Investigations of ephrin ligands during development

A thesis submitted for the degree of Doctor of Philosophy

Molecular Biosciences, Adelaide University

By

Paul Tosch, B.Sc (Hons.)

Department of Molecular Bioscience,

Adelaide University, Australia,

Adelaide, South Australia, 5005

May 2002

Declaration

This work contains no material, which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge contains no material previously published or written by another person, except where due reference has been made in the text.

I give full consent for this copy of my thesis to be made available for loan and photocopying when deposited in the university library.

Paul Tosch,

May 2002

I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right.

Albert Einstein (1879–1955), German-born U.S. theoretical physicist.

Acknowledgments

Well what an amazing experience my PhD has been, I am not sure if I have grown more as a person or more as an academic. It is obvious that I could not have achieved any of this without the contributions of a large number of people. I would like to thank the following people:
Paul Moir for driving me across the desert and giving me somewhere to live, he is like the brother I never had (it's getting a bit sand duney around here hey). Jacinta Makin for putting up with and listening to countless theories, and teaching me to ride horses.

My two supervisors Dr Simon Koblar and Dr Robert Saint for providing me with the resources to perform this work.

The following postdocs for providing challenging my mind and pushing me beyond my limits: Dr Tetyana Shandala, Dr Daniel Kortschak, Dr Rory Hope, Dr Tim Cox.

The members of the Koblar lab past and present: Agnes Stokowski, Edwina Ashby, Chathurani Jayesena (Chaza), Anthony Condina, Dr Warren Flood (Wazza), Dr Robert Moyer , and Rebecca McLennon

The members of the Hope lab: Dave Wheeler, Scott Spargo

The members of the Saint lab: Julianne Camerotto, Volkan Evci, Greg Somers (Gregor), Peter Smibert, Donna Crack, Jane Sibbons, Michelle Coulson,

Member of the genetics department: Velta Vingelis, Itty Bitty Kitty

People who have helped out at various times: Rebecca Patrick, Vera Shandala, Ian Miller, Calluna Denwood, James Brooks, Matthew Kelcey, Dr Peter Gunn, and Martin Blanchard.

Albert Einstein, Gene Roddenberry, to me these are the two most inspirational minds to live this century.
Dr Anthony Koutoulis (Kouts) for introducing me to Molecular biology, you are a legend.

I think it is obvious that none of this would have been possible without a large number of my friends and colleagues. In fact, so many people have assisted me over the years that it is impossible to thank them all, but rest assured that if you have talked to me for any length within the last 27 years then you have assisted me, and I thanks you. I am a simple man with a simple mind, so I ask many simple questions; but the whole really is greater than the sum of its parts.

Abstract

The Eph receptors and their ligands, the ephrins, are the largest subfamily of receptor tyrosine kinases (RTKs). Functional studies of this kinase family have demonstrated their importance in various aspects of embryonic development. These include migration of precursor neural crest cells (Krull *et. al.*, 1997; Krull, 1998; Koblar *et. al.*, 2000), vascular development (Wang *et. al.*, 1998), axonal guidance and bundling of nerves in pathway formation in the nervous system, and hindbrain segmentation in embryonic development (Wilkinson, 2001).

The overall aim of this thesis was to isolate ephrin ligands from *Drosophila melanogaster,* and to analyse their involvement in *Drosophila* development. In addition, the potential of ephrin-B1 as a causative gene in the human condition Aicardi's syndrome was also investigated.

The release of the *Drosophila* genome in 1999 revealed one full length EST coding for a gene that contained an ephrin core domain, designated *d-ephrin*. *In situ* analysis shows that *d-ephrin* mRNA is restricted to the CNS of *Drosophila* embryos at the time of axon pathfinding, suggesting that *d-ephrin* could play a role in nervous system development.

Bioinformatics analysis on d-ephrin was unable to assign d-ephrin to either of the recognised ephrin-A or ephrin-B subclasses. Tissue culture experiments demonstrated that d-ephrin has an affinity for the only currently known *Drosophila* Eph (Dek). Analysis of misexpressed *d-ephrin* in *Drosophila* embryos showed axon guidance defects in the ventral nerve cord of *Drosophila*.

Detailed evolutionary analyses of all the currently known ephrin genes are in agreement with the current system of nomenclature derived by structural and functional studies. Furthermore, these evolutionary analyses placed the invertebrate ephrins equidistant from the ephrin-A and ephrin-B subclass, respectively.

Aicardi's syndrome is an X-linked dominant disorder. Patients with this disorder typically present with a distinct collection of symptoms, including callosal agenesis, and retinal pigmentary defects (Aicardi *et. al.*, 1969). Genetic evidence taken from mouse models suggests that *ephrin-B1* is involved in callosal formation (Henkemeyer *et. al.*, 1996). DNA sequence analysis of the *h-ephrin-B1* genomic region in six known Aicardi's patients was performed to determine if h-ephrin-B1 is the principal disease causative gene of this disorder, however no mutations in the five exons of *h-ephrin-B1* were found.

Publications resulting from the work in this thesis

Analysis of the X-linked ephrin-B1 gene as a candidate For Aicardi syndrome (Submitted, American Journal of Medical Genetics, 2002)
Paul Tosch[1], Simon Koblar[1], Tim Cox[1]

Characterisation of a *Drosophila* ephrin gene (*d-ephrin*) and molecular evolution of the ephrin gene family. (Submitted, Genes Development and Evolution, 2002)
Paul Tosch[1], Rory Hope[2], Tetyana Shandala[1], Daniel Kortschak[1], Robert Saint[3], Simon Koblar[1]

---

[1] Centre for the Molecular Genetics of Development, Department of Molecular Biosciences, Adelaide University, South Australia 5005, Australia.
[2] Laboratory of Molecular Evolution, Department of Molecular Biosciences, Adelaide University, South Australia 5005, Australia.
[3] Research School of Biological Sciences, Institute of Advanced Studies, Australian National University, Canberra ACT, 0200 Australia

Table of Contents

List of Figures and Tables

Figures

Tables

Chapter 1: Introduction

Introduction

Receptor Tyrosine Kinases (RTKs)

The question of how a multicellular organism grows from a few cells into specialised tissue types and ultimately a complete organism is fundamental to the study of biology.

The receptor tyrosine kinases (RTKs) are a large group of proteins that serve as key components in the transduction of extracellular signals across the cell membrane to mediate cell-cell interactions within a developing organism. These receptors together with their corresponding ligands regulate growth and tissue morphogenesis during embryonic development. There are a large number of RTKs that have been characterised, which all share some common features including: an extracellular ligand binding domain, a transmembrane region, and an intracellular kinase domain. These receptors are classified into families according to the structure of their extracellular regions, although the structure of the intracellular tyrosine kinase domain remains fairly conserved in all RTKs (reviewed in Van Der Geer *et. al.*, 1994).

The Eph/ephrin receptor tyrosine kinases

The Eph subfamily of RTK currently consists of 14 receptors. The founding member, EphA1 was isolated in a low stringency screen for tyrosine kinases from an erythropoietin producing human hepatocellular cell line, and thus named Eph (Hirai *et. al.*, 1987). On the basis of their structural similarities and binding affinities, the Ephs and ephrins (Eph family receptor interacting proteins) are subdivided into two subclasses, EphA and EphB and their corresponding ligands ephrin-A and ephrin-B respectively. The EphA receptors in the main bind to ephrin-A ligands. Similarly the EphB receptors bind to the B subclass. There is considerable promiscuous binding within each subclass of Eph-ephrin, but no significant binding between subclasses with the exception of EphA4, which also binds ephrin-B2 and ephrin-B3 ligands (Figure 1) (Bartley *et. al.*, 1994; Beckmann *et. al.*, 1994; Cheng and Flanagan, 1994; Davis *et. al.*, 1994; Bennett *et. al.*, 1995; Bergemann *et. al.*, 1995; Cerretti *et. al.*, 1995; Kozlosky *et. al.*, 1995; Brambilla *et. al.*, 1996; Gale *et. al.*, 1996a; Gale *et. al.*, 1996b; Sakano *et. al.*, 1996; Monschau *et. al.*, 1997; Park and Sanchez, 1997; Menzel *et. al.*, 2001).

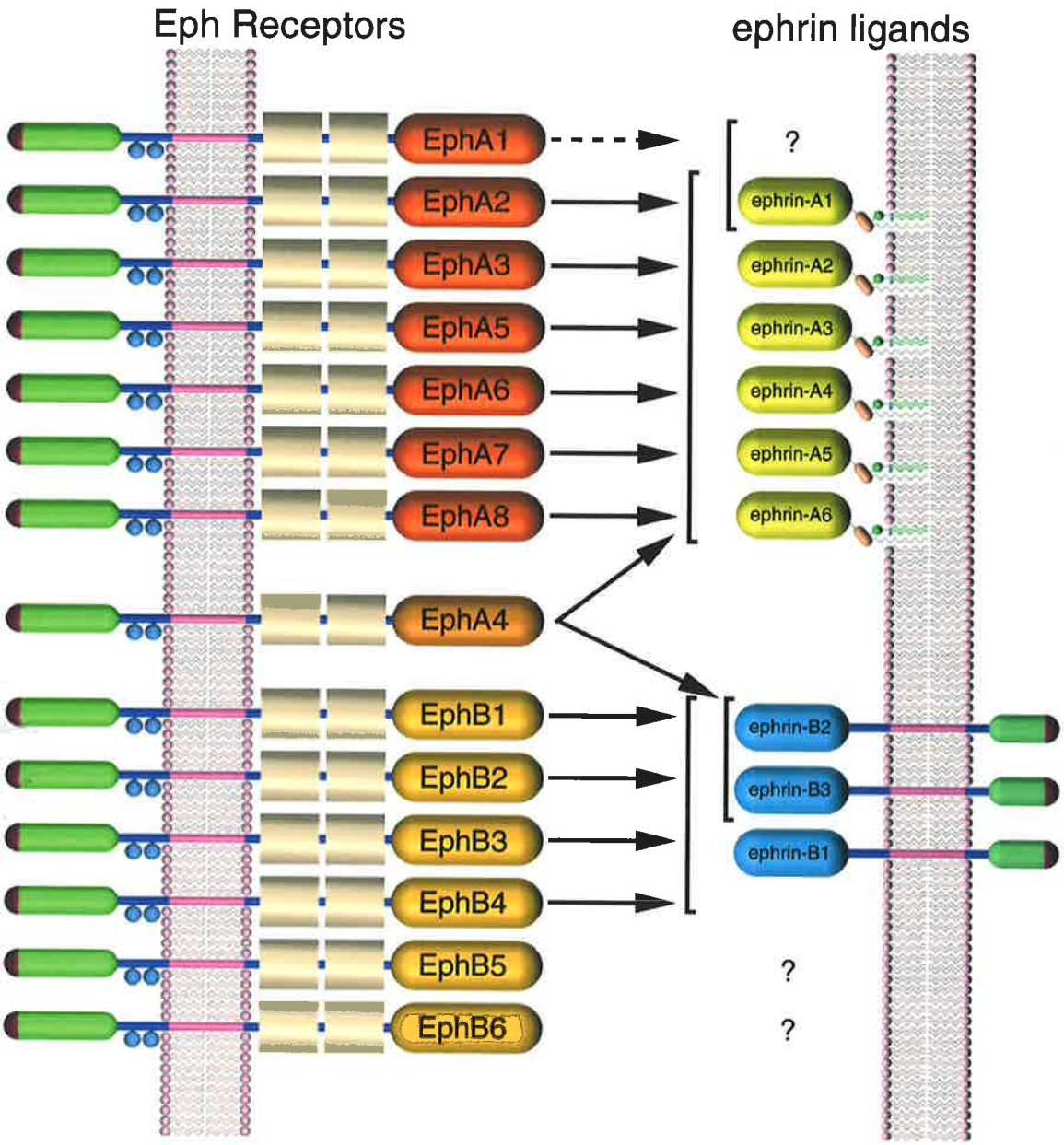


Figure 1 Summary of the currently known Eph receptors and their corresponding ligands, the ephrins, in vertebrate species. The arrows between the receptors and ligands denote reported significant (K_d<1m) binding interactions. Binding is promiscuous within the EphA/ephrin-A and EphB/ephrin-B subclasses, but not between subclasses with the exception of EphA4 which does show significant binding affinity to ephrin-B2 and ephrin-B3. EphA1 binds only to ephrin-A1 with a low affinity (broken arrow), suggesting there may be another ligand to be isolated. Also EphB5 and EphB6 show no significant binding affinity to any known ephrin ligand again suggesting that new ligands remain to be isolated (binding data from Bartley *et. al.*, 1994; Beckmann *et. al.*, 1994; Cheng and Flanagan, 1994; Davis *et. al.*, 1994; Bennett *et. al.*, 1995; Bergemann *et. al.*, 1995; Cerretti *et. al.*, 1995; Kozlosky *et. al.*, 1995; Brambilla *et. al.*, 1996; Gale *et. al.*, 1996a; Gale *et. al.*, 1996b; Sakano *et. al.*, 1996; Monschau *et. al.*, 1997; Park and Sanchez, 1997; Menzel *et. al.*, 2001).

Eph/ephrins across species

To date most of the Eph and ephrins have been isolated in *Homo sapiens*, *Gallus gallus*, and *Mus musculus*. In warm-blooded animals the overall number of receptors and ligands has been mostly conserved. In cold-blooded animals, *Xenopus laevis* and *Danio rerio*, fewer Eph and ephrins have been isolated to date. At least some of these appear to be orthologs of particular mammalian proteins, and it been suggested that similar numbers may exist (Flanagan and Vanderhaeghen, 1998). However, in lower species the number of Eph/ephrins is significantly reduced. In *Drosophila melanogaster* one Eph receptor (Dek) (Scully *et. al.*, 1999) and one ligand has been isolated (d-ephrin) (Tosch *et. al.*, 2002). In *Caenorhabditis elegans*, one Eph (VAB-1) (George *et. al.*, 1998) and four ligands (EFN-1 to 4) (Chin-Sang *et. al.*, 1999; Wang *et. al.*, 1999) are currently known. The presence of Eph and ephrins in *Drosophila melanogaster* and *Caenorhabditis elegans* indicates that the family is ancient, perhaps dating back to the origins of the metazoans (Flanagan and Vanderhaeghen, 1998).

Eph/ephrin binding

In vitro binding studies

Various binding studies using IgG or alkaline phosphatase tagged soluble ligands or receptor binding domains show that on the whole the binding specificities of the Eph receptors to the ephrin ligands fall into two main classes, corresponding to anchorage mechanisms of the ligands (Gale *et. al.*, 1996a; Gale *et. al.*, 1996b). Firstly, the EphA receptors show an affinity for the ephrin-A ligands (Figure 1). Secondly, the EphB receptors show an affinity for the ephrin-B ligands (Figure 1). Receptors from each subclass generally show an affinity for all ligands within that same subclass (Gale *et. al.*, 1996a; Gale *et. al.*, 1996b). However, there are exceptions to the rule, such as EphA1, which only binds to ephrin-A1 with the lowest affinity within the group (K_d=2.67 m), suggesting that there may be another ligand yet to be found for EphA1 (Gale *et. al.*, 1996b), as EphA4 has a much higher affinity for ephrin-A1 (K_d=0.39m). Furthermore, in the B subclass, EphB5 and B6 show no binding affinities to any of the known ephrin-B ligands, again suggesting that within this subclass new ligands remain to be isolated. Also of particular note is the ability of EphA4 to bind to ephrin-B2 (Gale *et. al.*, 1996b) and ephrin-B3 (Gale *et. al.*, 1996a).

The significance of *In vitro* binding studies

To date binding studies have been assayed with one member of the binding pair in the soluble form. When the receptor-ligand interactions occur *in vivo* the binding specificities are likely to be much more stringent owing to the fact that both the ligand and receptor are confined to the interacting cell surfaces. However, the lower affinities may also be significant due to increased concentration of Ephs and ephrins via sequestering into lipid raft micro domains within the cell membrane (Wilkinson, 2000).

The significance of the binding specificity of ligands to their corresponding receptors has also been correlated with their ability to stimulate Eph receptor tyrosine kinase activity. Ligands from the different subclasses usually fail to activate receptors from the other subclass. The binding affinity studies, together with the tyrosine activation studies, indicate that the promiscuous binding of the Eph/ephrins within each subclass at least, is likely to be biologically significant (Brambilla *et. al.*, 1995; Brambilla *et. al.*, 1996). The high degree of promiscuity of Eph-ephrin binding, suggests that there may be a high level of redundancy and functional overlap, as shown with the double EphB2 and EphB3 mouse knockout, which had a markedly more severe phenotype than those with a mutation in only EphB2 and EphB3, indicating that the two receptors act in a partially redundant fashion (Orioli *et. al.*, 1996).

In another study, mice that were homozygous for a mutation in EphA2 did not exhibit any discernable phenotype. Western analysis and *in vitro* kinase assays showed that EphA2 null mice are deficient for EphA2. However the lack of phenotype suggests that other members of the Eph family of receptor tyrosine kinases can functionally compensate for the loss of EphA2 (Chen *et. al.*, 1996).

Isolation of the Eph/ephrins

The Eph receptors have so far been identified by approaches based on either conservation of sequence or catalytic activity in the kinase domain (Flanagan and Vanderhaeghen, 1998). EphA1 was cloned using a cDNA probe for a tyrosine kinase region, which is highly conserved in a number of the receptor tyrosine kinase subfamilies (Hirai *et. al.*, 1987). EphA2 was isolated by a degenerate PCR technique with the kinase region amino acid sequence HDLAAR that is also highly conserved (Lindberg and Hunter, 1990). EphA3 was isolated from 10-day-old chicken embryos using anti-phosphotyrosine antibodies (Sajjadi *et. al.*, 1991). Cross hybridisation or PCR of kinase domain encoding sequences was subsequently used to isolate most of the other members of the Eph receptors (Flanagan and Vanderhaeghen, 1998).

The identification of the Eph receptor ligands lagged behind the Eph receptors (Pandey *et. al.*, 1995a). Typically ligands for RTK have weak sequence conservation, making it impractical to use sequence homology to identify ligand families (Flanagan and Vanderhaeghen, 1998). The ligands for the Eph family of receptors were initially identified using affinity chromatography or expression library screening with the Eph extracellular domain. Ephrin-A1 was purified by using an affinity column containing the extracellular domain of EphA2 (Bartley *et. al.*, 1994). Ephrin-A2 was isolated from an expression library using alkaline phosphatase tagged EphA3 extracellular domain (Cheng and Flanagan, 1994). Similarly, ephrin-A3 and ephrin-A4 were first isolated using an EphA3 extracellular domain IgG fusion as a probe (Davis *et. al.*, 1994; Kozlosky *et. al.*, 1995). Ephrin-A5 was isolated simultaneously via an EphA5-IgG fusion (Winslow *et. al.*, 1995; Lackmann *et. al.*, 1996) and in a two dimensional electrophoresis as repulsive axonal guidance signal (RAGS), which was found to be enriched in the posterior tectum (Drescher *et. al.*, 1995). In a screen for EphA4 ligands, ephrin-A6 was identified and found to be expressed in chicken embryonic retina (Menzel *et. al.*, 2001). Ephrin-B1 and ephrin-B2 were isolated using EphB1-IgG and EphB4-IgG affinity probes respectively (Beckmann *et. al.*, 1994; Davis *et. al.*, 1994; Shao *et. al.*, 1994; Bennett *et. al.*, 1995). Ephrin-B2 was also independently cloned using a probe of an EST from human chromosome 13 which showed a 57% homology to ephrin-B1 (Bergemann *et. al.*, 1995; Cerretti *et. al.*, 1995). Ephrin-B3 was also isolated using a probe from an EST fragment having significant homology to the Eph ligands (Gale *et. al.*, 1996a; Nicola *et. al.*, 1996). Currently there are 45 orthologous (Appendix A) ephrin ligands described in databases, aside from the initial ephrin ligands which were isolated via receptor affinity methods, the majority of the orthologs were isolate by taking advantage of the high degree of sequence homology unique to the ephrin ligand subfamily.

Eph structure

The extracellular region of the Eph receptors is characterised by a number of key domains, including a cysteine rich region with 20 conserved cysteine residues, two fibronectin type III motifs (FNIII) (Pasquale, 1991), a hydrophobic transmembrane region (TM) and an intracellular phosphotyrosine kinase domain (Kinase), which is involved in cell signaling upon ligand binding (Figure 2) (Zisch *et. al.*, 1997). Originally the N-terminal region of the Ephs was thought to contain an immunoglobulin like structure (Labrador *et. al.*, 1997). However, subsequent analysis using a series of deletion and domain substitution mutants, disputed the presence of an immunoglobulin structure (Labrador *et. al.*, 1997). This work showed that the EphB2 globular domain is the ligand-binding domain of the ephrin-B1 ligand. Subsequent X-ray crystallography experiments resolved the extracellular domain structure, and determined that it is comprised of 11-anti-parallel β-sheets, with a ligand interacting loop (Himanen *et. al.*, 1998). This initial structural analysis provided an insight into a potential ligand interacting domain of the Eph extracellular region (discussed latter).

Ephrin structure

From 1994 nine ephrin ligands have been cloned, both the ephrin-A and the ephrin-B proteins share a highly conserved extracellular core region corresponding to the receptor-binding domain, which contains four invariantly positioned cysteine residues (Himanen *et. al.*, 2001). There is a high degree of sequence identity, with up to 70% within the core region of the ephrins (Flanagan and Vanderhaeghen, 1998). This high degree of sequence homology among the ephrin ligands is unique to this family of RTK ligands, with most RTK ligand families showing little or no sequence homology (Van Der Geer *et. al.*, 1994). The two subclasses of ephrins utilise a different anchorage mechanism, the ephrin-As tethered to the cell membrane by a glycosylphosphatidylinositol (GPI) anchor, while the ephrin-B ligands are transmembrane proteins, with an intracellular region that is 98% conserved on the cytoplasmic tail. This conserved tail can serve as a receptor and becomes tyrosine phosphorylated upon receptor binding, suggesting that ephrin-Bs are involved in signal transduction within the cell (Figure 2) (Holland *et. al.*, 1996).

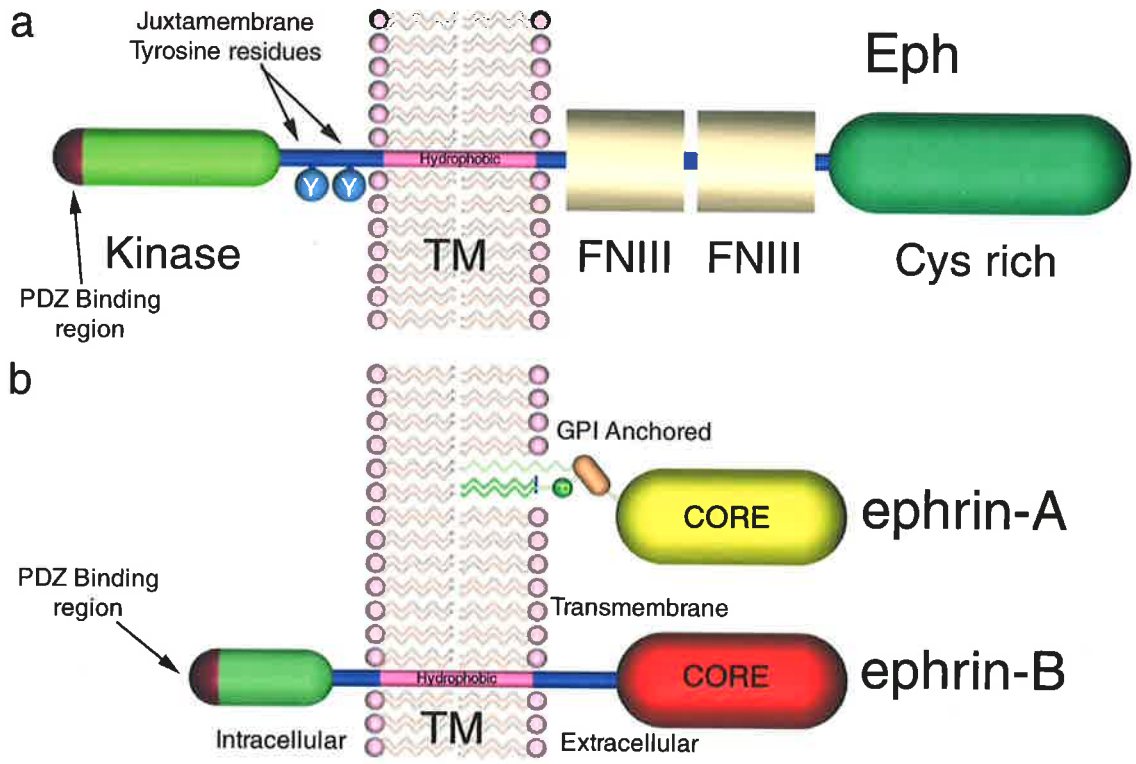


Figure 2 Structure of the Eph receptors and their ligands the ephrins. (a) Structural features common to all Eph proteins include: a Cysteine rich region (Cys rich) with 18 conserved cysteine residues which form the globular ligand binding domain, two fibronectin type III motifs (FNIII), a hydrophobic transmembrane region (TM), an intracellular phosphotyrosine kinase domain (Kinase) which is involved in cell signalling upon ligand binding, and a PDZ binding region on the C-terminus. (b) Cell membrane anchorage of the GPI linked ephrin-As and the transmembrane ephrin-Bs. The ephrin-As are linked on the outer leaflet of the cell membrane via covalent linkage to a glycosylphosphatidylinositol (GPI) moiety. The ephrin-Bs are transmembrane proteins with both an intracellular and extracellular globular domains linked by a membrane spanning hydrophobic region. Furthermore, the intracellular region of the ephrin-Bs has a PDZ binding domain. The ephrin ligands all share a common extracellular receptor-binding core with four conserved cysteine residues (core).

Intracellular signalling in ephrins

Transmembrane ligand Signalling

Analysis of VAB-1 (Eph) in the nematode (George *et. al.*, 1998) and EphB2 in the mouse (Henkemeyer *et. al.*, 1996) indicates that Eph receptors have a kinase-dependent and kinase-independent function (i.e. Ephs are acting as a ligand), raising the possibility that ephrin-Bs are able to propagate intracellular signals. Also, ephrin-B1 and ephrin-B2 become phosphorylated on the highly conserved intracellular tail (Figure 3) upon binding of the EphB2 extracellular domain *in vitro* (Holland *et. al.*, 1996; Bruckner *et. al.*, 1997). Furthermore, expression of the cytoplasmic tail of *Xenopus* ephrin-B, leads to loss of cell adhesion, as does expression of an ephrin ligand missing the extracellular domain (Jones *et. al.*, 1998). Taken together these data suggest that the cytoplasmic domain of ephrin-Bs have an important role in signaling within the cell, via interaction with other cellular proteins.

Furthermore, phosphorylation of the tyrosine residues within the ephrin-B tail may also modulate the binding of intracellular proteins in a positive or negative fashion (Holland *et. al.*, 1996).

PDZ binding motifs at the C-terminus of membrane proteins interact with other proteins containing a PDZ domain to form macromolecular complexes involved in signal transduction pathways (reviewed in Saras and Heldin, 1996; Craven and Bredt, 1998). The PDZ domain was initially identified as a common homology region in postsynaptic density protein PSD-95, the *Drosophila* discs-large tumour suppressor protein DlgA, and the tight junction protein ZO-1 (PDZ) (Woods and Bryant, 1991; Cho *et. al.*, 1992; Woods and Bryant, 1993; Kennedy, 1995; Kornau *et. al.*, 1995).

The C-terminus of all ephrin-B proteins contains the motif (Tyr-Try-Lys-Val, Songyang *et. al.*, 1997), which is a known PDZ binding motif (Figure 3), suggesting a mechanism by which ephrin-Bs may interact with cytoplasmic proteins. Work by Lin *et. al.,* (1999) tested this hypothesis by screening a 10.5 day old mouse embryonic expression library with a biotinylated peptide corresponding to the carboxyl terminus of ephrin-B3 (YYKV). From this screen, cDNAs which coded for GRIP, syntenin, and PHIP, which are all proteins with multiple PDZ domains, were isolated (Lin *et. al.*, 1999). Also, *in vitro* studies using GST fusions in COS-1 cells demonstrated that FAP-1 (a tyrosine phosphatase with a PDZ binding domain) and syntenin bind ephrin-B1 via the carboxyl terminus. Furthermore, in cells co-transfected with ephrin-B1 and syntenin, immunoprecipitation of ephrin-B1 specifically co-precipitated syntenin (Lin *et. al.*, 1999). In a yeast two hybrid screen with EphB2 (which also contains a PDZ binding motif, Figure 2) and ephrin-B1 carboxyl terminus, two cDNAs

corresponding to PICK1 and GRIP were isolated, suggesting a potential interaction (Torres *et. al.*, 1998). The specificity of these interactions was tested using GST pulldown assays, which showed that PICK1 or GRIP co-immunoprecipitated EphB2 and ephrin-B1. Conversely, expression of deleted versions of EphB2 or ephrin-B1 lacking the 3 C-terminal residues prevented an interaction with PICK1 and GRIP, indicating that the PICK1, GRIP interactions are dependent on the C-terminal residues. Taken together these data suggest that the carboxyl-terminal motif of B ephrins provides a binding site for specific PDZ domain-containing proteins, providing a mechanism for the localisation and signaling of ephrin-B ligands.

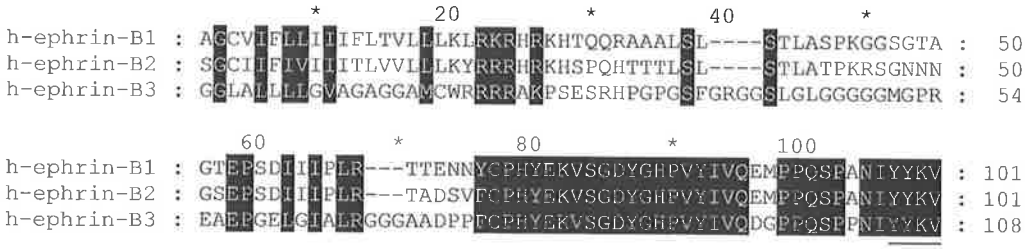


Figure 3 Clustalw protein alignment of the cytoplasmic domain of the human ephrin-Bs. Conserved residues are in black. The last 35 amino acids of the cytoplasmic tail are 98% identical for all known ephrin-B orthologs, suggesting an important function in cell signaling. The YYKV PDZ binding motif on the COOH terminus is underlined.

GPI anchored ligand signaling

The glycosylphosphatidylinositol (GPI) anchorage mechanism utilised by the *ephrin-As* is common to a large number of protein families, such as the neural cell adhesion molecules (NCAM) (Walsh and Doherty, 1996) and a number of cell surface hydrolases such as Alkaline phosphatase (Apase) and Acetylcholinesterase (Ferguson and Williams, 1988). Early research into NCAM's found that the gene anchorage mechanism could be modified to be transmembrane or GPI anchored in different splice variants of the same gene. It was suggested that if a cell needed to indicate its position, the GPI version of the NCAM would be utilised, whereas if the cell needed to propagate a signal to the host cell the transmembrane version of the NCAM with an intracellular globular domain would be utilised (Walsh and Doherty, 1991). However, this notion was challenged by work showing that GPI anchored proteins present on hematopoietic cells can activate cellular signaling responses upon binding of their natural ligands (Brown, 1993). This work by Brown (1993) showed that GPI-anchored proteins associate with tyrosine kinases of the SRC family, indicating possible signaling mechanisms for GPI proteins which were previously only thought to be able to act as ligands.

Direct evidence for ephrin-A mediated signaling came first from a null mutation in VAB-1, the only currently known Eph receptor in *C.elegans* (George *et. al.*, 1998). VAB-1 has been

shown to disrupt the movement and organization of neuroblasts that express this receptor. Furthermore, VAB-1 mutants have defects in ventral epidermis closure, indicating a role for VAB-1 signaling in cell adhesion (George *et. al.*, 1998). However, partial mutants with the VAB-1 kinase domain inactivated, exhibit a markedly less severe phenotype (George *et. al.*, 1998). One explanation for this is that the *C.elegans* Eph (VAB-1) acts as a "ligand", signaling to an ephrin "receptor". The four C.elegans ephrins (EFN 1-4), are all GPI anchored (ephrin-A). Mutations in EFN1-4 have been found to enhance the phenotype of the kinase dead VAB-1, again suggesting a signaling mechanism or "receptor" function for the *C.elegans* ephrins (George *et. al.*, 1998; Chin-Sang *et. al.*, 1999; Wang *et. al.*, 1999).

The exact mechanism by which ephrin-A transduces signals across the plasma membrane is not well characterised. However, other GPI-anchored molecules have been shown to be involved in signal transduction (reviewed in Brown, 2002), one possible mechanism for signal transduction in ephrin-A molecules could be via sequestering in membrane raft micro domains, which are known to carry high concentrations of SRC tyrosine kinases and other signaling molecules (Figure 4) (Bruckner *et. al.*, 1999). Membrane rafts are highly organised lipid domains, which provide sub compartments within the membrane itself. These rafts are enriched in sphingolipids and cholesterol. They are also known to incorporate GPI anchored proteins. Furthermore, membrane rafts have been proposed to function as platforms for the assembly of cytoplasmic and membranous signaling molecules (reviewed in Brown and London, 1998; Brown, 2002).

Caveolae-like microdomains are a type of lipid raft characterised by the presence of caveolin-1, a 22-kD protein that forms the structural basis of the caveolae domains. Biochemical analysis of mouse fibroblasts ectopically expressing ephrin-A5 showed that it is localised to caveolae-like plasma membrane microdomains. Furthermore, when activated with EphA5, ephrin-A5 is able to induce a signaling event within the microdomains. The mechanism of this signaling event is still unclear (Davy *et. al.*, 1999). However, Davy *et. al.*, (1999) showed that upon EphA5 binding an increase in tyrosine phosphorylation occurred in the ephrin-A5-expressing cell line. Also, it was shown that the Src-family kinase, Fyn, is recruited to the same caveolae like microdomains, suggesting a role for this kinase in transducing the signal downstream of ephrin-A5 (Davy *et. al.*, 1999).

Prediction and detection of GPI anchored proteins

The GPI anchorage mechanism is a common posttranslational modification employed by a large number of extracellular proteins, whereby a GPI lipid moiety is covalently attached to the C-terminal end of the protein to be anchored (Figure 5)(Cross, 1990). One of the most common ways to detect a GPI anchored protein is via enzymatic or chemical cleavage with

either phosphatidylinositol-specific phospholipase C (PI-PLC) or Nitrous acid respectively (Figure 6), followed by a subsequent assay to detect the cleavage such as a size shift on a western (Cross, 1990). However, resistance to cleavage may not eliminate the protein from potential GPI anchorage. A more recent method has been developed to assist in coping with the *in vivo* problems of detecting GPI and anchorage, but also to address the large numbers of new proteins coming from the various sequencing efforts.

Due to the accumulation of a large number of sequences with GPI anchors, it has been possible to analyse these sequences to allow a predictive tool to "learn" some of the properties of a GPI proprotein motif. Work by Eisenhaber *et. al.,* (1999) has developed a putative model for predicting GPI anchorage in protein sequence. By analysing the physical properties of certain amino acids at various regions (1-4) near the ω-site[4] Eisenhaber (1999) constructed a model of the active site of the putative transamidase complex (Eisenhaber *et. al.*, 1999a). By searching the COOH terminus of proteins for two factors a prediction of potential GPI anchorage of a protein can be made. The first factor is the presence of a hydrophobic region. The second factor is the nature of the properties of the amino acids around the ω-site (Eisenhaber *et. al.*, 2000).

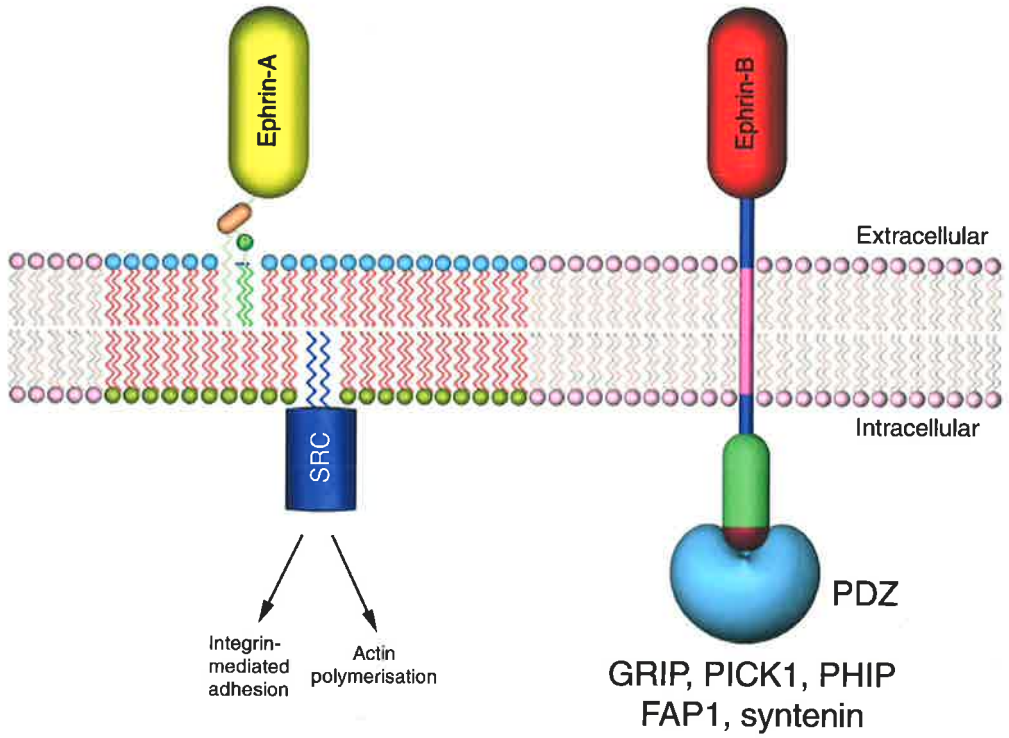


Figure 4 Proteins involved in downstream signalling of Eph/ephrin molecules. The GPI linked ephrin-As (left) are thought to associate with SRC kinases in a membrane raft, and presumably propagate intracellular signals, although the exact mechanisms are unclear. The transmembrane ephrin-Bs have been shown to cluster via PDZ adapter proteins.

[4] The site of polypeptide cleavage which allows the attachment of a GPI moiety.

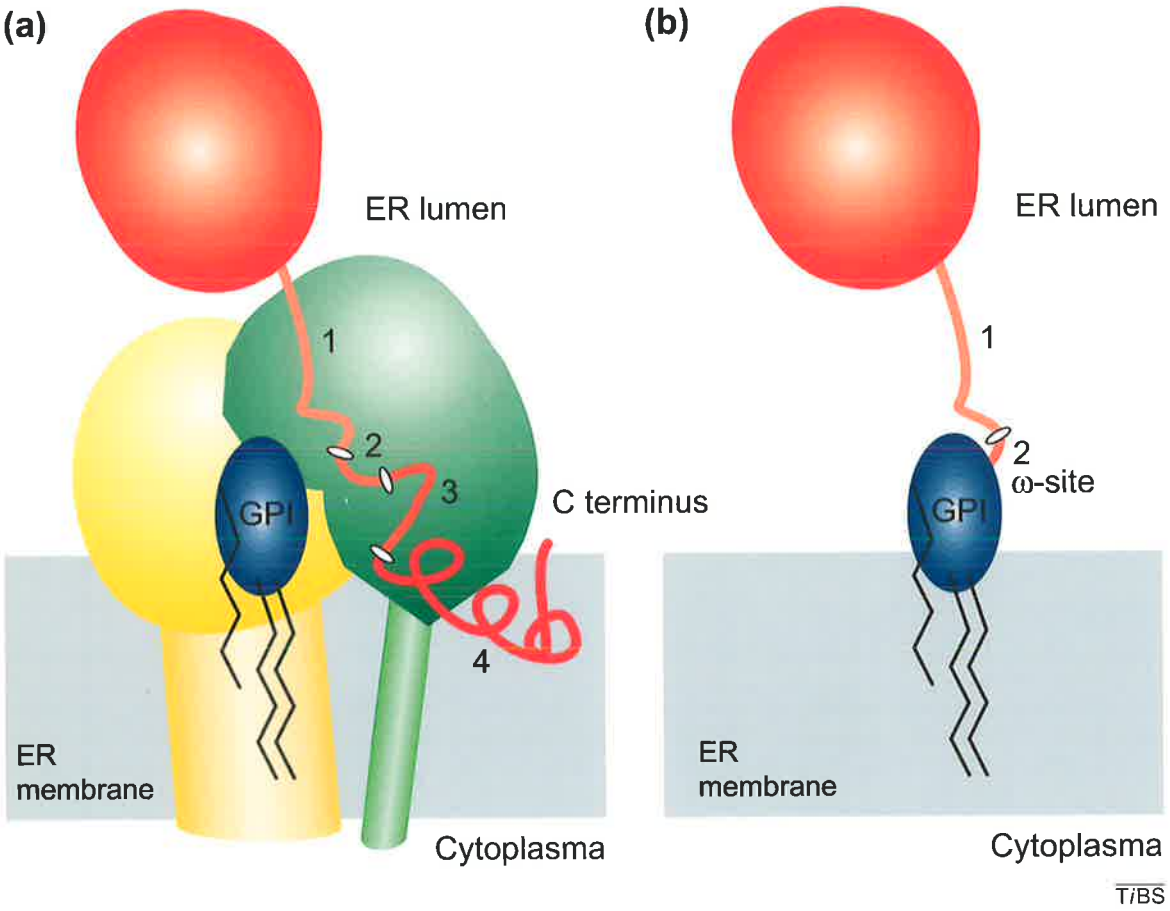


Figure 5 Glycosylphosphatidylinositol (GPI) anchoring of an extracellular protein. a) Attachment of the GPI moiety (blue) occurs after proteolytic cleavage at the ω-site of the substrate protein (red) via a transamidase complex (green) in the endoplasmic reticulum. b) After release from the transamidase complex, the protein is covalently linked to the GPI moiety and ready to be transported to the outer leaf of the cell membrane (after Eisenhaber *et. al.*, 2000).

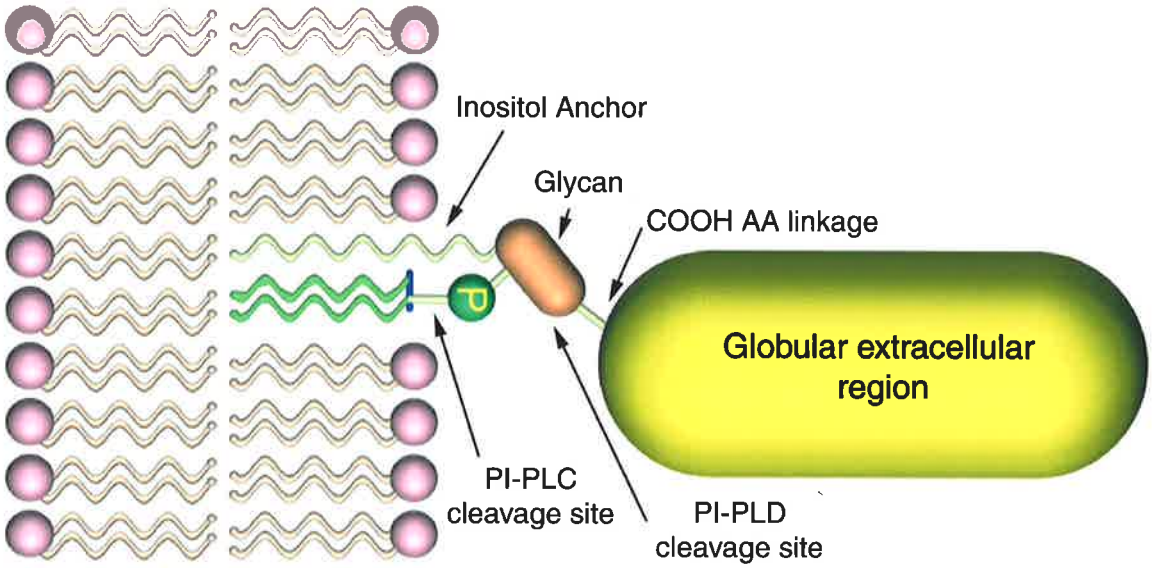


Figure 6 An illustration of the minimal core structure of a GPI anchor on the outer leaflet of cell membrane. This structure appears to be common to all currently known eukaryotic GPI anchors characterised to date. The substrate regions for the various enzymes used in detection of GPI anchored proteins *in vivo* are also indicated. PI-PLC: site for cleavage with the bacterial phosphatidylinositol-specific phospholipase C (Ferguson and Williams, 1988). PI-PLD: site for cleavage with phosphatidylinositol-specific phospholipase D (Ferguson and Williams, 1988).

Receptor-Ligand binding facilitates Repulsion?

The interaction between the Eph receptors and their respective ligands is well characterized and occurs with a high affinity, even if promiscuous between Eph family members (Gale and Yancopoulos, 1997). This fact presents something of a paradox; how does high affinity binding facilitate cell repulsion? The answer lies with work done by Hattori *et al,* (2000). This work found that upon binding of clustered EphA3-Fc, ephrin-A2 extracellular domains disappear from the cell membrane fraction, and a smaller peptide appeared in the cell lysate. However unclustered EphA2-Fc had no effect, suggesting that ephrin-A2 is cleaved from the cell membrane in a mechanism regulated by binding the clustered receptor (Hattori *et. al.*, 2000). The size of the cleaved ephrin-A2 was shown to be smaller than the size of an ephrin-A2 which is cleaved by PI-PLC, indicating that ephrin release was not due to phospholipase activity and was due to cleavage within the polypeptide (Hattori *et. al.*, 2000).

Hattori *et. al.,* (2000) performed *in situ* hybridisation analysis on the metalloproteinase Kuzbanian, which showed that *kuzbanian* is expressed widely throughout the nervous system of E18 mouse embryos, with high expression in the posterior midbrain, diminishing toward the anterior midbrain. This pattern is similar to the graded midbrain expression seen for ephrin-A2 and ephrin-A5 (Frisen *et. al.*, 1998; Feldheim *et. al.*, 2000). A role for extracellular proteases in axonal path finding was supported by the finding that mutations in the *Drosophila* gene *kuzbanian* cause axons to stall with many failing to extend through the nerve cord (Fambrough *et. al.*, 1996).

When cells were co-transfected with *kuzbanian* and *ephrin-A2*, an increase in the truncated forms of ephrin-A2 was observed (Hattori *et. al.*, 2000). Also, Hattori *et. al.,* (2000) found that ephrin-A2 forms a stable complex with Kuzbanian, involving interactions outside the cleavage domain and protease domain (Hattori *et. al.*, 2000). Hattori *et. al.,* (2000) used a MEME search/alignment algorithm to find potential metalloproteinase sites within ephrin-A2 and ephrin-B1, based on known cleavage sites from other proteins cleaved by ADAM and Kuzbanian (Hattori *et. al.*, 2000). Mutation of a potential metalloproteinase site in ephrin-A3 resulted in the delayed retraction of motor axons expressing EphA3, when exposed to substrate cells expressing the mutated ephrin-A2, suggesting that ephrin-A2 is cleaved by a metalloproteinase after receptor binding (Hattori *et. al.*, 2000). Also expression of a dominant negative form of *kuzbanian* delayed axonal retraction, directly linking this metalloproteinase to ephrin cleavage. This finding provides an important insight into the role of regulated ephrin cleavage *in vivo*, suggesting that it leads to the retraction of axons after ligand binding (Figure 7).

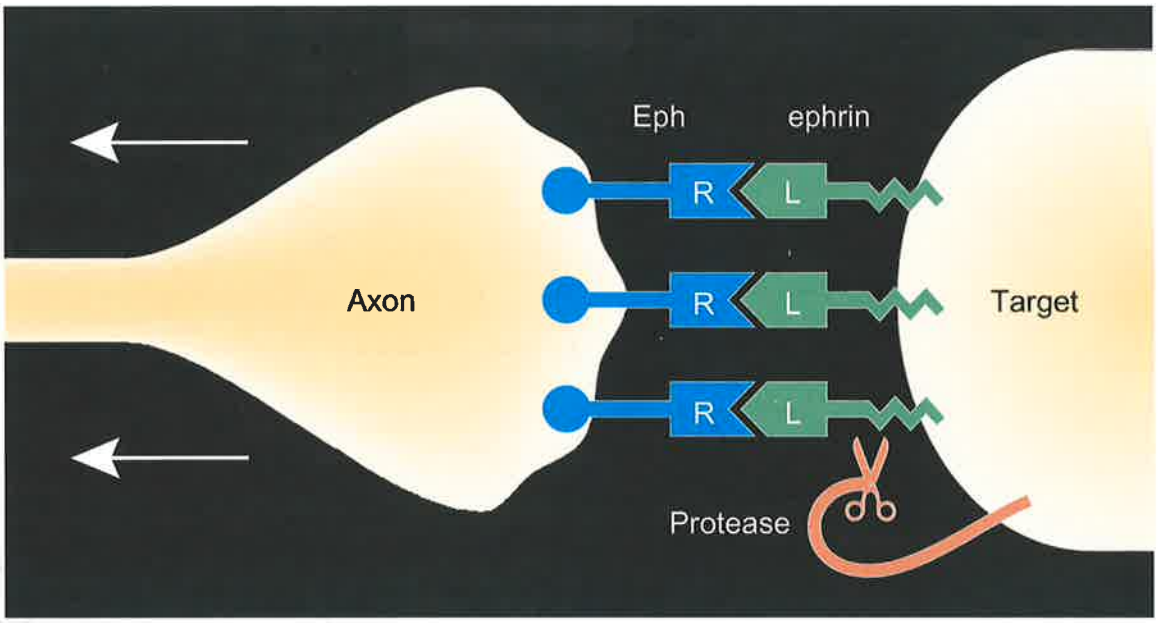


Figure 7 Model of ephrin protease cleavage following Eph receptor binding. Work by Hattori *et. al.*, (2000) showed conclusively that ephrin-A2 is cleaved by a metalloproteinase following receptor binding. There are also a number of potential metalloproteinase cleavage sites in other ephrin proteins (after Collins, 2000).

Eph/ephrin binding crystal structure domain

EphB2 crystal structure

The first Eph ligand binding domain structure to be resolved was EphB2, which was shown to be composed of 11 antiparallel β-strands folded into a jellyroll β-sandwich (Himanen *et. al.*, 1998). This ligand binding domain is unique to the Eph receptors, sharing no significant homology with other known proteins (Van Der Geer *et. al.*, 1994). The crystal structure of EphB2 also reveals a large well-ordered loop between β-strands H-I, which was shown by mutagenesis to be the ligand interaction domain (Himanen *et. al.*, 1998; Himanen *et. al.*, 2001). Furthermore this loop was also shown to be responsible for the difference in receptor-ligand specificity between EphA/ephrin-A and EphB/ephrin-B, with the EphB globular domain containing four additional residues than EphA binding domains (Himanen *et. al.*, 1998). The resolution of the EphB ligand-binding domain, gives some insight into the chemical nature of the binding promiscuity within but not between subclasses. However, the structural mechanism for the promiscuous binding of EphA4 between subclasses remains to be elucidated (Himanen *et. al.*, 1998).

Ephrin-B2 crystal structure

The ephrin-B2 receptor binding domain has been shown to consist of a β-barrel structure composed of eight strands around a hydrophobic core (Toth *et. al.*, 2001). By analysis of the solvent accessible regions highly conserved across ephrin-A and ephrin-B molecules, two potential Eph receptor interaction regions were identified on the surface of the ephrin ectodomain, which were designated Area I and Area II (Toth *et. al.*, 2001). Area I was shown to be a small concave pocket of hydrophobic residues surrounded by charged amino acids, Toth *et. al.*, (2001) suggested that Area I was involved in Eph receptor binding and is critical for distinguishing between EphA and EphB subclasses (Toth *et. al.*, 2001). Furthermore, the amino acid conservation within Area I, determines the degree of binding promiscuity within each subclass of ephrins (Toth *et. al.*, 2001). Area II facilitates the formation of ephrin-B2 ectodomain dimers, by an intricate packing of large loop structures present in each ephrin monomer, known as the G-H loop (according to the β-sheet designations) (Toth *et. al.*, 2001). This is the first indication that ligand dimerisation can occur via the interaction of ephrin ligand binding regions directly, as opposed to intracellular clustering mechanisms (Toth *et. al.*, 2001).

EphB2/ephrin-B2 complex crystal structure

In an effort to further elucidate the chemical nature of the interactions between Eph receptors and ephrin ligands, work by Himanen *et. al.,* (2001) resolved the crystal structure of an EphB2/ephrin-B2 complex (Himanen *et. al.*, 2001; Himanen and Nikolov, 2002). The structure of this interacting complex confirmed the chemical mechanisms of receptor-ligand interaction. Furthermore, the ligand-receptor dimer interfaces shown provide insight into the basis of Eph/ephrin subclass specificity. For example, ephrin-B molecules contain a bulky polar residue at position 109 (glutamine or leucine), which is hydrogen bonded to the conserved Thr 38 found in the EphB subclass of receptors (Himanen *et. al.*, 2001). Whereas in the EphA subclass of receptors position 38 is occupied by a bulky residue (glutamine or methionine), and the ephrin-A ligands have a serine or alanine at position 109 (Himanen *et. al.*, 2001). Such an arrangement would maintain favourable contact between EphA/ephrin-A and EphB/ephrin-B complexes, but not between A+B complexes (Himanen *et. al.*, 2001). Similarly, Thr 114 of the ephrin-Bs is in van der Waals contact with Val 54 of the EphBs, whereas a residue with a larger hydrophobic side chain occupies the EphA position 54 (isoleucine or methionine), which is in contact with the smaller side chain of serine at position 114 of the ephrin-As (Himanen *et. al.*, 2001). Interestingly, EphA4, which binds promiscuously to both ephrin-A and ephrin-B subclasses, is the only EphA that has a valine at position 54, providing some insight into the chemical mechanism of the binding between ephrin subclasses (Himanen *et. al.*, 2001).

The EphB2-ephrin-B2 crystal complexes showed that EphB2 also interfaces with EphB2 to form a duplex. Similarly, ephrin-B2 interfaces with ephrin-B2 to form a duplex. These self-recognition regions of the Eph and ephrins are known as tetramerisation regions, owing to the formation of a tetramer generated from two Eph-ephrin dimers joining together (Himanen *et. al.*, 2001). The absence of this tetramer interface between receptor-receptor and ligand-ligand in the crystal of unbound proteins (Himanen *et. al.*, 1998; Toth *et. al.*, 2001) suggests that these interfaces become energetically favourable after heterodimerisation between Eph-ephrin resulting in the subsequent tetramerisation of the Eph-ephrin dimers, which may help to explain the formation of the higher order clusters observed for Ephs and ephrins (Himanen *et. al.*, 2001). However, in addition to the Eph-ephrin binding domains, other regions of the Eph and ephrin are probably involved in the final positioning of bi-directional signaling complexes, including the cysteine rich linker (Lackmann *et. al.*, 1998), the intracellular SAM domain of the receptors (Stapleton *et. al.*, 1999; Thanos *et. al.*, 1999), and the C-terminal PDZ-binding domain found in Eph receptors and ephrin-B ligands (Hock *et. al.*, 1998; Torres *et. al.*, 1998; Lin *et. al.*, 1999; Lu *et. al.*, 2001).

EphActivation

Work by Davis *et. al.,* (1994) has shown that unlike other RTK ligands, soluble ephrin ligands do not activate Eph receptors, even though the soluble ligands bind to the Eph receptors (Davis *et. al.*, 1994). However, soluble ephrins, that are artificially clustered with IgG-Fc to produce dimers are able to activate receptors (Davis *et. al.*, 1994). This suggests, that membrane anchorage of ephrins may enable clustering, either before or after Eph receptor binding (Wilkinson, 2001). Furthermore, in contrast to other RTK ligands such as Kit or the epidermal growth factor receptor (EGFR) family, which can act in soluble form, Eph-ephrin interactions mediate contact dependant cell interactions (Gale and Yancopoulos, 1997; Wilkinson, 2001). Interestingly, alternative splice variants of some ephrins do generate soluble forms, although an *in vivo* role is yet to be demonstrated for these soluble ligands (Weinstein *et. al.*, 1996). The binding of soluble ephrins to their Eph receptors has been utilised to investigate the function of the receptors by antagonizing receptor activation and blocking endogenous ephrin binding. Although soluble ephrin-FC dimers can cause phosphorylation of specific Eph receptors, some Eph-ephrin combinations only result in a low level of activation. This low level of activation can be increased by the formation of anti-FC higher order clusters of ephrin-FCs, which leads to an increase in receptor stimulation (Davis *et. al.*, 1994). suggesting an important role for ephrin multimerisation in Eph receptor activation. As it lacks kinase activity EphB6 may also act as a ligand triggering unidirectional signaling via the ephrin-B proteins (Gurniak and Berg, 1996; Matsuoka *et. al.*, 1997).

Ephrin activation

The intracellular tail of all currently known ephrin-B proteins contains a 98% conserved stretch of 35 amino acids (Figure 3) raising the possibility that the ephrin-Bs may be involved in intracellular signaling themselves or act as "receptors". Biochemical evidence for this was found in cell tissue culture experiments where ephrin-B1 could transduce signals via phosphorylation, which is induced by interaction with clustered and soluble EphB2, indicating that clustered receptor binding leads to ephrin activation (Holland *et. al.*, 1996; Bruckner and Klein, 1998). Furthermore, when *EphB2*-expressing cells were mixed with *ephrin-B1* expressing cells, not only the Eph receptors were phosphorylated, but also the tyrosine residues of the ephrin. Since ephrin-B lacks a tyrosine kinase domain, their phosphorylation presumably involves recruitment of a cytoplasmic kinase (Bruckner and Klein, 1998). Src kinase can also phosphorylate ephrins *in vivo*. However, activation by Src kinases is much slower suggesting that other kinases are also involved *in vivo* (Holland *et. al.*, 1996; Bruckner *et. al.*, 1997). As discussed earlier, the ephrin-As may also transduce signals, even though they lack cytoplasmic domains. This may occur by a mechanism similar to that of other GPI

anchored proteins, such as via a co-receptor protein or by localisation to lipid rafts in the plasma membrane (Brown and London, 1998).

Key areas of Eph/ephrins in development

In the developing embryo, axonal growth cones and neural precursor cells navigate long distances (as much as several centimetres), along appropriate pathways to find their targets (Tessier-Lavigne and Goodman, 1996). Once at the target region, an appropriate "address" must be found within that region. With over 1 x 10^{12} neurons in the adult human brain, that on average make connections to over 1000 cells, the process of wiring the nervous system would seem to be an impossible task. However, with the involvement of a complex array of cell signaling molecules (both diffusible and cell anchored), all of which operate through a number of mechanisms, the task of wiring the nervous system appears less daunting (for a review of the molecular mechanisms involved including other RTK families see Tessier-Lavigne and Goodman, 1996).

It is generally accepted that there are four main types of guidance mechanisms working in concert to guide growth cones during migration; they are chemoattraction, chemorepulsion, contact attraction, and contact repulsion (reviewed in Tessier-Lavigne and Goodman, 1996). Eph-ephrin interactions facilitate contact attraction and contact repulsion, as they all occur at cell-cell interfaces. This discussion has focused on the Eph-ephrin family, and the mechanisms by which they can mediate cell communication. In order to illustrate how these mechanisms are utilised in the developing embryo to facilitate cell sorting and axon guidance, a number of specific mechanisms will now be discussed.

Functional studies of the Eph and ephrin molecules have been demonstrated to be essential in several areas of embryonic development: Eph/ephrins have been demonstrated to be involved in: (1) defining axon pathways and fasciculation via repulsive and attractive cues, (2) topographic mapping within axonal target regions, (3) maintenance of segmental boundaries within the hindbrain, also via a repulsion and attraction mechanism, (4) coordination of segmentation of the peripheral nervous system (PNS), and (5) help define the vasculature during angiogenesis.

Fasciculation

In vitro studies of cortical neurons cultured on astrocytes have implicated EphA5 and ephrin-A5 in axonal fasciculation. When cortical neurons are cultured on astrocytes, the differentiation of the neurons reaches an advanced stage similar to that seen in late brain development, when axonal tracts and fibres are laid down, characterised by the formation of large bundles of axons (Caras, 1997). Work by Winslow *et. al.,* (1995) showed that when a

soluble EphA5-IgG Fc (which will bind endogenous ephrin-As) was added to the co-cultures of cortical neurons and astrocytes, significantly fewer axonal bundles were formed, suggesting that EphA5 activation is required for axonal fasciculation (Winslow *et. al.*, 1995). Soluble ephrin-A5, which can bind EphA5 but not cause phosphorylation, also blocks axon fasciculation in cortical neuron/astrocyte cultures. Also, localisation experiments showed that *EphA5* is expressed on cortical neuron axons, while ephrin-A5 is expressed on the surface of astrocytes (Winslow *et. al.*, 1995). These data suggest that direct contact between *EphA5* and *ephrin-A5* causes axon fasciculation. One suggestion for the mechanism behind this is that EphA5 may regulate the expression of adhesion molecules, such as fasciclin or NGCAM, on the axon surface (Caras, 1997). Support for this contention comes from the finding that in *Xenopus* EphA4 mutants there is a disruption in cell adhesion, possibly by regulating the expression of cadherins (Winning *et. al.*, 1996). Alternatively, ephrin-A5 may be acting in a repulsive manner, similar to its role in the retinotectal system (discussed later) during chick development (Drescher *et. al.*, 1995). Ephrin-A5 expressed on the astrocytes may act to repel the EphA5-expressing cortical neurons, forcing them to form bundles (Caras, 1997). This is supported by growth cone collapse assays, which show that ephrin-A5-IgG can induce growth cone collapse of cortical neurons, suggesting that ephrin-A5 is a repellent axon guidance molecule for cortical neurons (Meima *et. al.*, 1997). Overall these data suggest a potential link between axon guidance and axon fasciculation, both of which are possibly caused by repulsive factors such as ephrin-A5 (Caras, 1997).

Axon pathfinding- Roles in commissural tract formation

Targeted mutations of genes in mice provide *in vivo* evidence that Eph/ephrins are involved in directing axon migration in the central nervous system (Henkemeyer *et. al.*, 1996). EphB2 knockout mice have a marked reduction in the formation of the posterior tract of the anterior commissure (AcP), which is a major interhemispheric connection between the two temporal lobes of the cerebral cortex (Henkemeyer *et. al.*, 1996). EphB2 is expressed in the embryonic forebrain ventral to the commissural axons at stage E14.5, which corresponds to the time of axon migration and pathfinding in mice. In mice mutants that have had the intracellular kinase domain of EphB2 replaced by a *lacZ* gene, the anterior commissure is normal, indicating that EphB2 is acting as a ligand. Furthermore, antibody stains using the extracellular domain of ephrin-B1 found that it was expressed on the axons of the anterior commissure. These results suggest that EphB2 plays a unique role in the guidance of cortical axons in the anterior commissure via binding with its cognate ligand ephrin-B1, triggering a repulsion response and prevents axons from moving ventrally into the forebrain (Henkemeyer *et. al.*, 1996).

Work by Orioli *et. al.,* (1996) showed that EphB3 knockout mice also have defects in another commissural tract, the corpus callosum, with EphB3-/- mice failing to form the corpus callosum (callosal agenesis). *EphB3* expression occurs in tissue adjacent to the developing commissure. Axons in EphB3-/- mice project out in the correct orientation towards the midline, but they fail to cross the midline and instead form large bundles, known as Probst's bundles (Orioli *et. al.*, 1996). This suggests that EphB3 acts a ligand to play a repulsive role in guiding callosal axons across the midline. Mice double mutants of EphB2 and EphB3 have even more severe phenotypes in both the corpus callosum and the anterior commissure, which can be explained by the overlapping expression patterns of EphB2 and EphB3. This provides evidence that EphB2 and EphB3 have cooperative roles in axon repulsion, with EphB2 responsible for the anterior commissure and EphB3 responsible for the corpus callosum (Orioli *et. al.*, 1996).

EphA4 deficient mice exhibit a loss of co-ordination and a characteristic kangaroo like hop, resulting from a disruption of the correct assembly of the corticospinal tract (CST). Spinal cord sections showed that the EphA4 mutants have a shallower dorsal funiculus, through which the CST descends in rodent spinal cords. EphA4 null mutant mice also exhibit a loss of the anterior commissure similar to that found in the EphB2 and EphB3 mice knockouts (Dottori *et. al.*, 1998). Unexpectedly, EphA4 is not expressed on the CST axons, however levels of expression are higher within the intermediate and ventral regions of the spinal cord. Furthermore, ephrin-B3, which has a high binding affinity for EphA4, is expressed in the sensorimotor cortex, at E 18.5, suggesting that ephrin-B3 is present on CST axons at the time of growth through the brain and spinal cord (Dottori *et. al.*, 1998). These experiments suggest a role for EphA4 and ephrin-B3 in the guidance of CST axons through the central nervous system. Further analysis by Kullander *et. al.,* (2001) used mouse knockins to introduce a kinase-reduced and a kinase-dead form of the EphA4. These studies show that the kinase domain of EphA4 is required for the formation of the CST (EphA4 as a receptor), whereas the guidance of AC axons occurs without a functional EphA4 kinase domain (EphA4 as a ligand) (Kullander *et. al.*, 2001).

Segmentation of the hindbrain

The hindbrain is subdivided into repeated morphological units termed rhombomeres, which underlie a segmental organization of cranial nerves and neural crest cells that migrate into the branchial arches and form the face and its final nervous system innervation. The segmental boundaries are established by periodic expression of *krox-20*, and by *hox* genes, which confer anteroposterior gradients (Trainor and Krumlauf, 2000; Xu *et. al.*, 2000; Trainor and Krumlauf, 2001). Studies in chick embryos have demonstrated that after rhombomere

segmentation, cell movement between adjacent segments is restricted, subdividing the hindbrain into five segments corresponding to rhombomeres r2-r6 (Fraser *et. al.*, 1990). The mechanism of restricting cell intermingling between rhombomeres may be via cell adhesion molecules, which underlie differential adhesion of cells in odd versus even numbered rhombomeres. However adhesion proteins with alternating expression had not been previously discovered (Xu *et. al.*, 2000). Thus although the transcriptional molecular mechanism was largely characterised the effector proteins underlying rhombomere border formation were unknown. Initial evidence for the involvement of the Eph receptors and their ephrin ligands in the restriction of cell intermingling between rhombomeres was suggested by their expression patterns. *EphA4*, *EphB2* and *EphB3* are expressed at high levels in rhombomeres r3 and r5 (Nieto *et. al.*, 1992; Becker *et. al.*, 1994; Henkemeyer *et. al.*, 1994). The corresponding ligands *ephrin-B1, ephrin-B2,* and *ephrin-B3* are expressed at high levels in r2/r4/r6 (Bergemann *et. al.*, 1995; Flenniken *et. al.*, 1996; Gale *et. al.*, 1996a).

Work by Xu *et. al.,* (1995 & 1996) directly showed the involvement of the Eph and ephrins in segmental tissue patterning, by expression of a dominant negative form of EphA4 lacking kinase activity. A disruption of cell boundaries was observed in *Danio rerio* and *Xenopus* embryos expressing the dominant negative form of EphA4. Here, cells usually located in r3/r5 were found to be present in r2/r4/r6, and in some embryos, a complete fusion of r3 and r5 territories was observed (Xu *et. al.*, 1995). This disruption was rescued by the expression of a full length EphA4. These experiments suggest that the blocking of Eph receptor activity caused some cells with r2/r4/r6 identity to switch to r3/r5 identity, or blocked normal switches in identity when cells intermingle between presumptive odd and even segments. Alternatively, there could be a disruption of the normal restriction of intermingling between odd and even segments causing cells to mix freely between segments (Xu *et. al.*, 2000). In order to distinguish between these possibilities, Xu *et. al.,* (1999) analysed whether mosaic expression of ephrin-B2, which activates EphA4, EphB2 ,and EphB3, could lead to changes in the identity or movement of cells within r3/r5 (Xu *et. al.*, 1999). Full-length *ephrin-B2* and *lacZ* genes were co-injected into one cell of eight cell stage zebrafish embryos to give mosaic expression. The result of this is that EphA4 and EphB receptors are activated in r3/r5. Ectopic *ephrin-B2/lacZ*-expressing cells were restricted to the boundaries of r3/r5, whereas *ephrin-B2/lacZ*-expressing cells in r2/r4/r6 were scattered through the rhombomere. Furthermore, expression patterns marking r2/r3 were unaltered in the *ephrin-B2/lacZ* injected embryos, which suggested that mosaic expression of *ephrin-B2* does not alter the identity of the expressing or adjacent cells. A similar result was obtained when *ephrin-B2* without the intracellular domain, which can activate Eph receptors but not transduce a signal within the cell, was introduced in the same manner (Xu *et. al.*, 1999). This suggested that the mosaic

activation of Eph receptors facilitates a cell sorting mechanism. Similarly, expression of a truncated EphA4, which can activate ephrin-B but not transduce a signal, causes expressing cells to sort at the boundaries of r2/r4/r6, which endogenously express *ephrin-B*, whereas ectopic *EphA4* cells in r3/r5 are restricted to the central regions of r3/r5 (as opposed to the rhombomere boundary, shown with *ephrin-B2* cells) (Xu *et. al.*, 1999). These experiments indicate that mosaic activation of either Eph receptors or ephrin-B ligands can facilitate cell sorting, without altering cell identity. Furthermore, unidirectional signaling (i.e. truncated Eph or ephrin) causes cells to move towards the rhombomere boundaries. A possible mechanism for this is that the interactions of endogenous Ephs and ephrin at rhombomere boundaries create a region of lower cell affinities than those occurring within the rhombomeres. In a repulsive mechanism the cells expressing ligand (i.e. the truncated Eph or ephrin) have a lower affinity for their neighbours and thus sort to create the rhombomere boundaries (Xu *et. al.*, 1999). However, bi-directional signaling may still play a role at the interfaces of Eph-ephrin expression, as suggested by the truncated EphA4 experiments described earlier (Xu *et. al.*, 1995).

To establish the role of bi-directional activation in the establishment of rhombomere boundaries work by Mellitzer *et. al.,* (1999) established an *in vitro* assay where an Eph receptor and ephrin ligand are expressed in adjacent cell populations. This was achieved by co-injecting one cell stage zebrafish embryos with a fluorescent tracer and either an *Eph* or *ephrin* gene. The injected embryo cell cap is then dissected at the 1000 cell stage. Upon juxtaposition, the two cell populations fuse to form a "fishball", within which cell lineage can by traced using the fluorescent die (Mellitzer *et. al.*, 1999). After overnight culture the fishball can be assayed for cell intermingling using fluorescent microscopy. In control embryos the two cell populations were shown to mix, however cell populations expressing *EphA4/ephrin-B2*, *EphB2/ephrin-B2*, or *EphB2/ephrin-B1* (which are all known to interact, Figure 1) do not mix. Interestingly, *EphA4/ephrin-B1* (which do not interact, Figure 1) cell populations do mix in a similar manner to the control cell populations, suggesting a biological significance to EphA4s affinity to the ephrin-A and ephrin-B subclasses. Furthermore, when cell populations expressing a truncated form of either the Eph receptor or ephrin ligand (unidirectional signaling) are cultured, significant cell mixing occurs in a similar fashion to the control cell populations. Suggesting that bi-directional signaling restricts intermingling of adjacent cell populations (Mellitzer *et. al.*, 1999). These observations suggest a model in which Eph receptors and ephrin-B ligands each trigger a cell repulsion or de-adhesion response, such that bi-directional signaling maintains a restriction on two cell populations intermingling (Mellitzer *et. al.*, 1999).

Eph and ephrins have also been shown to be involved in the restriction of gap junctions (Mellitzer *et. al.*, 1999). Gap junctions are intercellular channels known as connexons, formed by connexins, which allow the passage of small proteins between cells; gap junctions can be detected by diffusion of Lucifer yellow through these channels (Mellitzer *et. al.*, 2000). Work by Mellitzer *et. al.,* (1999) used the fishball assay to show that Eph and ephrin expression will prevent the formation of gap junctions. When two cell populations with Lucifer yellow and rhodamine dextran die are allowed to intermingle, a mixing of die between cell populations indicates the presence of gap junctions. However, when *EphA2* or *EphB2* and *ephrin-B2* are expressed in cell populations, no Lucifer yellow mixing is observed, indicating a lack of Gap junction formation (Figure 8a). Furthermore, when truncated forms of EphA2, EphB2, or ephrin-B2 were allowed to mix, no gap junctions were formed, regardless of cell intermingling (Mellitzer *et. al.*, 1999). This suggests that unidirectional signaling in Eph/ephrin interactions is responsible for the blocking of gap junction formation, underling a restriction in cell-cell communication via gap junctions (Figure 8b) (Mellitzer *et. al.*, 1999). These data suggest that in hindbrain, the restriction of cell communication and intermingling between rhombomeres at the boundaries is facilitated by endogenous expression of Eph in r3/r5 and ephrin in r2/24/r6, with communication acting in a bi-directional manner (Mellitzer *et. al.*, 1999).

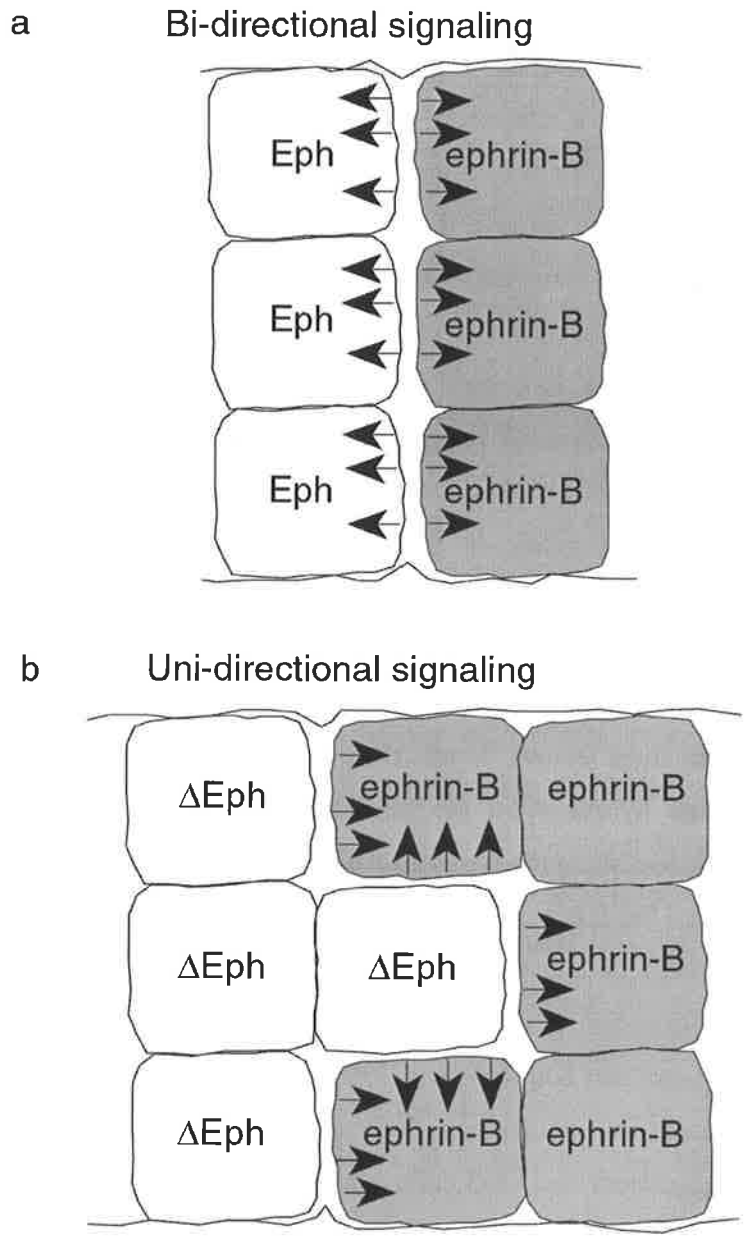


Figure 8 Cell intermingling and communication in rhombomeres is restricted via a repulsive mechanism facilitated by Eph and ephrin expression. (a) Bi-directional signaling between Eph receptors and ephrin ligands leads to repulsion between two cells restricting cell contact and communication (gap junctions), also resulting in a restriction of cell intermingling. (b) In a unidirectional signaling mechanism cell populations are still repulsed, indicating that repulsion of one cell population is sufficient to inhibit gap junction formation, however the ligand (truncated) expressing cell can now invade receptor (intracellular signaling) cell territory, resulting in intermingling, which occurs if the ephrin is truncated (after Mellitzer *et. al.*, 1999).

Roles of Eph/ephrins in establishing the anterior-posterior axis in the chick retinotectal map

The principal of topographic mapping refers to the spatial organization of neuronal connections being precisely conserved onto various neural fields within a neural system. The establishment of topographic maps is thought to involve long-range cues to guide axons to their target areas and local cues to "fine tune" the target region. The best-characterised topographic map is the chicken retinotectal system, where visual information is sent from the retina to the optic tectum via the retinal ganglion axons (reviewed in Goodhill and Richards, 1999). These retinal ganglion axons are spatially highly organised, so that the nasal axons terminate posteriorly and the temporal axons terminate anteriorly in the optic tectum (Figure 9a). The prevailing hypothesis for the implementation of a topographic map of this kind, first formulated by Sperry (1963), is known as the chemoaffinity theory. This classic model proposed that molecules expressed in gradients in the tectum would provide positional information, to incoming retinal axons with corresponding receptors also expressed in a gradient to allow correct targeting in the tectum (Sperry, 1963). Several studies have implicated the Eph/ephrin family in providing the gradient information first proposed by Sperry (1963).

Ephrin-A5 was first isolated in a screen for GPI anchored proteins that were expressed in a gradient fashion in the chick tectum. Northern analysis showed that *ephrin-A5* is expressed at E7-E13, while *in situ* hybridisation analysis showed that *ephrin-A5* is expressed in an anterior-posterior gradient along the posterior part of the tectum at the time of retinal ganglion innervation (E9, Figure 9b) (Drescher *et. al.*, 1995). By utilising two different *in vitro* assays Drescher *et. al.,* (1995) demonstrated the possible involvement of ephrin-A5 in axonal guidance within the formation of the retinotectal map. Firstly, in collapse assay experiments (Cox *et. al.*, 1990; Raper and Kapfhammer, 1990), growth cones from retinal explants were repelled when exposed to membranes derived from *ephrin-A5* expressing COS cells. Conversely, control untransfected COS cells were unable to repel retinal axons (Drescher *et. al.*, 1995). Secondly, stripe assays were utilised to test the ability of *ephrin-A5* to guide retinal axons *in vitro*. In these experiments alternating stripes of membranes derived from *ephrin-A5* COS cells and mock-transfected cells are prepared on laminin-coated filters. The laminin is used to encourage axonal outgrowth. When retinal explants were allowed to grow on these stripes both nasal and temporal axons were repelled from the *ephrin-A5* stripes, and extended solely on the mock-transfected stripes (Drescher *et. al.*, 1995). Furthermore, *ephrin-A5* induces growth cone collapse and repulsion of both nasal and temporal retinal ganglion cells (Drescher *et. al.*, 1995). However, *ephrin-A5* was later shown to have a

stronger repulsion for temporal axons than nasal axons, in a manner directly related to the concentration of *ephrin-A5*, which again was demonstrated by a stripe assay approach (Monschau *et. al.*, 1997). Taken together, this work suggested that ephrin-A5 is a candidate for axon guidance in the formation of retinotectal connections.

In a separate study, *ephrin-A2* was also found to be expressed in an anterior to posterior pattern within the optic tectum (Cheng *et. al.*, 1995; Nakamoto *et. al.*, 1996). It was also found that *ephrin-A2* has a bimodal effect which could repel temporal but not nasal axons *in vivo*, suggesting that *ephrin-A2* determines nasal verses temporal specificity (Nakamoto *et. al.*, 1996). Also, *EphA3,* which is a receptor to *ephrin-A2,* is expressed in the retina in a nasal to temporal gradient, consistent with the proposal that the *ephrin-A* gradients mediate the formation of a topographical map. When *ephrin-A2* was expressed ectopically in the tectum, retinal axons avoid ectopic *ephrin-A2* patches and map to abnormally anterior positions, providing direct evidence for the involvement of *ephrin-A2* in retinotectal map formation (Nakamoto *et. al.*, 1996).

Both *ephrin-A2* and *ephrin-A5* are expressed in a high posterior to a low anterior gradient. However, *ephrin-A5* appears to be in a steeper gradient and is expressed at high levels immediately posterior to the tectum (Monschau *et. al.*, 1997). It has been proposed that retinotectal mapping occurs via a difference in sensitivity of retinal axons to the repulsive cues of *ephrin-A2* and *ephrin-A5* in the tectum due to the graded expression of *EphA3* in the retinal axons (Cheng *et. al.*, 1995; Drescher *et. al.*, 1995; Nakamoto *et. al.*, 1996; reviewed in Drescher *et. al.*, 1997; Monschau *et. al.*, 1997; Feldheim *et. al.*, 1998; Goodhill and Richards, 1999; Knoll and Drescher, 2002). Therefore, nasal retinal ganglion axons with low EphA3 expression will project to posterior targets with a high level of *ephrin-A2* and *ephrin-A5* expression. Conversely, temporal retinal ganglion axons with a high level of *EphA3* expression will be more sensitive to ephrin-A mediated repulsion and project to the anterior portion of the tectum (Figure 9).

So why are both *ephrin-A2* and *ephrin-A5* needed to establish this gradient? Both *ephrin-A2* and *ephrin-A5* have distinct effects on nasal versus temporal axons, which may be due to differences in binding affinities to the *Eph* receptor expressed by the retinal ganglion axons, therefore it has been suggested that the ephrin-A2 and ephrin-A5 gradients act additively to establish a gradient of repulsion along the optic tectum (Monschau *et. al.*, 1997).

A more recent study found a new ephrin (ephrin-A6), in a screen for ligands to EphA4 that are expressed in the chick retina. *Ephrin-A6* has been found to be expressed in a high nasal to low temporal gradient in the retina, similar to *ephrin-A2* and *ephrin-A5*, at stages E6-E8 at the time that retinal axons reach their tectal targets (Menzel *et. al.*, 2001). Furthermore, ephrin-A6 has been shown to have an affinity for EphA4 and can elicit a response in growth

cone collapse assays of retinal explants, suggesting a possible role for ephrin-A6 in the formation of the retinotectal map of chick.

Another mechanism of graded sensitivity of the retinal axons to tectal *ephrin-A* expression has also been characterized. *EphA4* is uniformly expressed in the retina, along with a high nasal to low temporal graded expression of *ephrin-A2* and *ephrin-A5* in the retina (Connor *et. al.*, 1998; Hornberger *et. al.*, 1999). This overlap in expression causes phosphorylation of EphA4 in nasal axons (high *ephrin-A5*), which desensitises axons to repulsion by ephrin-A ligands. Indeed, a graded phosphorylation of EphA4 has been observed, with low levels of phosphorylation in temporal axons and high levels in nasal axons (Connor *et. al.*, 1998). Furthermore, EphA4 has been shown to be required for repulsive guidance of nasal (high *ephrin-A5*) but not of temporal axons (low *ephrin-A5*), suggesting that the overlapping expression is important for the graded sensitivity of nasal retinal axons (Walkenhorst *et. al.*, 2000).

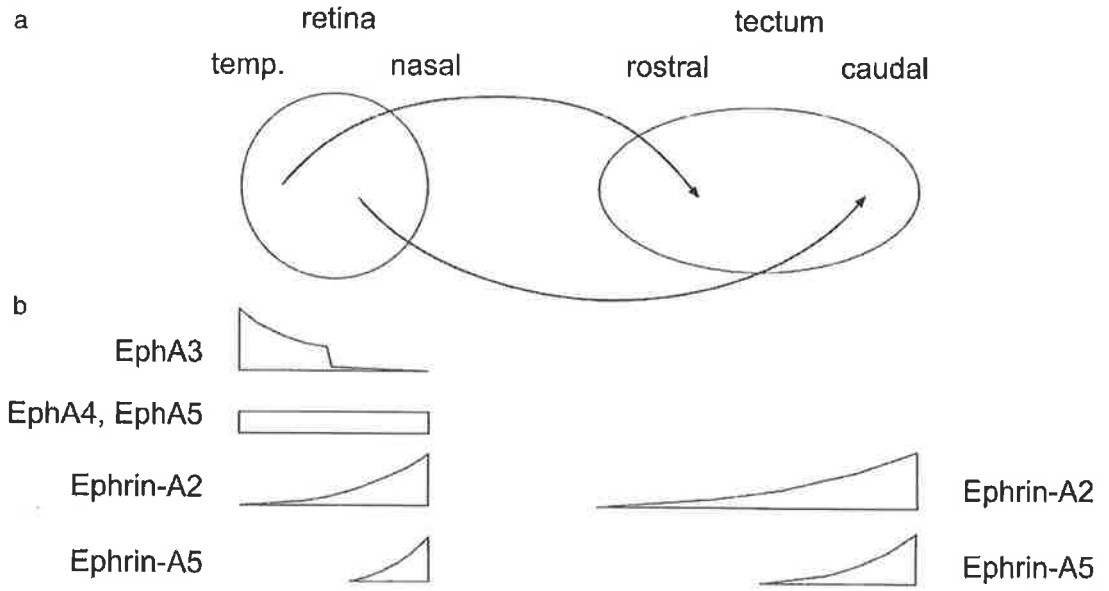


Figure 9 Eph and ephrin gradients in chick retinotectal mapping. (a) Nasal and Temporal axons in the retina project to the posterior (caudal) and anterior (rostral) of the optic tectum respectively. (b) This is facilitated by a gradient of expression from high to low, of EphA3 in temporal to nasal axons in the retina. There is a corresponding low to high ephrin-A2 and ephrin-A5 expression gradient in the optic tectum target region. Nasal axons with low receptor expression are less sensitive to ephrin-A repulsion and project further in to the optic tectum, whereas temporal retinal axons are more sensitive and do not project as deep into the optic tectum. Also, EphA4 and EphA5 are expressed at uniform levels throughout the retina; studies have shown that EphA4 becomes increasingly phosphorylated in a temporal (low ephrin-A) to nasal (high ephrin-A) gradient. (after Hornberger *et. al.*, 1999).

Segmentation in peripheral nervous system (PNS)

Studies have shown that during development of the PNS, neural crest (NC) cells migrate through the rostral, but not the caudal half of developing somites, suggesting that the caudal half contains an inhibitory or repulsive guidance cue. Using ephrin-B1-Fc proteins, it was shown that *EphB* receptors are expressed in the rostral portion of mature somites in stage 15 chick embryos, corresponding to the onset of neural crest migration. *In situ* analysis subsequently showed that *EphB3* is expressed on neural crest cells during their migration through the rostral half of the somite. Similar experiments using EphB2-Fc showed that ligands were present in the caudal somite. *In situ* hybridisation experiments showed these to be *ephrin-B1* (Krull *et. al.*, 1997). *In vitro* stripe assays show that boundaries of ephrin-B1 and ephrin-B2 proteins repel explants of neural crest cells (Krull *et. al.*, 1997; Wang and Anderson, 1997). The addition of soluble ephrin-B1-Fc to developing embryo trunks, which results in ectopic activation of EphB3 on the neural crest cells, blocks this inhibition, allowing neural crest cells to migrate through the caudal half of the somites. This suggests, that Eph receptors and ephrin ligands are involved in interactions between neural crest cells and sclerotome cells (Krull *et. al.*, 1997). However, mutation of the mouse EphB2 (which is highly expressed on neural crest cells) fails to disrupt neural crest cells, suggesting that other Eph receptors are also expressed in neural crest cell migration (Wang and Anderson, 1997).

Spinal motor axons also move through the rostral half of the somite and through the dermamyotome, in a similar manner to that of neural crest cells. Ephrin-B1 is also expressed in the caudal somite during motor axon navigation through the rostral somite. *In vitro* stripe assays demonstrated that axons growing from spinal cord explants avoid stripes containing ephrin-B1 or ephrin-B2, suggesting that ephrin-B ligands, like ephrin-A ligands, could also have a repulsive role in axon migration (Wang and Anderson, 1997). Work by Koblar *et. al.*, (2000) investigated the potential of ephrin-B1-Fc to perturb motor axon outgrowth in the developing neural trunk. Although neural crest cells migrated aberrantly as seen previously, motor axons continue to be restricted to the rostral half of the somite indicating that the interaction between ephrin-B1 and EphB2 alone is not required for patterning spinal motor axon segmentation (Koblar *et. al.*, 2000).

Roles in venous/arterial specification during Angiogenesis

In the vertebrate system, arteries and veins are defined by the direction of blood and by anatomical and functional differences (Wang *et. al.*, 1998). The intricate network of arteries and veins form from a remodelling of a primitive vascular network of endothelial cells, a process that is termed vasculogenesis (Mellitzer *et. al.*, 2000). Subsequently, the network of thin tubules, that results from vasculogenesis then undergoes a succession of morphogenetic events involving sprouting, splitting, and remodelling, collectively called angiogenesis (Wang *et. al.*, 1998).

The Ephs and ephrins have been implicated in angiogenesis. EphrinA1 has been shown to stimulate angiogenesis via a tumour necrosis factor-α (TNF-α) alpha-dependant pathway (Pandey *et. al.*, 1995b). Also, ephrin-B1 has been shown to induce endothelial cells to form capillary like networks *in vitro* (Daniel *et. al.*, 1996; Stein *et. al.*, 1998). Work by Wang *et. al.,* (1998) shows that there is a molecular distinction between veins and arteries in the earliest stages of angiogenesis, which is demarcated by the expression of *ephrin-B2* on arterial cells and its cognate receptor on *EphB4* on venous cells. Targeted disruption of the ephrin-B2 gene in the mouse does not disrupt the formation of the primitive vascular network, but the remodelling of veins and arteries in angiogenesis fails, suggesting a bi-directional signaling mechanism is in operation, where ephrin-B2 transduces the signal required for the angiogenesis of arteries and EphB4 transduces the signal for veins to develop (Wang *et. al.*, 1998). A targeted mutation in EphB4 results in the same phenotype as the ephrin-B2 mouse mutant (Gerety *et. al.*, 1999). Taken together the data from the ephrin-B2 and EphB4 mutant mice clearly indicate that EphB4/ephrin-B2 interactions are required to allow vascular endothelial cells to undergo angiogenesis.

Aims of this study

The purpose of this research project was to isolate and characterise ephrin ligands from *Drosophila melanogaster* and to analyse the structure and function of any *ephrin* found in *Drosophila* development. Furthermore, the potential of *ephrin-B1* as a causative gene in the human condition Aicardi's syndrome was investigated.

Chapter three outlines approaches to isolate *Drosophila* ephrins using a unique PCR methodology, which utilises one conserved gene region to isolate new gene family members. This chapter also shows how *d-ephrin* was subsequently obtained from the Berkley *Drosophila* Genome Project (BDGP). Also, bioinformatic analysis of *d-ephrin* was performed to elucidate the structural organization of *d-ephrin*.

Chapter four details an *in vitro* cell assay, set up to determine the affinity of *d-ephrin* for the *Drosophila* Eph receptor *dek*. Spatial and temporal expression analysis was performed using *in situ* hybridisation. Also, misexpression studies are used to determine the role of d-ephrin *in vivo*.

Chapter five shows a detailed evolutionary study of the phylogenetic relationship of all currently known ephrin orthologs (45 across mouse, human, chicken, rat, zebrafish, nematode and fruit fly). This was performed in an effort to clarify the current nomenclature of the ephrin family and to place the invertebrate ephrins (1 from *Drosophila melanogaster* and 4 from *Caenorhabditis elegans*) within a specific subclass of ephrin A or B.

Chapter six describes the analysis of *h-ephrin-B1* as a potential candidate in human Aicardi's syndrome. In an attempt to determine if a mutation in human ephrin-B1 was the principle disease gene for Aicardi's syndrome in humans, six patients known to carry the genetic defect causing Aicardi's syndrome were analysed.

Chapter 2: Materials and methods

Materials and methods

Abbreviations

Abbreviations used are as described in "Instructions to Authors", *Journal of Biochemistry* (1978) 169, 1-27. In addition;

aa	amino acids
AP	alkaline phosphatase
APS	ammonium persulphate
bisacrylamide	N,N`-methylene-bisacrylamide
Blotto	5% skim milk powder in PBT
BDGP	Berkley *Drosophila* Genome Project
BSA	bovine serum albumen
CIP	calf intestinal phosphatase
DAB	diaminobenzidine
ddH_2O	double distilled water (i.e. distilled and Millipore filtered)
DIC	differential inference contrast
ECL	enhanced chemiluminescense
EDTA	ethylenediaminetetraacetic acid
EtBR	ethidium bromide
HRP	horse-radish peroxidase
IPTG	isopropyl b-D-thiogalactopryanoside
kb	kilobases
NT	nucleotides
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
PBS	phosphate buffered saline
RO water	reverse osmosis water
r.p.m.	revolutions per minute
RTK	receptor tyrosine kinase
RT	room temperature
SDS	sodium dodecyl sulphate
TEMED	N,N,N`,-tetra ethylenediamine
UAS	upstream activating sequence
UV	ultraviolet
WT	wild type
X-gal	5-bromo-4-4chloro-3-indolyl-β-D-galactoside

Materials

Enzymes

Enzymes were obtained from the following sources

Restriction enzymes	New England Biolabs
Ligases	New England Biolabs
CIP	Roche
T7, SP6, T3 polymerases	Roche
Taq Polymerases	Gene works

Kits

Nucleic Acid Purification Kits

Qiagen - QIAquick Gel Extraction Kit (50)	Cat # 28704
Qiagen - QIAquick PCR Purification Kit (50)	Cat # 28104
Qiagen - QIAprep Spin Miniprep Kit (50)	Cat # 27104
Qiagen - QIAfilter Plasmid Midi Kit (25)	Cat # 12243

Poly A Purification Kits

Qiagen - Oligotex Direct mRNA Mini Kit (12)	Cat # 72022
Promega - PolyATtract mRNA Isolation System III	Cat # Z5300

Reverse Transcription

Qiagen - Sensiscript RT Kit (50)	Cat # 205211
Gibco BRL - Superscript II	Cat # 18064-014

Antibodies

Primary Antibodies

Qiagen - Anti·His Antibody Selector Kit	Cat # 34698
Cell Signaling - Myc-Tag 9B11 Monoclonal Antibody	Cat # 2276
Cell Signaling - His-Tag Polyclonal Antibody	Cat # 2365
Cell Signaling - HA-Tag 262K Monoclonal Antibody	Cat # 2362
Rabbit - α - cEphA4	(gift from David Wilkinson)
mouse - α - 22C10	was developed by (Benzer, S)[5]
mouse - α - BP 102.	was developed by (Goodman, C)[5]

[5] **This antibody was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biological Sciences, Iowa City, IA 52242.**

Secondary Antibodies

α-mouse-CY3 conjugate	Jackson Laboratories
α-rabbit-HRP	Jackson Laboratories
α-mouse-HRP	Jackson Laboratories
α-rabbit-AP	Jackson Laboratories
α-mouse-AP	Jackson Laboratories

Radiolabelled nucleotides

^{32}P-dATP	Geneworks

Antibiotics

Ampicillin	Sigma

Molecular Weight Markers

DNA: 1KB Plus Ladder	Promega
Protein: Standard	Gibco BRL

Bacterial Strains

DH5α: F$^-$,f80, *lac*ZΔM15, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17, (r_{K-}, M_K+), *sup*E44, *rel*A1, *deo*R, Δ(*lac*ZYA-*arg*F) U169 (Hanahan, 1983)

JM109 electro competent cells Promega

***Drosophila* strains**

w^{1118} were obtained from laboratory stocks (Lindsay and Zimm, 1992).

w; +/*cyo* ; *xb3*/*tmb6* balancer stocks were obtained from laboratory stocks (Lindsay and Zimm, 1992).

Plasmids

pGEM T easy	Promega
pSK+	Stratagene
pUAST	gift from Dr Andrea Brand
Δ2-3 Transposase	gift from Dr Andrea Brand

Buffers and solutions

25X TAE

121g Trizma base

9.3g Na_2EDTA

pH to 8.2 with ~26ml glacial acetic acid

q.s. to 1 litre, store at 21°C.

Agarose gel loading buffer 10X

2.5ml glycerol

2.4ml 0.5M Na_2EDTA

0.025g bromophenol blue (SIGMA) or orange G (Sigma)

q.s. to 10 ml store at 21°C.

N.B. - in any gel from 0.5%-1.5% agarose, the bromophenol blue moves at the equivalent of 300 bp DNA, whereas orange G migrates equivalent to 150bp. At higher concentrations of agarose, the relationship between DNA and dye doesn't hold.

Ampicillin

50µg/ml for liquid culture

100µg/ml for agar plates

Drosophila media

10% treacle, 20% yeast, 1% agar, 10% polenta, 2.5% tegosept and 1.5% propionic acid

L-Broth (LB)

10g Tryptone

5g Yeast Extract

5g NaCl

q.s. to 1 litre Autoclave

PBS

7.5mM Na_2HPO_4, 2.5mM NaH_2PO_4, 145mM NaCl

PBT

1 x PBS, 0.1% Tween 20

Protein gel running buffer

1.5% Tris Base, 7.2% Glycine, 0.5% SDS

Protein gel sample buffer

10% glycerol, 2% SDS, 5% β-mercapoethanol, 0.05% bromophenol blue, 12.5% xylene cyanol 0.5M Tris-HCL pH6.8

Oligonucleotides

MD-PCR degenerate primers

ephrin-B: CGA CDA TRT ANA CNG GRT GNC CRT ART CNC

ephrin-A: CGA GGR CAD WHN AYR TCN AVR TAR TCR TT

Sequencing primers

pUAST-F: GAA GAG AAC TCT GAA TAG GG

pUAST-R: GTC ACA CCA CAG AAG TAA GG

T7: TAA TAC GAC TCA CTA TAG GG

SP6: TAT TTA GGT GAC ACT ATA G

T3: ATT AAC CCT CAC TAA AGG GA

M13F: GTT TTC CCA GTC AGC AC

M13R: CAG GAA ACA GCT ATG AC

In Situ Probe primers

DekF-T7: TGT AAT ACG ACT CAC TAT AGG GTA TGT ACA CAG AGA CTT GGC

DekR-T7: TGT AAT ACG ACT CAC TAT AGG GAA CCT GGA TAT CTG TTG AGC

C43F-T7: TGT AAT ACG ACT CAC TAT AGG GGG CAA TCT TGC ATT TGA GTT

C43R-T7: TGT AAT ACG ACT CAC TAT AGG GGG ATT GGA ATC CGA TTA AAC

Epitope Tagged expression construct primers

LD1109-NotI-F: AAG CGG CCG CAT GCA AGA ACG ATC AAA GC

LD6HISTGA-KpnI-R: GGG GTA CCC CTC AGT GAT GGT GAT GGT GAT GCC GGT CAT ATT CAA TAG TGC C

LD6TGA-KpnI-R: GGG GTA CCC CTC ACC GGT CAT ATT CAA TAG TGC C

NTERM-R: ATG GAT CCA TGC ATG TAA AAA GTC TTG GCG C

NTERM6HIS-F: ATG GAT CCC ATC ACC ATC ACC ATC ACT GGA ACA CAT CGA ACA G

NTERMHA-F: ATG GAT CCT ACC CAT ACG ATG TTC CAG ATT ACG CTT GGA ACA CAT CGA ACA G

LD-2CHA-F: ATG GAT CCT ACC CAT ACG ATG TTC CAG ATT ACG CTG CAG ATC CGC GAG TAA TAG C

DEP-NCMYC-F: ATG GAT CCT GCG AAC AAA AAC TTA TTT CTG AAG AAG ATC TGT GGA ACA CAT CGA ACA G

DEP-NCMYC-R: ATG GAT CCC AGA TCT TCT TCA GAA ATA AGT TTT TGT TCG CAA TGC ATG TAA AAA GTC TTG GCG C

Methods

Electrocompetent cells

DH5α bacterial cells were grown to an optical density of approximately 0.6 (600nm). The remaining work was performed at 4°C. Cells were placed in four 50ml centrifuge tubes and pelleted by centrifugation @ 4300rpm for 10 minutes. The pellet was resuspended in ice-cold ddH_2O and re-pelleted 3 times to remove as much salt as possible. The cells were resuspended in 10% glycerol, pelleted and finally resuspended in 10% glycerol. Electrocompetent cells were stored @ -70°C.

Electroporation of bacteria with plasmid DNA

Electroporation was performed using a Biorad gene pulser I with a Biorad pulse controller plus as per the manufacture's instructions (Biorad). Cuvettes (Biorad 0.1cm) were prechilled on ice water. Electro competent bacteria (DH5α), were allowed to thaw on ice. Approximately 10ng of plasmid DNA was mixed with 40μl bacteria and electroporated (2.5V, 200Ω, 25 μFD; which gave a time constant in a blank cuvette of approximately 4.98msecs). Immediately after electroporation, 1ml of SOC + glucose was added to the cuvette, the suspension was transferred to an eppendorf tube and cells were incubated for 30 minutes @ 37°C in a water bath which permitted expression of the ampicillin resistance gene. The electroporated cells were plated onto LB + ampicillin to select for successful transformants.

Isolation of mRNA and cDNA synthesis

RNA was isolated from zebrafish and *Drosophila* using selective LiCl precipitation of total RNA (Peale *et. al.*, 1998). In zebrafish we initially isolated RNA from 16-20 hour embryos, which spans an important neuro-developmental period. Similarly, in *Drosophila* we chose 0-15 hour embryos. Following total RNA isolation, mRNA was selected using magnasphere-d(T) beads (Promega) to increase our probability of isolating transcribed sequences. The isolated mRNA was then reverse transcribed into complementary DNA (cDNA) using MMLV-reverse transcriptase.

MD-PCR protocol

The MD-PCR protocol utilises single stranded cDNA as a template (Figure 10). Single stranded cDNA is digested with a restriction enzyme that exhibits single stranded cutting activity (i.e. HaeIII). The use of a single stranded restriction enzyme to digest the first strand cDNA results in products of different sizes due to the distribution of enzyme sites overcoming anther problem of conventional degenerate PCR protocols.

A common oligonucleotide of known sequence is then ligated onto the 5' end of the digested cDNA fragments using a bridging oligonucleotide and diedeoxy-blocking to prevent concatamers of the oligonucleotide forming (Figure 11). Amplification is then carried out using the ligated cDNA as a template with primers complementary to the common oligonucleotide and complementary to the conserved region of the gene family of interest. This PCR reaction is carried out with radio labelled nucleotides allowing the products to be easily resolved on a polyacrylamide gel and eluted for cloning. The MD-PCR method is summarised in Figure 10. A detailed description of MD-PCR is outlined in (Peale *et. al.*, 1998).

Generation of recombinant plasmids

Plasmid vector was digested with the appropriate restriction enzyme for 2-4hrs. CIP was then added directly to the restriction enzyme mix and further incubated for 1hr. The entire reaction was then run on a gel and purified using a gel extraction column (Qiagen). Insert DNA was also digested and purified from a gel fragment. The two fragments were then mixed at equi molar ratios and ligated overnight @18°c. The ligation mixture was then phenol/chloroform extracted and ethanol precipitated (using glycogen as a carrier), resuspended in 10µl of MQ-water in preparation for transformation into electro competent bacteria.

Expression constructs

Blast analysis showed that the *Drosophila* genome project EST LD11109 contains the entire coding region of a putative *Drosophila* ephrin gene. The entire coding region of *d-ephrin* was inserted into pUAST expression vector (Brand and Perrimon, 1993) by PCR amplification from the EST LD11109, introducing a *Not*I site with the primer 5`-AAGCGGCCGCATGCAAGAACGATCAAAGC-3` and a *Kpn*I site with the primer 5`-GGGGTACCCCTCACCGGTCATATTCAATAGTGCC-3`. The restriction endonuclease sites in the primers are underlined. Constructs were confirmed by sequencing across vector insert boundaries and the modified sites. The UAS-Dek-Myc construct was kindly provided to us by John Thomas (Scully *et. al.*, 1999).

Maintenance of *Drosophila* fly stocks

All flies were maintained in either 18°c or 25°c constant temperature rooms with controlled humidity.

Transformation of *Drosophila*

High purity DNA for injection was prepared using either the Qiagen Plasmid Mini or Qiagen Plasmid Midi kits. The pUAST construct of interest was used at 700ng/μl and combined with the transposase activity plasmid (Δ2-3) at 300ng/μl in injection buffer. w^{1118} embryos, staged between 15-30 minutes were dechorinated in 100% bleach for 1 minute 15 seconds and washed thoroughly with 0.7% NaCl and 0.3% Tween. The dechorinated embryos were then aligned on a strip of non-toxic rubber cement (Earth) within 5min to prevent desiccation, and then covered with light paraffin oil. The posterior end of each embryo was then micro-injected with the above DNA mixture and the embryos were left at 25°C in a humidified chamber to hatch and crawl into a yeast paste encircling the embryos. 1st instar larvae were then retrieved and allowed to develop in *Drosophila* medium in normal vials. Adults that survived the injection procedure were individually crossed to w^{1118} virgins allowing transformed lines to be identified amongst the progeny by the w^+ eye colour marker, which is contained on pUAST plasmid (Appendix E). The eye colours obtained varied from pale orange to red, but were consistent with each sex for each independent event (Appendix F). Independent transformants were crossed to the doubly balanced stock, w^{1118};+/CyO;*Df*(3R)ro^{XB3}/TM6b, *Hu*. Male transformants carrying the CyO and TM6b chromosomes were selected and backcrossed to w^{1118} virgins in the next generation. The progeny of this cross were scored to determine whether the P-element insert was segregating from either the second chromosome by the absence of w^+ Cy progeny, or the third chromosome by the absence of the w^+ *Hu* progeny. P-element insertions were identified by the absence of w^+ male progeny. Once the chromosome containing the insert was identified, stable lines were generated by selecting for P-element insert homozygotes, or if this was lethal, stable lines were produced by maintaining the insertion over a balancer chromosome such as CyO or TM6b.

Automated sequencing

DNA to be sequenced was prepared using a Qiagen Gel extraction kit or a Qiagen PCR cleanup kit as required. Big Dye or Die Terminator (Perkin-Elmer) chemistry was used to generate sequence fragments essentially as per the manufactures instructions. Reactions were cycled through 25 cycles of: 95°C for 10 seconds, 50°C for 5 seconds, and 60°C for 5 minutes in a PTC-200 DNA engine (MJ research). 20µl sequence reactions were then cleaned up by precipitating in 80µl freshly prepared 70% isopropanol for 15 minutes, reactions were then pelleted and washed in 250µl of 70% isopropanol, re-pelleted and dried @ 100°C for 30 seconds. Automated sequencing equipment at the department of molecular pathology, IMVS, Adelaide, was used for all sequencing reactions.

***In Situ* Hybridisation analysis of *Drosophila* embryos**

An overnight lay (0-17hrs) of embryos was collected on grape agar plates from w^{1118} stock, de-chorionated, and fixed in 3.7% w/v formaldehyde for 20mins. Embryos were then devitellinsed under methanol according to standard procedures (Patel and Goodman, 2000).

Antisense and sense DIG-labelled RNA probes corresponding to the d-ephrin core domain were generated using the DIG RNA labelling kit (Roche), with the exception that DNA templates were generated using PCR to introduce a T7 promoter region on both the 5` and 3` ends of the DNA used for probe transcription (Appendix F). *In situ* hybridisation analyses were carried out essentially as outlined in the Non-radioactive *In Situ* Hybridisation Application Manual (Roche, 2000). Bound probe was detected using a DIG Nucleic Acid Detection Kit (Roche).

Whole mount antibody staining of *Drosophila* embryos

Fixed embryos were rehydrated by replacing Methanol with a 50% Methanol/PBT mix, embryos were rinsed several times with PBT followed by a single rinse of PBT on a nutator. The embryos were then blocked in 1ml of PBT containing 5% blotto and 0.2% BSA for at least 1 hour. The blocking solution was then removed and primary antibody diluted in up to 200µl fresh blocking solution was added. The embryos were incubated overnight at 4°C on a nutator, the following day the antibody solution was removed and embryos were washed several times with PBT over 2 hours. Embryos were then incubated with secondary antibodies (conjugated to either HRP or AP) in fresh blocking solution for at least two hours at room temperature with gentle nutation. Excess secondary antibody solution was then washed with PBT over a period of 2 hours as per the primary antibody. Antibodies were then detected using either ECL (HRP) or NBT/BCIP (AP), with the colour reaction monitored

under a dissection microscope. The colour reaction was the stopped by rinsing embryos in PBT with 10mM EDTA. Following incubating in STOP, embryos were rinsed thoroughly in PBT and allowed to clear in 80% glycerol in PBT. Embryos were then dissected under a dissection microscope using two 36g needles, and mounted in 80% glycerol in PBT on glass slides with the cover slip sealed by nail polish. Dissected embryos were then examined using a Zeiss Axiophot microscope and photographed using a CCD camera. Images were then edited using Adobe Photoshop 6.0.

S2 Cell culture and Transfection

The S2 cell line (Schneider, 1972) was grown in S2 media (Gibco BRL) supplemented with 10% v/v foetal calf serum (FCS; Gibco BRL), 100U/ml penicillin and 100µg/ml streptomycin. S2 cells were grown until log phase and approximately 1 x 10^6/ml were transfected with a 19:1 ratio calcium phosphate DNA precipitate of UB-UAS: d-ephrin-GAL4 plasmid respectively, in 1ml per 5ml of culture, as described by (Fehon *et. al.*, 1990). Typically, transfection efficiencies of 12-14% were obtained using this protocol. Ub-GAL4 (a gift from Thomas Kornberg) was co-transfected with the UAS constructs to drive the UAS expression vectors continuously. Following transfection cells were allowed to recover for 16hrs in S2 media + 10% v/v FCS.

Aggregation Assays and Immunofluorescence

Equal cell numbers from separate transfection experiments, expressing either Dek-Myc or d-ephrin (Figure 24) were mixed and allowed to aggregate in a roller bottle for 48hrs, as described in (Fehon *et. al.*, 1990). The cell aggregate mixture was then allowed to settle for 30 minutes, onto poly-lysinated slides. Cells were then washed with phosphate buffered saline + 0.1% w/v Tween (PBT). The immobilised cell aggregates were then fixed for 15mins with 3.7% w/v formaldehyde. Cells were washed with PBT, blocked for 30mins in 2% w/v Bovine Serum Albumin (BSA) and 5% v/v sheep serum, and incubated for 2hrs with primary antibody. Excess antibody was washed off with PBT. Cells were then exposed to the appropriate secondary fluorescent conjugate. Cell aggregates were analysed under epifluorescence on an Olympus Provis microscope. Dek^+ and d-$ephrin^+$ cells were counted by eye for their ability to aggregate to other expressing cells. For detection of Myc tagged Dek, anti Myc-Tag 9B11 (Cell Signaling Cat# 2276) was diluted 1:2000, and subsequently detected with CY3 conjugated to anti-mouse IgG. We detected d-ephrin expressing cells by co-transfection with GFP (gift from Barry Dickson) (Figure 24).

Protein gel electrophoresis and Western analysis

The BIO-RAD mini protean III apparatus was used in protein gel electrophoresis, according to the manufacturers instructions. Gels were run at 180-200V until the bromophenol blue in the sample buffer had reached the bottom of the gel. Western blotting of proteins onto nitrocellulose membranes was performed as described in Harlow and Lane (1998). Nitrocellulose blots were washed thoroughly with PBT and then blocked for 1 hour in 5% BLOTTO. Primary and secondary antibody incubations were carried out overnight at 4°c and for 45 minutes at RT respectively, with the appropriate dilutions of antibody in blocking solution. The secondary antibodies were always conjugated to HRP (Jackson), therefore detection was either by ECL or colorimetric detection using nickel enhanced DAB staining (Harlow and Lane, 1988).

Phylogenetic analysis of all currently known *ephrin* genes

Ephrin DNA and protein sequences were extracted from GenBank using exhaustive searches based on gene names and key words. The initial list of sequences was augmented following BLASTN and BLASTP searches using known *ephrin* sequences. Redundancy was eliminated by i) rejecting *ephrin* ESTs, many of which were incomplete and of variable sequence reliability, and ii) rejecting copies of sequences that appeared in the database as cDNA and also as part of genomic sequences. Where multiple database entries existed for similarly identified *ephrin* cDNAs from the same species, comparisons revealed that these sequences were either identical or differed by only several nucleotides. In such cases, a single representative sequence was chosen for further study. (A full list of all sequences used is given in the Supplementary Data). Only coding regions of *ephrin* cDNAs were used for phylogenetic analysis and these sequences were truncated after alignment to include only the *ephrin* core domain (PFAM: 00812). Preliminary alignments indicated that the region of cDNA corresponding to the signal peptide is of variable length and is poorly conserved amongst *ephrins*, making the accuracy of alignment problematic. This region of the cDNA was therefore omitted from phylogenetic analyses.

DNA and protein sequence alignments were carried out with CLUSTALW using the default parameters (Thompson *et. al.*, 1994). Sequence alignments were edited and displayed using GeneDoc (Nicholas *et. al.*, 1997). For phylogenetic analysis, ephrin protein sequences were aligned and then converted to cDNA, maintaining the protein sequence alignment and without interrupting codons. Aligned sequences were bootstrapped using the program SEQBOOT, and strict consensus maximum parsimony trees were derived using the programs DNAPARS and CONSENSE (Felsenstein, 1993)

Maximum Likelihood trees were derived using the BAMBE package (Simon and Larget, 2001). The markov chain Monte Carlo algorithm was applied as described, to the ephrin core cDNA data set as used in the Parsimony analysis. The Tamura and Nei (1993) maximum likelihood model was applied using separate rate characteristics for the third, second, and third nucleotide positions of each codon (Tamura and Nei, 1993). Using the initial estimated parameters, three randomly seeded sets of 1,000,000 topologies were generated. These individual runs where then summarised and analysed independently and in conjunction to confirm consistency. The final tree is a consensus compiled from the results generated from three different random starts. The log likelihood's of trees sampled during each of the four runs plateaued at similar points, supporting the final tree as belonging to a true maximum, not a local maximum. A total of 3 million sampling iterations were included in the final summation. Branch lengths of trees are drawn to scale, and posterior probabilities are indicated at internal nodes. The program TreeView was used to display trees (Page, 1996).

Chapter 3: Isolation and bioinformatics analysis of *d-ephrin*

Isolation and bioinformatics analysis of d-ephrin

Introduction

Affinity binding studies (Gale *et. al.*, 1996a; Gale *et. al.*, 1996b) have shown that some of the 14 known Eph receptors show no significant binding to any of the currently identified ephrins. This suggests that there may exist new and orthologous members of the ephrin subfamily that still remain to be discovered (see Chapter 1, Figure 1).

When this thesis was undertaken, there were no known invertebrate ephrin genes. One Eph receptor was characterised in *C.elegans* in 1998 (George *et. al.*, 1998), suggesting that *ephrin* genes remained to be isolated for *C.elegans*. These were subsequently isolated and characterised (e-ephrin-1-4, see Chapter 1) with the *C.elegans* genome project (Wang *et. al.*, 1999). (Scully *et. al.*, 1999)) reported the isolation of an *Eph* receptor in *Drosophila*, suggesting that an *ephrin* gene also remained to be isolated in *Drosophila*. With this in mind, a new degenerate PCR protocol was employed in order to isolate *Drosophila* ephrin orthologs. Subsequently, the Berkeley *Drosophila* Genome Project released an EST, which encoded for a *Drosophila* ephrin, which is characterised here. This chapter describes the outcomes of these experiments.

Results

Multiplex Display PCR (MD-PCR)

One approach that has been successfully utilised in the past to isolate new members of the Eph/ephrin family is degenerate PCR (Zhou, 1998). There are many degenerate PCR protocols, which attempt to utilise PCR to amplify orthologous or homologous genes. These techniques can amplify multiple related sequences, both known and novel from specific gene families (Wilks, 1989). However, this requires that the target gene family contain two regions of sequence homology, to enable sense and anti-sense (i.e. 5' -> and <- 3') primers to be designed. Any members of a gene family without one of these regions will not be amplified (Peale *et. al.*, 1998). Furthermore, gene products that are amplified are typically of the same length, making separation and sequence determination difficult.

MD-PCR is a technique developed by Peale *et. al.,* (1998), which is designed to overcome the limitations of conventional degenerate PCR techniques by utilising only one degenerate primer specific to a gene family region. Amplification by PCR is achieved by targeting a "common" primer to a restriction enzyme site (Figure 10), facilitated by the use of a bridging oligo (Figure 11). The cDNA fragments that contain only the common ligated primer will be amplified in a linear fashion (i.e. one direction), whereas those cDNA fragments with both the

common primer and the gene family-specific region will be amplified exponentially (i.e. both directions). Also, the length of the fragments will be determined by the distribution of the restriction enzyme sites, which allows for easy separation on a polyacrylamide gel (Peale *et. al.*, 1998).

The MD-PCR approach was utilised in an attempt to find new members of the Eph/ephrin family using single highly conserved domains. Work by Peale *et. al.,* (1998) demonstrated the power of MD-PCR over other PCR techniques in cloning five novel genes from other gene families that had previously been considered to be investigated to saturation (Hox and K^+ channel voltage gated gene families).

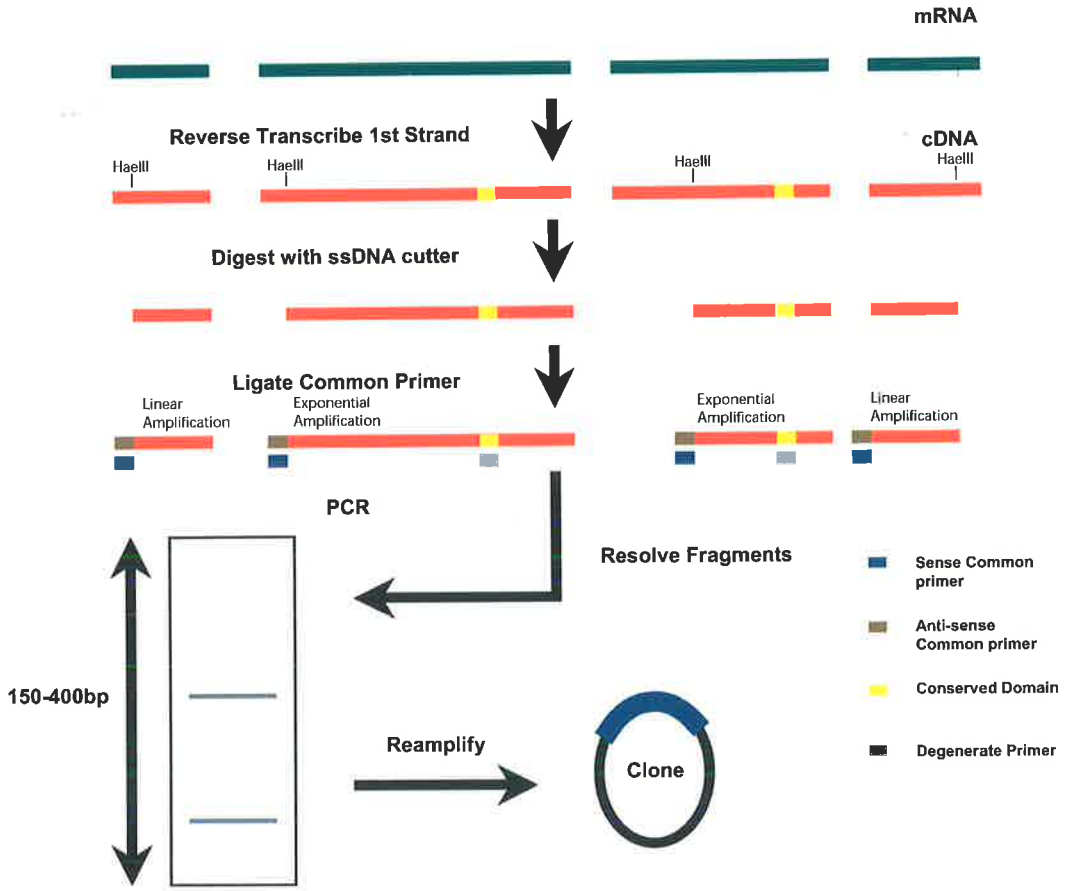


Figure 10 Outline of Multiplex Display PCR (MD-PCR) protocol. MD-PCR is a PCR method that requires only one conserved domain within a gene family. The first strand of cDNA is synthesised and is then digested with an enzyme that exhibits single stranded cutting activity. This results in fragments of different length, dependant on the distribution of the restriction site. A common primer is then ligated onto the 5` end of the cDNA fragments. Fragments that contain both the conserved domain and the common primer will amplify exponentially, whereas those cDNA fragments that contain only the common primer will amplify linearly.

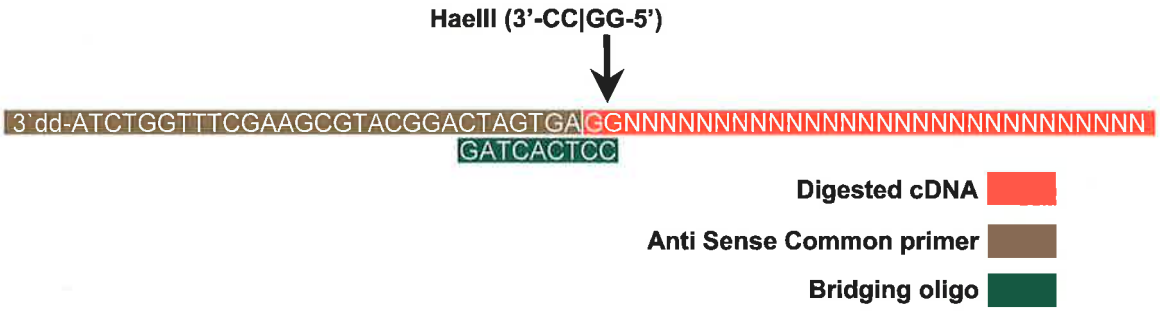


Figure 11 Use of a bridging oligo to facilitate the single stranded ligation reaction between HaeIII digested cDNA and the common primer. Following coupling with the Antisense common primer (brown), to the bridging oligo (green), the cDNA (red) is annealed to the two base pair overhang of the bridging oligo. The cDNA is then covalently linked to the common primer by ligation. Following ligation the bridging oligo is removed from the reaction, prior to PCR amplification.

Conserved regions of ephrin proteins

Ephrin B proteins contain a conserved cytoplasmic tail

Sequence alignments show that all currently known orthologs of the ephrin-B subclass of molecules contain a highly conserved cytoplasmic tail (Figure 12). This region is an excellent candidate for a degenerate primer for the MD-PCR protocol. A degenerate primer for this region was generated for *Drosophila* by back-translating the peptide sequence CPHYEKVSGD using *Drosophila* preferred codons, giving the sequence CGA CDA TRT ANA CNG GRT GNC CRT ART CNC. This primer was then tested for its ability to amplify ephrin-B coding regions by testing the primers in a PCR using plasmid specific promoters (data not shown).

The conserved ephrin core signature of ephrin-A proteins

The ephrin core contains a number of conserved regions, which are similar between both ephrin-A and ephrin-B subgroups. The peptide sequence SEKFQRFT occurs within the ephrin-A core (Figure 12). This region was chosen for degenerate primers as it showed the most conservation in the ephrin-A core that was least similar to sequences within the ephrin-B core. As with the ephrin-B degenerate primer, the sequence CGA GGR CAD WHN AYR TCN AVR TAR was generated by back translating the peptide sequence SEKFQRFT with *Drosophila* codon tables.

MD-PCR with the ephrin-A and ephrin-B degenerate primers

In three separate MD-PCR reactions utilising the highly conserved ephrin-B degenerate primer, no radiolabelled PCR products were seen on a PAGE gel (data not shown). However, when the less specific (in terms of protein coding region) degenerate ephrin-A primer was used in the MD-PCR reaction seven distinct bands were detected on a PAGE gel (Figure 13b). These bands were excised, re-amplified separately and resolved in a 2% agarose gel. Only five of the excised fragments re-amplified successfully (Figure 13b). These fragments were then cloned into pGEM-T easy (Promega). The sequence of the cloned fragments were then determined (see Chapter 2). Blast search analysis revealed that none of the five clones contained any significant similarity to a putative *ephrin* gene (Table 1).

Figure 12 Protein Sequence alignments showing all human and mouse ephrins (other species have been removed for clarity). Amino acids that are conserved between ephrin-A and B subgroups are shown in Black. Amino acids conserved for ephrin-A subgroup only are shown in red. Amino acids conserved in the ephrin-B subgroup only are shown in yellow. The ephrin-A-conserved region used to design the degenerate primer is within the ephrin core domain (green line). The ephrin-B-conserved region used to design a degenerate primer is in the highly conserved cytoplasmic tail (blue line).

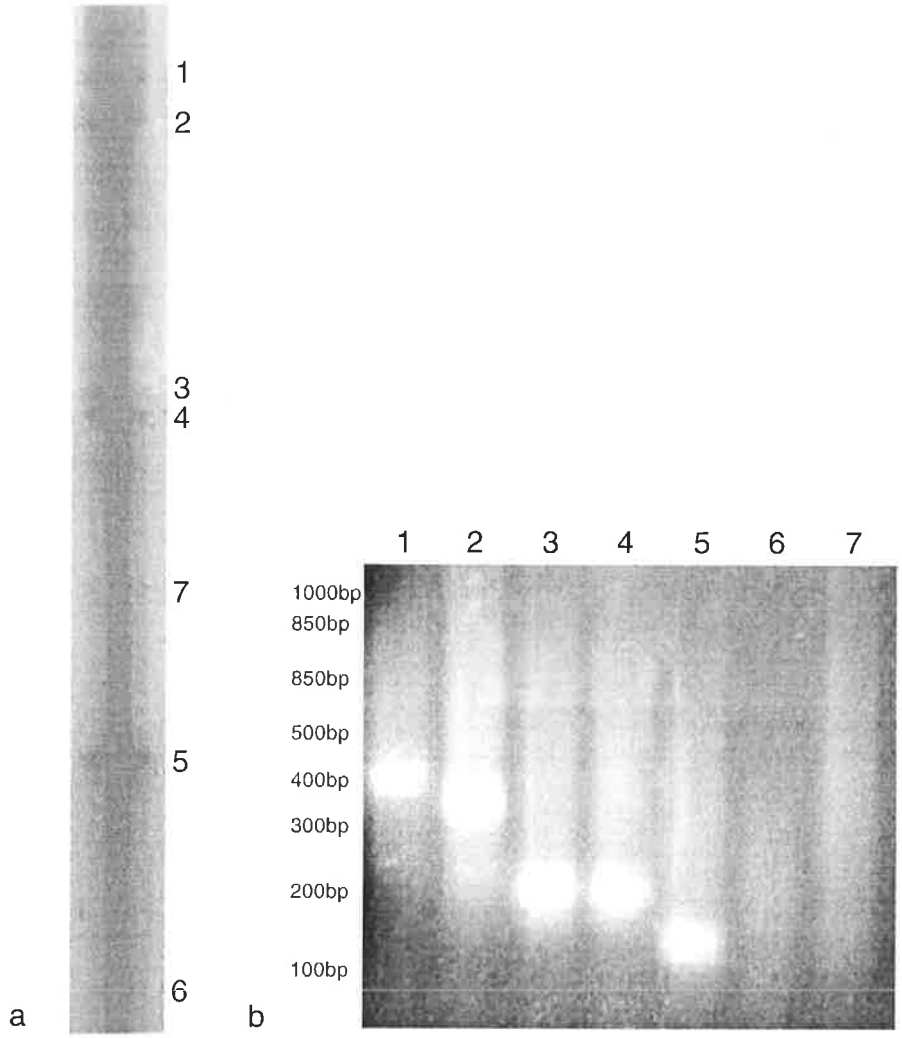


Figure 13 (A) PAGE Gel showing MD-PCR products generated with the degenerate primer for ephrin-A. The brightest and clearest fragments (1-7) were then excised from the corresponding region in the PAGE gel and re-amplified using the common primer and the degenerate ephrin-A primer from the original reaction. (B) 2% Agarose gel with the seven re-amplified fragments. Only five fragments re-amplified and were subsequently cloned.

1	No significant match
2	*Drosophila* Cosmid - NO EST
3	Plasmid
4	*Drosophila* Laminin A
5	No significant match

Table 1 Blast search analysis showed no significant match to any Drosophila ephrins.

EST Isolation of a putative *Drosophila* ephrin gene

While attempting to identify a *Drosophila* ephrin gene using MD-PCR, the first release of the *Drosophila* genome project[6] occurred. Searching these sequences directly with BlastP (Altschul *et. al.*, 1997) using the *C.elegans* (AF201079) ephrin core region as a query, revealed that the EST GH24276 encodes for a gene with approx 50% similarity to the ephrin core region of the four known *C.elegans* ephrins.

The second release of the *Drosophila* genome project showed the entire genomic arrangement of the putative *Drosophila* ephrin (CG1862), which maps to 102C5-102C5 on chromosome 4 and spans 5740bp (Figure 14, Appendix D). This putative *Drosophila* ephrin gene product will be referred to as *d-ephrin* (*Drosophila* ephrin) in accordance in the naming conventions for the ephrin ligand family (Eph Nomenclature Committee, 1997; Lemke, 1997).

D-ephrin is encoded by five exons, similar to m-ephrin-B1 (Fletcher *et. al.*, 1994), h-ephrin-A2, m-ephrin-A3, and m-ephrin-A4 (Cerretti and Nelson, 1998). Furthermore, the boundaries for exons 3-4 and 4-5 (which encode the ephrin core region) are conserved between d-ephrin and the aforementioned vertebrate orthologs (Figure 14, blue stars), indicating a potential conserved function in this region of the gene.

The genome-viewing tool GADFLY[6] showed that the EST LD11109 spans the entire genomic fragment of *Drosophila* ephrin with LD11109 5prime sequence adjacent to the start of the first exon, and LD11109 3prime sequence flanking the end of translation (Figure 14). The putative d-ephrin protein is available on GenBank with Accession Number AAF28394.

Analysis of the open reading frames (ORF) of LD11109 indicates two potential ATG start sites. The first of these corresponds to the first Methionine in the open reading frame, which gives a protein of 652aa (Figure 14). The second potential start site contains the *Drosophila* translation start consensus sequence CAAA**ATG** (Cavener, 1987), suggesting that this may be the start of translation, giving a predicted protein size of 472aa (Figure 14).

A multiple sequence alignment generated with Clustalw (Thompson *et. al.*, 1994), of d-ephrin with all other known ephrins, showed that d-ephrin only shares significant homology with the ephrin core region, with the four invariant cysteine residues also present (Figure 15). However, the rest of the predicted protein shows no significant homology to any currently known ephrin. The predicted d-ephrin protein of (472aa) is much larger than other ephrin-As (200-238aa) and ephrin-Bs (330-340aa).

---

[6] For an overview of the genomic and EST data contained in the *Drosophila* genome collection see (Adams *et. al.*, 2000; Rubin *et. al.*, 2000). The entire genome database is available at http://flybase.bio.indiana.edu/, the tools available to search this database are outlined in (Flybase, 1999; Flybase, 2002).

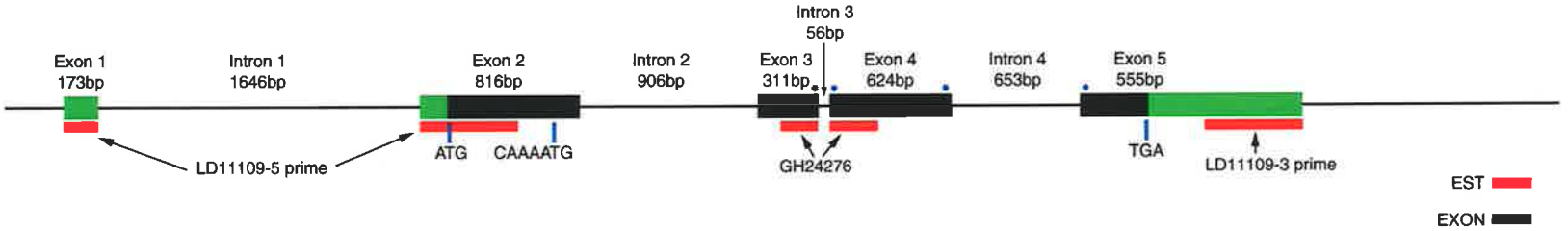
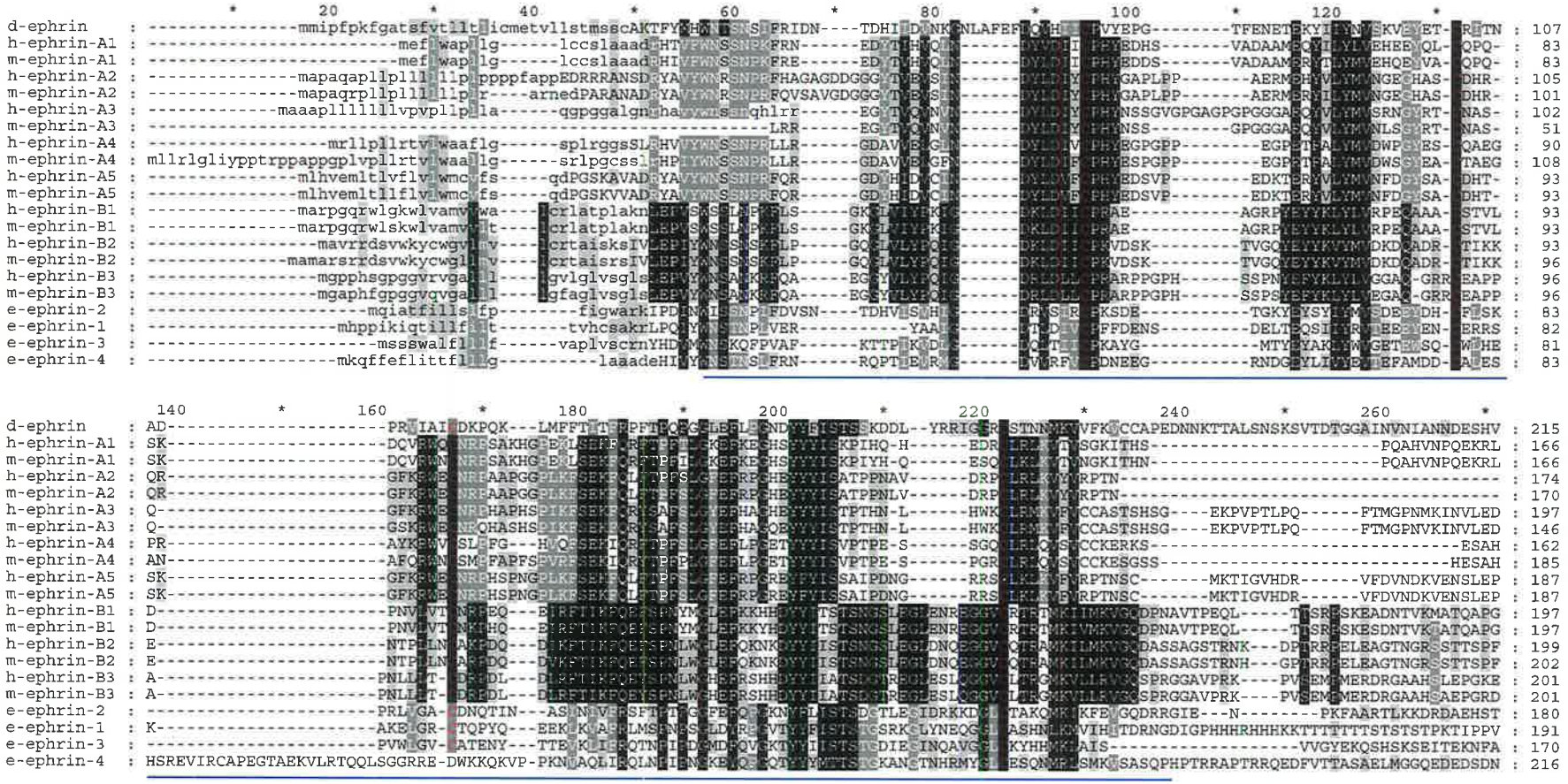


Figure 14 The genomic structure of d-ephrin consists of five exons, which span 5740bp. D-ephrin exons are shown as black boxes. Corresponding exon and intron sizes are indicated. Expressed sequence tags (EST), which overlap the genomic region, are also shown with red boxes. Potential start sites are indicated with green lines. The first in frame stop codon is shown with a blue line. For the full-length genomic sequence see Appendix D.



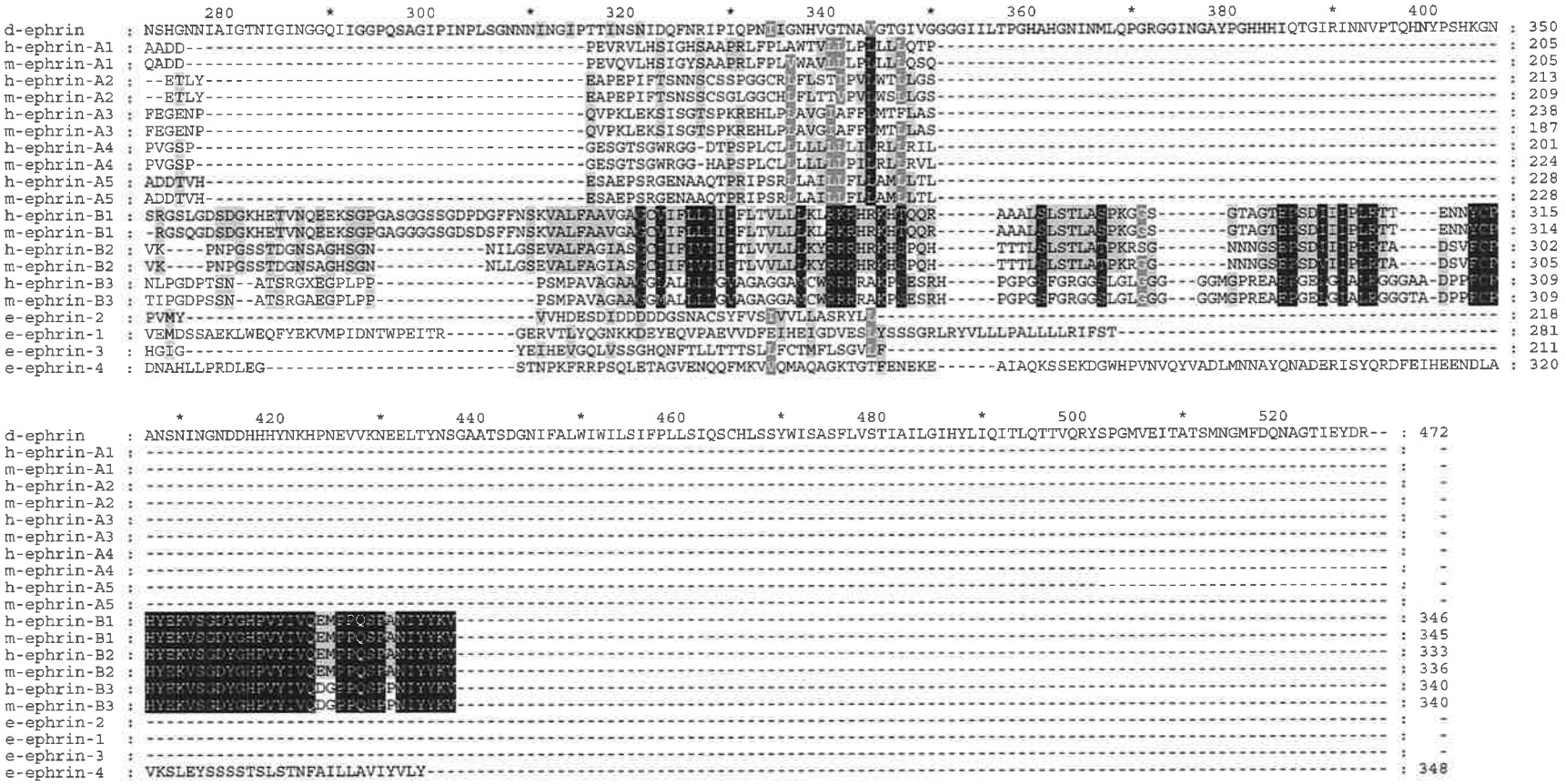


Figure 15 D-ephrin only shares the ephrin core region with other known ephrin orthologs. Clustalw (Thompson *et. al.*, 1994) alignment of *Drosophila* ephrin with *homo sapiens*, *mus musculus*, and *C.elegans* ephrins. The signal peptides for each ephrin shown are in lower case (for a complete list of signal peptides see Appendix B**). The ephrin core region is underlined with a blue line. The four invariant cysteines are shown in red. Shading is grouped according to anchorage mechanism and invertebrates and vertebrates (i.e. ephrin-A, ephrin-B and invertebrates). 100% similarity within groups is shaded black, 80% similarity within groups is dark grey with black letters, and 60% similarity is light grey with white letters (alignment drawn with Nicholas *et. al.*, 1997).**

MEME analysis of d-ephrin

MEME is a bioinformatics tool that searches for motifs in a set of unaligned unrelated sequences (Bailey and Elkan, 1994). A statistical model is applied to determine if the likelihood that the amino acid pattern within a given window (i.e. 28 aa) is sufficiently conserved to constitute a motif. The parameters of the program can be modified to search for any number of motifs, within a given size (window).

Work by Hattori *et. al.,* (2000), used MEME analysis to determine if a common motif existed between proteins (Delta, TNF-α, and APP) which are known to be cleaved by the metalloproteinase *kuzbanian* and the eight vertebrate ephrins in an effort to find a conserved metalloproteinase domain in ephrins. A conserved motif was found (Table 2a), to be located centrally in the ephrin core region (Hattori *et. al.*, 2000). *In vitro* analysis showed that this conserved region is involved in regulated cleavage of the ephrin ectodomain after receptor binding via *kuzbanian*, which facilitates axon retraction (see Chapter 1).

In an effort to determine whether d-ephrin also contained a similar motif, a MEME analysis was conducted with d-ephrin and m-ephrin-B1 and m-ephrin-A2 orthologs (Table 2b). This analysis showed that d-ephrin contains a putative metalloproteinase site, consistent with the study done by Hattori *et. al.,* (2000), with the same motif being derived via an independent MEME analysis. This motif is represented in MEME analyses with the majority of ephrins vs. d-ephrin (data not shown), suggesting that similar metalloproteinase mechanisms exist for all ephrin orthologs, although the functional mechanisms of this putative motif in d-ephrin requires analysis *in vitro*.

a

Sequence name	Start	Site
m-ephrin-A2	104-131	RPAAPGGPL KFSEKFQLFT PFSLGFEFR
m-ephrin-B1	59-86	TCNKPHQEI RFTIKFQEFS PNYMGLEFK
mDelta	25-52	LLCQVWSSG VFELKLQEFV NKKGLLGNR
mAPP	606-633	MDAEFGHDS GFEVRHQKLV FFAEDVGSN
mTNF	161-188	YVLLTHTVS RFAISYQEKV NLLSAVKSP
Consensus		KFEIKFQEFV
		R VRY
		L

b

Sequence name	Start	Site
d-ephrin	305-335	ICDKPQKLMF FTITFRPFTP QPGGLEFLPG
m-ephrin-B1	110-140	TCNKPHQEIR FTIKFQEFSP NYMGLEFKKY
m-ephrin-A2	122-152	RPAAPGGPLK FSEKFQLFTP FSLGFEFRPG

Table 2 (a) MEME (Bailey and Elkan, 1994) analysis of m-ephrin-A2 and m-ephrin-B1 with other mouse proteins known to contain a metalloproteinase binding site (data taken from Hattori *et. al.*, 2000). (b) An independent MEME analysis with the same two ephrins and d-ephrin shows the same site, suggesting d-ephrin may also contain a metalloproteinase recognition site.

Identification of potential Signal Peptides using SignalP

Signal sequences are responsible for directing a protein polypeptide across the membrane of the endoplasmic reticulum (Plath *et. al.*, 1998). A signal peptide is comprised of an N-terminal signal peptide, which is cleaved while the protein is translocated through the membrane. The presence of a signal peptide is an indication of membrane association or secretion of a protein.

In order to determine if d-ephrin contains a signal peptide, an analysis of the two potential start sites discussed earlier was undertaken using the SignalP (Nielsen *et. al.*, 1999) prediction software. This software can predict a potential signal peptide with up to 90% accuracy, using two distinct methods. The first was a neural network approach (NN), whereby the software "learns" to recognise signal peptides by training on proteins that have had their signal peptides determined experimentally (Nielsen *et. al.*, 1997b). The second uses a Hidden Markov model (HMM) to determine the likelihood of a signal peptide based on a statistical model (Nielsen and Krogh, 1998).

Analysis of the first 70aa of the open reading frame of d-ephrin with SignalP (Nielsen *et. al.*, 1997b; Nielsen and Krogh, 1998) gives no significant evidence for a signal peptide using either a NN or HMM approach (Figure 16). However, analysis of the 70aa after the Methionine (181aa into the ORF) adjacent to the Cavener sequence (Figure 14), gives significant evidence for a signal peptide with a cleavage site between SSC-AK (p=0.965) indicating a potential membrane targeted protein (Figure 17). This region also corresponds to the beginning of sequence homology with other ephrin proteins (Figure 15). Taken together these data suggest a putative start of translation for d-ephrin at Methionine 181 from the start of the ORF.

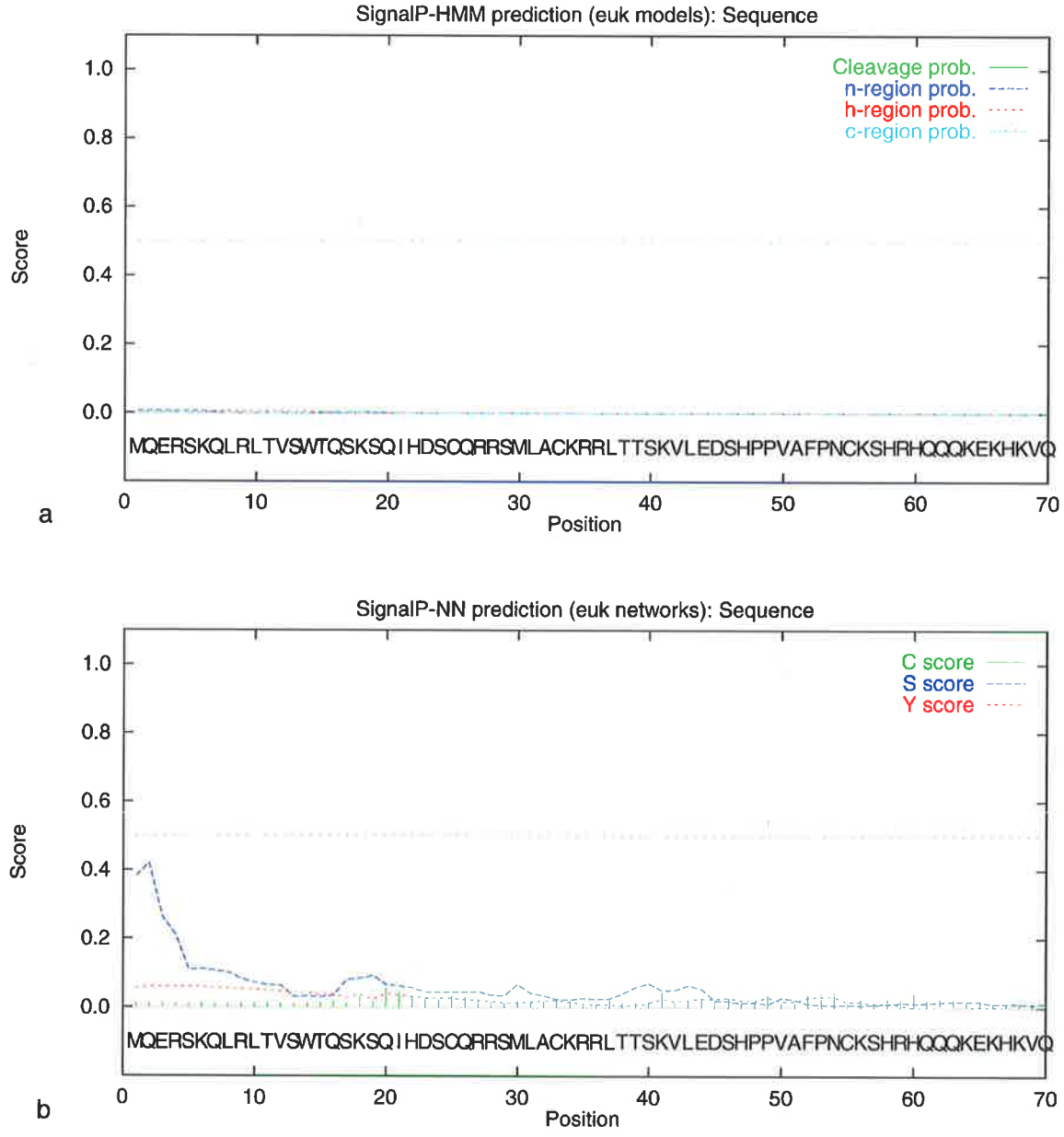


Figure 16 The 5`aa sequence of the full open reading frame (ORF) of *d-ephrin* does not encode a signal peptide. a) Hidden Markov model analysis of 70aa from the full ORF of *d-ephrin* shows no significant signal peptide (maximum cleavage probability 0.005). b) Neural network prediction also shows no significant signal peptide in the first 70aa of the ORF (mean S-score 0.402)

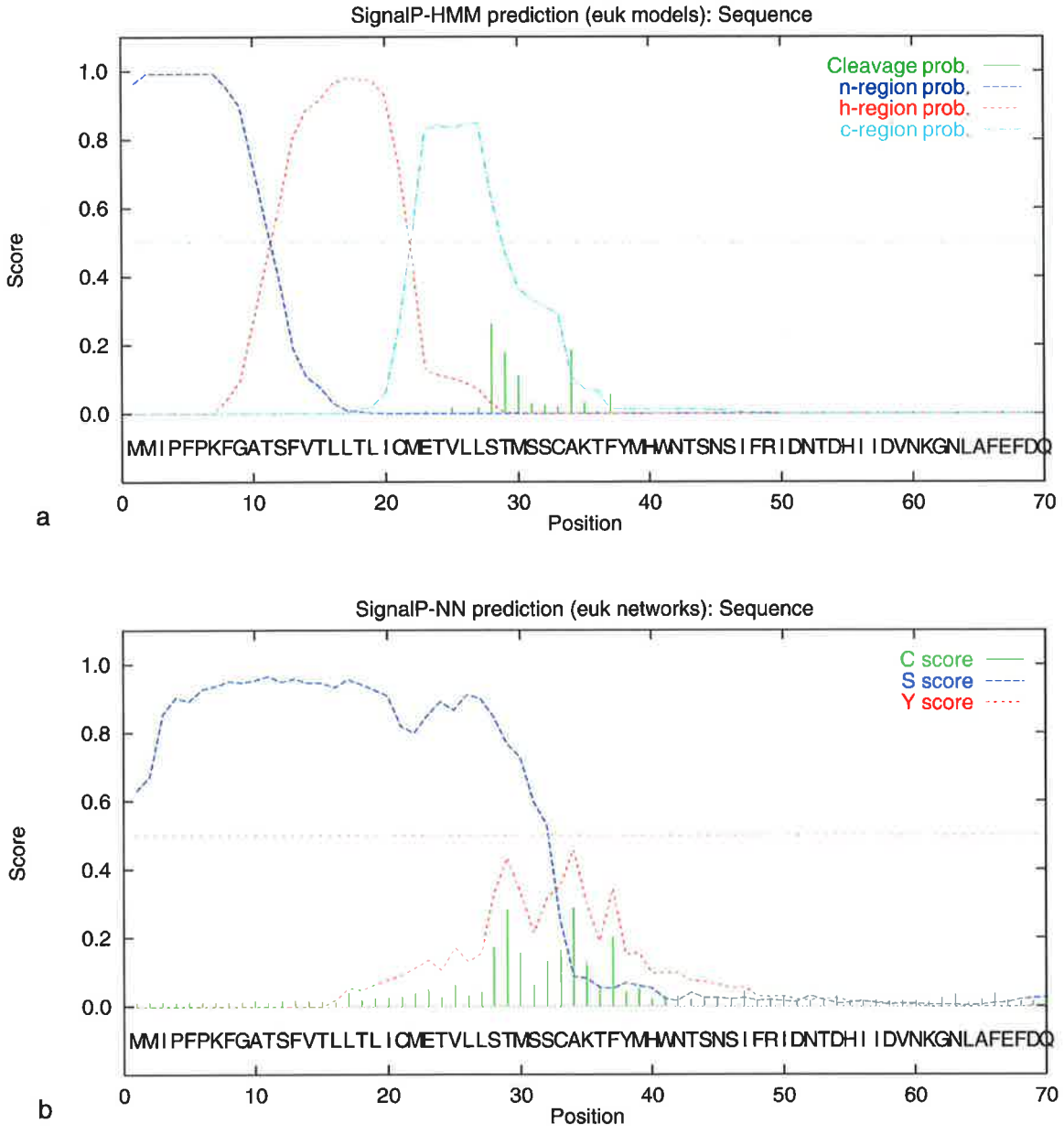


Figure 17 Analysis from the Methionine (180-250aa) adjacent to the Cavener translation start sequence (Cavener, 1987) shows a significant potential Signal peptide. a) Hidden Markov model analysis of 70aa from the Methionine adjacent to the Cavener shows a significant signal peptide (c-score = cleavage probability = 0.965). b) Neural network prediction also shows a significant signal peptide (mean S-score 0.844).

Prediction of potential transmembrane regions in d-ephrin with TMHHM

Following identification of a putative signal peptide, it was important to determine whether d-ephrin contained residues consistent with a transmembrane protein. Transmembrane regions are characterised by unusually long stretches of hydrophobic residues, allowing the accurate prediction (up to 95%) of transmembrane regions in proteins (Moller *et. al.*, 2001). Furthermore, the orientation or topology of proteins within a membrane is characterised by the presence of the positively charged residues arginine and lysine on the cytoplasmic side of the transmembrane region which is often referred to as the 'positive inside rule' (Wallin and Von Heijne, 1998). The bioinformatics program TMHMM allows accurate prediction of potential transmembrane spanning regions and protein topology by applying a hidden Markov model approach to membrane prediction (Krogh *et. al.*, 2001).

The hydrophobic transmembrane spanning regions of ephrin-B proteins are easily detected in all ephrin-B members, with the corresponding extracellular region and conserved intracellular regions correctly predicted by TMHMM (Figure 18a). Similarly, ephrin-A proteins all share a typical hydrophobicity plot with a cluster of hydrophobic residues at the C-terminus of the polypeptide chain, which is cleaved during posttranslational attachment of the GPI moiety (Figure 18b). Also common to both ephrin-A and ephrin-B proteins is the hydrophobic signal peptide which is also detected by TMHMM, which corresponds to the signal peptide predicted by SignalP.

In order to determine if any potential transmembrane regions are present within the ORF a TMHMM analysis of the predicted d-ephrin protein was performed.

Analysis of the full ORF of d-ephrin with TMHMM shows three potential transmembrane regions ($p > 0.9$). Amino acids 188-210 which is similar to the predicted signal peptide discussed earlier (Figure 17). In addition, amino acids 570-587, and 592-614 are both predicted as potential transmembrane regions (Figure 19a). The two latter membrane-spanning regions predicted by TMHMM suggest the presence of an ephrin-B-type anchorage mechanism with an intracellular region for d-ephrin. However, the topology of the protein is predicted with the ephrin core region to be intracellular, when it should be extracellular.

Further, analysis of a truncated form of d-ephrin, with the 5'-untranslated region up to the predicted signal peptide (Met 181) and the last 53 aa of the ORF corresponding to the last hydrophobic region removed, gives an ephrin-A type profile in terms of protein topology and probability of the hydrophobic regions, which is lower when the noise of the untranslated regions is removed (Figure 19b). This truncated version may be a true profile of the d-ephrin membrane anchorage, giving a putative *Drosophila* ephrin-A type protein. Clearly, this needs to be tested *in vitro*.

In summary, the membrane anchorage of d-ephrin is unable to be accurately detected by TMHMM analysis, however this program can accurately predict transmembrane regions for all ephrin-B gene or the lack thereof for all ephrin-A genes. Although the transmembrane regions predicted for the full ORF are a good candidate for membrane anchorage.

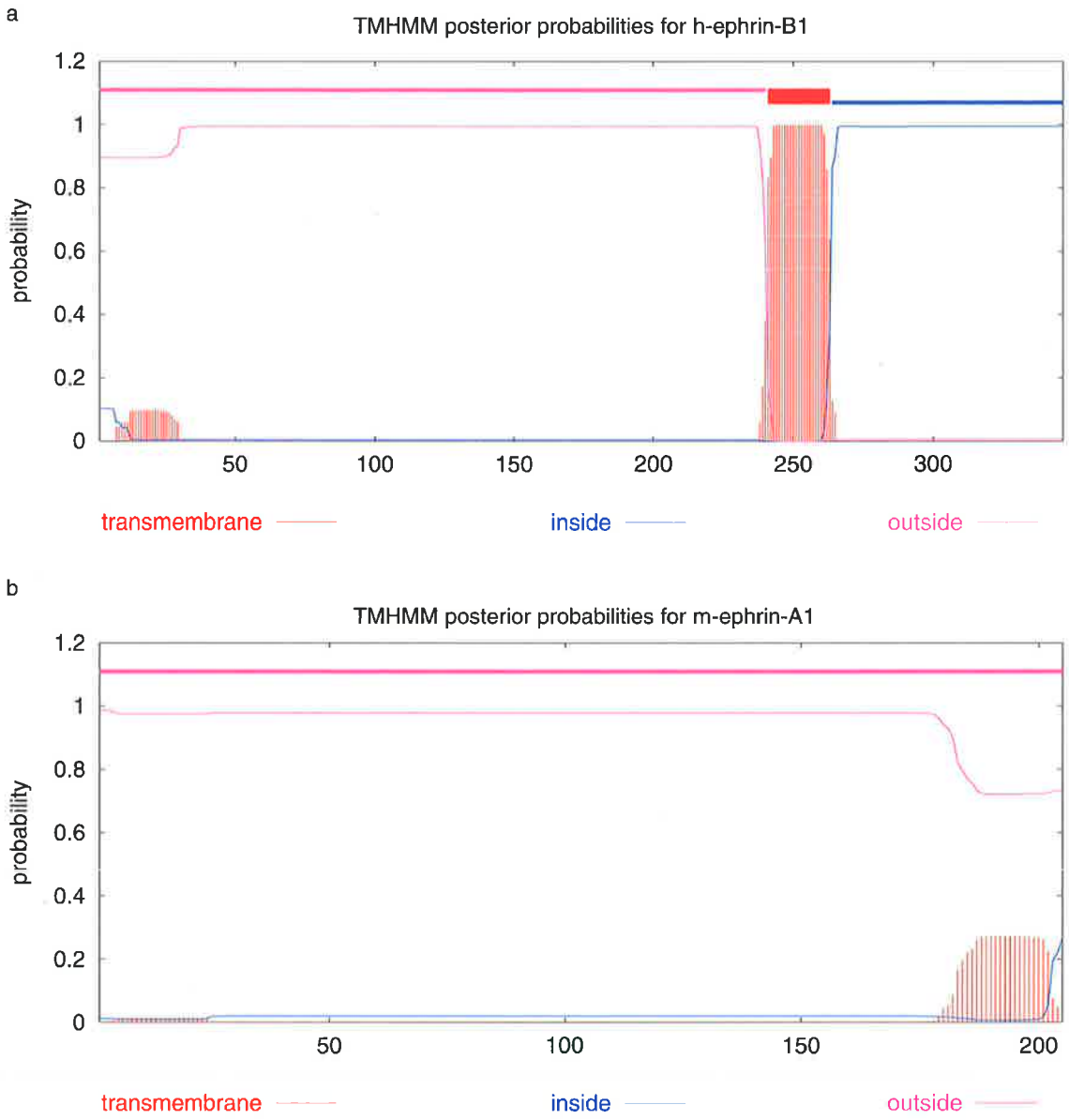


Figure 18 Typical TMHMM profiles for ephrin molecules. (a) TMHMM profile for h-ephrin-B1. The small hydrophobic peak in the first 20aa is a typical signal peptide signature ($p < 0.2$). The large hydrophobic peak at 250aa is the transmembrane region of the ephrin-B molecule ($p > 1$). The intracellular region is shown in blue, the extracellular domain is in red. This profile is typical of all ephrin-B proteins. (b) TMHMM profile for m-ephrin-A1. The small hydrophobic peak in the first 20aa is a typical signal peptide signature ($p < 0.2$). The large hydrophobic peak at 200aa is the hydrophobic GPI anchorage region, which is cleaved after GPI moiety attachment ($p > 0.2 < 0.4$). This profile is typical of ephrin-A proteins.

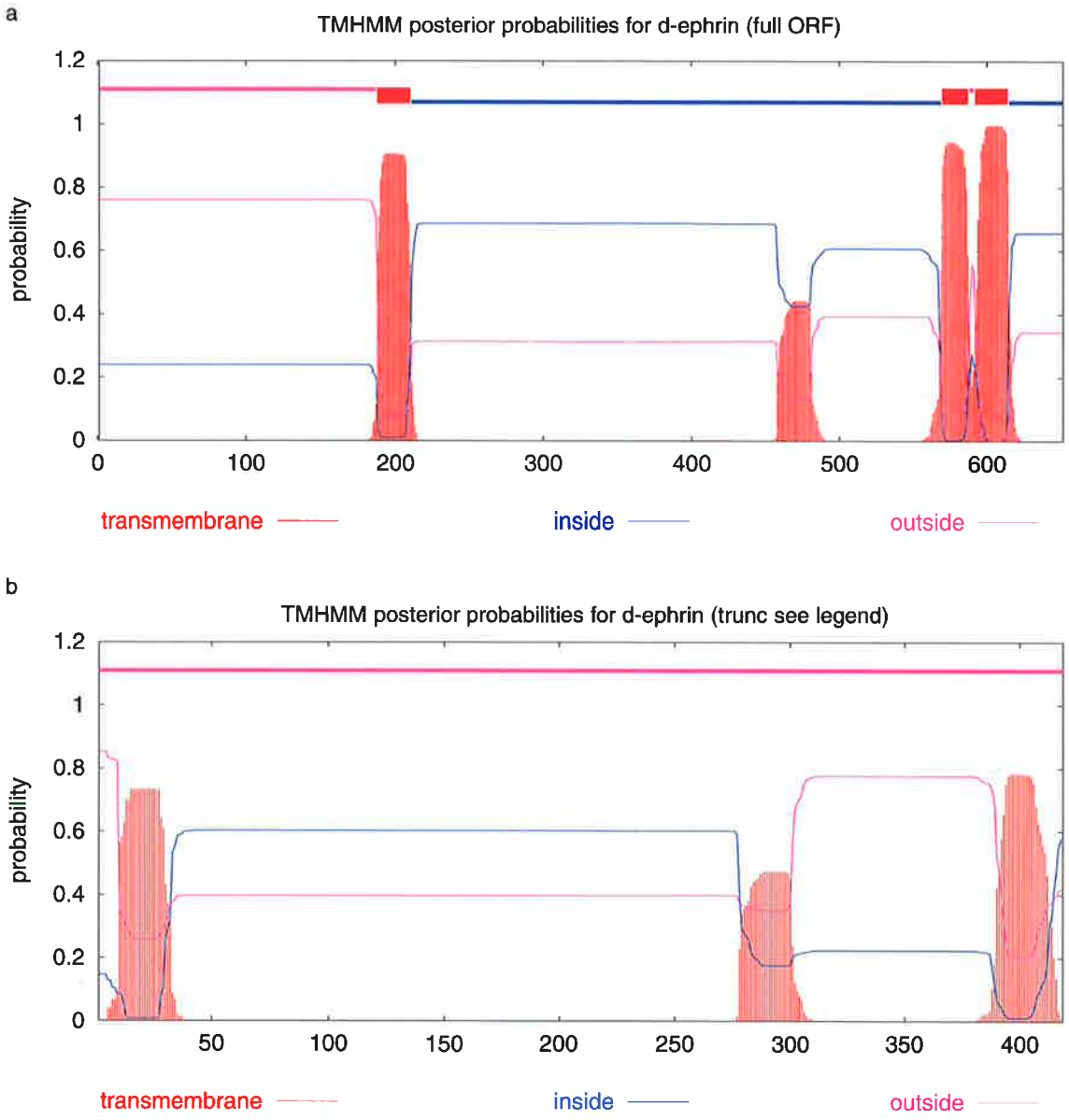


Figure 19 TMHMM predictions for d-ephrin does not fit either a typical ephrin-A or ephrin-B profile. (a) Full 652aa open reading frame of d-ephrin showing three potential transmembrane regions (TM) (p>0.9), the first predicted region at 200aa corresponds to the signal peptide predicted in Figure 17. The second TM region around 475 aa is discounted by TMHMM due to the low probability value. The membrane topology for d-ephrin is also incorrect with the putative extracellular core (220-359 aa) predicted as intracellular. (b) Removal of the untranslated regions of d-ephrin results in an ephrin-A like profile. In this profile the d-ephrin extracellular core (44-183 aa) is correctly orientated and all hydrophobic regions are rejected as potential transmembrane regions (p < 0.8).

Detection of GPI posttranslational modification sites using Big PI and DGPI

The analysis of transmembrane regions with TMHMM did not conclusively show that d-ephrin is a transmembrane protein. The ephrin-A type profile shown in the TMHMM analysis of the truncated d-ephrin protein (Figure 19b) prompted an analysis of potential GPI modification sites for the entire open reading frame of d-ephrin. This was done using both the Big PI (Eisenhaber *et. al.*, 1999b) GPI modification predictor discussed in Chapter 1 and DGPI (Buloz and Kronegg, 2001), which is a similar program, trained on a different set of known GPI proteins.

The 652aa open reading frame of d-ephrin does not contain a potential GPI modification site. Conversely, if the ORF of LD11109 is truncated at aa 594 a potential GPI modification site is predicted with both Big PI and DGPI, indicating the presence of a potential GPI anchorage mechanism for *Drosophila* ephrin (Figure 20). The caveat to this is that if h-ephrin-B1 (XP_038809) is truncated at the hydrophobic transmembrane region (253 aa), BDGPI will predict a potential GPI modification, although Big PI does not detect a potential GPI modification in the 253 aa portion of h-ephrin-B1. This demonstrates that caution must be used in assigning GPI anchorages by sequence analysis. However, both Big PI and DGPI are able to detect all GPI anchorage sites for the vertebrate ephrin-As as well as the four currently known GPI anchored *C.elegans* ephrins.

Protein	d-ephrin truncated from Signal Peptide to second hydrophobic region
Length	414
Data from DGPI	**N-term Signal :** there is a N-term signal (1..35 in green) maximal score=3.72 **C-term Hydrophobicity profile :** hydrophobe length (low-pass filter)=29 hydrophile length (low-pass filter)=11 hydrophobe length (median filter)=7 hydrophile length (median filter)=2 average hydrophobe length = 18.0 (in blue) average hydrophile length = 6.5 **Cleavage site (w) :** There's a GPI-anchor near 386 (7 aa after hydrophobic tail) There's a potential cleavage site at 386 (score=0.1462) detected by w,w+2 rule. There's a potential cleavage site at 382 (score=0.64800006) detected by w,w+2 rule. There's a potential cleavage site at 381 (score=2.5800002) detected by w,w+2 rule. There's a potential cleavage site at 380 (score=0.9366999) detected by w,w+2 rule. The best cleavage site is 382 **Conclusion :** This protein is GPI-anchored (signal, hydrophobic & hydrophilic tail present). There is a potential cleavage site at 382 (w, w+1, w+2 in red)
Sequence annotated by DGPI	MMIPFPKFGATSFVTLLTLICMETVLLSTMSSCAKTFYMHWNTSNSIFRIDNTDHIIDVNK GNLAFEFDQVHIICPVYEPGTFENETEKYIIYNVSKVEYETCRITNADPRVIAICDKPQKL MFFTITFRPFTPQPGGLEFLPGNDYYFISTSSKDDLYRRIGGRCSTNNMKVVFKVCCAPED NNKTTALSNSKSVTDTGGAINVNIANNDESHVNSHGNNIAIGTNIGINGGQIIGGPQSAGIP INPLSGNNNINGIPTTINSNIDQFNRIPIQPNIIGNHVGTNAVGTGIVGGGGIILTPGHAHGNI NMLQPGRGGINGAYPGHHHIQTGIRINNVPTQHNYPSHKGNANSNINGNDDHHHYNKH PNEVVKNEELTYNSGAATSDGNIFALWIWILSIFPLLSIQSCHLSSY

Figure 20 Prediction of a potential GPI anchorage in the truncated form of d-ephrin. A truncated form of the full open reading frame (653 aa) gives the potential GPI anchorage point (GAA). The hydrophobic C-terminus, which is cleaved during attachment of the GPI moiety, is shown in blue, and a signal peptide which corresponds to the signal peptide predicted with SignalP is shown in green. These data were generated using DGPI (Buloz and Kronegg, 2001). Big PI (Eisenhaber *et. al.*, 1999b) also predicts a potential GPI modification at GAA when presented with the same sequence.

Discussion

When this work was undertaken, the *Drosophila* Eph, *dek,* had been isolated and characterised (Scully *et. al.*, 1999), although no phenotype had been described. Recently, *dek* has been found to be involved in the formation of a topographic map in a kinase dependant fashion, in a similar manner as the vertebrate *ephrin* genes (Dearborn *et. al.*, 2002). This work provides compelling evidence for the function of *d-ephrin* in the nervous system of *Drosophila.* In an attempt to isolate a *Drosophila* ortholog of an *ephrin* gene, the MD-PCR technique presented here was utilised.

A primer for ephrin-Bs was designed based on the premise that all known orthologs of ephrin-B proteins contained a highly conserved intracellular tail (Figure 12). Similarly, a primer for ephrin-A proteins was designed based on the most highly conserved region of the ephrin-A core (Figure 12). However, initial attempts to isolate ephrins using the *ephrin-B* degenerate primer showed no products on the PAGE gel of the PCR reaction. Following release of the *Drosophila* genome sequence, it is evident that d-ephrin shows no intracellular ephrin-B like tail, to which the ephrin-B degenerate primers were targeted. This explains why no bands were found on the polyacrylamide gel.

When MD-PCR was performed with the *ephrin-A* degenerate primers, 7 bands were amplified. The five that could be reamplified showed no similarity to any currently known *ephrin* genes. The isolation of Laminin-A in this screen appears to be due to a low stringency, as there are no regions within Laminin-A that are conserved with ephrins or the primer. One potential solution to this problem would have been to increase the stringency of these reactions, or to redesign the ephrin-A degenerate primers.

However, during the implementation of these procedures the *Drosophila* genome project was released, revealing a putative *Drosophila ephrin* gene. As a result the optimisation of the MD-PCR protocol ceased and work became focused on the characterisation of the *Drosophila* ephrin.

Bioinformatics analysis is now recognized as a valuable part of gene characterization and the initial characterization of *d-ephrin* was approached in this manner. The ORF of the *d-ephrin* gene is much larger than any other known ephrin. A bioinformatics approach was utilised, in an effort to elucidate potential protein structure and anchorage mechanisms of d-ephrin. The presence of a signal peptide adjacent to a translation initiation site (Cavener, 1987) gives clear evidence for the start of translation of d-ephrin at this region (Figure 21), although this needs to be confirmed by peptide sequencing. Initiation of translation at this site gives d-ephrin a 472aa protein product, still larger than any currently known vertebrate or invertebrate ephrin.

What is still unclear is the anchorage mechanism employed by d-ephrin, as analysis with TMHMM of both the full ORF and the truncated *d-ephrin* gives no definitive indication of a transmembrane region, as it does for any other known ephrin protein. Furthermore, detection of GPI anchorage regions is possible by truncating the protein, although the same is true for h-ephrin-B1 which is well characterised as a transmembrane protein. However, this does not completely rule out the possibility that d-ephrin is GPI anchored. There are four other known invertebrate ephrins all in *C.elegans*, which are all GPI anchored (i.e. ephrin-A like in terms of anchorage mechanism), therefore it is tempting to speculate that d-ephrin too is GPI anchored, albeit with an additional posttranslational modification preceding GPI attachment. Ultimately, a definitive answer to this question can only be obtained in an *in vitro* system, such as detection via PI-PLC cleavage (see Chapter 1).

Interestingly, the ephrin core region remains conserved in *d-ephrin* although the rest of the protein does not bear any resemblance to either ephrin-A or ephrin-B molecules. Furthermore the intron-exon boundaries are conserved for the ephrin core between d-ephrin and invertebrate ephrins (Figure 21). Conversely, the *Drosophila* Eph receptor *dek* (Scully *et. al.*, 1999), appears to be "prototypic" in terms of sequence homology, with the amino acid sequence of *dek* being equally homologous to both vertebrate EphA and EphB receptors.

All these apparent differences of d-ephrin may simply be due to a sequence error in LD11109, which results in an error of the predicted gene in the BDGP, however the full length of LD11109 was sequenced to confirm the predicted ORF listed on the BDGP and no errors were found.

The results shown in this chapter confirm that the *Drosophila ephrin* gene, contains a conserved ephrin core region. Also, the ephrin core contains a highly conserved motif, which may be a cleavage site for a metalloproteinase, facilitating axonal retraction as in vertebrate axon guidance systems.

However, the rest of the protein appears to be unique, and bioinformatic analysis is unable to determine any motifs outside the ephrin core region. Furthermore, bioinformatics analysis is unable to determine if d-ephrin is an ephrin-B or an ephrin-A gene. A summary of all the result presented in this chapter is presented in Figure 21.

The question of the anchorage mechanism still remains unclear. The next chapter outlines experiments, that address this question *in vitro*. Also, the *in vivo* function of d-ephrin is also addressed.

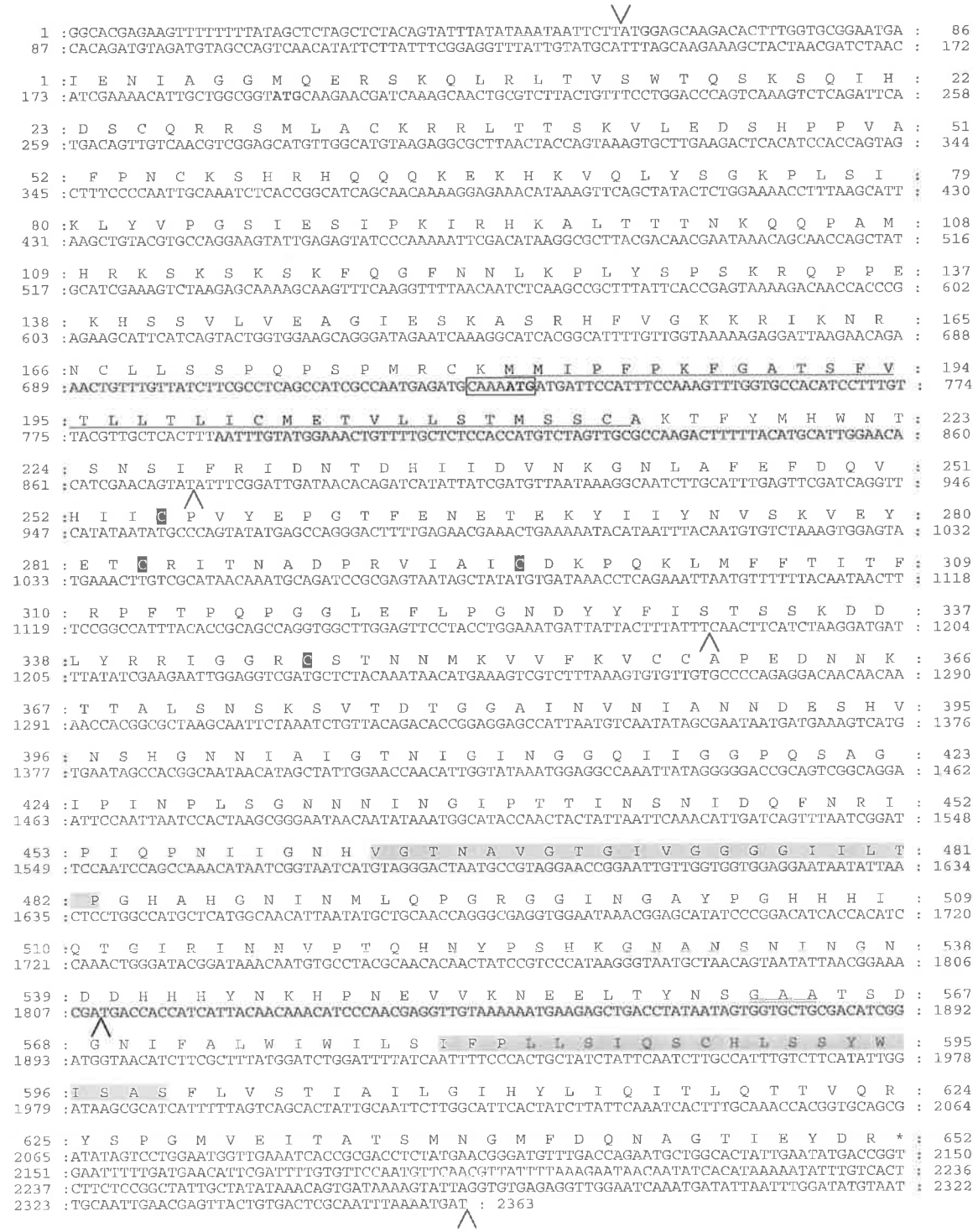


Figure 21 Structure of the *Drosophila* ephrin gene. The potential signal peptide sequence is underlined; the four conserved cysteine residues of the ephrin core region are boxed; the intron exon boundaries are indicated with arrow heads; the exon boundaries which are conserved with eukaryotic ephrins are marked with an asterisk; the two potential transmembrane regions are shaded in grey; the potential GPI anchorage point is double underlined

Chapter 4: Characterisation of d-ephrin

Characterisation of d-ephrin

Introduction

The previous chapter outlines the identification of the first *Drosophila* ephrin gene (*d-ephrin*). The full open reading frame of *d-ephrin* encodes a putative protein of 652 aa. Analysis of the full open reading frame shows a signal peptide adjacent to a translational start site, which gives a protein size of 472aa. This protein product only shares similarity with the ephrin core region of other ephrin genes, the rest of the protein exhibits no clear intracellular domain or functional motifs, suggesting either that the protein is GPI anchored or that the intracellular region has no function. A reliable prediction of either of these scenarios using bioinformatics was not possible with d-ephrin.

In an effort to gain a clear understanding of the anchorage mechanism employed by d-ephrin an *in vitro* approach is needed. To achieve this, *d-ephrin* was expressed in *Drosophila* S2 (Schneider, 1972) cells. The binding of *d-ephrin* to bind *dek* was investigated by assaying the ability of *d-ephrin* expressing S2 cells to bind to *dek* expressing S2 cells (i.e. to form aggregates).

In order to understand the function of *d-ephrin in vivo*, the spatial and temporal expression pattern of *d-ephrin* was determined by *in situ* analysis. Furthermore, a misexpression approach was utilised in order to analyse d-ephrin function *in vivo*. This chapter describes the results of these experiments.

Results

Visualisation of d-ephrin protein using epitope tags

In an effort to provide a mechanism for the visualisation of d-ephrin protein, a number of epitope tagged constructs were generated and cloned into pUAST expression vector (see Chapter 2). An epitope tagged approach to this problem was chosen to expedite the work on this protein, as opposed to the generation of an antibody to d-ephrin.

The *dek* expression construct used in all experiments performed was kindly provided by Scully *et. al.,* (1999). It encodes for a MYC tagged Dek protein (*UAS-dek-MYC*). Therefore, it was decided to generate d-ephrin constructs, which coded either for 6HIS or for HA epitope tags, as these would be most useful when used in experiments that contain the *dek-MYC* construct.

A *UAS-d-ephrin-6HIS* tagged construct was generated by the addition of 6 histidines to the COOH terminus of the full ORF of d-ephrin (Figure 23) using PCR (see Chapter 2). This modified d-ephrin was cloned into pUAST and sequenced in full to determine any errors

introduced by PCR. Sequence analysis showed no errors within the reading frame of the *UAS-d-ephrin-6HIS* construct. However, when this construct was introduced into S2 cells and expressed by induction with *UB-GAL4* (which constantly drives expression of GAL4), no protein products were detectable on a western blot with α-mouse-6HIS monoclonal or α-rabbit-6HIS polyclonal antibodies. Also, α-mouse-6HIS monoclonal or α-rabbit-6HIS polyclonal antibodies were unable to detect d-ephrin-6HIS directly *in vitro* when expressed in S2 cells. The *UAS-d-ephrin-6HIS* construct was able to induce cell aggregates when mixed with *dek* S2 cells (discussed later), suggesting that the extracellular region of the d-ephrin-6HIS protein is intact and functional. Therefore, the C-terminal 6HIS tag is either not present, not functional, or unable to be detected by the antibodies used. It is possible that the intracellular portion of d-ephrin is cleaved and proteolysed, which may account for the lack of the 6HIS epitope in *UAS-d-ephrin-6HIS* cells.

In an attempt to circumvent the potential cleavage and breakdown of the intracellular portion of *UAS-d-ephrin-6HIS*, three *UAS-d-ephrin* epitope tagged constructs were made. The first was, an N-terminal 6HIS tagged d-ephrin (*UAS-6HIS-d-ephrin*), the second was an N-terminal HA tagged d-ephrin (*UAS-HA-d-ephrin*) and the third was a construct which contained two MYC tags constructed in an effort to amplify the signal (*UAS-2MYC-d-ephrin*) (Figure 23). These constructs were made by introducing the epitope tag to the region between the start of the d-ephrin core and the signal peptide. This was done with PCR by amplifying the region upstream of the epitope tag and the region downstream of the epitope tag to introduce the epitope. These two fragments were then ligated together with a BamHI site that was introduced via PCR (Figure 23). The epitopes were inserted at this region to prevent disruption of the signaling peptide or the ephrin core region, which could perturb protein transport within the cell, or ephrin function. The N terminal tagged d-ephrin constructs were then cloned into pUAST (see Chapter 2) and fully sequenced to ensure no errors had been introduced by PCR.

When these constructs were expressed in S2 cells, none could be detected on either a western blot or directly in S2 cells, with α-mouse-6HIS, α-rabbit-6HIS, α-mouse-HA, or α-mouse-MYC (antibodies used on respective constructs as appropriate). Furthermore, none of these constructs were able to facilitate cell aggregation between S2 cells expressing an N terminal tagged *UAS-d-ephrin* construct and *UAS-Dek* expressing cells. This suggests either that the protein localisation has been perturbed by the addition of an epitope tag immediately downstream from the signaling peptide or that the protein folding of d-ephrin is such that the epitopes are not accessible to the antibody. It also suggests that protein folding/structure may be disrupted enough to prevent receptor (Dek) recognition. Also, the possibility that the

epitopes do not allow sufficient antibody to bind to allow detection cannot be ruled out, although in the case of the N-terminal-MYC tagged construct this is unlikely, due to the successful use of MYC both in western analysis (Figure 22) and directly in S2 cells (Figure 24). The most likely reason for the failure of the N-terminal tagged d-ephrin constructs to be detected is therefore a protein trafficking or a protein folding defect that was caused by the addition of the epitope to the N-terminal portion of d-ephrin.

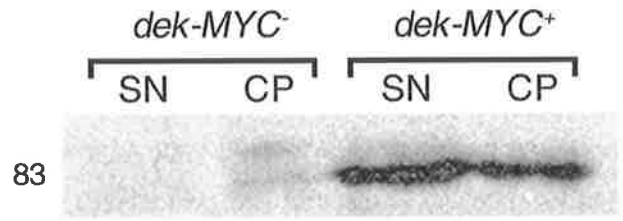


Figure 22 Western blot probed with α-mouse-MYC, showing S2 cells expressing *UAS-dek-MYC* (*dek-MYC*$^{+}$) in both the supernatant (SP) and cell pellet (CP). The non induced (*dek-MYC*$^{-}$) controls are shown on the right. Interestingly this blot shows Dek protein is present in the cell supernatent, which is unexpected, as Dek is membrane anchored. This may be due to the large amount of Dek being expressed.

The inability of any of the d-ephrin epitopes to be detected on either a western blot or in S2 cells means another method of reliably detecting d-ephrin on the cell surfaces is needed. The most effective method to facilitate this would be to generate an antibody to d-ephrin. However, due to time constraints, this was unable to be accomplished.

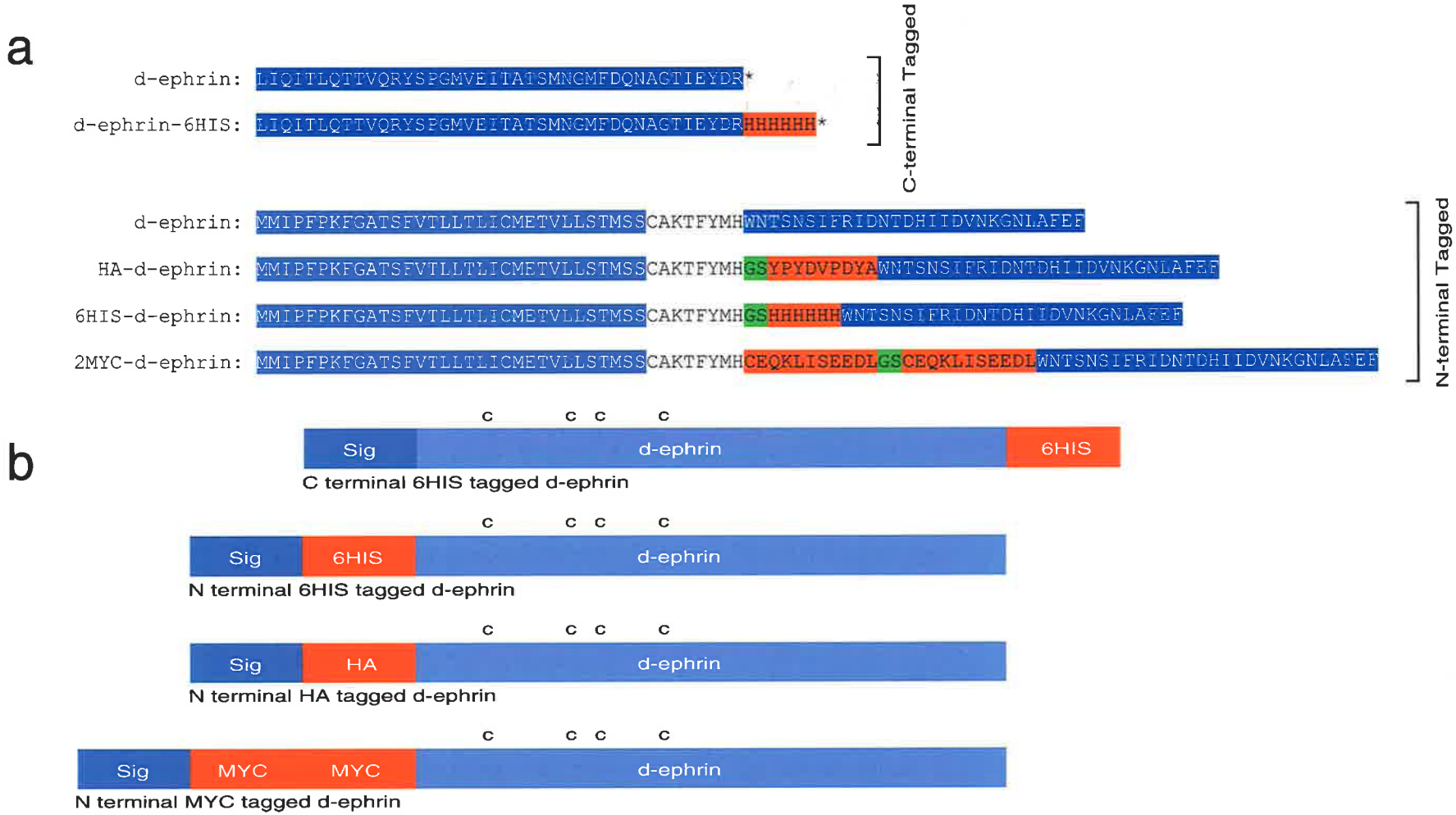


Figure 24 Epitope tagged constructs of d-ephrin. (a) Amino acids adjacent to epitope tags. Normal d-ephrin is shown first. d-ephrin was tagged with a 6HIS epitope on the COOH terminus of the open reading frame. For the N-terminal tagged versions, d-ephrin was tagged with a 6HIS, a HA, or 2MYC epitope tags on the N terminal end of the protein. The tags on the N terminal end were inserted between the signal peptide (Dark blue) and the ephrin core region (light blue), to prevent disruption of both protein transport to the cell membrane and folding of the ephrin core. The GS residues (green) result from the introduction of the BamHI site used in the ligation of the PCR products (see Chapter 2). (b) Schematic of epitope tag constructs showing relative positions of the epitope tags (red) from the Signal peptide (dark blue), and the d-ephrin ORF (light blue). The four conserved cysteines of the ephrin core are indicated (C).

S2 cell Aggregation assays

To determine if d-ephrin is a Dek ligand, the ability of d-ephrin to interact with dek-MYC was examined by expressing these proteins in *Drosophila* S2 cells (Schneider, 1972). S2 cells were chosen for this assay for a number of reasons: Firstly, they are readily transformed with transfection efficiencies of up to 10% with $CaPO_4$ mediated transfection (see Chapter 2). Also, S2 cells are relatively non-adhesive making them ideal for use in such assays. In an aggregation assay two cell lines expressing different putative interacting proteins are mixed and their ability to aggregate into large clumps of cells is assayed. This provides evidence for the interaction of the two surface proteins in question. This type of assay has been used successfully in the past. For example, this approach was used to show that fasciclin III is capable of mediating adhesion in a homophilic, Ca^{2+} independent manner (Snow *et. al.*, 1989). In another study, *Notch*-expressing S2 cells were shown to form aggregates with *Delta*-expressing S2 cells in a Ca^{2+} dependent fashion, suggesting that *Notch* and *Delta* interact on the cell surface via their extracellular domains (Fehon *et. al.*, 1990).

In the experiments outlined here, the GAL4::UAS *trans*-activation system is used to drive expression of an appropriate transgene in S2 cells. This is achieved by GAL4 mediated induction of the gene of interest in the UAS (upstream activation sequence) vector (Brand and Perrimon, 1993; Phelps and Brand, 1998). In the *in vitro* experiments outlined here GAL4 is under the control of the ubiquitin promoter, which ubiquitously drives expression of GAL4 (UB-GAL4).

Due to the lack of an available antibody or epitope tag for d-ephrin, its expression was followed by co-transfection with UAS-GFP. Cell populations with equal mixtures of *dek-MYC*-expressing cells (visualised with CY3, Figure 25a) and *d-ephrin*-expressing cells (visualised by co-transfection with *GFP*, Figure 25a) consistently form aggregates of up to 10 cells (Figure 25c-d). Importantly, a marked reduction in the number of cell aggregates is seen in cell populations expressing only the d-ephrin ligand, or only the Dek receptor (Figure 26, Table 3). Furthermore, cell populations expressing only *GFP* or *GFP*-expressing cells mixed with *dek*-expressing cells do not form aggregates (Figure 25b, Table 3).

To assay which part of the d-ephrin protein is responsible for interaction with Dek, a truncated form of d-ephrin missing the first two of the four conserved cysteine residues from the ephrin core (d-ephrin Δ2C) was used in an aggregation assay. S2 cells expressing this construct did not form aggregates with Dek S2 cells (Figure 27). The results of these two experiments suggest that Dek is able to recognise and bind to d-ephrin; giving compelling evidence that d-ephrin is a Dek ligand.

In order to determine if the binding promiscuity of the Eph/ephrins observed with other vertebrate Ephs, is observed with *Drosophila* Eph/ephrin interactions, the full ORF of both chick (c) cEphA4 (Q07496) and c-ephrin-A5 (X90377) were cloned into pUAST and expressed in S2 cells. Expression of these avian Eph/ephrins in the tissue culture assay was used to obtain a non-quantitative estimate of the relative binding affinities between Dek/d-ephrin and cEphA4/c-ephrin-A5.

cEphA4 (Chicken EphA4) expressing S2 cells form aggregates with c-ephrin-A5 expressing S2 cells (Figure 28b, d). Furthermore, d-ephrin expressing cells are able to form aggregates with cEphA4 (Figure 28c), which is not unexpected due to the promiscuous binding affinity of EphA4 which is known to bind both ephrin-B and ephrin-A type ligands. Interestingly, Dek cells do not form aggregates with c-ephrin-A5 (data not shown), which suggests that Dek prefers ephrin-B type ligands, although the reciprocal aggregation experiment (i.e. Dek assayed with a c-ephrin-B) is needed to test this.

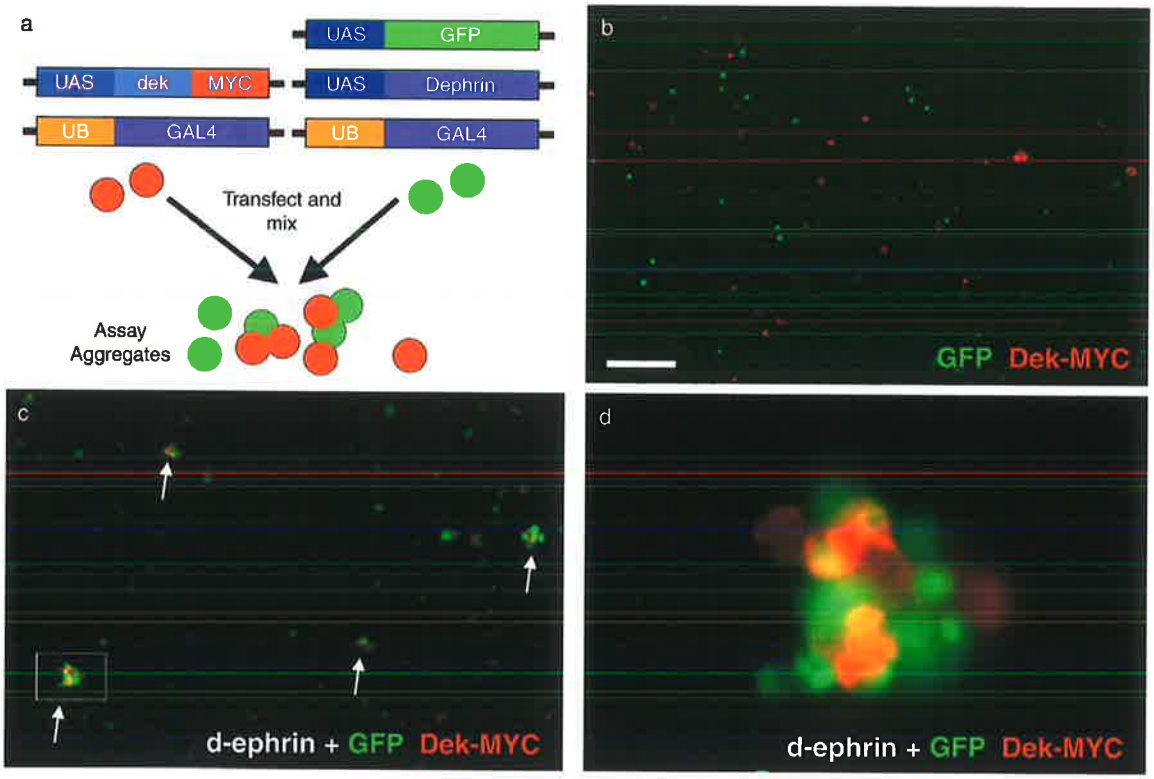


Figure 25 *dek*$^+$ Schneider cells (S2) form large aggregates when cultured with *d-ephrin*$^+$ S2 cells. (a) UAS constructs used in S2 aggregation assays. One cell line was transfected with *UAS-GFP*, *UAS-d-ephrin*, and the *UB-GAL4* driver plasmid, at a 19: 1 ratio of UAS : GAL4 The other cell line was transfected with UAS-Dek and the *UB-GAL4* driver. These expressing cell populations were then mixed and assayed for their aggregation potential. For detailed methods see Chapter 2. (b) Cells expressing *GFP* only do not form aggregates with *dek*$^+$ cell lines. Scale bar is 50μm. (c) Cell populations transfected with both *d-ephrin* and *GFP* form large aggregates with *dek* cells (arrows). (d) High resolution image of *d-ephrin* + *GFP* aggregate with *dek* (box from (c)).

	Dek	GFP	d-ephrin / GFP	GFP ~ Dek	d-ephrin/ GFP ~ Dek
Experiment 1	9%	0%	13%	3%	76%
Experiment 2	13%	0%	16%	4%	65%
Experiment 3	-	2%	-	0%	56%
Experiment 4	-	0%	-	0%	50%

Table 3 Percentage of *dek* and *d-ephrin* expressing cells aggregates of four or more. Minimum of 200 expressing cells counted. (/) Co-transfected vectors in same cell line; (~) Mixed cell lines. All UAS lines were driven with UB-GAL4 at a 19:1 UAS:GAL4 ratio.

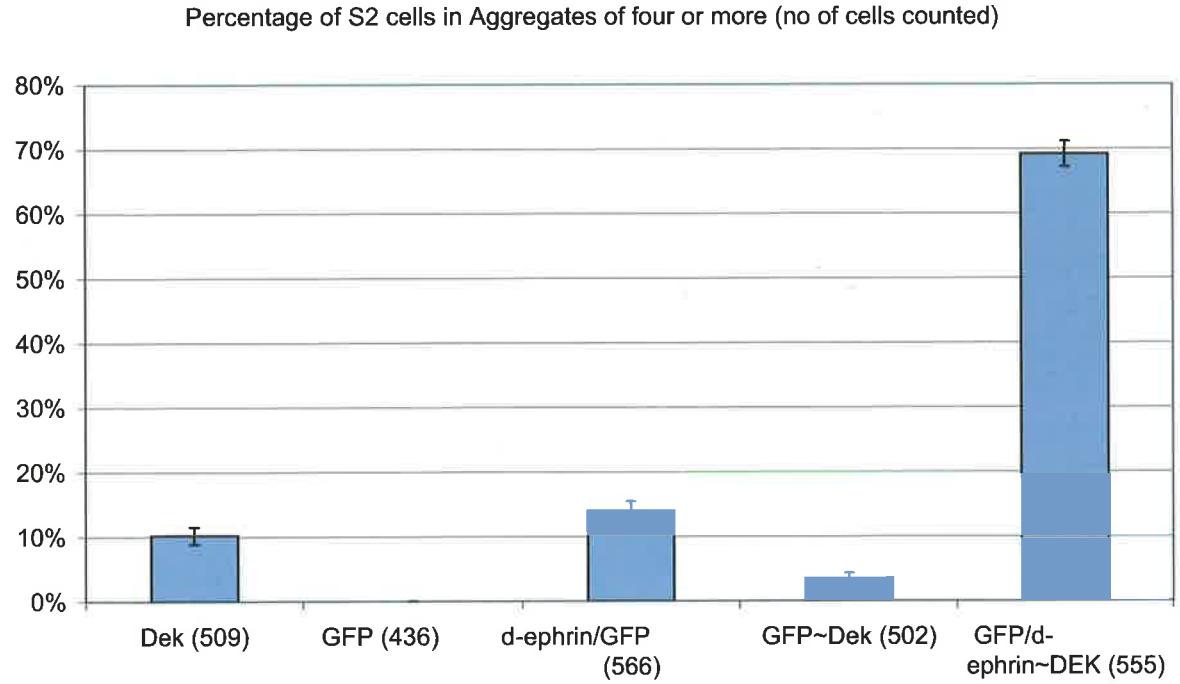


Figure 26 S2 cells expressing *dek* receptor show a significant increase in cell clumping when mixed with S2 cells expressing *d-ephrin* ligand. Combined results from two separate S2 aggregation experiments (experiments 1 and 2 in Table 3) Control experiments with cells expressing only *dek*, *GFP*, *d-ephrin/GFP*, or *GFP* cells aggregated to *dek* cells (*GFP ~ dek*) showed markedly less clumps of four or more cells, than cells expressing *GFP/d-ephrin* mixed with *dek* cells.

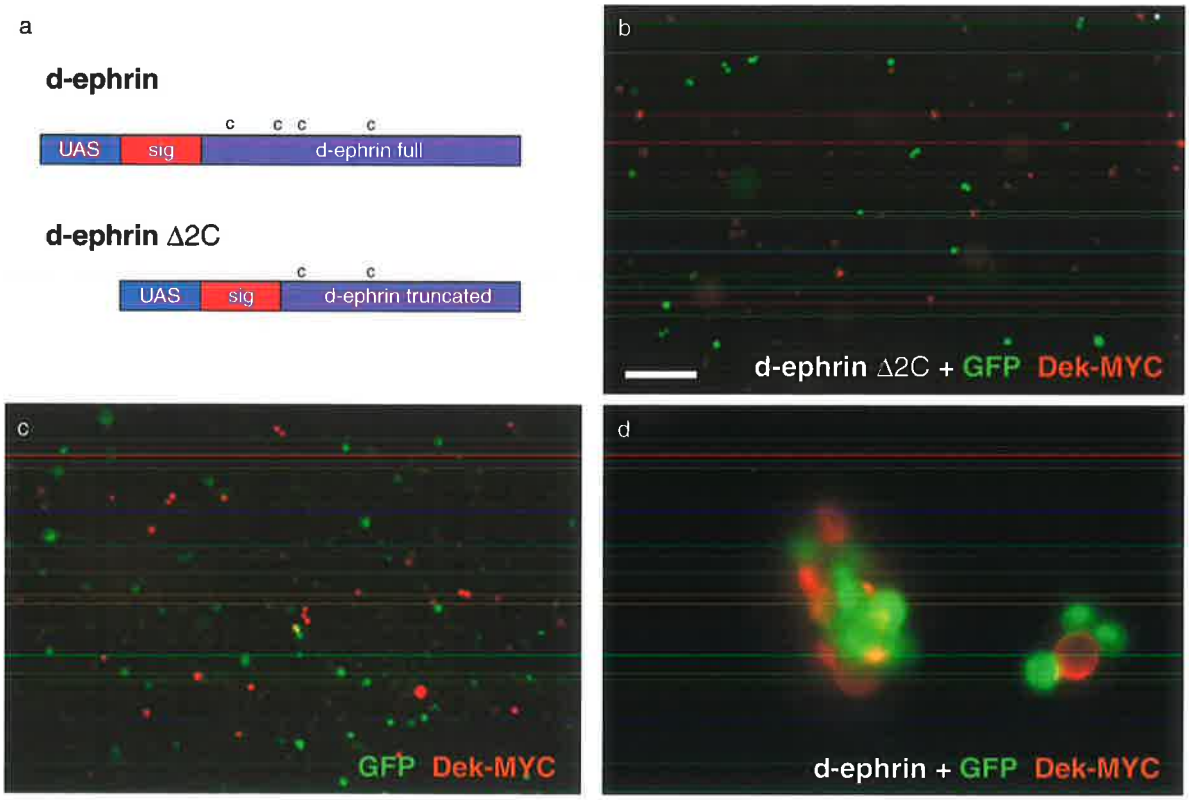


Figure 27 *d-ephrin* causes S2 cells to aggregate to *dek* cells. (a) Full length UAS-d-ephrin was modified by PCR to remove two of the four-cysteine residues in the ephrin core region while leaving the cell membrane targeting Signal peptide (sig) intact. (b) *d-ephrinΔ2C*$^+$ cells do not form aggregates with *dek* expressing cells. (c) *GFP*$^+$ cells also do not form aggregates with *dek*$^+$ S2 cells. (d) *d-ephrin*$^+$ cells form large aggregates with *dek*$^+$ cells.

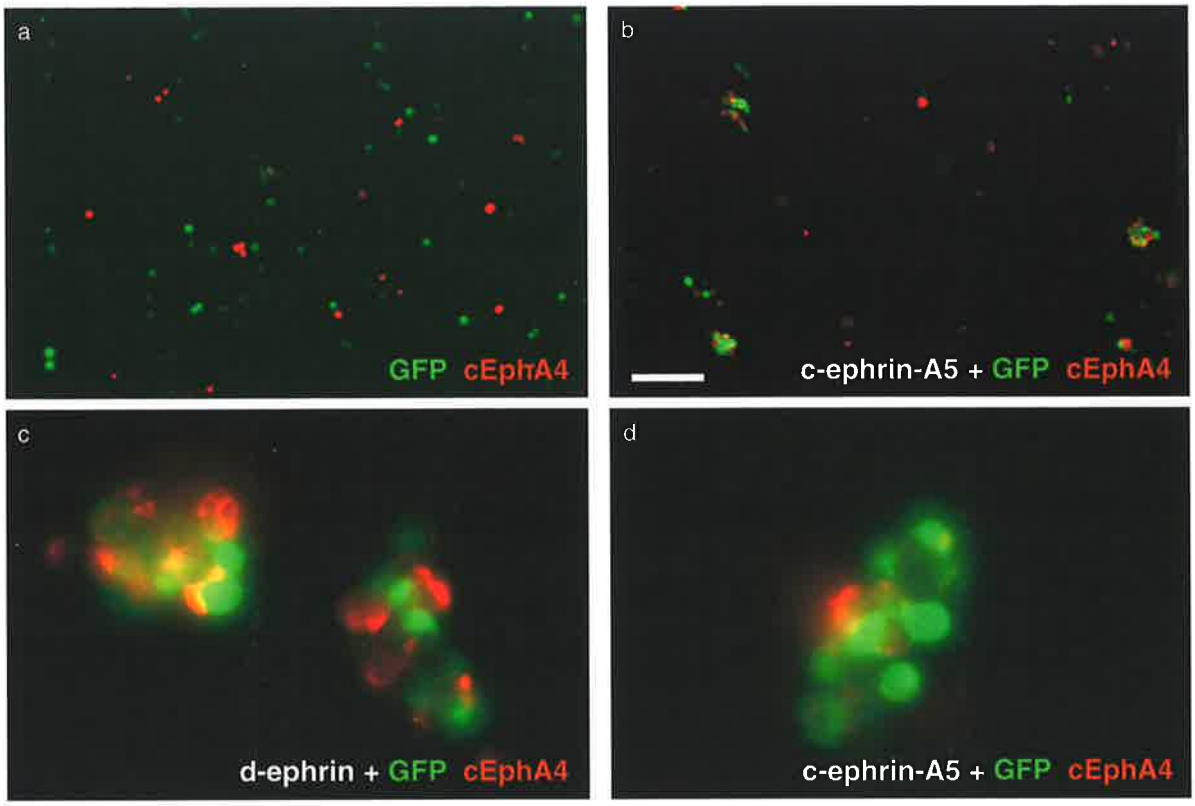


Figure 28 *d-ephrin*$^+$ cells form aggregates with *cEphA4*$^+$ expressing cells. (a) *GFP*$^+$ cells do not form aggregates with *cEphA4*$^+$ cells. (b) *c-ephrin-A5*$^+$ cells form large aggregates with *cEphA4*$^+$ cells. (c) High power image of *d-ephrin*$^+$ cells mixed with *cEphA4*$^+$ showing aggregates of up to 20 cells. (d) High power image of *c-ephrin-A5*$^+$ and *cEphA4*$^+$ cell aggregates.

Expression of *d-ephrin* during *Drosophila* embryogenesis

In order to determine whether *d-ephrin* RNA is present *in vivo* and to characterise the spatial and temporal expression pattern of *d-ephrin*, *in situ* hybridisation of whole mount embryos was performed. A 642bp fragment of *d-ephrin* corresponding to the ephrin core region (241-455aa of the open reading frame) was transcribed using T7 RNA polymerase for both the sense and anti sense orientation. This was achieved by adding a T7 promoter to the 642bp fragment on the 5` end (sense) or the 3` end (anti-sense) by PCR. These two separate PCR products were then transcribed with T7 RNA polymerase (Chapter 2, Appendix F).

In situ analysis with these probes revealed the presence of *d-ephrin* transcripts in the lateral region of the ventral nerve cord (VNC) and the brain of stage 14-17 embryos (Figure 29a, c; Figure 30). Sense controls did not show any staining above background levels (Figure 29b, d). The presence of *d-ephrin* in the developing nervous system during the stage of axonal pathfinding and refinement, suggests that it may be involved in neural system development.

Previously, *in situ* analysis done by Scully *et. al.,* (1999) had reported the presence of *dek* transcripts in the brain and cortex of the ventral nerve cord, in a pattern overlapping with *d-ephrin*. Furthermore, antibody stains with α-Dek showed that Dek antigens are present on the extending axons in the ventral nerve cord (Scully *et. al.*, 1999).

In order to confirm the co-localisation of *dek* and *d-ephrin* in the CNS, a 608bp probe corresponding to the intracellular SAM domain (848-1050aa) of *dek* was used to generate RNA probes with a T7 promoter in the sense and anti sense direction, in the same manner as the *d-ephrin* probe (see Chapter 2, Appendix F). *In situ* analysis with these probes confirmed the results seen by Scully *et. al.,* (1999), with *dek* transcripts present in the brain and ventral nerve cord (Figure 29e, g). Sense controls did not show levels of staining above background (Figure 29f, h).

The results of the *d-ephrin* and *dek in situ* experiments show that these two genes are spatially and temporally oriented in a manner consistent with ligand-receptor interactions. Taken together with the *dek* protein localisation on the projecting axons, shown by Scully *et. al.,* (1999), an *in vivo* ligand-receptor relationship between *dek* and *d-ephrin* is supported by these observations.

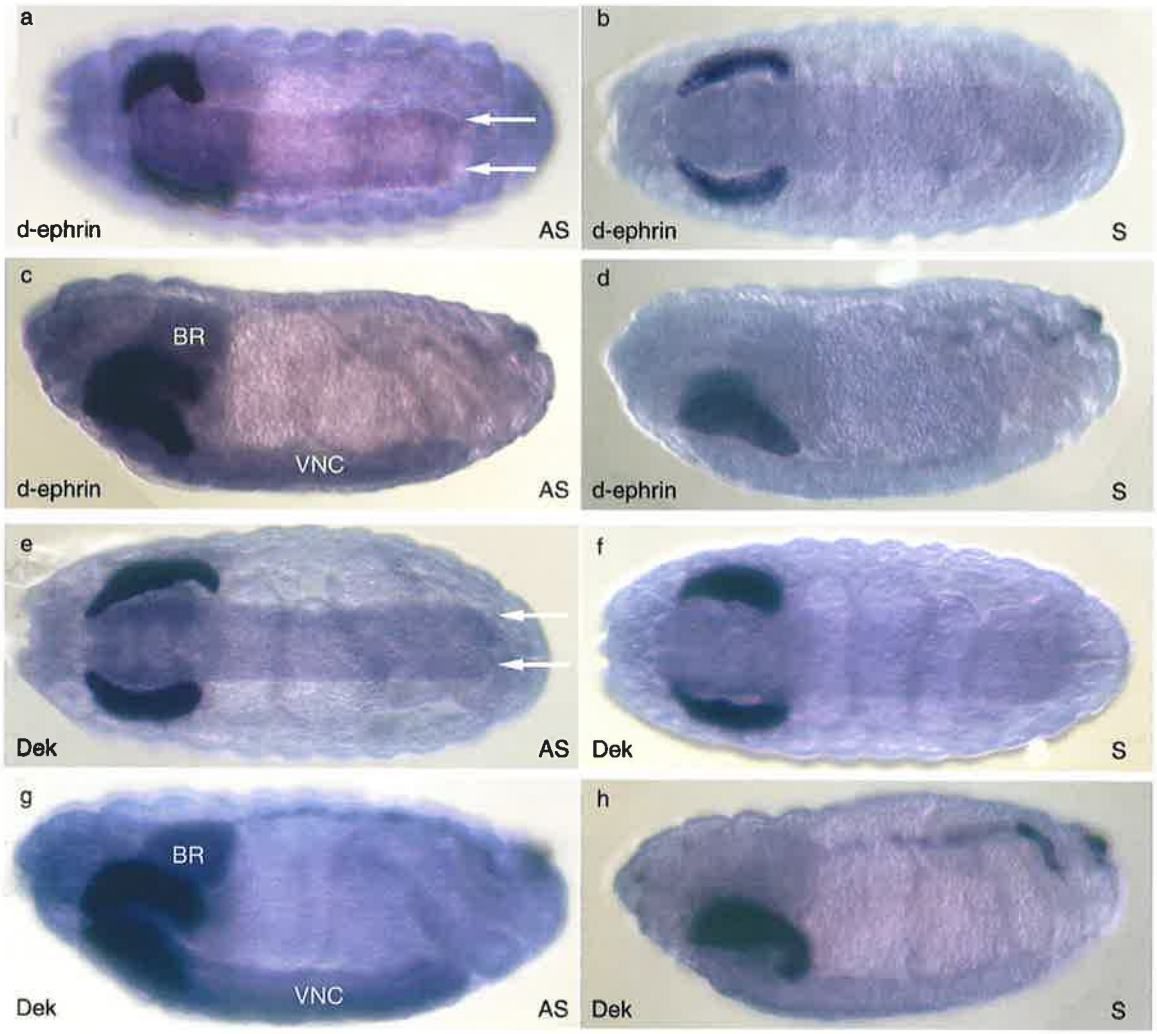


Figure 29 *In situ* hybridisation analyses of *d-ephrin* (a-d) and *dek* (e-h) in 16hr w^{1118} embryos. *d-ephrin* transcripts are present in the lateral edge of the ventral nerve cord (VNC, white arrows), and in the brain lobes (BR) (a-d). Similarly, *dek* transcripts are present in the ventral nerve cord (VNC) and brain (BR) (e-h). The dark structures present in both the sense (b, d, f, h) and anti sense (a, c, e, g) experiments are salivary glands, which frequently show artifactual staining in *in situ* hybridisation experiments. AS = Anti Sense probe; S = Sense probe. Anterior is to the left (a-h), Ventral view (a, b, e, f), lateral view (c, d, g, h).

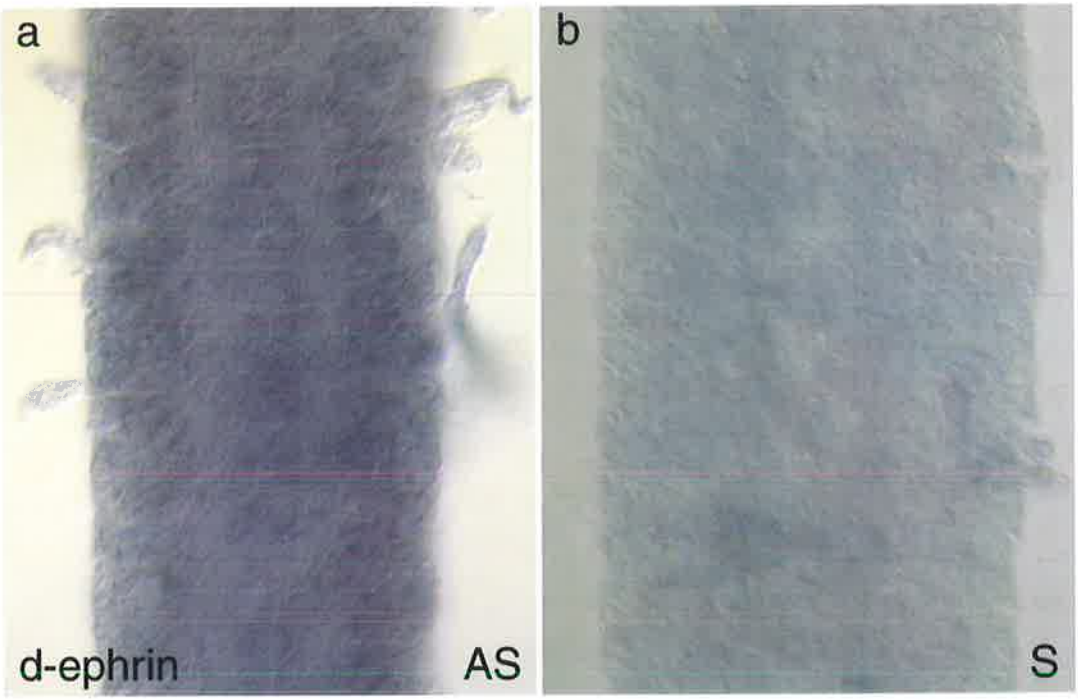


Figure 30 Dissected Ventral nerve cords from *in situ* hybridisation analyses of d-ephrin (a-b) in approximately 16-hour w^{1118} embryos shows the presence of d-ephrin transcripts in the lateral regions of the ventral nerve cord. AS = Antisense probe; S = Sense probe.

In vivo tissue specific UAS – Gal 4 over expression

In an attempt to determine whether *d-ephrin* has a function *in vivo*, the GAL4-UAS system was used to over express *d-ephrin* in neural tissues (Brand and Perrimon, 1993; Phelps and Brand, 1998). The GAL4-UAS trans-activation system is a powerful tool for misexpressing genes of interest in order to elucidate their biological function. In this system the gene of interest is placed under direct control of an *S.cerevisiae GAL4* upstream activation sequence (UAS). Upon introduction into the *Drosophila* germline by genetic transformation, expression of the gene can be induced by the presence of the yeast GAL4 protein (Brand and Perrimon, 1993). Genetic transformation in *Drosophila* is facilitated by the presence of *P*-element inverted repeats on the pUAST vector, which allow insertion of the construct into the *Drosophila* germline in the presence of *P*-element transposase (Rubin and Spradling, 1982; Spradling and Rubin, 1982). The *UAS-d-ephrin* expression constructs used in the S2 aggregation, were injected into the pole cells of *Drosophila* embryos and after subsequent crosses four independent transgenic lines were established (see Chapter 2). The chromosome in which the *P*-element was inserted in these lines was established via genetic mapping (see Chapter 2). One line mapped to the second chromosome, and three mapped to the third chromosome. All except one of these lines is homozygous viable.

Here the GAL4-UAS system was used to constitutively misexpress d-ephrin in tissue regions or cell types that should interact with *Dek*.

Misexpression of d-ephrin in midline glial cells does not perturb pathfinding

As mentioned above Scully *et. al.,* (1999) showed that *dek* expression is first detected in the neural ectoderm from stage 10, later Dek protein localises to the lateral connectives and commissures of *Drosophila* stage 16 embryos. Therefore, if *d-ephrin* ligand is ectopically expressed in the midline from stage 10, when it is not normally expressed, the Dek positive commissural axons crossing the midline may be perturbed.

In this experiment, *sim-GAL4* (Scholz *et. al.*, 1997) was used to drive expression of d-ephrin in all midline cells from stage 10 onwards. This was achieved by crossing flies homozygous for *sim-GAL4* with flies homozygous for *UAS-d-ephrin*. This cross was carried out at 30°C to provide higher levels of expression of the *sim-GAL4* thereby increasing the d-ephrin protein levels present in the midline. The progeny of this cross was collected after 17 hours (giving 0-17hr embryos), fixed and subsequently stained with the monoclonal antibody BP102, which stains neural cell membranes by recognizing a nervous system-specific isoform of the neuroglian protein.

Analysis of *sim-GAL4*::*UAS-d-ephrin* embryos stained with BP102 showed normal formation of the anterior and posterior commissures, and longitudinal connectives, at all stages when compared to w^{1118} embryos (Figure 30). Furthermore, embryos from this cross were viable through to adult flies.

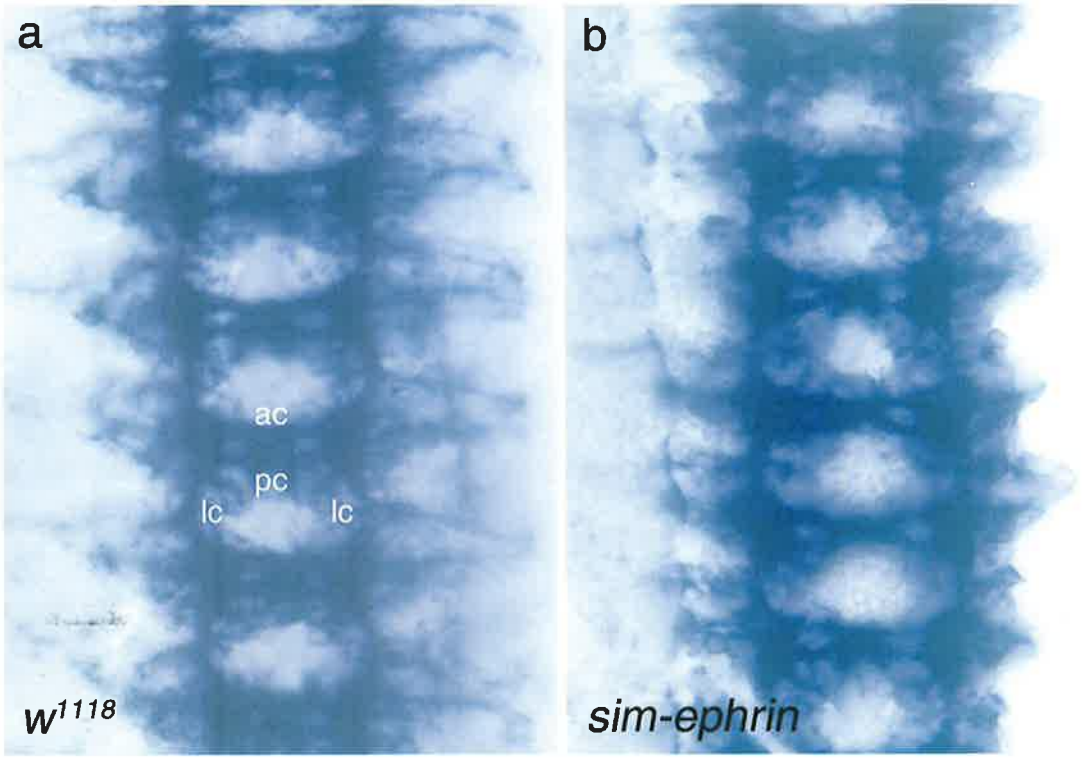


Figure 31 Embryos expressing *d-ephrin* in midline glial cells show no significant defects in CNS development. Flat preparation of a stage 13 wild type ventral nerve cord (VNC) stained with the monoclonal antibody BP102, which labels anterior commissure (ac), posterior commissure (pc), and longitudinal connectives (lc). (a) wild-type embryo. (b) *sim-GAL4*::*UAS-d-ephrin* embryo. Expression of *d-ephrin* in the midline cells does not give any significant perturbation of axons crossing the midline. Anterior is up.

Ectopic expression of d-ephrin in all neural cells does not perturb axon pathfinding

The *in situ* patterns for *d-ephrin* and *dek* suggest that this ligand and receptor may be involved in development of axonal scaffold, as are their vertebrate counterparts (Chapter 1). Therefore, misexpression of either *dek* or *d-ephrin* in the VNC at the time of axonal pathfinding may result in a perturbation of the regular segmental pattern of the *Drosophila* VNC.

Here, *elav-GAL4* was used to drive expression of d-ephrin in all cells that have adopted a neural cell fate (Lin and Goodman, 1994; Luo *et. al.*, 1994), in order to examine the effect of over expressing *d-ephrin* during differentiation of neural tissues. The *elav-GAL4* line used in this experiment is on the X chromosome, therefore *elav-GAL4* was introduced by crossing homozygous females with male *UAS-d-ephrin*, ensuring that all progeny are *elav-GAL4*::*UAS-d-ephrin*. Embryos from this cross were analysed with the monoclonal antibody 22C10, which stains the inner surface of the cell membrane of all PNS neurons and a subset of CNS neurons.

Embryos from this cross, stained with 22C10, showed no gross morphological defects with segment boundaries appearing normal. However, within the nervous system itself their appears to be a misrouting of some axons in *elav-GAL4*::*UAS-d-ephrin* embryos compared with w^{1118} embryos (Figure 32). The intersegmental nerve roots of *elav-GAL4*::*UAS-d-ephrin* embryos appear to fuse more proximal than those of wild-type embryos (Figure 32a,b). Furthermore, the lateral connectives of *elav-GAL4*::*UAS-d-ephrin* embryos appear to route further from the midline of the VNC than the lateral connectives of wild-type embryos (Figure 32a,b). In the peripheral nervous system (PNS) the gross morphological structure of neuron cells appears relatively normal, although some thickening of cell clusters is apparent (Figure 32c,d), embryos from this cross were viable, through to adult flies.

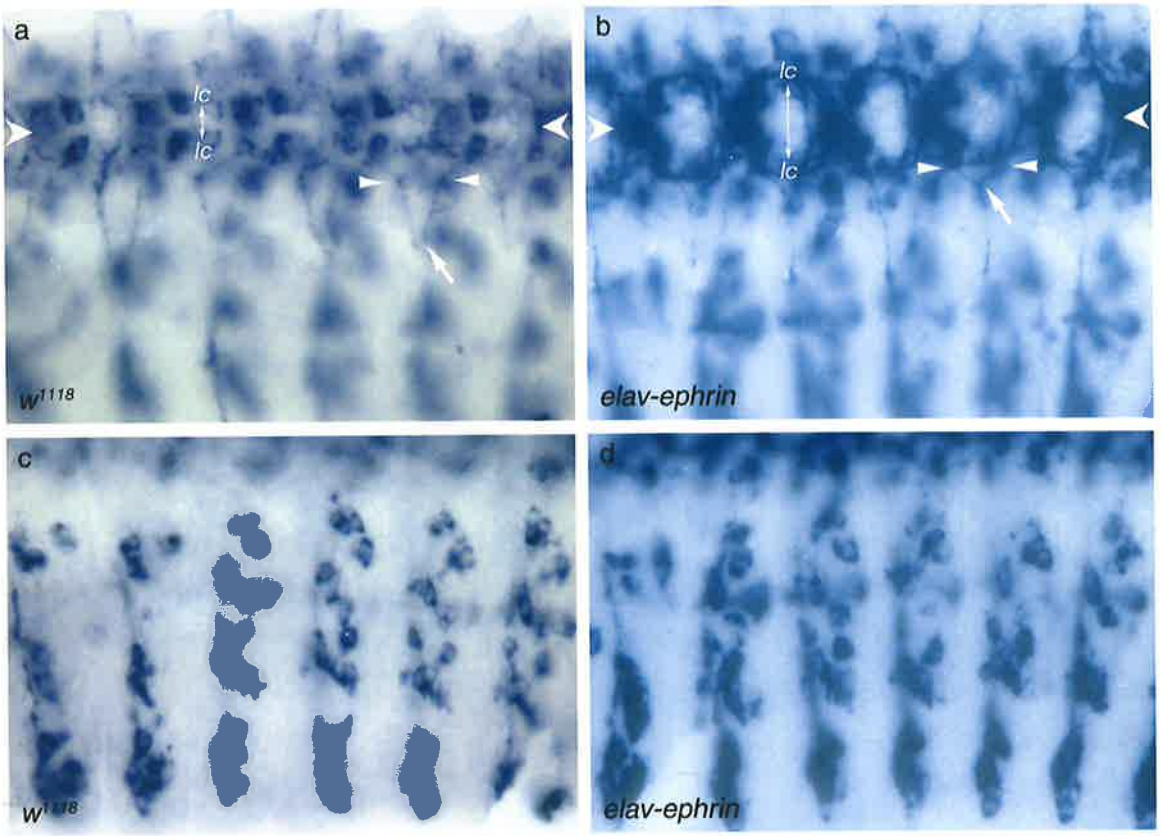


Figure 31 Expression of d-ephrin in all neural cells shows a mild disruption of lateral axonal exit points in the peripheral nervous system (PNS). (a) A wild type stage 13-14 embryo stained with the 22C10 antibody, showing the central nervous system (CNS). There are two intersegmental nerve roots (white arrowheads) fusing to form one axonal tract at the very lateral edge of ventral nerve cord (VNC) (white arrow) (b) A stage 13-14 embryo *elav-GAL4* :: *UAS-d-ephrin* embryo: ectopic expression of d-ephrin in differentiating neural cells appears to cause misrouting of pioneer axons which results in the longitudinal connectives (lc) forming further from midline than wild-type (double headed arrow). Moreover, the two intersegmental nerve roots (white arrowheads) fuse more proximally relative to midline (white arrow). (c) A lateral view of the same stage 14 embryo shown in (a) with the PNS in focus. (d) A lateral view of an *elav-GAL4* :: *UAS-d-ephrin* embryo shown in (b) showing the PNS: there appears to be no dramatic phenotype resulting from the over expression of *d-ephrin* in all neural cells. Some cell clusters within each segment do appear slightly larger. Ventral is up. Anterior is left.

Discussion

This chapter aimed to elucidate the function of d-ephrin *in vivo*. The S2 aggregation assays shown in this chapter demonstrate that d-ephrin binds to Dek, indicating that d-ephrin is a putative ligand for Dek. This was confirmed by removing two of the four conserved cysteine residues of d-ephrin. Unlike its wild-type counterpart, this extracellular truncated version of d-ephrin was unable to induce aggregation with Dek in S2 cells.

In separate S2 aggregation experiments d-ephrin, was shown to be able to bind chicken EphA4. This is not surprising given the well-characterised binding promiscuity of EphA4 EphA4. Also it was shown that Dek was unable to bind chicken ephrin A5, suggesting that Dek cannot bind ephrin-A type ligands, although, it could be that other ephrin-A ligands can be recognised by a *Drosophila* Eph (Dek). It would be interesting to investigate the reciprocal experiment, to see if Dek could recognise any of the known ephrin-B orthologs. If Dek was found to recognise an ephrin-B ligand, this may suggest that d-ephrin is an ephrin-B type ligand.

Unfortunately, d-ephrin could not be detected by tagging with an epitope, which means that the posttranslational size of the protein could not be determined. If the size of the protein could be determined, the nature of any post-translational cleavage might be able to be inferred. Furthermore, given the ability to detect the extracellular region of the d-ephrin core, the question of GPI or transmembrane anchorage could be addressed by testing the ability of GPI-specific PI-PLC to remove the d-ephrin extracellular region from the surface of S2 cells. This type of assay has been used in the past to confirm GPI anchorage in most of the currently known ephrin-A proteins (Winslow *et. al.*, 1995; Kozlosky *et. al.*, 1997; Davy *et. al.*, 1999).

The analysis of *d-ephrin* using *in situ* hybridisation showed that it is expressed in the ventral nerve cord and the embryonic brain. The localisation of d-ephrin to the nervous system of *Drosophila* suggests that it could play a role in axon guidance, like the vertebrate ephrin orthologs. Furthermore, the *d-ephrin* expression pattern parallels that of *Drosophila* Eph kinase (Dek) (Scully *et. al.*, 1999). This suggests a ligand receptor relationship between the d-ephrin and Dek. Interestingly, Dek protein is localised to the developing axons of *Drosophila* embryos. It would be useful to analyse where the d-ephrin protein localises in the developing *Drosophila* embryo. D-ephrin may localised to the cortex region of the ventral nerve cord (VNC), which would be expected if it was providing a repulsive mechanism to prevent axons from aberrantly exiting the VNC.

The misexpression of d-ephrin in the midline with *sim-GAL4* did not perturb the longitudinal connectives or commissures of the VNC. This result is unexpected, as *sim-GAL4* should be driving expression of *d-ephrin* from stage 10 onwards, which should preface *dek* expression from stage 14 onwards. Therefore, an abundance of d-ephrin protein should be present on the midline of the VNC, when commissural axons are crossing the midline.

It is possible that the levels of ectopic expression of *d-ephrin* with *sim-GAL4* were not sufficiently high to perturb axon pathfinding *in vivo*. An alternative approach would be to generate *Drosophila* lines with more than one copy of the UAS-*d-ephrin* construct, which may increase the levels of d-ephrin protein when driven with *sim-GAL4*, thus enhancing any mild phenotype. It is also possible that *sim-GAL4* does not ectopically express *d-ephrin* at a stage, which corresponds to endogenous *dek* expression. Confirmation of this could be obtained by *in situ* analysis or antibody stains against d-ephrin, to ensure that in these experiments *sim-GAL4* is driving *d-ephrin* appropriately.

The premature fusion of the intersegmental nerve roots upon misexpression of *d-ephrin* in all neural cells with *elav-GAL4* could be explained by axons that endogenously express *dek* becoming desensitised to endogenous *d-ephrin* present in the periphery of the VNC, due to the ectopic expression of *d-ephrin* on the same axons. This is similar to the mechanism of axonal desensitisation seen in the chicken retinotectal map (Drescher *et. al.*, 1995), where chicken retinal axons co-express EphA4 receptor and ephrin-A5 ligand (Chapter 1). The result of this is that axons appear to be clustering together soon after leaving the CNS, effectively fusing the intersegmental nerve roots, as opposed to fusion at the periphery of the VNC.

The apparent misrouting of the aCC and pCC axons in the CNS may also be explained by desensitisation of axons endogenously expressing *dek* to ectopic levels of *d-ephrin* on the same axons. Endogenous *d-ephrin* is present on the lateral VNC (In situ), it appears that axons have stopped at the border of endogenous d-ephrin protein. This suggests that pioneering axons may be still sensitive to repulsion from the endogenous d-ephrin. Another mechanism, which defines the trajectory of pioneer axons relative to the midline, is repulsion between Slit and Robo (Roundabout) (Rajagopalan *et. al.*, 2000). The position of first axonal track suggests that overexpression of d-ephrin in the developing neurons somehow overrides repulsion via Robo receptor tyrosine kinase.

The longitudinal projections in the *elav-GAL4::UAS-ephrin* embryos seem to be slightly thicker than in wild-type embryos, which may be a result of increased axonal branching. This can be tested by single cell DiI labelling, a method used successfully to show an increase in axonal branching in *robo* embryos (Murray and Whitington, 1999).

While these results do show a potential phenotype, the expression levels of *d-ephrin* need to be tested in the misexpressed embryos, with either antibody stains or *in situ* analysis to increase confidence in the phenotypes shown, due to time constraints this was not done.

Work by Scully *et. al.,* (1999) found no significant phenotype when misexpressing *dek* with various GAL4 drivers, including *elav-GAL4* and *A80-GAL4* (which drives expression in a subset of midline glial cells, similar to that of *sim-GAL4* driver used in this study). Scully *et. al.,* (1999) also generated *Drosophila* lines which contained multiple copies of *UAS-Dek* and *UAS-Dek-MYC*. These lines also showed no significant phenotype when driven with $elav^{c155}$, *sca*, *elav*, ftz_{ng}, *A80*, and *GMR*. However, a more recent study by Dearborn *et. al.,* (2002), has shown that *dek* is involved in *Drosophila* eye development. Dearborn *et. al.,* (2002) used both RNAi and misexpression analysis with the UAS-GAL4 system to modify *dek* expression in *Drosophila* eye development. Disruption of *dek* expression in both photoreceptor and medulla cortical cells results in defects in the axon projections of these cell types (Dearborn *et. al.*, 2002). Expression analysis of *d-ephrin* by Dearborn *et. al.,* (2002) also shows that it is present in the developing eye disc, as would be expected if it is a ligand for *dek*.

These two studies suggest that *dek* may only play a role in the developing eye and might have a redundant function the VNC. However, the results shown in this chapter indicate that d-ephrin and Dek interactions may play a role in the formation of the VNC. It would be interesting to pursue the role of *d-ephrin* in the VNC with misexpression studies and RNAi to try to shed some light on the role of *dek/d-ephrin* interactions in the developing *Drosophila* VNC.

Chapter 5: Molecular phylogenetic analysis of the ephrin gene family

Molecular phylogenetic analysis of the ephrin gene family

Introduction

The results presented so far were unable to determine whether *d-ephrin* is an ephrin-A or an ephrin-B. Discussed in this chapter is another approach to determining the relationship of the invertebrate ephrins to the vertebrate ephrins. These results also clarify the current nomenclature of all known ephrin orthologs.

Based on cell binding interactions and receptor activation studies the Eph/ephrin molecules were classified into EphA and EphB subclasses which bind to the ephrin-A and ephrin-B subclasses respectively (Gale *et. al.*, 1996b; reviewed in Gale and Yancopoulos, 1997). Subsequently, the Eph nomenclature committee assigned orthologous Eph/ephrin gene names to the large number of Eph/ephrin genes known at the time, by sequence similarity of the proteins (Eph Nomenclature Committee, 1997; Lemke, 1997). While this approach is not necessarily incorrect, phylogenetic analysis is a more rigorous method for inferring phylogenies from molecular sequence data. Using a phylogenetic approach, all useful evolutionary data contained within an aligned set of sequences can be used to infer an evolutionary tree, which represents a "best estimate" of the evolutionary history of the genes being analysed. Therefore, In order to better understand the phylogenetic relationships between the ephrin genes and to provide support for a robust system of gene and protein nomenclature, a phylogenetic analysis was performed on all currently known members of the ephrin gene family (Appendix A).

Two popular methods for carrying out such a phylogenetic analysis were used in this analysis, maximum parsimony and a bayesian likelihood approach

Maximum parsimony involves the identification of a phylogeny, that minimises the number of nucleotide substitutions required to explain the observed differences in the data (outlined in Page and Holmes, 1998; Graur and Wen-Hsiung, 1999). This approach is commonly used because it is computationally favourable over other phylogenetic inference methods. The confidence levels of the branching order in the maximum parsimony tree can then be assessed using techniques such as Bootstrapping, which is a technique that measures the robustness of the most parsimonious tree, but also allows a distinction between equally parsimonious trees. This is achieved by random sampling from the original dataset to produce pseudo-datasets, which are then applied to maximum parsimony analysis to produce phylogenetic trees. The trees produced by the analysis of many pseudo-datasets are then combined to form a consensus tree, which will consist of groupings represented in the greatest number of phylogenies.

Maximum likelihood trees are derived by searching for the most probable tree, given the observed dataset, based on a model of character state changes (e.g. nucleotide sequences). The advantage with this technique is many different evolutionary models can be applied to the same data set (outlined in Page and Holmes, 1998; Graur and Wen-Hsiung, 1999). Maximum likelihood methods have proven very effective in molecular evolutionary analyses. However, the computational requirements can be prohibitive, particularly if a bootstrapping method is employed to estimate confidence levels. In recent years a Bayesian approach to likelihood analysis has allowed a more computationally feasible approach to estimating likelihood trees. Bayesian analysis determines the posterior probabilities of parameters; the posterior probability is a fraction of the prior distribution and the likelihood distribution, which allows the identification of the most probable value for any parameter, including tree topology (Huelsenbeck *et. al.*, 2001; Simon and Larget, 2001). Bayesian analysis involves the estimation of a posterior distribution, using a Markov-chain Monte Carlo (MCMC) algorithm for a given data set (Yang and Rannala, 1997; Larget and Simon, 1999). All inferences, including tree topology and branch lengths, are drawn from the posterior distribution. The probability of a given parameter, such as branching order, is based on its relative likelihood; if all the most likely trees share a branching order, this order has a high probability. If two or more different branching orders are nearly equally likely, the probability of the most likely order is diminished (Huelsenbeck *et. al.*, 2001; Simon and Larget, 2001).

Results and Discussion

A protein alignment of the ephrin core region (PFAM 00812), which is the most conserved between the two subclasses, was constructed with Clustalw (Thompson *et. al.*, 1994). This protein sequence alignment (Figure 32) was used to generate a multiple nucleotide sequence alignment with the corresponding ephrin cDNA sequences (Appendix A), using the program Tip/Top (Kortschak, 2000).

A Bayesian tree was identified using the program BAMBE (Simon and Larget, 2001) with three randomly seeded runs of 1,000,000 iterations (see Chapter 2). The final unrooted tree (Figure 34) was the result of the summation of all three independent sampling iterations.

The Maximum parsimony tree describing the same data set (see Chapter 2), had an identical topology to that of the BAMBE tree, with bootstrap support > 70% (Figure 33).

Excluding those isolated from *Drosophila melanogaster* and *Ctenophorus elegans*, the ephrin genes were arranged into eight monophyletic clusters indicated using different colours in Figure 34. Each group contains one of the human ephrins A1, A2, A3, A4, A5, B1, B2 and B3. This finding suggests that gene and protein nomenclature that has previously been

applied is generally supported by the molecular phylogeny. The single exception is chicken ephrin-A6, which groups with the human and mouse ephrin-A4s with high probability, suggesting that it should be renamed chicken ephrin-A4. Within each of the eight monophyletic groups the phylogenetic arrangement, in general, is in accord with known species relationships. This would be expected if the genes within each group were orthologous. The exceptions to this generalisation (eg the arrangement of the *Danio rerio* and *Xenopus laevis* ephrin-B3 genes) tend to be accompanied by low probabilities.

The two major subgroups of genes, the ephrin-A and ephrin-B genes, form distinct clades that show reciprocal monophyly. That is to say, the ephrin A genes are more closely related to one another, than they are to any of the ephrin B genes, and *vice versa*. This finding is not unexpected, given the known structural and functional differences between the ephrins in these major groups. Within the ephrin-A group, the phylogenetic arrangement is (A4+A1)((A3)(A2+A5)). As there are just three members of the ephrin B group, only a single unrooted tree is possible ((B1)(B2)(B3)).

Also of interest in this analysis is the relationship of the invertebrate (*Drosophila melanogaster* and *Ctenophorus elegans*) ephrins with their vertebrate counterparts. The membrane anchorage mechanism of d-ephrin remains unclear, with two potential transmembrane regions and a potential GPI modification site, albeit both requiring posttranslational modifications before they could function (discussed earlier). The *Ctenophorus elegans* ephrin orthologues are GPI anchored. However the genes for these proteins, and the d-ephrin gene, show no closer evolutionary affinities to the ephrin A or B genes (Figure 33, Figure 34). This could suggest that the invertebrate ephrins are equally diverged from the ephrin genes now found throughout the vertebrates and perhaps similar to the ancestral prototype gene. However, the caveat to this hypothesis is that all the invertebrate ephrins are placed with low bootstrap values and posterior probabilities in the maximum parsimony and BAMBE trees respectively. Thus the positions of the invertebrate phyla remain largely unresolved in this evolutionary analysis. Unfortunately, this also means that no conclusions can be drawn as to the potential anchorage mechanisms of d-ephrin using these data. However, investigations into the anchorage mechanisms of the *Drosophila* ephrin could provide some insight as to the origins of ephrin B and A molecules.

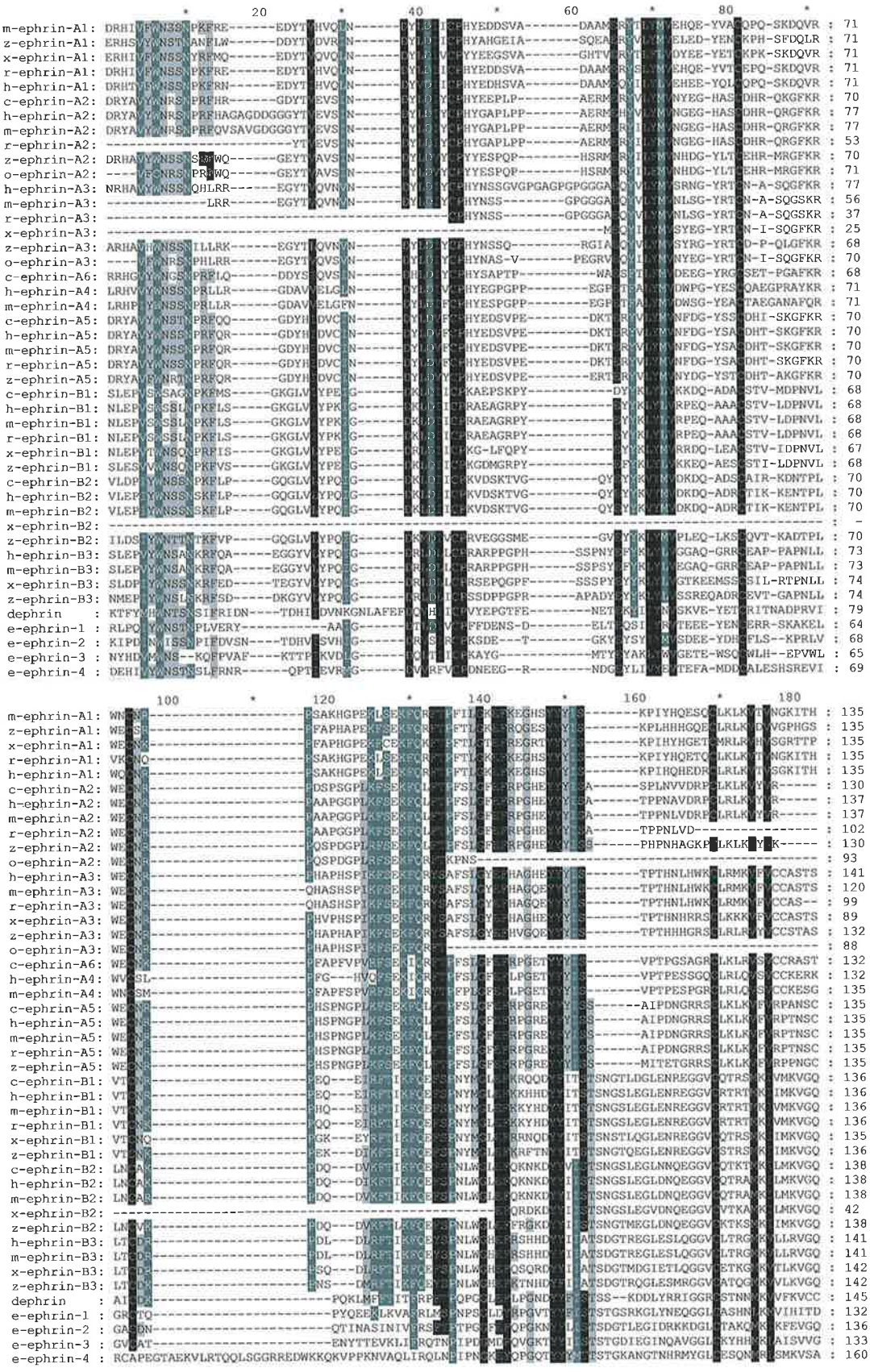


Figure 33 Protein sequence alignment of the ephrin core region generated using a profile Clustalw (Thompson *et. al.*, 1994). An alignment of the corresponding cDNA sequences, based on the protein alignment, was produced using the program Tip/Top (Kortschak, 2000). The cDNA sequences were then used as data for phylogenetic analysis using both maximum parsimony and maximum likelihood techniques.

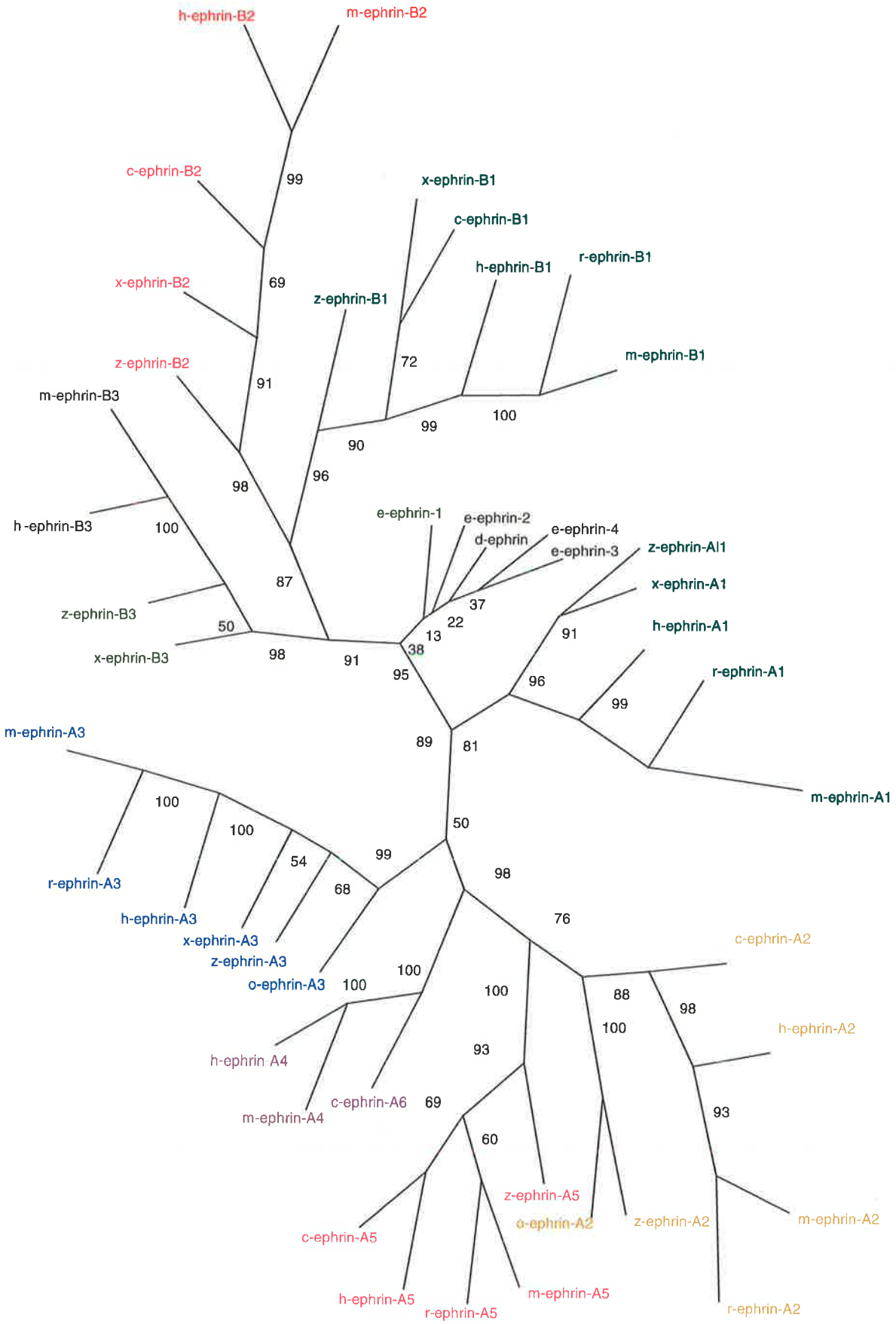


Figure 34 A consensus Maximum Parsimony Tree of all currently know ephrin orthologs, generated from 1000 bootstrap datasets using DNAPARS from the PHYLIP package (Felsenstein, 1993). Branch Lengths are not drawn to scale.

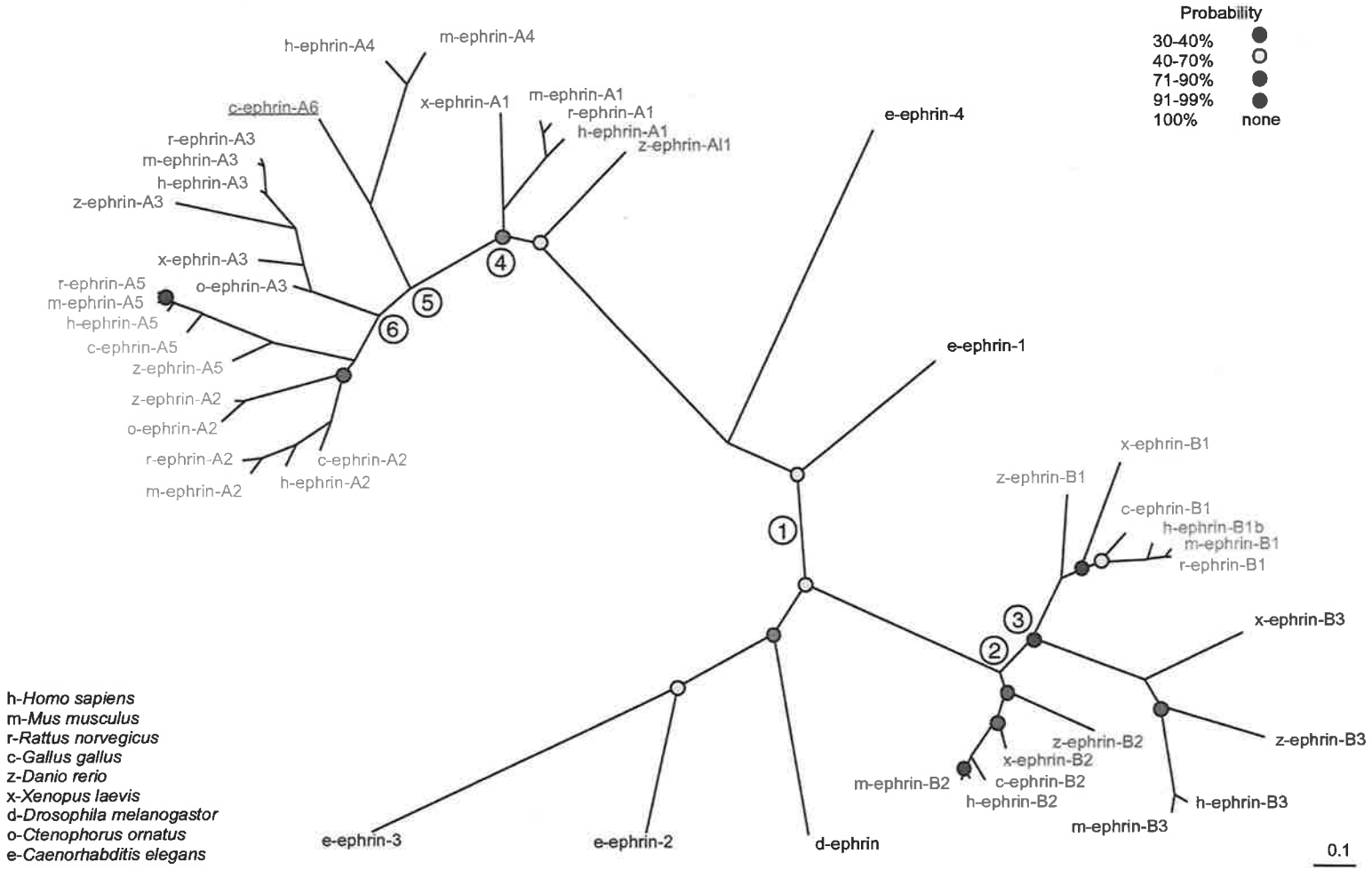


Figure 36 Phylogenetic tree based on cDNA alignments of all the currently known ephrin core region orthologs aligned with a profile Clustalw, branch lengths are drawn to scale and posterior probabilities are colour coded according to the probabilities. Individual paralogous groups are colour coded. Possible duplications are shown with white circles

Chapter 6: Genomic organization and analysis of the X-linked *ephrin-B1* gene as a candidate for Aicardi syndrome

Genomic organization and analysis of the X-linked ephrin-B1 gene as a candidate for Aicardi syndrome.

Introduction

Aicardi syndrome was first described in 1965 (Aicardi *et. al.*, 1965). This disease is considered to be an X-linked dominant disorder because it has been reported in more than 200 females and one male who had an XXY karyotype (Ropers *et. al.*, 1982). Patients with this disorder typically present with a distinct collection of symptoms, namely callosal agenesis (Figure 35), infantile spasms, retinal pigmentary defects, and cleft lip and palate. Survival into adolescence or early adulthood is rare in Aicardi's syndrome with death usually the result of pulmonary infection, which may be aggravated by kyphoscoliosis. Aicardi's Syndrome is not due to a deficiency mutation but results from a functionally altered gene product (Ropers *et. al.*, 1982). One potential Aicardi's patient has been shown to contain a balanced *de novo* translocation X/3, suggesting that the affected gene is located on the short arm of the X chromosome (Ropers *et. al.*, 1982). However, this case has been disputed as truly representing Aicardi syndrome, due to the lack of the diagnostic chorioretinal lacunae (Aicardi and Chevrie, 1994). Most non-disputed Aicardi's patients show normal karyotypes, making the disease not amenable to genetic analysis. Molina et al (1989) showed the first examples that Aicardi's syndrome can be familial, with two non-zygotic sisters presenting with the disorder.

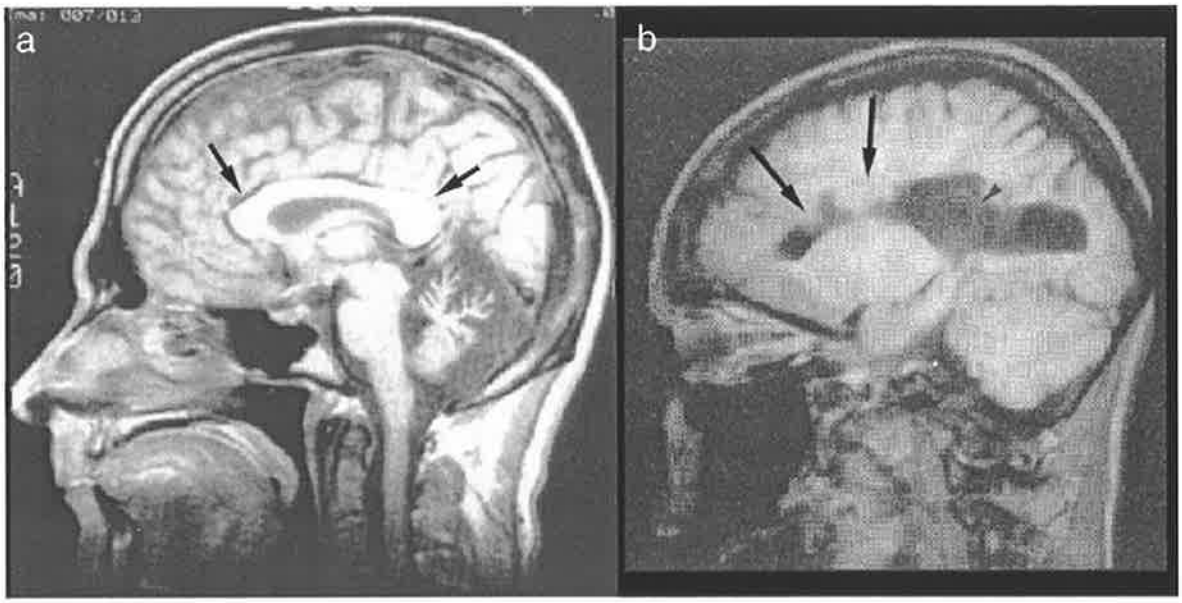


Figure 35 Magnetic Resonance Imaging of a sagittal Section through the human brain showing the different degrees of callosal agenesis. (a) The normal corpus callosum is seen as a large white structure (arrows). (b) An MRI showing agenesis of the corpus callosum, which is typical in Aicardi patients (arrows) (after Lassonde and Jeeves, 1994).

Ephrin-B1 as a candidate for Aicardi Syndrome

Genetic evidence taken from mouse knockout models suggests that *ephrin-B1* is involved in the development and/or formation of the corpus callosum (Henkemeyer *et. al.*, 1996) (discussed in Chapter 1). Mice deficient for either of the two Eph receptors, EphB2 and EphB3, also exhibit callosal agenesis and cleft palate (Orioli *et. al.*, 1996). Mice deficient in EphB2 alone exhibit defects in anterior commissure axon path finding (Henkemeyer *et. al.*, 1996), while *EphB3* mutants have defects in callosal formation (Orioli *et. al.*, 1996). However, mice that are defective in both EphB2 and EphB3 show complete callosal agenesis and a subsequent block in commissural axon migration between the hemispheres of the brain. The *h-ephrin-B1* locus, which is a known ligand to the EphB2 and EphB3 receptors, is the only known ephrin (or indeed Eph) to be located on the X chromosome (Xq12). Furthermore ephrin-B1, has been shown to be expressed on the commissural axons (Orioli *et. al.*, 1996).

The other two known ligands of EphB2 and EphB3 are ephrin-B2 and ephrin-B3 and the loci for these are located on chromosomes 13 and 17 respectively. Since ephrin-B1 has been shown to be involved in corpus callosum development in the mouse and is located on the X-chromosome, it is an ideal candidate for the Aicardi's syndrome locus. This is further supported by the fact that mice deficient in EphB2 and EphB3 demonstrate an Aicardi's-like phenotype implicating ephrin-B1 or another Eph/ephrin as a causative gene of this disease.

The results presented here attempt to determine if a mutation in *ephrin-B1* is the principle disease gene for Aicardi's syndrome, to achieve this the genomic structure of the human *ephrin-B1* gene was first determined.

Results

Genomic DNA was prepared from blood samples from six Aicardi's Patients. PCR primers were designed for the human exon boundaries of *h-ephrin-B1* (NM_004429) based on mouse intron exon boundaries for *m-ephrin-B1* (NM_010110), which shares 89% coding region identity with the human homologue (Figure 36). Reaction conditions for these primers were optimised for $MgCl_2$ concentration and annealing temperature (Table 4). Cycling parameters were: 94 ^{0}C, 120 sec; 94 ^{0}C, 30 sec; T_A ^{0}C, 60 sec; 74 ^{0}C, 60 sec; 35 cycles. Introns 2-4 were amplified using exon - exon primer combinations. Sequence data for the coding regions and small introns (five exons and introns 2-4) of six patients with Aicardi's syndrome was compared to normal patients. No base changes in any of the five coding regions of the *h-ephrin-B1* gene or introns 2-4 were found, when compared to normal patients. Reactions were all sequenced individually and used to construct a contig to elucidate the genomic structure of *h-ephrin-B1* (Figure 36). This was subsequently confirmed when the human genome project was released. Intron exon/boundary sequences are shown in Table 5. The genomic structure of *ephrin-B1* was similarly determined by constructing a contig using sequence data generated in this investigation as well as database entries for genomic *h-ephrin-B1* regions and cDNA sequences (data not shown). The *ephrin-B1* sequence extends over 12kb of genomic DNA.

Exon	PCR product length (bp)	Name	Forward Primer	Name	Reverse Primer	Annealing Temperature (°c)
1	187	1F	AGGCGAGCTTTGGTGAGGAG	1R	TGGAGCTCCCTCAACCCCAA	62
2	278	2F	GTTCCTGAGTGGGAAGGGCT	2R	GCACCATGATTACTACATTACCT	60
3	93	3F	CAACATCCAATGGAAGCCTG	3R	CATCATGAAGGTTGGGCAAG	60
4	129	4F	ATCCCAATGCTGTGACGCCT	4R	TGACTCTGATGGCAAGCATG	60
5	509	5F	AGACTGTGAACCAGGAAGAG	5R	CCTTGCCAGCTGTGCCCAC	60

Table 4 PCR Primers used for amplification of the coding exons of the *h-ephrin-B1* gene.

Exon no	3' Splice Site	5' Splice Site	Exon Size (bp)
1	CGGGCTCGAT	TCAACCCCAAgtgagtaact	829
2	ttctgggcagGTTCCTGAGT	TACATTACCTgtgagtcccg	278
3	cttcctgcagCAACATCCAA	GTTGGGCAAGgtgagtgcct	93
4	ccattcttagATCCCAATGC	GGCAAGCATGgtaagtgtat	129
5	ccttcctcagAGACTGTGAA	GCTGCCTGGG	413

Table 5 Intron-Exon Junctions of the *h-ephrin-B1* gene.

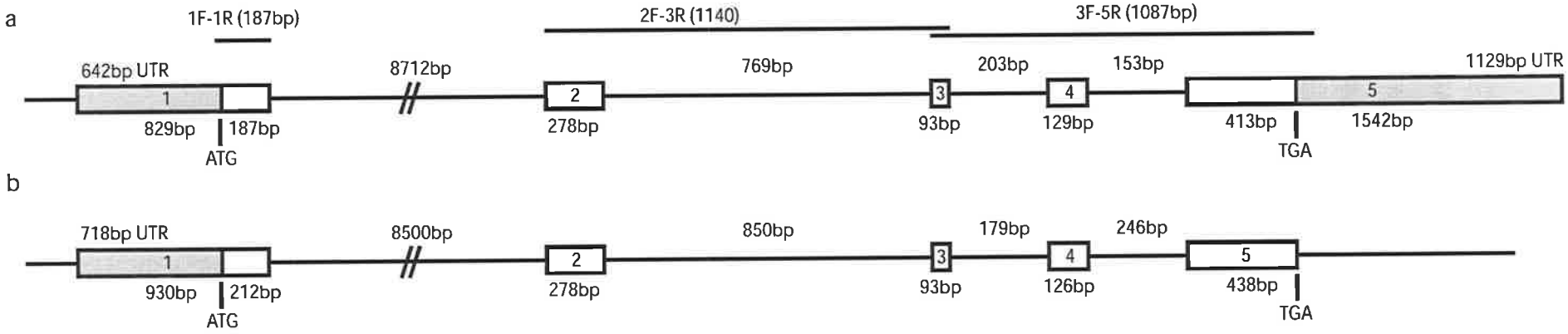


Figure 37 The intron-exon boundaries of ephrin-b1 are highly conserved in human and mouse, exons are indicated by rectangles. (a) Genomic organization of the *h-ephrin-B1* gene. Intron sizes are generated from genomic data (NT_019696). (b) *m-ephrin-B1* genomic structure (adapted from (Fletcher *et. al.*, 1994)). Untranslated regions are shown in grey. The primer combinations used to amplify the Aicardi's patients are indicated above the h-ephrin-B1.

Discussion

Sequence analysis of the *h-ephrin-B1* locus in non-disputed Aicardi's patients showed no mutations in any of the five exons. Nonetheless, physiological evidence does suggest that the *h-ephrin-B1* gene is involved in callosal formation, implicating ephrin-B1 and its regulatory elements or a downstream interactor as a putative Aicardi's causative gene. While no mutation in the coding region of the gene itself was found, DNA sequence variations in the regulatory region of *h-ephrin-B1* remain to be evaluated. Also, regulatory elements upstream of exon 1 may significantly influence expression of *h-ephrin-B1*, which could also be putative sites for mutation. Chromosome walking using STS sites across the X chromosome region containing *h-ephrin-B1* could identify these possible elements. While the Ephs and their ligands have been proposed to have a number of potential therapeutic applications (for example in cancer and nerve cord regeneration), no member of this subfamily of receptor tyrosine kinases has been implicated directly in disease causation. Should any mutation be detected in either the coding or non-coding region of the Aicardi's patients, it would be the first heritable disease linked with the Eph/ephrin family. Such a result would also provide an invaluable insight into the causation of Aicardi's syndrome.

Chapter 7: General Discussion

General Discussion

Introduction

This thesis encompasses investigations into the ephrins, which are ligands to the Eph receptor tyrosine kinases. The vertebrate Ephs are divided into two subclasses EphA and EphB according sequence homologies and their binding specificities to the ephrin-A and ephrin-B ligands respectively (Figure 1,Figure 37a).

In *C.elegans* there is only one Eph (Vab1), which shows equivalent sequence similarity to vertebrate EphA and EphB (George *et. al.*, 1998). Vab-1 also contains all domains found in vertebrate Ephs (Figure 37b). The *C.elegans* genome encodes for four GPI anchored ephrin ligands (e-ephrin-1-4, Figure 37b) that all are able to activate the Vab-1 receptor *in vivo*. Mutations in e-ephrin-1, e-ephrin-2, or e-ephrin-3, have defects in epidermal cell organization (Wang *et. al.*, 1999). Mutations in e-ephrin-1, e-ephrin-2, and e-ephrin-3, lead to defects in head morphology and dorsal epidermal closure (Wang *et. al.*, 1999), similar to those found with Vab-1 (George *et. al.*, 1998).

In *Drosophila* there is only one Eph receptor (*dek*) (Figure 37c) that has been shown to be involved in the developing *Drosophila* eye, which acts by controlling photoreceptor and cortical receptor axonal topography (Dearborn *et. al.*, 2002). In a study of the ventral nerve cord of *Drosophila,* misexpression studies of *dek* showed no perturbation of axonal pathfinding (Scully *et. al.*, 1999).

Initially, the aim of this thesis was to isolate a ligand to the only known *Drosophila* Eph receptor (Dek). This ligand was subsequently released in the *Drosophila* genome project and named *d-ephrin* according to naming conventions (Figure 37c). Investigations then turned to characterising the function *d-ephrin* both *in vitro* and *in vivo*. The, role of human ephrin-B1 as a principle disease gene of Aicardi's syndrome was also investigated.

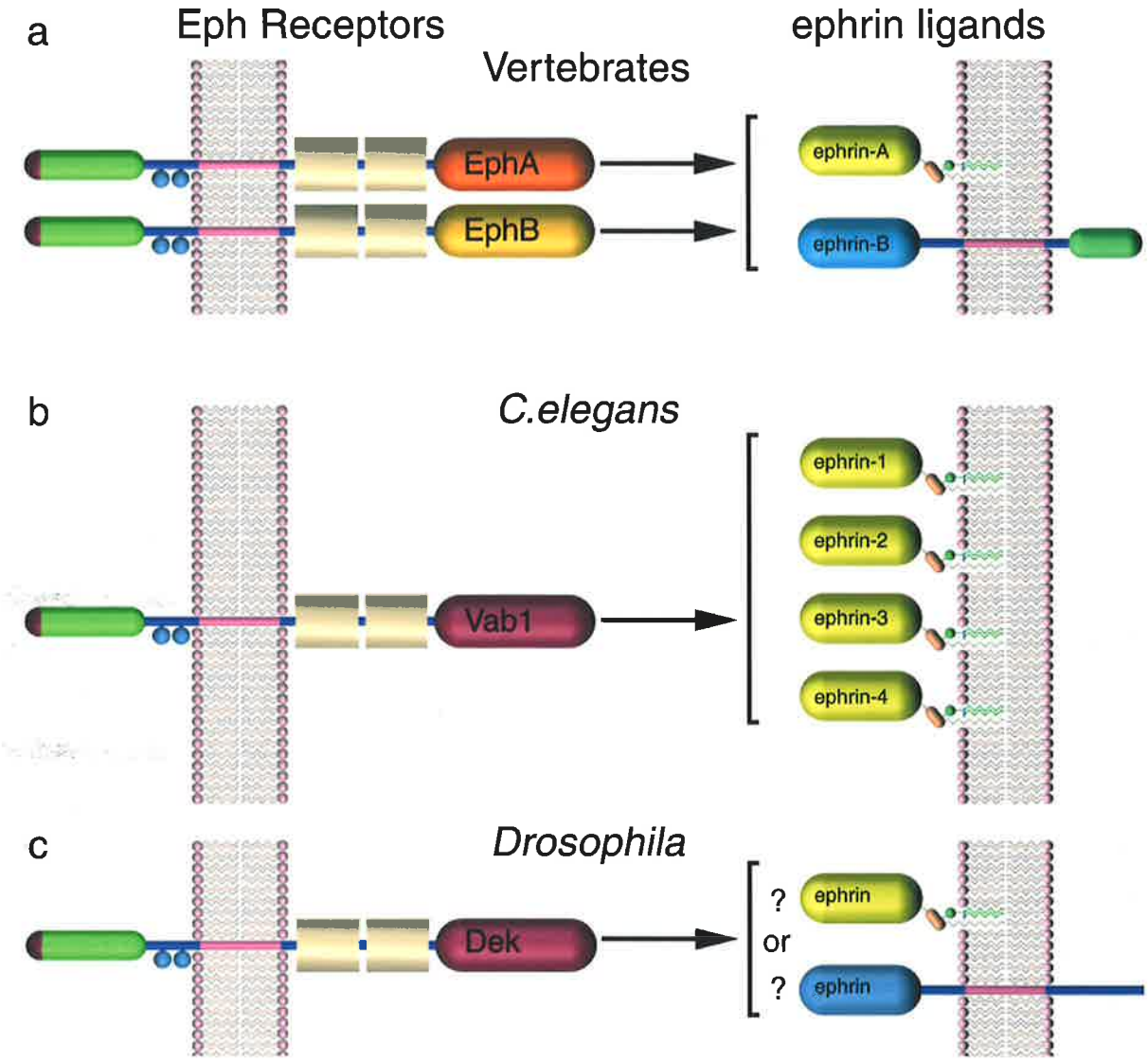


Figure 39 All currently known Eph receptors and ephrin ligands in vertebrates and invertebrates. (a) Canonical vertebrate EphA and EphB, and ephrin-A and ephrin-B subclasses are found in a number of vertebrate species including: *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Gallus gallus*, *Danio rerio*, and *Xenopus laevis*. Other members of each subclass are removed for clarity. (b) Vab1 is the only currently known Eph receptor in *Ctenophorus ornatus*, which is equally similar to the verterbrate EphA and EphB proteins in terms of sequence homology. The four ephrins in the *C.elegans* genome are all GPI linked to the membrane (ephrin 1-4, also referred to as EFN1-4 in the literature). (c) Dek is the only currently known Eph receptor in *Drosophila melanogaster*, which is also equally similar to the invertebrate EphA and EphB proteins in terms of sequence similarity. The only currently known *Drosophila* ephrin, which has been the major topic of this thesis could be either transmembrane or GPI anchored. However the presence of a GPI anchored ephrin in *C.elegans* may suggest that there remains an ephrin-A type gene in *Drosophila*.

Summary of results

The *Drosophila* genome encodes for at least one ephrin protein

The d-ephrin protein identified by the genome project only shares the ephrin core with other vertebrate ephrins. The protein contains two putative membrane anchorage regions both of which give the d-ephrin protein an intracellular region with no apparent function. Although, cleavage of d-ephrin gives a potential GPI modification site, which means that a GPI anchored version of this gene cannot be ruled out.

At the time of axon pathfinding *d-ephrin* expression is restricted to the developing CNS, specifically in the lateral regions of the VNC and the mushroom bodies in the brain. This expression pattern mirrors that of the only currently known *Drosophila* Eph (Dek).

In vitro S2 cell aggregation assays indicate that d-ephrin is a potential ligand to the Dek receptor, with cells expressing both the ligand and receptor forming aggregates when mixed together.

Phylogenetic analysis of all currently known ephrin orthologs

At the time of this analysis there were 45 known ephrin orthologs across *Rattus norvegicus*, *Gallus gallus*, *Danio rerio*, *Xenopus laevis*, *Drosophila melanogaster*, *Ctenophorus ornatus*, *and Caenorhabditis elegans*. Phylogenetic analysis showed that on the whole all vertebrate ephrin-A orthologs form a clade and all vertebrate ephrin-B orthologs form a clade. Within each clade ephrin-A1 to ephrin-A5 are all monophyletic, similarly the ephrin-B1 to ephrin-B3 are also monophyletic. The single exception to this is c-ephrin-A6, which consistently groups with ephrin-A4, suggesting that it may be a chicken ortholog of ephrin-A4. The invertebrate ephrins (e-ephrin-1-4, d-ephrin) group equidistant from the ephrin-A and ephrin-B subgroups, which may suggest a prototypic relationship between the invertebrate and vertebrate ephrins.

The coding regions of h-ephrin-B1 are normal in Aicardi's patients

The experiments outlined in this thesis show no mutation in the coding regions of *h-ephrin-B1* of six Aicardi's patients. However, the regulatory and non-coding regions of the h-ephrin-B1 gene still remain to be evaluated. The possibility of another axon guidance factor on the X-chromosome cannot be ruled out as a causative gene in Aicardi syndrome.

Future experiments

Does Kuzbanian cleave d-ephrin following receptor activation?

The results outlined in Chapter 3 indicate that the *Drosophila* ephrin contains a potential metalloproteinase site, which could facilitate axon retraction upon Dek/d-ephrin interactions, in a manner similar to that of vertebrate Eph/ephrin interactions (discussed in Chapter 1). In order to test this *in vitro* a disaggregation assay could be performed. Where the addition of *UAS-Kuz* to the *UAS-d-ephrin* cell line may cause perturbation of the cell aggregates shown to be formed between *d-ephrin* and *dek* expressing S2 cells (Figure 39). As a control for this experiment a UAS construct with the putative metalloproteinase site of d-ephrin mutated could be used (*UAS-ΔME-d-ephrin*).

Should the disaggregation of *dek* and site directed mutagenesis of the KUZ site of *d-ephrin*, it would be interesting to generate flies carrying the *UAS-ΔME-d-ephrin*, which may act in a dominant negative manner when misexpressed in tissues where *dek* expressing axons are extending growth cones. This is similar to experiments done by Fambrough *et. al.,* (1996), where *kuzbanian* mutant flies, had major defects in the formation of the longitudinal connectives and commissures (Fambrough *et. al.*, 1996).

Posttranslational processing of d-ephrin and anchorage mechanisms?

The results outlined in this chapter do not conclusively show if *d-ephrin* codes for a protein, which is transmembrane anchored, or GPI anchored. With an antibody available to d-ephrin, this could be tested by expressing d-ephrin in S2 cells and exposing these cells to GPI specific PI-PLC. Protein lysates run on a western blot should see a shift in the size of the d-ephrin extracellular region, between PI-PLC exposed and non exposed cell populations, if d-ephrin is GPI anchored. This assay is similar to that which has been used in the past to determine if other ephrin-As are GPI anchored (Drescher *et. al.*, 1995; Winslow *et. al.*, 1995; Davy *et. al.*, 1999).

Function of *d-ephrin in vivo*?

The experiments outlined in Chapter 4 show compelling evidence for a VNC phenotype when *d-ephrin* is misexpressed in all cells with a neural fate. Obviously, this requires further investigation. This could be done with a number of approaches including RNAi or site directed mutagenesis of *d-ephrin*. It would be very interesting to explore the role of d-ephrin in the development of the *Drosophila* VNC or in the developing *Drosophila* eye. Furthermore, future experiments in *Drosophila* to elucidate the signaling pathway of d-ephrin could prove extremely useful. In vertebrate systems the downstream effectors of Eph/ephrin

signaling are beginning to be elucidated (see Chapter 1). *Drosophila* is highly amenable to genetic analysis, which could make testing for upstream and downstream genes in the Dek/d-ephrin signaling pathway much easier than in vertebrate systems.

Concluding remarks

This thesis has surveyed the Eph/ephrin family of receptor tyrosine kinases, with a particular focus on the ephrin ligands. The *Drosophila* ephrin characterised, d-ephrin, was shown to have a potential function in the development of the CNS during fly embryogenesis. In particular, it seems evident that d-ephrin may be involved in the arrangement of longitudinal connectives relative to the midline. This is consistent with studies in vertebrate systems, where Eph/ephrin interactions have been implicated in a large array of developmental mechanisms in the developing embryo, including many that are involved in the guidance of axons both ipsilateral and contralateral to the midline. Collectively, this evidence inferred that there may be some involvement of Eph/ephrin interactions in humans, in particular with commissure formation. This was addressed by looking at a particular human disease, Aicardi syndrome, which is characterised by a commissural defect, namely callosal agenesis. However, ephrin-B1 was not found to be a causative gene for this disorder.

Currently, there is a large body of literature accumulating in regards to the involvement of Eph/ephrins in cancer. It will be of interest to see if Ephs or ephrins are directly linked as causative agents in human genetic disorders, as well as cancer, in the future.

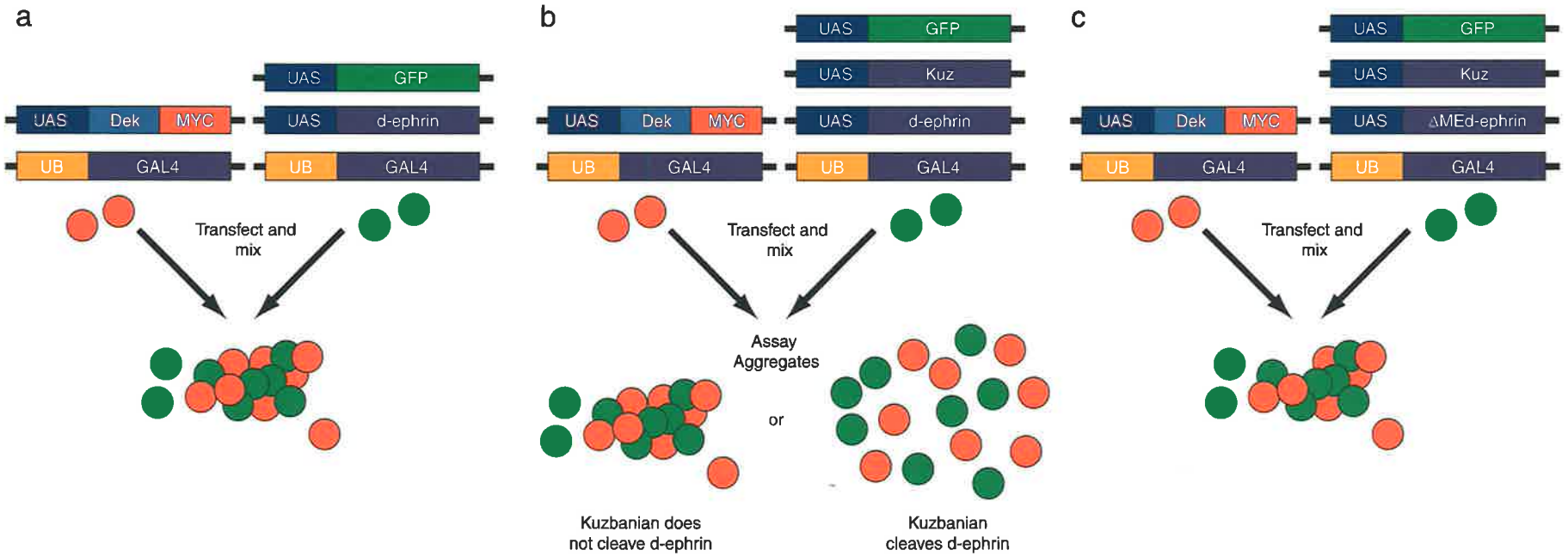


Figure 41 Disaggregation assay with Kuzbanian to test if *d-ephrin* contains a metalloproteinase site as predicted in Chapter 3. (a) Aggregation experiment as shown in Chapter 4, where *d-ephrin* expressing S2 cells (green) are able to aggregate with *dek* expressing S2 cells (red). (b) In this assay the ability of Kuzbanian to perturb cell aggregation of *dek* expressing cells with *d-ephrin* expressing cells is tested by co-expressing *kuzbanian (kuz)* in the ligand cell. If *kuz* can cleave d-ephrin upon receptor binding no cell aggregates should be seen (i.e dis aggregate). (c) A control experiment expressing *d-ephrin* with the putative metalloproteinase site mutated (*UAS-ΔMEd-ephrin*), should also prevent disaggregation.

References

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., *et. al.,* (2000). The genome sequence of *Drosophila* melanogaster. *Science* **287** (5461): 2185-2195.

Aicardi, J. and Chevrie, J. J. (1994). The Aicardi syndrome. Calosal Agenesis: A natural split in the brain? M. Lassonde and Jeeves, M. A. New York, Plenum Press: 7-17.

Aicardi, J., Chevrie, J. J. and Rousselie, F. (1969). Le syndrome spasmes en flexion, agenesic calleuse, anomalies chorio-retiniennes. *Arch Franc Pediat* **26**: 1103-1120.

Aicardi, J., Lefebvre, J. and Lerique, A. (1965). Spasms in flexion, callosal agenesis, ocular abnormalities: a new syndrome. *Electroencephalogr Clin Neurophysiol* **19**: 609-622.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped blast and psi-blast - a new generation of protein database search programs. *Nucleic Acids Res* **25** (17): 3389-3402.

Bailey, T. L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, Menlo Park, California, AAAI Press.

Bartley, T. D., Hunt, R. W., Welcher, A. A., Boyle, W. J., Parker, V. P., Lindberg, R. A., Lu, H. S., Colombero, A. M., Elliott, R. L., Guthrie, B. A., *et. al.,* (1994). B61 is a ligand for the ECK receptor protein-tyrosine kinase. *Nature* **368** (6471): 558-560.

Becker, N., Seitanidou, T., Murphy, P., Matt'ei, M. G., Topilko, P., Nieto, M. A., Wilkinson, D. G., Charnay, P. and Gilardi Hebenstreit, P. (1994). Several receptor tyrosine kinase genes of the Eph family are segmentally expressed in the developing hindbrain. *Mech Dev* **47** (1): 3-17.

Beckmann, M. P., Cerretti, D. P., Baum, P., Vanden Bos, T., James, L., Farrah, T., Kozlosky, C., Hollingsworth, T., Shilling, H., Maraskovsky, E., *et. al.,* (1994). Molecular

characterization of a family of ligands for eph-related tyrosine kinase receptors. *Embo J* **13** (16): 3757-3762.

Bennett, B. D., Zeigler, F. C., Gu, Q., Fendly, B., Goddard, A. D., Gillett, N. and Matthews, W. (1995). Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine kinase Htk. *Proc Natl Acad Sci U S A* **92** (6): 1866-1870.

Bergemann, A. D., Cheng, H. J., Brambilla, R., Klein, R. and Flanagan, J. G. (1995). ELF-2, a new member of the Eph ligand family, is segmentally expressed in mouse embryos in the region of the hindbrain and newly forming somites. *Mol Cell Biol* **15** (9): 4921-4929.

Brambilla, R., Bruckner, K., Orioli, D., Bergemann, A. D., Flanagan, J. G. and Klein, R. (1996). Similarities and differences in the way transmembrane-type ligands interact with the Elk subclass of Eph receptors. *Mol Cell Neurosci* **8** (2-3): 199-209.

Brambilla, R., Schnapp, A., Casagranda, F., Labrador, J. P., Bergemann, A. D., Flanagan, J. G., Pasquale, E. B. and Klein, R. (1995). Membrane-bound LERK2 ligand can signal through three different Eph-related receptor tyrosine kinases. *Embo J* **14** (13): 3116-3126.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118** (2): 401-415.

Brown, D. (1993). The tyrosine kinase connection: how GPI-anchored proteins activate T cells. *Curr Opin Immunol* **5** (3): 349-354.

Brown, D. (2002). Structure and function of membrane rafts. *Int J Med Microbiol* **291** (6-7): 433-437.

Brown, D. A. and London, E. (1998). Functions of lipid rafts in biological membranes. *Annu Rev Cell Dev Biol* **14**: 111-136.

Bruckner, K. and Klein, R. (1998). Signaling by Eph receptors and their ephrin ligands. *Curr Opin Neurobiol* **8** (3): 375-382.

Bruckner, K., Pablo Labrador, J., Scheiffele, P., Herb, A., Seeburg, P. H. and Klein, R. (1999). EphrinB ligands recruit GRIP family PDZ adaptor proteins into raft membrane microdomains. *Neuron* **22** (3): 511-524.

Bruckner, K., Pasquale, E. B. and Klein, R. (1997). Tyrosine phosphorylation of transmembrane ligands for Eph receptors. *Science* **275** (5306): 1640-1643.

Buloz, D. and Kronegg, J. (2001). Automatic detection of GPI-anchored proteins: http://129.194.186.123/GPI-anchor/auteurs_en.html. Geneva.

Caras, I. W. (1997). A link between axon guidance and axon fasciculation suggested by studies of the tyrosine kinase receptor EphA5/REK7 and its ligand ephrin-A5/AL-1. *Cell Tissue Res* **290** (2): 261-264.

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res* **15** (4): 1353-1361.

Cerretti, D. P., Bos, T. V., Nelson, N., Kozlosky, C. J., Reddy, P., Maraskovsky, E., Park, L. S., Lyman, S. D., Copeland, N. G., Gilbert, D. J., *et. al.,* (1995). Isolation of LERK-5: A ligand of the eph-related receptor tyrosine kinases. *Mol Immunol* **32** (16): 1197-1205.

Cerretti, D. P. and Nelson, N. (1998). Characterization of the genes for mouse LERK-3/Ephrin-A3 (Epl3), mouse LERK-4/Ephrin-A4 (Epl4), and human LERK-6/Ephrin-A2 (EPLG6): conservation of intron/exon structure. *Genomics* **47** (1): 131-135.

Chen, J., Nachabah, A., Scherer, C., Ganju, P., Reith, A., Bronson, R. and Ruley, H. E. (1996). Germ-line inactivation of the murine Eck receptor tyrosine kinase by gene trap retroviral insertion. *Oncogene* **12** (5): 979-988.

Cheng, H. J. and Flanagan, J. G. (1994). Identification and cloning of ELF-1, a developmentally expressed ligand for the Mek4 and Sek receptor tyrosine kinases. *Cell* **79** (1): 157-168.

Cheng, H. J., Nakamoto, M., Bergemann, A. D. and Flanagan, J. G. (1995). Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* **82** (3): 371-381.

Chin-Sang, I. D., George, S. E., Ding, M., Moseley, S. L., Lynch, A. S. and Chisholm, A. D. (1999). The ephrin VAB-2/EFN-1 functions in neuronal signaling to regulate epidermal morphogenesis in C. elegans. *Cell* **99** (7): 781-790.

Cho, K. O., Hunt, C. A. and Kennedy, M. B. (1992). The rat brain postsynaptic density fraction contains a homolog of the *Drosophila* discs-large tumor suppressor protein. *Neuron* **9** (5): 929-942.

Collins, P. (2000). Cleavage: resolving a sticky problem. *NATURE REVIEWS NEUROSCIENCE* **1**: 9.

Connor, R. J., Menzel, P. and Pasquale, E. B. (1998). Expression and Tyrosine Phosphorylation of Eph Receptors Suggest Multiple Mechanisms in Patterning of the Visual System. *Developmental Biology* **193** (1): 21-35.

Cox, E. C., Muller, B. and Bonhoeffer, F. (1990). Axonal guidance in the chick visual system: posterior tectal membranes induce collapse of growth cones from the temporal retina. *Neuron* **4** (1): 31-37.

Craven, S. E. and Bredt, D. S. (1998). PDZ proteins organize synaptic signaling pathways. *Cell* **93** (4): 495-498.

Cross, G. A. (1990). Glycolipid anchoring of plasma membrane proteins. *Annual Reviews in Cell Biology* **6**: 1-39.

Daniel, T. O., Stein, E., Cerretti, D. P., L, S. T. P., Robert, B. and Abrahamson, D. R. (1996). ELK and LERK-2 in developing kidney and microvascular endothelial assembly. *Kidney International* **57**: S73-S81.

Davis, S., Gale, N. W., Aldrich, T. H., Maisonpierre, P. C., Lhotak, V., Pawson, T., Goldfarb, M. and Yancopoulos, G. D. (1994). Ligands for EPH-related receptor tyrosine kinases that require membrane attachment or clustering for activity. *Science* **266** (5186): 816-819.

Davy, A., Gale, N. W., Murray, E. W., Klinghoffer, R. A., Soriano, P., Feuerstein, C. and Robbins, S. M. (1999). Compartmentalized signaling by GPI-anchored ephrin-A5 requires the Fyn tyrosine kinase to regulate cellular adhesion. *Genes Dev* **13** (23): 3125-3135.

Dearborn, R., He, Q., Kunes, S. and Yong, D. (2002). Eph Receptor Tyrosine Kinase-Mediated Formation of a Topographic Map in the *Drosophila* Visual System. *The Journal of Neuroscience* **22** (4): 1338–1349.

Dottori, M., Hartley, L., Galea, M., Paxinos, G., Polizzotto, M., Kilpatrick, T., Bartlett, P. F., Murphy, M., Kontgen, F. and Boyd, A. W. (1998). EphA4 (Sek1) receptor tyrosine kinase is required for the development of the corticospinal tract. *Proc Natl Acad Sci U S A* **95** (22): 13248-13253.

Drescher, U., Bonhoeffer, F. and Muller, B. K. (1997). The Eph family in retinal axon guidance. *Current Opinion In Neurobiology* **7** (1): 75-80.

Drescher, U., Kremoser, C., Handwerker, C., Loschinger, J., Noda, M. and Bonhoeffer, F. (1995). In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell* **82** (3): 359-370.

Eisenhaber, B., Bork, P. and Eisenha (1999a). Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase.

Eisenhaber, B., Bork, P. and Eisenhaber, F. (1999b). Prediction of potential GPI-modification sites in proprotein sequences. *J Mol Biol* **292** (3): 741-758.

Eisenhaber, B., Bork, P., Yuan, Y., Loffler, G. and Eisenhaber, F. (2000). Automated annotation of GPI anchor sites: case study C. elegans. *Trends Biochem Sci* **25** (7): 340-341.

Fambrough, D., Pan, D., Rubin, G. M. and Goodman, C. S. (1996). The cell surface metalloprotease/disintegrin Kuzbanian is required for axonal extension in *Drosophila*. *Proc Natl Acad Sci U S A* **93** (23): 13233-13238.

Fehon, R. G., Kooh, P. J., Rebay, I., Regan, C. L., Xu, T., Muskavitch, M. A. and Artavanis-Tsakonas, S. (1990). Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in *Drosophila*. *Cell* **61** (3): 523-534.

Feldheim, D. A., Kim, Y. I., Bergemann, A. D., Frisen, J., Barbacid, M. and Flanagan, J. G. (2000). Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25** (3): 563-574.

Feldheim, D. A., Vanderhaeghen, P., Hansen, M. J., Frisen, J., Lu, Q., Barbacid, M. and Flanagan, J. G. (1998). Topographic guidance labels in a sensory projection to the forebrain. *Neuron* **21** (6): 1303-1313.

Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package). Seattle, Department of Genetics, University of Washington.

Ferguson, A. J. and Williams, A. F. (1988). Cell-surface anchoring of proteins via Glycosyl-PhosphatidylInositol structures. *Annual Reviews in Biochemistry* **57**: 285-320.

Flanagan, J. G. and Vanderhaeghen, P. (1998). The ephrins and Eph receptors in neural development. *Annu Rev Neurosci* **21**: 309-345.

Flenniken, A. M., Gale, N. W., Yancopoulos, G. D. and Wilkinson, D. G. (1996). Distinct and overlapping expression patterns of ligands for Eph-related receptor tyrosine kinases during mouse embryogenesis. *Develop Biol* **179** (2): 382-401.

Fletcher, F. A., Renshaw, B., Hollingsworth, T., Baum, P., Lyman, S. D., Jenkins, N. A., Gilbert, D. J., Copeland, N. G. and Davison, B. L. (1994). Genomic organization and chromosomal localization of mouse Eplg2, a gene encoding a binding protein for the receptor tyrosine kinase elk. *Genomics* **24** (1): 127-132.

FlyBase (1999). The FlyBase database of the *Drosophila* Genome Projects and community literature. The FlyBase Consortium. *Nucleic Acids Res* **27** (1): 85-88.

FlyBase (2002). The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res* **30** (1): 106-108.

Fraser, S., Keynes, R. and Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344**: 431-435.

Frisen, J., Yates, P. A., McLaughlin, T., Friedman, G. C., O'Leary, D. D. and Barbacid, M. (1998). Ephrin-A5 (AL-1/RAGS) is essential for proper retinal axon guidance and topographic mapping in the mammalian visual system. *Neuron* **20** (2): 235-243.

Gale, N. W., Flenniken, A., Compton, D. C., Jenkins, N., Copeland, N. G., Gilbert, D. J., Davis, S., Wilkinson, D. G. and Yancopoulos, G. D. (1996a). Elk-L3, a novel transmembrane ligand for the Eph family of receptor tyrosine kinases, expressed in embryonic floor plate, roof plate and hindbrain segments. *Oncogene* **13** (6): 1343-1352.

Gale, N. W., Holland, S. J., Valenzuela, D. M., Flenniken, A., Pan, L., Ryan, T. E., Henkemeyer, M., Strebhardt, K., Hirai, H., Wilkinson, D. G., *et. al.,* (1996b). Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. *Neuron* **17** (1): 9-19.

Gale, N. W. and Yancopoulos, G. D. (1997). Ephrins and their receptors: a repulsive topic? *Cell and Tissue Research* **290** (2): 227-241.

George, S. E., Simokat, K., Hardin, J. and Chisholm, A. D. (1998). The Vab-1 Eph Receptor Tyrosine Kinase Functions in Neural and Epithelial Morphogenesis in C-Elegans. *Cell* **92** (5): 633-643.

Gerety, S. S., Wang, H. U., Chen, Z. F. and Anderson, D. J. (1999). Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. *Molecular Cell* **4** (3): 403-414.

Goodhill, G. J. and Richards, L. J. (1999). Retinotectal maps: molecules, models and misplaced data. *Trends Neurosci* **22** (12): 529-534.

Graur, D. and Wen-Hsiung, L. (1999). Fundamentals of molecular evolution. Sunderland, Massachusetts, Sinauer Associates.

Gurniak, C. B. and Berg, L. J. (1996). A new member of the Eph family of receptors that lacks protein tyrosine kinase activity. *Oncogene* **13** (4): 777-786.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166** (4): 557-580.

Harlow, E. and Lane, D. (1988). Antibodies: A laboratory Manual. New York, Cold Spring Harbour Laboratory.

Hattori, M., Osterfield, M. and Flanagan, J. G. (2000). Regulated cleavage of a contact-mediated axon repellent. *Science* **289** (5483): 1360-1365.

Henkemeyer, M., Marengere, L. E., McGlade, J., Olivier, J. P., Conlon, R. A., Holmyard, D. P., Letwin, K. and Pawson, T. (1994). Immunolocalization of the Nuk receptor tyrosine kinase suggests roles in segmental patterning of the brain and axonogenesis. *Oncogene* **9** (4): 1001-1014.

Henkemeyer, M., Orioli, D., Henderson, J. T., Saxton, T. M., Roder, J., Pawson, T. and Klein, R. (1996). Nuk controls pathfinding of commissural axons in the mammalian central nervous system. *Cell* **86** (1): 35-46.

Himanen, J. P., Henkemeyer, M. and Nikolov, D. B. (1998). Crystal structure of the ligand-binding domain of the receptor tyrosine kinase EphB2. *Nature* **396** (6710): 486-491.

Himanen, J. P. and Nikolov, D. B. (2002). Purification, crystallization and preliminary characterization of an Eph-B2/ephrin-B2 complex. *Acta Crystallogr D Biol Crystallogr* **58** (Pt 3): 533-535.

Himanen, J. P., Rajashankar, K. R., Lackmann, M., Cowan, C. A., Henkemeyer, M. and Nikolov, D. B. (2001). Crystal structure of an Eph receptor-ephrin complex. *Nature* **414** (6866): 933-938.

Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. and Takaku, F. (1987). A novel putative tyrosine kinase receptor encoded by the eph gene. *Science* **238** (4834): 1717-1720.

Hock, B., Bohme, B., Karn, T., Yamamoto, T., Kaibuchi, K., Holtrich, U., Holland, S., Pawson, T., Rubsamen-Waigmann, H. and Strebhardt, K. (1998). PDZ-domain-mediated interaction of the Eph-related receptor tyrosine kinase EphB3 and the ras-binding protein AF6 depends on the kinase activity of the receptor. *Proc Natl Acad Sci U S A* **95** (17): 9779-9784.

Holland, S. J., Gale, N. W., Mbamalu, G., Yancopoulos, G. D., Henkemeyer, M. and Pawson, T. (1996). Bidirectional signaling through the EPH-family receptor Nuk and its transmembrane ligands. *Nature* **383** (6602): 722-725.

Hornberger, M. R., Dutting, D., Ciossek, T., Yamada, T., Handwerker, C., Lang, S., Weth, F., Huf, J., Wessel, R., Logan, C., *et. al.,* (1999). Modulation of EphA receptor function by coexpressed ephrinA ligands on retinal ganglion cell axons. *Neuron* **22** (4): 731-742.

Huelsenbeck, J. P., Ronquist, F., Nielsen, R. and Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294** (5550): 2310-2314.

Jones, T. L., Chong, L. D., Kim, J., Xu, R. H., Kung, H. F. and Daar, I. O. (1998). Loss Of Cell Adhesion In Xenopus Laevis Embryos Mediated By the Cytoplasmic Domain Of Xlerk, an Erythropoietin-Producing Hepatocellular Ligand. *Proc Natl Acad Sci USA* **95** (2): 576-581.

Kennedy, M. B. (1995). Origin of PDZ (DHR, GLGF) domains. *Trends Biochem Sci* **20** (9): 350.

Knoll, B. and Drescher, U. (2002). Ephrin-As as receptors in topographic projections. *Trends Neurosci* **25** (3): 145-149.

Koblar, S. A., Krull, C. E., Pasquale, E. B., McLennan, R., Peale, F. D., Cerretti, D. P. and Bothwell, M. (2000). Spinal motor axons and neural crest cells use different molecular guides for segmental migration through the rostral half-somite. *J Neurobiol* **42** (4): 437-447.

Kornau, H. C., Schenker, L. T., Kennedy, M. B. and Seeburg, P. H. (1995). Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science* **269** (5231): 1737-1740.

Kortschak (2000). Tip/Top, A program to convert protein sequence alignments in corresponding cDNA alignments. Adelaide University.

Kozlosky, C. J., Maraskovsky, E., McGrew, J. T., VandenBos, T., Teepe, M., Lyman, S. D., Srinivasan, S., Fletcher, F. A., Gayle, R. B. r., Cerretti, D. P., *et. al.,* (1995). Ligands for the receptor tyrosine kinases hek and elk: isolation of cDNAs encoding a family of proteins. *Oncogene* **10** (2): 299-306.

Kozlosky, C. J., VandenBos, T., Park, L., Cerretti, D. P. and Carpenter, M. K. (1997). LERK-7: A ligand of the Eph-related kinases is developmentally regulated in the brain. *Cytokine* **9** (8): 540-549.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305** (3): 567-580.

Krull, C. E. (1998). Inhibitory interactions in the patterning of trunk neural crest migration. *Ann N Y Acad Sci* **857**: 13-22.

Krull, C. E., Lansford, R., Gale, N. W., Collazo, A., Marcelle, C., Yancopoulos, G. D., Fraser, S. E. and Bronner-Fraser, M. (1997). Interactions of Eph-related receptors and ligands confer rostrocaudal pattern to trunk neural crest migration. *Curr Biol* **7** (8): 571-580.

Kullander, K., Mather, N. K., Diella, F., Dottori, M., Boyd, A. W. and Klein, R. (2001). Kinase-dependent and kinase-independent functions of EphA4 receptors in major axon tract formation in vivo. *Neuron* **29** (1): 73-84.

Labrador, J. P., Brambilla, R. and Klein, R. (1997). The N-terminal globular domain of Eph receptors is sufficient for ligand binding and receptor signaling. *Embo J* **16** (13): 3889-3897.

Lackmann, M., Bucci, T., Mann, R. J., Kravets, L. A., Viney, E., Smith, F., Moritz, R. L., Carter, W., Simpson, R. J., Nicola, N. A., *et. al.,* (1996). Purification of a ligand for the EPH-like receptor HEK using a biosensor-based affinity detection approach. *Proc Nat Acad Sci USA* **93** (6): 2523-2527.

Lackmann, M., Oates, A. C., Dottori, M., Smith, F. M., Do, C., Power, M., Kravets, L. and Boyd, A. W. (1998). Distinct subdomains of the EphA3 receptor mediate ligand binding and receptor dimerization. *J Biol Chem* **273** (32): 20228-20237.

Larget, B. and Simon, D. L. (1999). Markov Chain Monte Carlo Algorithms for the Bayesian Analysis of Phylogenetic Trees. *Molecular Biology and Evolution* **16** (6): 750-759.

Lassonde, M. and Jeeves, M. A. (1994). Calosal Agenesis: A natural split in the brain? New York, Plenum Press.

Lemke, G. (1997). A coherent nomenclature for Eph receptors and their ligands. *Molecular and Cellular Neuroscience* **9** (5-6): 331-332.

Lin, D., Gish, G. D., Songyang, Z. and Pawson, T. (1999). The carboxyl terminus of B class ephrins constitutes a PDZ domain binding motif. *J Biol Chem* **274** (6): 3726-3733.

Lin, D. M. and Goodman, C. S. (1994). Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* **13** (3): 507-523.

Lindberg, R. A. and Hunter, T. (1990). cDna cloning and characterization of eck, an epithelial cell receptor protein-tyrosine kinase in the eph/elk family of protein kinases. *Mol Cell Biol* **10** (12): 6316-6324.

Lindsay, D. L. and Zimm, G. G. (1992). The Genome of *Drosophila Melanogaster*. San Diego, Academic Press.

Lu, Q., Sun, E. E., Klein, R. S. and Flanagan, J. G. (2001). Ephrin-b reverse signaling is mediated by a novel pdz-rgs protein and selectively inhibits g protein-coupled chemoattraction. *Cell* **105** (1): 69-79.

Luo, L., Liao, Y. J., Jan, L. Y. and Jan, Y. N. (1994). Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev* **8** (15): 1787-1802.

Matsuoka, H., Iwata, N., Ito, M., Shimoyama, M., Nagata, A., Chihara, K., Takai, S. and Matsui, T. (1997). Expression of a kinase-defective Eph-like receptor in the normal human brain. *Biochemical and Biophysical Research Communications* **235** (3): 487-492.

Meima, L., Moran, P., Matthews, W. and Caras, I. W. (1997). Lerk2 (ephrin-B1) is a collapsing factor for a subset of cortical growth cones and acts by a mechanism different from AL-1 (ephrin-A5). *Molecular and Cellular Neuroscience* **9** (4): 314-328.

Mellitzer, G., Xu, Q. and Wilkinson, D. G. (1999). Eph receptors and ephrins restrict cell intermingling and communication. *Nature* **400** (6739): 77-81.

Mellitzer, G., Xu, Q. and Wilkinson, D. G. (2000). Control of cell behaviour by signaling through Eph receptors and ephrins. *Curr Opin Neurobiol* **10** (3): 400-408.

Menzel, P., Valencia, F., Godement, P., Dodelet, V. C. and Pasquale, E. B. (2001). Ephrin-A6, a new ligand for EphA receptors in the developing visual system. *Dev Biol* **230** (1): 74-88.

Molina, J. A., Mateos, F., Merino, M., Epifanio, J. L. and Gorrono, M. (1989). Aicardi syndrome in 2 sisters. *Journal of Pediatrics* **115**: 282-283.

Moller, S., Croning, M. D. R. and Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17** (7): 646-653.

Monschau, B., Kremoser, C., Ohta, K., Tanaka, H., Kaneko, T., Yamada, T., Handwerker, C., Hornberger, M. R., Loschinger, J., Pasquale, E. B., *et. al.,* (1997). Shared and distinct functions of RAGS and ELF-1 in guiding retinal axons. *Embo J* **16** (6): 1258-1267.

Murray, M. J. and Whitington, P. M. (1999). Effects of roundabout on growth cone dynamics, filopodial length, and growth cone morphology at the midline and throughout the neuropile. *J Neurosci* **19** (18): 7901-7912.

Nakamoto, M., Cheng, H. J., Friedman, G. C., McLaughlin, T., Hansen, M. J., Yoon, C. H., D, O. L. D. and Flanagan, J. G. (1996). Topographically specific effects of ELF-1 on retinal axon guidance in vitro and retinal axon mapping in vivo. *Cell* **86** (5): 755-766.

Neidich, J. A., Nussbaum, R. L., Packer, R. J., Emanuel, B. S. and Puck, J. M. (1990). Heterogeneity of clinical severity and molecular lesions in Aicardi syndrome. *Journal of Pediatrics* **116**: 911-917.

Nicholas, K. B., Jr., N. H. B. and Deerfield, D. W. (1997). GeneDoc: Analysis and Visualization of Genetic Variation. EMBNEWNEWS. **4:** 14-15.

Nicola, N. A., Viney, E., Hilton, D. J., Roberts, B. and Willson, T. (1996). Molecular cloning of two novel transmembrane ligands for eph-related kinases (LERKS) that are related to LERK-2. *Growth Factors* **13** (1-2): 141-149.

Nielsen, H., Brunak, S. and von Heijne, G. (1999). Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Eng* **12** (1): 3-9.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997a). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10** (1): 1-6.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997b). A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* **8** (5-6): 581-599.

Nielsen, H. and Krogh, A. (1998). Prediction of signal peptides and signal anchors by a hidden Markov model. Sixth International Conference on Intelligent Systems for Molecular Biology, Menlo Park, California, AAAI Press.

Nieto, M. A., Gilardi, H.-P., Charnay, P. and Wilkinson, D. G. (1992). A receptor protein tyrosine kinase implicated in the segmental patterning of the hindbrain and mesoderm. *Development* **116** (4): 1137-1150.

Orioli, D., Henkemeyer, M., Lemke, G., Klein, R. and Pawson, T. (1996). Sek4 and Nuk receptors cooperate in guidance of commissural axons and in palate formation. *Embo J* **15** (22): 6035-6049.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12** (4): 357-358.

Page, R. D. and Holmes, E. C. (1998). Molecular evolution: A phylogenetic approach. Oxford, Blackwell Science.

Pandey, A., Lindberg, R. A. and Dixit, V. M. (1995a). Cell signaling. Receptor orphans find a family. *Curr Biol* **5** (9): 986-989.

Pandey, A., Shao, H., Marks, R. M., Polverini, P. J. and Dixit, V. M. (1995b). Role of B61, the ligand for the Eck receptor tyrosine kinase, in TNF-alpha-induced angiogenesis. *Science* **268** (5210): 567-569.

Park, S. and Sanchez, M. P. (1997). The Eek receptor, a member of the Eph family of tyrosine protein kinases, can be activated by three different Eph family ligands. *Oncogene* **14** (5): 533-542.

Pasquale, E. B. (1991). Identification of chicken embryo kinase 5, a developmentally regulated receptor-type tyrosine kinase of the Eph family. *Cell Regul* **2** (7): 523-534.

Patel, N. and Goodman, C. (2000). Detection of even-skipped transcripts in *Drosophila* embryos with PCR/DIG-labeled DNA probes. Nonradioactive In Situ Hybridization Application Manual. Roche, Roche: 162-164.

Peale, F. V., Jr., Mason, K., Hunter, A. W. and Bothwell, M. (1998). Multiplex display polymerase chain reaction amplifies and resolves related sequences sharing a single moderately conserved domain. *Anal Biochem* **256** (2): 158-168.

Phelps, C. B. and Brand, A. H. (1998). Ectopic gene expression in *Drosophila* using GAL4 system. *Methods* **14** (4): 367-379.

Plath, K., Mothes, W., Wilkinson, B. M., Stirling, C. J. and Rapoport, T. A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* **94** (6): 795-807.

Rajagopalan, S., Vivancos, V., Nicolas, E. and Dickson, B. J. (2000). Selecting a longitudinal pathway: Robo receptors specify the lateral position of axons in the *Drosophila* CNS. *Cell* **103** (7): 1033-1045.

Raper, J. A. and Kapfhammer, J. P. (1990). The enrichment of a neuronal growth cone collapsing activity from embryonic chick brain. *Neuron* **4** (1): 21-29.

Roche (2000). Nonradioactive In Situ Hybridization Application Manual, Roche Biochemicals.

Ropers, H. H., Zuffardi, O., Bianchi, E. and Tiepolo, L. (1982). Agenesis of the Corpus Callosum, Ocular, and skeletal Anomalies (X-linked Dominant Aicardi's Syndrome) in a girl with Balanced X/3 Translocation. *Human Genetics* **61**: 364-368.

Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. and Harvey, D. A. (2000). A *Drosophila* complementary DNA resource. *Science* **287** (5461): 2222-2224.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposible element vectors. *Science* **218**: 348-353.

Sajjadi, F. G., Pasquale, E. B. and Subramani, S. (1991). Identification of a new eph-related receptor tyrosine kinase gene from mouse and chicken that is developmentally regulated and encodes at least two forms of the receptor. *New Biol* **3** (8): 769-778.

Sakano, S., Serizawa, R., Inada, T., Iwama, A., Itoh, A., Kato, C., Shimizu, Y., Shinkai, F., Shimizu, R., Kondo, S., *et. al.,* (1996). Characterization of a ligand for receptor protein-tyrosine kinase HTK expressed in immature hematopoietic cells. *Oncogene* **13** (4): 813-822.

Saras, J. and Heldin, C. H. (1996). PDZ domains bind carboxy-terminal sequences of target proteins. *Trends Biochem Sci* **21** (12): 455-458.

Schneider, I. (1972). Cell lines derived from late embryonic stages of *Drosophila* melanogaster. *J Embryol Exp Morphol* **27** (2): 353-365.

Scholz, H., Sadlowski, E., Klaes, A. and Klambt, C. (1997). Control of midline glia development in the embryonic *Drosophila* CNS. *Mech Dev* **64** (1-2): 137-151.

Scully, A. L., McKeown, M. and Thomas, J. B. (1999). Isolation and characterization of Dek, a *Drosophila* eph receptor protein tyrosine kinase. *Mol Cell Neurosci* **13** (5): 337-347.

Shao, H., Lou, L., Pandey, A., Pasquale, E. B. and Dixit, V. M. (1994). cDna cloning and characterization of a ligand for the Cek5 receptor protein-tyrosine kinase. *J Biol Chem* **269** (43): 26606-26609.

Simon, D. and Larget, B. (2001). Bayesian analysis in molecular biology and evolution (BAMBE). Duquesne University.

Snow, P. M., Bieber, A. J. and Goodman, C. S. (1989). Fasciclin III: a novel homophilic adhesion molecule in *Drosophila*. *Cell* **59** (2): 313-323.

Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M. and Cantley, L. C. (1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275** (5296): 73-77.

Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc Natl Acad Sci U S A* **50**: 703-710.

Spradling, A. C. and Rubin, G. M. (1982). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218** (4570): 341-347.

Stapleton, D., Balan, I., Pawson, T. and Sicheri, F. (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat Struct Biol* **6** (1): 44-49.

Stein, E., Lane, A. A., Cerretti, D. P., Schoecklmann, H. O., Schroff, A. D., Vanetten, R. L. and Daniel, T. O. (1998). Eph Receptors Discriminate Specific Ligand Oligomers to Determine Alternative Signaling Complexes, Attachment, and Assembly Responses. *Genes & Development* **12** (5): 667-678.

Tamura, K. and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10** (3): 512-526.

Tessier-Lavigne, M. and Goodman, C. S. (1996). The molecular biology of axon guidance. *Science* **274** (5290): 1123-1133.

Thanos, C. D., Goodwill, K. E. and Bowie, J. U. (1999). Oligomeric structure of the human EphB2 receptor SAM domain. *Science* **283** (5403): 833-836.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22** (22): 4673-4680.

Torres, R., Firestein, B. L., Dong, H., Staudinger, J., Olson, E. N., Huganir, R. L., Bredt, D. S., Gale, N. W. and Yancopoulos, G. D. (1998). PDZ proteins bind, cluster, and synaptically colocalize with Eph receptors and their ephrin ligands. *Neuron* **21** (6): 1453-1463.

Tosch, P., Koblar, S. and Saint, R. (2002). Isolation and Characterisation of d-ephrin. *in press.*

Toth, J., Cutforth, T., Gelinas, A. D., Bethoney, K. A., Bard, J. and Harrison, C. J. (2001). Crystal structure of an ephrin ectodomain. *Dev Cell* **1** (1): 83-92.

Trainor, P. A. and Krumlauf, R. (2000). Patterning the cranial neural crest: hindbrain segmentation and Hox gene plasticity. *Nat Rev Neurosci* **1** (2): 116-124.

Trainor, P. A. and Krumlauf, R. (2001). Hox genes, neural crest cells and branchial arch patterning. *Curr Opin Cell Biol* **13** (6): 698-705.

Unified nomenclature for Eph family receptors and their ligands, the ephrins. (1997). *Cell* **90** (3): 403-404.

van der Geer, P., Hunter, T. and Lindberg, R. A. (1994). Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu Rev Cell Biol* **10**: 251-337. California 91320.

Walkenhorst, J., Dutting, D., Handwerker, C., Huai, J., Tanaka, H. and Drescher, U. (2000). The EphA4 receptor tyrosine kinase is necessary for the guidance of nasal retinal ganglion cell axons in vitro. *Mol Cell Neurosci* **16** (4): 365-375.

Wallin, E. and von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7** (4): 1029-1038.

Walsh, F. S. and Doherty, P. (1991). Glycosylphosphatidylinositol ancohored recognition molecules that function in axonal fasiculation, growth and guidance in the nervous system. *Cell Biology International Reports* **15** (11): 1151-1167.

Walsh, F. S. and Doherty, P. (1996). Cell adhesion molecules and neuronal regeneration. *Curr Opin Cell Biol* **8** (5): 707-713.

Wang, H. U. and Anderson, D. J. (1997). Eph family transmembrane ligands can mediate repulsive guidance of trunk neural crest migration and motor axon outgrowth. *Neuron* **18** (3): 383-396.

Wang, H. U., Chen, Z. F. and Anderson, D. J. (1998). Molecular Distinction and Angiogenic Interaction Between Embryonic Arteries and Veins Revealed By Ephrin-B2 and Its Receptor Eph-B4. *Cell* **93** (5): 741-753.

Wang, X., Roy, P. J., Holland, S. J., Zhang, L. W., Culotti, J. G. and Pawson, T. (1999). Multiple ephrins control cell organization in C. elegans using kinase- dependent and - independent functions of the VAB-1 Eph receptor. *Molecular Cell* **4** (6): 903-913.

Weinstein, D. C., Rahman, S. M., Ruiz, J. C. and HemmatiBrivanlou, A. (1996). Embryonic expression of eph signaling factors in Xenopus. *Mech Dev* **57** (2): 133-144.

Wilkinson, D. G. (2000). Eph receptors and ephrins: regulators of guidance and assembly. *Int Rev Cytol* **196**: 177-244.

Wilkinson, D. G. (2001). Multiple roles of EPH receptors and ephrins in neural development. *Nat Rev Neurosci* **2** (3): 155-164.

Wilks, A. F. (1989). Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. *Proc Natl Acad Sci U S A* **86** (5): 1603-1607.

Winning, R. S., Scales, J. B. and Sargent, T. D. (1996). Disruption of cell adhesion in Xenopus embryos by Pagliaccio, an Eph-class receptor tyrosine kinase. *Develop Biol* **179** (2): 309-319.

Winslow, J. W., Moran, P., Valverde, J., Shih, A., Yuan, J. Q., Wong, S. C., Tsai, S. P., Goddard, A., Henzel, W. J., Hefti, F., *et. al.,* (1995). Cloning of AL-1, a ligand for an Eph-related tyrosine kinase receptor involved in axon bundle formation. *Neuron* **14** (5): 973-981.

Woods, D. F. and Bryant, P. J. (1991). The discs-large tumor suppressor gene of *Drosophila* encodes a guanylate kinase homolog localized at septate junctions. *Cell* **66** (3): 451-464.

Woods, D. F. and Bryant, P. J. (1993). ZO-1, DlgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions. *Mech Dev* **44** (2-3): 85-89.

Xu, Q., Mellitzer, G., Robinson, V. and Wilkinson, D. G. (1999). In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. *Nature* **399** (6733): 267-271.

Xu, Q., Mellitzer, G. and Wilkinson, D. G. (2000). Roles of Eph receptors and ephrins in segmental patterning. *Philos Trans R Soc Lond B Biol Sci* **355** (1399): 993-1002.

Xu, Q. L., Alldus, G., Holder, N. and Wilkinson, D. G. (1995). Expression of truncated Sek-1 receptor tyrosine kinase disrupts the segmental restriction of gene expression in the Xenopos and zebrafish hindbrain. *Development* **121** (12): 4005-4016.

Yang, Z. and Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Mol Biol Evol* **14** (7): 717-724.

Zhou, R. P. (1998). THE EPH FAMILY RECEPTORS AND LIGANDS. *Pharmacology & Therapeutics* **77** (3): 151-181.

Zisch, A. H., Stallcup, W. B., Chong, L. D., DahlinHuppe, K., Voshol, J., Schachner, M. and Pasquale, E. B. (1997). Tyrosine phosphorylation of L1 family adhesion molecules: Implication of the Eph kinase Cek5. *Journal Of Neuroscience Research* **47** (6): 655-665.

Appendices

Appendix A ephrin gene, cDNA and protein sequences

cDNA and gene sequences with Accession Numbers		**Protein sequences with Accession Numbers**	
c-ephrin-A2	L40932	c-ephrin-A2	AAC42229.1
c-ephrin-A5	X90377	c-ephrin-A5	CAA62027.1
c-ephrin-B1	U72394	c-ephrin-B1	AAC07986.1
c-ephrin-B2	AF180729	c-ephrin-B2	AAD53948.1
c-ephrin-A6	AF317286	c-ephrin-A6	AAK00944.1
d-ephrin	AF216287	d-ephrin	AAF28394.1
e-ephrin-1	AF201079	e-ephrin-1	AAF25647
e-ephrin-2	Wang *et al.* (1999)	e-ephrin-2	Wang *et al.* (1999)
e-ephrin-3	Wang *et al.* (1999)	e-ephrin-3	Wang *et al.* (1999)
e-ephrin-4	AF410936	e-ephrin-4	AAL05561
h-ephrin-A1	NM_004428	h-ephrin-A1	NP_004419.1
h-ephrin-A2	NM_001405	h-ephrin-A2	NP_001396.1
h-ephrin-A3	XM_001787	h-ephrin-A3	XP_001787.1
h-ephrin-A4	XM_001784	h-ephrin-A4	XP_001784.1
h-ephrin-A5	XM_003914	h-ephrin-A5	XP_003914.2
h-ephrin-B1	NM_004429	h-ephrin-B1	NP_004420.1
h-ephrin-B2	NM_004093	h-ephrin-B2	NP_004084.1
h-ephrin-B3	XM_008230	h-ephrin-B3	XP_008230.1
o-ephrin-A2	AF209776	o-ephrin-A2	AAF19443.1
o-ephrin-A3	AF209777	o-ephrin-A3	AAF19444.1
m-ephrin-A1	BC002046	m-ephrin-A1	AAH02046.1
m-ephrin-A2	NM_007909	m-ephrin-A2	NP_031935.1
m-ephrin-A3	U92885	m-ephrin-A3	AAC39961.1
m-ephrin-A4	NM_007910	m-ephrin-A4	NP_031936.1
m-ephrin-A5	U90664	m-ephrin-A5	AAB50239.1
m-ephrin-B1	NM_010110	m-ephrin-B1	NP_034240.1
m-ephrin-B2	NM_010111	m-ephrin-B2	NP_034241.1
m-ephrin-B3	NM_007911	m-ephrin-B3	NP_031937.1
r-ephrin-A1	NM_053599	r-ephrin-A1	NP_446051.1
r-ephrin-A2	AF131912	r-ephrin-A2	AAD33515.1
r-ephrin-A3	AYO45577	r-ephrin-A3	AAK92219.1
r-ephrin-A5	U69279	r-ephrin-A5	NP_446355
r-ephrin-B1	NM_017089	r-ephrin-B1	NP_058785.1
x-ephrin-A1	U31204	x-ephrin-A1	AAA74485.1
x-ephrin-A3	AW200648	x-ephrin-A3	AW200648
x-ephrin-B1	U31427	x-ephrin-B1	AAC35995.1
x-ephrin-B2	AF128844	x-ephrin-B2	AAD32610.1
x-ephrin-B3	AJ236866	x-ephrin-B3	CAB65511.1
z-ephrin-A2	Y09668	z-ephrin-A2	CAA70863.1
z-ephrin-A3	AB051678	z-ephrin-A3	BAB55891
z-ephrin-A5	Y09669	z-ephrin-A5	CAA70864.1
z-ephrin-AL1	AJ006838	z-ephrin-AL1	CAA07264.1
z-ephrin-B1	AAK64274	z-ephrin-B1	AAK64274
z-ephrin-B2	AJ004863	z-ephrin-B2	CAA06168.1
z-ephrin-B3	AF375227	z-ephrin-B3	AAK64277.1

Species abbreviations: h-*Homo sapiens*, m-*Mus musculus*, r-*Rattus norvegicus*, c-*Gallus gallus*, z-*Danio rerio*, x-*Xenopus laevis*, d-*Drosophila melanogaster*, o-*Ctenophorus ornatus*, e-*Caenorhabditis elegans*

Appendix B SignalP predictions using neural networks (NN) and hidden markov models (HMM)

	neural networks (NN) Predictions (Nielsen *et. al.*, 1997b; Nielsen *et. al.*, 1997a)	hidden Markov models (HMM) Predictions (Nielsen and Krogh, 1998)
h-ephrin-A1	>huma1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 19 0.697 0.33 YES max. Y 19 0.749 0.32 YES max. S 12 0.966 0.82 YES mean S 1-18 0.873 0.47 YES # Most likely cleavage site between pos. 18 and 19: AAA-DR	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.751 at 19
m-ephrin-A1	m-ephrin-A1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 19 0.697 0.33 YES max. Y 19 0.745 0.32 YES max. S 12 0.967 0.82 YES mean S 1-18 0.866 0.47 YES # Most likely cleavage site between pos. 18 and 19: AAA-DR	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.687 at 19
r-ephrin-A1	r-ephrin-A1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.294 0.33 NO max. Y 17 0.460 0.32 YES max. S 12 0.971 0.82 YES mean S 1-16 0.914 0.47 YES # Most likely cleavage site between pos. 16 and 17: SLA-AV	Prediction: Signal peptide Signal peptide probability: 0.996 Signal anchor probability: 0.000 Max cleavage site probability: 0.581 at 17
x-ephrin-A1	x-ephrin-A1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 29 0.593 0.33 YES max. Y 29 0.505 0.32 YES max. S 10 0.949 0.82 YES mean S 1-28 0.630 0.47 YES # Most likely cleavage site between pos. 28 and 29: AQG-ER	Prediction: Signal peptide Signal peptide probability: 0.496 Signal anchor probability: 0.437 Max cleavage site probability: 0.293 at 29
z-ephrin-A11	z-ephrin-A11 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.830 0.33 YES max. Y 21 0.795 0.32 YES max. S 6 0.975 0.82 YES mean S 1-20 0.854 0.47 YES # Most likely cleavage site between pos. 20 and 21: ASA-ER	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.809 at 21
h-ephrin-A2b	h-ephrin-A2b length = 70 # Measure Position Value Cutoff signal peptide? max. C 25 0.486 0.33 YES max. Y 25 0.599 0.32 YES max. S 11 0.996 0.82 YES mean S 1-24 0.875 0.47 YES # Most likely cleavage site between pos. 24 and 25: PFA-PP	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.612 at 25

m-ephrin-A2	m-ephrin-A2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 23 0.650 0.33 YES max. Y 23 0.691 0.32 YES max. S 13 0.979 0.82 YES mean S 1-22 0.876 0.47 YES # Most likely cleavage site between pos. 22 and 23: ARN-ED	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.646 at 21
r-ephrin-A2	r-ephrin-A2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 28 0.105 0.33 NO max. Y 20 0.034 0.32 NO max. S 41 0.107 0.82 NO mean S 1-19 0.057 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 0
c-ephrin-A2	c-ephrin-A2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.404 0.33 YES max. Y 21 0.519 0.32 YES max. S 7 0.912 0.82 YES mean S 1-20 0.766 0.47 YES # Most likely cleavage site between pos. 20 and 21: VWS-DD	Prediction: Signal peptide Signal peptide probability: 0.992 Signal anchor probability: 0.008 Max cleavage site probability: 0.925 at 21
>ndraga2	>ndraga2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 57 0.060 0.33 NO max. Y 14 0.055 0.32 NO max. S 6 0.243 0.82 NO mean S 1-13 0.167 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 23
>ndraga2	>ndraga2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 57 0.060 0.33 NO max. Y 14 0.055 0.32 NO max. S 6 0.243 0.82 NO mean S 1-13 0.167 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 23
z-ephrin-A2	z-ephrin-A2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 19 0.600 0.33 YES max. Y 19 0.645 0.32 YES max. S 4 0.943 0.82 YES mean S 1-18 0.816 0.47 YES # Most likely cleavage site between pos. 18 and 19: VWS-DD	Prediction: Signal peptide Signal peptide probability: 0.993 Signal anchor probability: 0.001 Max cleavage site probability: 0.930 at 19
h-ephrin-A3	h-ephrin-A3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 31 0.488 0.33 YES max. Y 23 0.562 0.32 YES max. S 9 0.998 0.82 YES mean S 1-22 0.969 0.47 YES # Most likely cleavage site between pos. 22 and 23: LLA-QG	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.780 at 23
m-ephrin-A3	m-ephrin-A3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 51 0.564 0.33 YES max. Y 51 0.171 0.32 NO max. S 47 0.195 0.82 NO mean S 1-50 0.052 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 0
r-ephrin-A3	r-ephrin-A3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 32 0.564 0.33 YES max. Y 32 0.175 0.32 NO max. S 28 0.195 0.82 NO mean S 1-31 0.068 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.001 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 32

>ndraga3	>ndraga3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 57 0.062 0.33 NO max. Y 13 0.049 0.32 NO max. S 1 0.279 0.82 NO mean S 1-12 0.151 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 0
x-ephrin-A3	x-ephrin-A3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 17 0.059 0.33 NO max. Y 17 0.057 0.32 NO max. S 13 0.173 0.82 NO mean S 1-16 0.090 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 17
z-ephrin-A3	z-ephrin-A3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 24 0.747 0.33 YES max. Y 24 0.770 0.32 YES max. S 8 0.997 0.82 YES mean S 1-23 0.912 0.47 YES # Most likely cleavage site between pos. 23 and 24: VTA-AR	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.666 at 24
h-ephrin-A4	h-ephrin-A4 length = 70 # Measure Position Value Cutoff signal peptide? max. C 23 0.312 0.33 NO max. Y 23 0.471 0.32 YES max. S 9 0.989 0.82 YES mean S 1-22 0.857 0.47 YES # Most likely cleavage site between pos. 22 and 23: LRG-GS	Prediction: Signal peptide Signal peptide probability: 0.995 Signal anchor probability: 0.001 Max cleavage site probability: 0.404 at 23
m-ephrin-A4	m-ephrin-A4 length = 70 # Measure Position Value Cutoff signal peptide? max. C 44 0.501 0.33 YES max. Y 44 0.552 0.32 YES max. S 31 0.881 0.82 YES mean S 1-43 0.512 0.47 YES # Most likely cleavage site between pos. 43 and 44: CSS-LR	Prediction: Signal peptide Signal peptide probability: 0.979 Signal anchor probability: 0.000 Max cleavage site probability: 0.662 at 44
h-ephrin-A5	h-ephrin-A5 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 1.000 0.33 YES max. Y 21 0.855 0.32 YES max. S 9 0.994 0.82 YES mean S 1-20 0.959 0.47 YES # Most likely cleavage site between pos. 20 and 21: VFS-QD	Prediction: Signal peptide Signal peptide probability: 0.986 Signal anchor probability: 0.013 Max cleavage site probability: 0.865 at 21
m-ephrin-A5c	m-ephrin-A5c length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 1.000 0.33 YES max. Y 21 0.858 0.32 YES max. S 7 0.995 0.82 YES mean S 1-20 0.966 0.47 YES # Most likely cleavage site between pos. 20 and 21: VFS-QD	Prediction: Signal peptide Signal peptide probability: 0.993 Signal anchor probability: 0.007 Max cleavage site probability: 0.918 at 21
r-ephrin-A5	r-ephrin-A5 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 1.000 0.33 YES max. Y 21 0.858 0.32 YES max. S 7 0.995 0.82 YES mean S 1-20 0.966 0.47 YES # Most likely cleavage site between pos. 20 and 21: VFS-QD	Prediction: Signal peptide Signal peptide probability: 0.993 Signal anchor probability: 0.007 Max cleavage site probability: 0.918 at 21

c-ephrin-A5	c-ephrin-A5 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.979 0.33 YES max. Y 21 0.853 0.32 YES max. S 5 0.961 0.82 YES mean S 1-20 0.864 0.47 YES # Most likely cleavage site between pos. 20 and 21: VRG-QE	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.984 at 21
z-ephrin-A5	z-ephrin-A5 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.825 0.33 YES max. Y 21 0.774 0.32 YES max. S 7 0.973 0.82 YES mean S 1-20 0.890 0.47 YES # Most likely cleavage site between pos. 20 and 21: VFS-QE	Prediction: Signal peptide Signal peptide probability: 0.845 Signal anchor probability: 0.145 Max cleavage site probability: 0.764 at 21
c-ephrin-A6	c-ephrin-A6 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.965 0.33 YES max. Y 21 0.776 0.32 YES max. S 5 0.991 0.82 YES mean S 1-20 0.734 0.47 YES # Most likely cleavage site between pos. 20 and 21: VRG-RR	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.989 at 21
h-ephrin-B1b	h-ephrin-B1b length = 70 # Measure Position Value Cutoff signal peptide? max. C 30 0.318 0.33 NO max. Y 30 0.503 0.32 YES max. S 19 0.976 0.82 YES mean S 1-29 0.857 0.47 YES # Most likely cleavage site between pos. 29 and 30: PLA-KN	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.381 at 30
m-ephrin-B1	m-ephrin-B1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 30 0.351 0.33 YES max. Y 30 0.532 0.32 YES max. S 19 0.986 0.82 YES mean S 1-29 0.867 0.47 YES # Most likely cleavage site between pos. 29 and 30: PLA-KN	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.365 at 30
r-ephrin-B1	r-ephrin-B1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 30 0.351 0.33 YES max. Y 30 0.532 0.32 YES max. S 19 0.986 0.82 YES mean S 1-29 0.867 0.47 YES # Most likely cleavage site between pos. 29 and 30: PLA-KN	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.365 at 30
c-ephrin-B1	c-ephrin-B1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 26 0.323 0.33 NO max. Y 26 0.496 0.32 YES max. S 14 0.981 0.82 YES mean S 1-25 0.920 0.47 YES # Most likely cleavage site between pos. 25 and 26: PLA-KS	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.518 at 28
x-ephrin-B1	x-ephrin-B1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 23 0.770 0.33 YES max. Y 23 0.797 0.32 YES max. S 12 0.989 0.82 YES mean S 1-22 0.925 0.47 YES # Most likely cleavage site between pos. 22 and 23: ALG-KN	Prediction: Signal peptide Signal peptide probability: 0.999 Signal anchor probability: 0.000 Max cleavage site probability: 0.624 at 23

z-ephrin-B1	z-ephrin-B1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 45 0.390 0.33 YES max. Y 29 0.300 0.32 NO max. S 22 0.834 0.82 YES mean S 1-28 0.491 0.47 YES # Most likely cleavage site between pos. 28 and 29: AKS-LE	Prediction: Non-secretory protein Signal peptide probability: 0.466 Signal anchor probability: 0.000 Max cleavage site probability: 0.204 at 25
h-ephrin-B2	h-ephrin-B2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 26 0.363 0.33 YES max. Y 26 0.456 0.32 YES max. S 7 0.874 0.82 YES mean S 1-25 0.661 0.47 YES # Most likely cleavage site between pos. 25 and 26: AIS-KS	Prediction: Signal peptide Signal peptide probability: 0.895 Signal anchor probability: 0.000 Max cleavage site probability: 0.659 at 26
m-ephrin-B2	m-ephrin-B2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 29 0.265 0.33 NO max. Y 29 0.399 0.32 YES max. S 10 0.896 0.82 YES mean S 1-28 0.675 0.47 YES # Most likely cleavage site between pos. 28 and 29: AIS-RS	Prediction: Signal peptide Signal peptide probability: 0.987 Signal anchor probability: 0.000 Max cleavage site probability: 0.531 at 29
c-ephrin-B2	c-ephrin-B2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 28 0.655 0.33 YES max. Y 28 0.568 0.32 YES max. S 25 0.948 0.82 YES mean S 1-27 0.606 0.47 YES # Most likely cleavage site between pos. 27 and 28: ALA-KS	Prediction: Signal peptide Signal peptide probability: 0.994 Signal anchor probability: 0.000 Max cleavage site probability: 0.940 at 28
x-ephrin-B2	x-ephrin-B2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 0.061 0.33 NO max. Y 21 0.028 0.32 NO max. S 5 0.094 0.82 NO mean S 1-20 0.039 0.47 NO	Prediction: Non-secretory protein Signal peptide probability: 0.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.000 at 24
z-ephrin-B2	z-ephrin-B2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 25 0.667 0.33 YES max. Y 25 0.578 0.32 YES max. S 4 0.756 0.82 NO mean S 1-24 0.586 0.47 YES # Most likely cleavage site between pos. 24 and 25: SRA-LI	Prediction: Non-secretory protein Signal peptide probability: 0.382 Signal anchor probability: 0.000 Max cleavage site probability: 0.313 at 25
h-ephrin-B3	h-ephrin-B3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 28 0.588 0.33 YES max. Y 28 0.593 0.32 YES max. S 20 0.975 0.82 YES mean S 1-27 0.729 0.47 YES # Most likely cleavage site between pos. 27 and 28: VSG-LS	Prediction: Signal peptide Signal peptide probability: 0.983 Signal anchor probability: 0.016 Max cleavage site probability: 0.441 at 28
m-ephrin-B3	m-ephrin-B3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 28 0.785 0.33 YES max. Y 28 0.722 0.32 YES max. S 18 0.948 0.82 YES mean S 1-27 0.768 0.47 YES # Most likely cleavage site between pos. 27 and 28: VSG-LS	Prediction: Signal peptide Signal peptide probability: 0.985 Signal anchor probability: 0.013 Max cleavage site probability: 0.615 at 28

x-ephrin-B3	x-ephrin-B3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 25 0.558 0.33 YES max. Y 25 0.608 0.32 YES max. S 12 0.931 0.82 YES mean S 1-24 0.794 0.47 YES # Most likely cleavage site between pos. 24 and 25: ISA-LS	Prediction: Signal peptide Signal peptide probability: 0.744 Signal anchor probability: 0.021 Max cleavage site probability: 0.521 at 25
z-ephrin-B3	z-ephrin-B3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 26 0.841 0.33 YES max. Y 26 0.825 0.32 YES max. S 14 0.965 0.82 YES mean S 1-25 0.871 0.47 YES # Most likely cleavage site between pos. 25 and 26: ITA-TN	Prediction: Signal peptide Signal peptide probability: 0.961 Signal anchor probability: 0.033 Max cleavage site probability: 0.465 at 26
d-ephrin	d-ephrin length = 70 # Measure Position Value Cutoff signal peptide? max. C 34 0.288 0.33 NO max. Y 34 0.458 0.32 YES max. S 11 0.965 0.82 YES mean S 1-33 0.844 0.47 YES # Most likely cleavage site between pos. 33 and 34: SSC-AK	Prediction: Signal peptide Signal peptide probability: 0.965 Signal anchor probability: 0.026 Max cleavage site probability: 0.262 at 28
e-ephrin-1	>efn-1 length = 70 # Measure Position Value Cutoff signal peptide? max. C 23 0.617 0.33 YES max. Y 23 0.687 0.32 YES max. S 5 0.919 0.82 YES mean S 1-22 0.827 0.47 YES # Most likely cleavage site between pos. 22 and 23: CSA-KR	Prediction: Signal peptide Signal peptide probability: 0.974 Signal anchor probability: 0.013 Max cleavage site probability: 0.651 at 23
e-ephrin-2	>efn-2 length = 70 # Measure Position Value Cutoff signal peptide? max. C 19 0.972 0.33 YES max. Y 19 0.923 0.32 YES max. S 6 0.971 0.82 YES mean S 1-18 0.937 0.47 YES # Most likely cleavage site between pos. 18 and 19: GWA-RK	Prediction: Signal peptide Signal peptide probability: 0.996 Signal anchor probability: 0.000 Max cleavage site probability: 0.803 at 19
e-ephrin-3	>efn-3 length = 70 # Measure Position Value Cutoff signal peptide? max. C 20 0.944 0.33 YES max. Y 20 0.913 0.32 YES max. S 9 0.993 0.82 YES mean S 1-19 0.963 0.47 YES # Most likely cleavage site between pos. 19 and 20: VSC-RN	Prediction: Signal peptide Signal peptide probability: 1.000 Signal anchor probability: 0.000 Max cleavage site probability: 0.938 at 20
e-ephrin-4	>efn-4 length = 70 # Measure Position Value Cutoff signal peptide? max. C 21 1.000 0.33 YES max. Y 21 0.920 0.32 YES max. S 9 0.984 0.82 YES mean S 1-20 0.936 0.47 YES # Most likely cleavage site between pos. 20 and 21: AAA-DE	Prediction: Signal peptide Signal peptide probability: 0.994 Signal anchor probability: 0.006 Max cleavage site probability: 0.894 at 21

Appendix C Genomic sequence of h-ephrin-B1

Primers used in the analysis are underlined, the start and stop codons are in bold, and intronic sequence is in lowercase.

CGGGCTCGATCGCCCGGGAGCCAGGACTCGGCGACGCGAGGCTGCCGGGCTACCCGGCCGAGGCTTCGGGGGCGC
AAACTAATGGGACTGGCTCGCTCGGCAGCATCTCCCCGCTCTTCTAAGTACACTGAGCAGGGCCCGCGCTGAAGT
AGAAGCTGTCCGGGGGCGCGTAGCCCGGAGTCCCAGTGTGGCCCGGAGGAACGGAGCCCGTGCCAGGGCGGCCCA
GTCGGGAGCCCGGGGACCGAGCTTGTGCTGTGGGGAAACCCCCACTTCTTCCAAGGGACAGCGATCCCGGGACGG
TCGAGGCGTCGGGGCGGTCACCGAGACCTCTGCGGGAAGACCCCGTCGGGGAGAGGGCGCGCAGCCCCGAAGCGT
CTCGGGAAGTCGAGCGGAATCGGGCGGGATCACCCGGGGGCGCAGAGCCCCCGTCGCGCCTCGTGCGGCAGCGGA
GAGCCCAGGAGAACGAGCCCTCGGGGGCCCGAAGCCCATGCCCGGGTTGGGGGCGGCTGCCCAGTGAGTCCTCCTG
GCCGGCCGGGCGGAGAAGAGCGACACCGAAGCCGGCGGGAGGGGAGCACTTCAAGGCCGGCGGCTGCGGAGGATG
GGCGCCTGAGCGGCTCCGAGCGCAGCGCGGCAGAGGAAGGCGAGGCGAGCTTTGGTGAGGAGGCGCCAAGGGATC
CCGAAGTGCAGTCTGCCCCCGGGAAG**ATG**GCTCGGCCTGGGCAGCGTTGGCTCGGCAAGTGGCTTGTGGCGATGG
TCGTGTGGGCGCTGTGCCGGCTCGCCACACCGCTGGCCAAGAACCTGGAGCCCGTATCCTGGAGCTCCCTCAACC
CCAAgtgagtaacttatctcctctggacgctggggtgggaggcactccttcagggtgaggccgcacgccccggag
tgcatgtggggaggtcttcggaggaggagcggcgcctcattttgtctccggctttttcgagtgttttcctgcggg
cgggcggatggagcagggtgcggcggggtggggtaggggcacttggctgggttgtgacccccgggcttcccacc
ccctgcctctgcacgtcttggatgaagtcgaggtggtcgcacaggcggagctggggagtccgggcgcccgagagc
gcaccggagaactgagtggaggcgagaacgcagctccctccggggcccctgggaacgtggtgattttcgccgaga
gccgctccaggagcgagcttgactcgggcgagggacgaggccagaacacgggttcctcgtttgcgtagctgtgga
gggaacacagagcgggacccggcgcttggggtgggtggcagcgccaacatccccactcgggcttctcccattttc
atttctctttcttccctgttttcctccctggctttttgctttgctgtctctttttgccccccgcccctcctttct
ccctgccggagcggtagaaaggcgggcgctgctcccaggctcacaaaggcagcgggtggagaggggcggagtggg
agagtggtcgttgggtagtgggggttcgaaccagtgcccgtcaggggtccaatggagccggtaacagagcttcaa
gggcccgccagggcggagagtggcggccgagctgttgagggaggtgggtggagaaggagcgattcctgcccaggc
aggtccccttggcccccgagtctgtcccccagattgggaagctgtgcttagttttccccaaggggttggggcggg
gagtgggcagctatataaaggacatcctttcttttctttcttctccttcctctgcacaactaccccaccaccacc
acacacacaaccccctccaaacaacacgcacacctttcagggcggtggctcagtggtttctgggcttttagtggga
agtcgctggggctgttgtagtgacctctggaccttctcttttggcatccaggactcaggttggaaccaagggtca
aaattgcccttcctctcaattttctggaggtcccagcaatgtgcgaacaaccccacttagcaggaaactggggag
ctttgacctgctccatccgtgatcactgattcctgtatctcaaatttcttcctgtccactcttgctgccccacgc
cgcctccccccccagccccgccctgccccgggtaacttcccaggctcctgcctaccttggcctcttgggtcccagg
gacgtttgatctctcagccctggagtaggagaaggcagctttagttttcttctcctcgattaaggtagagtcccta
ttactggttggtgggatggctaaaccagggttggggtggagggctctccctgggtacaaggaaagggattggcaa
gaatttctttcctacccagaagccaaacggggtacaacactctcctaacagcaggggctcaaaatgctctgggag
ctctgggctgcttgagcccaggggggaacgttgtgaagtgctcagtggttgtgaagtgctcagtggttgtgaagt
gctcagttaccctggtaagggtggaggctggaaatgttctgtcttgagccaggagagttgaaatccgcctggggt
agcccatcatagggggctcatgggagtgacctaattgctctctcgggtctctaaccagagatagcatgcagctcca
cagctggaaaggagccctcctgcagagttgggggtgctggaagaaacacgatctgtctgtgcttcccccttttttg
tgctcagagactttagtgctctggactggcttacaggttttgggggggtggtatcctctctgacccccctcctcat
ctacctcaagctgagggcaggcattgtggggtgaggatgggggggtggctgaagactctctagaaactgtccgaat
tccattctgtaattgggtagatcctgggaaggatcagagactgatcctggtgcaccctccttccatccaggctca

ggggtctggaatgcagccactgtctacatgtgttggctcggggtggtgggggcaaatggacaggaaggaggcact
gccctcagctctccttgagctttgggggctgcttctgccaccgcgcccaaccctcaggctgccacagcctttcaa
ataaacagtccctctctgtctccctccctctctgcctcctgcccgagttctcctctcaggcccgtgtcagttcgc
cctctcataccctccgcgagcagcgctcacttcccctctccagctaggctcgcctgctctgccttgctctcacac
tccctctcactcttagtcactctctgcctctttttctcaagtccttttgtttctcgcacatgcccgtgctctccc
tctcaggctcattttctctcttccttcccctcttcctctcctggcttttggtctctctcctcagccctgtcagaa
gctggcaacccccctcccaaaaaagaaaatctcccagtgcctataaaccctgcttaattgcttccttcttgggaa
aaaaaaaatcaaaataaaagaagtggtggggctgggggtccgcttccaggttccaaggtgcgtgcggcccggccc
cagccctgactgagaggggaagcgggaatggctggcccaagactcccaagcctctttcccaaattggtgctgggg
ccttctgggatgctagtccctgagtggcaggagcttctgcctccttggcccctcctgtcagatttgccccatgag
cgcaaggtgggggatgggtttggtcctgctagagttttttttaggaacctctgtgtgtgcgtgtgtgtgtgtgt
gtgtgtgtgtgtgtctgagttgggggggctgttatctaaatggtcctaggcgacccctcccgcagcctcttccc
ctgggaacttggggcagccggcaggcctcacacttgagtcctcagccaagtccagctgctgcagttcagtctctc
agtcccgtcctgacccctcctgcccccccactcaggttcccccacctccagggagtcagaatgtgcctgagatca
cattacaaacattccctacaaggggtgggggtgggggggttcagcttggcatccaggacccttgactccagcat
gggtgggggcagctgcagaccccaacccatgcttgcctgctccaagccaaagtgttggccccagccccacacagc
tgttctagctcacagtcctggggaggccagctcaggggccctttccggggtggttagggacctgacctcttct
actttgtggctgattggccacaggtgggagcctgagagtgatgggaatccagcagctcctaaagtcccctgttcc
cctgcctcccctctcccacatagggcagccaggcaagggcaagaaggacctgtgaggctggggtggggactgaat
caagggttccctttcttcccttgggctctctgttgtctcatccgctcaggtgaatggggatgcgtgggccctgga
tactatggtgcatcactgttttctggggttagggactaaaggagatagagtgggccagccaggaggctgctagtt
gccatgggaacaccaggtttgggggagagtccactctgtgactattagaggagctgagatacgggggtgctcacc
agccacattctgggttccttaccccgcagaccatagcagcccctcttgggcgtgtttcacgatcattgttttggg
tcagggtgctagcagccaggaatgtaagctccttcagcttttacttctaaataatgctgtgtttggcatgctaa
ctaataggcaggcggcaccttgctacttacaatgcattctcatatatgatctcatcgaatgctcccaacaaccct
gggaagtaggtattggtcccatttccagataagaaggcatatgcccggacaagggaagtgacttgtctaaggcc
agtcaatgacttgcctatgggggcagagctgtgttgtgtgccattcattcctgctgccttgaaatgtaggagaga
caggaattcaaggtggggtttgggaagaggccctggctccctgcccctgagttcccctcttacctcgcaggtgtg
caagcatctgtccctgtgactcaggccctgtctgtgtggggtgagcagtgctccaacttcctcatttgacttttc
taggcctaccttccagggctgactgcttgggtgacctgaactaaaggcagccaactcactgcctggcacactccc
tgcggagctctgagtggcacctttggttcctctctgagctaagaatctctcttctcagggaccccacagaaagca
gacatcgacatcgaggaggggggggtattaaattcataatggacacgagccttagccaggtggggcagagggaagg
gcatagttcagtgtagatggggagggtgacttcctccaggactctcctggcttccctgcctctgacagagctgtg
tgtgccaggacccagctgctgttcctgcccacccacgcctgggcgggcactgctgggacatgccaggcctgctcc
tgccagcctcgggggctgggtggagctgggggagttaaagaggagcctttaatttgggggctccccaccctagct
gggatgagaacagacttgagctctggcttggggtctgtcctccattggtcttgccaccctgggaggtgcaggtgc
agtggtcgtggggtcactgaatcgaggttgaggctagaggcttctaccagacttgggcttgcatggcctagaacc
aggctccccctttggtcattctctgaccacaagttgctagagcctgagagcagggtgacagggaaagaagtatag
acctggtgggatgtgggcagctttctgggcaagaccctcctagctgcaagctctccctgccccctgtaaattgta
ttcacatgggtggagaagggcttatccggagaaaacaaggaggggctgttggttagttggtttctggccagtcc
agggccttggaggctactgtccccaacccccttcaggcttttcatatcccaggattctttgtggcagtgatggcat
cgtggtgcttctaagctaggtagaaaggctgaaggtagaacttgggttagggtagaggggggaagagctggaatag
cttaggccgcctcccccaagtccttttactccctggggtggcgccctggggtcccagggctatctgaggctggag
agggaaggcctcaggggttaccatggtgacctgcttcaggcttcctgctgagccatgtcaatccctttatttttc

agaaaagggctttgtctgggggtggccccagacagaaaagcttggcttggtcctagggtgggggggaagtaaggtg
gaggataaagagtttcacttcgggatatctcttcctcctcttcccttactcctactcctaatctggtaagagggg
agaactgtagggtccccctttaaagtatgactagtgttctacctagctttggccttcactctcactggctataca
cccatcccttatggcccgttaccactctcttgagcttctgacaggtcaagacgaggtttgcccagtggtcgggaa
gatgctaaggttaccagtatgggggatcagggtcagggaaggcgggtctcatcaggctccgctcccttctgcact
tcaggggaaggtgagtttccctctgccagctgccatgcaaatgaactaatgaatatttatgaagtctgttcactg
aggttgaagagactctgaggggttttgggggcaatagttctctcttctcccccctcctccatcatgtagacatct
gtttcctcaatgcctctggcttcaaggtagaaccccaattagctagaagccctgttccaattccagggctctgga
acctgggtggcaatgggagactgcacttggggtagacaggactccctggtgagtaagccagaaccgagtggttag
aaagctgagctggcctgtctccctttccgtgccctctatgtagctgcatatatatgtgttgttaagatgcaacag
tacaaggattcaggtaatgttttagggcctcttagacccatcgggttattgtcctgggtaggtttagacccctct
tttcaatgggaacattcatctgatcattcattccctcacatgaagggcctggactctaaggttcctgctggctct
aagatgtgagcagttgtctaatcctgccatctggatggaaccgttgtgagcactggaaagggaaaaaatgcctgg
agaattctagaggctttggggaaacttgttggctttgtatggtgcatgggacaggggtattccatctgaaatgtt
tatccaggcttccatcccatgcctacttcctttctgggccctctatagttaggtaccagtggggaggccatctca
ttctggcttttaccctcctgcacatcaagatggcagctcccaatttggtggagcctcactcgacatggaggaaag
gctgcttttcttcgtccacttcatccatgtggtcagtttccctcttcgcctacttgggtatcatttccctccctc
agatgaggacacaggtttggcattgcaaaggcatttgcccagagttcccagtaggccggtattgcttgtagtacc
tttggtgtaatgggccctcaggaaatggaaccactcttttctcagatgattgtggtagcctgtgaactggtgcat
gacactgtcagcttcattctctttaagcactgcctgaagccttccattgtccttcctgctctgagagcctgtctc
caatgggatgaaacccaccctgcttgctgtgacccttgaagcctttcagtctggccccaggtgaacttttcacct
cttcttctgtttctctcttctcactcccagccaagctggaccaatcactggcctgggctttgtgttgtagccact
gctgtctactgtctcagtgtgttgacttcttattcctcacctaagacccactttcatcctgcctttcctgtggat
acatctctgaccatcctggtcagatgagctatctctttcttctccctgcctccagcccttgctgcttgaatcatc
atcctcacacatacctacagatagcctggtgttgtctatcactcgagtatgggttatttacttttctgacccaaa
ctggacagtaaagtccttgacgatggagccatatcttctcttcctgtttcccctccaaaatatgagagtgagctc
ataggtacttagtgttagtatctgacctcctgcctcttgggagactgcttctaagaggcagtcaaacatgttgaa
cgcctattgtgtgctaggtgttctccatgtgctgccttgtctccccaccacaatattggaaggtaggggctcttg
ccacccccactttacaagagaagaaactgaagctcagaacaggcagctcatccaaggtcacagagttagtaaggg
cctagctttggaatctgattctaaagtctgtgttctttctcctacatggttctgccttgctctcagcaccagccc
atccagtctcagaaatggccaatggagcaccattttctttctgtgtgcttaggcaaggacagagtcctatggata
ggtggtcctgggctgctgaccctccctgctgtcccccagtgcctggccttctggattagaatatgaacctggctc
aaaaggcatttgtccatctcctctcaggagctggagtaggagcaacatagaggtgacaggttaggatacagatga
ggccctatcttctatgatctgttgtctctggttgagccccttctcttctcctctgttctttttaagaatgagatg
aatcaagtttaggaattcaggctcttgcattttgagggtgggatgaattgccagtaggacaataggcagtgggag
tgagcagagctgcacctggccagagactgtggccatcatccctttgctccaggcgatgcctgttcctatctttcc
atcatgaccaagattagtaaggatgggggagaagacatgggtcccagacagattgagctcccccagcctttggct
gggagctgcggggagtggtgactaagacatgtgtgggcttgctgtgtgtgtgtgtatgtgtgtgtgtgtgtgtgt
gtgtgtgtgtgtgtgtgtgtgcgcgcgcgcgcgcgcacgtgtgtatgggtataggagaggtagtagcagcggt
catgaaggaatctttgttgggggtgggggtatataagactgccacctccaggtagcatgggcctgacaccttggt
acccagcccattggccacgtcagccccctgcggaaatgccacattaggacaaagggctcccccagccaggcagtg
ctccaccctgccgagcagtcagagcctggcagccttgcccatgggccacccccttccacactctcctggtagtgtg
gttctctctccccaccccccagcctgaggatggaggaagggcagaaggcttgctcttgccatctccacccagtagg
cccagcccggctcttgtccgcttccctggttctggaatggcctggggccacccccaaccctgaggctgaccatct

tcttccttctgggcagGTTCCTGAGTGGGAAGGGCTTGGTGATCTATCCGAAAATTGGAGACAAGCTGGACATCA
TCTGCCCCCGAGCAGAAGCAGGGCGGCCCTATGAGTACTACAAGCTGTACCTGGTGCGGCCTGAGCAGGCAGCTG
CCTGTAGCACAGTTCTCGACCCCAACGTGTTGGTCACCTGCAATAGGCCAGAGCAGGAAATACGCTTTACCATCA
AGTTCCAGGAGTTCAGCCCCAACTACATGGGCCTGGAGTTCAAGAAGCACCATGATTACTACATTACCTgtgagt
cccgcccatcccatcctctggctctctccctgggcttaactctttcctctcctgtagtagtgggagcttctaagt
ggtgcaatgctattgcatgtagttaagaccctgctggatctgatccccttgaagactggcatgttctcctctta
gcctggccttggaattctcgccccaacatttaccaagctcaccttggctccaggggttgggcaggagaggtctcc
aattcgtgttgcttctctcttattttctttgcccacctaagttctaatattgggaatggtaacatatgccaggcc
ctttgggatcagccagctattcagccttttcttcaggggaaaccaaggccccaaaaggtgaagggacttcttgc
cttagatcacacagtgagttagagatagggtcaggactccaacccagttctcctgattcctactccagaattcct
ttcagtctatgtagaagcccccattatgatcccagtgaggaggcagaccttaccgaggggccatggcctgctcgt
gacagaggagcagtgcccctgggggtggggcttgttcttggcctgggctggactaggccattttgttcctagtgg
gagggaagaaaatgaaaaggttcactggtagatgggagctgcttgcttctcccgactgcagtttctctccctgtt
ggctgaagcagaatgggagtttctgggtaatgctagtagacctttctctcctcctgacttctctgacttctctgg
cctcttcctgcagCAACATCCAATGGAAGCCTGGAGGGGCTGGAAAACCGGGAGGGCGGTGTGTGCCGCACACGC
ACCATGAAGATCATCATGAAGGTTGGGCAAGgtgagtgcctagtctgagggtcccctcaccccaccctgtttgac
ctttggaacagatgttcctggctgggtgcatgtgtattaggagtgggggagcaggcgtagggttacagtatccag
gccattcttggcccaccttgatgactgagggcacctatgctggccgggtccctgcctctcacctgttctgtctc
cattcttagATCCCAATGCTGTGACGCCTGAGCAGCTGACTACCAGCAGGCCCAGCAAGGAGGCAGACAACACTG
TCAAGATGGCCACACAGGCCCCTGGTAGTCGGGGCTCCCTGGGTGACTCTGATGGCAAGCATGgtaagtgtatgt
gtttcccagaggtcaggagccattgctctgtcaccttgttaggccctgtccctgaagaaatgcaagctgggcctg
gcctgaaatctgctgtgtgtccctgggacccctggctgactgttccttcccctttcccttcctcagAGACTGTGA
ACCAGGAAGAGAAGAGTGGCCCAGGTGCAAGTGGGGGCAGCAGCGGGGACCCTGATGGCTTCTTCAACTCCAAGG
TGGCATTGTTCGCGGCTGTCGGTGCCGGTTGCGTCATCTTCCTGCTCATCATCATCTTCCTGACGGTCCTACTAC
TGAAGCTACGCAAGCGGCACCGCAAGCACACACAGCAGCGGGCGGCTGCCCTCTCGCTCAGTACCCTGGCCAGTC
CCAAGGGGGGCAGTGGCACAGCGGGCACCGAGCCCAGCGACATCATCATTCCCTTACGGACTACAGAGAACAACT
ACTGCCCCCACTATGAGAAGGTGAGTGGGGACTACGGGCACCCTGTCTACATCGTCCAAGAGATGCCGCCCCAGA
GCCCGGCGAACATCTACTACAAGGTC**TGA**GTGCCCGGCACGGCCTCAGGCCCCCGAGGGACAGTCGGCCTGGACC
GGACCTCTCCTTTCGCCCCCACACCCCCTCCCCTTGCCAGCTGTGCCCACCTTTGTATTTAGTTTTGTAGTTTCT
TGGCTTTTATAATCCCCCTTTTTCCCTGCCCCCTGGGCTTCGGAGGGGGGTGCTTGTGCCCCTAACCCCCATGCT
CTTGTGCCTTCCCCCTCTGGCCAGGCCTCTGGGCTCCGTGGGGGCGCCCCTTCTTGGAAGGCAGGGCTGGACACT
GATGGACAGCAGGCAGGGAGACAGTCCCCTGGCCCTGCCCCTCCCTCGCCCCCCTTGCCACCTTCCCAGGACTGC
TTGTCCGCTATCATCACTGTTTTTAATGCTTTTGTGTTCATTTTTTAGCTGTCAACTCATTTTCATCTGTTTTTT
GAAGAAAAATGGAAAAATGTAAAAGGCAGCCCCTCCCCAGGCTTTGTGAGCCTGGCCCAAGCCAGTACAAGAGGG
CCTGGGGCACGATGTGGTCAGCCAGGAAGCATAGGATGCCATTTCTTTTATAGATTCCTTGGTATTTCTGGTGGG
GTAAGGGGCAGGCCAGGGCTGTTCACGCCCATGAGGGAAGAGGAAAGTGCCACTGGGCAAGGTGTCCCACCCTCC
CCTCCTGACCCTCCTACGAGGCTTATCCTGGCAATGGGGTAGTCACTGCCACCCTTCCACACACACACACACACA
CACACACACAAAAAAAAAATCCCTTCCTTGTGGGATTCTTGGGCATCTCCTGCCTCCCTCACTCTCACGGTAATTA
ATGTCTTAATTGGCTGTTGCCTGGGGAACAGGAGAGCTGCTGCAGGCAGATGACCTCATGGGGGGGTGGAGGGAGG
TGAGGTGCCCAGGTGGCTATTTGCCCTGCAGAGCTGGGAGTTTCACCCCCACCCCCCACCCTGTTCTCTCCTTAC
CTTTGGCATCCTTTGGCCTGGTGGGGAAACAGAGGCCCAGGGTGGAGACCTAAGCGGGTATAAGACCAGGTGGCC
TGCTCCTTTTCTGGGCCCTAGCACAGGTGGGTAACCCCCACCCAACCCAGCTCCTGCTGCTGTCCCAGTCTTGGG
CTGGGGCCTGGAAAGAGGAAGAGGCTGCCTGGG

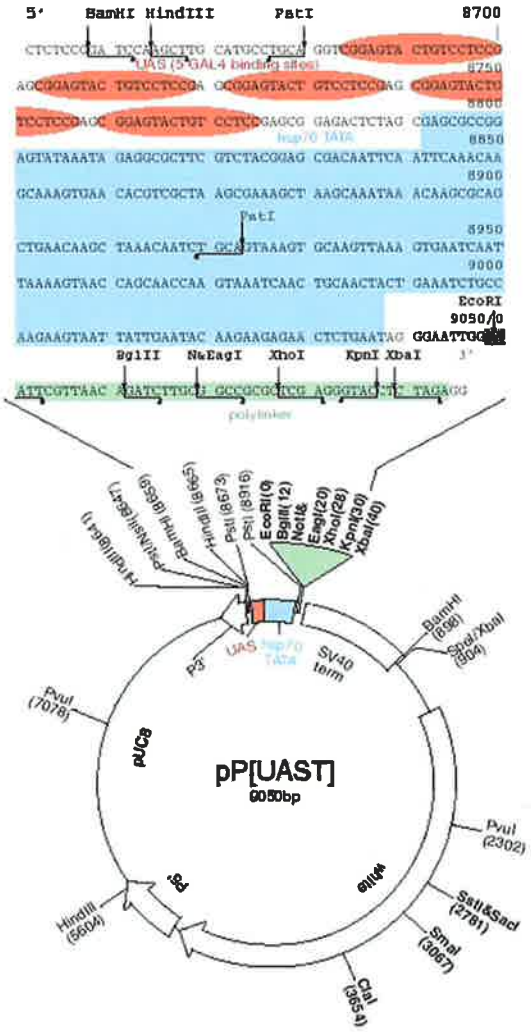
Appendix D Genomic structure of d-ephrin

Start and stop codons are in bold, and intronic sequence is in lowercase. Also the Cavener start of translation site is highlighted.

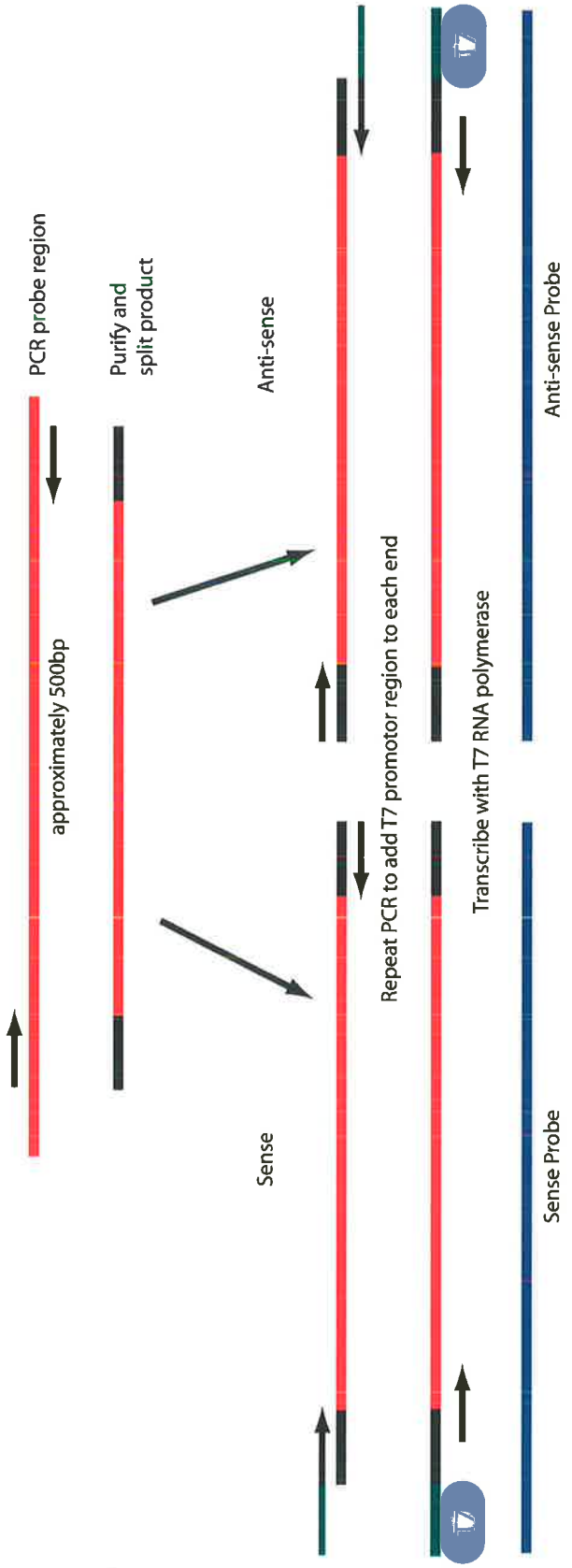
TTTTTTGTTCCAGCTCTAAACAAGTTTTTTTTATAGCTCTAGCTCTACAGTATTTATATAAATAATTCTTGTATG
TAGTCCAAGTGTAGAAAAAACTAATTTAAATGAGCTGTTTTCGAGTATTTCTCTTCAGACAAAAACCTGTTTGGA
ACAATTTTTAATCATGTGAAAATgtaattttgaaaaagatgtaccaaaaaaggaagacggtgttatgggcatata
tgtatattttgtgtttttatatgtcaactatgttataccaacatatttgtaccggattgttaatcgtatcgaata
agaatgtttccccgagagtaagatgtttgtgtatatatagagtttcttctgtctcctcggagagtgtgtcgtcac
tgtgtatgtatacagaaatgtcttttggtgtatacatacatatgtaattgtataatgctgttgggatgttttatc
acttaaacatacaatttttaagaaatatattgattgttccgttataaattgaatgtcaagaactgcagatccgct
atatattgttgatcttatgccatgttgattgtttatgactatgaatgttaaaaaatatattaatatttcgccaag
ccattaccaaccctgtttgtatagaaaatgtaccatacatgtattaaaatgtagtatcccacttaaattgataat
tggtttgccggcctgttaaataagtatttttgatcaatcagggtcgaatcaaatcaattaggcggtttccaaagt
taaagcgaagtacttttaaaaatctcgtaatctgtcaggtggcggtagtcaacatataaatattgttgatcagca
ttaacagctgcttatttcgccgactgccgccgaacgcggcgcaaacaggatccctcagtcttaggctgtttgttt
taaagatcaacttgtataaactttacataaaaatataaatacgcataatatgctactgttacattaaaatgttaa
cgcgtaatgatctcaaagcttagaatgagtagatataatatgtagttagattttgcaatgggtaaactaatttca
gtcaggattgtcaacgaagtaaaggaggaaatgtcaatagaagaacacatattcttgatcagtttacacagccat
gtcgatatatttttttgtctttctttccgtttgccaaacgaccttcgaattgaaggttgaaatccttttatttttgt
agttcatatacatatacgggtcgaaaaagaatgccgttaaagtattgtttctttttttaactgcctagtaaacttt
ttgacgagataaattgaaaagtataattccgctctaaaaaagtcgggttaaattttctaagaaccctaggtatta
tatcaatttcattttcagtatcaaaatttaatcttgaagaggtgaaaactaagtctacgttaaaaagaatactag
acctaaatcacagatgaaaaagaaacaataatttaacggccttttcaatttaatttttcggattgttcctatat
atgcatatatatttatatagtgatatatatgcacaaaagttgtattatttaactattgtcatcaaaaattaaatt
gcacctgactttataacgcactacatataatttactcatttataagaaatcttttaaagggtaaagtaaacaacc
aaaccgcatagtagtctttagtctgttggtgcgcttcgtaactataccgtccaaagtgagaaaatcatgttacaa
gttacactgcagtttcagtgttttttttttatgcccagaattcgtacacggctacagcgttttgttgtaaaaatt
gttattttattgattgcagATGGAGCAAGACACTTTGGTGCGGAATGACACAGATGTAGATGTAGCCAGTCAACA
TATTCTTATTTCGGAGGTTTATTGTATGCATTTAGCAAGAAAGCTACTAACGATCTAACATCGAAAACATTGCTG
GCGGTATGCAAGAACGATCAAAGCAACTGCGTCTTACTGTTTCCTGGACCCAGTCAAAGTCTCAGATTCATGACA
GTTGTCAACGTCGGAGCATGTTGGCATGTAAGAGGCGCTTAACTACCAGTAAAGTGCTTGAAGACTCACATCCAC
CAGTAGCTTTCCCCAATTGCAAATCTCACCGGCATCAGCAACAAAAGGAGAAACATAAAGTTCAGCTATACTCTG
GAAAACCTTTAAGCATTAAGCTGTACGTGCCAGGAAGTATTGAGAGTATCCCAAAAATTCGACATAAGGCGCTTA
CGACAACGAATAAACAGCAACCAGCTATGCATCGAAAGTCTAAGAGCAAAAGCAAGTTTCAAGGTTTTAACAATC
TCAAGCCGCTTTATTCACCGAGTAAAAGACAACCACCCGAGAAGCATTCATCAGTACTGGTGGAAGCAGGGATAG
AATCAAAGGCATCACGGCATTTTGTTGGTAAAAAGAGGATTAAGAACAGAAACTGTTTGTTATCTTCGCCTCAGC
CATCGCCAATGAGATGCAAAATGATGATTCCATTTCCAAAGTTTGGTGCCACATCCTTTGTTACGTTGCTCACTT
TAATTTGTATGGAAACTGTTTTGCTCTCCACCATGTCTAGTTGCGCCAAGACTTTTTACATGCATTGGAACACAT
CGAACAGTATgtaagtactcaaatatagtggtatacttaagcagcggtattataatttattgaagttaaaaatt
gtagatttttatttttctctacaaatgtattatgttttattacaattattacagttttaagtcagtagtatgag
gattattattattcgtttaggggatttgttatttcaaataaaattttctttccaataataagactgctggccaat
aatatattgataaattaatccctaatgtcgcgtataaacacgctgggcttttataaactttcccagcaattagta
atgcacatgagtcattttcgatagaataactattgtcataatttgttgtgcaagtgagcagacgatgattaggcg

atcccattgatccctaaggatcatgattagtaacttttttggtaatcgcatcagggcatatatggcctactcatt
ttgtctggttttctctttctcgatataaacatatccatatcatctataccatccatattgcaatgggacattaag
tttgttaggataatgttacttccaaaatcggacaacaaaccccaaaaacagaagacaaatgcacgttttcgtagc
caaggcattaaacaataaagtctagatcagataaaaacaacccaattttccattttcaaaatgttaacttcagca
ttttcataaactgttgcttaaactattctgtgcgaactatttcagcattttccaaatctaaaattatatatgata
tgtagttccttgaaccctacgcgttcagtgctcatcgttactatgtgtgatgttttgcacattaattggcccaaa
tttggtatttcgatttgcgcttatttattgaaaatcgttttctgactttctcaaaacaaaatgtaatttgattgc
actcttatatccacagATTTCGGATTGATAACACAGATCATATTATCGATGTTAATAAAGGCAATCTTGCATTTG
AGTTCGATCAGGTTCATATAATATGCCCAGTATATGAGCCAGGGACTTTTGAGAACGAAACTGAAAAATACATAA
TTTACAATGTGTCTAAAGTGGAGTATGAAACTTGTCGCATAACAAATGCAGATCCGCGAGTAATAGCTATATGTG
ATAAACCTCAGAAATTAATGTTTTTTACAATAACTTTCCGGCCATTTACACCGCAGCCAGGTGGCTTGGAGTTCC
TACCTGGAAATGATTATTACTTTATTTgtgagttcgtttgcattccttctatttaaaaaataatatttttaatat
ttatttagCAACTTCATCTAAGGATGATTTATATCGAAGAATTGGAGGTCGATGCTCTACAAATAACATGAAAGT
CGTCTTTAAAGTGTGTTGTGCCCCAGAGGACAACAACAAAACCACGGCGCTAAGCAATTCTAAATCTGTTACAGA
CACCGGAGGAGCCATTAATGTCAATATAGCGAATAATGATGAAAGTCATGTGAATAGCCACGGCAATAACATAGC
TATTGGAACCAACATTGGTATAAATGGAGGCCAAATTATAGGGGGACCGCAGTCGGCAGGAATTCCAATTAATCC
ACTAAGCGGGAATAACAATATAAATGGCATACCAACTACTATTAATTCAAACATTGATCAGTTTAATCGGATTCC
AATCCAGCCAAACATAATCGGTAATCATGTAGGGACTAATGCCGTAGGAACCGGAATTGTTGGTGGTGGAGGAAT
AATATTAACTCCTGGCCATGCTCATGGCAACATTAATATGCTGCAACCAGGGCGAGGTGGAATAAACGGAGCATA
TCCCGGACATCACCACATCCAAACTGGGATACGGATAAACAATGTGCCTACGCAACACAACTATCCGTCCCATAA
GGGTAATGCTAACAGTAATATTAACGGAAACGgtacgtaatggccccatccaaatccacaaaaggagcgaacttt
tgaagcttacaggcaaaaatgtactccttgcgtacgctctattcaaactgtttttatttgaggcaatcggttac
ttatttgctaaattacatttcctacttagaaccaatattttcacgattccagtgctttttaaaacaaaaatttta
gctgaaaaactgtcttttaattctaaagaaagttacaaacgagattttaatgtttttttgtttcctatgtctgcct
ttaatcgaatctgcatgccaatttccaactttctagcttttatagttcctgagattcgactttcatacggacaca
caggcagacgaaatgactttagtaatacaaatgatttatttgttatttataatatcagaaaaggtatatagata
aagaaacacatagagcacatgaaaaactgaaaaaactgaacactttacataaaatagcccgattcaaagtatttt
tatttaaatttttttcaagtgtatacactataaattgttcataacgatctaaatctgaatctgaaagtcataacat
tcgactaaattgctgtactgtgtttgcgatataaaatatacttcatattgtttatatttttaatatttataataa
atattttcagATGACCACCATCATTACAACAAACATCCCAACGAGGTTGTAAAAAATGAAGAGCTGACCTATAAT
AGTGGTGCTGCGACATCGGATGGTAACATCTTCGCTTTATGGATCTGGATTTTATCAATTTTCCCACTGCTATCT
ATTCAATCTTGCCATTTGTCTTCATATTGGATAAGCGCATCATTTTTAGTCAGCACTATTGCAATTCTTGGCATT
CACTATCTTATTCAAATCACTTTGCAAACCACGGTGCAGCGATATAGTCCTGGAATGGTTGAAATCACCGCGACC
TCTATGAACGGGATGTTTGACCAGAATGCTGGCACTATTGAATATGACCGGTGAATTTTTGATGAACATTCGATT
TTGTGTTCCAATGTTCAACGTTATTTTAAAGAATAACAATATCACATAAAAATATTTGTCACTCTTCTCCGGCTA
TTGCTATATAAACAGTGATAAAAGTATTAGGTGTGAGAGGTTGGAATCAAATGATATTAATTTGGATATGTAATT
GCAATTGAACGAGTTACTGTGACTCGCAATTTAAAATGAT

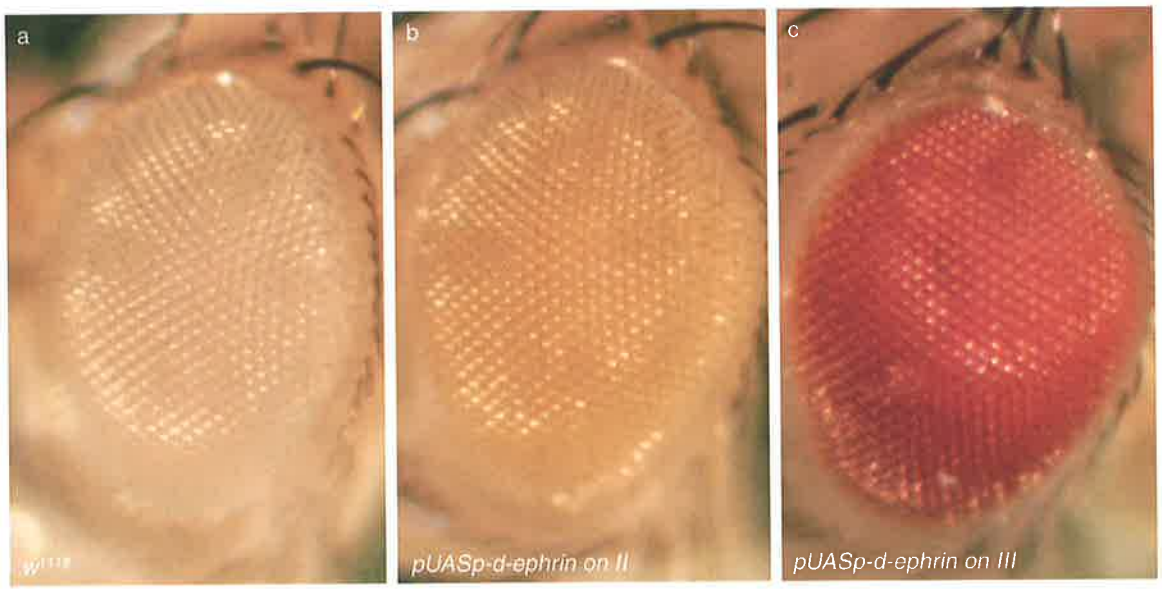
Appendix E Restriction map of pUAST expression vector (Brand and Perrimon, 1993).



Appendix F Making Sense and Anti-sense *in situ* probes with the same promoter via PCR



Appendix G Variation in eye colour obtained in injected flies



(a) w^{1118} white eyes (pw-). (b-c) pw+ expression from two separate flies transformed with pUAST-d-ephrin.

Accept the challenges so that you may feel the exhilaration of victory.

My mate James

Addendum to thesis Entitled "Investigations of ephrin ligands during development"

Chapter 1
Page 34: Receptor-Ligand binding facilitates repulsion
It should be emphasised at the end of this section that receptor ligand binding facilitates **retraction** of the axonal growth cone not inhibition of cell growth. This was shown by Hattori *et. al.*, 2000, where the expression of a dominant negative form of *Kuzbanian* delayed axonal retraction. Therefore giving a clear mechanism by which receptor-ligand binding can facilitate axonal repulsion.

Chapter 2
Antibody concentrations
Primary Antibodies
Qiagen - Anti·His Antibody Selector Kit (1/4000)
Cell Signaling - Myc-Tag 9B11 (1/2000)
Cell Signaling - His-Tag (1/4000)
Cell Signaling - HA-Tag 262K (1/4000)
Rabbit - α - cEphA4 (1/5000)
mouse - α - 22C10 was developed by (1/2000)
mouse - α - BP 102. was developed by (1/2000)

Secondary Antibodies
α-mouse-CY3 conjugate Jackson Laboratories (1/2000)
α-rabbit-HRP Jackson Laboratories (1/2000)
α-mouse-HRP Jackson Laboratories (1/2000)
α-rabbit-AP Jackson Laboratories (1/2000)
α-mouse-AP Jackson Laboratories (1/2000)

Chapter 3
Page 84: Last sentence "Clearly, this needs to be tested *in vitro*". For a description of some proposed *in vitro* studies please see page 91

Page 85:
The sentence "In summary, the membrane anchorage of d-ephrin is unable to be accurately detected by TMHMM analysis, however this program can accurately predict transmembrane regions for all ephrin-B gene or the lack thereof for all ephrin-A genes",
Should be changed to:
In summary, the membrane anchorage of d-ephrin is unable to be accurately detected by TMHMM analysis, despite the fact that this program can accurately predict transmembrane regions for all ephrin-B gene or the lack thereof for all ephrin-A genes.

The Sentence "Although the transmembrane regions predicted for the full ORF are a good candidate for membrane anchorage.
Should be changed to:
Nonetheless the transmembrane regions predicted for the full ORF are a good candidate for membrane anchorage.

Page 91: After "All these apparent differences of d-ephrin may simply be due to a sequence error in LD11109, which results in an error of the predicted gene in the BDGP, however the full length of LD11109 was sequenced to confirm the predicted ORF listed on the BDGP and no errors".
Add. The sequencing done by me confirmed the *Drosophila* genome database LD11109 sequence, ruling out any errors.

Chapter 4
Page 108: The sentence "Analysis of *sim-GAL4*::*UAS-d-ephrin* embryos stained with BP102 showed normal formation of the anterior and posterior commissures, and longitudinal connectives, at all stages when compared to *w1118* embryos (Figure 30). **Is referring to** Figure 31.

Page 110:
The sentence "The intersegmental nerve roots of *elav-GAL4*::*UAS-d-ephrin* embryos appear to fuse more proximal than those of wild-type embryos (Figure 33a,b)"

Should be changed to
The intersegmental and segmental nerve roots of *elav-GAL4*::*UAS-d-ephrin* embryos appear to fuse more proximal than those of wild-type embryos (Figure 33a,b).

Page 111
Change Figure 31 to Figure 32.

The sentence "There are two intersegmental nerve roots (white arrowheads) fusing to form one axonal tract at the very lateral edge of ventral nerve cord (VNC) (white arrow)
Should be changed to
There are two intersegmental and segmental nerve roots (white arrowheads) fuse to form one axonal tract at the very lateral edge of ventral nerve cord (VNC) (white arrow)

Page112
The sentence "This suggests a ligand receptor relationship between the d-ephrin and Dek.
Should be changed to
This is consistent with a ligand receptor relationship between d-ephrin and Dek.

The sentence "D-ephrin may localised to the cortex region of the ventral nerve cord (VNC), which would be expected if it was providing a repulsive mechanism to prevent axons from aberrantly exiting the VNC"
Should be deleted

Page 113
The sentence "This result is unexpected, as *sim-GAL4* should be driving expression of *d-ephrin* from stage 10 onwards, which should preface *dek* expression from stage 14 onwards"
Should be changed to
This result would be unexpected if *d-ephrin* were normally acting to repel axons away from the midline fromw stage 10 onwards which should preface *dek* expression from stage 14 onwards.

Second paragraph add
Another possibility is that d-ephrin plays no role in repelling axons away from the midline.

Page 114
The sentence "While these results do show a potential phenotype, the expression levels of *d-ephrin* need to be tested in the misexpressed embryos, with either antibody stains or *in situ* analysis to increase confidence in the phenotypes shown, due to time constraints this was not done"
Should be changed to
While these results do show a potential phenotype, the expression levels of *d-ephrin* need to be tested in the misexpressed embryos, using anti-d-ephrin antibodies.

Chapter 5
Throughout this chapter the species name *Ctenophoris elegans* **should read** *Caenorhabditis elegans.*

Chapter 7
Page 132
The sentence The vertebrate Ephs are divided into two subclasses EphA and EphB according sequence homologies and their binding specificities to the ephrin-A and ephrin-B ligands respectively (Figure 1,Figure 41a).
Should be changed to
The vertebrate Ephs are divided into two subclasses EphA and EphB on the basis of sequence homologies and their binding specificities to the ephrin-A and ephrin-B ligands respectively (Figure 1,Figure 41a).

Page 134
The sentence The *Drosophila* genome encodes for at least one ephrin protein
Should be changed to
The *Drosophila* genome encodes at least one ephrin protein

Page 135
The sentence Where the addition of *UAS-Kuz* to the *UAS-d-ephrin* cell line may cause perturbation of the cell aggregates shown to be formed between *d-ephrin* and *dek* expressing S2 cells (Figure 42).
Should be changed to
The addition of *UAS-Kuz* to the *UAS-d-ephrin* cell line may cause perturbation of the cell aggregates shown to be formed between *d-ephrin* and *dek* expressing S2 cells (Figure 42).

Page 136
The sentence Currently, there is a large body of literature accumulating in regards to the involvement of Eph/ephrins in cancer.
Should be changed to
Currently, there is a large body of literature accumulating in regards to the involvement of Eph/ephrins in cancer (reviewed in Dodelet, 2000; Easty,2000).

References Add these two references
Dodelet VC, Pasquale EB.
Eph receptors and ephrin ligands: embryogenesis to tumorigenesis. Oncogene. 2000 Nov 20; 19(49):5614-9. Review.
Easty DJ, Bennett DC. Protein tyrosine kinases in malignant melanoma.Mela noma Res. 2000 Oct; 10(5):401-11. Review.