Chapter 1: Introduction
Chapter 1: Introduction

1.0. INTRODUCTION

The mammalian cerebral cortex is a most remarkable product of brain evolution, and is the structure that most distinctively delineates the human species from others (Northcutt and Kaas, 1995; Rakic, 1988). Neurons within the cortex underlie our most sophisticated cognitive and perceptual abilities and are thus highly specialised for reception of multiple sensory inputs and generation of motor outputs. Furthermore, neurons in the adult brain are organised into cytoarchitectonic areas (see Figure 1.1), defined by distinct biochemical, morphological and physiological characteristics (Rakic, 1988). Remarkably, this structure of multiplicate complexity is generated from an outwardly simple neuroepithelium in which cells appear morphologically indistinguishable.

What are the mechanisms that direct neuron formation and subsequent functional-parcellation of the cerebral cortex? Key to the study of this process is an understanding of neuronal fate determination. During development, there is a precise coordination of intrinsic, cell-autonomous signalling networks, coupled with reception of non cell-autonomous, short-range and long-range extracellular signals which guide cells through distinct developmental paths. Accumulating evidence demonstrates an intrinsic programming potential by neuronal progenitors within subdomains of the developing cerebral cortex that is instructive for proper corticogenesis. These regional boundaries are demarcated by expression of certain transcription factors, including members of the Helix-Loop-Helix (HLH) family of proteins (Puelles and Rubenstein, 1993; Schuurmans and Guillemot, 2002).

While studies in mice have convincingly demonstrated that neurogenic HLH proteins such as NeuroD (Lee et al., 1995) and Mash1 (Casarosa et al., 1999) are intimately involved in neuronal fate determination and terminal differentiation, the role of the ubiquitously expressed HLH protein, E12 (Murre et al., 1989b), remains ambiguous. Numerous descriptive studies have implicated a role for E12 in the development of the nervous systems of humans (Rutherford and LeBrun, 1998), rodents (Perez-Moreno et al., 2001; Roberts et al., 1993) and the nematode Caenorhabditis elegans (Krause et al., 1997), but its precise neurogenetic contribution may not be easily elucidated through classical genetic dissection, including loss-of-function approaches (Bain et al., 1994; Zhuang et al., 1994). Since the function of E12 is, in part, coded
Figure 1.1. Schematic representation of functional subdivisions of the adult rodent cortex, viewed from the lateral aspect. Ob = olfactory bulb; Fr1 = frontal 1; Fr2 = frontal 2; Sm1 = primary sensorimotor; Sm2 = secondary somatosensory; Occ = occipital; Te = auditory; Pir = piriform; Prh = perirhinal; Ent = entorhinal; Ins = insular (Reproduced from Tan et al., 2002)

Figure 1.2. Transverse sections of the embryo during early development. The blastula stage embryo comprises a hollow sphere of cells (A), which then undergoes further divisions and invaginations to give rise to the multi-layered gastrula (B).
by its capacity for protein dimerisation, elucidation of its role in corticogenesis requires not only an appreciation of the spatial and temporal expression pattern of this gene, but also an evaluation of key neurogenetic signalling pathways that are triggered by this transcription factor.

1.1.1. The Development of the Mammalian Central Nervous System

Embryogenesis begins with the zygote, a product of fertilization of an egg by penetration of a sperm. Following fertilization, the zygote divides into blastomeres which partition to form a hollow sphere of cells called the blastula. In the next developmental step, further invaginations at the blastula stage ultimately give rise to the three germ layers of gastrula-stage embryo: endoderm, the inner most layer, of which the gut, lungs and liver are derived; mesoderm, the middle layer, which contributes connective tissue, muscle, kidneys and the vascular system; and ectoderm, the outermost layer, which gives rise to the all the major tissues of the central and peripheral nervous system, as well as the epidermis (see Figure 1.2).

Local interactions between ectoderm and underlying mesoderm along the dorsal midline of the embryo instruct these cells to form a neuroepithelium, rather than to differentiate into epidermis. Known as the neural plate, this specialised region of the ectoderm becomes committed to formation of the nervous system through further interactions with mesodermal tissue, including the notochord, in a process known as neural induction. Soon after this, the neural plate begins to fold at its lateral edges, finally fusing at its dorsal most extreme to form a neural tube. The resulting neuroepithelium that lines the walls of the neural tube generate all cell components of the central nervous system, while the peripheral nervous system arises from a distinct group of migratory precursor cells along the dorsolateral border of the neural folds, known as neural crest cells.

Development of the neural tube follows a rostrocaudal progression of maturation, and early specification of brain structures, observed as disproportionate cell proliferation at the rostral end, give rise to three distinct swellings called the forebrain vesicle (prosencephalon), the midbrain vesicle (mesencephalon) and hindbrain vesicle (rhombencephalon) (Figure 1.3). The forebrain vesicle ultimately gives rise to the telencephalon, a structure from which the bihemispheric cerebral cortex, basal ganglia, hippocampal...
Figure 1.3. Development of the neural tube. Following neural induction, the nervous system may be represented as a hollow tube of cells with three brain vesicles at the anterior end (A). These vesicles eventually undergo disproportionate expansion to give rise to the forebrain (composed of the telencephalon and diencephalon), midbrain (mesencephalon) and hindbrain (composed of the metencephalon and myelencephalon) (B).
formation and amygdala originate; as well as the diencephalon, composed of the thalamus, subthalamus, hypothalamus and promordial retine (Martin and Jessell, 1991). It is the development of the cerebral cortex that is the focus of this study.

1.1.2. Histogenesis of the mammalian cerebral cortex

A hallmark of the cerebral cortex has been its rapid and selective expansion during the course of evolution; this expansion is visible in the early telencephalic vesicles of the mammalian embryonic brain, which balloon outward from the rostral forebrain to generate a large area of cortical progenitor cells (McConnell, 1995). It is the disproportionate expansion and exceptional growth of the forebrain rudiment which gives rise to a cerebral cortex that dwarfs all other structures of the central nervous system, comprising up to two-thirds of the neuronal mass and containing about three-quarters of all synapses in humans (Rakic, 1988).

Neurogenesis of the cerebral cortex is essentially a tripartite process, beginning with growth of a large pool of proliferative precursor cells, predominantly taking place in the germinal matrix. Next, waves of postmitotic neurons migrate away from the germinal matrix, moving to more superficial layers of the cortical anlage. Finally, these neurons undergo terminal differentiation, settle at their final destination in the cortex and form mature connections.

Neurons of the cerebral cortex are mainly generated prenatally (Bayer and Altman, 1991; Super et al., 1998; Uylings and van Eden, 1990) and arise from a germinal neuroepithelium that undergoes approximately 11 integer cycles of cell division initiated by the founder progenitor population (Caviness et al., 1995). The first terminally postmitotic neurons are observable on embryonic day e11.5 in the mouse, as the formative cortex comprises two distinct components, namely the neuroepithelium (ne), and the preplate (pp), also known as the primordial plexiform layer (see Figure 1.4). The preplate principally contains the precociously mature Cajal-Retzius neurons, a pioneer cell population which are the earliest neurons to differentiate and is indispensable for proper neurogenesis of the cerebral cortex (for a recent review see Super et al., 1998). Also among the first cells in the developing cortex with a differentiated phenotype are the radial glial cells; typical long bipolar cells that expand across the whole cerebral wall from the
Figure 1.4. Diagrammatic representation of histogenesis of the developing cerebral cortex. Among the proliferative neuroblasts lie the pioneer neurons which guide successive waves of postmitotic precursors to their final laminar destinations, producing a six-layered cortical plate (Future Cortex). Abbreviations: V, ventricular zone; PP, preplate; IZ, intermediate zone; SV, subventricular zone; SP, subplate; CP, cortical plate (from Uylings et al, 1990).

Figure 1.5. Coronal section of telencephalon illustrating two modes of migration into the cerebral cortex (Cx). Radially oriented cortical neurons (*) from the dorsal telencephalon migrate towards the pial surface, giving rise to a columnar organization. These neurons are excitatory in nature, using glutamate as a neurotransmitter. On the other hand, tangentially migrating neurons (#) from the ganglionic eminence (GE) in the ventral telencephalon give rise to inhibitory neurons in the cortex which express GABA.
ventricular zone to the pial membranes. As postmitotic neurons leave the germinal neuroepithelium, they migrate along the radial glia scaffold and insert in the newly formed cortical plate within the preplate. In this process, the preplate is split into a superficial component, the marginal zone (MZ), and a deep component, the subplate (SP) and the intermediate zone (IZ) (Figure 1.4).

The cortical plate is first observed around e12.5 in developing mouse brain (Bayer and Altman, 1991), and is constructed in an inside-out fashion; successive waves of postmitotic neurons leave the ventricular zone and migrate toward the pial surface, moving through earlier-born neurons. This orchestrated process of neuronal migration finally gives rise to the six-layered cortex; the marginal zone constitutes Layer I, while the intermediate zone eventually shrinks and is replaced by the cortical white matter, composed mainly of incoming cortical afferents. Similarly, there is a concomitant shrinkage of the ventricular zone to an ependymal layer.

1.1.3. There exists two modes of neuronal migration in the developing telencephalon

The observation that the majority of postmitotic neurons migrate radially into the cortex led to the hypothesis that the ventricular zone represents a “protomap” as a genetic blueprint for the cytoarchitectonic areal map of the cerebral cortex, thereby facilitating its construction (Rakic, 1988; Super et al., 1998 and references herein). This model, however, does not account for the origin of tangentially orientated cells in the intermediate zone of the developing cerebral cortex, as revealed histologically (Bayer and Altman, 1991; Marin and Rubenstein, 2001; and references herein). The first indication of non-radial migratory behaviour in the developing cortex came from retroviral cell lineage studies that demonstrated the capacity for clonally related neurons to disperse tangentially across functionally different areas of the cortex (Austin and Cepko, 1990; Reid et al., 1995; Walsh and Cepko, 1988; Walsh and Cepko, 1992). Complementary approaches (O'Rourke et al., 1992), including the use of X-inactivated transgenic mosaics (Tan and Breen, 1993; Tan et al., 1995) arrived at a similar conclusion, and provided immutable evidence for this migratory behaviour. The existence of two modes of migration in the developing cortex was eventually recognised as a developmental phenomenon that facilitated integration of two different neuronal progenitor cell populations from the dorsal and ventral telencephalon. Shown in Figure 1.5, radially migrating neurons...
from the dorsal telencephalon give rise to excitatory neurons which express glutamate as a neurotransmitter, while tangentially dispersed neurons originate from the ventral telencephalon, and these constitute the majority of inhibitory interneurons of the dorsal cortex which express GABA ($\gamma$-aminobutyric acid) (for a review see Marin and Rubenstein, 2001). This heterogeneity in cell composition of the cortex is, in part, sculpted by intrinsic mechanisms, such as differential patterns of gene expression observed in dorsal- versus ventral-progenitor pools (Marin and Rubenstein, 2001; Schuurmans and Guillemot, 2002).

1.1.4. Defined genetic subdivisions within the developing cerebral cortex

An early description of distinct genetic subdivisions in the embryonic mouse forebrain was presented by Puelles and Rubenstein, who demonstrated an organisation of the developing neuroepithelium into transverse domains and longitudinal domains (Puelles and Rubenstein, 1993). In the developing telencephalon, separate dorsal and ventral progenitor pools are demarcated by boundaries of transcription factor gene expression within these fields. These genetic subdivisions are crucial for the generation of distinct neuronal subtypes of the adult cerebral cortex. For example, the helix-loop-helix (HLH) transcription factors $\textit{ngn1/2}$ and $\textit{Mash1}$ are required for proper development of the dorsal and ventral telencephalon, respectively (see below; for a review see Schuurmans and Guillemot, 2002). Further, cross-regulatory interactions between dorsal and ventral progenitor domains participate in the maintenance of progenitor cell identity (Schuurmans and Guillemot, 2002; also see text below).

The HLH proteins are a diverse family of protein transcription factors which exert profound influences over a wide variety of developmental programs such as neurogenesis, haematopoiesis, myogenesis and skeletogenesis. These transcription factors homo- or heterodimerise and bind cognate promoter DNA sequences (known as E-boxes) upstream of target genes to enhance or repress transcription. While some HLH proteins exhibit specific roles in development of the nervous system (for a review see Bertrand 2002), others, including the ubiquitously expressed E proteins, display widespread, overlapping patterns of gene expression and may orchestrate the development of multiple cell lineages, including progenitor cells of the nervous system, in a redundant manner. One such member of the E protein subfamily is E12, a
transcription factor first identified as an important transcriptional regulator of immunoglobulin genes (Murre et al., 1989b). While expression studies show that E12 is expressed in many nascent organs of the developing mouse embryo, including the developing nervous system, evidence for a definitive role in neurodevelopment is lacking. It is, therefore, the focus of the following review to formulate a working model which ascribes a function for E12 in cortical development.

1.2.1. Helix-loop-helix proteins

The HLH family of transcription factors comprises more than 400 members identified in organisms from yeast to humans (reviewed in Littlewood and Evan, 1998; Massari and Murre, 2000). This list is ever-growing with yet more discovered through exhaustive searches for genes with conserved HLH domains in whole genomes (Moore et al., 2000; Peyrefitte et al., 2001), as well as through conventional molecular biological approaches (Narumi et al., 2000; Zhou et al., 2000). In metazoan species, HLH proteins function in the coordinate regulation of gene expression, with roles in cell cycle regulation (Norton, 2000), lineage commitment (for a review see Bertrand et al., 2002) and cell differentiation (see Massari and Murre, 2000 for a review).

The highly conserved HLH region common to all family members comprises two amphiphatic $\alpha$ helices separated by a short intervening loop of variable length. Additionally, the majority of HLH proteins also contain a stretch of basic amino acid residues adjacent to the HLH motif that directly contacts DNA. The HLH domain primarily mediates homo- or heterodimerisation, a function essential for DNA binding and transcriptional regulation by these proteins (see Figure 1.6).

Confirmation of the existence of the HLH dimerisation motif was first provided through solution of the crystal structure of the HLH protein Max (Ferre-D'Amare et al., 1993). Independent verification of these observations arose from analysis of the three dimensional structure of another HLH protein, E47, at 2.8 Å resolution (Ellenberger et al., 1994). This dimeric protein complex is bound to its E-box oligonucleotide recognition site, CACCTG, as a parallel, four-helix bundle which allows the basic region to contact the major groove of the duplexed DNA target (Ellenberger et al., 1994). In addition to the basic region,
Figure 1.6. Diagrammatic representation of the crystal structure of a HLH protein dimer contacting DNA in the major groove. The helix-loop-helix region mediates protein dimerisation, a necessary function for DNA-binding by these transcription factors.

Figure 1.7. A functional classification scheme for HLH proteins. Multiple sequence alignment of candidate HLH proteins by Massari and Murre (Massari and Murre, 2000) resulted in the segregation of family members into seven groups (I to VII). Class I and II factors encode bHLH domains, while Class III and IV contain an additional leucine zipper (LZ) in their C-termini. Proteins grouped to Class V lack a basic domain for DNA binding, and Class VI proteins encode a subfamily of bHLH factors with a conserved proline residue within the basic domain. Finally, Class VII proteins contain a conserved PAS domain adjacent to the second helix of the bHLH region.
residues in the loop and helix 2 also make contact with DNA, and Van der Waals interactions between conserved hydrophobic residues of the helical segment permit formation of a stable HLH dimer. The E47 homodimer is centred over the E-box, with each monomer interacting with either a CAC or CAG half-site. A glutamate present in the basic region of each subunit makes contact with the cytosine and adenine bases in the E-box half-site. An adjacent arginine residue stabilises the position of the glutamate by direct interaction with these nucleotides and the phosphodiester backbone. Both the glutamate and arginine residues are conserved in most HLH proteins containing a basic amino acid domain, consistent with a role in specific DNA binding (Atchley and Fitch, 1997; Ellenberger et al., 1994).

1.2.2. Classification of HLH proteins into functional groups

The enormous number of HLH proteins presents a genuine challenge for their appropriate classification, and two schemes have been devised. Experimental biologists have grouped HLH proteins with respect to structure/function relationships (Massari and Murre, 2000; Murre et al., 1994), while computational biologists have focused on evolutionary conservation of polypeptide sequence and DNA-binding specificities (Atchley and Fitch, 1997; Atchley et al., 2000).

The approach by Murre and coworkers (Murre et al., 1994) to devise a classification scheme for HLH proteins focused on evaluating candidates for tissue distribution, capacity for dimerisation, as well as specificity for DNA-binding (Murre et al., 1994). An alignment of these HLH proteins using the Clustal W algorithm (Thompson et al., 1994) segregated candidates into seven categories (Figure 1.7). Class I HLH proteins comprise the E proteins, such as E12 (Murre et al., 1989b), E47 (Murre et al., 1989b), Daughterless (Caudy et al., 1988b), HEB (Henthorn et al., 1990), and E2-2 (Henthorn et al., 1990), which exhibit widespread tissue distribution, and can homo- or heterodimerise with other HLH proteins. Dimerisation is essential for transcriptional regulation of downstream genes through binding to E-box hexanucleotide promoter sequences (Murre et al., 1994). Members of the Class II HLH proteins primarily exhibit restricted tissue distribution, and includes myogenic (MyoD, myogenin), haematopoietic (Tal1) and neurogenic (Drosophila achaete, scute, MASH-1, neurogenin) HLH factors (see Massari and Murre, 2000). The majority of Class II HLH factors preferentially heterodimerise with Class I proteins to bind canonical...
and noncanonical E-box sites (Murre et al., 1994 and references therein). Members of the Class III subgroup contain a leucine-zipper (denoted as “LZ”) domain adjacent to the HLH motif, and include the Myc family of transcription factors (Grandori et al., 2000 for a review), TFE3 (Roman et al., 1991; Roman et al., 1992), SREBP-1 (Athanikar and Osborne, 1998), and the microphthalmia-associated transcription factor, Mi (Hemesath et al., 1994). The Class IV proteins also encode bHLH-LZ domains, and members include dimerisation partners to Myc proteins such as Mad (Ayer et al., 1993) and Mxi (Zervos et al., 1994), which, additionally, have the capability to dimerise with each other (Grandori et al., 2000). Members of the Class V HLH proteins, such as *extramacrochaete* (emc) and mammalian IDs, lack a stretch of basic amino acids, and can form inactive heterodimers with Class I (reviewed in Norton, 2000) and II HLH factors (Langlands et al., 1997) which do not bind DNA, hence are classified as negative regulators of signalling cascades (reviewed in Norton, 2000). The HLH proteins which belong to Class VI are distinguished by a conserved proline residue in their basic amino acid region, and members include the *Drosophila* Hairy (Rushlow et al., 1989) and Enhancer of Split (Klambt et al., 1989), as well as their cloned mammalian homologues HES-1 and HES-3 (reviewed in Kageyama et al., 1997). Finally, members of the Class VII subfamily contain a conserved bHLH-PAS domain, and include the aromatic hydrocarbon receptor (AHR), the AHR nuclear translocator (Arnt) and hypoxia-inducible factor 1α (for reviews see Crews, 1998; Crews and Fan, 1999).

An alternative strategy was employed by computational biologists whom devised a classification scheme for HLH proteins with particular emphasis on evolutionary relationships (Atchley and Fitch, 1997; Atchley et al., 2000). Critical to this scheme was an assessment of amino acid sequence conservation within the HLH domain of candidate proteins, as well as their sequence preference for DNA-binding. This allowed for a delineation of HLH members into four subgroups, each with a predicted HLH consensus sequence; these four ancestral sequence motifs could be further collapsed to derive a hypothetical bHLH precursor motif (Atchley and Fitch, 1997). Importantly, these consensus sequences may be utilised as probes to search for yet undiscovered HLH motifs in sequenced genomes, thereby allowing codification of whole-genome complements of HLH-encoding genes. This “*in silico*” approach to search for HLH genes has been employed with some success (Moore et al., 2000; Peyrefitte et al., 2001).
While the latter approach to classification may also be more useful when new members need to be incorporated to the family, another approach using dedicated phylogenetic trees for small functional groups of HLH genes has also been informative. This was demonstrated in a recent description of HLH factors important for programming neurogenesis (Bertrand et al., 2002). For the purposes of this review, the former classification scheme developed by Murre and colleagues will be adopted, since discussion of HLH proteins herein includes an evaluation of their functional relationships, as well as their patterns of gene expression (see below).

1.2.3 The role of HLH genes in neurodevelopment

1.2.3.1 Identification of Class II HLH genes involved in *Drosophila* neurogenesis

The study of neural development in *Drosophila melanogaster* has been important for understanding signalling pathways involving HLH genes, and how these genes direct uncommitted precursor cells to a multitude of cell types of the adult nervous system (reviewed in Campuzano and Modolell, 1992; Garcia-Bellido and Moscoso del Prado, 1979; Ghysen and Dambly-Chaudiere, 1988). In particular, a small number of HLH proteins have important roles as ‘proneural genes’, and are both necessary and sufficient to initiate neuronal development.

Genetic analysis of mutant flies which lacked subsets of peripheral sense organs revealed convergence to a complex of genes, known as the achaete-scute complex, which is crucial for the regulation of the early steps of neural development (Campuzano and Modolell, 1992; Garcia-Bellido and Moscoso del Prado, 1979). Subsequent molecular cloning approaches led to the identification of four genes of this complex, namely *achaete* (*ac*), *scute* (*sc*), *lethal of scute* (*lsc*) and *asense* (*asc*) (Gonzalez et al., 1989; Villares and Cabrera, 1987). Remarkably, analysis of the coding sequence of all four genes revealed that each encoded a bHLH (basic Helix-Loop-Helix) domain. Additional screens for genes containing bHLH sequences related to *achaete-scute* complex genes (Campuzano and Modolell, 1992) led to cloning of *atonal* (*ato*) (Jarman et al., 1993) and structurally related genes which significantly contribute to development of the peripheral nervous system (Goulding et al., 2000a; Goulding et al., 2000b; Huang et al., 2000).
A considerable body of evidence indicates a proneural role for members of the *asc* and *ato* family. Firstly, the majority of these genes are expressed in proneural clusters prior to any observable phenotypic changes to the wing imaginal disc. Additionally, gain-of-function (GOF) approaches indicate these genes are necessary and sufficient to promote generation of neural progenitors from undifferentiated ectoderm (Bertrand et al., 2002; Jan and Jan, 1993; and references herein). Importantly, analysis of loss-of-function (LOF) mutants that harbour single deletions of *asc* or *ato* genes reveal dissimilarities in their potential for programming neurogenesis; while *sc* and *lsc* are indispensable for neuronal fate acquisition in undifferentiated ectoderm, *ase* is required for proper differentiation of committed neuroblasts, (Bertrand et al., 2002; Jan and Jan, 1993). Hence, these observations illustrate separate roles for proneural bHLH genes in fate determination (selection and programming of uncommitted progenitors) and proper differentiation (proper lineage specification of neuroblasts).

Furthermore, seminal work by Chien and colleagues (Chien et al., 1996) revealed encryption of neuronal type information in the bHLH domains of *asc* and *ato* genes. Analysis of mutant flies harbouring a chimeric gene comprising the basic domain of *ato* fused to the HLH domain of *sc* demonstrated that the basic domain of *ato* directs generation of stretch receptors, through a mechanism that possibly involves interaction with cofactors (Chien et al., 1996). Strikingly, this functional nonequivalence is also observed in vertebrate homologues of the *asc* and *ato* gene families (see below). Together, these observations demonstrate a functional role for each of these bHLH genes in programming sub-phenotype specification in neurons of the mature nervous system.

**1.2.3.2. Vertebrate homologues of Drosophila proneural genes behave as determination and differentiation factors**

Neurogenesis in vertebrates exhibits notable differences from the invertebrate model. In vertebrate embryos, neuroepithelial cells are pluripotent and generate precursors to either neurons or glia before differentiating. This is in contrast to *Drosophila* neurogenesis, whereby ectodermal precursors are directly programmed to generate neurons or glia. Some neurodevelopmental mechanisms, however, are common to
both models, such as the role of Notch signalling for proper construction of both vertebrate and invertebrate nervous systems (Artavanis-Tsakonas et al., 1999; de la Pompa et al., 1997). This drove speculation that vertebrate neurogenesis may be programmed through analogous signalling systems, and could involve proneural HLH genes paralogous to members of the achaete-scute complex. To this end, several molecular cloning strategies were successful in identifying more than 18 vertebrate homologues of asc and ato (Bertrand et al., 2002; Lee, 1997). Low-stringency hybridisation screens for HLH-like cDNAs led to cloning of NSCL (Begley et al., 1992), while a polymerase chain reaction (PCR)-based method led to cloning of Mash1 (mammalian homologue of achaete-scute; (Guillemot and Joyner, 1993; Sasai et al., 1992) and Math genes (mammalian homologues of atonal; Akazawa et al., 1995; Shimizu et al., 1995). In an alternative approach, yeast 2-hybrid screens which exploited the capacity for heterodimerisation by HLH proteins led to cloning of NeuroD (Lee et al., 1995) and Math4A/ngn2 (Gradwohl et al., 1996).

While in vitro assays are useful to demonstrate “neurogenic” capacities for HLH genes (Farah et al., 2000), their expression profiles suggest separate roles for corticogenesis in vivo (for a review see Lee, 1997). For example, the determination factors Mash1 (Guillemot and Joyner, 1993) and ngn2 (Ma et al., 1996) are expressed in progenitor cells of the developing ventral and dorsal telencephalon, respectively, and are indispensable for endowing a neuronal phenotype in postmitotic neuroblasts. In contrast, the neural differentiation factors, such as NeuroD (Lee et al., 1995) and Math-2 (Shimizu et al., 1995) exhibit delayed expression in neurons of the dorsal telencephalon that have left the ventricular zone. This sequential expression pattern in the developing cortex suggests specific cascades of HLH genes are required for proper neurogenesis of dorsal and ventral telencephalic progenitors. Indeed, functional experiments in Xenopus embryos confirmed this differential capacity for activating downstream effectors (Chien et al., 1996; Talikka et al., 2002). Studies of the Xenopus HLH genes Xash1 and Xngnr1 demonstrate that while ectopic expression of both genes can program neurogenesis in uncommitted cells, only Xngnr1 can activate expression of XneuroD and Xath3.

In a complementary approach, loss-of-function analyses reveal differing capacities for HLH “differentiation” factors in proper development of different neuronal subpopulations within the central
<table>
<thead>
<tr>
<th>LOF Mutant mouse</th>
<th>CNS Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mash1^{−/−}</td>
<td>Developmental defects in the ventral telencephalon, including a loss of specific classes of ventral neurons. Defects in postnatal development of the retina, as well as the peripheral nervous system (Casarosa et al., 1999; Horton et al., 1999)</td>
</tr>
<tr>
<td>Ngn2^{−/−}</td>
<td>Ventralisation of the dorsal telencephalon (Fode et al., 2000)</td>
</tr>
<tr>
<td>Ngn1^{−/−}/Ngn2^{−/−}</td>
<td>Pronounced ventralisation of telencephalon, compared to single-mutant mice (Fode et al., 2000)</td>
</tr>
<tr>
<td>Math1^{−/−}</td>
<td>Severe loss of cerebellar granule neurons (Ben-Arie et al., 1997)</td>
</tr>
<tr>
<td>Math2^{−/−}</td>
<td>No overt neurological phenotype (Schwab et al., 1998)</td>
</tr>
<tr>
<td>NeuroD^{−/−}</td>
<td>Loss of granule cells in hippocampus and cerebellum (Miyata et al., 1999; Liu et al., 2000)</td>
</tr>
<tr>
<td>NeuroD2^{−/−}</td>
<td>Loss of cerebellar and hippocampal neurons (Schwab et al., 2000; Olson et al., 2001)</td>
</tr>
<tr>
<td>Math2^{−/−}/NeuroD^{−/−}</td>
<td>Disruption of terminal differentiation in granule cells of the hippocampus (Schwab et al., 2000)</td>
</tr>
<tr>
<td>ID1^{−/−}/ID3^{−/−}</td>
<td>Early cell-cycle exit and premature neurogenesis (Lyden et al., 1999)</td>
</tr>
<tr>
<td>E2A^{−/−}</td>
<td>No overt neurological phenotype reported (Bain et al, 1994; Zhuang et al., 1994)</td>
</tr>
</tbody>
</table>

Figure 1.8. Summary of mutant mice harbouring null alleles for various HLH genes. Most mutant animals which lack Class II and V genes (such as Mash1 and Ids, respectively) exhibit defects in neurogenesis, while loss of the Class I HLH gene E2A (which encodes the polypeptides E12 and E47) does not result in any gross neurological abnormalities.

Figure 1.9. Specification of dorsoventral identity through modulation of neurogenic Class II HLH factor expression. Neurogenins 1 and 2 (Ngn 1/2) are expressed in the dorsal telencephalon, while Mash1 is detected exclusively in the ventral telencephalon. Loss of Ngn 1/2 in double mutant animals results in upregulation of Mash1 in the dorsal telencephalon, implying a role for Ngn 1/2 for suppression of Mash1 expression (reviewed in Schuurmans and Guillemot, 2001).
nervous system. A summary of mutant mice harbouring null alleles for various HLH genes is presented in Figure 1.8. As shown, loss of NeuroD results in deficits in the proliferation, differentiation and survival of granule cells of the hippocampus and cerebellum (Liu et al., 2000; Miyata et al., 1999). Similarly, mutations in NeuroD2 also result in defects in the differentiation and survival of cerebellar and hippocampal neurons (Olson et al., 2001; Schwab et al., 2000). Mutant mice which lack functional Math1 activity completely lack cerebellar granule cells, as well as hair cells in the inner ear and other non-neural cell types (Ben-Arie et al., 1997; Bermingham et al., 1999). Additionally, loss of Math1 function also results in the loss of a specific population of spinal interneurons, indicating a common role for Math1 in proper establishment of neuronal lineages in the central and peripheral nervous systems (Bermingham et al., 1999). Strikingly, there is significant conservation of function between atonal and its cloned mammalian homologue, Math1, as demonstrated mutant flies which contain a functional replacement of atonal with Math1 (Ben-Arie et al., 2000). The loss of chordotonal organs in atonal mutants is partially rescued by Math1, suggesting significant conservation in signalling networks between Drosophila and mammalian neurogenic programs (Ben-Arie et al., 2000).

1.2.3.3. Identification of a subset of neurogenic Class II HLH factors in the developing mouse telencephalon

Expression studies have defined a subset of neurogenic Class II HLH factors important for development of the telencephalon. An assessment of their temporal expression patterns, in combination with GOF and LOF analyses, further grouped these into determination and differentiation factors (Lee, 1997). The determination factors Mash1 and ngn1/2 are expressed early in germinal zones, and in non-complementary domains along the dorso-ventral axis. Mash1 is detected in ventral progenitors of the ganglionic eminence by e8.5 (Fode et al., 2000; Guillemot and Joyner, 1993), and ngn1/2 in dorsal progenitors (Fode et al., 2000; Sommer et al., 1996). Further, mutational analyses reveal that the exclusivity in domains of Mash1 and ngn1/2 expression is governed by distinct cross-regulatory mechanisms which restrict neuroprogenitors to a regional identity, while actively suppressing alternative cell fates within their respective domains (see Figure 1.9). For example, loss of ngn2 in mutant mice result in overexpression of Mash1 in the dorsal telencephalon, and this effect is significantly enhanced in ngn1/2 double homozygous mutant mice (Fode et
al., 2000). Importantly, functional replacement of the *ngn2* locus with *Mash1* induces ventral markers such as *Dlx-1* and *GAD67* (Fode et al., 2000), demonstrating differing capacities for *Mash1* and *ngn2* in programming ventral and dorsal progenitor cell fates, respectively. Additionally, NSCL expression is also detected in ventricular zone cells of the developing dorsal telencephalon (Begley et al., 1992; Murdoch et al., 1999; Uittenbogaard et al., 1999), though its precise role in corticogenesis is unclear.

In contrast to the determination factors, the differentiation factors (NeuroD, Math2, NDRF, NSCL, BETA3) may be more important for the maintenance rather than initiation of neuronal fates, since expression of these genes is traced to postmitotic neurons that are already committed to the neuronal phenotype, and are on their way to the cortical plate. NeuroD expression follows detection of Ngn expression in dorsal telencephalon (Fode et al., 2000; Sommer et al., 1996), and is localised to more superficial layers of the developing cortex, but is absent in cells of the ventricular zone (Fode et al., 2000; Sommer et al., 1996). Similarly, both Math2 (also known as NEX) and a NeuroD-related HLH factor, NDRF, are also detected in postmitotic neurons of the intermediate zone and cortical plate (Bartholoma and Nave, 1994; Shimizu et al., 1995), and their expression persists to adulthood.

Recently, a novel Class II HLH factor, BETA3, with atypical dimerisation behaviour (Peyton et al., 1996) was shown to exhibit a highly restricted expression pattern in the developing dorsal, not ventral, telencephalon (Kim et al., 2002). In particular, BETA3 was predominantly found in the cortical plate and subventricular zones at e15.5, with a progressive restriction to superficial layers of the cortex in the postnatal brain (Kim et al., 2002). Taken together, these observations indicate possible redundant roles for these HLH factors in signalling terminal differentiation in cortical neurons. However, in consideration of overlapping expression patterns, their precise downstream signalling mechanisms may not be readily elucidated through LOF analyses, as seen in analysis of Math2 (Schwab et al., 1998) and Math2/NeuroD double-mutant mice (Schwab et al., 2000) which present no obvious phenotype in the cortex (Figure 1.8).
1.3.1. The role of inhibitory Class V HLH factors in neuroprogenitor formation and timing of differentiation

Members of the Class V factors (according to Massari and Murre, 2000), otherwise known as the ID proteins (reviewed in Norton, 2000; and references herein) constitute a distinct subfamily of HLH proteins which lack a stretch of basic amino acids adjacent to the HLH region (see Figure 1.7). These proteins recognise and bind HLH factors of the Class I and II families, forming inactive ID-HLH heterodimers that are unable to contact DNA. It has been assumed that the principle role of ID proteins is to antagonise the functions of Class I and II HLH factors as dominant negative regulators of cell differentiation. However, recent reports indicate a role for ID proteins in lineage commitment and timing of differentiation during neurogenesis (reviewed in Norton, 2000) and gliogenesis (see review by Rowitch et al., 2002).

The first description of an ID protein emerged from studies of wing morphogenesis and sensory organ development in *Drosophila* (Garrell and Modolell, 1990). The *extramacrochaete* (*emc*) locus encodes an HLH protein which lacks a basic domain important for DNA-binding. The EMC protein heterodimerises with and antagonises several bHLH proteins, including those encoded by the *achaete-scute* complex genes involved in neurogenesis and sex determination (Ellis et al., 1990; Garrell and Modolell, 1990; Jan and Jan, 1993). To date, four ID proteins (ID1-ID4) have been isolated in mammals (Benezra, 1994; Biggs et al., 1992; Christy et al., 1991; Riechmann et al., 1994; Sun and Baltimore, 1991). All four proteins vary in length, from 119 amino acids (ID3) to 161 residues (ID4), and display high amino acid sequence identity within the HLH domain; there is extensive sequence divergence beyond the HLH domain. As expected, ID proteins antagonise Class I HLH signalling cascades by forming inactive dimers with these, thereby negatively regulating cell differentiation signals initiated by these (reviewed in Norton, 2000). For example, *Id3* is crucial for proper maturation of B-lymphocytes, a function which presumably involves E2A signalling in progenitor cells (Pan et al., 1999; Quong et al., 1999 see below). Further, overexpression of *Id1* in transgenic mice blocks B-lymphocyte maturation at the pro-B-cell stage (Sun, 1994). Similarly, *Id3* is required for proper selection of T-cells in the cell-mediated immune response in mammals (Bain et al., 1999; Blom et al., 1999; Rivera et al., 2000).
1.3.2. All four \textit{Id} genes are expressed in the developing mouse brain

During neurogenesis, all four mammalian \textit{Id} genes exhibit overlapping but non-identical expression patterns in the developing mouse forebrain (Jen et al., 1997). The earliest description of \textit{Id} expression in the neocortex is at embryonic day 9.5, when \textit{Id4} transcripts are detected in scattered cells of the prosencephalic vesicles (Jen et al., 1997). By embryonic day e11.5, \textit{Id1}, \textit{Id2}, and \textit{Id3} transcripts are detected in the medial aspect of the prosencephalon (which give rise to hippocampus and cingulate cortex), while \textit{Id4} is found throughout the entire telencephalon as neurogenesis progresses (Jen et al., 1997). In addition, it was observed that \textit{Id1} and \textit{Id3} predominantly marked mitotically active and less-differentiated neuroblasts located in the ventricular zone, and levels decrease as the progenitor pool diminishes during development. On the other hand, \textit{Id2} and \textit{Id4} transcripts were detected throughout the entire ventricular zone early in development, and followed a progressive temporal restriction to postmitotic neurons of the intermediate zone and cortical plate (Jen et al., 1997; Neuman et al., 1993b). Proper regulation of bHLH signalling by Id2 may also be required for genetic control of cortical regionalization in the mouse (Rubenstein et al., 1999). In the developing mouse neocortex, \textit{Id2} transcripts are localised in discrete rostro-caudal areal boundaries as early as day 18.5; this expression pattern remains distinct in layer V cells even when neocortical neurogenesis and migration are complete (Rubenstein et al., 1999). Together, these observations suggest, in addition to a role for \textit{Ids} in precursor cells, \textit{Id2} and \textit{Id4} may have ancillary roles in differentiated neurons of the cerebral cortex. Certainly, analysis of the \textit{Id2} proximal promoter sequence revealed multiple E-boxes which were regulated by Class I and II HLH factors (Neuman et al., 1995), suggesting that proper expression of \textit{Id2} is important for regulation of neurogenic HLH signalling cascades.

1.3.3. The role of IDs in neurogenesis

The overlapping expression pattern of \textit{Ids} in proliferating ventricular zone cells of the embryonic forebrain suggests these signal in a redundant manner. However, while analyses of mutant mice harbouring single deletions of \textit{Id} genes reveal subtly different contributions by these to the development of the immune system, the nervous systems of these mice exhibit no overt phenotype. For example, mice which lacked of \textit{Id2} (Yokota et al., 1999) or \textit{Id3} (Pan et al., 1999) exhibited defects in lymphopoiesis, while loss of \textit{Id1} (Yan et al., 1997) did not produce any overt morphological abnormalities and appeared normal.
Significantly, none of these strains of single-mutant mice displayed obvious neurological phenotypes, thereby complicating analyses of their neurogenetic contribution in isolation. In light of these observations, Lyden and colleagues generated homozygous double mutant mice harbouring deletions in both \textit{Id1} and \textit{Id3} to uncover their role in neurogenesis, vasculogenesis and angiogenesis (Lyden et al., 1999). Analysis of double mutant embryos revealed these had smaller brains, a feature attributable to a reduction in the numbers of proliferating cells in the telencephalon, and not excessive apoptosis (Lyden et al., 1999). This was observed as a decrease in the numbers of proliferating cells which expressed a marker for cell proliferation, Ki-67, as well as a concomitant increase in expression of cyclin-dependent kinase inhibitors (CDKIs) such as p16 (Lyden et al., 1999). Importantly, these mice also exhibited premature neural differentiation in the developing cortex, observed as precocious expression of neurogenic Class II HLH differentiation factors such as \textit{NeuroD1} and \textit{Math2} (Lyden et al., 1999). Taken together, these observations demonstrate a requirement for both \textit{Id1} and \textit{Id3} in regulating neuroblast proliferation and timing of neuronal differentiation in a redundant manner during neurogenesis.

In the developing avian nervous system, ectopic expression of \textit{Id2} in ectodermal precursors directs cells to adopt a neural rather than epidermal fate (Martinsen and Bronner-Fraser, 1998). While a precise signalling pathway for neural crest specification has yet to be elucidated, overexpression of \textit{Id2} in these precursors presumably represses an intact HLH signalling cascade which normally signals epidermialization in ectodermal precursors. Additionally, ID proteins are also important for timing for oligodendrocyte differentiation (Kondo and Raff, 2000a; Kondo and Raff, 2000b; Wang et al., 2001). These observations indicate that a precise temporal and spatial expression of ID proteins is crucial for proper coordination of cellular proliferation and timing of differentiation in multiple cell lineages.

1.3.4. The role of IDs as negative regulators of signalling by HLH and Pax transcription factors

Numerous studies have demonstrated biochemical (Ellis et al., 1990; Langlands et al., 1997; Neuman et al., 1995; Peverali et al., 1994; Prabhu et al., 1997; Sun et al., 1991) and functional (Lyden et al., 1999; Rivera et al., 2000; Sun, 1994; Yan et al., 1997) evidence for ID proteins in regulating HLH signalling cascades in a dominant-negative manner. Inhibition involves formation of inactive heterodimers to Class I (Ellis et al.,
1990; Langlands et al., 1997; Peverali et al., 1994; Prabhu et al., 1997; Sun et al., 1991) or Class II (Ellis et al., 1990; Langlands et al., 1997) HLH factors. While all ID proteins readily heterodimerise with E proteins (Langlands et al., 1997; Loveys et al., 1996; Sun et al., 1991), implying multiple redundant functions by these, only ID1 and ID2, but not ID3, heterodimerises with the Class II myogenic bHLH factors MyoD and Myf-5 (Langlands et al., 1997). Furthermore, ID1, 2 and 3 could not bind Class II haematopoietic factors such as Tal-1, Tal-2 and Lyl-1, highlighting different signalling mechanisms by ID proteins in regulating myogenesis and haematopoiesis, respectively. That is, control of muscle differentiation programs by ID proteins may involve direct binding to myogenic Class II factors, while IDs modulate haematopoiesis through binding and sequestration of Class I HLH factors, thus regulating the pool of Class I monomers available for heterodimerisation to Tal-1, Tal-2 and Lyl-1 (Langlands et al., 1997). While experiments with neurogenic Class II factors were not performed, these data demonstrate different IDs may negatively regulate developmental programs through specific interactions with Class I and/or II HLH factors \textit{in vivo}.

Recently, it was reported that ID proteins could interact with non-HLH transcription factors, suggesting a more widespread regulatory influence over developmental programs (reviewed in Norton, 2000). Work by Roberts and colleagues (Roberts et al., 2001) demonstrated that ID1, ID2 and ID3 proteins bound the paired-domain transcription factors Pax-2, Pax-5 and Pax-8 to antagonise transcriptional activation and DNA binding (Roberts et al., 2001). Importantly, this inhibitory role for IDs is mediated by their HLH domain, suggesting cross-modulatory roles for ID proteins in antagonism of HLH and Pax signalling cascades.

1.3.5. IDs regulate cell cycling through interactions with HLH factors and the Retinoblastoma protein

The role of ID proteins in cell cycle regulation is implied through analysis of \textit{Id} expression in synchronised cells which reveals a periodicity in gene expression, with high levels in proliferating cells followed by a progressive decrease upon differentiation (Loveys et al., 1996). These observations are recapitulated in models of neurogenesis \textit{in vitro} which demonstrate down-regulation of \textit{Id} gene expression in neuroblasts.
which exit the cell cycle and express neuronal markers (Einarson and Chao, 1995; Jogi et al., 2002; Nagata and Todokoro, 1994; Neuman et al., 1993b).

Evidence for a role for ID proteins in cell cycle regulation arose from GOF and LOF approaches in vitro (Alani et al., 1999) and in vivo (Lyden et al., 1999; Sun, 1994; Wice and Gordon, 1998; Yan et al., 1997). Overexpression of \( Id1 \), and not \( Id2 \) or \( Id3 \), can immortalise primary human keratinocytes, indicating that \( Id1 \) can function as a dominant oncogene (Alani et al., 1999). Similarly, enforced expression of \( Id1 \) targeted to intestinal epithelial cells is associated with development of adenomas (Wice and Gordon, 1998), while constitutive expression impairs differentiation of B cells, possibly through HLH cascades which involve E proteins (Sun, 1994). Conversely, loss of \( Id \) results in reduced tumorigenicity of xenografts to nude mice (Lyden et al., 1999), an observation that complements overexpression analyses.

While these functional experiments implicate a role for ID proteins in cell cycle regulation and tumorigenesis, accumulating biochemical evidence suggests these signal through modulation of HLH signalling cascades, as well as through interaction with retinoblastoma protein (pRB) (reviewed in Norton, 2000). Transient transfection assays with E proteins demonstrate a role in programming growth suppression through activation of CDKI expression (Pagliuca et al., 2000; Peverali et al., 1994 see below). Predictably, ID proteins can antagonise transcriptional activation potential of these factors through formation of inactive heterodimers (Pagliuca et al., 2000; Peverali et al., 1994). Significantly, this capacity for HLH dimerisation is regulated by phosphorylation of IDs by cyclin-dependent kinase CDK2 (Deed et al., 1997; Hara et al., 1997).

In addition to their capacity to negatively regulate CDKI expression through antagonism of Class I HLH factors, IDs can also regulate G1-S phase transition through a non-HLH signalling pathway. Work by Lasorella and colleagues revealed that ID2 and ID4 bind the G1-S checkpoint factors pRB, p107 and p130 (Lasorella et al., 2002; Lasorella et al., 1996), thereby releasing these from their association with positive regulators of S-phase genes, such as the ETS-domain transcription factor E2F (Yates et al., 1999). This convergence in signalling between ID2 and pRB is further demonstrated in analysis of mutant mice that
lack both pRB and Id2 (Lyden et al., 1999). While embryos that lacked pRB alone did not survive beyond e14.5, a phenotype attributable to dysregulation of cellular proliferation, mutation of Id2 in these pRB mutant mice rescued the lethal phenotype (Lyden et al., 1999).

A summary of the key roles for ID proteins in proliferation and differentiation is presented in Figure 1.10. During proliferation, ID proteins can bind pRB, thereby releasing pRB-bound E2F to allow an upregulation of E2F-responsive genes important for S-phase progression. Concomitantly, ID proteins repress transcription-factor dependent activation of differentiation programs, through direct antagonism of HLH factors as well as Pax proteins. Importantly, the potential for protein-protein interaction by IDs is regulated through phosphorylation of key residues on the polypeptide, a role which is regulated by cyclin-dependent kinases (CDKs) and their associated inhibitors (CDKIs).

Taken together, these observations suggest a dual role for IDs in regulating G1-S cell cycle progression through modulation of two key signalling pathways: (i) negative regulation of E protein-mediated CDKI expression; and (ii) sequestration of Rb and associated proteins to promote S phase entry. In light of their exclusivity for regulation of pRB and associated proteins, the activities of Id2 may be similar to Id4, but these are different to the activities of Id1 and Id3. This dichotomy is also observed in their patterns of gene expression: during embryogenesis, all four Ids are expressed in overlapping domains, but transcripts encoding Id2 and Id4, and not Id1 nor Id3, persists in postmitotic neurons of the developing cortex, indicating these (Id2 and Id4) proteins are involved in signalling proliferation, as well as instructing neurodifferentiation. These data highlight crucial roles for IDs during development, and are important considerations when ascribing a role for the related Class I HLH factors in vivo.

1.4.1. The role of Class I HLH proteins in neurogenesis is not well understood

While Class II and V factors exhibit demonstrable capabilities for programming neuroprogenitor formation and proper differentiation, the role of Class I HLH factors, also known as E proteins, remains ambiguous. The Class I HLH factor E12 is found in flies, rodents, humans and nematodes. Although this transcription factor is expressed in proliferating cells of the developing nervous systems of these organisms, its precise
Figure 1.10. Class V HLH proteins promote cell cycle progression whilst suppressing cell differentiation. These inhibitory ID proteins negatively regulate the transcriptional activities of HLH and Pax proteins by forming inactive dimers with these. Similarly, ID proteins regulate G1-S phase progression by binding the retinoblastoma protein (pRB), thereby releasing E2F for activation of downstream genes to signal S-phase progression (Adapted from Norton, 2000). This capacity for protein-protein interaction by ID is regulated through phosphorylation by cyclin dependent kinase 2 (CDK2).
contribution to neurogenesis is still loosely defined. Furthermore, a direct neurogenic role for E12 may not easily be uncovered through classical genetic dissection, since LOF analyses in mammals are complicated by multiple E proteins exhibiting complementary patterns of gene expression, as well as redundant signalling roles. Rather, accumulating evidence suggests E12 signals neurogenesis through heterodimerisation with Class II or VI factors to control neuroprogenitor proliferation and differentiation in a redundant manner.

1.4.2. The role of Class I HLH factors in neurogenesis: lessons from flies and worms

In Drosophila, there is only one Class I HLH protein, encoded by the daughterless (da) gene. Owing to its exclusivity, daughterless is a significant contributor to a multitude of developmental programs, including sex determination and neurogenesis (reviewed by Massari and Murre, 2000; Moore et al., 2000). During sex determination in the embryo, Da protein is required for the early transcriptional activation of Sex lethal (Sxl) as a heterodimer with another bHLH protein, sis-b, to produce female progeny (Deshpande et al., 1995). In female progeny, da is additionally required for proper oogenesis (Smith and Cronmiller, 2001). Da is also important for the establishment of the proneural field that gives rise to the central (CNS) and peripheral (PNS) nervous systems (Caudy et al., 1988a). An investigation of male Drosophila embryos which lack Da protein revealed a role in neuronal precursor differentiation, and not neuronal precursor formation (Vaessin et al., 1994). While lack of Da in these embryos did not significantly affect the numbers of neuroblasts which expressed the proneural markers scute and hunchback, expression of other neuronal markers were severely perturbed; the PNS-specific gene couch potato was undetectable, while levels of the proneural genes deadpan and asense were drastically reduced or undetectable (Hassan and Vaessin, 1997; Vaessin et al., 1994). In addition, the expression of cyclin A was reduced in CNS precursor cells, and completely absent in neuronal precursors of the PNS (Hassan and Vaessin, 1997; Vaessin et al., 1994). These observations provided the first indication that the function of E proteins was critical to proper development of the CNS and PNS.

Defects in neurogenesis are also observed in nematodes which harbour targeted deletions of HLH genes (Portman and Emmons, 2000). While the C. elegans genome lacks HLH proteins of the ID class (Ruvkun
and Hobert, 1998), nematodes require Class I and II HLH genes for proper neurogenesis. Expression studies of the homologue for Daughterless in *C. elegans*, known as CeE/DA (Krause et al., 1997) or HLH-2 (Portman and Emmons, 2000), show that most neuronal precursors in 350-cell stage embryos express CeE/DA, with a progressive down-regulation as differentiation and morphogenesis occurs (Krause et al., 1997). This dynamic and restricted expression of CeE/DA in neuronal precursors strongly implicate a role in pattern formation and neurodevelopment in nematodes (Krause et al., 1997).

An accessory role for CeE/DA/HLH-2 in programming neurogenesis was subsequently uncovered through analysis of genes important for neuronal subtype specification of the male tails of *C. elegans*. Animals with targeted mutations in the coding regions of the Class II HLH factor *lin-32*, a member of the atonal family of HLH genes, exhibited defects in lineage specification of multiple neuronal subtypes in the developing male tails. While mutations in *hlh-2* alone did not affect male tail development, loss of *hlh-2* function augmented the neuronal phenotype of *lin-32* mutants, as *lin-32/hlh-2* double-mutant animals completely lacked tail sensory rays (Portman and Emmons, 2000). Significantly, these *lin-32/hlh-2* double-mutant animals harboured point-mutations in their HLH domains, resulting in proteins which failed to dimerise and bind an E-box sequence (Portman and Emmons, 2000). From these observations, it is clear that E proteins are necessary as cofactors for proper development of the nervous system in nematodes and flies, a function that requires an intact HLH domain. Additionally, these observations also indicate that resolution of the entire spectrum of binding partners to E proteins in neuronal cells may uncover their true role in neurodevelopment, when LOF approaches are uninformative.

1.4.3. A neurological phenotype is not reported in mice which lack functional Class I HLH genes

While the roles of E proteins, such as E2A, E2-2 and HEB, in lymphopoiesis and haematopoiesis are well understood (for recent reviews see Engel and Murre, 2001; Quong et al., 2002; Rivera and Murre, 2001), their precise contribution to neurogenic signalling cascades remains to be elucidated. For example, mice which lack a functional E2A gene exhibit defects in B- (Bain et al., 1994; Zhuang et al., 1994) and T- (Yan et al., 1997) cell lineage commitment and differentiation in a dose-dependent manner. Additionally, these mutant mice exhibit high frequencies of postnatal death as well as a severe reduction in the population of
mature B- and T- cells. The capacity for E2A to program B-cell differentiation is a cell autonomous effect, since wild type haematopoietic stem cells (HSCs) can provide full radioprotection when grafted into E2A mutant mice (Zhuang et al., 1998). Furthermore, despite earlier descriptions of E2A gene expression in the developing central nervous system which implicated a role for neurogenesis (Roberts et al., 1993 see below), the brains of these mutant mice did not exhibit any gross abnormalities (Bain et al., 1994; Zhuang et al., 1994). Interestingly, there is a significant under-representation of female homozygous mutant pups which survive the first week, and these female offspring are sterile. These observations are reminiscent of the role of daughterless in sex determination (Deshpande et al., 1995; Yang et al., 2001) and proper oogenesis in female progeny (Smith and Cronmiller, 2001), but provide no direct evidence for a neurogenic role for E2A gene products. As a corollary, in vitro differentiation assays using ES cells harbouring hypomorphic alleles for E2A did not show any obvious defects in skeletal or cardiac myogenesis, erythropoiesis, chondrogenesis and neurogenesis (Zhuang et al., 1992). In a different approach, the mouse embryocarcinoma cell line P19 was employed to evaluated the potential for neurodifferentiation upon overexpression of Class I and II HLH proteins (Farah et al., 2000). While transient expression of Class II HLH factors such as Mash1 could program neurogenesis in uncommitted P19 cells, overexpression of E12 failed to program neurogenesis in these (Farah et al., 2000). Taken together, these observations indicate that at the cellular level, LOF and GOF approaches may be uninformative when addressing the role of E2A gene products in neurodevelopment.

Mutations in other Class I HLH genes highly similar to E2A also display similar phenotypes in mutant mice. For example, loss of E2-2 (Bergqvist et al., 2000; Zhuang et al., 1996) or HEB (Zhuang et al., 1996) results in mice which exhibit deficiencies in the B-cell compartment, and demonstrate a high incidence of postnatal lethality, akin to E2A null mutant mice. Once again, the brains of these mice appear outwardly normal, though a precise neurological phenotype remains undocumented.

Most strikingly, redundancy in signalling by Class I HLH factors is revealed in mutant mice harbouring a functional replacement of the E2A gene with the human Class I factor HEB (Zhuang et al., 1998). Expression of human HEB in the E2A locus (denoted as E2A^{HEB}) can fully compensate for E2A-dependent
programming of B-cell differentiation. This compensatory function by HEB gene expression occurs in a
dose-dependent manner, whereby transheterozygous E2A^{HEB}/E2A^{HEB} mice harbour significantly fewer
mature B cells when compared to homozygous E2A^{HEB}/E2A^{HEB} mice (Zhuang et al., 1998). This extreme
model of functional redundancy demonstrates sufficient conservation of Class I HLH protein domains
which signal B-cell differentiation. In consideration of their combinatorial roles in signalling
lymphopoiesis, the structural similarities of these polypeptides, as well as their overlapping expression
patterns in developing cortex (see below), it is likely that these Class I HLH factors direct neurogenesis in a
redundant manner. Therefore, the lack of observable neurological phenotype in the aforementioned mutant
mice suggests that these alone may not be useful tools for dissection of the neurogenic role for Class I HLH
factors. Nevertheless, breeding of double homozygous mutants which lack multiple Class I HLH factors
may uncover subtle contributions by these genes to corticogenesis. Alternatively, close examination of the
biochemical function of these factors, including elucidation of key signalling pathways that involve Class I
factors, may prove a more informative approach.

1.4.4. The Class I HLH factor E12 contains functionally separable domains

Members of the Class I HLH factors display extremely high polypeptide sequence conservation within
(Henthorn et al., 1990; Murre et al., 1989a; reviewed in Littlewood and Evan, 1998; Murre et al., 1994), as
well as beyond (Goldfarb et al., 1998; Henthorn et al., 1990; Massari et al., 1999; Murre et al., 1994), their
HLH domains. The Class I HLH factor E12 was originally cloned by Murre and coworkers as one of two
HLH proteins which were important transcriptional regulators of immunoglobulin genes (Murre et al.,
1989b). Both proteins, E12 and E47, could bind the kE2 E-box sequence of the immunoglobulin kappa
chain enhancer (Murre et al., 1989b), a function which requires the basic amino acid region for DNA
binding (Murre et al., 1989b) as well as E-box sequence selection (Vitola et al., 1996). Importantly, both
E12 and E47 are bound to E-boxes as homo- or heterodimeric complexes (Voronova and Baltimore, 1990),
and this function is mediated through conservation of key residues within the HLH domain (Shirakata et
al., 1993). A diagrammatic representation of E12 polypeptide is provided in Figure 1.11. The bHLH
domain of E12 is encoded in amino acids 540 to 601. Additionally, a minimal domain adjacent to the
second helix of the HLH domain (amino acids 598 to 618) is indispensable for in vivo dimerisation. This
**Figure 1.11. Functional domains within the E12 polypeptide.** In addition to the well characterised bHLH domain, the adjacent C domain can regulate dimerisation. E12 homomerises poorly, owing to an inhibitory domain which is acidic in nature. Two transcriptional activation domains (ADI and ADII) in its N-terminus mediate target gene expression upon DNA binding. Additionally, an LFDS domain within ADI signals chromatin condensation through recruitment of histone acetyltransferases. This polypeptide is localised to the nucleus by a conserved NLS signal.

<table>
<thead>
<tr>
<th>Binding partner to E12</th>
<th>Description of protein</th>
<th>Evidence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mash1</td>
<td>Class II HLH factor</td>
<td>EMSA; Reporter Assays; Yeast 2-hybrid assay</td>
<td>Johnson et al., 1992; Farah et al., 2000</td>
</tr>
<tr>
<td>Mash2</td>
<td>Class II HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Johnson et al., 1992; Farah et al., 2000</td>
</tr>
<tr>
<td>Math1</td>
<td>Class II HLH factor</td>
<td>EMSA;</td>
<td>Akazawa et al, 1995</td>
</tr>
<tr>
<td>Math2</td>
<td>Class II HLH factor</td>
<td>EMSA;</td>
<td>Shimizu et al., 1995</td>
</tr>
<tr>
<td>ID1</td>
<td>Class V HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Sun and Baltimore, 1991</td>
</tr>
<tr>
<td>ID2</td>
<td>Class V HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Sun and Baltimore, 1991</td>
</tr>
<tr>
<td>ID3</td>
<td>Class V HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Loveys et al., 1997</td>
</tr>
<tr>
<td>UBC9/UBCE2A</td>
<td>Ubiquitin conjugating enzyme</td>
<td>Yeast 2-hybrid assay; In vitro ubiquitination assays</td>
<td>Loveys et al., 1997; Kho et al., 1997</td>
</tr>
<tr>
<td>BETA3*</td>
<td>Class II HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Peyton et al., 1996</td>
</tr>
<tr>
<td>Hes1*</td>
<td>Class VI HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Sasai et al., 1992</td>
</tr>
<tr>
<td>Hes5*</td>
<td>Class VI HLH factor</td>
<td>EMSA; Reporter Assays</td>
<td>Akazawa et al., 1992</td>
</tr>
</tbody>
</table>

**Figure 1.12. Summary of known binding partners to E12.** Positive interactions were detected through various biochemical approaches, including electrophoretic mobility shift assays (EMSA) using an E-box oligonucleotide sequence, yeast 2-hybrid assays or heterologous reporter assay systems. (*) denotes probable interaction partners to E12, implied through their capacity to bind to E47.
adjacent segment, named Domain C, is highly conserved in E proteins and appears to influence the *in vivo* conformation of the adjacent HLH domain through a mechanism which involves chaperonins (Goldfarb et al., 1998). Nuclear localisation of this transcription factor is coded in amino acids 170 to 176 (Deed et al., 1996).

Analyses of E12 and E47 cDNA sequences revealed these transcripts were encoded by alternative splicing of the E2A gene (Bain et al., 1994; Murre et al., 1989b). The functional significance of alternative splice forms was subsequently clarified through characterisation of an inhibitory domain only found on E12 that prevents homodimerisation (Sun and Baltimore, 1991), and not DNA binding (Shirakata and Paterson, 1995). This domain was fine-mapped to comprise 19 amino acids N-terminal to the basic domain of E12 (see Figure 1.11), and is acidic in nature. Mutation analysis of this domain uncovered the importance of the negative charge conferred by this domain for inhibition of E12 homomerisation, since substitution to positively charged amino acids abolished this function (Sun and Baltimore, 1991). The presence of the inhibitory domain, however, does not preclude homodimerisation by E12. Homodimers of E12 may arise through disulphide linkage between monomers (Benezra, 1994), a function which requires ubiquitously expressed protein disulphide isomerases (Markus and Benezra, 1999). Nevertheless, oxidised homodimers of E12 represent only a small fraction in physiologically relevant systems (Sloan et al., 1996), indicating that the inhibitory domain confers a preference for E12 to form heterodimers with other HLH factors *in vivo*.

In contrast to the inhibitory domain on E12, homo- and heterodimerisation may also be regulated through phosphorylation of key serine residues adjacent to the bHLH domain. Phosphorylation of serines 514 and 529 on E12 and E47 can, in part, prevent DNA binding by homodimers and not heterodimers (Sloan et al., 1996). Further, binding studies indicate that heterodimers of E47 and Class II HLH factors bind with 10-fold greater affinity, compared with homodimers (Sun et al., 1991). Taken together, these data indicate that E2A proteins signal predominantly as heterodimeric complexes *in vivo*. 
The E12 polypeptide encodes two transcriptional activation domains (Figure 1.11). The first, named ADI, is encoded by amino acids 1 to 100 and is rich in serine and glycine residues (Aronheim et al., 1993). This domain resembles an \( \alpha \)-helical structure, and is indispensable for programming growth arrest in NIH3T3 fibroblasts (see below). Furthermore, within the ADI domain lies a conserved LFDS domain (amino acids 21 to 24) which can recruit histone acetyltransferases, thereby facilitating chromatin decompaction of target genomic DNA through direct histone acetylation (Massari et al., 1999). The second transactivation domain, ADII, rich in serines and leucine residues, is encoded in amino acids 323 to 478 (Aronheim et al., 1993). Curiously, while ADI functions efficiently in a variety of mammalian cells, the second transactivation domain, ADII, resembles a leucine zipper and demonstrates preferential transactivation potential in pancreatic beta cell lines, suggesting cell-specificity by the latter domain (Aronheim et al., 1993). Mechanistically, this specificity may be coded through preferential heterodimerisation to select Class II factors in a lineage-restricted cell type, resulting in repression, or augmentation, of target gene transcription through these distinct transactivation domains (Markus et al., 2002).

The ADII domain was fine-mapped by Quong and coworkers to a conserved element comprising amino acids 279 to 345, and demonstrated to have conserved transactivation potential in mammalian cells and even in yeast cells (Quong et al., 1993). These transcriptional activation domains play a critical role in cellular transformation as chimeric oncoproteins which arise from chromosomal translocations, giving rise to E2A-HLF and E2A-Pbx1 in patients with hereditable acute lymphoblastic leukaemia (Inaba et al., 1992; Lu et al., 1994; Van Dijk et al., 1993; Yoshihara et al., 1995). Additionally, work by Markus and coworkers (Markus et al., 2002) characterised a repressor domain which was conserved in multiple E protein species, including E12. This repressor domain is a charged 30 amino acid region which is highly conserved among vertebrate E proteins, and can specifically silence both AD domains of E proteins.

### 1.4.5. Biochemical characterisation of E12 reveals a role for cell proliferation and differentiation

#### 1.4.5.1. E12 heterodimerises with HLH and non-HLH factors

Owing to a lack of observable neurological defects in mutant mice which lack E2A function, biochemical analyses, including an evaluation of HLH dimerisation partners important for brain development, may be
more informative for elucidating its true signalling roles during neurogenesis. A number of studies have been conducted and these are summarised in Figure 1.12. E12 has been shown to heterodimerise with Class II, V, VI HLH factors, including Mash1 (Johnson et al., 1992a), Mash2 (Johnson et al., 1992a), Math1 (Akazawa et al., 1995), Math2 (Shimizu et al., 1995), Id1 (Sun et al., 1991), Id2 (Sun et al., 1991), Id3 (Loveys et al., 1996). In addition, E12 probably heterodimerises with BETA3 (Peyton et al., 1996), Hes1 (Sasai et al., 1992) and Hes5 (Akazawa et al., 1992). Importantly, the neurogenic Class II HLH factors homodimerise poorly, as evidenced through electrophoretic mobility shift assays performed with Math1 (Akazawa et al., 1995), Math2 (Shimizu et al., 1995), NeuroD (Peyton et al., 1996), Mash1 (Gradwohl et al., 1996) and MATH4A/ngn2 (Gradwohl et al., 1996). These observations indicate that heterodimerisation to Class II HLH factors may be important for proper signalling by E2A proteins, albeit in a redundant fashion (Bain et al., 1994; Zhuang et al., 1994; see below). In addition, the HLH domain of E12 also mediates dimerisation to non-HLH proteins, such as the ubiquitin-conjugating enzyme UBC9/UbcE2A (Kho et al., 1997; Loveys et al., 1997). This enzyme binds the HLH domain of E2A proteins for proper ubiquitination and targeting for subsequent proteasome degradation (Loveys et al., 1997).

### 1.4.5.2. A direct role for E12 in cell cycle regulation

Speculation of a role for E12 in cell cycle regulation was fuelled by the periodicity of gene expression observed in cell culture models (Nielsen et al., 1992; Walker et al., 1990). Experiments with Balb/c 3T3 cells demonstrated that E2A transcripts are expressed in quiescent cells and are detectable upon serum stimulation, with marked down-regulation seen at 2-4 hours of culture, but increase as cells approach a second cell cycle (Loveys et al., 1996). Direct evidence for a growth suppressor role was demonstrated through functional experiments using transiently transfected NIH3T3 fibroblasts (Pagliuca et al., 2000; Peverali et al., 1994; Prabhu et al., 1997). Overexpression of E2A proteins led to growth arrest, a function which requires direct transcriptional activation of cyclin-dependent kinase inhibitors (CDKIs), such as p21, p15 and p16 (Pagliuca et al., 2000; Prabhu et al., 1997). Cyclin-dependent kinase inhibitors bind cyclin-CDK complexes, resulting in the inhibition of kinase activities essential for all phases of cell cycle transition (Elledge, 1996; Sherr and Roberts, 1995). Moreover, this potential for E12 to signal growth
suppression may be modulated through phosphorylation of key serine residues by cyclins (Chu and Kohtz, 2001).

Direct transcriptional regulation of CDKIs was validated upon evaluation of multiple E-boxes on the cis-promoter of p21 for binding by E2A proteins (Pagliuca et al., 2000; Prabhu et al., 1997). Deletion of any of these E-boxes negatively affects expression of a luciferase reporter under the control of this promoter region. Furthermore, overexpression of Id2 can suppress growth inhibition mediated by E2A, and silence activation of a luciferase reporter under the control of the p21 cis-promoter sequence. Similarly, deletion of the HLH domain precludes E47-mediated growth suppression of transiently transfected fibroblasts (Peverali et al., 1994).

In addition to direct regulation of growth suppression through CDKI expression, E12 may also regulate the expression of Id2 in vivo. Analysis of the proximal promoter of this gene identifies several E-boxes which regulate its expression, and are bound by HLH factors such as NSCL, as well as the E proteins ME1 and ME2 (Neuman et al., 1995). Coexpression of Id2 abrogates activation by active HLH dimers, indicating a negative feedback loop that eventually silences Id2 expression (Neuman et al., 1995). While the transactivation potential of E12 was not tested on this promoter, E-box sequences present on the Id2 promoter are in good correspondence with consensus sequences for E12 binding and activation (Neuman et al., 1995). Taken together, these findings demonstrate a direct role for E12 in programming growth suppression through activation of CDKIs, and this function is mediated by heterodimerisation to inhibitory HLH proteins. Further, E12 may directly regulate expression of its inhibitory partner Id2, thereby establishing a negative feedback loop which abrogates the growth suppressor function of E12, and maintain expression of Id2 in postmitotic neurons (Neuman et al., 1993b).

1.4.5.3. E12 may behave as a molecular chaperone for Id proteins

While it is clear that ID proteins elicit a role in the nucleus through antagonism of E12-mediated target gene transcription, surprisingly, all ID proteins lack a nuclear localisation signal. This discrepancy was clarified through experiments that demonstrated a role for E proteins as a molecular chaperone which
facilitated nuclear entry of IDs (Deed et al., 1996). The authors demonstrated that while IDs were predominantly found in the cytoplasmic compartment, co-expression with E47 resulted in their nuclear localisation (Deed et al., 1996). Furthermore, nuclear import by E47 increased the stability of IDs, while augmenting degradation of E47, possibly through a ubiquitin-mediated pathway (Kho et al., 1997; Loveys et al., 1996; Loveys et al., 1997).

The biological relevance of these findings is demonstrated through correlation of the subcellular localisation of ID proteins with timing of neuroprogenitor cell differentiation (Wang et al., 2001). Studies of oligodendrocyte precursor cells (OPC) induced to differentiate upon withdrawal of platelet-derived growth factor (PDGF) demonstrate that while ID2 is predominantly localised to the nucleus of precursor cells, there is progressive accumulation in the cytoplasm upon differentiation (Wang et al., 2001). These findings raise the possibility that a similar mode of regulation may exist in neuronal progenitors.

1.4.5.4. The role of E2A proteins in direct regulation of cadherin-mediated cell adhesion

Recently, a direct role for E2A proteins in mediating cell adhesion was demonstrated through transcriptional repression of the calcium-dependent cell adhesion molecule, E-cadherin (Perez-Moreno et al., 2001). Screening for transcription factors which bound E-boxes on the repressor element of the E-cadherin promoter led to cloning of E47 through a yeast one-hybrid screen (Perez-Moreno et al., 2001). Experiments with an epithelial cell line demonstrated that overexpression of E47 resulted in a dramatic reduction in E-cadherin protein expression in stably transfected cells. Further, these transformed cells exhibited metastatic and invasive properties not observed in mock transfected cells, with demonstrable tumorigenicity in xenograph experiments with nude mice (Perez-Moreno et al., 2001).

These observations suggest that E12 may also have a direct role in mediating cadherin-dependent adhesion in ventricular zone cells of the developing telencephalon. It is well documented that during neurulation, prospective neural tissue separates from surface ectoderm and, in the process, changes cell adhesion properties through switching expression from E-cadherin to N-cadherin (reviewed in Redies, 2000). This switch is crucial for proper neurogenesis through promotion of a migratory phenotype, allowing
neuroprogenitor cells to loose their adhesive properties in order to invade surrounding areas for proper localisation of mature neurons (reviewed in Larue et al., 1996; Redies, 2000). Certainly, E12 could play a direct role in repressing E-cadherin expression in neuronal progenitors, allowing for migration of postmitotic neurons away from the ventricular zone.

1.4.5.5. Expression studies suggest a role for E2A early in neurogenesis

Studies of the expression of E2A during embryodevelopment implicate a role for E12 in the early stages of cerebral corticogenesis. *In situ* hybridisation techniques reveal E2A expression as early as embryonic day e8.5 in the mouse, where transcripts are detected in all tissues except non-neural ectoderm, heart primordium and extraembryonic membranes apart from the allantois (Perez-Moreno et al., 2001). The strongest expression is in the proliferative zones lining the cerebral ventricles (Roberts et al., 1993). A similar pattern of E2A protein expression was also observed in fetal kidney, lung, intestine and testis of humans (Rutherford and LeBrun, 1998). Additionally, in fetal human spinal cord, E2A protein is expressed in the subventricular zone, and partially overlaps expression of the cell cycle marker Ki-67 (Rutherford and LeBrun, 1998). Expression of Ki-67 is highest in the ventricular zone, suggesting that E2A marks precursor cells which are still cycling, but are physically removed from the germinal zone in the developing spinal cord.

As neurogenesis proceeds in the embryo, E2A expression is progressively downregulated in most tissues apart from the ventricular zone of the developing cortex. By adulthood, E2A is only weakly detectable in adult brain as diffuse staining in the ependymal cell layer lining the walls of the cerebral ventricles, and lowly expressed in other areas in the forebrain, such as the piriform cortex, olfactory tubercle, and the islands of Calleja (Roberts et al., 1993). A weak signal was also apparent over the magnocellular neurons of the hypothalamic paraventricular supraoptic nuclei, in the granule cell and pyramidal cell layers of the hippocampus, and in Purkinje and granule cells of the cerebellum (Roberts et al., 1993). Taken together, these data indicate that E2A may be more important for early neurogenetic signalling events in proliferating cells of the developing cortex.
While these data demonstrate restriction of E2A expression predominantly to proliferative zones, the use of probe sequences common to both splice forms did not allow for an evaluation of relative levels of E12 and E47 in all tissues tested. To this end, several groups have reported that while the predominant mRNA species from the E2A locus is E47 (Shen and Kadesch, 1995; Watada et al., 1995), there are differences in the relative levels of E12 and E47 in adult organs (Watada et al., 1995). An investigation by Watada and coworkers (Watada et al., 1995) indicates that there is a 2:1 ratio of E47 transcripts compared to E12 in adult muscle, brain, spleen, heart and lung. Further, experiments using a cell culture model of myogenesis demonstrated that while E2A levels decrease during differentiation, the ratio of E47 to E12 remained unchanged (Watada et al., 1995). Similarly, E47 is the majority species in B cells and Hela cells (Shen and Kadesch, 1995). Conversely, the ratio of E47 to E12 is no greater than 1.5:1 in pancreas, liver and testis. While the physiological relevance for distinct stoichiometric ratios in different organs is unclear, these observations may suggest isoform preference with respect to organs exhibiting different levels of E47, compared with E12. For example, organs with a 2:1 ratio of E47 to E12 may preferentially signal through E47, perhaps as homodimers, while the majority of E2A proteins in pancreas, liver and testis (which exhibit a ratio of <1.5:1) signal as heterodimers to other HLH factors in these tissues.

In an alternative approach, Zhuang and colleagues (Zhuang et al., 1998) generated mutant mice which targeted deletion of E47, and not E12, to elucidate its role in B-cell development. While mutant mice lacking both E47 and E12 exhibit severe deficiencies in B-cell development in juveniles and adults (Bain et al., 1994; Zhuang et al., 1994), mutant mice which only lack E47 exhibit deficiencies in B-cell production in neonates, but not in adulthood (Zhuang et al., 1998). These observations could suggest that the different isoforms of E2A may perform non-identical roles in signalling B-cell lineage commitment at different stages of development, though their roles in neurogenesis are still largely unknown.

1.4.5.6. Overlapping expression patterns of Class I HLH genes in the developing brain imply a redundant role in signalling neurogenesis

Expression analyses indicate that, in addition to E12, other Class I HLH factors, such as ME1 (ME1 is mouse HEB; HEB aka HTF4 or TCF12, Unigene cluster Hs. 326198) and ME2, are expressed in the
developing nervous system in overlapping but non-identical domains. During embryogenesis, ME1 expression closely mimics the distribution of E2A transcripts, and is detected in ventricular zone cells of the developing cortex, with progressive restriction as development proceeds (Chiaramello et al., 1995). In adult brain, ME1 is undetectable in the cerebral cortex, but expression persists in cells of the hippocampus and cerebellum (Chiaramello et al., 1995).

The expression of the Class I HLH factor ME2 shows many differences from the expression of ME1 and E2A (Soosaar et al., 1994). RNA analysis indicates ME2 is expressed as a single 6kb species, and is detected in mouse embryos by e12.5, with staining in the ventricular zone of the telencephalon, as well as in cerebella primordia, pons, medulla, myotome (where E2A is absent), spinal cord and developing limbs. At e16.5, staining is evident in the cortical plate and ventricular zone, but not in the subventricular zone. In addition, transcripts are detected in the ventricular zone of the ganglionic eminence. At e18.5, staining is strongest in the ventricular zone and subependymal regions of the lateral ventricle and extending to the olfactory bulb, with low signal in cerebral cortex. Following birth, ME2 expression becomes progressively restricted to granule cells of the cerebellum and hippocampus (Soosaar et al., 1994).

In addition to overlapping expression patterns of E2A and ME1 in the developing central nervous system, alternative transcripts for ME1/HEB have been identified, named ME1a and ME1b (Neuman et al., 1993a). These differ by an additional 24 amino acid ankyrin repeat domain encoded in transcripts for ME1a (Chiaramello et al., 1995; Neuman et al., 1993a), though the transactivation potential of both proteins appear identical (Chiaramello et al., 1995). Importantly, the ME1a isoform is enriched in neuronal cells; quantitation of relative levels of ME1 during neurogenesis of PCC7 cells reveals an increase in the ratio of ME1a to ME1b as neurogenesis progresses, though the significance of this observation is unclear (Neuman et al., 1993a). Ostensibly, ME1a may be the predominant transcript expressed in mature neurons of the adult brain. Once again, these observations suggest multiple redundant roles in programming neurogenesis by Class I HLH factors, and may provide clues to explain the lack of obvious defects in the brains of single-mutant mice.
1.5. Defining a “co-neurogenic” role for E12 in developing cortex

In consideration of all available evidence presented, the role of E12 in corticogenesis still remains ambiguous. While expression studies suggest a role for E12 in neuroprogenitor formation (Roberts et al., 1993; Perez-Moreno et al., 2001; Rutherford and LeBrun, 1998), GOF assays employing uncommitted P19 cells demonstrate that E12 does not encode intrinsic neurogenic signalling potential (Farah et al., 2000). Furthermore, the lack of overt neurological phenotype in E2A null mutant mice suggests a redundant signalling pathway that involves this transcription factor (Bain et al, 1994; Zhuang et al., 1994). Nevertheless, the affected faculties in E2A mutant mice, including the B- and T- cell compartments, represent cellular systems involving molecular circuits which absolutely require E12 and E47 signalling.

Since biochemical analyses reveals that E12 predominantly elicits its function through heterodimerisation with Class II and VI HLH factors, then it should follow that a survey of the expression patterns of its binding partners during cortical development may delineate key neurogenetic signalling pathways that involve this transcription factor. To clarify this, an outline of expression patterns for known HLH factors in the developing cortex is presented in Figure 1.13. As shown, neurogenesis is divided into three phases: (i) neuroblast proliferation; (ii) neuronal migration and (iii) terminal differentiation. Available expression data reveals distinct waves of HLH gene expression at different stages of corticogenesis that precedes expression of neuron-specific genes such as tyrosine hydroxylase (TH) and glutamic acid decarboxylase (GAD67). For example, expression of the determination HLH factors Mash1 and ngn1/2 precedes activation of NeuroD and Math2. Additionally, expression of the novel Class II HLH factor BETA3 follows Mash1 and ngn1/2 expression, suggesting a role for proper terminal differentiation of neurons. In contrast, E12 is only detected in mitotically active cells of the ventricular zone, and its expression coincides with Ids in proliferative neuroblasts. These observations indicate that E12 may be more important for the early steps of progenitor formation and lineage determination, rather than for terminal differentiation of cortical neurons.
Neuroblast formation, Migration of Postmitotic Neurons, Terminal Differentiation

A role for E12 in early neurogenesis?

Figure 1.13. Distinct waves of HLH gene expression during cortical development. Neurogenesis is divided into three phases: (i) proliferation of neuroblasts within a germinal matrix, followed by (ii) migration of postmitotic neurons towards their final laminar destination. Finally, these cells undergo (iii) terminal differentiation. All available expression data indicates E12 may be important for the early stages of neurogenesis, and could regulate neuroblast formation as well as the early steps of neurodifferentiation, possibly through direct interaction with Class II (Mash1, Ngn1/2) and Class V (ID1-4) proteins in vivo.
The following diagram in Figure 1.14 summarises available biochemical and functional evidence for relevant pathways that involve E12 signalling during corticogenesis. Taken together, I propose that E12 functions as a “co-neurogenic” factor, facilitating neurodifferentiation programs signalled through Class II HLH factors, while signalling cells to stop dividing. As shown, E12 behaves as a multifunctional molecular chaperone during neurodevelopment (Figure 1.14). In neuroblasts, E12 may signal as a growth suppressor through transcriptional activation of CDKIs expression, a function that is regulated through phosphorylation by cyclin-dependent kinases CDK4/6 and CDK2. Concomitantly, the Class II HLH factors may heterodimerise with E12 to upregulate expression of downstream neuronal genes, such as NeuroD, in dorsal progenitors. This activity is balanced through co-stimulation of Id2 levels by E12 heterodimers which, upon expression, binds E12 to regulate the pool of available monomers for dimerisation to Class II HLH factors, thereby serving as a negative feedback loop. Similarly, the cellular pool of E12 is regulated through ubiquitin-mediated proteasome degradation.

In addition to roles in cell cycle exit and lineage commitment, E12 heterodimers may recruit histone acetyltransferases to specific loci, thereby facilitating chromatin decompaction for accessibility to accessory transcriptional machinery. Finally, E12 may be important for proper endowment of a migratory phenotype in neuroprogenitors through direct transcriptional repression of E-cadherin expression, thereby decreasing their adhesiveness to the germinal matrix. These postulated roles for E12 during cortical development must be clarified through positive identification of protein-protein heterodimerisation partners in the tissue milieu. To do this, a search for binding partners for E12 in the developing cortex was conducted using a yeast 2-hybrid interaction assay.

1.6. A functional screen for interacting partners to E12 through a yeast 2-hybrid interaction assay.

It is clear that protein dimerisation is essential for E12 function. Since the C-terminal HLH domain of E12 mediates the capacity for heterodimerisation, a function important for DNA binding and transactivation as a dimeric complex, this domain may be exploited to search for potential binding partners in the developing mouse cortex, and using a yeast 2-hybrid assay. This functional screen for protein-protein interaction, first described by Fields and Song (Fields and Song, 1989), exploits the modular nature of eukaryotic
Figure 1.14. A proposed “co-neurogenic” role for E12 in the developing cortex. E12 heterodimerises with neurogenic Class II HLH factors (CII*) to activate transcription of downstream genes (such as NeuroD) for proper terminal differentiation, as well as to regulate chromatin decompaction through recruitment of histone acetyltransferase (HAT) complexes to target DNA. Similarly, E12-expressing neuroblasts loose their adhesiveness through repression of E-cadherin, a function which may be important for endowing a migratory phenotype. Concomitantly, E12 inhibits cell cycle re-entry through upregulation of the CDKI protein, p21. Finally, E12 upregulates ID2 expression to establish a negative feedback loop for regulating dimerisation to Class II HLH factors. The available pool of E12 monomers in vivo is directly regulated by proteasome degradation (P#) through a ubiquitin-mediated pathway which involves Ubc9/UBCE2A.
transcription factors such as GAL4 in yeast. Important for galactose metabolism, the yeast transcription factor GAL4 encodes separable domains for contacting DNA (Keegan et al., 1986) and transcriptional activation (Ma and Ptashne, 1987). It is this separation of function, coupled with a capacity for transcriptional activation domains to signal as chimeric fusion proteins (Brent and Ptashne, 1985; Ma and Ptashne, 1988) that allows for the evaluation of candidate protein domains for interaction through this assay (see Figure 1.15). Firstly, plasmids are constructed that encode hybrid proteins for assessment of protein interaction. One plasmid encodes a polypeptide comprising the DNA binding domain (DBD) of GAL4 fused to test protein “X”, while a second construct produces a fusion protein composing of the transcriptional activation domain (TAD) of GAL4 fused to test protein “Y”. These plasmids are introduced into a modified strain of the yeast *Saccharomyces cerevisiae* which encodes nutritional (the histidine biosynthetic enzymes) and/or colorimetric (*LacZ* gene encoding β-galactosidase) reporter genes under the control of GAL4. Introduction of either constructs in isolation does not activate transcription of reporter genes, since fusion protein X may bind GAL4 sequences but lack transactivation potential (Figure 1.15B), while fusion protein Y does not localise to reporter genes through lack of the DBD (Figure 1.15C). However, a positive interaction between protein X and protein Y results in recruitment of the GAL4 TAD polypeptide to this site, thereby activating transcription of these reporter genes for positive selection that survive in the absence of histidine, and turn blue in the presence of the chromogenic substrate X-gal (Figure 1.15D).

Initially developed as an assay for genes in isolation, the utility of this system has been extended to permit screening of activation domain libraries for fusion proteins that interact with a target polypeptide (Chakraborty et al., 1992; Chien et al., 1991; Dalton and Treisman, 1992; Staudinger et al., 1993). High efficiency cloning procedures employing bacteriophage lambda (Dotto et al., 1984; Short et al., 1988) allow for cloning of complex pools of complementary DNA (cDNA) molecules corresponding to expressed genes, thereby generating libraries of chimeric cDNAs encoding fusion proteins with the GAL4 TAD domain. These plasmids are then co-transformed with the test protein, fused to GAL4 DBD, for rapid isolation of interacting genes from the TAD library. Essentially, this approach may be envisaged as a “molecular fishing trip” whereby the “bait protein”, encoded as a fusion protein to GAL4 DBD, specifically
Figure 1.15. Schematic of the yeast 2-hybrid interaction assay. (A) The utility of this system impinges on the modular nature of transcription factors, such as GAL4, which comprise separable domains for DNA binding (DBD) and transcriptional activation (TAD). A “bait” fusion protein containing the DBD is fused to test protein “X”, while a “prey” fusion is also prepared, and encodes the TAD fused to protein “Y”. While protein X may bind GAL4 regulatory sequences (shown in B), the chimeric protein “Y” does not localise its TAD to the reporter genes (shown in C). A positive interaction between “X” and “Y” results in activation of reporter genes in the host cell (D).
bends a “prey protein” to activate GAL4 promoter-driven reporter genes. Shown in Figure 1.15D, the bait protein “X” specifically binds the prey protein “Y”, and not to prey proteins “Z”. Most importantly, when positive interactions are identified through this screening procedure, coding sequence for the interacting prey protein is immediately available upon plasmid purification and sequencing. While there are clear disadvantages and advantages for the use of this assay to clone protein-protein interacting pairs (reviewed in Bartel et al., 1993), its utility for cloning HLH-containing polypeptides (Chakraborty et al., 1992; Gradwohl et al., 1996; Ravassard et al., 1997; Staudinger et al., 1993), as well as others (Brakeman et al., 1997; Dong et al., 1997; Kho et al., 1997; Loveys et al., 1996; Trommsdorff et al., 1998; Zukerberg et al., 2000) is well documented.

All available evidence describing a role for E12 suggests it signals early in corticogenesis, predominantly in the proliferative neuroblasts of the germinal epithelium. Therefore, it follows that the search for binding partners to E12 should be focused on genes coincidently expressed in the early stages of cortical development that interact with this transcription factor. To do this, yeast 2-hybrid prey libraries will be constructed from genes derived from embryonic brains isolated at two developmental time-points: at e11.5 and e15.5 in the mouse. These time-points for embryonic brain dissection represent two key phases in early development of the cerebral cortex: (i) a proliferative phase and (ii) a peak neurogenesis phase. At e11.5, the mouse forebrain is essentially a homogeneous sheet of mitotically active precursor cells which generate sufficient numbers of neuroblasts for proper corticogenesis. On the other hand, the developing cortex encounters a peak in neurogenesis at e15.5, when the number of departed neurons is matched by the number of remaining proliferating cells (Takahashi et al., 1996a; Takahashi et al., 1996b). Careful dissection of brain tissue from these embryonic mice should allow for construction of dedicated prey libraries from early (e11.5) and peak (e15.5) corticogenesis.

Following their construction, these prey libraries will then be systematically screened for binding partners using to an E12 bait protein. The search for binding partners to E12 will be invaluable for elucidating key neurogenic signalling pathways which involve known HLH factors, such as Mash1 and ID2. Additionally, this approach allows for cloning of possible novel binding partners to E12, thereby increasing our understanding of the contribution of this transcription factor to neurogenic signalling cascades.
Chapter 2: Materials and Methods
2.1. Animals

All animal procedures were carried out in compliance with institutional guidelines. C57BL/6 mice from an inbred colony were used for all experiments. The mid-day of the vaginal plug was considered as embryonic day 0.5 (written as “e0.5”). Adult mice were sacrificed by cervical dislocation, while embryos were decapitated before further manipulation. Forebrain tissue from e11.5 embryos was collected as follows: the meninges was carefully peeled away from the forebrain, and the telencephalic vesicles (including the underlying diencephalon) were dissected from the rest of the head and immediately plunged into denaturing solution for RNA isolation. To isolate cerebral cortical tissue from e15.5 embryonic brain, the telencephalic vesicles were first dissected, and the meninges peeled away. Next, the medial (the hippocampal primordium) and ventrolateral tissues (which gives rise to the striatum) were dissected away, and the remaining cerebral cortical tissue was harvested for RNA isolation (see Figure 2.1).

2.2. RNA isolation

All RNA isolations were carried out using either conventional methods employing guanidine thiocyanate and acid phenol (Chomczynski and Sacchi, 1987), or with TRIZOL (GIBCO-BRL Life Technologies, Inc., Gaithersburg, Md.). The integrity of RNA was determined using denaturing formaldehyde agarose gel electrophoresis; good quality RNA being assessed by visual inspection of intact 28S and 18S ribosomal RNA bands. Messenger RNA was isolated using the MESSAGEMAKER mRNA isolation system (GIBCO-BRL Life Technologies, Inc., Gaithersburg, Md.). Alternatively, mRNA was isolated using a conventional oligo-d(T) cellulose method according to the following procedure: oligo-d(T) cellulose (Catalogue number 15940-018, GIBCO-BRL Life Technologies, Gaithersburg, Md.) was pre-swollen in WB1 (Wash Buffer 1; 0.5 M NaCl, 20 mM Tris at pH 7.5) for at least 1 hour, then pelleted. All centrifugation steps were carried out on a tabletop swing-bucket centrifuge (Beckman Instruments, Palo Alto, CA) at 1500 RPM for 5 minutes at room temperature. The supernatant was then removed, and the cellulose pretreated with alkaline buffer (0.2 M NaOH, 1 mM EDTA) for 5 minutes at room temperature. Following this, the cellulose was recovered and washed four times with WB1, then finally resuspended in WB1 at a final concentration of 10mg/ml. All following procedures for mRNA isolation were carried out
Figure 2.1. **Source of embryonic brain tissue for RNA isolation.** Diagrammatic representation of a coronal section of embryonic mouse cortex at day e11.5 and e15.5. Regions shaded grey indicate sources of tissue for RNA isolation according to procedures detailed in Chapter 2.2. Cortical tissue was dissected from both telencephalic vesicles of the e15.5 mouse brain. NCx = neocortex; GE = ganglionic eminence; H = hippocampal primordium; LV = lateral ventricle.
Chapter 2: Materials and Methods

2.3. Construction of a yeast 2-hybrid cDNA prey library

Complementary DNA (cDNA) libraries were constructed using the HybriZAP-2.1 Two-Hybrid cDNA Gigapack Cloning Kit (Stratagene; catalogue number 235612) according to the manufacturer’s instructions. Detailed protocols are available (http://www.stratagene.com/manuals/235612.pdf), and a brief description of the procedure is as follows: Five µg of messenger RNA (mRNA) was reverse transcribed using a Moloney murine leukemia viral reverse transcriptase and a XhoI-tailed oligo-d(T) primer. In addition, methyl-dCTPs were included in the reaction mix to generate hemi-methylated, double stranded cDNA molecules. It is important to note that a 3’ bias is introduced through the use of an oligo-d(T) primer for reverse transcription. These cDNAs were then blunt-ended with Pfu DNA polymerase, ligated to EcoRI linkers, then digested with XhoI restriction enzyme to generate cDNAs with non-complementary ends (an EcoRI site on the 5’ end, and a XhoI site on the 3’ end). Finally, these tailed cDNA molecules were separated from residual linker-primer and other short DNA products by column chromatography using Sepharose CL-2B filtration medium. The size-fractionated cDNA is then precipitated and ligated to the lambda phage cloning vector HybriZAP-2.1. The lambda library is packaged at high efficiency using packaging extracts included in the kit, then host bacteria (XL1-Blue MRF’ strain of E. coli) are infected with viable phage to produce the primary library. At this stage, individual phage plaques can be isolated and their insert cDNA amplified by PCR (see Appendix 10.3) to determine the size distribution of cDNA inserts in individual library clones. Finally, mass in vivo excision of the phagemid vector (harbouring the insert cDNAs) is performed using helper phage and the E. coli cell strain XLOLR. Phagemids were harvested using the Midiprep plasmid isolation kit (Qiagen GmbH, Germany). The prey library plasmid DNA stock was stored at -20°C.

The number of β-actin clones in the library was assessed by colony hybridisation using standard procedures. Representative colonies (approximately 5000) from each library were transferred onto
nitrocellulose filters, then colony hybridisation performed with a radiolabelled β-actin cDNA probe. Signals were detected via exposure to autoradiographic film or a phosphor imaging plate.

2.4. Yeast cell lines

Yeast strains HF7c (Feilotter et al., 1994) and PJ69-4A (James et al., 1996) were used for all experiments. The HF7c cell line harbours HIS3 and βgal genes as reporters, while the PJ69-4A cell line harbours HIS3, βgal and ADE2 reporter genes which allow for selection in the absence of histidine as well as adenine. Prior to carrying out large-scale yeast 2-hybrid interaction screens, an assessment was made of the vigour of HF7c cells, compared to PJ69-4A. It was observed that the PJ69-4A cells grew twice as quickly as the HF7c in complete medium, and also recovered in half the time following plasmid transformation using lithium acetate (data not shown). As a result, all yeast 2-hybrid assays were performed with PJ69-4A cells. The histidine biosynthesis inhibitor 3-aminotriazole (3-AT) was titrated to a final concentration of 8mM for all selection procedures.

2.5. Yeast cell transformations

Baffled flasks were washed with boiling MilliQ water and autoclaved before use. Detergent was never used to clean any glassware for yeast manipulation. Yeast cells were transformed with bait and/or prey plasmid DNA using a standard lithium acetate method. For 2-hybrid library transformation experiments, 30 micrograms of library plasmid DNA was used according to the following protocol: an overnight suspension of bait-transformed PJ69-4A (5ml for SD-trp) was added to 100 ml of SD-trp and grown overnight. The next day, cells were recovered by centrifugation (2400 x g for 5 min at room temperature), inoculated into 200ml of YETD to give an OD₆₀₀ of 0.3 (corresponding to ~ 1 x 10⁷ cells/ml) and grown for 3-5 hours with shaking (225 rpm) at 30 °C. When the OD₆₀₀ is ~ 0.8, cells were pelleted and washed once with sterile water, then with TL (100mM LiAc, TE pH 7.5) and finally resuspended in 12 ml of PTL (40% PEG₃₃₅₀, 100 mM LiAc, TE, pH 7.5). Five Falcon (15ml) tubes were prepared for the transformation reaction, each containing 6 µg of prey library DNA, 400 µg of sonicated salmon sperm DNA and 2.4 ml of cell suspension. These tubes were incubated for 30 min with shaking (225 rpm) at 30°C, then 250 µl of dimethylsulphoxide was carefully added. Finally, the cells were heat-shocked at 42°C for 7.5 min and
Transformation reactions were then pooled and added to recovery medium (100 ml of SD-tl) for 1-24 hours on a shaking platform (225 rpm, 30°C). Finally, the cells were recovered by centrifugation, resuspended in TE (pH 7.5) and plated in solid agar medium lacking trp, leu, his (complemented with 3-AT at 8mM) and incubated at 30°C. Colonies typically appeared after 5 days, and these were then replated on medium lacking trp, leu, his and ade, and supplemented with 3-AT (SD -tlha +3-AT) and assessed for β-gal activity using conventional methods (Bartel et al., 1993).

**2.6. Recovery of plasmid DNA from yeast**

Plasmid-transformed yeast cells (1ml of confluent culture) were harvested (9600 x g for 5 min at room temperature) and resuspended in 200µl breaking buffer (100 mM NaCl, 2% Triton X-100, 1% SDS, TE, pH 7.5). Next, 150µl of phenol (pH 8.0), 150µl of chloroform and 200 mg of acid washed 425-600µm glass beads (Sigma-Aldrich Chemicals, Australia) were added and mixed. This preparation was vortexed for 2 min at 1500 rpm at room temperature using a multitube vortexer (IKA-VIBRAX GmbH, Germany), then centrifuged for 5 min at 20800 x g at room temperature. The aqueous phase was then collected, re-extracted with phenol:chloroform, then precipitated with an equal volume of isopropanol. The final DNA pellet was resuspended in 5µl of sterile water, and 1µl was used to transform bacteria.

**2.7. Plasmid DNA constructions**

Full-length cDNA for mouse E12 (AK017617) was provided by RIKEN (Yokahama Institute, Japan). The plasmid pBSIJKS-A1(A7)HLH (D29919) containing an insert encoding amino acids 491 to 651 of the mouse E12 gene was kindly provided by Dr. Yoshitaka Kajimoto (Osaka University School of Medicine, Japan). The pBD-muE12505-651 bait plasmid was cloned by ligating mouse E12 cDNA sequence (encoding amino acids 505 to 651) into pBD-Gal4-2.1 (Stratagene, La Jolla, California, USA). Bait plasmids that encoded truncations of the E12 cDNA are described in Appendix 10.5. The two-hybrid prey plasmid pAD-GRIPE includes 3768 nucleotides of the 3’ end of GRIPE cDNA, and encodes an in-frame fusion product of 785 amino acids. The plasmid pEGFP-GRIPE<sup>701-1485</sup> was constructed by ligating the BamHI fragment of GRIPE cDNA downstream of EGFP in the mammalian expression plasmid pEGFP-C1 (Clontech, Palo Alto, California, USA), producing a C-terminal EGFP-GRIPE<sup>701-1485</sup> fusion protein. Full-length E12
protein was cloned into pcDNAF, generating FLAG-E121-651, a Flag-tagged protein of 651 amino acids. All mammalian expression plasmids are described in Appendix 10.7.

2.8. Animal cell culture

The following adherent cell lines were used for all cell culture experiments: HEK293T human embryonic kidney cell line (DuBridge et al., 1987), COS-7 monkey kidney epithelium (Gluzman, 1981), and P19 mouse embryocarcinoma cell line (McBurney, 1993; McBurney and Rogers, 1982). All cell lines were maintained in DMEM medium containing 5% fetal calf serum (heat-inactivated), 2 mM L-glutamine, 0.5 U/ml penicillin and 0.5µg/ml of streptomycin. Cells were incubated at 37°C in a 5% CO₂ humidified chamber.

Frozen stocks were prepared as follows: Cells from a 25 cm² flask (90% confluent) were pooled, resuspended in 1 ml of freezing medium (20% FCS, 10% DMSO in DMEM) then transferred to cryogenic tubes. These tubes were placed in a slow-freeze isopropanol chamber and kept at -80°C overnight, then transferred to liquid N₂ storage the next day.

2.9. Mammalian cell transfection

Complete medium lacking antibiotics was used for all cell transfection experiments. Approximately 2 x 10⁵ cells were plated in each well of a 12-well plate, and were allowed to recover overnight before DNA transfection the next day. All transfection procedures employed Lipofectamine 2000 (GIBCO-BRL Life Technologies, Inc., Gaithersburg, Md.) according to the manufacturer’s instructions. Transfected cells were allowed to recover for 30 hours before further manipulation.

2.10. Coimmunoprecipitation of epitope-tagged fusion proteins.

The use of epitope-tagged fusion proteins for coimmunoprecipitation allows for an assessment of protein-protein interaction that does not require native antibodies to the target polypeptides. A diagrammatic representation of this approach is shown in Figure 2.2, with two hypothetical interacting proteins X and Y. Protein X is a fusion protein with Green Fluorescent Protein (GFP), while protein Y is a FLAG epitope-
tagged protein. Lysate containing these proteins are incubated with antibodies to GFP (Fig. 2.2A); these recognise and bind GFP fusion protein X, which is complexed with protein Y. This immune-complex is subsequently precipitated through addition of protein A sepharose beads (Sigma-Aldrich Chemicals, Australia). Since protein X and Y form a complex, protein Y is precipitated through this association (Fig. 2.2B). After repeated washes to remove unbound proteins in solution, the immunoprecipitated products are resolved on a polyacrylamide gel and blotted on a nitrocellulose membrane. Western blotting is then performed with FLAG antibodies to detect protein Y in the preparation. As a control experiment, protein Y is tested for non-specific binding to GFP protein alone.

![Figure 2.2](image)

**Figure 2.2. Outline of co-immunoprecipitation experiment.** (A) Fusion proteins X and Y are incubated with antibodies to GFP, and protein A sepharose. (B) Protein Y is immunoprecipitated with protein X (Test), while the GFP protein does not bind protein Y in a separate reaction (Control).

Coimmunoprecipitation experiments discussed herein were carried out using protein lysates from transiently transfected HEK293T cells. Following transfection, cells were harvested by scraping and washed with PBS at room temperature, then resuspended in 3ml of ice-cold lysis buffer (PBS containing PMSF, leupeptin, pepstatin). All steps were performed at 4°C. The cells were gently lysed and incubated for 10 minutes on a rotating wheel. The lysate was then cleared by centrifugation at 10 000 rpm for 30 min, and the cleared supernatant incubated with 15µl of a 50:50 slurry of protein A sepharose (Sigma-
Aldrich Chemicals, Australia) in PBS for 1 hour on a rotating wheel. Samples were then centrifuged at 2000 rpm for 1 minute and the supernatant transferred to fresh tubes containing 1µl of anti-GFP antibody (Cat no. 8372-1; Clontech, Palo Alto, CA, USA), 15µl of Protein A sepharose beads and incubated for 2 h on a rotating wheel. The beads were then washed (4X) with ice cold PBS, bound complexes harvested and separated by SDS-PAGE before electroblotting and immunodetection using the methods described below.

2.11. Western blot analysis.

Protein samples were fractionated on 12.5% SDS-PAGE gels and electroblotted onto Hybond C extra nitrocellulose membrane (Amersham plc, Buckinghamshire, UK). The blots were rinsed briefly in Tris-buffered saline containing 0.1% Tween 20 (TBST) and incubated overnight at 4°C in blocking buffer containing 5% dry milk powder in TBST. Membranes were incubated with the appropriate primary antibodies prepared in TBST (polyclonal anti-E12, 1:200 dilution, Santa Cruz Biotechnology, CA; anti-Flag M1 mouse monoclonal, 1:1000 dilution, Sigma Australia; anti-GFP peptide antibody raised in rabbit, 1:200 dilution, Clontech, Palo Alto, CA) for 1 hour at room temperature. Following washing (3 washes, 15 minutes each), membranes were incubated with secondary antibodies in TBST (1:4000 dilution; horse radish peroxidase-conjugated, Silenus Australia) for 1 h at room temperature. Following washing, protein bands were visualised by chemiluminescence using ECL Western blotting detection reagents (Amersham plc, Buckinghamshire, UK).

2.12. Immunocytochemistry.

Transfected cells were washed with cold PBS then fixed with PBS containing 2% paraformaldehyde/0.1% Triton-X for 30 min on ice. Following three washes with ice-cold PBS (5 min for each wash), a blocking solution consisting of 5% normal horse serum (in PBS) was added to the cells and incubated for 30 minutes at room temperature. Cells were then washed three times with cold PBS and incubated with a mouse anti-FLAG antibody (1:1000 dilution in PBS) for 1 h at room temperature. Following three washes in PBS containing 1mM calcium chloride, a 1:1000 dilution of donkey anti-mouse Cy3-conjugated antibody (Chemicon, Australia) was added for 1 hour at room temperature. Following three washes with PBS, cells
were briefly incubated with a solution of bisbenzimide (1mg/ml in PBS) to reveal cell nuclei, then subjected a final wash before mounting and examined under fluorescence optics.

2.13. Luciferase Assay.

Assays were performed in triplicate using transiently transfected HEK293T cells in 12-well plates. The reporter construct, p(µE5 + µE2),Luc+, encodes a luciferase reporter gene under the control of 6 copies of the E-box (µE5 + µE2) region of the IgH enhancer (Ruezinsky et al., 1991), upstream of a liver/bone/kidney alkaline phosphatase minimal promoter (a gift from Dr. Cornelis Murre). Transfections were performed as follows: each transfection preparation comprised 2.2 µg of DNA (made up with pcDNA3 vector) and 6 µl of Lipofectamine. Similar (200 ng) amounts of plasmid DNA encoding full-length FLAG-E12{1,651}, as well as p(µE5 + µE2),Luc+ reporter vector, were added to prescribed samples. Increasing amounts of EGFP-GRIPE{701-1485} or EGFP were added as prescribed. A βgal plasmid (200 ng) was also included in all samples to control for transfection efficiency. Following transfection, cells were lysed with GloLysis buffer (Promega, Madison WI) and diluted before addition of an equal volume of Bright-Glo Luciferase Assay reagent (Promega, Madison WI). Chemiluminescence was measured through use of a Topcount scintillation counter (Packard Instrument Co., Meriden, USA), while βgal activity was measured through a standard liquid ONPG assay (Ogilvy et al., 1998). Representative samples from each preparation were subjected to Western analysis to detect FLAG-E12{1,651}, as well as EGFP-GRIPE{701-1485} or EGFP proteins in the lysate. These Western blots were stained with Ponceau red to confirm equal loading. All luciferase activity measurements are represented as a ratio of βgal activity. The βgal activity units were calculated as follows: raw reading multiplied by 380 then divided by length of incubation; the luciferase measurement was then divided by this reading to generate normalised activity units. Raw data are presented in Appendix 10.14.

2.14. Northern Analysis

Northern membranes were prepared as follows: 15µg of total RNA was separated on a denaturing formaldehyde-agarose gel before capillary transfer onto Hybond N+ (or Hybond C-Super) nylon membranes using standard techniques. Following overnight transfer, membranes were baked for 2 hours at
Chapter 2: Materials and Methods

80°C in a vacuum oven. Hybridisations to radiolabelled cDNA probes were carried out using either Rapidhyb (Amersham plc, Buckinghamshire, UK) or Aqueous hybridisation solution (5 x SSC, 5 x Denhardts, 1% SDS, 100µg/ml denatured salmon sperm DNA), and incubated overnight at 65°C with rotation in bottles. The next day, the membrane was washed (2 x SSC/0.1% SDS for 15 minutes at 65°C, 0.2 x SSC/0.5% SDS at 65°C) then exposed to autoradiographic film or a phosphor-imaging plate for visualisation of signals.

Complementary DNA probes for target genes (see Appendix 10.2 for description of cDNA probes) were labelled with $^{32}$P-dCTP using a hexanucleotide mix (Roche Biochemicals GmbH, Germany) according to standard random-prime labelling procedures. Labelled probes were purified by column chromatography with sephadex G-50 resin to remove un incorporated nucleotides. Up to $2 \times 10^6$/ml of incorporated counts was used for each hybridisation experiment.

Non-saturating RNA signals detected by phosphor-imaging were quantitated using the AIS 3.0 computer package (Imaging Research Inc., Canada). All signals were normalised with 18S rRNA to account for differences in loading, and the final value multiplied by 100. A single-factor ANOVA was performed to determine statistical significance of the observations. These data are presented in Appendix 10.13.

2.15. Reverse Transcriptase - Polymerase Chain Reaction

Reverse transcription (RT) reactions were prepared as follows: two micrograms of DNase treated total RNA was heated to 65°C for 5 minutes, then chilled on ice. The following components were then added to the 50µl reaction (50 µM of oligo-d(T)$_{12}$ primer, 500 µM dNTPs in reaction buffer) and incubated at 42°C for 2 minutes. Finally, 9 units of cloned AMV reverse transcriptase (Promega, Madison WI) were added and the tube incubated for 45 min at 42°C to synthesise first strand complementary DNA. As a control for genomic DNA contamination, parallel RT reactions were performed without reverse transcriptase enzyme and processed as above. Two microlitres of each RT reaction mix was then used as template for PCR in a 25µl reaction containing 2.5 mM dNTPs, 2.5 mM MgCl$_2$, 30 µM of each primer and 1 unit of cloned Taq polymerase (Perkin Elmer Life Science, Boston, MA). Negative control PCR reactions were performed
without Taq enzyme. The amplification conditions are as follows: preheat to 94°C for 2 minutes; 35 cycles of 94°C for 0.5 min, 58°C for 1 min, 72°C for 1 min; final extension step at 72°C for 10 min then store at 4°C. The PCR products were separated on a 1% TAE agarose gel.

### 2.16. In situ hybridisation

Whole mouse embryos were embedded in O.C.T. compound (Sakura Finetek Inc., USA) and frozen in a bath of liquid nitrogen-cooled isopentane. Tissue sections 10µm thick were cut using a freezing microtome, mounted on Superfrost Plus slides (Fisher Scientific, USA) and air-dried for at least 20 minutes at room temperature. These prepared slides were either stored at -20°C or used immediately. Ultrapure injection water (Baxter Healthcare, Australia) was used for all RNA hybridisation steps, while autoclaved water was used in procedures which followed ribonuclease treatment (see below). Coverslips were siliconized by dipping in 3% silane (Sigmacote; diluted in ethanol), followed by two immersions in chloroform, then twice in 100% ethanol before allowing these to air-dry.

Tissue sections were fixed in fresh 4% paraformaldehyde (made in PBS and filter-sterilised) for 10 minutes at room temperature. These slides were then washed 3 times with PBS (3 minutes each wash), and an acetylation reaction performed by incubation in buffer (13mM triethanolamine pH 8, 17.5 mM HCl, 2.5 ml/L acetic anhydride in injection water) for 10 minutes at room temperature. Following this, the slides were then washed three times with gentle agitation in PBS (5 minutes each wash).

Each slide was pre-incubated with hybridisation solution (50% deionised formamide, 5 x SSC, 5 x Denhardt’s, 250 mg/ml MRE 600 tRNA, 500 µg/ml herring sperm DNA) in a humidified chamber for 2 hours at room temperature. Following this, the hybridisation solution was removed and fresh probe solution (hyb buffer containing 200-300 ng of DIG-labelled cRNA, pre-heated for 5 minutes at 80°C) was added to the slides. Preparation of digoxigenin-labelled cRNA probes is described in section 10.9. A siliconized coverslip was carefully placed on top, and these sections were incubated overnight at 72°C. The next day, coverslips were removed by submerging in 5 x SSC at room temperature, and slides were subjected to two washes in 0.2 x SSC at 65°C for 1 hour. Following this, the slides were incubated in
preheated (37°C) RNAse buffer (0.5M NaCl, 10mM Tris pH 7.5, 5mM EDTA) for 5 minutes, followed by RNAse digestion buffer containing 10 µg/ml RNase (Roche Biochemicals, Gmbh, Germany) for 30 minutes at 37°C. The slides were immersed in 0.2 x SSC for 30 minutes at room temperature, then briefly washed in B1 buffer (0.1M Tris pH 7.5). Next, the sections were incubated with B1 buffer containing 10% heat-inactivated goat serum (GIBCO-BRL Life Technologies, Inc., Gaithersburg, Md.) at room temperature for 1 hour, then the buffer was replaced with antibody solution containing anti-DIG-AP Fabs (Roche Biochemicals Gmbh, Germany) and 10% normal goat serum, and incubated overnight at 4°C. On the final day, the sections were washed three times with buffer (0.1 M Tris pH 9.5, 0.1 M NaCl, 50 mM MgCl₂) at room temperature (5 minutes each wash), then incubated with buffer containing colour substrate (0.338 mg/ml NBT, 0.175 mg/ml BCIP, 0.24 mg/ml levamisole). Sections were finally rinsed in PBS before mounting with Mowiol.

2.17. Differentiation of P19 embryocarcinoma cells with all-trans retinoic acid

Differentiation of cells was performed as described by Rudnicki and McBurney (Rudnicki and McBurney, 1987). Briefly, 10⁵ cells/ml were added to bacteriological plates and cultured in media in the presence of 0.6 x 10⁻⁶ M all-trans retinoic acid (RA) (Cat. No. R2625; Sigma Chemicals, Australia) for 2 days. Cell aggregates were then harvested and allowed to sediment in a 15 ml Falcon tube before the media was aspirated, and cells replated in fresh medium containing RA for another 2 days. Following this, aggregates were pooled and cells washed once with Versene (PBS containing 0.5mM EDTA) before digestion with Trypsin (0.0025% in Versene) to generate a single-cell suspension. These cells were then plated on tissue culture grade dishes in complete media without RA (all subsequent steps involve complete media without RA). After 24 hours, cytosine arabinoside (Sigma-Aldrich Chemicals, Australia) was added to the cultures at a final concentration of 5µg/ml, and the media containing ara-c was refreshed every 48 hours. Cells were harvested at 0, 2, 4, 6 and 8 days in culture for RNA isolation with TRIZOL (Section 2.2) and immunocytochemistry (Section 2.12).
Chapter 3:
Performing a yeast 2-hybrid interaction assay
3.1. Construction of yeast 2-hybrid cDNA libraries from e11.5 forebrain and e15.5 cortex tissue

The cDNA libraries prepared from e11.5 forebrain and e15.5 cortex tissue isolates were constructed according to procedures detailed in section 2.3. Figure 3.1A shows the products of a cDNA synthesis reaction using 5 µg of mRNA from e15.5 cortex as template. As shown in Lane 2, the majority of double stranded cDNAs are between 4 kb and 0.5kb in size. Following XhoI digestion, the cDNAs were size fractionated and the high molecular weight species pooled for subsequent cloning. As shown in Fig. 3.1B, these high molecular weight cDNAs (fractions F1 to F4) were collected for cloning to the lambda vector HybriZAP-2.1. Following their construction, these libraries were evaluated as represented in Table 3.1.

<table>
<thead>
<tr>
<th></th>
<th>e11.5 forebrain cDNA library</th>
<th>e15.5 cortex cDNA library</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primary library complexity</td>
<td>$4.2 \times 10^6$</td>
<td>$2 \times 10^6$</td>
</tr>
<tr>
<td>2. Proportion of phage which harbour cDNA inserts</td>
<td>&gt;85% (12/14)</td>
<td>&gt;90% (10/11)</td>
</tr>
<tr>
<td>3. Distribution of insert sizes</td>
<td>0.5 kb – 3.2 kb</td>
<td>0.5 kb – 3.0 kb</td>
</tr>
<tr>
<td>4. Representation of the cDNA library</td>
<td>N.D.*</td>
<td>&lt;1% colonies harbour $\beta$-actin cDNA (22/5000)</td>
</tr>
</tbody>
</table>

* N.D. represents “not done”

**Table 3.1. Construction of the yeast 2-hybrid prey libraries**

A total of $4.2 \times 10^6$ and $2 \times 10^6$ independent phage clones were amplified in the e11.5 and e15.5 embryonic brain libraries respectively. To assess the size of cloned cDNAs in these libraries, individual phage clones were collected from each pool, and their insert cDNAs amplified by PCR (see 10.3 in appendix for PCR protocol). Figure 3.2A shows PCR-amplified insert cDNAs from 14 individual phage clones isolated from the e11.5 forebrain library. While an amplification product of 237 bp indicates a clone which lacks insert cDNA (not shown), twelve out of fourteen clones generated PCR products with a size range of 0.8 kb to 3.5 kb, indicating that these contain cDNA inserts ranging in size from 0.56 kb to 3.25 kb. No bands were produced from clones 8 and 11 indicating failed PCR amplification of these samples; hence it could not be determined if these clones harboured an insert cDNA. Nevertheless, more than 85% of clones in this library contained insert cDNAs. A similar experiment performed with 11 colonies isolated from the e15.5
**Figure 3.1. Preparation of cDNAs for prey library construction.** (A) Synthesis of complementary DNAs (cDNAs) from total RNA (isolated from e15.5 mouse cerebral cortex) using a Stratagene cDNA synthesis kit according to the instructions of the manufacturer. Radioactive dCTPs were incorporated into the cDNA strands for autoradiographic visualisation of first strand (lane 1) and second strand (lane 2) synthesis, respectively. The majority of cDNA species are approximately between 4kb and 0.5kb in size. (B) Size fractionation of XhoI-digested cDNAs using column chromatography. Lane “M” represents λDNA marker. Lanes F0-F12 represent 100µl fractions collected. Fractions F1 to F4 were pooled and used for further cloning experiments. Wells F0 to F2 were torn from the gel, but DNA quantitation with ethidium bromide revealed significant DNA eluted in these fractions. Similar analyses were carried out with cDNAs prepared from e11.5 forebrain mRNA.

**Figure 3.2. Evaluation of insert cDNAs by PCR.** Individual phage colonies were harvested and PCR performed as described in Appendix 10.3. A; The e11.5 forebrain prey library has a primary complexity of $4 \times 10^6$, and 10/14 (>85%) contain inserts. B; The e15.5 cortex prey library has a primary complexity of $2 \times 10^6$, and 9/10 (>90%) contain inserts. A positive control, “+”, confirms utility of this protocol for cDNA amplification, while DNA contamination is ruled out through a separate reaction which lacks template, denoted “-”.

47i
cortex library showed that 10 out of 11 (>90%) colonies contained an insert, with sizes ranging between 0.5 and 3kb (Figure 3.2B).

The quality of the e15.5 cortex cDNA library was then assessed using colony hybridisation with a radiolabelled probe for β-actin (see section 2.3 of Materials and Methods). This approach allows an assessment of the extent of amplification bias during library construction; an abnormal representation of high abundance genes such as β-actin would suggest that conditions may have favoured phage colonies which grew more quickly during primary library amplification. Using conventional library construction approaches, other researchers have observed that approximately 1% of library clones harbour β-actin cDNA (Klickstein, 1993). As shown in Fig. 3.3, less than 1% (22 out of 5000) of the clones in the library harbour β-actin cDNA, indicating that the representation of the library is comparable to those generated in other laboratories (Klickstein, 1993). This experiment was not carried out with the e11.5 forebrain library. Together, these libraries represent invaluable resources for identification of protein binding partners to bait constructs of choice.

3.2 Performing the yeast 2-hybrid interaction screen

3.2.1 Yeast 2-hybrid interaction screen for binding partners to human E12

Following the successful construction of yeast 2-hybrid prey libraries, large-scale transformation procedures were carried out to screen for interacting partners to a human E12 bait (Staudinger et al., 1993). The yeast expression plasmid pAS1-E12 encodes a chimeric fusion protein of the DNA binding domain of yeast GAL4 with the C-terminal bHLH domain of human E12. The cDNA for E12 encodes the last 147 amino acids of the C-terminus, and a diagrammatic representation of the protein domains is featured in Fig. 3.4. As shown, the HLH domain is responsible for protein dimerisation (Voronova and Baltimore, 1990), and is comprised of two helices (Helix 1 and 2), separated by a loop region. In addition, an inhibitory domain prevents homodimerisation of E12, as well as binding to its alternative splice form, E47 (Sun et al., 1991). In a yeast 2-hybrid assay, this inhibitory domain should preclude the cloning of E12 and E47 in subsequent library screens, thereby reducing the recovery of these binding partners as prey fusion proteins.
Figure 3.3. Estimation of the number of β-actin clones (indicated by arrow) in the cDNA library. Colony hybridisation reveals that 22 out of 5000 clones (<1%) harbour β-actin cDNA. The filter is defined by the dashed circle.

Figure 3.4. Sequence alignment of protein fragments of human (huE12, 147 aa) and mouse E12 (muE12, 146 aa) used to construct yeast 2-hybrid baits. An inhibitory region adjacent to the basic amino acid region prevents homodimerisation (Sun and Baltimore, 1991). The stretch of basic amino acids is crucial for contacting DNA, while the helix-loop-helix region dictates protein-protein interaction. There are 16 mismatches, and huE12 has one more amino acid than muE12. The raw amino acid identity within this region is 89.11%.
This bait plasmid was introduced into the yeast cell line PJ69-4A and maintained by plating transformed cells on media lacking tryptophan (the selection marker for the bait plasmid). These bait-transformed cells were then expanded in culture, transformed with prey library plasmid and finally plated on selection media to survey for reporter gene activation by interacting bait and prey proteins (described in detail in Section 2.5). Table 3.2 details the results of these screens:

<table>
<thead>
<tr>
<th></th>
<th>e11.5 forebrain library</th>
<th>e15.5 cortex library</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of independent clones 1 hour after recovery$</td>
<td>2.2 x 10^6</td>
<td>2.9 x 10^6</td>
</tr>
<tr>
<td>2. Coverage of cDNA library$</td>
<td>0.523 fold</td>
<td>1.45 fold</td>
</tr>
<tr>
<td>3. Probability that any clone in the primary library is not assessed for interaction with bait*</td>
<td>p = 1</td>
<td>p &gt; 0.69</td>
</tr>
<tr>
<td>4. Total number of colonies screened#</td>
<td>3.2 x 10^6 (5 hours recovery)</td>
<td>2 x 10^9 (overnight recovery)</td>
</tr>
</tbody>
</table>

$calculated as the number of clones recovered after 1 hour divided by the representation of the prey library.

*calculated as the representation of the prey library divided by the number of clones recovered after 1 hour.

#calculated as the number of colonies which grew on media lacking leucine and tryptophan plated at the end of the recovery phase (5-20 hours)

**Table 3.2. Performing yeast 2-hybrid interaction screens for binding partners to human E12.**

When performing a library screen, sufficient numbers of independent colonies must be screened to ensure that all cDNAs in the library are assessed for interaction with the bait protein. Ideally, a comprehensive screen would assess 10 times as many clones as was represented in the library. For example, at least 10⁷ independent clones should be surveyed from a library with a primary complexity of 10⁶, hence the theoretical probability (p) that any given cDNA from the pool is not screened is p = 0.1. As shown in Table 3.2, coverage of the e15.5 cortex library was only 1.45 fold, suggesting that there is a probability of p > 0.69 that any given primary clone in the library may not have been assessed for interaction with the bait protein. Similarly, the results of screening the e11.5 forebrain library show, almost certainly, that not all of the individual cDNAs were assessed for interaction with E12. In an attempt to address this, bait- and prey-plasmid transformed cells were allowed to recover for 5-20 hours before plating on selection media. This approach allows for a clonal expansion of double-transformed cells (compare row 1 with row 4), thereby increasing the penetrance of potential interacting pairs in the pool, since transfected cells may survive transformation but benefit from additional recovery time before plating on selection media. Interacting prey cDNAs were finally isolated and re-tested for their interaction with the bait. Those preys which activated
yeast reporter genes in the absence of the bait are deemed false positives and discarded, while the putative binding partners cloned from both libraries were pooled and described in Table 3.3.

<table>
<thead>
<tr>
<th>Prey cDNAs isolated*</th>
<th>Genbank Acc. No.</th>
<th>Description</th>
<th>Known binding partner?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATH2</td>
<td>D44480</td>
<td>bHLH protein, a mammalian helix-loop-helix factor structurally related to the product of the Drosophila proneural gene atonal</td>
<td>Yes</td>
</tr>
<tr>
<td>ID2</td>
<td>M69293</td>
<td>HLH protein, negative regulator of dimerisation, isoform 2</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3</td>
<td>M60523</td>
<td>HLH protein, negative regulator of dimerisation, isoform 3</td>
<td>Yes</td>
</tr>
<tr>
<td>NFI/x</td>
<td>U57636</td>
<td>Mus musculus DNA binding protein NFI-X (NFI/x)</td>
<td>Novel</td>
</tr>
<tr>
<td>GRIP1</td>
<td>U88572</td>
<td>Rattus norvegicus AMPA receptor interacting protein GRIP, isoform 1</td>
<td>Novel</td>
</tr>
<tr>
<td>PBP</td>
<td>U43206</td>
<td>Mus musculus phosphatidylethanolamine binding protein</td>
<td>Novel</td>
</tr>
<tr>
<td>HGCP</td>
<td>U61837</td>
<td>Homo sapiens putative cyclin G1 interacting protein</td>
<td>Novel</td>
</tr>
<tr>
<td>UBC9</td>
<td>X99739</td>
<td>Ubiquitin conjugating enzyme for E2A proteins</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*describes only prey cDNAs that were: (i) fused in-frame to the transactivation domain of GAL4; and (ii) could activate the reporter genes only in the presence of huE12 bait, and not on its own, nor with a spurious binding partner.

Table 3.3. Description of prey cDNAs isolated from e11.5 forebrain and e15.5 cerebral cortical libraries as binding partners to human E12.

Only prey cDNAs that were fused in-frame to the transactivation domain of GAL4 were considered as putative binding partners of E12. In addition, clones were screened for their ability to activate the reporter genes only in the presence of the human $E12$ bait, and not via autoactivation of reporter genes, nor with a spurious binding partner such as p53 (data not shown). Importantly, due to the low coverage of prey libraries screened with the human E12 bait (refer to Table3.2), isolated cDNAs from both libraries are pooled and discussed together, since it cannot be ruled out that any given prey cDNA isolated from one library may not have been cloned from the other by chance.

Yeast 2-hybrid screens of both prey libraries resulted in the cloning of several HLH genes, namely Math2, Id2 and Id3. A known binding partner of E12, the HLH gene Math2 is a proneural gene which was originally cloned as a mammalian homologue of the Drosophila transcription factor Atonal (Shimizu et al., 1995). The inhibitory HLH genes Id2 and Id3 are negative regulators of HLH gene cascades, and are known interacting partners to E12 (Christy et al., 1991). These cloned prey cDNAs encoded their entire
HLH domains (see Appendix 10.4). No prey cDNAs encoding E12 or E47 were cloned using this bait construct, thus validating the inclusion of the inhibitory domain in the bait construct which prevents homodimerisation.

In addition to the HLH genes cloned, several novel binding partners were also isolated. These non-HLH binding partners included a transcription factor (NFI/x), a PDZ protein (GRIP), a putative cyclin-associated protein (HGCP), a phosphatidylethanolamine binding protein (PBP) and a ubiquitin ligase (UBC9). With these in hand, it is noteworthy to consider that interactions between these mouse proteins and human E12 bait may represent non-physiological phenomena. That is, the identification of novel binding partners (such as NFI/x and PBP) to human E12 may not imply bona fide association with mouse E12. To address this, a mouse E12 bait was constructed to test for interaction with these prey constructs.

3.2.2. Assessing interaction of mouse E12 bait to mouse cDNAs isolated using human E12 bait.

A mouse E12 cDNA (GenBank Number D29919) encoding 167 amino acids of the carboxy terminus was obtained for the purposes of generating a bait construct (see Section 2.7 for details of cloning). Of this, only 146 amino acids of the carboxy terminus was used to produce a mouse E12\textsuperscript{505-651} bait so as to minimise differences to the human E12 bait used in library screens described above. An alignment of human and mouse E12 fragments is shown in Figure 3.4.

As shown, both proteins share >89% (130/147) identity in amino acid sequence in this defined region. Of the 16 mismatches and 1 deletion event observed, 7 of these represented conservative changes; that is, a difference in amino acid identity, but both confer analogous charge and solubility profiles. For example, position 9 shows that both glutamate (amino acid denoted as “E”) and aspartate (denoted as “D”) comprise acidic side chains, and are negatively charged at physiological pH. In addition, mouse E12 protein contains one less amino acid when compared with human E12. Taken together, these dissimilarities could manifest as differences in binding preferences between mouse and human E12 baits. To test this, the mouse E12 bait was assessed for interaction with preys isolated as binding partners to human E12. These results are shown in Table 3.4.
### Table 3.4. Assessing interaction of mouse E12 bait to interacting partners cloned with human E12.

Positive interactions are indicated “+”, while lack of reporter gene activation is denoted “-“.

<table>
<thead>
<tr>
<th>Bait</th>
<th>Prey</th>
<th>Growth on media –t,l,h,a, +3-AT?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>muE12</td>
<td>pSV40</td>
</tr>
<tr>
<td>2</td>
<td>muE12</td>
<td>MATH2</td>
</tr>
<tr>
<td>3</td>
<td>muE12</td>
<td>ID2</td>
</tr>
<tr>
<td>4</td>
<td>muE12</td>
<td>ID3</td>
</tr>
<tr>
<td>5</td>
<td>muE12</td>
<td>GRIP1</td>
</tr>
<tr>
<td>6</td>
<td>muE12</td>
<td>PBP</td>
</tr>
<tr>
<td>7</td>
<td>muE12</td>
<td>HGCP</td>
</tr>
<tr>
<td>8</td>
<td>muE12</td>
<td>UBC9</td>
</tr>
<tr>
<td>9</td>
<td>muE12</td>
<td>NFI/x</td>
</tr>
</tbody>
</table>

The mouse E12 bait construct was first evaluated for use in a yeast 2-hybrid assay. As shown in row 1 of Table 3.4, this bait does not activate the yeast reporter genes with a spurious binding partner such as SV40. However, known interacting partners to E12, such as Math2 (Shimizu et al., 1995), Id2 (Sun and Baltimore, 1991), Id3 (Christy et al., 1991) and Ubc9 (Kho et al., 1997; Loveys et al., 1997) could activate the reporter genes when co-transformed with the mouse E12 bait (rows 2, 3, 4 and 8 respectively). Conversely, while the preys GRIP, PBP, HGCP and NFI/x could bind human E12, these did not bind to the mouse E12 bait protein (rows 5, 6, 7 and 9), hence these were concluded to be spurious binding partners. Following from these observations, the mouse E12 bait was used for all subsequent yeast 2-hybrid interaction screens.

### 3.2.3. Yeast 2-hybrid interaction screen for binding partners to mouse E12

Upon discovery that human E12 and mouse E12 baits are not equivalent in their binding preferences, the mouse E12 bait was then used to screen for biologically relevant interacting partners in both e11.5 and e15.5 embryonic brain libraries. These results are shown in Table 3.5.
Chapter 3: Performing a yeast 2-hybrid interaction assay

Table 3.5. Performing yeast 2-hybrid interaction screens for binding partners to mouse E12 bait

<table>
<thead>
<tr>
<th>Prey cDNAs isolated</th>
<th>Genbank Acc. No.</th>
<th>Description</th>
<th>Known Binding partner?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASH1</td>
<td>X53725</td>
<td>bHLH protein, mammalian achaete-scute homologue</td>
<td>Yes</td>
</tr>
<tr>
<td>NSCL1</td>
<td>M82874</td>
<td>bHLH protein, neuronal stem cell ligand-1</td>
<td>Yes</td>
</tr>
<tr>
<td>ID2</td>
<td>M69293</td>
<td>HLH protein, negative regulator of dimerisation, isoform 2</td>
<td>Yes</td>
</tr>
<tr>
<td>ID3</td>
<td>M60523</td>
<td>HLH protein, negative regulator of dimerisation, isoform 3</td>
<td>Yes</td>
</tr>
<tr>
<td>UBC9</td>
<td>X99739</td>
<td>ubiquitin conjugating enzyme for E2A proteins</td>
<td>Yes</td>
</tr>
<tr>
<td>clone 26</td>
<td></td>
<td>Novel, Contains a putative GAP (GTPase Activating Protein) Domain</td>
<td>Novel</td>
</tr>
</tbody>
</table>

The low coverage of the e11.5 forebrain library (row 2) was partly compensated by an extended recovery phase before transformed cells were plated on selection media (row 4). This was not done for the screen with the e15.5 cortex library. The unique interacting preys isolated from screening both libraries are presented in Table 3.6. Once again, prey cDNAs isolated from both libraries were pooled and presented together, since it cannot be ruled out that any given interacting prey cDNA may not have been cloned from either library by chance.

Table 3.6. Summary of prey cDNAs isolated as interacting partners mouse E12.

- calculated as the number of transformant colonies which grew on media lacking leucine and tryptophan.
- calculated as the number of clones recovered after 1 hour divided by the representation of the prey library.
- calculated as the representation of the prey library divided by the number of clones recovered after 1 hour.
- calculated as the number of colonies which grew on media lacking leucine and tryptophan plated at the end of the recovery phase (5-20 hours)
Of the six unique interacting prey cDNAs isolated as interacting partners to mouse E12, three of these, namely ID2, ID3 and UBC9, had been independently cloned as binding partners to human E12. This may suggest that the human and mouse E12 bait proteins used in these experiments constitute a minimal consensus element which recognises and binds these three prey proteins. In addition to these, the neurogenic bHLH genes MASH1 and NSCL were isolated. These have previously been shown to heterodimerise with E12 and bind specific E-box promoters using electrophoretic mobility-shift assays (Gradwohl et al., 1996). Once again, all cloned prey cDNAs for HLH genes encoded their entire HLH domains (see Appendix 10.4).

Finally, a novel interacting partner was also isolated in the library screens. Named “Clone 26”, this prey plasmid contains a 3.8kb cDNA which encodes an in-frame fusion protein of 785 amino acids. Database homology searches using BlastP (Pearson, 1995) indicate that the protein contained a putative GAP (GTPase Activating Domain), and may be a novel member of the family of GAP proteins. A comprehensive discussion of this putative novel interacting partner to E12 will be presented later.

3.2.4. Deletion analysis of mouse E12 baits with cloned interacting partners

The interaction of mouse E12 with its cloned binding partners was investigated further, using deletion constructs of the mouse E12^{505-651} bait prepared according to detailed instructions in section 2.7. The bait muE12^{505-600} lacks a critical C-terminal domain which has been previously reported as an indispensable element for E12 heterodimerisation (Goldfarb et al., 1998), while muE12^{535-651} lacks the N-terminal amino acids adjacent to the bHLH domain. The third mutant bait encodes amino acids 505-571, and lacks helix 2 of the bHLH domain, as well as most of the adjacent loop region. These bait constructs were then assessed for interaction with MASH1, ID2, and UBC9, and the results presented in Figure 3.5. Deletion analysis was not performed with Math2, NSCL and Id3 preys, though these cDNAs encode HLH domains, hence are implied HLH-dimerisation partners to E12.

While all preys could interact with E12^{505-651} (row 1), no interaction could be detected with E12^{505-600} (row 2). These observations are in agreement with work by Goldfarb and coworkers who demonstrated the
<table>
<thead>
<tr>
<th>Bait</th>
<th>Mash1</th>
<th>Id2</th>
<th>Ubc9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) muE12^{505-651}</td>
<td>++</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2) muE12^{505-600}</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3) muE12^{534-651}</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4) muE12^{506-571}</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 3.5. E12 requires an intact HLH domain as well as the C terminus for binding Mash1, Id2 and Ubc9. Deletion mutants of E12 were prepared as described (section 2.7). While E12^{505-651} interacts with all three preys (row 1), deletion of the C-terminus in E12^{505-600} abolishes binding to these (row 2). An N-terminal truncation in E12^{534-651} does not bind Ubc9, but still interacts with Mash1 and Id2 (row 3). Disruption of the loop and helix 2 regions, as well as the adjacent C terminus in E12^{506-571} abolishes binding to all preys (row 4).

Figure 3.6. The human and mouse E12 baits demonstrate similar but non-identical binding specificities. Of the 11 prey cDNAs cloned, only three genes (ID2, ID3 and UBC9) were identified as binding partners to both human and mouse baits. The cloning of non-HLH binding partners to human E12, such as GRIP and NFI/x, represented non-biologically relevant interaction events, since mouse E12 could not bind these.
importance of the C-terminal domain for dimerisation with other bHLH partners such as ID2 and ID3 (Goldfarb et al., 1998). Conversely, the results in Row 3 indicate that the 29 amino acids adjacent to the amino terminus of the bHLH region in E12 is not required for binding to all interacting partners except UBC9. This suggests that E12$_{505-651}$ comprises the minimal C-terminal domain for binding to UBC9, and is in agreement with experiments conducted by Loveys and coworkers whom demonstrated similar observations (Loveys et al., 1997). Finally, all 3 preys do not interact with E12$_{505-571}$, implying that their interaction requires an intact helix 2 and loop region, hence a functional HLH domain, on E12 for binding. Taken together, these experiments demonstrate a requirement for the HLH domain, and adjacent C-terminus, of E12 for heterodimerisation.

3.4. Summary

3.4.1 The bHLH domain of mouse and human E12 interacts with different proteins

The role of E12 during neurogenesis is poorly understood. One approach to uncovering its function in brain development is to define its binding partners, thereby allowing for the elucidation of intracellular signalling pathways that involve E12. The yeast 2-hybrid interaction screen surveys for E12 binding partners through an _ex vivo_ approach; cDNAs encoding E12 and a potential partner are introduced into yeast cells, and the expressed protein products evaluated for binding through activation of yeast reporter genes. While this powerful technique permits detection of interacting protein pairs, the choice of bait protein is crucial to the outcome of the assay; in this case, an orthologous human E12 bait protein may not be a suitable surrogate to clone binding partners to mouse E12. Certainly, these orthologs may be traced to a common ancestral progenitor, but their gene products may assume analogous but non-equivalent roles. Evidence presented here demonstrates, in the context of a yeast 2-hybrid interaction screen, that the C-terminal HLH domains of human and mouse E12 are not equivalent in their binding specificities.

Due to the low coverage of library screens performed with both human and mouse E12 baits, it was not possible to compile separate lists of prey cDNAs isolated from each library. As a result, interacting preys for each bait construct were pooled and discussed together, since it could not be determined if any given prey cDNA may not have been cloned from either library by chance. While future screens may be
Chapter 3: Performing a yeast 2-hybrid interaction assay

conducted whereby each library is screened 10-fold (for example, at least $4.2 \times 10^7$ clones be screened from the e11.5 library, with a primary complexity of $4.2 \times 10^6$), the benefit of this exercise here was judged to be minimal for the extra work required.

An evaluation of binding partners to near identical human and mouse E12 baits revealed partial complementarity in the prey cDNAs cloned. Shown in Fig. 3.6., prey cDNAs encoding Math2, GRIP and NFI/x, PBP, and HGCP were only cloned with the human bait, while MASH1, NSCL, as well as the novel cDNA “clone 26”, were only cloned using mouse E12. Conversely, three prey cDNAs, namely ID2, ID3 and UBC9, were cloned as binding partners to both human and mouse E12 baits. Taken together, these observations indicate that seemingly minor differences in secondary structure between mouse and human E12 baits may significantly affect their binding preferences in a yeast 2-hybrid assay. In particular, MASH1 and NSCL may only interact with mouse E12, since these genes were not isolated as binding partners to human E12. Furthermore, ID2, ID3 and UBC9 may be refractory to dissimilarities between baits, hence are cloned by both.

Alternative explanations for these observations abound. Firstly, the low coverage of library screens with the human E12 bait may have precluded the isolation of MASH1 and NSCL by chance. This, however, is unlikely since transformed yeast cells were allowed an extended recovery phase, and multiple overlapping prey clones of each of these HLH genes were isolated from their respective library screens (data not shown). As a corollary, ID2, ID3 and UBC9 were isolated from both human and mouse E12 library screens, suggesting rather that subtle differences between human and mouse E12 bait proteins may account for the different subsets of cloned binding partners.

In view of results presented here, it is worthwhile to consider ancillary factors which could affect the outcome of future screens conducted with the aforementioned libraries. One consideration pertains to the strategy for prey library construction: genes encoding E12 binding proteins may harbour long 3’ UTRs and binding domains in their amino-termini, hence the probability of isolating 5’ sequences may be detrimentally affected by the 3’ bias imposed through the use of an oligo-d(T) for generating cDNAs.
Furthermore, complications with the yeast 2-hybrid approach may confound results as well, since prey fusion proteins may not be properly expressed in yeast, or prey plasmid-transformed yeast cells may not be viable. Alternatively, the mouse E12\textsuperscript{505-651} bait may be incompatible for binding to cognate protein targets, as additional N-terminal sequences not present on mouse E12\textsuperscript{505-651} may be required for interaction to the entire gamut of binding partners in the tissue milieu. Nevertheless, the cloning of interacting partners to E12 validates this approach, and further work must be carried out to deduce the biological relevance of their combinatorial roles with E12 during neurogenesis.

3.4.2 An important role for protein domains adjacent to the bHLH domain of E12

Fine-mapping of regions adjacent to the bHLH domain allowed for a definition of the minimal element which governed E12 heterodimerisation to cloned binding partners. It has been reported that the C-terminus of E12 comprises a highly conserved domain, and is observed in orthologous E proteins of flies and mice (Goldfarb et al., 1998). Results presented here demonstrate that deletion of this region abolishes binding to Mash1, ID2 and UBC9; an observation that corroborates evidence presented by Goldfarb and coworkers which include this region as an essential requirement for E12 heterodimerisation (Goldfarb et al., 1998).

Deletion analysis of the mouse E12/UBC9 interaction also provided important validation of observations by Loveys and colleagues (Loveys et al., 1997). While the authors had reported that the C terminus of E12 was indispensable for interaction to UBC9, Kho and colleagues (Kho et al., 1997) had demonstrated opposing results. Data presented in Figure 3.5 are in agreement with Loveys and coworkers, and provides independent confirmation that the C-terminus of E12 is an important determinant of dimerisation to UBC9. Taken together, these experiments further define the principal elements indispensable to E12 heterodimerisation. These considerations are important for the anticipation of phenotypes that result from genetic lesions which target \textit{E12}.
Chapter 4:
GRIPE is a novel binding partner to E12, and may regulate its dimerisation state in the nucleus
4.1. **Clone 26, hereby named GRIPE, is a novel interacting partner to E12 with a putative GAP domain.**

Using a yeast 2-hybrid interaction screen for binding partners to the C-terminal bHLH domain of mouse E12, a novel 3.8kb prey cDNA was cloned. Arbitrarily named “clone 26”, this cDNA encodes an in-frame fusion protein of 785 amino acids to the transactivation domain of GAL4. In addition, it encodes a 1.4kb non-coding 3’ untranslated region, including a polyadenylation tract. As shown in Table 4.1, this prey interacts with E12 to activate the reporter genes (row 1), but does not activate the yeast reporter genes in the absence of a bait protein (row 2). Furthermore, GRIPE does not interact with a spurious bait, p53 (row 3), and the E12 bait does not activate reporter genes in the presence of a spurious prey, SV40 (row 4).

Interrogation of the Conserved Domain Database (CDD; Marchler-Bauer et al., 2002) with the protein sequence for Clone 26 revealed a putative Rap/Ran-GAP (GTPase Activating Protein) domain encoded in its C-terminus (Fig 4.1). While its true GTPase activating potential remains to be tested, amino acid sequences for putative arginine fingers of this domain are in good correspondence with those observed in other known GAP proteins such as RapGAP (P47736) and mouse Tuberin (AAA86901). These arginine fingers are important features for the proper orientation and binding of small G-proteins to facilitate their hydrolysis of bound GTP (Scheffzek et al., 1998). In light of these observations, the cDNA for Clone 26 was named GRIPE, a GAP Related Interacting Partner to E12. Subsequent cloning experiments revealed that GRIPE is a 1485 amino acid protein encoded by at least 32 exons (see Section 6.1), hence Clone 26 will be renamed GRIPE\(^\text{701-1485}\). Results of experiments that led to the cloning of full-length GRIPE cDNA are discussed more completely in Section 6.1.

4.2. **GRIPE binds to the HLH domain of E12**

The association of GRIPE with E12 was further investigated using the yeast 2-hybrid assay. Deletion constructs of \(\text{E12}^{505-651}\) were generated to map its interacting domain with GRIPE (see Appendix 10.5 for details). As shown in Table 4.2, while \(\text{E12}^{505-651}\) binds GRIPE\(^{701-1485}\) (row 1), the C-terminal domain of E12 adjacent to the helix-loop-helix region is dispensable for interaction (row 2). More significantly, disruption of the second helix of the HLH domain in \(\text{E12}^{505-571}\) prevents their interaction (row 3), and
**Table 4.1.** GRIPE is a novel binding partner to E12. GRIPE is a novel 2.6kb prey cDNA which encodes a 785 amino acid protein that interacts with E12 (row 1), and does not autoactivate yeast reporter genes in the absence of a bait (row 2). Further, GRIPE does not spuriously activate the yeast reporter genes in the presence of a spurious bait, p53 (row 3). Similarly, the mouse E12501-651 bait construct does not activate the reporter genes in the presence of a spurious binding partner, SV40 (row 4).

<table>
<thead>
<tr>
<th>Bait</th>
<th>Prey</th>
<th>Growth on media (–tlha +3AT)</th>
<th>β-gal activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 muE12⁵⁰¹-⁶⁵¹</td>
<td>GRIPE</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>GRIPE</td>
<td>-</td>
</tr>
<tr>
<td>3 p53</td>
<td>GRIPE</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4 muE12⁵⁰¹-⁶⁵¹</td>
<td>SV40</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**GRIPE may be a novel GAP protein.** Alignment of GRIPE with a consensus sequence for a Rap/Ran-GAP (Conserved Domain pfam02145). Residues in red indicate sequence correspondence, while boxed arginine residues (R1368 and R1446 of GRIPE) could be important for the GAP activity of GRIPE.
demonstrates the importance of an intact helix 2 and loop structure for binding. This region of E12 also governs heterodimerisation to other HLH proteins such as Mash1 and Id2, demonstrated through evidence presented elsewhere (see Fig 3.7). In view of these findings, it is likely that GRIPE may compete or disrupt dimer formation of E12 to other HLH proteins in vivo.

4.3. Deletion analysis of GRIPE suggests that the entire 785 amino acid polypeptide, including the GAP domain, is essential for binding to the bHLH domain of E12.

Deletion constructs of GRIPE cDNA were generated to define the minimal domain required for binding to E12 (see Appendix 10.6). To do this, the entire 785 amino acid sequence was first subjected to homology searches to uncover possible conserved domains which could govern E12-binding. The utility of this approach has been demonstrated with the prediction of a putative GAP domain in amino acids 1295 to 1453 (see Fig 4.1). Protein homology searches using short segments of GRIPE sequence, and with the search program FastA (Pearson and Lipman, 1988), revealed weak but significant homology of GRIPE with a bHLH gene known as ESR-5 (see Fig 4.2). Curiously, amino acids 816 to 1000 of GRIPE displayed sequence homology to the HLH domain of ESR-5 (encoded in aa 14 to 120; Jen et al., 1999), suggesting that GRIPE may contain a HLH-like domain.

![Figure 4.2. Alignment of amino acids 816 – 1000 of GRIPE with the bHLH gene ESR-5 (AAD42782). The similari.jpg](attachment:image.png)
Chapter 4: GRIPE is a novel binding partner to E12, and may regulate its dimersation state in the nucleus

Based on this information, three deletion constructs of the GRIPE prey were designed with respect to the predicted GAP and HLH-like domains. The first deletion construct, named GRIPE$^{933-1485}$, encodes an N-terminal deletion of 232 amino acids and lacks the predicted HLH domain. The second prey is GRIPE$^{701-932}$, a C-terminal deletion construct which lacks the putative GAP domain. Finally, a small C-terminal deletion in the GRIPE$^{701-1458}$ prey was also prepared. All three preys could not activate yeast reporter genes in the presence of a spurious bait protein, p53 (data not shown).

Next, these preys were assessed for interaction with the mouse E12$^{505-651}$ bait, and the results presented in Table 4.2. All mutants of GRIPE, however, failed to demonstrate binding; the N-terminal deletion mutant GRIPE$^{933-1485}$ (row 4) failed to activate reporter genes in the presence of the E12$^{505-651}$ bait. Similarly, the C-terminal mutant GRIPE$^{701-932}$ (row 5) demonstrated a lack of reporter gene expression in the presence of E12$^{501-651}$. To rule out the possibility that GRIPE may bind E12 within a domain that was disrupted in these two preys, the mutant GRIPE$^{701-1458}$ was constructed which encoded a small 27 amino acid deletion in the C-terminus. This prey construct, however, also failed to bind E12 (row 6). In light of these observations, it is concluded that, within the context of the yeast 2-hybrid assay, GRIPE$^{701-1485}$ comprises the minimal domain that binds E12.

4.4. Coimmunoprecipitation experiments show that GRIPE is a bona fide interacting partner to E12.

To confirm the interaction between E12 and GRIPE observed in the yeast 2-hybrid interaction assay, a co-immunoprecipitation experiment was conducted. Because a native antibody to GRIPE was not available, these experiments were carried out using C-terminal fusion of GRIPE protein to Enhanced Green Fluorescent Protein (EGFP). Similarly, a Flag epitope-tagged E12$^{505-651}$ was constructed for this purpose (see Appendix 10.7). Transiently transfected HEK293T cells harbouring expression plasmids encoding E12$^{505-651}$ and EGFP-GRIPE$^{1-1485}$ were harvested, and immunoprecipitation then carried out as described in section 2.10. As a control experiment, HEK293T cells were transfected with FLAG-E12$^{505-651}$ and EGFP protein only to account for non-specific aggregation of E12 to EGFP. As shown in Fig 4.3, E12$^{505-651}$ protein is immunoprecipitated with the EGFP-GRIPE$^{701-1485}$ fusion protein (lane 1), but not with EGFP
Table 4.2. GRIPE binds to the helix-loop-helix domain of E12, as demonstrated using a yeast two-hybrid interaction assay. Truncated E12 baits were prepared as described in Appendix 10.7. The letter “b” denotes the basic amino acid region responsible for contacting DNA, while H1 (helix 1), L (loop), H2 (helix 2) denotes the helix-loop-helix domain of E12. The putative GAP domain of GRIPE is shaded black on the polypeptide. An intact HLH domain of E12 is required for binding to GRIPE (rows 1 to 3). Deletion analysis with GRIPE mutants demonstrates that the entire region of GRIPE701-1485 is required for binding in this assay (rows 4 to 6).

<table>
<thead>
<tr>
<th>Bait</th>
<th>Prey</th>
<th>β-gal activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) µE12505-651</td>
<td>GRIPE701-1485</td>
<td>++</td>
</tr>
<tr>
<td>2) µE12505-600</td>
<td>GRIPE701-1485</td>
<td>+++</td>
</tr>
<tr>
<td>3) µE12505-571</td>
<td>GRIPE701-1485</td>
<td>-</td>
</tr>
<tr>
<td>4) µE12505-651</td>
<td>GRIPE701-932</td>
<td>-</td>
</tr>
<tr>
<td>5) µE12505-571</td>
<td>GRIPE701-1458</td>
<td>-</td>
</tr>
<tr>
<td>6) µE12505-651</td>
<td>GRIPE701-1458</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 4.3. E12 associates with GRIPE in vivo. HEK293T cells were transfected with plasmids encoding EGFP-GRIPE747-1531 and FLAG- E12501-651 fusion proteins (1), or with EGFP and FLAG-E12501-651 plasmids (2). (A) FLAG-tagged E12501-651 protein was immunoprecipitated in the presence of EGFP-GRIPE747-1531 and to a very minor extent with EGFP (arrowhead). “N.S.” represents non-specific bands detected using antibodies to FLAG epitope. (B) Approximately equal amounts of detectable FLAG-E12501-651 were used in both protein preparations (arrowhead). (C) EGFP-GRIPE747-1531 (arrowhead) and EGFP (arrow) were detectable in a Western blot of their respective protein preparations.
protein alone (lane 2). To rule out the possibility that band intensities are due to differences in protein loading, Western blots were also performed using antibodies to Flag and EGFP epitopes. The results confirmed expression of all proteins tested (Fig. 4.3B, arrowhead; Fig. 4.3C, arrows). Further, coimmunoprecipitations incubated with buffers of increasing ionic strength show this interaction is abolished in the presence of 500mM sodium chloride (Fig. 4.4). Once again, Western analysis accounted for proper loading of E12 and EGFP-GRIPE proteins in those preparations. Taken together, these results indicate that GRIPE is a bona fide binding partner to E12.

4.5. Colocalisation of GRIPE with E12 in HEK293T cells suggests a nuclear function for GRIPE

To further demonstrate an interaction between GRIPE and E12, experiments were devised to observe the cellular localisation of expressed full-length E12 and EGFP-GRIPE\textsuperscript{701-1485} in HEK293T cells. Cells transfected with EGFP alone showed widespread expression in both cytoplasm and nucleus (Fig. 4.5A, arrows). When a EGFP-GRIPE\textsuperscript{701-1485} expression plasmid was introduced, the distribution of the fusion protein, revealed by EGFP fluorescence, was predominantly in the cytoplasmic compartment (Fig. 4.5B, arrows), with few cells (<1%) exhibiting pancellular distribution (Fig. 4.5B, arrowhead). To rule out the possibility that this pattern of EGFP fluorescence was peculiar to HEK293T cells, other cell lines were also tested. Transfection of the EGFP-GRIPE\textsuperscript{701-1485} plasmid into monkey COS-7 epithelial cells, as well as into mouse P19 embryocarcinoma cells also produced EGFP fluorescence predominantly in the cytoplasm (data not shown). In view of these findings, all subsequent transfections were carried out with HEK293T cells. Next, a mammalian expression vector encoding an N-terminal FLAG epitope-tagged E12\textsuperscript{1-651} construct was prepared according to detailed instructions (see Appendix 10.7). Utility of this construct was then confirmed by Western analysis of cell extracts harvested from transiently transfected cells. As shown in Fig. 4.5H, transfection of E12\textsuperscript{1-651} resulted in immunodetection of a predicted band by E12 and FLAG antibodies (lanes 1 and 2 represent two independent clones evaluated), and is absent in extracts of untransfected cells (lane “U”). Introduction of E12\textsuperscript{1-651} alone resulted in the expected nuclear localisation of this protein in transfected cells (Fig. 4.5C, arrows), and is attributable to the nuclear localisation sequence coded by amino acids 160 to 176, as predicted by PROSITE (http://kr.expasy.org/prosite/), and
Figure 4.6. **GRIPE may negatively regulate E12-dependent target gene transcription.** Lane 1: p(µE5+µE2)6-Luc+ only; Lane 2: p(µE5+µE2)6-Luc+ + FLAG-E12; Lane 3: p(µE5+µE2)6-Luc+ + FLAG-E12 + EGFP-GRIPE (x1); Lane 4: p(µE5+µE2)6-Luc+ + FLAG-E12 + EGFP-GRIPE (x2); Lane 5: p(µE5+µE2)6-Luc+ + FLAG-E12 + EGFP-GRIPE (x4); Lane 6: p(µE5+µE2)6-Luc+ + FLAG-E12 + EGFP (x1); Lane 7: p(µE5+µE2)6-Luc+ + FLAG-E12 + EGFP (x2). E12 protein was assayed by Western analysis using FLAG antibodies, while EGFP-GRIPE and EGFP proteins were detected using a GFP antibody. Luciferase activities are expressed relative to βgal expression (Raw data is presented in Appendix 10.14). Error bars represent S.E.M. of triplicate experiments.

**Figure 4.4. GRIPE associates with E12 in high ionic strength buffers.** Coimmunoprecipitations were performed in buffer containing increasing concentrations of sodium chloride, and with EGFP-GRIPE and E12 (1) or EGFP and FLAG-E12 (2). FLAG-E12 is immunoprecipitated by EGFP-GRIPE even in the presence of 250 mM NaCl (arrow), while negligible FLAG-E12 is bound by EGFP in all treatments. Equal amounts of FLAG-E12 and were detected in both preparations. Similarly, EGFP-GRIPE and EGFP were detectable in their respective preparations (data not shown).
Figure 4.5. **Localisation of EGFP-GRIPE<sup>701-1485</sup> or E12 in transfected cells.** (A) EGFP protein is distributed throughout the cell (arrows). (B) The majority of cells (>99%) which express EGFP-GRIPE fusion protein exhibit cytoplasmic staining (arrows), while a few cells show pan-cellular staining (arrowhead). (C) Cells transfected with FLAG-tagged E12<sup>1-651</sup> protein exhibit nuclear staining (arrow). (D-G) Arrows indicate doubly transfected cells that express EGFP-GRIPE (D) and FLAG-E12 (E). (F) An overlay of fluorescence signals reveals colocalisation of both signals in the nucleus. (G) Cell nuclei are stained with bisbenzamide. Scale Bar represents 20 µm. (H) Western blot extracts from transient transfections of two independent FLAG-E12 clones (lanes 1 and 2); a signal of predicted size is detected by FLAG and E12 antibodies, and is absent in untransfected cells (lane "U").
confirmed by others (Deed et al., 1996). Strikingly, however, double transfection of EGFP-GRIPE\textsuperscript{701-1485} and E12\textsuperscript{1-651} (with GFP and Flag epitopes respectively) revealed colocalization of both proteins in cell nuclei (Figs 4.5D – 3F). These results suggest that EGFP-GRIPE\textsuperscript{701-1485}, which is normally found in the cytoplasm, has the capability to be imported into the nucleus upon binding to full-length E12, and further demonstrates a physical interaction between these two proteins in vivo.

4.6. GRIPE negatively regulates E12-dependent transcription of a reporter gene

Following from the observations that $E12^{1-651}$ could be a molecular chaperone for GRIPE, it remains to be tested if GRIPE might interfere with the transcriptional activity of $E12$ upon nuclear entry. To test this, an expression assay was designed to observe the potential for GRIPE to influence expression of an E12-dependent luciferase reporter. The promoter element of this luciferase reporter plasmid, p($\mu E5 + \mu E2)_6$-Luc\textsuperscript{+}, encodes 6 tandem repeats of the ($\mu E5 + \mu E2$) region of the IgH promoter (Ruezinsky et al., 1991). This region contains several E-boxes, and is recognised and bound by $E12$ to activate target gene expression. Insignificant luciferase activity is detected in cells transfected with p($\mu E5 + \mu E2)_6$-Luc\textsuperscript{+} alone (Fig 4.6, lane 1), and co-transfection of EGFP\textsuperscript{701-1485} or EGFP with this construct did not affect reporter gene expression (data not shown). However, transfection of $E12$ with p($\mu E5 + \mu E2)_6$-Luc\textsuperscript{+} results in high levels of luciferase activity (lane 2). Introduction of increasing amounts of EGFP-GRIPE\textsuperscript{701-1485} results in a dose-dependent decrease in luciferase gene activation by $E12$ (lanes 3 to 5). This reduction is not attributable to differences in $E12$ protein expression in each preparation, as detected by Western analysis. To omit the possibility that EGFP may be responsible for this reduction in luciferase activity, parallel experiments with increasing amounts of EGFP demonstrated a negligible affect on $E12$-dependent luciferase expression (lanes 6 and 7). Taken together, these observations indicate that GRIPE can behave as a negative regulator of $E12$-dependent target gene transcription.
4.7. Summary

4.7.1. GRIPE is a novel interacting partner to E12

Using a yeast 2-hybrid interaction screen, GRIPE was cloned as a novel interacting partner to E12. This interaction requires an intact HLH domain of E12 for binding GRIPE. Unfortunately, experiments designed to fine-map the E12-binding domain of GRIPE were uninformative, since all deletion mutants of the GRIPE prey failed to activate yeast reporter genes in the presence of E12 bait. Therefore, in the context of the yeast 2-hybrid assay, the entire 785 amino acids of GRIPE is required for E12 binding.

Coimmunoprecipitations provided independent confirmation of protein-protein interaction first detected in the yeast 2-hybrid assay, and this behaviour \textit{in vitro} was manifested as a capacity for E12 to import GRIPE to the nucleus \textit{in vivo}. Further, experiments with an E12-dependent luciferase reporter demonstrated that GRIPE appears to negatively regulate target gene transcription by E12. This function may be attributable to the capacity for GRIPE to bind the HLH domain of E12, thus preventing the formation of functional HLH dimers with E12, which are crucial for binding E-boxes on target promoter DNA.

While these biochemical approaches demonstrate physical interaction between E12 and GRIPE, it remains to be determined if these two genes are coincidently expressed \textit{in vivo}, since a lack of co-expression would infer that the combinatorial role of \textit{E12} and \textit{GRIPE} may never eventuate in developing e11.5 mouse forebrain neurons. To this end, a comprehensive survey of mRNA expression by Northern analysis, as well as \textit{in situ} hybridisation experiments, was conducted. These results are presented in the following chapter.

4.7.2. Lack of activation by deletion mutants of GRIPE prey may be explained by limitations of the yeast 2-hybrid assay

An alternative explanation for the lack of yeast reporter activation in the deletion constructs of GRIPE prey pertain to prey protein expression; while all constructs were sequence-verified, these were not evaluated for protein expression when transformed in yeast. Once transformed, all yeast cell lines harbouring prey plasmids were viable, as judged by their growth on media lacking leucine, which selects only for maintenance of the prey plasmid. While expression of the prey-fusion gene is under the control of a
constitutively active alcohol dehydrogenase promoter, it was not determined if the encoded prey protein was properly transcribed and translated for interaction with the E12 bait protein. Certainly, complications in protein expression or processing, such as the mis-targeting of prey-fusion protein to organelles rather than the nucleus, may explain the lack of reporter activation by these constructs. To account for this, protein extracts could be harvested, and Western analysis and immunocytochemistry performed using antibodies which target the Gal4TAD portion of the fusion protein. Since this was not done, it cannot be ruled out that truncated forms of GRIPE\textsuperscript{701-1485} did not interact with E12 because of complications arising from proper protein expressed in prey-transformed yeast cells.
Author/s: Heng, Julian Ik Tsen

Title: Cloning and characterisation of gripe: a novel interacting partner of e12 during brain development

Date: 2002-10


Publication Status: Unpublished

Persistent Link: http://hdl.handle.net/11343/39471

File Description: p.1-64

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.