and loss of limb proprioceptive afferents. Cell **77**:503-512.

Evans, G.A., Ingraham, H.A., Lewis, K., Cunningham, K., Seki, T., Moriuchi, T., Chang, H.-C., Silver, J. and Hyman, R. 1984. Expression of the Thy-1 glycoprotein gene by DNA mediated gene transfer. Proc. Natl. Acad. Sci. USA **81**:5532-5536.

Evans, M.J. and Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature **292**:154-156.

Fairall, L., Schwabe, J.W., Chapman, L., Finch, J.T. and Rhodes, D. 1993. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. Nature **366**:483-487.

Feinberg, A.P. and Vogelstein, B. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal.Biochem. **132**:6-13.

Felsenfeld, G. 1992. Chromatin as an essential part of transcriptional mechanism. Nature **355**:219.

Felsenfeld, G. and McGhee, J.D. 1986. Structure of the 30 nm chromatin fiber. Cell **44**:375-377.

Ferguson, M.A. 1992. Glycosyl-phosphatidylinositol membrane anchors: the tale in the tail. Biochem. Soc.Trans. **20**:243-256.

Ferrin, L.J. and Camerini-Otero, R.D. 1991. Selective cleavage of human DNA: RecA-assisted restriction endonuclease (RARE) cleavage. Science **254**:1494-1496.

Fiering, S., Kim, C.G., Epner, E.M. and Groudine, M. 1993. An "in-out" strategy using gene targeting and FLP recombinase for the functional dissection of complex DNA regulatory elements: analysis of the β-globin locus control region. Proc. Natl. Acad. Sci. USA **90**:8469-8473.

Fleming, J., Rogers, M.J.C., Brown, S.D. and Steel, K.P. 1994. Linkage analysis of the *whirler* deafness gene on mouse chromosome 4. Genomics **21**:42-48.

Folger, K.R., Wong, E.A., Wahl, G. and Capecchi, M.R. 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. **5**:59-69.

Foulkes, N.S., Schlotter, F., Pevet, P. and Sassone-Corsi, P. 1993. Pituitary hormone FSH directs CREM functional switching during spermatogenesis. Nature **362**:264-267.

Fourney, R.M., Miyakoshi, J., Day III, R.S. and Paterson, M.C. 1988. Northern blotting: efficient RNA staining and transfer. Bethesda Res. Lab. Focus **10**:5-7.

Frankel, A.D. and Kim, P.S. 1991. Modular structure of transcription factors: implications for gene regulation. Cell **65**:717-719.

Freedman, L.P., Luisi, B.F., Korszun, Z.R., Basavappa, R., Sigler, P.J. and Yamamoto, K.R. 1988. The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. Nature **334**:543-546.

Freemont, P.S., Lane, A.N. and Sanderson, M.R. 1991. Structural aspects of protein-DNA recognition. Biochem. J. **278**:1.

French, P.W. and Jeffrey, P.L. 1986. Partial charcterisation of chicken Thy-1 glycoprotein by monolclonal antibodies. J. Neurosci. Res. **16**:479-489.

Freter, R.R., Irminger, J.-C., Porter, J.A., Jones, S.D. and Stiles, C. 1992. A novel 7-nucleotide motif located in 3'untranslated sequences of the immediate-early gene set mediates platelet-derived growth factor induction of the JE gene. Mol. Cell. Biol. **12**:5288-5300.

Friedrich, G. and 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.

Fujimoto, T. 1993. Calcium pump of the plasma membrane is localised in caveolae. J.Cell Biol. **120**:1147-1157.

Fujimoto, T., Nakade, S., Miyawaki, A., Mikoshiba, K. and Ogawa, K. 1993. Localisation of inositol 1,4,5-triphosphate receptor-like protein in plasmalemmal caveolae. J. Cell Biol. **119**:1507-1513.

Fujimura, F.K., Deininger, P.L., Friedmann, T. and Linney, E. 1981. Mutation near the polyoma DNA replication origin permits productive infection of F9 embryonal carcinoma cells. Cell **23**:809-814.

Garnett, D., Barclay, A.N., Carmo, A.M. and Beyers, A.D. 1993. The association of the protein tyrosine kinases $p56^{lck}$ and $p60^{fyn}$ with the glycosyl phosphatidylinositol-anchored proteins Thy-1 and CD48 in rat thymocytes is dependent on the state of cellular activation. Eur. J. Immunol **23**:2540-2544.

Gasser, S.M. and Laemmli, U.K. 1987. A glimpse at chromosomal order. Trends Genet. **3**:16-22.

Gearing, D.P., Gough, N.M., King, J.A., Hilton, D.J., Nicola, N.A., Simpson, R.J., Nice, E.C., Kelso, A. and Metcalf, D. 1987. Molecular cloning and expression of cDNA encoding a murine leukaemia inhibitory factor (LIF). EMBO J. **6**:3995-4002.

Giese, K.P., Martini, R., Lemke, G., Soriano, P. and Schachner, M. 1992. Mouse P_0 gene disruption leads to hypomyelination, abnormal expression of recognition molecules and degeneration of myelin and axons. Cell **71**:565-576.

Giguere, V., Isobe, K.-I. and Grosveld, F. 1985. Structure of the murine Thy-1 gene. EMBO J. **4**:2017-2024.

Gill, G., Pascal, E., Tseng, Z.H. and Tjian, R. 1994. A glutamine-rich hydrophobic patch in transcription factor SP1 contacts the $dTAF_{II}110$ component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91**:192-196.

Giniger, E. and Ptashne, M. 1987. Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. Nature **330**:670-672.

Goodman, C.S. 1994. The likeness of being: phylogenetically conserved molecular mechanisms of growth cone guidance. Cell **78**:353-356.

Goodman, C.S. and Shatz, C.J. 1993. Developmental mechanisms that generate precise patterns of neuronal connectivity. Cell **72**:77-98.

Goodrich, J.A., Hoey, T., Thut, C.J., Admon, A. and Tjian, R. 1993. Drosophila $TAF_{II}40$ interacts with both a VP16 activation domain and the basal transcription factor

TFIIB. Cell **75**:519-530.

Goodwin, L.O., Rocha, A.J. and Basch, R.C. 1986. Isolation of cell lines possessing functional and serological properties resembling those of thymocyte precursors. Nature **323**:166-169.

Gossler, A. and Zachgo, J. 1993. Gene and enhancer trap screens in ES cell chimeras. p. 181-227. In Joyner, A.L. (ed.), Gene targeting: a practical approach. IRL Press, Oxford.

Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci.USA **89**:5547-5551.

Grant, S.G. 1994. Gene targeting and synaptic plasticity. Curr. Opin. Neurobiol. **4**:687-692.

Grant, S.G. and Silva, A.J. 1994. Targeting learning. Trends Neurosci. **17**:71-75.

Grant, S.G.N., O'Dell, T.J., Karl, K.A., Stein, P.L., Soriano, P. and Kandel, E.K. 1992. Impaired long-term potentiation, spatial learning and hippocampal development in *fyn* mutant mice. Science **258**:1903-1910.

Graziadei, P.P. and Graziadei, G.A. 1978. The olfactory system: A model for the study of neurogenesis and axon regeneration in mammals. p. 131-153. In Cotman, C.W. (ed.), Neuronal plasticity. Raven Press, New York.

Greaves, D.R., Wilson, F.D., Lang, G. and Kioussis, D. 1989. Human CD2 3'-flanking sequences confer high-level, T-cell-specific, position independent gene expression in transgenic mice. Cell **56**:979-986.

Gross, D. and Garrard, W. 1988. Nuclease hypersensitive sites in chromatin. Ann. Rev. Biochem. **57**:159.

Grosveld, F., Blom van Assendelft, G., Greaves, D.G. and Kollias, G. 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell **51**:975-985.

Gu, H., Zou, Y.-R. and Rajewsky, K. 1993. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell **73**:1155-1164.

Guddat, U., Bakken, A.H. and Pieler, T. 1990. Protein-mediated nuclear export of RNA: 5S rRNA containing small RNPs in Xenopus oocytes. Cell **60**:619-628.

Guillemot, F., Lo, L.-C., Johnson, J.E., Auerbach, A.B., Anderson, D.J. and Joyner, A.L. 1993. Mammalian *achaete-scute* Homolog-1 is required for early development of olfactory and autonomic neurons. Cell **75**:463-476.

Hahn, S. 1993. Structure(?) and function of acidic transcription activators. Cell **72**:481.

Ham, J., Steger, G. and Yaniv, M. 1994. Cooperativity *in vivo* between the E2 transactivator and TATA box binding protein depends on core promoter structure. EMBO J. **13**:147-157.

Hamann, A., Arndt, R., Klein, P. and Thiele, H.-G. 1980. Isolation and characterisation of the thymus-brain antigen (analogous to Thy-1 antigen) from human brain. Biochem J. **187**:403-412.

Han, K. and Manley, J. 1993. Transcriptional repression by the *Drosophila even-skipped* protein: definition of a minimal repression domain. Genes Dev. **7**:491-503.

Handyside, A.H., O'Neill, G.T., Jones, M. and Hooper, M.L. 1989. Use of BRL-conditioned medium in combination with feeder layers to isolate a diploid embryonal stem cell line. Roux Archiv. Dev. Biol. **198**:48-56.

Haraguchi, T., Fisher, S., Olofsson, S., Endo, T., Groth, D., Tarantino, A., Borchelt, D.R., Teplow, D., Hood, L., Burlingame, A., Lycke, E., Kobata, A. and Pruisner, S.B. 1989. Asparagine-linked glycosylation of the scrapie and cellular prion proteins. Arch. Biochem. Biophys. **274**:1-13.

Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. Nature **353**:715.

Hasty, P. and Bradley, A. 1993. Gene targeting vectors for mammalian cells. p. 1-31. In Joyner, A.L. (ed.), Gene targeting: a practical approach. IRL press, Oxford.

Hasty, P., Bradley, A., Morris, J., Edmondson, D.G., Venuti, J.D., Olson, E.N. and Klein, W.H. 1993. Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. Nature **364**:501-506.

Hasty, P., Rivera-Perez, J. and Bradley, A. 1991a. The length of homology required for gene targeting in embryonic stem cells. Mol. Cell. Biol. **11**:5586-5591.

Hasty, P., Rivera-Perez, J., Chang, C. and Bradley, A. 1991b. Target frequency and integration pattern of insertion and replacement vectors in embryonic stem cells. Mol. Cell. Biol. **11**:4509-4517

Hasty, P., Ramirez-Solis, R., Krumlauf, R. and Bradley, A. 1991c. Introduction of a subtle mutation into the Hox 2.6 locus in embryonic stem cells. Nature **351**:234-246.

Hasty, P., Rivera-Perez, J. and Bradley, A. 1992. The role and fate of DNA ends for homologous recombination in embryonic stem cells. Mol. Cell. Biol. **12**:2464-2474.

Haupt, Y., Alexander, W.S., Barri, G., Klinken, S.P. and Adams, J.M. 1991. Novel zinc-finger gene implicated as *myc* collaborator by retrovirally accelerated lymphomagenesis in Eμ-*myc* transgenic mice. Cell **65**:753-763.

He, H.-T., Naquet, P., Caillol, D. and Pierres, M. 1991. Thy-1 supports adhesion of mouse thymocytes to thymic epithelial cells through a Ca^{2+}-independent mechanism. J. Exp. Med. **173**:515-518.

Heemskerk, J., DiNardo, S., R., K. and O'Farrell, P.H. 1991. Multiple nodes of *engrailed* regulation in the progression towards cell fate determination. Nature **352**:404-410.

Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkun, G. and Horovitz, H.R. 1988. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2 and *Caenorhabditis elegans* unc-86 gene products. Genes Dev. **2**:1513-1516.

Hill, R.L., Delaney, R., Fellows, R.E. and Lebovitz, H.E. 1966. The evolutionary origins of the immunoglobulins. Proc. Natl. Acad. Sci.USA **56**:1762-1769.

Hinnen, A., Hicks, J.B. and Fink, G.R. 1978. Transformation in yeast. Proc. Natl. Acad. Sci. USA **75**:1929-1933.

Hoey, T., Weinzierl, R.O., Gill, G., Chen, J.-L., Dynlacht, B.D. and Tjian, R. 1993. Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. Cell **72**:247-260.

Hogan, B., Beddington, R., Constantini, F. and Lacy, E. 1994. Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory Press.

Holliday, R. 1964. A mechanism for gene conversion in fungi. Genet. Res. **5**:282-304.

Homans, S.W., Ferguson, M.A., R.A., D., Rademacher, T.W., Anand, R. and Williams, A.F. 1988. Complete structure of the glycosyl phosphatidylinositol membrane anchor of rat Thy-1 glycoprotein. Nature **333**:269-272.

Honda, B.M. and Roeder, R.G. 1980. Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell **22**:119-126.

Hood, L., Kronenberg, M. and Hunkapiller, T. 1985. T cell antigen receptors and the immunoglobulin supergene family. Cell **40**:225-229.

Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature **326**:292-295.

Hooper, N.M. 1992. More than just a membrane anchor. Current Biology **2**:617-619.

Hosseini, R., Marsh, P., Pizzey, J., Leonard, L., Ruddy, S., Bains, S. and Dudley, K. 1994. Restricted expression of a zinc finger protein in male germ cells. J. Mol. Endocrinol. **13**:157-165.

Hseih, P., Meyn, M.S. and Camerini-Otero, C.S. 1986. Partial purification and characterisation of a recombinase from human cells. Cell **44**:885-894.

Huang, J.D., Schwyter, D.H., Shirokawa, J.M. and Courey, A.J. 1993b. The interplay between multiple enhancer and silencer elements defines the pattern of decapentaplegic expression. Genes Dev. **7**:694.

Huang, S., Hendriks, W., Althage, A., Hemmi, S., Bluethmann, H., Kamijo, R., Vilcek, J., Zinkernagel, R.M. and Aguet, M. 1993a. Immune response in mice that lack the interferon-γ receptor. Science **259**:1742-1745.

Hynes, M., Poulsen, K., Tessier-Lavigne, M. and Rosenthal, A. 1995. Control of neuronal diversity by the floor plate: contact-mediated induction of midbrain dopaminergic neurons. Cell **80**:95-101.

Iguchi-Ariga, S.M. and Schaffner, W. 1989. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev. **3**:612-619.

Imbalzano, A.N., Kwon, H., Green, M.R. and Kingston, R.E. 1994. Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature **370**:481-485.

Ingraham, H.A. and Evans, G.A. 1986. Characterisation of two atypical promoter elements and alternate mRNA processing in the mouse Thy-1.2 glycoprotein gene. Mol. Cell. Biol. **6**:2923-2931.

Ingraham, H.A., Lawless, G.M. and Evans, G.A. 1986. The mouse Thy-1.2 glycoprotein gene: complete sequence and identification of an unusual promoter. J. Immunology **136**:1482-1489.

Ish-Horowicz, D. and Burke, J.F. 1981. Rapid and efficient cosmid cloning. Nucl. Acids Res. **9**:2989-2998.

Ishida, A., Asano, H., Hasegawa, M., Koseki, H., Ono, T., Yoshida, M.C., Taniguchi, M. and Kanno, M. 1993. Cloning and chromosome mapping of the human *mel-18* gene which encodes a DNA-binding protein with a new "RING-finger" motif. Gene **129**:249-255.

Itakura, K., Hutton, J.J., Boyse, E.A. and Old, L.J. 1972. Genetic linkage relationships of loci specifying differential alloantigens in the mouse. Transplantation **13**:239-243.

Jacks, T., Fazeli, A., Schmitt, E., Bronson, R.T., Goodell, M.A. and Weinberg, R.A. 1992. Effects of an Rb mutation in the mouse. Nature **359**:295-300.

Jarman, A.P. and Higgs, D.R. 1988. Nuclear scaffold attachment sites in the globin gene complexes. EMBO J. **7**:3337.

Jaynes, J.B. and O'Farrell, P.H. 1991. Active repression of transcription by the engrailed homeodomain protein. EMBO J. **10**:1427-1433.

Jones, K.R., Farinas, I., Backus, C. and Reichardt, L.F. 1994. Targeted disruption of the brain-derived neurotrophic gene perturbs brain and sensory neuron but not motor neuron development. Cell **76**:989-1000.

Jones, N. 1990. Transcriptional regulation by dimerisation: two sides to an incestuous relationship. Cell **61**:9.

Joyner, A.L. 1991. Gene targeting and gene trap screens using embryonic stem cells: new approaches to mammalian development. Bioessays **13**:649-656.

Joyner, A.L. and Guillemot, F. 1994. Gene targeting and development of the nervous system. Curr. Opin. Neurobiol. **4**:37-42.

Joyner, A.L., Herrup, K., Auerbach, A., Davis, C.A. and Rossant, J. 1991. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the *En-2* homeobox. Science **251**: 1239-1243.

Joyner, A.L., Skarnes, W.C. and Rossant, J. 1989. Production of a mutation in mouse En-2 gene by homologous recombination in embryonic stem cells. Nature **338**:153-156.

Jung, S., Rajewsky, K. and Radbruch, A. 1993. Shutdown of class switch recombination

by deletion of a switch region control element. Science **259**:984-987.

Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. 1984. A short amino acid sequence able to specify nuclear location. Cell **39**:499-509.

Kandel, E.R., Schwartz, J.H. and Jessell, T.M. 1991. Principles of neural science. Elsevier, New York.

Keller, G.-A., Siegel, M.W. and Caras, I.W. 1992. Endocytosis of glycophospholipid-anchored transmembrane forms of CD4 by different endocytic pathways. EMBO J. **11**:863-874.

Kemshead, J.T., Carner, K.R., Maiarz, F.R. and Greaves, M.F. 1982. Human Thy-1 expression on the cell surface of neuronal and glial cells. Brain Res. **236**:451-461.

Kennedy, M.B. and Marder, E. 1992. Cellular and molecular mechanisms of neuronal plasticity. p. 463-495. In Hall, Z.W. (ed.), An introduction to molecular neurobiology. Sinauer Assoc, Inc., Sunderland, MA.

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M. and Tsui, L.C. 1989. Identification of the cystic fibrosis gene: genetic analysis. Science **245**:1073-1080.

Kilby, N.J., Snaith, M.R. and Murray, J.A. 1993. Site-specific recombinases: tools for genome engineering. Trends Genet. **9**:413-421.

Killeen, N., Stuart, S.G. and Littman, D.R. 1992. Development and function of T cells in mice with a disrupted CD2 gene. EMBO J. **11**:4329-4336.

Kim, H.-S. and Smithies, O. 1988. Recombinant fragment assay for gene targeting based

on the polymerase chain reaction. Nucl. Acid. Res. **16**:8887-8903.

Kim, J.G. and Hudson, L.D. 1992. Novel member of the zinc finger superfamily: a C_2-HC finger that recognises a glia-specific gene. Mol. Cell. Biol. **12**:5632-5639.

Kitamura, D., Roes, J., Kuhn, R. and Rajewsky. 1991. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin μ chain gene. Nature **350**:423-426.

Klein, R., Silos-Santiago, I., Smeyne, R.J., Lira, S.A., Brambilla, R., Bryant, S., Zhang, L., Snider, W.D. and Barbacid, M. 1994. Disruption of the neurotrophin-3 receptor gene *trkC* eliminates 1a muscle afferents and results in abnormal movements. Nature **368**:249-251.

Klein, R., Smeyne, R.J., Wurst, W., Long, L.K., Auerbach, B.A., Joyner, A.L. and Barbacid, M. 1993. Targeted disruption of the *trkB* neurotrophin receptor gene results in nervous system lesions and neonatal death. Cell **75**:113-122.

Klevit, R.E. 1991. Recognition of DNA by Cys_2His_2 zinc fingers. Science **253**:1367.

Klocke, W., Koster, M., Hille, S., Bouwmeester, T., Bohm, S., Pieler, T. and Knochel, W. 1994. The FAR domain defines a new *Xenopus laevis* zinc finger protein subfamily with specific RNA homopolymer binding activity. Biochim. Biophys. Acta **1217**:81-89.

Klug, A. and Lutter, L.C. 1981. The helical periodicity of DNA on the nucleosome. Nucl. Acids Res. **17**:4267.

Knochel, W., Poting, A., Koster, M., El-Baradi, T., Nietfeld, W., Bouwmeester, T. and Pieler, T. 1989. Evolutionary conserved modules associated with zinc fingers in

Xenopus laevis. Proc. Natl. Acad. Sci. USA **86**:6097-6100.

Kollias, G., Spanopoulou, E., Grosveld, F., Ritter, M., Beech, M. and Morris, R. 1987. Differenetial regulation of a Thy-1 gene in transgenic mice. Proc. Natl. Acad. Sci. USA **84**:1492-1496.

Kornberg, R.D. and Lorch, Y. 1995. Interplay between chromatin structure and transcription. Curr. Opin. Cell. Biol. **7**:371-375.

Koster, M., Kuhn, U., Bouwmeester, T., Nietfield, W., El-Baradi, T., Knochel, W. and Pieler, T. 1991. Structure, expression and *in vitro* functional charcterisation of a novel RNA binding zinc finger protein from *Xenopus*. EMBO J. **10**:3087-3093.

Koster, M., Pieler, T., Poting, A. and Knochel, W. 1988. The finger motif defines a multigene family represented in the maternal mRNA of *Xenopus laevis* oocytes. EMBO J. **7**:1735-1741.

Kouzarides, T. and Ziff, E. 1988. The role of the leucine zipper in the *fos-jun* interaction. Nature **336**:646-651.

Kowalczykowski, S. 1991. Biochemistry of genetic recombination: energetics and mechanism of DNA strand exchange. Annu. Rev. Biophys. Chem. **20**:539-575.

Kretzschmar, H.A., Pruisner, S.B., Stowring, L.E. and DeArmond, S.J. 1986. Scrapie prion proteins are synthesised in neurones. Am. J. Path. **122**:1-5.

Kroczek, R.A., Gunter, K.C., Germain, R.C. and Shevach, E.M. 1986. Thy-1 functions as a signal transduction molecule in T-lymphocyte and transfected B-lymphocytes. Nature **322**:181-184.

Krumlauf, R., Marshall, H., Studer, M., Nonchev, S., Sham, M.H. and Lumsden, A. 1993. *Hox* homeobox genes and regionalisation of the nervous system. J. Neurobiol. **24**:1328-1340.

Krumlauf, R. 1994. *Hox* genes in vertebrate development. Cell **78**:191-201.

Krumlauf, R., Chapman, V.M., Hammer, R.E., Brinster, R. and Tilghman, S.M. 1986. Differential expression of α-fetoprotein genes on the inactive X chromosome in extraembryonic and somatic tissues of a transgenic mouse. Nature **319**:224-226.

Kuchel, P.W., Campbell, D.G., Barclay, A.N. and Williams, A.F. 1978. Molecular weights of the Thy-1 glycoprotein from rat thymus and brain in the prescence and abscence of deoxycholate. Biochem. J. **169**:411-417.

Kucherlapati, R.S., Eves, E.M., Song, K.-Y., Morse, B.S. and Smithies, O. 1984. Homologous recombination between plasmids in mammalian cells can be enhanced by treatment of input DNA. Proc. Natl. Acad. Sci. USA **81**:3153-3157.

Kuehn, M.R., Bradley, A., Robertson, E.J. and Evans, M.J. 1987. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature **326**:295-298.

Kwon, H., Imbalzano, A.N., Khavarl, P.A., Kingston, R.E. and Green, M.R. 1994. Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. Nature **370**:477-481.

Laemmli, U.K. 1970. Cleavage of structural proteins during assembly at the head of bacteriophage T4. Nature **227**:680-685.

Laemmli, U.K., Cheng, S.M., Adolph, K.W., Paulson, J.R., Brown, J.A. and Baumbach,

W.R. 1977. Metaphase chromosome structure: the role of non-histone proteins. Cold Spring Harbor Symp. **42**:351-360.

Laird, P.W., Zijderveld, A., Linders, M.A., Rudnicki, R., Jaenisch, R. and Berns, A. 1991. Simplified mammalian DNA isolation procedure. Nucl. Acids Res. **19**:4293-4294.

Lakso, M., Sauer, B., Mosinger, B., Lee, E.J., Manning, R.W., Yu, S.-H., Mulder, K.L. and Westphal, H. 1992. Targeted oncogene activation by site-specific recombination in transgenic mice. Proc. Natl. Acad. Sci. USA **89**:6232-6236.

Landshultz, W.H., Johnson, P.F. and McKnight, S.L. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759-1764.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. 1992. CpG islands as gene markers in the human genome. Genomics **13**:1095-1107.

Le Mouellic, H., Lallemand, Y. and Brulet, P. 1992. Homeosis in the mouse induced by a null mutation in the *Hox-3.1* gene. Cell **69**:251-264.

Le Mouellic, H.L., Lallemand, Y. and Brulet, P. 1990. Targeted replacement of the homeobox gene *Hox-3.1* by the *Escherichia coli lacZ* in mouse chimeric embryos. Proc. Natl. Acad. Sci. USA **87**:4712-4716.

Lederman, B. and Burki, K. 1991. Establishment of a germ-line competent C57BL/6 embryonic stem cell line. Exp. Cell Res. **197**:254-258.

Lee, E.Y.-H., Chang, C.-Y., Hu, N., Wang, Y.-C.J., Lai, C.-C., Herrup, K., W-H., L. and Bradley, A. 1992. Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. Nature **359**:288-294.

Leiffer, D., Lipton, S.A., Barnstable, C.J. and Masland, R.H. 1984. Monoclonal antibodies to Thy-1 enhances regeneration of processes by rat retinal ganglion cells in culture. Science **224**:303-306.

Lemke, G. and Axel, R. 1985. Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. Cell **40**:501-508.

Lesley, J.F. and Lennon, V.A. 1977. Transitory expression of Thy-1 antigen in skeletal muscle development. Nature **268**:163-165.

Letarte-Muirhead, M., Barclay, A.N. and Williams, A.F. 1975. Purification of the Thy-1 molecule, a major cell surface glycoprotein of rat thymocytes. Biochem. J. **151**:685-697.

Leuther, K.K., Salmeron, J.M. and Johnston, S.A. 1993. Genetic evidence that the activation domain of GAL4 does not require acidity and may form a beta sheet. Cell **72**:575.

Levine, M. and Hoey, T. 1988. Homeobox proteins as sequence-specific transcription factors. Cell **55**:537-540.

Lewin, B. 1990. Commitment and activation at Pol II promoters: a tail of protein-protein interactions. Cell **61**:1161-1164.

Lewin, B. 1994. Chromatin and gene expression: constant questions but changing answers. Cell **79**:397-406.

Lewis, C.D. and Laemmli, U.K. 1982. High order metaphase chromosome structure: evidence for metalloprotein interactions. Cell **29**:171-181.

Lewis, J. and Bird, A. 1991. DNA methylation and chromatin structure. FEBS Lett. **285**:155-159.

Li, C., Tropak, M.B., Gerlai, R., Clapoff, S., Abranow-Newerly, W., Trapp, B., Peterson, A. and Roder, J. 1994b. Myelination in the abscence of myelin-associated glycoprotein. Nature **369**:747-750.

Li, E., Bester, T.H. and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell **69**:915-926.

Li, J., Knight, J.D., Jackson, S.P., Tjian, R. and Botchan, M.R. 1991. Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. Cell **65**:493-505.

Li, Y., Erzurumlu, R.S., Chen, C., Jhaveri, S. and Tonegawa, S. 1994a. Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. Cell **76**:427-437.

Licht, J.D., Grossel, M.J., Figge, J. and Hansen, U.M. 1990. Drosophila *Kruppel* protein is a transcription factor. Nature **346**:76-79.

Lichtsteiner, S., Wuarin, J. and Schibler, U. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell **51**:963-973.

Lin, F.-L., Sperle, K. and Sternberg, N. 1984. Homologous recombination in mouse L-cells. Cold Spring Harbor Symp. Quant. Biol. **49**:139-143.

Lin, F.L., Sperle, K. and Sternberg, N. 1985. Recombination in mouse L cells between DNA introduced into cells and homologous chromosomal sequences. Proc. Natl. Acad. Sci. USA **82**:1391-1395.

Lipton, S.A., Leiffer, D. and Barnstable, C.J. 1992. Selectivity of Thy-1 monoclonal antibodies in enhancing neurite outgrowth. Neurosci. Lett. **137**:75-77.

Lovering, R. and Trowsdale, J. 1991. A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell lines. Nucl. Acids Res. **19**:2921-2928.

Low, M.G. 1989a. Glycosyl-phosphatidylinositol: a versatile anchor for cell surface proteins. FASEB J. **3**:1600-1608.

Low, M.G. 1989b. Anchoring of Thy-1 to the cell membrane. p.71-83. In Reif, A.E. and Schlesinger, M. (ed.), Cell surface antigen Thy-1. Immunology, Neurology and Therapeutic applications. Marcel Dekker Inc., New York.

Low, M.G. and Kincade, P.W. 1985. Phosphatidylinositol is the membrane-anchoring domain of the Thy-1 glycoprotein. Nature **318**:62-64.

Lufkin, T., Dierich, A., LeMeur, M., Mark, M. and Chambon, P. 1992. Disruption of the *Hox-1.6* homeobox gene results in defects in a region corresponding to its rostral domain of expression. Cell **668**:1105-1119.

Lyon, M.F. and Searle, A.G. 1989. Genetic variants and strains of the laboratory mouse. Oxford University Press, Oxford.

Ma, J. and Ptashne, M. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847-853.

MacGibbon, G.A., Lawlor, P.A., Bravo, R. and Dragunow, M. 1994. Clozapine and haloperidol produce a differential pattern of immediate early gene expression in rat caudate-putamen, nucleus accumbens, lateral septum and islands of Calleja. Mol. Brain Res. **23**:21-32.

Mahanthappa, N.K. and Patterson, P.H. 1992. Thy-1 involvement in neurite outgrowth: perturbation by antibodies, phospholipase C and mutation. Dev. Biol. **150**:47-59.

Manson, J., West, J.D., Thomson, V., McBride, P., Kaufman, M.H. and Hope, J. 1992. The prion protein gene: a role in mouse embryogenesis? Development **115**:117-122.

Mansour, M.H. and Cooper, E.L. 1984a. Purification and characterisation of *Rana pipiens* brain Thy-1 glycoprotein. J.Immunol. **132**:2515-2523.

Mansour, M.H. and Cooper, E.L. 1984b. Serological and partial molecular characterisation of a Thy-1.2 homologue in tunicates. Eur.J.Immunol. **14**:1031-1039.

Mansour, S.L., Thomas, K.R. and Capecchi, M.R. 1988. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature **336**:348-352.

Mansour, S.L., Thomas, K.R., Deng, C. and Capecchi, M.R. 1990. Introduction of a *lacZ* reporter gene into the mouse *int-2* locus by homologous recombination. Proc. Natl. Acad. Sci. USA **87**:7688-7692.

Margolin, J.F., Friedman, J.R., Meyer, W.K.-H., Vissing, H., Thiesen, H.-J. and Rauscher III, F.J. 1994. Kruppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA **91**:4509-4513.

Mark, M., Lufkin, T., Vonesch, J.-L., ruberte, E., Olivo, J.-C., Dolle, P., Gorry, P., Lumsden, A. and Chambon, P. 1993. Two rhombomeres are altered in *Hoxa-1* mutant mice. Development **119**:319-338.

Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos culture in medium condtioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA

78:7634-7638.

Mazarakis, N., Michalovich, D., Karis, A., Grosveld, F. and Galjart, N. 1995. *Zfp-37* is a member of the KRAB zinc finger gene family and is expressed in neurons of the developing and adult CNS. Submitted

McGinnis, W., Levine, M., Hafer, E., Kuriowa, A. and Gehring, W. 1984. A conserved sequence in homeotic genes of *Drosophila* Antennapedia and Bithorax complexes. Nature **308**:428-433.

McKay, D.B. and Steitz, T.A. 1981. Structutre of the catabolite gene activator protein at 2.9A resolution suggests buinding to left-handed B-DNA. Nature **290**:744-749.

McKenzie, J.L. and Fabre, J.W. 1981a. Studies with a monoclonal antibody on the distribution of Thy-1 in the lymphoid and extracellular connective tissue of the dog. Transplantation **31**:275-282.

McKenzie, J.L. and Fabre, J.W. 1981b. Human Thy-1: unusual localisation and possible functional significance in lymphoid tissue. J. Immun. **126**:843-850.

McKnight, S. 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucl. Acids Res. **8**:5949-5964.

McLaren, A. 1976. Mammalian chimeras. Cambridge University Press, Cambridge.

McMahon, A.P. and Bradley, A. 1990. The *Wnt-1 (int-1)* proto-oncogene is required for development of a large region of the mouse brain. Cell **62**:1073-1085.

McMahon, A.P., Joyner, A.L., Bradley, A. and McMahon, J.A. 1992. The midbrain-hindbrain phenotype of *Wnt-1*$^{-}$/*Wnt-1*$^{-}$ mice results from stepwise deletion

of *engrailed*-expressing cells by 9.5 days postcoitum. Cell **69**:581-595.

Melton, D.W. 1990. The use of targeting to develop animal models for human genetic diseases. Biomed. Soc. Transac. **18**:1035-1039.

Melton, D.W., Konecki, D.S., Brennand, J. and Caskey, T.C. 1984. Structure, expression and mutation of the hypoxanthine phosphoribosyltransferase gene. Proc. Natl. Acad. Sci. USA **81**:2147-2151.

Mermod, N., O'Neill, E.A., Kelley, T.J. and Tjian, R. 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58**:741-753.

Meselson, M.S. and Radding, C.M. 1975. A general model for genetic recombination. Proc. Nat. Acad. Sci. USA **77**:503-507.

Millen, K.J., Wurst, W., Herrup, K. and Joyner, A.L. 1994. Abnormal embryonic cerebellar development and patterning of postnatal foliation in two mouse *Engrailed-2* mutants. Development **120**:695-706.

Miller, C.K. and Temin, H.M. 1983. High-efficiency ligation and recombination of DNA fragments by vertebrate cells. Science **220**:606-609.

Miller, I.J. and Bieker, J.J. 1993. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the *kruppel* family of nuclear proteins. Mol. Cell. Biol. **13**:2776-2786.

Miller, J., McLachlan, A.D. and Klug, A. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. **4**:1609-1614.

Millhouse, O.E. 1987. Granule cells of the olfactory tubercule and the question of the islands of Calleja. J. Comp. Neurol. **265**:1-24.

Mintz, B. and Ilmense, I. 1975. Normal genetically mosaic mice produced from malignant teratocarcinoma cells. Proc. Natl. Acad. Sci. USA **75**:3585-3589.

Mirkovitch, J., Mirault, M.-E. and Laemmli, U.K. 1984. Organisation of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell **39**:223-232.

Mitchell, P.J. and Tjian, R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371-378.

Molina, C.A., Foulkes, N.S., Lalli, E. and Sassone-Corsi, P. 1993. Inducibility and negative autoregulation of CREM: an alternative promoter directs expression of ICER, an early response repressor. Cell **75**:875-886.

Molina, T.J., Kishihara, K., Siderovski, D.P., Timms, E., Wakeham, A., Paige, C.J., Hartmann, K.-U., Veillette, A., Davidson, D. and Mak, T.W. 1992. Profound block in thymocyte development in mice lacking $p56^{lck}$. Nature **357**:161-164.

Mombaerts, P., Clarke, A.R., Hooper, M.L. and Tonegawa, S. 1991. Creation of a large genomic deletion at the T-cell antigen receptor β-subunit locus in mouse embryonic stem cells by gene targeting. Proc. Natl. Acad. Sci. USA **88**:3084-3087.

Mombaerts, P., Iacomini, J., Johnson, R., Hierrup, K., Tonegawa, S. and Papaioannou, V.E. 1992. RAG-1-deficient mice have no mature B and T lymphocytes. Cell **68**:869-877.

Monaghan, P., Warburton, M.U., Monaghan, P. and Ritter, M.A. 1983. Topographical

arrangement of basement membrane proteins in lactating rat mammary gland. Comparison of the distribution of type IV collagen, laminin, fibronectin and Thy-1 at the ultrastructural level. Proc. Natl. Acad. Sci. USA **80**:3344-3349.

Montag, D., Giese, K.P., Bartsch, U., Martini, R., Lang, Y., Bluthmann, H., Karthigasan, J., Kirschner, D.A., Wintergerst, E., Nave, K.-A., Zielasek, J., Toyka, K.V., Lipp, H.-P. and Schachner, M. 1994. Mice deficient for myelin-associated glycoprotein show subtle abnormalities in myelin. Neuron **13**:229-246.

Moreau, J.F., Donaldson, D.D., Bennett, F., Witek-Giannotti, J., Clark, S.C. and Wong, G.G. 1988. Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. Nature **336**:690-692.

Morgan, J.I. and Curran, T. 1989. Stimulus-transcription coupling in neurons: role of immediate early genes. Trends Neurosci. **12**:459-462.

Morgan, J.I. and Curran, T. 1991. Stimulus-transcription coupling in the nervous system: involvement of the inducible proto-oncogenes *fos* and *jun*. Annu. Rev. Neurosci. **14**:421-451.

Mori, N., Stein, R., Sigmund, O. and Anderson, D.J. 1990. A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. Neuron **4**:583-594.

Moriuchi, T., Chang, H.-C., Denome, R. and Silver, J. 1983. Thy-1 cDNA sequence suggests a novel regulatory mechanism. Nature **301**:80-82.

Morris, R.G., Anderson, E., Lynch, G.S. and Baudry, M. 1986. Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature **319**:774-775.

Morris, R.G.M. 1981. Spatial localisation does not require the presence of local cues. Learning Motiv. **12**:239-260.

Morris, R.G.M., Garrud, P., Rawlins, J.N. and O'Keefe, J. 1982. Place navigation in rats with hippocampal lesions. Nature **297**:681-683.

Morris, R.J. 1985. Thy-1 in developing nervous tissue. Dev. Neurosci. **7**:133-160.

Morris, R.J. 1992. Thy-1, the enigmatic extrovert on the neuronal surface. Bioessays **14**:715-722.

Morris, R.J. and Barber, P.C. 1983. Fixation of Thy-1 in nervous tissue for immunohistochemistry: a quantitative assessment of the effect of different fixation conditions upon retention of antigenicity and the cross-linking of Thy-1. J. Histochem. Cytochem. **31**:263-274.

Morris, R.J. and Beech, J. 1984. Differential expression of Thy-1 on the various components of connective tissue of rat nerve during postnatal development. Dev. Biol. **102**:32-42.

Morris, R.J. and Beech, J. 1987. Sequential expression of OX-2 and Thy-1 on the various glycoproteins on the neuronal surface during development: An immunohistochemical stufy of rat cerebellum. Dev. Neurosci. **9**:33-44.

Morris, R.J. and Grosveld, F.G. 1989. Expression of Thy-1 in the nervous system of the rat and mouse. p. 121-148. In Reif, A.E. and Schlesinger, M. (ed.), Cell surface antigen Thy-1. Immunology, Neurology and Therapeutic applications. Marcel Dekker Inc., New York.

Moser, M., Colello, R.J., Pott, U. and Oesch, B. 1995. Developmental expression of the

prion protein gene in glial cells. Neuron **14**:509-517.

Mountford, P., Zevnik, B., Duwel, A., nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I. and Smith, A. 1994. Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression. Proc. Natl. Acad. Sci. USA **91**:4303-4307.

Muller, U. and Kypta, R. 1995. Molecular genetics of neuronal adhesion. Curr. Opin. Neurobiol. **5**:36-41.

Muller-Sieburg, C.E., Whitlock, C.A. and Weissman, I.L. 1986. Isolation of two early B lymphocyte progenitors from mouse marrow: a committed pre-pre-B cell and clonogenic Thy-1^{lo} haematopoietic stem cell. Cell **44**:653-662.

Mullis, K.B. and Faloona, F.A. 1987. Specific synthesis of DNA *in-vitro* via a polymerase catalysed chain reaction. Methods Enzymol. **155**:335-350.

Murre, C., McCaw, P.S. and Baltimore, D. 1989. A new DNA binding and dimerisation motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell **56**:777-783.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S., Nonaka, I. and Nabeshima, Y.-I. 1993. *Myogenin* gene disruption results in perinatal lethality because of severe muscle defect. Nature **364**:532-535.

Nagy, A., Gocza, E., Diaz, E.M., Prideaux, V.R., Ivanyi, E., Markkula, M. and Rossant, J. 1990. Embryonic stem cells alone are able to support fetal development in the mouse. Development **110**:815-821.

Nagy, A. and Rossant, J. 1993. Production of completely ES cell-derived fetuses. p.

147-179. In Joyner, A.L. (ed.), Gene targeting: a practical approach. IRL Press, Oxford.

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. and Roder, J.C. 1993. Derivation of completely cell culture-derived mice from early-passage embryonic srem cells. Proc. Natl. Acad. Sci. USA **90**:8424-8428.

Nardelli, J., Gibson, T.J., Vesque, C. and Charnay, P. 1991. Base sequence discrimination by zinc-finger DNA-binding domains. Nature **349**:175-178.

Nasmyth, K.A. 1982. Molecular genetics of yeast mating type. Ann.Rev.Genet. **16**:439-500.

Nelki, D., Dudley, K., Cunningham, P. and Akhavan, M. 1990. Cloning and sequencing of a zinc-finger cDNA expressed in mouse testis. Nucl. Acids Res. **18**:3655.

Nichols, J., Evans, E.P. and Smith, A.G. 1990. Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. Development **110**:1341-1348.

Nomura, M., Takihara, Y. and Shimada, K. 1994. Isolation and characterisation of retinoic acid-inducible cDNA clones in F9 cells: one of the early inducible clones encodes a novel protein sharing highly homologous regions with a *Drosophila polyhomeotic* protein. Differentiation **57**:39-50.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. and Grosveld, F. 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature **316**:316-318.

O'Shea, E.K., Rutkowski, K.R. and Kim, P.S. 1989. Evidence that the leucine zipper is a

coiled coil. Science **243**:538.

Oas, T.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J. and Gronenborn, A.M. 1990. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. Biochemistry **29**:2891.

Oesch, B., Westaway, D., Walchi, M., McKinley, M.P., Kent, S.B., Aebersold, R., Barry, R.A., Teplow, D.B., Tempst, D.B., Hood, L.E., Pruisner, S.B. and Weissmann, C. 1985. A cellular gene encodes scrapie PrP 27-30 protein. Cell **40**:735-746.

Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J. and Gronenborn, A.M. 1993. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. Science **261**:438.

Ondek, B., Shepard, A. and Herr, W. 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**:1017-1025.

Ono, K., Tomasiewicz, H., Magnuson, T. and Rutishauser, U. 1994. N-CAM mutation inhibits tangential neuronal migration and is phenocopied by enzymatic removal of polysialic acid. Neuron **13**:595-609.

Orban, P.C., Chui, D. and Marth, J. 1992. Tissue- and site-specific DNA recombination in transgenic mice. Proc. Natl. Acad. Sci. USA **89**:6861-6865.

Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA **78**:4417-4421.

Pabo, C.O. and Lewis, M. 1982. The opperator binding domain of the lambda repressor: structure and DNA recognition. Nature **298**:443-447.

Pan, T. and Coleman, J.E. 1990. GAL4 transcription factor is not a "zinc finger" but forms a Zn(II)2Cys6 binuclear cluster. Proc. Natl. Acad. Sci.USA **97**:2077.

Pandolfi, P.P., Roth, M.E., Karis, A., Leonard, M.W., Dzierzak, E., Grosveld, F.G., Engel, J.D. and Lindenbaum, M.H. 1995. Targeted disruption of the *GATA3* gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. Nature Genet. **11**:40-44.

Papaioannou, V. and Johnson, R. 1993. Production of chimeras and genetically defined offspring. p. 107-146. In Joyner, A.L. (ed.), Gene targeting: a practical approach. IRL Press, Oxford.

Papaioannou, V.E., McBurney, M.W., Gardner, R.L. and Evans, M.J. 1975. Fate of teratocarcinoma cells injected into early mouse embryos. Nature **258**:70-73.

Parekh, R.B., Tse, A.G., Dwek, R.A., Williams, A.F. and Rademacher, T.W. 1987. Tissue specific N-glycosylation, site-specific oligosaccharide patterns and lentil lectin recognition of rat Thy-1. EMBO J. **6**:1233-1244.

Parkhurst, S.M., Harrison, D.A., Remington, M.P., Spana, C., Kelley, R.L., Coyne, R.S. and Corces, V.G. 1988. The *Drosophila su(Hw)* gene, which controls the phenotypic effect of gypsy transposable element encodes a putative DNA-binding protein. Genes Dev. **2**:1205-1215.

Paro, R. 1990. Imprinting a determined state into the chromatin of *Drosophila.* Trends Genet. **6**:416-421.

Parsons, C.A., Tsaneva, I., Lloyd, R.G. and West, S.C. 1992. Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. Proc. Natl. Acad. Sci. USA **89**:5452-5456.

Paul, L.C., Resenke, H.G., Milford, E.L. and Carpenter, C.B. 1984. Thy-1.1 in glomeruli of rat kidneys. Kidney Int. **25**:771-772.

Pavletich, N.P. and Pabo, C.O. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1A. Science **252**:809-817.

Pavletich, N.P. and Pabo, C.O. 1993. Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. Science **261**:1701-1707.

Pelham, H.R. and Brown, D.D. 1980. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc. Natl. Acad. Sci. USA **77**:4170-4174.

Pelletier, J. and Sonenberg, N. 1988. Internal initiation of translation of eukaryotic mRNA by a sequence derived from poliovirus RNA. Nature **334**:320-325.

Pengue, G., Calabro, V., Bartoli, P.C., Pagliuca, A. and Lania, L. 1994. Repression of transcriptional activity at a distance by the evolutionary conserved KRAB domain present in a subfamily of zinc finger proteins. Nucl. Acids Res. **22**:2908-2914.

Peterson, C.L., Dingwall, A. and Scott, M.P. 1994. Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. Proc. Natl. Acad. Sci. USA **91**:2905-2908.

Phi-Van, L. and Stratling, W.H. 1988. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J. **7**:655.

Phi-Van, L., von Kries, J.P., Ostertag, W. and Stratling, W.H. 1990. The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected

genes. Mol. Cell. Biol. **10**:2302.

Pieler, T. and Bellefroid, E. 1994. Perspectives on zinc finger protein function and evolution-an update. Mol. Biol. Rep. **20**:1-8.

Porter, A.C. 1989. Designer genomes. Technique **1**:53-65.

Pruisner, S.B. 1994. Molecular biology and genetics of prion diseases. Phil. Trans. R. Soc. Lond. **343**:447-463.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. Nature **335**:683-689.

Ptashne, M. and Gann, A.A. 1990. Activators and targets. Nature **346**:329-331.

Pugh, F. and Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell **61**:1187-1197.

Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R. and Bradley, A. 1993. Hoxb-4 (Hox-2.6) Mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. Cell **73**:279-294.

Rathjen, P.D., Toth, S., Willis, A., Heath, J.K. and Smith, A.G. 1990. Differentiation inhibiting activity is produced in matrix-associated and diffusible forms that are generated by alternate promoter usage. Cell **62**:1105-1114.

Reif, A.E. and Allen, J.M. 1964. The AKR thymic antigen and its distribution in leukemias and nervous tissue. J. Exp. Med. **120**:413-433.

Reif, A.E. and Allen, J.M. 1966a. Mouse thymic iso-antigens. Nature **209**:521-523.

Reif, A.E. and Allen, J.M. 1966b. Mouse nervous tissue iso-antigens. Nature **209**:523.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. 1984. Structure of the nucleosome core particle at 7 A resolution. Nature **311**:532-537.

Rigby, P.W., Dieckmann, M., Rhodes, C. and Berg, P. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237-251.

Riggs, A.D. and Pfeifer, G.P. 1992. X-chromosome inactivation and cell memory. Trends Genet. **8**:169-174.

Ritter, M.A., Gordon, L.K. and Goldschneider, I. 1978. Distribution and identity of Thy-1 bearing cells during ontogeny in rat haematopoietic and lymphoid tissue. J. Immunol. **121**:2463-2471.

Ritter, M.A., Morris, R.J. and Goldschneider, I. 1980. Hidden Thy-1 antigen in subpopulation of mouse bone marrow cells. Immunology **39**:375-384.

Ritter, M.A., Sauvage, C.A. and Cotmore, S.F. 1981. The human thymus microenviroment: in vivo identification of thymic nurse cells and other antigenically distinct subpopulations of epithelial cells. Immunol. **44**:439-446.

Ritter, M.A., Sauvage, C.A. and Delia, D. 1983. Human Thy-1 antigen: cell surface expression on early T- and B-cell lymphocytes. Immunology **49**:555-564.

Robertson, E.J. 1987. Isolation of embryonic stem cells. p. 71-112. In Robertson, E.J. (ed.), Teratocarcinomas and embryonic stem cells: a practical approach. IRL Press, Oxford.

Robertson, E.J., Bradley, A., Kuehn, M. and Evans, M. 1986. Germline transmission of genes introduced into cultured pluripotential cells by retroviral vectors. Nature **323**:445-448.

Robertson, L.M., Kerppola, T.K., Montserrat, V., Luk, D., Smeyne, R.J., Bocchiaro, C., Morgan, J.I. and Curran, T. 1995. Regulation of *c-fos* expression in transgenic mice requires multiple interdependent transcription control elements. Neuron **14**:241-252.

Robinson, P.J. 1991. Phosphatidylinositol membrane anchors and T-cell activation. Immunol. Today **12**:35-41.

Robinson, P.J., Millrain, M., Antoniou, J., Simpson, E. and Mellor, A.L. 1989. A glycophospholipid anchor is required for Qa-2-mediated T-cell activation. Nature **342**:85-87.

Roeder, R.G. 1991. The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trend Biochem.Sci. **16**:402-498.

Rosati, M., Marino, M., Franze, A., Tramontano, A. and Grimaldi, G. 1991. Members of the zinc finger gene family sharing a conserved N-terminal module. Nucl. Acids Res. **19**:5661-5667.

Roseman, R.R., Pirotta, V. and Geyer, P. 1992. The *su(Hw)* protein insulates expression of the *Drosophila melanogaster white* gene from chromosomal position effects. EMBO J. 435-442.

Rosenberg, V.B., Schroeder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. and Jackle, H. 1986. Structural homology of the product of *Drosophila Kruppel* gene with *Xenopus* transcription factor IIIA. Nature **319**:336-339.

Rossant, J. 1991. Gene disruptions in mammmals. Curr. Opin. Genet. Dev. **1**:236-240.

Rostas, J.A., Shevenan, T.A., Sinclair, C.M. and Jeffrey, P.L. 1983. The purification and characterisation of a Thy-1-like glycoprotein from chicken brain. Biochem. J. **213**:143-152.

Roth, D.B., Porter, T.N. and Wilson, J.H. 1985. Mechanisms of nonhomologous recombination in mammalian cells. Mol. Cell. Biol. **5**:2599-2607.

Roth, D.B. and Wilson, J.H. 1985. Relative rates of homologous and nonhomologous recombination in transfected DNA. Proc. Natl. Acad. Sci. USA **82**:3355-3359.

Roy, A.L., Malik, S., Meisterernst, M. and Roeder, R.G. 1993. An alternative pathway for transcription initiation involving TFII-I. Nature **365**:355-359.

Roy, A.L., Meisterernst, M., Pognonec, P. and Roeder, R.G. 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature **354**:245-248.

Rubinstein, M., Japon, M.A. and Low, M.J. 1993. Introduction of a point mutation into the mouse genome by homologous recombination in embryonic stem cells using a replacement type vector with a selectable marker. Nucl. Acid. Res. **21**:2613-2617.

Rudnicki, M.A., Braun, T., Hinuma, S. and Jaenisch, R. 1992. Inactivation of *MyoD* in mice leads to up-regulation of the myogenic HLH gene *Myf-5* and results in apparently normal muscle development. Cell **71**:383-390.

Rudnicki, M.A., Schnegelsberg, P.N.J., Stead, R.H., Braun, T., Arnold, H.-H. and Jaenisch, R. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. Cell **75**:1351-1359.

Rutishauser, U. 1992. NCAM and its polysialic acid moiety: a mechanism for pull/push regulation of cell interactions during development. Development Suppl.:99-104.

Rutishauser, U., Acheson, A., Hall, A.K., Mann, D.M. and Sunshine, J. 1988. The neural cell adhesion molecule (NCAM) as a regulator of cell-cell interactions. Science **240**:53-57.

Saitoh, Y. and Laemmli, U.K. 1994. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. Cell **76**:609-622.

Sakano, H., Rogers, J.H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. 1979. Domains and the hinge region of an immunoglobulin heavy chain are encoded in seperate DNA segments. Nature **277**:627-633.

Saltzman, A.G. and Weinmann, R. 1989. Promoter specificity and modulation of RNA polymerase II transcription. FASEB J. **3**:1723-1733.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press.

Santerre, R.F., Allen, N.E., Hobbs, J.N., Rao, R.N. and Schmidt, R.J. 1984. Expression of prokaryotic genes for hygromycin B and G418 resistance as dominant-selection markers. Gene **30**:147-156.

Sargiacomo, M., Sudol, M., Tang, Z. and Lisanti, M.P. 1993. signal transducing molecules and glycosyl-phosphatidylinositol-linked proteins form a caveolin-rich insoluble complex in MDCK cells. J.Cell Biol. **122**:789-807.

Sassone-Corsi, P. 1994. Rhythmic transcription and autoregulatory loops: winding up the biological clock. Cell **78**:361-364.

Schafer, U., Rausch, O., Bouwmeester, T. and Pieler, T. 1994. Sequence-specific recognition of a repetitive DNA element by a C_2H_2 zinc-finger protein in *Xenopus*. Eur. J. Biochem **226**:567-576.

Scheid, M., Boyse, E.A., Carswell, E.A. and Old, L.J. 1972. Serologically demonstrable alloantigens of mouse epidermal cells. J. Exp. Med. **135**:938-955.

Schlesinger, M. and Hurvitz, D. 1969. Characterisation of cytotoxic isoantisera produced in RIII mice. Transplantation **7**:132-141.

Schneider-Maunory, S., Topilko, P., Seitanidou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C. and Charnay, P. 1993. Disruption of *Krox-20* results in alteration of rhombomeres 3 and 5 in the developing hindbrain. Cell **75**:1199-1214.

Schneider-Schaulies, J., von Brunn, A. and Schachner, M. 1990. Recombinant peripheral myelin protein Po confers both adhesion and neurite outgrowth-promoting properties. J. Neurosci. Res. **27**:286-297.

Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. and Pachnis, V. 1994. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. Nature **367**:380-383.

Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, K. and Jackle, H. 1986. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Kruppel*, a Drosphila segmentation gene. Cell **47**:1025-1035.

Schwartzberg, P.L., Goff, S.P. and Robertson, E.J. 1989. Germ-line transmission of a *c-abl* mutation produced by targeted gene disruption in ES cells. Science **246**:799-803.

Schwartzberg, P.L., Robertson, E.J. and Goff, S.P. 1990. Targeted gene disruption of the endogenous *c-abl* locus by homologous recombination with DNA encoding a selectable fusion protein. Proc. Natl. Acad. Sci. USA **87**:3210-3214.

Scott, M.P. and Weiner, A.J. 1984. Structural relationship among genes that control development: sequence homology between Antennapedia, Ultrabithorax, Fushi-tarazu loci of *Drosophila*. Proc. Natl. Acad. Sci.USA **81**:4115-4119.

Seki, T., Moriuchi, T., Chang, H.-C., Denome, R. and Silver, J. 1985a. Structural organisation of the rat Thy-1 gene. Nature **313**:485-487.

Seki, T., Spurr, N., Obata, F., Goyert, S. and Goodfellow, P. 1985b. The human Thy-1 gene: structure and chromosomal location. Proc. Natl. Acad. Sci. USA **82**:6657-6661.

Seki, T., Chang, H.-C., Moriuchi, T., Denome, R., Ploegh, H. and Silver, J. 1985c. A hydrophobic transmembrane segment at the carboxyl terminus of Thy-1. Science **227**:649-651.

Serwe, M. and Sablitzky, F. 1993. V(D)J recombination in B cells is impaired but not blocked byb targeted deletion of the immunoglobulin heavy chain intron enhancer. EMBO J. **12**:2321-2327.

Shaw, G. and Kamen, R. 1986. A conserved sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**:659-657.

Sheldrake, R.F., Crocker, B.D., Husband, A.J. and Rostas, J.A. 1985. Purification and partial characterisation of a Thy-1 like glycoprotein from sheep brain. Neurosci. Lett. **19**:97.

Shen, B., Kim, J., K. and Dorsett, D. 1994. The enhancer-blocking *suppressor of Hairy-wing* zinc finger protein of *Drosophila melanogaster* alters DNA structure. Mol. Cell. Biol. **14**:5645-5652.

Shenoy-Scaria, A.M., Kwong, J., Fujita, T., Olszowy, M.W., Shaw, A.S. and Lublin, D.M. 1992. Signal transduction through decay-accelerating factor. Interaction of glycosyl-phosphatidylinositol anchor and protein tyrosine kinases $p56^{lck}$ and $p59^{fyn}$. J. Immunol **149**:3535-41.

Shepard, J.C., McGinnis, W., Carrasco, A.E., DeRobertis, E.M. and Gehring, W.J. 1984. Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. Nature **310**:70-71.

Sheseley, E.G., Kim, H.-S., Shehee, R.W., Papayannopoulou, T., Smithies, O. and Popovich, B.W. 1991. Correction of a human β^s-globin gene by gene targeting. Proc. Natl. Acad. Sci. USA **88**:4294-4298.

Shiba, T., Iwasaki, H., Nakata, A. and Shinagawa, H. 1991. SOS-inducible DNA repair proteins, RuvA and RuvB, of *Escherichia coli*: functional interactions between RuvA and RuvB for ATP hydrolysis and renaturation of the cruciform structure in supercoiled DNA. Proc. Natl. Acad. Sci. USA **88**:8445-8449.

Shigekawa, K. and Dower, W., J. 1988. Electroporation of eukaryotes and prokaryotes: A general approach to the intoduction of macromolecules into cells. Biotechniques **6**:742-751.

Shinkai, Y., Rathbun, G., Lam, K.-P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M. and Alt, F.W. 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell **68**:855-867.

Shinohara, A., Ogawa, H. and Ogawa, T. 1992. Rad 51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. Cell **69**:457-470.

Silva, A.J., Stevens, C.F., Tonegawa, S. and Wang, Y. 1992. Deficient hippocampal long-term potentiation in alpha-calcium-calmodulin kinase II mutant mice. Science **257**:201-206.

Simon, J. 1995. Locking in stable states of gene expression: transcriptional control during *Drosophila* development. Curr. Opin. Cell. Biol. **7**:376-385.

Simon, J., Chiang, A.C. and Bender, W. 1992. Ten different *Polycomb* group genes are required for spatial control of the *abdA* and *abdB* homeotic products. Development **114**:493-505.

Smeyne, R.J., Klein, R., Schnapp, A., Long, L.K., Bryant, S., Lewin, A., Lira, S.A. and Barbacid, M. 1994. Severe sensory and sympathetic neuropathies in mice carrying a disrupted *trk*/NGF receptor gene. Nature **368**:246-249.

Smith, A.G., Heath, J.G., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M. and Rogers, D. 1988. Inhibition of pluripotential stem cell differentiation by purified polypeptide. Nature **336**:688-690.

Smith, A.G. and Hooper, M.L. 1987. Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine emrbyonal carcinoma and embryonic stem cells. Developmental Biol. **121**:1-9.

Smith, J.C. 1994. Hedgehog, the floor plate, and the zone of polarizing activity. Cell **76**:193-196.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D.,

Fujimoto, E.K., Goeke, N.M., Olson, B.J. and Klenk, D.C. 1985. Measurement of protein using bicinchoninic acid. Anal. Biochem. **150**:76-85.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapatai, R.S. 1985. Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. Nature **317**:230-234.

Snider, W.D. 1994. Function of neurotrophins during nervous system development: what the knockouts are teaching us. Cell **77**:627-638.

Soriano, P., Montgomery, C., Geske, R. and Bradley, A. 1991. Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. Cell **64**:693-702.

Southern, E.M. 1975. Detection of specific DNA sequences among DNA fragments seperated by gel electrophoresis. J. Mol. Biol. **98**:503-517.

Southern, P.J. and Berg, P. 1982. Transformation of cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1**:327-341.

Spangrude, G.J. and Brooks, D.M. 1992. Phenotypic analysis of mouse hematopoietic stem cells shows a Thy-1-negative subset. Blood **80**:1957-1964.

Spanopoulou, E., Giguere, V. and Grosveld, F. 1988. Transcriptional unit of the murine Thy-1 gene: different distribution of transcription initiation sites in the brain. Mol. Cell. Biol. **8**:3847-3856.

Spanopoulou, E., Giguere, V. and Grosveld, F. 1991. The functional domains of the murine *Thy-1* gene promoter. Mol. Cell. Biol. **11**:2216-2228.

Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A. and Melton, D.W. 1994. Use of a double-replacement gene targeting to replace the murine α-lactalbumin gene with its human counterpart in embryonic stem cells and mice. Mol. Cell. Biol. **14**:1009-1016.

Stahl, N., Baldwin, M.A., Teplow, D.B., Hood, L., Gibson, B.W., Burlingame, A.L. and Pruisner, S.B. 1993. Structural studies of the scrapie prion protein using mass spectrometry and amino acid sequencing. Biochemistry **32**:1991-2002.

Stahl, N., Borchelt, D.R., Hsiao, K. and Pruisner, S.B. 1987. Scrapie prion protein contains a phosphatidylinositol glycolipid. Cell **51**:229-240.

Stefanova, I., Horejsi, V., Ansotegui, I.J., Knapp, W. and Stockinger, H. 1991. GPI-anchored cell-surface molecules complexed to protein tyrosine kinases. Science **254**:1016-1018.

Stehle, J.H., Foulkes, N.S., Molina, C.A., Simonneaux, V., Pevet, P. and Sassone-Corsi, P. 1993. Adrenergic signals direct rhythmic expression of transcriptional repressor CREM in the pineal gland. Nature **365**:314-320.

Steitz, T.A., Ohlendorf, D.H., Mckay, D.B., Anderson, W.F. and Mathews, B.W. 1982. Structural similarity in the DNA binding domain of catabolite gene activator and cro repressor proteins. Proc. Natl. Acad. Sci.USA **79**:3097-3100.

Stern, L.C. 1973. θ-alloantigen on mouse and rat fibroblasts. Nature New Biol. **246**:76-78.

Stevens, L.C. 1973. A new inbred subline of mice (129/terSV) with a high incidence of spontaneous congenital testicular teratomas. J. Natl. Cancer Inst. **50**:235-242.

Stief, A., Winter, D.M., Stratling, W.H. and Sippel, A.E. 1989. A nuclear DNA attachment

element mediates elevated and position-independent gene activity. Nature **341**:343-345.

Struhl, K. 1994. Duality of TBP, the universal transcription factor. Science **263**:1103-1104.

Studisky, V.M., Clark, D.J. and Felsenfeld, G. 1994. A histone octamer can step around a transcription polymerase without leaving the template. Cell **76**:371-382.

Su, B., Waneck, G.L., Flavell, R.A. and Bothwell, A.L.M. 1991. The glycosyl phosphatidylinositol anchor is critical for Ly-6A/E-mediated T cell activation. J. Cell. Biol. **112**:377-384.

Suda, Y., Suzuki, M., Ikawa, Y. and Aizawa, S. 1987. Mouse embryonic stem cells exhibit indefinite proliferative potential. J. Cell. Physiol. **133**:197-201.

Sullivan, C.H., Norman, J.T., Borras, T. and Grainger, R.M. 1989. Developmental regulation of hypomethylation of delta-crystallin genes in chicken embryo lens cells. Mol. Cell. Biol. **9**:3132-3135.

Summers, M.F., South, T.L., Kim, B. and Hare, D.R. 1990. High-resolution structure of a HIV zinc finger like domain via a new NMR-based distance geometry approach. Biochemistry **29**:329-340.

Sun, H., Treco, D. and Szostak, J.W. 1991. Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination site. Cell **64**:1155-1161.

Sutherland, R.J., Kolb, B. and Whishaw, I.Q. 1982. Spatial mapping: definitive disruption by hippocampal or medial frontal cortical damage. Neurosci. Lett. **31**:271-276.

Swiatek, P.J. and Gridley, T. 1993. Perinatal lethality and defects in hindbain development in mice homozygous for a targeted mutation of the zinc finger gene *Krox20*. Genes. Dev. **7**:2071-2084.

Szostak, J.W., Orr-Weaver, T.L. and Rothstein, R.J. 1983. The double-strand-break repair model for recombination. Cell **33**:25-35.

Talanian, R.V., McKnight, C.J. and Kim, P.S. 1990. Sequence-specific DNA binding by a short peptide dimer. Science **249**:769.

Talbot, D., Collis, P., Antoniou, M., Vidal, M., Grosveld, F. and Greaves, D.R. 1989. A dominant control region from the human β-globin locus conferring integration site independent gene expression. Nature **338**:352-355.

Tanese, N., Pugh, B.F. and Tjian, R. 1991. Coactivators for a proline activator purified from the multisubunit human TFIID complex. Genes Dev. **5**:2212-2224.

Tarkowski, A.K. 1961. Mouse chimeras developed from fused eggs. Nature **184**:1286-1287.

Tautz, D., Lehmann, H., Schnurch, P., Schuh, P., Seifert, E., Keinlin, A., Jones, J. and Jackle, H. 1987. Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. Nature **327**:383-389.

te Riele, H., Maandag, E.R. and Berns, A. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc. Natl. Acad. Sci. USA **89**:5128-5132.

Theunissen, O., Rudt, F., Guddat, U., Mentzel, H. and Pieler, T. 1992. RNA and DNA binding zinc fingers of Xenopus TFIIIA. Cell **71**:679-690.

Thoma, F., Koller, T.H. and Klug, A. 1979. Involvement of histone H1 in the organisation of the nucleosome and of the H1-dependent superstructures of chromatin. J. Cell. Biol. **83**:403.

Thomas, K.R. and Capecchi, M.R. 1987. Site directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell **51**:503-512.

Thomas, K.R. and Capecchi, M.R. 1990. Targeted disruption of the murine *int-1* proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. Nature **346**:847-850.

Thomas, K.R., Deng, C., Capecchi, M.R. 1992. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. Mol. Cell. Biol. **12**:2919-2923.

Thomas, K.R., Folger, K.R. and Capecchi, M.R. 1986. High frequency targeting of genes to specific sites in the mammalian genome. Cell **44**:419-428.

Thomas, P.M. and Samelson, L.E. 1992. The glycophosphatidylinositol anchored Thy-1 molecule interacts with the $p60^{fyn}$ protein tyrosine kinase in T cells. J. Biol. Chem. **267**:12317-12322.

Thompson, E.M., Christians, E., Stinnakre, M.-G. and Renard, J.-P. 1994. Scaffold attachment regions stimulate hsp70.1 expression in mouse preimplantation embryos but not in differentiated tissues. Mol .Cell. Biol. **14**:4694.

Tiveron, M.-C., Barboni, E., Rivero, B., Gormley, A., Seeley, P.J. and Grosveld, F. 1992. Selective inhibition of neurite outgrowth on mature astrocytes by Thy-1 glycoprotein. Nature **355**:745-748.

Tiveron, M.-C., Nosten-Bertrand, M., Jani, H., Garnett, D., Hirst, E.M., Grosveld, F. and

Morris, R.J. 1994. The mode of anchorage to the cell surface determines both the function and the membrane location of Thy-1 glycoprotein. J. Cell Sci. **107**:1738-1796.

Tomasiewicz, H., Ono, K., Yee, D., Thompson, C., Goridis, C., Rutishauser, U. and Magnuson, T. 1993. Genetic deletion of a neuronal cell adhesion molecule variant (N-CAM-180) produces distinct defects in the central nervous system. Neuron **11**:1163-1174.

Travers, A.A. 1994. Keeping the writhe. Curr. Biol. **4**:659-661.

Tsaneva, I.R., Muller, B. and West, S.C. 1992. ATP-dependent branch migration of Holliday junctions promoted by the RuvA and RuvB proteins of E.coli. Cell **69**:1171-1180.

Tse, A.G., Barcaly, A.N., Watts, A. and Williams, F. 1985. A glycophospholipid tail at the carboxyl terminus of the Thy-1 glycoprotein of neurons and thymocytes. Science **230**:1003-1008.

Tsukiyama, T., Becker, P.B. and Wu, C. 1994. ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. Nature **367**:525-532.

Valancius, V. and Smithies, O. 1991. Testing an "In-Out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. Mol. Cell. Biol. **11**:1402-1408.

van der Lugt, N.M., Domen, J., Linders, K., van Room, M., Robanus-Maandag, E., te Riele, H., van der Valk, M., Deschamps, J., Sofroniew, M., van Lohuizen, M. and Berns, A. 1994. Posterior transformation, neurological abnormalities and severe

haematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. Genes Dev. **8**:757-769.

van Deursen, J. and Wieringa, B. 1992. Targeting of the creatine kinase B gene in embryonic stem cells using isogenic and nonisogenic vectors. Nucl. Acids Res. **20**:3815-3820.

van Doorninck, J.H., French, P.J., Verbeck, E., Peters, R.H., Morreau, H., Bijman, J. and Scholte, B.J. A mouse model for the cystic fibrosis ΔF508 mutation. EMBO J. *In press.*

van Hoy, M., Leuther, K.K., Kodadek, T. and Johnston, S.A. 1993. The acidic activation domains of GCN4 and GAL4 proteins are not alpha helical but form beta sheets. Cell **72**:587.

van Lohuizen, M., Frasch, M., Wientjens, E. and Berns, A. 1991. Sequence similarity between the mammalian *bmi-1* proto-oncogene and *Drosophila* regulatory genes *Psc* and *Su(z)2*. Nature **353**:353-355.

van Rijs, J., Giguere, V., Hurst, J., van Agthoven, T., van Kessel, A.G., Goyert, S. and Grosveld, F. 1985. Chromosomal localisation of the human Thy-1 gene. Proc. Natl. Acad. Sci. USA **82**:5832-5835.

Vidal, M., Morris, R.J., Grosveld, F. and Spanopoulou, E. 1990. Tissue-specific control elements of the Thy-1 gene. EMBO J. **9**:833-840.

Vinson, C.R., Sigler, P.B. and McKnight, S.L. 1989. Scissor-grip model for DNA recognition by a family of leucine zipper proteins. Science **246**:911.

Volarevic, S., Burns, C.M., Sussman, J.J. and Ashwell, J.D. 1990. Intimate association of

Thy-1 and the T-cell antigen receptor with the CD45 tyrosine phosphatase. Proc. Natl. Acad. Sci. USA **87**:7085-7089.

Walsh, F. and Ritter, M.A. 1981. Surface antigen differentiation during human myogenesis in culture. Nature **289**:60-64.

Ware, L.M. and Axelrad, A.A. 1972. Inherited resistance to N- and B-tropic murine leukaemia viruses in vitro: evidence that congenic mouse strains SIM and SIM.R differ at the FV-1 locus. Virology **50**:339.

Watson, F.J. 1992. The murine Thy-1 gene: homologous recombination in embryonic stem cells and transcriptional analysis. Ph.D. Thesis. University of London.

Watt, F. and Molloy, P.L. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev. **2**:1136-1143.

Weintraub, H. and Groudine, M. 1976. Chromosomal subunits in active genes have altered conformation. Science **193**:848.

West, S. 1992. Enzymes and molecular mechanisms of genetic recombination. Annu. Rev. Biochem. **61**:603-640.

Wharton, R.P., Brown, E.L. and Ptashne, M. 1984. Substituting an alpha-helix switches the sequence-specific DNA interactions of a repressor. Cell **38**:361.

White, P.L. and Streilein, J.W. 1983. Thy-1 antigen is present on B and T lymphocytes of Syrian hamster. J. Immunol. **131**:29903-2907.

Widom, J. and Klug, A. 1985. Structure of the 300A chromatin filament: X-ray diffraction

from orientated samples. Cell **43**:207.

Wilkinson, D.G., Bailes, J.A. and McMahon, A.P. 1987. Expression of the proto-oncogene *int-1* is restricted to specific neural cells in the developing mouse embryo. Cell **50**:79-88.

Wilkinson, D.G., Bhatt, S., Chavrier, P., Bravo, R. and Charnay, P. 1989. Segment specific expression of a zinc-finger gene in the developing nervous system of the mouse. Nature **337**:461-464.

Wilkinson, D.G. and Krumlauf, R. 1990. Molecular approaches to segmentation of the hindbrain. Trends Neurosci. **13**:335-338.

Williams, A.F. and Barclay, A.N. 1988. The immunoglobulin superfamily: domains for cell surface recognition. Ann. Rev. Immunol. **6**:381-405.

Williams, A.F. and Gagnon, J. 1982. Neuronal cell Thy-1 glycoprotein: homology with immunoglobulin. Science **216**:696-703.

Williams, A.F., Parekh, R.B., Wing, D.R., Willis, A.C., Barclay, A.N., Dalchau, R., Fabre, J.W., Dwek, R.A. and Rademacher, T.W. 1993. Comparative analysis of the N-glycans of rat, mouse and human Thy-1. Site specific oligosaccharide patterns of neural Thy-1, a member of the immunoglobulin superfamily. Glycobiology **3**:339-48.

Williams, L.R., Hilton, D.G., Pease, S., Wilson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A. and Gough, N.M. 1988a. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature **336**:684-687.

Williams, T., Adman, A., Luscher, B. and Tjian, R. 1988b. Cloning and expression of AP-2, a cell type transcription factor that activates inducible enhancer elements. Genes Dev. **2**:1557-1562.

Wilson, C., Bellen, H.J. and Gehring, W.J. 1990. Position effects on eukaryotic gene expression. Ann. Rev. Cell Biol. **6**:679.

Winoto, A. and Baltimore, D. 1989. αβ Lineage-specific expression of the α T cell receptor gene by nearby silencers. Cell **59**:649-655.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. and Bonventre, J.V. 1994. The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. USA **91**:4514-4518.

Wood, S.A., Pascoe, W.S., Schmidt, C., Kemler, R., Evans, M.J. and Allen, N.D. 1993. Simple and efficient production of embryonic stem cell-embryo chimaeras by coculture. Proc. Natl. Acad. Sci. USA **90**:4582-4585.

Workman, J.L. and Roeder, R.G. 1987. Binding of transcription factor TFIID to the major late promoter *in vitro* nucleosome assembly potentiates subsequent initiation by polymerase II. Cell **51**:613-622.

Wurst, W., Auerbach, A.B. and Joyner, A.L. 1994. Multiple developmental defects in *Engrailed-1* mutant mice: an early mid-hindbrain deletion and patterning defects in forelimbs and sternum. Development **120**:2065-2075.

Wurst, W. and Joyner, A.L. 1993. Production of targeted embryonic stem cell clones. p. 33-61. In Joyner, A.L. (ed.), Gene targeting: a practical approach. IRL Press, Oxford.

Xue, G.P. and Morris, R.J. 1992. Expression of the neuronal surface glycoprotein Thy-1 does not follow appearance of its mRNA in developing mouse Purkinje cells. J. Neurochem. **58**:430-440.

Xue, G.P., Calvert, R.A. and Morris, R.J. 1990. Expression of the neuronal surface glycoprotein Thy-1 is under post-transcriptional control, and is spatially regulated, in the developing olfactory system. Development **109**:851-864.

Xue, G.P., Rivero, B.P. and Morris, R.J. 1991. The surface glycoprotein Thy-1 is excluded from growing axons during development: a study of the expression of Thy-1 during axogenesis in hippocampus and hindbrain. Development **112**:161-176.

Yagi, T., Ikawa, Y., Yoshida, K., Shigetani, Y., Takeda, N., Mabuchi, I., Yamamoto, T. and Aizawa, S. 1990. Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diptheria toxin A-fragment gene in negative selection. Proc. Natl. Acad. Sci. USA **87**:9918-9922.

Yamada, T., Pfaff, S.L., Edlund, T. and Jessell, T.M. 1993. Control of cell pattern in the neural tube: motor neuron induction by diffusible factors from notochord and floor plate. Cell **73**:673-686.

Yamada, T., Placzek, M., Tanaka, H., Dodd, J. and Jessell, T.M. 1991. Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and notochord. Cell **64**:635-647.

Ying, G.P., Anderson, R.G. and Rothberg, K.G. 1992. Each caveola contains multiple glycosylphosphatidylinositol anchored membrane proteins. Cold Spring Harbor Symp. Quant. Biol. **57**:593-604.

Zhang, H., Hasty, P. and Bradley, A. 1994. Targeting frequency for deletion vectors in

embryonic stem cells. Mol. Cell. Biol. **14**:2404-2410.

Zhang, H., Hasty, P., Brenneman, M.A., Grompe, M., Gibbs, R.A., Wilson, J.H. and Bradley, A. 1991. Fidelity of targeted recombination in human fibroblasts and murine embryonic stem cells. Proc. Natl. Acad. Sci. USA **88**:8067-8071.

Zijlstra, M., Li, E., Sajjadi, F., Subramani, S. and Jaenisch, R. 1989. Germ-line transmission of a disrupted β_2-microglobulin gene produced by homologous recombination in embryonic stem cells. Nature **342**:435-438.

Zimmer, A. 1992. Manipulating the genome by homologous recombination in embryonic stem cells. Annu. Rev. Neurosci. **15**:115-137.

Zimmer, A. and Gruss, P. 1989. Production of chimaeric mice containing embryonic stem (ES) cells carrying a homeobox *Hox 1.1* allele mutated by homologous recombination. Nature **338**:150-153.

Zuber, C., Lackie, P.M., Catterall, W.A. and Roth, J. 1992. Polysialic acid is associated with sodium channels and the neural cell adhesion molecule N-CAM in adult brain. J.Biol.Chem. **267**:9965-9971.