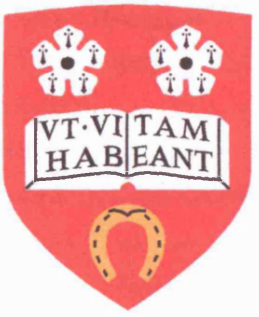
The role of *Enhancer of split* and *groucho* during neurogenesis in *Drosophila melanogaster* and *Musca domestica*

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester



by

Barry Denholm

Department of Genetics

University of Leicester

February 2000

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

CONTENTS

The role of *Enhancer of split* and *groucho* during neurogenesis in *Drosophila melanogaster* and *Musca Domestica*.

The bHLH genes of the *E(spl)-C* and *groucho* are required for allocation of appropriate cell fate in several developmental contexts, where they act to repress specific fates within a group of equivalent cells. To better understand the function of these genes, the work carried out in this study has investigated the role of the *E(spl)* and *groucho* genes during neural fate commitment in the fly. It is known that the carboxyl-terminal tryptophan-arginine-proline-tryptophan (WRPW) motif of E(spl) binds Groucho to form a complex, which represses the transcripton of target genes. The importance of specific residues within WRPW has been investigated by generating a number of mutant derivatives containing single amino acid substitutions within the motif. It has been found that changes in WRPW abolish the *in vivo* function of the protein, and attenuate interaction with the Groucho protein. To determine the mode of E(spl)-mediated regulation, a series of co-expression assays were performed. *E(spl)* has been ectopically co-expressed with proneural genes *scute* or *daughterless* during allocation of imaginal SOP cells. It was found that E(spl) did repress the neural fate in the context of the co-expression assay, suggesting that, in addition to transcriptional repression of the proneural genes, post-transcriptional modes of regulation also occur. The requirement for an intact WRPW motif further suggest that this mode of repression may involve Groucho. Finally, a region of the groucho gene from the housefly *(Musca domestica)* has been cloned which encodes the C-terminal WD40 repeats and part of the variable region and displays a high degree of identity with *Drosophila* Groucho in these regions. In the *Musca* blastoderm embryo *groucho* mRNA is ubiquitously expressed, but later becomes confined to the developing CNS. A preliminary functional analysis using the technique of RNA interference suggests that *groucho* plays a role during neurogenesis in the *Musca* embryo.

ACKNOWLEDGEMENTS

Many thanks firstly to Fred, for support, encouragement and excellent, much needed supervision, and who has taught me most of what is found between the two covers of this thesis. Additional thanks for the bed and the loan of the computer. Much thanks also to: Those who have provided clones and flies; Yiorgos Apidianakis in Christos Delidakis' lab and Ze'ev Paroush for yeast plasmids and Juan Modollel and Uwe Hinz for fly lines, without these much of the work could not have been done. To everyone in 136, for help and advice and making life in the lab a pleasent and fun place to work over the past three years. Special thanks to Ian for the yeast two-hybrid stuff and Alistair for the *Musca* maggots and his fragrantly challenged incubators. The kitchen staff, Keith and Maria, for generating a mountain of gourmet fly food (9 out of 10 flies preferred it). The 'Bossa club', you know who you are, with whom far too many cocktails have been consumed. To Audrey for support, encouragement, comments on intial drafts (often candid but always useful), and for helping with the printing out. Finally, to Mum and Dad for their continuing, unwavering support without which all of this would not have been possible.

ABBREVIATIONS

Genes

ac	*achaete*	*E(spl)*	*Enhancer of split*
AS-C	*Achaete-Scute Complex*	*E(spl)-C*	*Enhancer of split complex*
ase	*asense*	*gro*	*groucho*
ato	*atonal*	*kuz*	*kuzbanian*
Brd	*Bearded*	*l'sc*	*lethal of scute*
da	*daughterless*	*N*	*Notch*
dl	*delta*	*sc*	*scute*
dpn	*deadpan*	*Su(H)*	*Suppressor of Hairless*
emc	*extramachrochaete*		

Reagents

ATP	adenosine triphosphate	EtBr	ethidium bromide
amp	ampicillin	EtOH	ethanol
β-gal	β-galactosidase	IPTG	isopropyl-β-D-thiogalactopyranoside
dATP	2' deoxyadenosine triphosphate	ONPG	*o*-nitrophenyl-β-D-galactoside
dCTP	2' deoxycytidine triphosphate	PEG	polyethyline glycol
dGTP	2' deoxyguanosine triphosphate	RNA	ribonucleic acid
		SDS	sodium dodecyl sulphate
dTTP	2' deoxythymidine triphosphate	TBE	tris-borate EDTA
		TEMED	N,N,N'N'-tetramethylenediamine
DIG	digoxigenin		
DNA	2' deoxyribonuleic acid	Tris	Tris (Hydroxymethyl) aminomethane
dsRNA	double stranded RNA		
DTT	dithiothreitol	X-gal	5-bromo-4-chloro-3-indolyl-β -D-galactopyranoside
EDTA	ethylene diamine tetra-acetic acid (disodium salt)		

Units

bp	base pair	nmol	nanomoles
g	grammes	pH	acidity [$-\log_{10}$(molar concentration of H^+ ions)]
kb	kilobase pairs (10^3 bp)		
l	litres	pmol	picomoles
mg	milligrammes	rpm	revolutions per minute
min	minutes	sec	second
ml	millilitres	U	units
M	Molar	µg	microgrammes
mM	millimolar	µl	microlitres
µM	micromolar	vol	volume
ng	nanogrammes	W	watts
nm	nanometres		

Miscellaneous

AD	activation domain	ORF	open reading frame
BD	binding domain	PCR	polymerase chain reaction
bHLH	basic-Helix-Loop-Helix	PNS	peripheral nervous system
cDNA	complementary DNA	PNE	procephalic neuroectoderm
ch organs	chordotonal organs	RNAi	RNA interference
CNS	central nervous system	SO	sense organ
DOP	drop-out powder/plate	SOP	sensory organ precursor
es	external sense organs	UAS	upstream activation sequences
GMC	ganglion mother cell		
GOF	gain of function	UV	ultra violet
HLH	Helix-Loop-Helix	VNE	ventral neuroectoderm
LOF	loss of function	VNS	ventral nervous system
mRNA	messenger RNA	WRPW	tryptophan-arginine-proline-tryptophan
NB	neuroblast		
OD	optical density		

"Time flies like an arrow, fruit flies like a banana"

Groucho Marx (1932)

Chapter 1

Introduction

1.1. Introduction.

As the development of a multicellular animal proceeds, cells become different from one another as they gradually become allocated to more and more precisely determined fates. At the root of these differences is the activation of different genes in different cells. How orchestration of differential gene expression, and hence allocation of cells to alternative fates is accomplished is a question that has long interested developmental biologists.

In the fly *Drosophila melanogaster*, spatially organised information is deposited in and around the egg by the mother and this information provides the co-ordinates that will initiate pattern formation in the early embryo. This maternally provided information will interact with the genetic information carried in the zygotic nuclei of the fertilised egg, and because the maternally derived information is not symmetrically distributed, gene activation in different nuclei becomes dependent upon position within the egg. Once cellularization occurs however, other mechanisms which control differential gene expression also become important. One of these mechanisms involves cell-cell interactions. Throughout development cells communicate with one another, and as a result of receiving and sending signals gene activity within the cells is altered. In some instances, the consequence of this cell-cell communication may result in the adoption of differential cell fates. An example that highlights the importance of cell communication for patterning during development is exemplified by the induction of the R7 photoreceptor fate by the R8 photoreceptor cell during *Drosophila* eye development. The adult *Drosophila* eye is built from many twenty-cell units called ommatidia, each containing eight photoreceptor cells and twelve accessory cells. The cells that make up the eye initially all appear to be members of one equivalence group (Lawrence and Green, 1979) and specific cell fates are then sequentially allocated by local cell-cell interactions between the cells of the ommatidium. The R7 photoreceptor is the last of the eight photoreceptors to be recruited and is induced by an inductive signal from the R8 photoreceptor. In the absence of signal, the presumptive R7 cell is not induced and adopts an alternative fate (Reinke and Zipursky, 1988). In this example, the fate of one particular cell type is wholly dependent upon a signal generated by a neighbouring cell.

In both the developing central and the peripheral nervous system (CNS and PNS) of the fly initially equipotent cells in the ectoderm adopt either one of two fates: the neural fate or the epidermal fate. The allocation of differential fate in this context, as discussed above, is a consequence of differential gene activation which, in turn, is a consequence of cellular communication between cells in the ectoderm. Historically, CNS and PNS development in *Drosophila* have provided important models in which to study cell fate determination, both at the level of changes in gene expression and the cell communication processes by which these changes are brought about.

The work described in this study, following in the tradition of a large body of previous experiments some of which are described below, utilises the adult fly PNS as a model with which to study the mechanisms of cell fate acquisition.

1.2. Development of the central nervous system.

1.2.1. Development of the embryonic CNS.

In insects, neurons are generated by the proliferation of progenitor cells called neuroblasts that develop from a particular region of the embryo known as the neuroectodermal region. In *Drosophila*, the neuroectoderm becomes morphologically distinct shortly after gastrulation at a time approximately 4 hours into development (corresponding to stage 8 of embryonic development, Campos-Ortega and Hartenstein, 1985) as the cells in this region begin to swell and increase in size. There are approximately 2000 cells in the neuroectodermal region, of which approximately 500 will develop as neuroblasts. The remaining 1500 cells become epidermoblasts, the progenitor cells of the ventral and head epidermis. The neurogenic region can be divided into two: the ventral neuroectoderm (VNE) from which the neuroblasts of the ventral nerve cord develop, and the procephalic neuroectoderm (PNE) from which the brain hemispheres emerge. The two parts of the neuroectoderm demonstrate differences in terms of both the kinds of constituent cells and their behaviour. The VNE has been the most intensively studied and a summary of the morphogenetic events during neurogenesis in this region is given below.

At stage 8 of embryonic development the cells in the neuroectoderm enlarge as they adopt the primary cell fate, the neural fate. Starting at stage 9 the neuroblasts begin to delaminate from the neuroectoderm and move into the interior of the embryo, coming to reside between cells which will ultimately form mesodermal and ectodermal derivatives. Segregation of the full neuroblast complement takes approximately three hours and occurs in five discrete waves giving rise to five subpopulations of neuroblasts. During neuroblast segregation the prospective epidermal progenitor cells undergo characteristic cell shape changes and establish intimate contact with the neuroblasts. This contact between the cells, before and during lineage segregation, enables cell-cell interactions to be established which are important to control the correct allocation of the neural and epidermal cell fate. After neuroblast segregation is complete the remaining cells in the neuroectoderm shrink and ultimately follow the epidermal fate (Campos-Ortega and Hartenstein, 1985).

Immediately after leaving the neuroectoderm the neuroblasts enter mitosis and divide asymmetrically to produce a single ganglian mother cell (GMC). Each GMC then undergoes one equal division to produce two neurons. The GMCs and neurons come to reside between the neuroblasts and the mesoderm and form an irregular layer of increasing thickness on top of the neuroblasts. The main period of neuroblast mitosis occurs between stages 9 and 13 (between 4.5 hours and 11 hours). The number of neuroblast mitoses appears to vary, some neuroblasts divide only once whereas other may divide up to 10 or 11 times (Technau and Campos-Ortega, 1986; Prokop and Technau, 1991). At the end of these divisions the embryonic neuroblasts shrink and cannot be distinguished from the remaining cells of the CNS. The neuroblasts will begin to enlarge again at the first larval instar stage to produce the neurons that become incorporated into the imaginal disc CNS (Truman and Bate, 1988). Neuronal differentiation begins at stage 13 when a population of neurons lays down a scaffold of fibres on the dorsal surface of the CNS along which later appearing axons are guided.

1.2.2. Development of the adult CNS.

Unlike most other larval organs, a proportion of the CNS persists into the adult stages, therefore most motor neurons and large interneurons of the adult nervous system are of embryonic origin. To this set of embryonically derived neurons, a large number of neurons are added during the larval and pupal stages. The neuroblasts that generate these postembryonic neurons are the same neuroblasts that populated the larval CNS in the embryo, which appear in the first larval instar on the outer surface of the CNS and resume their proliferative activity. In a similar manner to embryonic development the neuroblasts divide to produce single GMCs. The progeny of the GMCs, the presumptive adult neurons, remain undifferentiated until pupal stages. A review of adult CNS development is provided in more detail in Truman *et al.* (1993).

1.3. Development of the peripheral nervous system.

The peripheral nervous system (PNS) is comprised of both sensory and motor components. The dendrites and cell bodies of sense organs are located in the periphery of the animal with axons projecting into the CNS. Whereas the cell bodies of the motor neurons are located within the CNS and extend their axons to the periphery to innervate the muscles. In contrast to the CNS, the larval PNS degenerates during metamorphosis and the adult PNS forms *de novo* during the late third instar and early pupal development. The following description focuses on the development of the sensory components of PNS.

1.3.1. Development of the embryonic PNS.

In the embryo percursor cells of the larval sensilla, the sensory organ precursor (SOP) cells, can be first identified in the ectoderm at stage 11 of embryonic development. The SOPs develop from specialised parts of the ectoderm called proneural clusters containing 7-15 cells, which are delineated by the expression of the proneural genes (section 1.4). Within the cluster cell-cell interactions take place to select a single SOP; the remaining cells within the cluster become epidermal precursors and subsequently develop as epidermal cells. The selection of the single SOP from the group of cells is under control of the neurogenic genes.

These interactions are analogous to the type of interactions which took place earlier in development to determine the fate of the neuroblast in the neuroectoderm. Once selected the SOP divides during embryonic stages 11 and 12 to produce the progeny cells that constitute the mature sense organ (figure 1.1). Typically, a sensillium contains one bipolar sensory neuron and a number of accessory cells. Two major classes of sensilla exist: the external sensilla which include mechano- and chemo-sensory receptors such as bristles, and the internal chordotonal organs which are stretch receptors. In the case of the external sensillia, the accessory cells remain in the epidermis and produce specialised cuticular apparatus responsible for receiving sensory stimuli (figure 1.1). In the case of the chordotonal organs both the neurons and accessory cells lie subepidermally. Sensilla start to differentiate during embryonic stages 13, 14 and 15 as the sensory axons begin to migrate toward the CNS.

1.3.2. Development of the adult PNS.

The complete adult PNS forms *de novo* during late third instar and early pupal development from the imaginal discs and histoblast nests. The mechanism by which the sense organs develop is identical to the mechanism occurring during larval PNS development, in that a single SOP cell is initially chosen from a group of equivalent cells, the proneural cluster, and then this SOP subsequently divides to produce the progeny which constitute the mature sense organ (figure 1.1). The SOP cells begin to form during the late third larval instar and appear in a strict temporal sequence. Among these early SOP cells are some destined to become macrochaetae of the head and thorax (figure 1.2), the chordotonal organs of the legs and wings, the large canpaniform sensilla of the wing blade, some of the chemoreceptors of the proboscis, leg and wing margin, and some of the mechanoreceptors of the legs. SOP cells which are born after puparium formation include the microchaete of the head and notum, many sensilla of the proboscis and antenna, and most of the mechanoreceptive sensilla of the wing margin, legs, haltere and genitalia. The sensilla of the abdomen appear last at approximately 32 hours after puparium formation. SOP cell division does not begin until after puparium. Both division and differentiation of the sense organs in the adult is similar to the embryonic pattern, in that the first sensory axons send pioneer tracks towards the CNS along which later axons are guided.

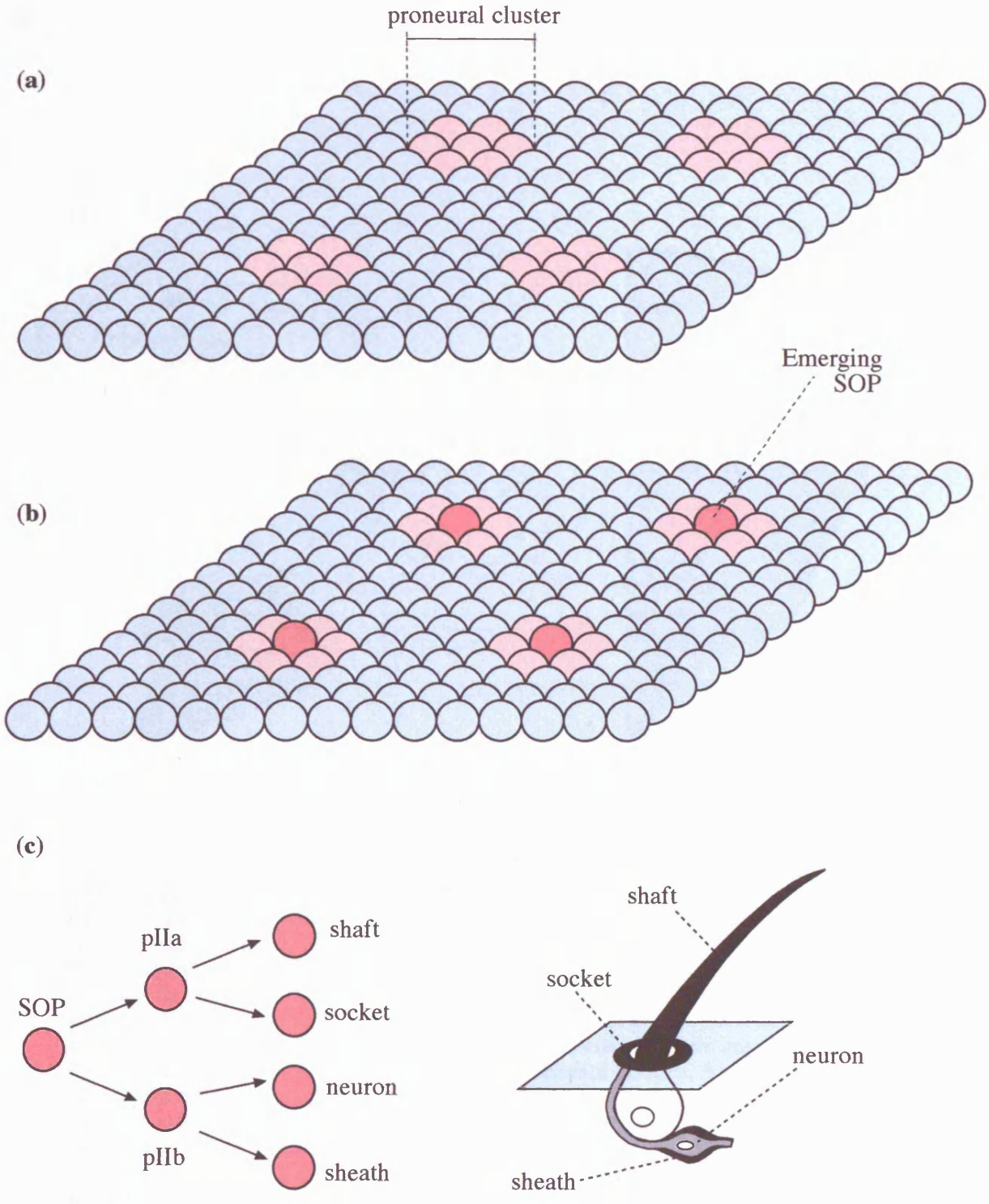


Figure 1.1. Progressive specification of the sensory organ precursor (SOP).
(a) Expression of the proneural genes (pink) delimits the proneural cluster and endows the cells, in which expression occurs, with the competence to become SOPs. (b) Cell-cell interactions mediated by the neurogenic genes allows one cell to be singled out from the proneural cluster to form a SOP (dark pink). (c) The SOP divides twice to give rise to the four different cell which constitute the mature sense organ, in this case a sensory bristle.

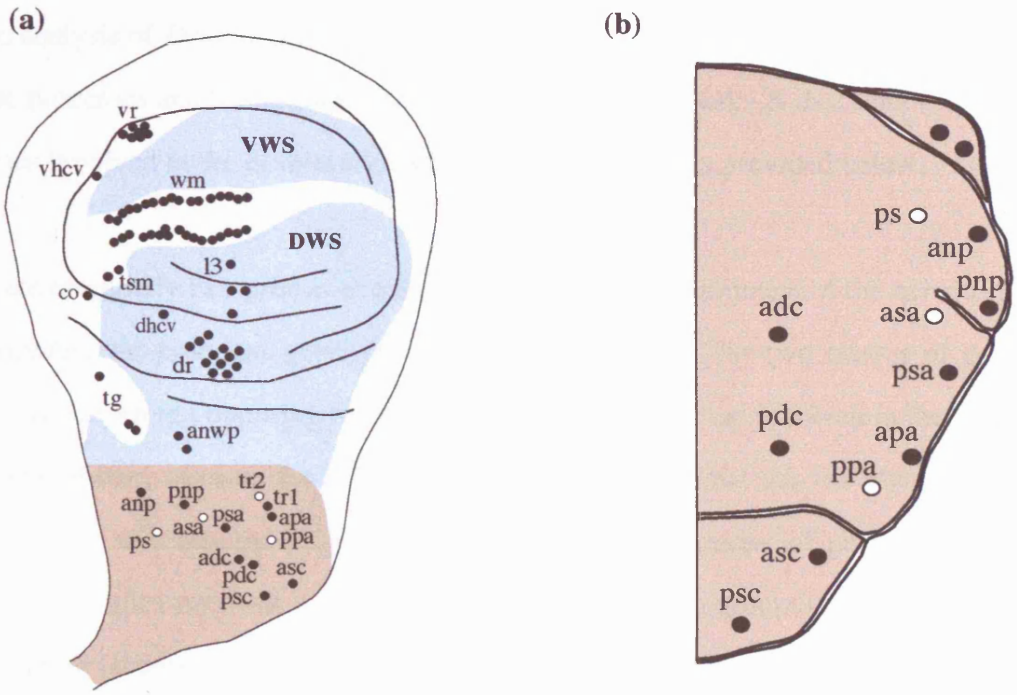


Figure 1.2. Relationship between the position of SOP cells in the wing imaginal disc and the macrocheate bristles on the notum of the adult.
(a) Schematic drawing of a wing imaginal disc from a late third instar larva showing the location of SOP cells corresponding to the adult sensory organs. SOP cells arising after pupation are indicated as open circles. The region of the disc that will give rise to the notum is indicated in brown, and the regions that will give rise to the wing blades are indicated in light blue. (b) Schematic drawing of the adult heminotum showing the position of the macrocheate bristles. Abbreviations: adc anterior dorso-central, anp anterior notopleural, apa anterior post-alar, asa anterior supra-alar, asc anterior scutellar, co costa, dhcv dorsal campaniform sensillum of the humeral cross vein , dr dorsal radius, DWS dorsal wing surface, l3 sensillum campaniform of the third wing vein, pdc posterior dorso-central, pnp posterior notopleural, ppa posterior post-alar, ps presutural, psa posterior supra-alar, psc posterior scutellar, tg tegula, tr trichoideum, tsm twin sensilla of the margin, vhcv ventral campaniform sensillum of the humeral cross vein, vr ventral radius, VWS ventral wing surface, wm wing margin.

The relative simplicity of its organisation, the invariance in position of its elements and the well characterised descriptions of its development have made the *Drosophila* PNS an extremely useful model to study a whole range of developmental mechanisms including cell fate determination, pattern formation, transcriptional activation and repression, equivalence group formation, asymmetric division and lateral inhibition. In addition, the amenable genetic analysis of *Drosophila* has allowed the identification of many of the genes involved in these processes and has allowed their function to be analysed. A description of some of the genes involved in the development of the nervous system is provided below.

There are essentially two groups of genes involved in the development of the nervous system in *Drosophila*; the proneural genes and the neurogenic genes. The two classes of genes can be seen as performing opposing functions; the proneural genes are involved in the promotion of the neural fate, whereas the neurogenic genes are responsible for the repression of the neural fate. It will become evident, however, that the two groups of genes are intimately linked in a complex network of interactions. Each of the two groups of genes is described below in turn (section 1.4 and 1.5) and then the mechanism which connects the two groups of genes, the process of lateral inhibition, will be described in detail in section 1.7.

1.4. The proneural genes.

The first step in neurogenesis is the definition of a region of neural competence from which selection of the neural precursors cells occur. No major differences exist in terms of the genes and mechanisms involved in the process of this selection between the neural progenitor cells of the CNS and PNS.

The term 'proneural' was first suggested by Ghysen and Modolell to describe genes that define a state that makes cells competent to become neuronal precursors (Ghysen and Dambly-Chaudière, 1989; Romani *et al.*, 1989; Simpson and Carteret, 1990). Proneural gene expression defines both the neuroectodermal region of the embryo and the proneural clusters in the developing PNS, providing the cells in which expression occurs, competence

to adopt a neural fate. Proneural function is encoded by the genes of the *achaete-scute* complex (*AS-C*) which consists of three adjacent proneural genes *achaete* (*ac*), *scute* (*sc*) and *lethal of scute* (*l'sc*) and the neural precursor gene *asense* (*ase*). In general, *ac* and *sc* are mainly involved in imaginal disc neural specification (Garcia Bellido and Santamaria, 1978; Garcia Bellido, 1979) whereas *l'sc* is involved in the development of the embryonic CNS. The fourth gene of the complex, *ase*, is most probably not a proneural gene but is involved in differentiation rather than determination of the SOP cells (Brand *et al.*, 1993). In addition, a number of other unlinked loci also encode proteins with proneural function. The first gene, *atonal* (*ato*) is a proneural gene which is involved in the specification of the chordotonal organs and the photoreceptors of the compound eye (Jarman *et al.*, 1993; Jarman *et al.*, 1995). Another gene, *daughterless* (*da*) is often considered as a proneural gene on the basis of its mutant phenotype and its genetic and physical interactions with the proneural genes of the *AS-C* (Dambly-Chaudière *et al.*, 1988; Brand and Campos-Ortega, 1989). However, there is now some debate as to whether *da* is necessary to determine the SOP fate or is just required to maintain SOP survival (Vässin *et al.*, 1994).

Loss of function mutations in the proneural genes lead to an inability to develop neural precursors and ultimately result in neural hypoplasia; deletions of the whole *AS-C* locus eliminate most of the CNS in the embryo and in *AS-C* mutant clones in the adult most of the sensory bristles are absent. In contrast, gain of function mutations such as the *Hairy-wing* mutant allele of the *achaete* locus give rise to supernumerary neural elements at the expense of epidermal tissue (Garcia-Alonso and Garcia-Bellido, 1986; Balcells *et al.*, 1988; Giebel *et al.*, 1997).

All of the proneural genes encode proteins containing the basic-Helix-Loop-Helix (bHLH) domain characteristic of a family of transcriptional regulators which also includes the vertebrate MyoD and myc factors (Murre *et al.*, 1989). The bHLH domain is required for DNA binding and the formation of dimeric complexes with other HLH proteins. The AS-C proteins Ac, Sc and L'sc are able to heterodimerize between themselves and with the Da protein. These dimers specifically bind to DNA *in vitro* (Vaessin *et al.*, 1990; Van Doren *et*

al., 1991) and regulate transcription of reporter genes containing the target binding sites for these proteins (Cabrera and Alonso, 1991). Da is ubiquitously expressed whereas the proteins of the *AS-C* are spatially regulated (see below), therefore formation of AS-C protein-Da heterodimers is dependent on where *AS-C* genes are expressed.

The proneural genes at the *AS-C* exhibit functional redundancy. Phenotypic analyses have uncovered this redundancy with respect to their function in the determination of embryonic neuroblasts as well as in the development of the larval and adult PNS. This functional redundancy is shown in experiments in which ectopic expression of the *AS-C* genes is performed in different *AS-C* mutant backgrounds. After heat-shock driven expression of *sc* in an *ac* mutant background, the proneural activity of *sc* can substitute for the function of *ac* by inducing bristle development of bristles which in the wild-type are *ac*-dependent (Cubas and Modolell, 1992; Rodriguez *et al.*, 1990). Similarly heat-shock induced expression of *ase* can compensate for removal of *sc* activity (Brand *et al.*, 1993; Dominguez and Campuzano, 1993). Remarkably, even ectopic *l'sc* expression can promote development of external sense organs in the adult notum even though the endogenous gene is not usually involved in the determination of SOPs in the adult (Hinz *et al.*, 1994).

The genetic complexity of the *AS-C* became evident in the very early studies by Muller in 1935 and by Garcia-Bellido and Santamaria in 1978. Using a set of *AS-C* chromosomal rearrangements (inversions, large deletions and translocations) it was inferred that a number of *cis*-acting, site-specific enhancers were scattered on both sides of the *ac* and *sc* genes (Ruiz-Gómez and Modolell, 1987; Leyns *et al.*, 1989). It appears that many of these enhancers are shared by the promoters of both genes and that the *ac* and *sc* promoters interact one at a time with the site specific enhancers that mediate activation in a particular proneural cluster to accomplish co-expression of *ac* and *sc* (Gómez-Skarmeta *et al.*, 1995). To further complicate matters the products from the *AS-C* are believed to both cross- and self-regulate one another (Van Doren *et al.*, 1992; Martínez *et al.*, 1993; Culí and Modolell, 1998).

Unsurprisingly, given the complexity of the regulatory regions of the genes of the *AS-C*, the genes comprising the locus are expressed in complex patterns during development. The genes are expressed during embryonic and adult development in partially, and sometimes completely overlapping patterns of cell clusters or rows of cells from which individual neural precursor cells arise. In the embryonic CNS the expression pattern of *ac*, *sc* and *l'sc* partially overlap in a checkerboard pattern of cell clusters that foreshadow the regions where the neuroblasts arise (Cabrera *et al.*, 1987; Skeath and Carroll, 1992). In the embryonic PNS, the dynamic pattern of *ac*, *sc* and *l'sc* expression create a grid of cell clusters from which the SOPs develop. Similarly, in the adult PNS, the *ac* and *sc* genes are expressed in complex patterns in the imaginal discs which prefigure the sites at which SOP cells arise (figure 1.2). This was elegantly demonstrated by Cubas *et al.* (1991) who performed a double staining experiment in which the SOP cells were revealed by staining for β-galactosidase activity in the A101 enhancer trap line and *sc* expression was revealed by *in situ* hybridisation. The A101 line is a *lacZ* enhancer trap inserted in the *neuralized* locus (a neural precursor gene expressed in all SOP cells) which expresses *lacZ* in the majority of SOP cells. In the double labelling experiment single SOP cells were shown to arise from small groups of cells expressing *sc* (Cubas *et al.*, 1991).

The exact identity of the *AS-C* transregulators and how these factors bring about the complex patterns of expression observed *in vivo* are not fully known or understood. It is proposed that these *AS-C* enhancers interact with local combinations of transcription factors and activate expression of the proneural genes at the sites containing the appropriate combination of these proteins. Together, these factors would constitute a prepattern and provide positional information which would be transduced by the enhancers of the *AS-C* genes into gene expression. In the embryo the prepattern genes responsible for regulating expression of the *AS-C* genes are the pair rule and segment polarity genes along the anterior-posterior axis (Skeath and Carroll, 1992) and the dorso-ventral genes along the D-V axis. The E(spl) (see section 1.5.4) and Hairy proteins are responsible for negative regulation of the *AS-C* genes, as derepression of the *AS-C* occurs in embryos mutant for either *E(spl)* or *hairy* (Skeath and Carroll, 1992). Our present knowledge of the regulation of the *AS-C* genes

during imaginal development is still relatively rudimentary. *hairy* is known to be a negative transcriptional regulator of *ac* during the early pupal stages (Ohsako *et al.*, 1994). In contrast, the recently discovered homeodomain encoding genes of the *iroquois* complex, *caupolican* and *araucan*, are direct transcriptional activators of the *ac* and *sc* genes and appear to establish a prepattern that prefigures the future site of macrochaetae in the notum (Gómez-Skarmeta *et al.*, 1996). The *pannier* gene, which is expressed in a complimentary fashion to that of the *caupolican* and *araucan* genes, has also been implicated in the regulation of a complimentary subset of macrochaetae in the adult notum (Ramain *et al.*, 1993; Heitzler *et al.*, 1996a).

Although proneural gene expression delimits the region where neuroblasts and SOP cells are born, proneural activity is also subject to regulation at a postranscriptional level. The involvement of topological factors in positioning the development of the neural precursors is best shown in an experiment where homogeneous *sc* activity is provided in the wing imaginal disc in an ac^{-} sc^{-} background. Expression of a *hsp70-sc* transgene in an ac^{-} sc^{-} background is sufficient to induce the formation of a small number of notum macrochaete, which remarkably form around the sites where wild-type bristles would normally develop (Rodríguez *et al.*, 1990). This indicates that cells at the sites where SOPs are normally born have an increased ability to respond to the neuralizing effects of *sc* and it is therefore these cells which become SOPs. The agent responsible for this differential responsiveness of cells to proneural function is the product of the *extramacrochaetae* (*emc*) locus. *emc* is expressed in the wing disc in a dynamic and complex pattern and SOPs often arise in places where *emc* expression is lowest. Partial loss of function *emc* mutations result in additional sensory structures at ectopic positions. Moreover, extra doses of *AS-C* linearly enhance the *emc* mutant phenotype and, conversely, extra doses of emc^{+} suppress the excess neural structures induced by over expression of *ac* or *sc*. Molecular analysis of *emc* (Ellis *et al.*, 1990; Garrell and Modolell, 1990) revealed the mechanistic basis of these titration effects. *emc* encodes a protein containing a HLH dimerizing domain but lacks the adjacent basic domain necessary for DNA binding. Emc is therefore able to dimerize with other HLH domain proteins but these dimers do not bind DNA. In this way, Emc complexes with the Ac, Sc or Da proteins

in vivo, rendering them unable to bind DNA and thus titrating their activity (Van Doren *et al.*, 1991; Van Doren *et al.*, 1992; Martínez *et al.*, 1993). As Ac and Sc are only effective in generating SOPs when sufficient activity is present to overcome the threshold imposed by the local concentration of Emc, the Emc protein can be seen as being responsible for generating greater accuracy for SOP positioning.

The genes targeted for activation by the proneural proteins remain stubbornly elusive. Two functions can be envisioned for these proneural target genes: (i) a subset of genes are likely to be involved in the maintenance of the SOP cell and (ii) a subset of genes will be involved in the implementation of the neural developmental pathway, of which many are likely to be additional regulatory genes. Another role of the proneural proteins, as we shall see in section 1.7, is to initiate the transmission of inhibitory signals to neighbouring cells which brings about the selection of a single neural precursor from a group of competent cells in a process called lateral inhibition (Haenlin *at al.*, 1994; Kunisch *et al.*, 1994; Hinz *et al.*, 1994).

1.5. The Notch signal transduction pathway.

The list of developmental processes requiring *Notch* is ever-increasing, so much so that it is difficult to name any tissue in the fly that is without dependence on *Notch* function at some stage in its development (for review a see Artavanis-Tsakonas *et al.*, 1995). With few exceptions (Lecourtois and Schweisguth, 1995; Wang *et al.*, 1997; Ligoxygakis *et al.*, 1998) these processes utilise the core *Notch* signalling pathway (which is described in detail below). The *Notch* signalling pathway, despite its apparent simplicity, is very flexible in terms of the developmental mechanisms in which it is involved. The type of processes that require *Notch* fall into at least three different categories: lateral inhibition (for example, neuro-epidermal fate specification, section 1.7), lineage decisions (for example, during lineage decisions in the muscle and neural precursors) and boundary formation (for example, during wing dorso-ventral boundary formation). The role of the *Notch* signalling pathway in lateral inhibition is described in detail in section 1.7, for a review of the other two

mechanisms see Jan and Jan (1993a) and Irvine and Vogt (1997). Furthermore, the components of the core *Notch* pathway have been identified in many other species and appear to act in an analogous manner, illustrating that this pathway is highly conserved (Artavanis-Tsakonas *et al.*, 1995).

Mutations in several loci produce similar neurogenic phenotype and exhibit dosage sensitive interactions (Lehmann *et al.*, 1983; Vassin *et al.*, 1985; de la Concha *et al.*, 1988; Xu *et al.*, 1990). These genes have been grouped together and are collectively known as the neurogenic genes. The products of the neurogenic loci make up the components of a cell-cell communication system which, as will be described in section 1.7, also includes the products of the proneural genes. The phenotype associated with mutations at each of the neurogenic loci are very similar and are manifest by an increase of neural tissue at the expense of epidermal tissue (this phenotype is the converse phenotype associated with mutations in the proneural genes). In animals with mutations at the neurogenic loci, all cells within the neuroectoderm or proneural cluster continue to express the proneural genes at a high level and develop as neural precursors; consequently there are no cells remaining to form epidermis. It should be noted that, in neurogenic mutants, only those cells expressing proneural protein can develop as ectopic neural precursors. Neurogenic mutant cells which are additionally mutant for the genes of the *AS-C* generate epidermis. Thus the neurogenic genes are not required for the differentiation of either cell type but are needed for the correct allocation of neuro-epidermal cell fate. Five neurogenic loci are described below, namely *Delta* (*Dl*), *Notch* (*N*), *Suppressor of Hairless* (*Su(H)*), *Enhancer of split* (*E(spl)*), which together constitute the core *Notch* signalling pathway, and *groucho* (*gro*). Figure 1.3 shows a schematic diagram of the *Notch* signal transduction pathway.

1.5.1. Notch.

Notch was discovered in the early days of mutagenesis screening in 1923 by Otto Mohr but it was Poulson (1937) who first described the neurogenic phenotype associated with lesions at the locus. Molecular analysis has revealed that *Notch* encodes a large transmembrane receptor protein composed of 2703 amino acids (Wharton *et al.*, 1985; Kidd *et al.*, 1986).

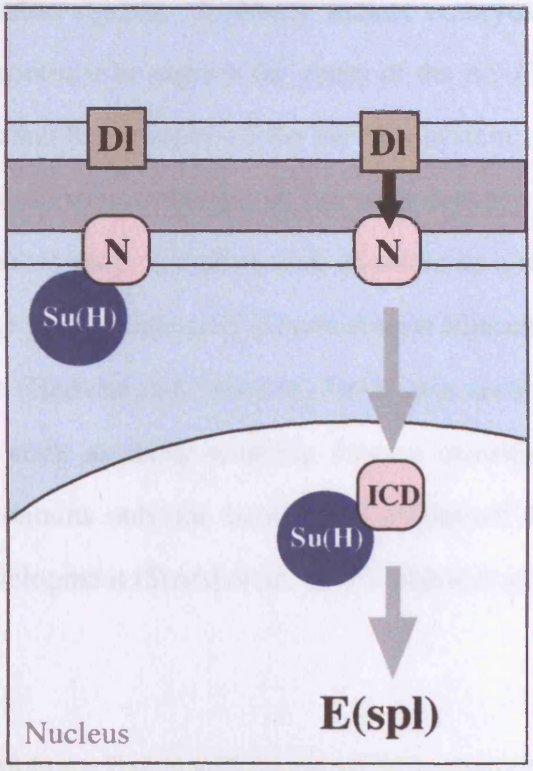


Figure 1.3. The *Notch* signal transduction pathway.
The core *Notch* signal transduction pathway comprises four main components: a transmembrane ligand, Delta (Dl); a transmembrane receptor, Notch (N); a transcription factor, Suppresor of Hairless (Su(H)) and; Enhancer of split (E(spl)). Initially, Su(H) is tethered at the cell membrane by interactions with the intracellular portion of Notch. Activation of the pathway is initiated by the binding of the ligand Delta to the Notch receptor on an adjacent cell. This interaction results in the nuclear localisation of Su(H) and possibly an intracellular portion of Notch (ICD). Nuclear Su(H), possibly in association with the Notch ICD, activates the transcription of the genes of the *E(spl)-C*.

The extracellular domain contains 36 epidermal growth factor (EGF)-like repeats and three copies of another cysteine-rich motif. Only the EGF repeats 11 and 12 have so far been assigned a function in the ligand binding process (Rebay *et al.*, 1991). The intracellular domain contains five copies of an ankyrin-like repeat (Kidd *et al.*, 1986), a RAM domain, a number of *opa* repeats, a 30 amino acid polyglutamine stretch, a PEST sequence and two putative nuclear localisation signals. In *Notch* mutant embryos virtually all cells in the ventral neuroectoderm continue to express the genes of the *AS-C* and adopt the neuroblast fate, which leads to a lethal hypertrophy of the nervous system (Poulson, 1937). Mosaic analysis of *Notch* activity during formation of the adult sensory bristles has demonstrated that *Notch* acts cell autonomously, in accord with its status as a receptor, and that relatively modest differences in the genetic dosage of *Notch* between adjacent cells is sufficient to bias the selection of cell fate (Heitzler and Simpson, 1991) (see section 1.8). Gain of function phenotypes of *Notch*, such as those resulting from a constitutively active form of the receptor N^{intra} (which contains only the intracellular portion of Notch), result in complete repression of neural development (Struhl *et al.*, 1993; Lieber *et al.*, 1993).

1.5.2. Delta.

Delta encodes a transmembrane ligand with an extracellular domain containing nine EGF-like repeats, a single hydrophobic transmembrane domain and an intracellular domain with no significant sequence similarity to the intracellular domain of any other described protein. Based on genetic and molecular interactions, *Delta* has been identified as encoding a ligand for *Notch* (Heitzler and Simpson, 1991; Fehon *et al.*, 1990). In contrast to cells mutant for *Notch*, cells mutant for *Delta* behave non-autonomously and can differentiate epidermis when adjacent to wild-type cells, demonstrating that Delta is not required for the reception of the signal, and is consistent with a role for Delta as the Notch ligand. Loss of function mutations in *Delta* cause the same cell fate transformations as do null mutations in *Notch*, which indicates that the Notch receptor is activated upon binding of Delta. There is now considerable evidence that *Delta* transcription is upregulated by the proneural proteins (Hinz *et al.*, 1994; Haenlin *at al.*, 1994). Furthermore, in reporter gene assays where the upstream control regions of the *Delta* gene have been linked to *lacZ* it has been shown that

transcriptional activation of this reporter construct is dependant upon the proneural proteins (Kunisch *et al.*, 1994).

1.5.3. Suppressor of Hairless.

Suppressor of Hairless (*Su(H)*) encodes an evolutionary conserved transcription factor which is highly related to the mammalian RBP-Jκ, CBF1, and KBF2 transcription factors. *Su(H)* is required for normal neurogenesis and clones of *Su(H)* mutant cells have phenotypes essentially indistinguishable from those lacking *Notch* (Schweisguth and Posakony, 1992; Schweisguth and Posakony, 1994; Schweisguth, 1995). Furthermore, genetic interactions between *Su(H)* and *Notch* alleles are consistent with the two genes functioning in the same pathway (Fortini and Artavanis-Tsakonas, 1994). Although it is clear that Su(H) is a component of the *Notch* signalling pathway and it seems likely that its requirement is involved with transcriptional activation of target genes of the pathway (see below) the molecular events of signal transduction once Notch has been activated by Delta are presently unclear. Two models for the transduction of the signal have been proposed.

The first model proposes that the Su(H) protein is located in the cytoplasm of the cell tethered to an intracellular portion Notch (Fortini and Artavanis-Tsakonas, 1994; Wettstein *et al.*, 1997). Upon activation by Delta, Notch releases Su(H) which then translocates into the nucleus where it activates transcription of target genes (Fortini and Artavanis-Tsakonas, 1994). More recently, an alternative model has been proposed stating that, binding of Delta to Notch results in the proteolytic cleavage to release an intracellular portion of Notch (N^{ICD}). N^{ICD} then enters the nucleus, associates with Su(H), and activates transcription of target genes (Jarriault *et al.*, 1995). A major obstacle in the acceptance of this model has been an inability to detect the N^{ICD} in the nucleus. Recent data however, demonstrates that Notch activity can be found in the nucleus (Struhl and Adachi, 1998; Lecourtois and Schweisguth, 1998), and further support for this hypothesis comes from a number of studies which demonstrate that a number of Notch proteolytic processing events, both ligand-dependant and ligand-independent, occur *in vivo* (Kopan *et al.*, 1996; Schroeter *et al.*, 1998; Blaumueller *et al.*, 1997; Logeat *et al.*, 1998).

Irrespective of the molecular basis of transduction of the Notch signal from the membrane to the nucleus there is strong evidence suggesting that nuclear Su(H) is directly involved in the transcriptional activation of the genes of the *Enhancer of split* complex (*E(spl)-C*) (see section 1.5.4).

1.5.4. Enhancer of split.

The *E(spl)* locus was initially identified as a result of a dominant mutation [$E(spl)^D$] that enhances the eye phenotype of the recessive viable *split* allele of *Notch* (Welshons, 1956). Phenotypic analysis of a number of deletion alleles generated by reverting the $E(spl)^D$ allele subsequently revealed the involvement of *E(spl)* in neural development (Lehmann *et al.*, 1983). Embryos with large deletions at the *E(spl)* locus are inviable and exhibit a strong neurogenic phenotype similar to that produced by mutations in *Notch*. Further genetic analysis of *E(spl)* revealed a complex pattern of heteroallelic complementation and indicated that a region spanning approximately 60 kb was required for *E(spl)* function (Ziemer *et al.*, 1988). This suggested that a number of genes may be responsible for *E(spl)* function. The cloning of the *E(spl)* genomic region provided molecular evidence for genetic complexity, as mutations phenotypically related to *E(spl)* were found to affect different genes (Knust *et al.*, 1987). At least 13 transcription units were found in this region (Knust *et al.*, 1987), seven of which encoded the closely related bHLH proteins mβ, mγ, mδ, m3, m5, m7, and m8 (Klämbt *et al.*, 1989; Knust *et al.*, 1992; Delidakis *et al.*, 1992). One of these, E(spl)-m8, was altered in the $E(spl)^D$ allele (Tietze *et al.*, 1992) (figure 1.4).

The products encoded by the seven bHLH *E(spl)* genes belong to a group of structurally and functionally related proteins known as the Hairy-related protein family which also includes the proteins encoded by the *hairy* and *deadpan* loci. These proteins share a number of highly conserved domains: the bHLH domain; a putative amphipathic helical domain (the Orange domain), and; the WRPW motif (for review see Fisher and Caudy, 1998) (figure 1.5). The HLH region mediates homo- and hetero-dimeric protein interactions and the basic region is important for DNA-binding. The putative amphipathic helical domain (Knust *et al.*, 1992) which Dawson *et al.* (1995) have referred to as the Orange domain, may contribute to

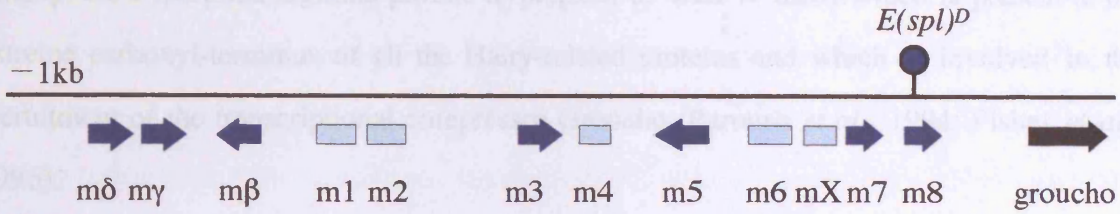


Figure 1.4. Genomic organisation of the *E(spl)* locus.
Schema of the genomic region of the *E(spl)-C*. Proximal is to the left and distal to the right. The 13 transcription units of the locus are shown. The seven bHLH *E(spl)* genes *E(spl)-mδ*, *-mγ*, *-mβ*, *-m3*, *-m5*, *-m7* and *-m8* are indicated in blue, the orientation of transcription is indicated by the direction of the arrow. The non-bHLH genes are indicated in light blue. Immediately distal to the *E(spl)-m8* gene is the *groucho* locus (black arrow). The *E(spl)D* allele, from which the locus got its name, is due to a frameshift mutation in the *E(spl)-m8* (lollipop).

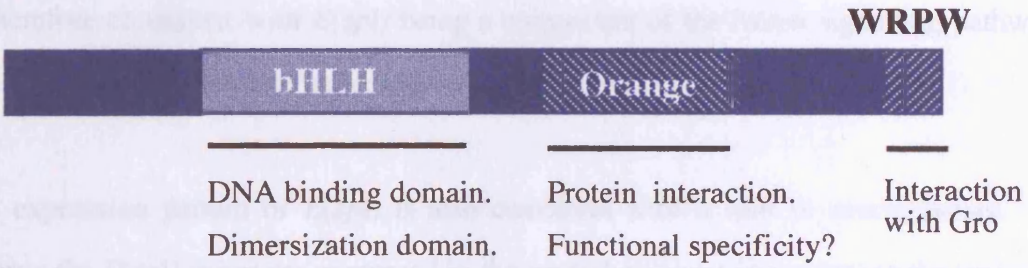


Figure 1.5. Shematic diagram of the domain structure of the Hairy-related proteins
The Hairy-related protein family includes the *Drosophila* proteins Hairy, E(spl) and Deadpan as well as the related vertebrate proteins Hes, Her and ESR. These proteins share three conserved domains: the bHLH domain, the Orange domain, and the WRPW motif. Given below are the functions which have been assigned to the particular domains (Murre *et al.*, 1989; Paroush *et al.*, 1994; Fisher *et al.*, 1996; Dawson *et al.*, 1995).

functional diversity between the proteins of the family, based on differences in the behaviour between the Orange domain of the Hairy and the E(spl) proteins (Dawson *et al.*, 1995). Finally, the tryptophan-arginine-proline-tryptophan or WRPW motif which is present at the extreme carboxyl-terminus of all the Hairy-related proteins and which is involved in the recruitment of the transcriptional corepressor Groucho (Paroush *et al.*, 1994; Fisher *et al.*, 1996).

Embryos mutant for the *E(spl)-C* exhibit a similar phenotype to embryos mutant for *Notch*. In *E(spl)-C* mutants proneural gene expression persists in all cells of the ventral neuroectoderm and results in all cells segregating as neuroblasts and producing embryos with an extremely hypertrophied nervous system. In clones of *E(spl)-C* mutant cells in the adult, the majority of the cells of the proneural cluster adopt the SOP fate and develop as bristles (de Celis *et al.*, 1991; Tata and Hartley, 1995). This LOF phenotype is cell autonomous. In contrast, when *E(spl)* is artificially expressed in the presumptive neural precursors the opposite phenotype is produced, a loss of neural tissue, which is similar to the phenotype produced by activated forms of *Notch* (N^{intra}). When ectopic expression of *E(spl)-m5* or *E(spl)-m8* is induced in the imaginal disc during SOP cell allocation a loss of sensory bristles occurs (Tata and Hartley, 1995). The data from the LOF and GOF analyses is therefore consistent with *E(spl)* being a component of the *Notch* signalling pathway and that the role of the E(spl) proteins is to repress the neural fate.

The expression pattern of *E(spl)* is also consistent with a role in neurogenesis. In the embryo the *E(spl)* genes are expressed in the ventral neuroectoderm during the period when neurogenesis is occurring and appears to correlate with the places where neuroblasts are delaminating from the ectoderm (Knust *et al.*, 1987; Knust *et al.*, 1992). Moreover, mRNA only accumulates in cells which remain in the ectoderm and not in the delaminating neuroblast itself suggesting that *E(spl)* is only expressed in the cells which are inhibited from adopting the neural fate. As explained earlier (section 1.2.1), neurogenesis in the embryonic CNS occurs in several waves and expression of the *E(spl)* genes during this time is dynamic. After the first wave of neuroblast delamination expression can be seen in most

cells remaining in the ectoderm, but expression is rapidly modulated as a new wave of neurogenesis begins (Jennings *et al.*, 1994). Unlike the expression pattern of the individual *E(spl)* genes in the embryonic neuroectoderm, which appear very similar if not identical, the expression patterns of the individual *E(spl)* genes during imaginal development are distinct (de Celis *et al.*, 1996). For example, in the wing imaginal disc, *E(spl)-m8* and *E(spl)-m7* mRNA is detected in clusters of cells that correspond to the locations where SOPs develop, whereas *E(spl)-mβ* is expressed most prominently in the wing blade in several proximodistal stripes crossing the dorsoventral boundary and in a complex pattern elsewhere in the disc with no simple association with developing sensory organs (de Celis *et al.*, 1996). In contrast, some *E(spl)* genes are not expressed in the wing disc at all, such as *E(spl)-m5*, or are expressed at very low levels, such as *E(spl)-m3*. Distinct expression patterns between the different *E(spl)* genes have also been observed in other imaginal discs such as the eye and leg discs (de Celis *et al.*, 1996).

The pattern and timing of *E(spl)* expression suggest that the genes are activated in response to *Notch* signalling activity. The dependence of E(spl) protein accumulation on *Notch* signalling can be seen from the effects of *Notch* and *Delta* mutations (Jennings *et al.*, 1994). Embryos mutant for either *Notch* or *Delta* express no detectable E(spl) protein. Conversely, the presence of an activated form of *Notch* (N^{intra}) results in ectopic expression of the *E(spl)* genes in both the embryos and imaginal discs (Jennings *et al.*, 1994; Jennings *et al.*, 1995). Epistatic studies using N^{intra} have placed *E(spl)* function downsteam of *Notch*, consistent with the LOF analyses which show that *E(spl)* function is required cell autonomously (Lieber *et al.*, 1993; de Celis *et al.*, 1991; Tata and Hartley, 1995). Collectively, these data suggest that *E(spl)* is the final nuclear response to *Notch* signalling.

As outlined above genes of the *E(spl)-C* are under direct transcriptional control of Su(H)/Notch^{ICD}. Evidence for this come from two observations. Firstly, ectopic accumulation of the *E(spl)* gene products induced by the use of a constitutively activated form of *Notch* requires *Su(H)* activity (Bailey and Posakony, 1995). Secondly, *Su(H)* binding sites have been located in the proximal upstream region of a number of the *E(spl)*

genes (*E(spl)-m4, -m5* and *-m8*) and the integrity of these sites is essential for the activity of each individual promoter during early neurogenesis in the embryo and during SOP selection in the imaginal disc proneural clusters (Lecourtois and Schweisguth, 1995; Bailey and Posakony, 1995).

During neurogensis, it appears that the primary role of the *Notch* pathway is to antagonise the neural fate. As *E(spl)* are the nuclear effector of the *Notch* pathway the molecular basis of this antagonism can be determined by examining the functional role of *E(spl)*. The structure of the E(spl) proteins suggest that they are transcription factors since all share a HLH dimerization domain and an adjacent basic domain found in a family of DNA binding proteins (Murre *et al.*, 1989). All E(spl) proteins have now been shown to bind DNA *in vitro* (Oellers *et al.*, 1994; Jennings *et al.*, 1999) and be capable of repressing the transcription of reporter genes in cultured cells (Oellers *et al.*, 1994; Heitzler *et al.*, 1996). The ability to function as transcriptional repressors involves interaction with the co-repressor protein Gro (Paroush *et al.*, 1994; Fisher *et al.*, 1996) (see below). The exact targets of E(spl) are currently unknown. There is considerable evidence that the genes of the *AS-C* are one set of candidate genes that are likely to be negatively regulated by E(spl). As previously discussed, derepression of the *AS-C* genes occurs in embryos mutant for *E(spl)* (Skeath and Carroll, 1992; Martin-Bermudo *et al.*, 1995). In line with this, binding sites for E(spl) and Hairy proteins have been identified upstream of the *achaete* gene (Van Doren *et al.*, 1994; Oellers *et al.*, 1994) and a number of E(spl) proteins have been shown to repress basal transcription from reporter genes containing the control regions of the *achaete* gene (Van Doren *et al.*, 1994; Heitzler *et al.*, 1996). It is possible that E(spl) is involved in directly regulating the transcription of other, as yet unidentified, targets genes.

Until this point the seven bHLH genes of the *E(spl)-C* have been considered as though they all behave identically. Evidence for redundancy in function comes from an inability to generate LOF mutations in any of the individual genes at the locus. Aside from the original mutant allele, $E(spl)^{D}$, isolated in 1956 by Welshons, mutations in single bHLH *E(spl)* genes have not been isolated. Further evidence comes from the observation that a number of genes

from the locus can be deleted without phenotypic consequence (Delidakis *et al.*, 1991; Schrons *et al.*, 1992). This functional redundancy is somewhat paradoxical because the genes and organisation of the *E(spl)* locus has been evolutionary conserved in the distantly related species *Drosophila hydei.* In line with this there has been a more recent shift from the idea that the E(spl) proteins exhibit total functional redundancy and it has been suggested that the individual proteins possess subtly different activities. Thus, differences in imaginal expression (de Celis *et al.*, 1996), preferences in protein interaction (Gigliani *et al.*, 1996; Alifragis *et al.*, 1996), preferences in DNA binding capabilities (Jennings *et al.*, 1999) and differential activities in various development contexts (Ligoxygakis *et al.*, 1999) between the individual E(spl) proteins have all been described, supporting the notion that the genes from the locus may not be completely functionally redundant.

Genes closely related to the *Drosophila E(spl)* genes have now been identified in other species and are known variously as the *Hes*, *Her* or *ESR* genes (Sasai *et al.*, 1992; Takebayashi *et al.*, 1995). These genes, like *E(spl)*, are often expressed in response to Notch activity (de la Pompa *et al.*, 1997) and are involved in the implementation of cell fate decisions such as the selection of cells to become neural precursors (Ishibashi *et al.*, 1994; Ishibashi *et al.*, 1995). It therefore appears that the function of *E(spl)*, and indeed the *Notch* signalling pathway, has been evolutionary conserved.

That completes the description of the components of the *Notch* signal transduction pathway. The E(spl) proteins are the nuclear effectors of this pathway and may function at several levels in the repression of the neural fate (section 1.10), including direct transcriptional repression of target genes. The E(spl) proteins are, however, unable to perform this function alone and require the activity of the transcriptional co-repressor Groucho protein.

1.5.5. groucho.

The *gro* gene was initially identified on the basis of its physical proximity to and genetic interactions with the genes of the *E(spl)-C* (Hartley *et al.*, 1988; Delidakis *et al.*, 1991). The *groucho* (*gro*) locus corresponds to the transcription unit immediately distal to *E(spl)-m8* in

the *E(spl)-C*. *gro* originally takes its name from the phenotype of an adult viable allele which resulted in thick tufts of bristles over the eyes resembling the bushy-eyebrowed comedian Groucho Marx (Lindsley and Grell, 1968). Subsequent phenotypic analysis has shown however that mutations in the *gro* locus produce a neurogenic phenotype similar to those associated with *E(spl)-C* deficiencies and those associated with mutations in other components of the *Notch* signalling pathway (Priess *et al.*, 1988; Delidakis *et al.*, 1991; Schrons *et al.*, 1992). The zygotic mutant *gro* phenotype is more penetrant when the maternal contribution is additionally compromised (Schrons *et al.*, 1992), suggesting that at least some *gro* activity is maternally provided.

gro encodes a nuclear protein, structurally unrelated to the bHLH genes of the *E(spl)-C,* and which was initially shown to exhibit weak similarity to the β-subunit of G-protein transducin (Hartley *et al.*, 1988). The *gro* gene has now been isolated from several species and comparisons between the amino acid sequence has revealed two conserved domains within the proteins; a highly conserved amino terminus and a carboxyl-terminal WD40 repeat region (for discussion of WD40 repeats, see Neer *et al.*, 1994) separated by a variable region (figure 1.6). The WD40 domain is thought to mediate protein-protein interaction and a region in the amino terminus of *Drosophila* Gro is required for homotetramerization which, in turn, has been shown to be necessary for repression function (Chen *et al.*, 1998). The *gro* transcript is initially expressed ubiquitously at high levels throughout the syncitial and cellular blastoderm embryo but later becomes progressively more restricted to the developing CNS and, by the time the germband has retracted (stage 13), the transcript is found exclusively in the CNS (Hartley *et al.*, 1988; Delidakis *et al.*, 1991). Early embryonic expression is believed to be of maternal contribution (Hartley *et al.*, 1988), This is in agreement with the requirement for maternal *gro* activity for normal development (Schrons *et al.*, 1992).

The exact role of the *gro* product remained elusive until Paroush *et al.* (1994) performed a yeast two-hybrid screen using the Hairy protein as a bait with which to fish for interacting proteins (see introduction to chapter 3 and 4). From this screen the Gro protein was

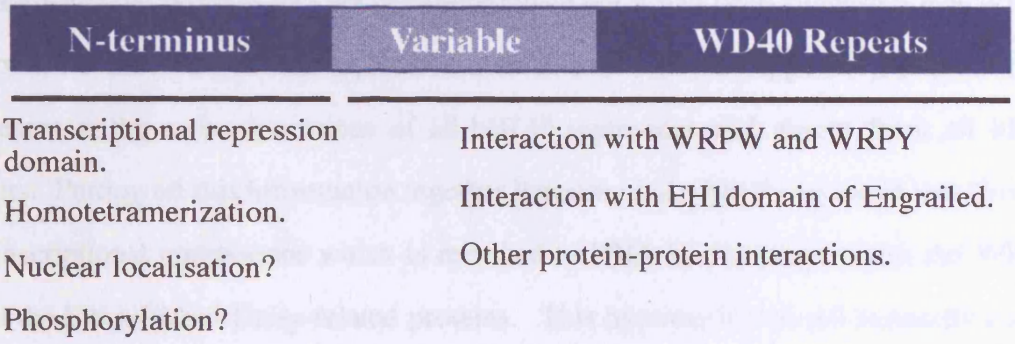


Figure 1.6. Schematic diagram of the domain structure of the Gro protein. Gro proteins consist of three domains, the amino terminus, a series of seven carboxyl-terminal WD40 repeats, and a variable region that separates the two and shows poor sequence similarity among family members. Given below are the functions that have been assigned to the particular regions of the protein (Paroush *et al.*, 1994; Fisher *et al.*, 1996; Jimenez *et al.*, 1997; Chen *et al.*, 1998).

identified as an interacting partner of Hairy. In addition to Hairy, physical interaction with the Deadpan and the E(spl) proteins (both members of the Hairy-related protein family) were also demonstrated and it was determined that the carboxl-terminal WRPW motif of these proteins was critical for mediating interaction with Gro (see chapters 3 and 4). Paroush *et al.* (1994) were also able to show that (i) *gro* function is required during sex determination, segmentation and neurogenesis, three developmental processes regulated by Deadpan, Hairy and E(spl) respectively and (ii) in the absence of *gro* the known target genes of the Hairy-related proteins become ectopically expressed. At this stage the function of the WRPW motif of the Hairy-related proteins had not been determined but it had been suggested that WRPW was a putative candidate for the repression domain based on the supposition that the motif was present at the carboxl-terminus of all bHLH repressors and absent from all bHLH activators. Putting all this information together Paroush *et al.* (1994) suggested that Gro acts as a transcriptional corepressor which is recruited to DNA by interaction with the WRPW motif of the DNA-bound Hairy-related proteins. This hypothesis was subsequently ratified in a study by Fisher *et al.* (1996) who demonstrated that the repressor activity of the Hairy-related proteins is provided by Gro.

This model of repression invoking a DNA binding protein which recruits a non-DNA binding corepressor is well documented. In yeast, repression of the a1-specific gene transcription by the α2 mating type protein depends on the *TUP1* and *SSN6* genes. The Matα2p-Mcm1p complex binds to specific operator sites and targets specific genes for repression by recruiting the corepressor Ssn6p-Tup1p complex to the promoter. Tup1 provides the repressor activity of the corepresser complex (Keleher *et al.*, 1992). Parallels between Tup1 and Gro are often made, not only because they are functionally related, but because they also share structural similarities in that both proteins contain the WD40 repeat motif. Structural differences in other regions however suggest that the two proteins may not be true homologues and therefore it may be more accurate to consider the two proteins as functionally analogous. Models of transcriptional repression involving corepressors have also been demonstrated in higher vertebrates. One example in the mouse are the mSin3A and mSin3B proteins which act as corepressors for MAD and MXI1. Both the MAD and MXI1

proteins heterodimerize with MAX, bind DNA and subsequently recruit mSin3A and mSin3B resulting in transcriptional repression of target genes (Ayer *et al.*, 1995).

It is now becoming clear that the role of Gro as a transcriptional corepressor is not confined to the bHLH proteins of the Hairy-related protein family. A wider range of transcriptional regulators which are active in the early embryo such as Engrailed, Dorsal and Runt also rely on Gro for at least some of their transcriptional repression activity (Jiménez *et al.*, 1997; Dubnicoff *et al.*, 1997; Aronson *et al.*, 1997; Häder *et al.*, 2000; for review see Mannervik *et al.*, 1999).

The Gro proteins from *Drosophila, C. elegans,* rats, mice and humans all share a similar primary sequence structure (Hartley *et al.*, 1988; Stifani *et al.*, 1992; Miyasaka *et al.*, 1993; Schmidt and Sladek, 1993; Choudhury *et al.*, 1997; Pflugrad *et al.*, 1997; Sharief *et al.*, 1997). In line with this, conservation of *gro* function from different species has been observed in a number of experiments. Firstly, in the repression analyses performed by Fisher *et al.* (1996) the human TLE1 protein was able to actively repress transcription when fused to a heterologous DNA-binding domain. Secondly, heterospecific transgenic experiments have shown that a transgene containing the human *TLE1* gene is able to rescue the phenotype of mutations at the *unc-37* locus, the *C. elegans gro* homologue (Pflugrad *et al.*, 1997).

1.6. Additional components of the Notch signalling pathway.

Notch signalling is more complex than the reader has been led to believe. Other components involved in *Notch* signalling pathway have now been identified. These include other ligands such as Serrate, proteins which modify the response of Notch to ligands in certain developmental contexts such as the Fringe protein (Panin *et al.*, 1997), and a host of factors which bind the intracellular portion of Notch and modify the receptors activity, such as Deltex (Diederich *et al.*, 1994; Matsuno *et al.*, 1995), Numb (Guo *et al.*, 1996), Dishevelled (Axelrod *et al.*, 1996) and Notchless (Royet *et al.*, 1998). Deltex appears to positively

enhance Notch activity whereas the other factors, Numb, Dishevelled and Notchless appear to antagonise Notch function. These modifiers of Notch activity are, at least in part, responsible for the differential output of the *Notch* signalling pathway in different developmental contexts (for a review, see Bray, 1998). It has also become clear over the last few years that proteolytic processing of Notch is critically required to generate the mature Notch receptor (Logeat *et al.*, 1998) and for the signalling process itself (Schroeter *et al.*, 1998). In line with this, the product of the *kuzbanian* (*kuz*) gene, a metalloprotease, (Rooke *et al.*, 1996) has also been implicated in *Notch* signalling and may be involved in the protolytic cleavage of Notch to produce the mature receptor. More recently, it has been reported that Kuz may also be responsible for the processing of Delta, to produce an extracellular fragment capable of binding Notch and acting as an agonist of Notch activity (Qi *et al.*, 1999).

1.7. Lateral inhibition.

The initial step in neural development is the demarcation of cells which possess neural competence. The adoption of neural competence is bestowed on a group of cells by the expression of the proneural genes. Every cell within this subset of cells has the potential to become a neural progenitor cell and will do so unless otherwise instructed. In the wild-type situation most cells within the group are instructed otherwise such that only one or a small number of cells within the competent population will retain the neural fate and go on to produce the neural progenitor cell. The remaining cells are prevented from adopting this fate and are held in a state of non-commitment, as such they are later free to adopt an alternative fate, such as the epidermal fate or adopt the neural fate in a subsequent wave of neurogenesis. It is the *Notch* signalling pathway that is critically required for this instructive function to restrict neighbouring cells from all committing to the neural fate; mutations in the genes involved in this pathway result in all competent cells adopting the neural fate. What then is the mechanism which brings about assortment of cells into two populations and what is the role of the Notch signalling pathway in this mechanism?

Intercellular signalling can occur in two distinct ways; inductive signalling and lateral signalling (figure 1.7). Inductive signalling occurs between two non-equivalent cell types, such that the signalling cell and receiving cell begin the process with different properties, including their repertoires of cell surface receptors and ligands. One cell produces the signal and the other cell will receive it, and as a consequence of signal reception the cell will adopt an alternative fate to the signalling cell. Lateral inhibition on the other hand, involves a number of equivalent cells, each of which have the same properties at the beginning of the process, which signal back and forth to one another. A feedback mechanism (described below) ensures that one cell becomes the 'sender' and the remaining cells become the 'receivers'. The process results in the 'sender' adopting one fate and the 'receivers' adopting an alternative fate. It transpires that the *Notch* signalling pathway during neurogenesis in *Drosophila* is involved in lateral signalling.

The term 'lateral inhibition' was coined by Wigglesworth in 1940 to denote a process responsible for the generation of an even density of bristles in the larval stages of the insect *Rhodnius prolixus* (Wigglesworth, 1940). He postulated that that an existing bristle exerts an inhibitory influence on surrounding cells, such that a new bristle cell can only develop at a distance outside the inhibitory influence of the existing bristles. The model postulated by Wigglesworth invokes an inhibitory substance produced by a bristle cell that diffuses and inhibits surrounding epidermal cell from developing as bristles. At a certain distance from the bristle the concentration of the inhibitory substance will drop below the threshold required for inhibition and a epidermal cell will develop into a bristle and begin to produce the substance itself, thereby inhibiting nearby cells from adopting an equivalent fate. The first experimental evidence that supported the lateral inhibition model came from the studies of Wilcox *et al.* (1973) and Wolk and Quine (1975) describing the one dimensional spacing of heterocyst cells among intervening vegetative cells in the filamentous blue-green alga *Anabaena.* This alga consists of a linear filament of vegetative cells dispersed with heterocyst cells. The vegetative cells divide and the filament lengthens and as it does so occasional cells convert to form new heterocyst cells. In the lateral inhibition model an inhibitory substance is produced by the heterocyst cell which diffuses along the filament and

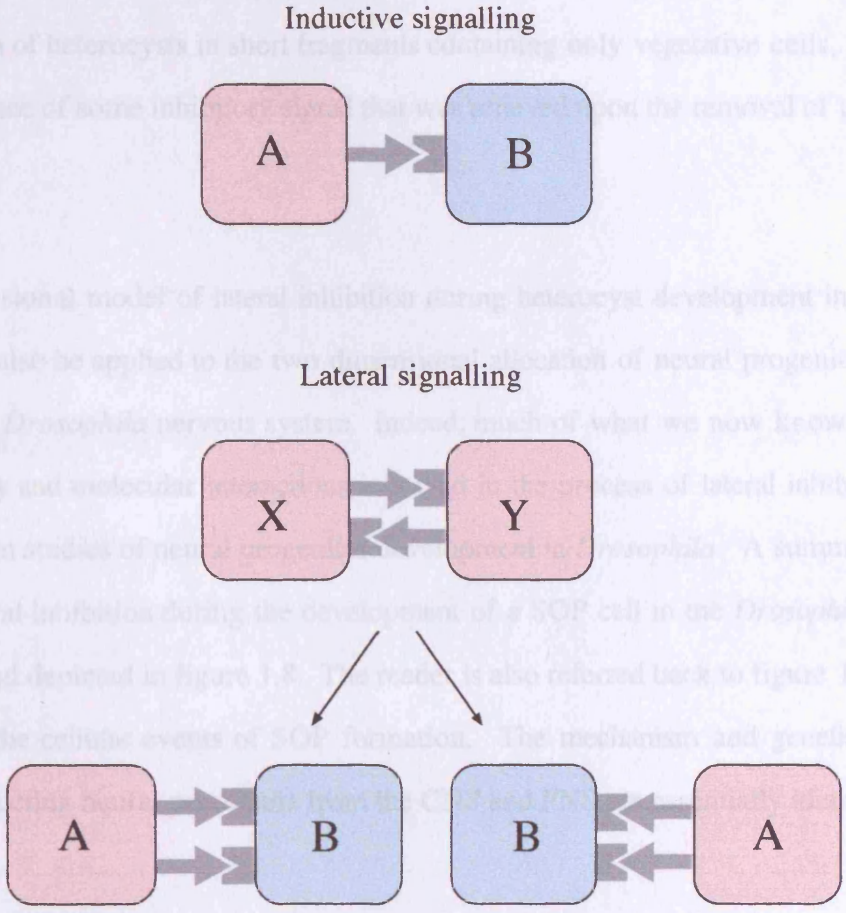


Figure 1.7. Inductive and lateral signalling.
Inductive signalling involves two different cell types, one cell (A) produces the signal which induces the other cell (B), that contains the receptor for the signal, to adopt the B cell fate. Lateral signalling involves signalling between the equivalent cell (X and Y) that each produce both signal and receptor. Only one of the cells however, adopts the A cell fate, the other cell adopts the B cell fate. The outcome is random, half of the time X adopts fate A, half the time Y adopts fate A.

inhibits the other cells from developing likewise. As the filament lengthens some cells, by virtue of their distance from the emanating inhibitory signal, are released from inhibition and begin to develop towards the heterocyst state, start to produce the inhibitory signal themselves and thereby set up a new inhibitory field in which surrounding cells are restricted from developing. Wolk and Quine (1975) showed that disrupting these filaments permitted rapid formation of heterocysts in short fragments containing only vegetative cells, consistent with the existence of some inhibitory signal that was relieved upon the removal of the mature heterocyst.

The one dimensional model of lateral inhibition during heterocyst development in the blue-green alga can also be applied to the two dimensional allocation of neural progenitor cells in the developing *Drosophila* nervous system. Indeed, much of what we now know about the genetic circuitry and molecular interactions involved in the process of lateral inhibition have been learnt from studies of neural progenitor development in *Drosophila*. A summary of the process of lateral inhibition during the development of a SOP cell in the *Drosophila* PNS is given below and depicted in figure 1.8. The reader is also referred back to figure 1.1a and b which shows the cellular events of SOP formation. The mechanism and genetic circuitry involved in selecting neural precursors from the CNS and PNS are essentially identical.

Positional information in, for example, the wing imaginal disc initiates transcription of the *AS-C* genes in a subset of cells called the proneural cluster (in figure 1.1a cells expressing the proneural genes are shown in pink). The AS-C proteins provide neural competency; in mutant clones without *AS-C* activity no sense organs will develop. The expression of the AS-C proteins initially provides every cell within the cluster with an equal potential to develop as the SOP cell. In addition to providing neural competency however, the proteins of the AS-C are also involved in the generation of a signal which inhibits neural competency in neighbouring cells. The inhibitory signal is initiated by the proneural proteins by direct transcriptional activation of the *Delta* gene. It will be recalled that *Delta* encodes the transmembrane ligand for the Notch receptor. The level of Delta on the membrane of any one cell is therefore proportional to the level of proneural protein within that cell. The role of

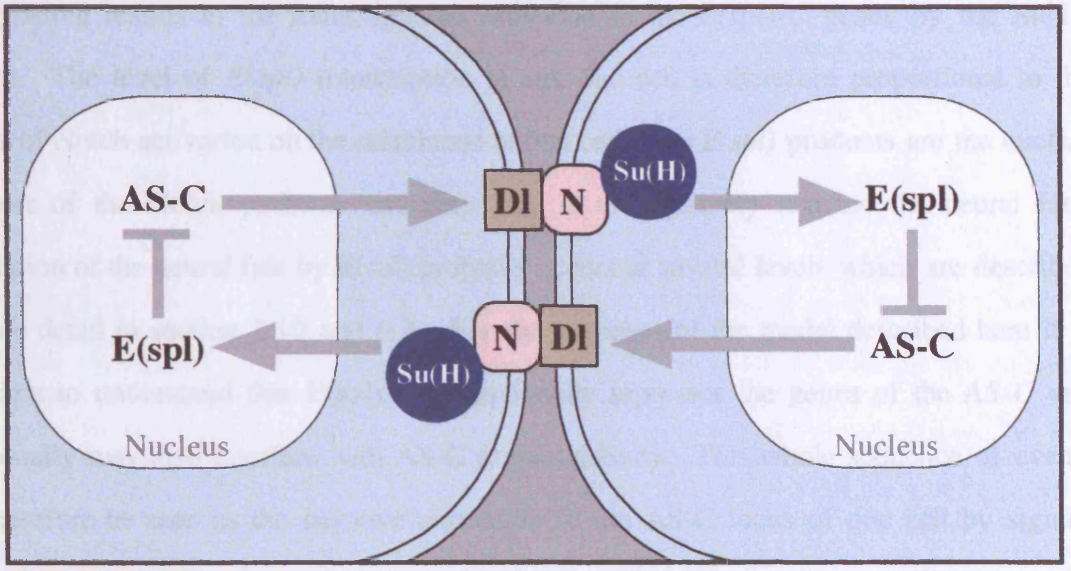


Figure 1.8. Lateral inhibition.
All cells initially express the *AS-C* genes, *Notch* and *Delta*. Binding of the Delta ligand to the transmembrane Notch receptor induces translocation of the Su(H) protein to the nucleus, where it acts as a transcription factor to activate the genes of the *E(spl)-C*. The proteins of the *E(spl)-C* will, in turn, inhibit the synthesis and/or activity of the products encoded by the *AS-C* locus. The transcription of *Delta* is controlled by AS-C activity, thus closing the feedback loop between *Notch* and *Delta*.

Delta is to active the receptor Notch on neighbouring cells. As both the ligand and receptor are membrane bound activation of Notch by Delta requires that communicating cells be in direct contact with one another. Activation of Notch by Delta on a neighbouring cell initiates transduction of a signal to the nucleus of the receiving cell. The nature of this signal is not fully determined but is likely to involve a cleaved portion of the intracellular domain of Notch itself and/or the Su(H) protein. Irrespective of how the signal is transduced the consequence of signalling results in the transcriptional activation of the *E(spl)-C* genes by the Su(H) protein. The level of *E(spl)* transcription in any one cell is therefore proportional to the degree of Notch activation on the membrane of that cell. The *E(spl)* products are the nuclear effectors of the Notch pathway and their role is to negatively regulate the neural fate. Regulation of the neural fate by E(spl) probably occurs at several levels which are described in more detail in section 1.10 and 6.3. For the purposes of the model described here it is sufficient to understand that E(spl) transcriptionally represses the genes of the *AS-C* and additionally may also interfere with AS-C protein activity. This whole sequence of events can therefore be seen as the negative regulation of the AS-C locus of one cell by signals issued from a neighbouring cell (figure 1.8). Importantly, this negative signalling alters the capacity of the recipient cell to send out its own negative regulatory signals to neighbouring cells because a consequence of a reduction in the level of AS-C proteins is a reduction in the levels of *Delta* transcription. The process is therefore self reinforcing: the more inhibited a cell becomes the less it is able to inhibit its neighbours, and vice versa, the more inhibitory a cell becomes the less inhibited it will be by its neighbours. The final outcome of this process is that one cell will retain *AS-C* gene expression and adopt one fate (the dark pink cell in figure 1.1b) and the others cells will loose *AS-C* gene expression and adopt an alternative cell fate.

The neuro-epidermal fate choice in the proneural cluster can be seen as being controlled at a multicellular level, the decision about which cell adopts the neural fate is controlled by the relative level of AS-C activity in all cells within the cluster. Integration is achieved by building cellular connectivity into the control of the key element, the level of *AS-C* gene expression.

1.8. Lateral inhibition during neural fate specification, the evidence.

There are a number of lines of evidence supporting a model of lateral inhibition for determination of cell fate during neural development. The first indication that a small group of cells (what we would now call the proneural cluster), rather than a single cell, possess neural competency was demonstrated in the classical genetic experiments of Stern (1954). Stern (1954) generated gynandromorphic flies that were genetically mosaic for mutations in the *achaete* gene. The *ac* gene is required cell autonomously for the development of a subset of the 13 invariantly positioned macrochaetae in the fly heminotum (see section 1.4). In line with this, Stern demonstrated in his analysis that whenever the ac^+ territory included the site of a bristle, the bristle would always develop and whenever the ac^- territory included the site of a bristle the bristle did not develop. Importantly, Stern's analysis showed that if cells at the precise location of the normal bristle were ac^- but nearby cells were ac^+, a bristle could develop in the ac^+ territory, slightly displaced from its normal position. This analysis was the first evidence that several cells in the region where a bristle develops are capable of producing a bristle. Furthermore, the fact that only one cell produces the bristle, in the wild-type situation, supports a model whereby the developing bristle cell inhibits its neighbouring cells from developing likewise. A similar conclusion was drawn from laser ablation experiments carried out in the developing CNS of the locust embryo (Doe and Goodman, 1985). It was shown that laser ablation of the emerging neuroblast resulted in a nearby presumptive epidermoblast switching fates to adopt the role of the eliminated neuroblast, demonstrating that a cell that chooses a neural fate is one of several cells that could do so. Evidence also comes from phenotypic analysis of mutations of the neurogenic loci. Mitotic clones mutant for *Notch* or *Delta*, for example, which overlap the territory of a proneural cluster produce a bristle cluster containing several bristles, again demonstrating that neural competence is bestowed on a number of cells and suggesting that competition between these cells in the wild-type results in only one adopting the neural fate.

In an elegant series of experiments performed by Heizler and Simpson the role of *Notch* during lateral inhibition was investigated with the express aim of determining whether the

amount of *Notch* product influenced the selection of the SOP from the proneural cluster (Heizler and Simpson, 1991; Heizler and Simpson, 1993). The effects of *Notch* dosage on bristle selection was determined by creating mosaic flies with adjacent territories carrying either one or two, or else two or three, copies of the wild-type $Notch^{+}$ allele (Heizler and Simpson, 1991). The genotype of the bristles on the borders of these clones was then examined (border bristles have a higher probability of arising from proneural clusters which are genetically mosaic for *Notch* dosage) (figure 1.9a). They found that the border bristles were always of the genotype corresponding to the lowest dosage of *Notch*. Therefore a cell with two doses of $Notch^{+}$ (the wild-type dose) will always adopt the epidermal fate when adjacent to cell expressing lower levels of *Notch*, but will adopt the neural fate when adjacent to cells with higher levels of *Notch*. This experiment shows that the behaviour of wild-type cells depends on the other cells in the cluster, in other words the fate of an individual cell is the result of a comparison within the group of which it is part. Further evidence that Notch and Delta mediate the competitive interactions between cells comes from a similar experiment in which $Notch^{-}$ clones were generated and the genotype of bristles on the borders of these clones was then examined (figure 1.9b). Similar to the results of the experiment described above, the border bristles of clones mutant for *Notch*, are always of the mutant genotype. This means that the fate of cells that were wild-type for *Notch* in the original cluster is no longer random, wild-type cells adjacent to mutant cells that autonomously produce neural precursors are forced to adopt the epidermal fate. This data therefore suggests that the mutant cell strongly inhibit the wild-type ones. Evidence that this inhibitory signal is via Delta is shown by observations of clones doubly mutant for *Notch* and *Delta* which can no longer influence the fate of the neighbouring wild-type cells (Heitzler and Simpson, 1991; Heitzler and Simpson, 1993).

As stated above, in order for the model of lateral inhibition to work requires that contacts be made between all cells in the equivalence group. This stipulation sets an upper limit on the number of cells to approximately seven. This may be the case in the ventral neuroectoderm where proneural gene expression is restricted to a small number of cells (section 1.2). However, during PNS development in the imaginal discs proneural clusters with as many as

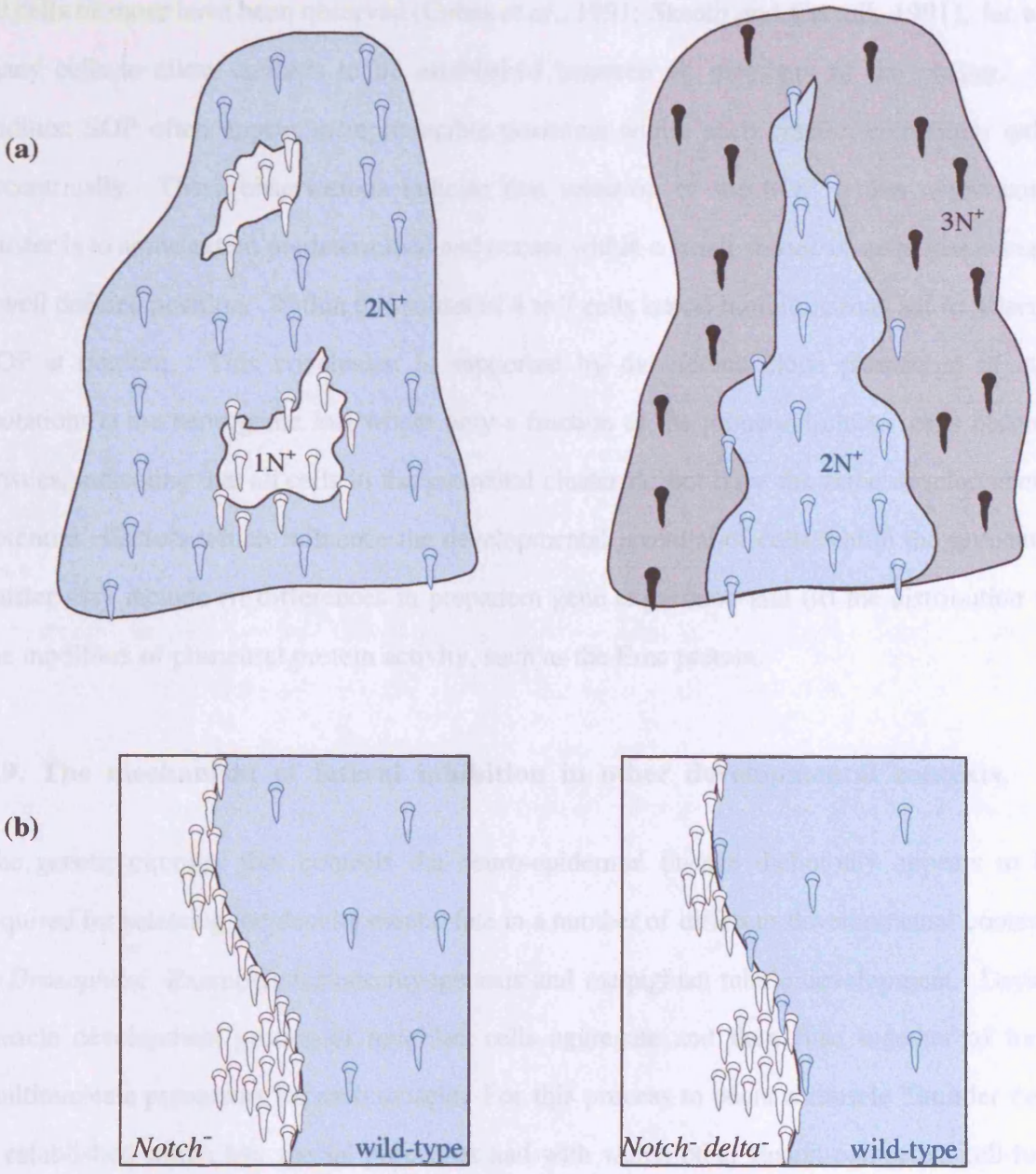


Figure 1.9. The effects of *Notch* dosage on the neuro-epidermal fate choice.
(a) Mosaic flies with adjacent territories carrying either one or two doses (left), or else two or three doses of *Notch* (right). Bristles which develop on the border of these clone always correspond to the cells with the lowest dosage of *Notch*. Thus adoption of the neural fate is dependant upon the relative levels of *Notch* between all the cells in the proneural cluster (figure adapted from Lawrence, 1992).
(b) Mosaic flies with patches of *Notch*$^-$ tissue in a wild-type background (left pannel). Cells mutant for *Notch* form bristles in the centre (not all bristles are drawn) and along the border of the clone. They thus display cell autonomy and are not influence by the presence of wild-type neighbouring cells. The wild-type cells are influenced by the adjacent *Notch*$^-$ cells however because, along the border of the clone, wild-type cells only differentiate as epidermal cells and never as bristle cells. The *Notch*$^-$ cells therefore inhibit neighbouring cells from becoming bristle cells. Evidence that the source of this inhibitory signal is Delta is shown in *Notch*$^-$ *delta*$^-$ clones (right pannel). Doubly mutant cells are no longer able to inhibit the adjacent wild-type cells from forming bristles.

30 cells or more have been observed (Cubas *et al.*, 1991; Skeath and Carroll, 1991), far too many cells to allow contacts to be established between all members of the cluster. In addition SOP often appear in reproducible positions within each cluster, sometimes quite eccentrically. These observations indicate that selection of the SOP within a proneural cluster is to some extent predetermined and occurs within a small subset of cells that occupy a well defined position. Within this subset of 4 to 7 cells lateral inhibition may act to select a SOP at random. This conclusion is supported by the mitotic clone phenotype of null mutations at the neurogenic loci where only a fraction of the proneural cluster cells become bristles, indicating that all cells in the proneural cluster do not have the same developmental potential. Factors which influence the developmental potential of cells within the proneural cluster may include (i) differences in prepattern gene expression and (ii) the distribution of the modifiers of proneural protein activity, such as the Emc protein.

1.9. The mechanism of lateral inhibition in other developmental contexts.

The genetic circuitry that controls the neuro-epidermal lineage dichotomy appears to be required for selecting the developmental fate in a number of different developmental contexts in *Drosophila.* Examples include myogenesis and malpighian tubule development. During muscle development groups of myoblast cells aggregate and then fuse together to form multinucleate precursors for each muscle. For this process to occur a muscle 'founder cell' is established which has special properties and with which other fusion competent cell fuse (Bate, 1990). The founder cells are distinct at the molecular level by the expression of specific genes such as S59 and *vestigial.* In embryos mutant for *Notch* and *Delta* an excess of founder cells are generated resulting in an increase in the number of cell committing to the muscle fate (Corbin *et al.*, 1991; Bate *et al.*, 1993). During malpighian tubule development the growth of the four tubules is controlled by a special type of cell called the tip cell which is found at the extreme tip of each tubule (Skaer, 1989). Segregation of the tip cells involve the singling out of a tip mother cell from a group of cells expressing the proneural genes (Hoch *et al.*, 1994). In the absence of *Notch* or *Delta* proneural gene expression persists in up to 12 cells that also all express tip cell markers.

The genetic circuitry is apparently active in other animals, homologues of the most important genes of the network have been described for a number of vertebrates and evidence indicates that the function of these genes is also to select cells from groups of equivalent ones (Chitnis *et al.*, 1995).

1.10. Repression of the neural fate.

So far it has been described that transcriptional activation of the genes of the *E(spl)-C* is the nuclear response to *Notch* pathway activation. The products of the *E(spl)* genes are essential to implement many of the cell fate decisions mediated by *Notch* signalling, not only in neural development as we have seen above, but also in other developmental contexts such as myogenesis and malpighian tubule development. Thus, a knowledge of the functional characteristics of the E(spl) proteins should lead to a greater understanding of how the activation of Notch mediates cell fate decisions via changes in gene transcription. During neural development the E(spl) proteins are responsible for the repression of the neural fate, which is probably mediated largely at the level of negative regulation of *AS-C* genes and/or their products. Much of the work carried out in this study is connected to the processes by which E(spl) negatively regulates the neural fate. Described below are four potential models by which a repressor molecule, such as E(spl), may function. Specific models of E(spl) mediated repression and the targets of this repression are considered in more detail in chapter 6. It is sufficient for present purposes to highlight the fact that these modes of repression are by no means mutually exclusive and it is possible, and indeed likely, that a number of modes of repression may be involved in antagonism of the neural fate.

(i) Protein-protein interaction between repressor and activator molecules.

Interactions between a repressor protein and an activator protein can antagonise the function of the activator by, for example, preventing it from binding DNA. An example of this kind of repression are the members of the Emc/Id class of HLH proteins which, as discussed previously (section 1.4), dimerize with bHLH activators to form inactive complexes which

are then rendered unable to bind DNA (Van Doren *et al.*, 1991; Cabrera *et al.*, 1994) (see section 1.5).

(ii) Competition for DNA target sequences.

Overlapping or shared DNA target sequences between repressor and activator proteins can result in competition for binding sites. Competitive binding as a means of repression has been proposed for homeodomain-containing proteins, which as a group have relatively poor DNA-binding specificity, and repressors may compete for generic homeobox sequences that are also bound by homeodomain activators (Han *et al.*, 1989).

(iii) Quenching.

Many repressors can co-occupy DNA with activators and prevent the activator from functioning. The simplest model to explain this would involve the repressor interacting with the activator to mask or in some other way compromise its activation surface. This mechanism of repression is often referred to as 'quenching'. Transcriptional quenching is a short range mechanism of repression which usually works over distances of less than 100bp, and generally does not affect basal transcription. The fly proteins Snail, Krüppel, Giant and Knirps have all been shown to quench the activity of local activators (Arnosti *et al.*, 1996; Gray *et al.*, 1994). As this mode of repression occurs at the local level within a small segment of DNA it does not affect the situation in neighbouring segments and thus allows multiple enhancers to function autonomously within a complex promoter.

(iv) Promoter silencing.

This mode of repression involves the repressor protein either (i) interacting directly with the basal transcriptional machinery itself to prevent it from reaching a transcriptionally competent state by, for example, disrupting assembly, structure or accessibility, or (ii) affecting the local chromatin configuration in the promoter to thus block access of basal transcription factors (for reviews see Johnson, 1995; Ashraf and Ip, 1998). Promoter silencing differs from the other modes of repression in two ways. Firstly, with respect to the range of action as this mode of repression can occur at distances up to and over 1kb and secondly,

repression mediated in this way is dominant to ensure that the promoter is repressed completely. An example of a transcriptional repressor which functions in this way is the *Drosophila* Hairy protein. The Hairy protein has been shown to mediate transcriptional repression of a reporter gene at distances of 1kb and experiments suggest that Hairy repression may involve direct interaction with one or more components of the basal transcription complex (Barlo and Levine, 1997).

An important question, addressed explicitly in chapter 6, is which of these mechanisms is used by E(spl) in repression of the neural fate? There are varying degrees of evidence for all four modes of repression. Firstly, E(spl) possesses a HLH domain which mediates interactions with other HLH proteins including some of the activator proneural proteins (Gigliani *et al.*, 1996; Alifragis *et al.*, 1997). Protein interactions of this kind may be employed *in vivo* to regulate the level of proneural gene activity. The potential for this mode of repression to occur certainly exists because E(spl) and the proneural proteins are expressed in coincident patterns in the developing nervous system.

The E(spl) bHLH proteins are site specific DNA-binding proteins which bind DNA via the basic domain (Oellers *et al.*, 1994). The DNA sites to which the E(spl) proteins have traditionally been considered to bind are known as N-box sequences (CACNAG) (Tietze *et al.*, 1992; Oellers *et al.*, 1994), which is distinct from the E-box sequence (CACGTG) that proneural protein dimers bind. N-box sequences and E-box sequences are found, often closely linked, in the upstream regions of a number of the proneural genes and also in the upstream regions of the neural precursor gene and proneural target *ase* (González *et al.*, 1989). These data indicate that competitive binding between the E(spl) and proneural proteins could occur for adjacent sites in the DNA and suggest that the E(spl) proteins may prevent transcriptional activation by sterically blocking activator access to DNA sites. More recently however, it has been shown that the preferred target binding site for the E(spl) proteins contains, at its core, an E-box sequence (Jennings *et al.*, 1999) rather than the N-box sequence previously described. The optimal target sequence differs from the optimal target sequence of the proneural proteins but there is at least some overlap in the sequences

recognised by the two classes of protein. This data suggests that the E(spl) and proneural proteins may compete for identical DNA sites in target genes.

Fisher *et al.* (1996) have demonstrated that a number of Hairy-related proteins including Hairy, E(spl)-m7 and Hes-1 are able to repress activated transcription in a cell culture assay. In these experiments, a reporter gene containing a modified *ac* promoter was activated by either Ac/Da or Sc/Da heterodimers. Efficient repression of this activated reporter gene was observed when either Hairy, E(spl)-m7 or Hes-1 were co-expressed with Ac/Da or Sc/Da in the cells. This experiment shows that the bHLH repressor proteins can repress activated transcription, in line with the possibility that a quenching mechanism may be functioning, but it does not however exclude other modes of repression.

The work of Paroush *et al.* (1994) and Fisher *et al.* (1996) have led to the current widely held view that the Hairy-related proteins, including E(spl), associate with Gro to form a transcriptional repression complex which, possibly in association with other unidentified factors, represses the transcription of target genes. Experiments have shown that the Hairy-Gro complex behaves as a long-range dominant repressor on synthetic promoters and suggest that repression is mediated by contacts with the basal transcriptionary machinery (Barlo and Levine, 1997). As the E(spl) proteins are also able to interact with Gro (Paroush *et al.*, 1994; Fisher *et al.*, 1996) it is likely that the mechanisms of E(spl) function will be similar to that of Hairy. Furthermore, in cell transfection assays a number of the E(spl) proteins are able to repress basal transcription from reporter genes containing the *ac* regulatory regions (Van Doren *et al.*, 1994; Heitzler *et al.*, 1996b). Additional support for a dominant long range mode of repression is provided by data from two recent experiments. Firstly, physical interaction between Gro and amino-terminal domain of histone H3 has been demonstrated suggesting that transcriptional repression complexes containing Gro may associate with chromatin and may promote the formation of a repressive chromatin configuration in the vicinity of binding. (Palaparti *et al.*, 1997). Secondly, physical and genetic interactions between Gro and the histone deacetylase Rpd3 have recently been demonstrated, suggesting that enzymes which modulate chromatin structure may also be

recruited to the complex and further support the idea that Gro mediated repression involves changes in chromatin configuration (Chen *et al.*, 1999).

1.11. Aims and experimental approach.

The work carried out in this study is based broadly on the function of *E(spl)* and *gro* during neural fate specification. There are five results chapters which can be grouped under three separate headings: (i) an analysis of the interaction domains of the E(spl) and Groucho proteins, (ii) an analysis of the modes by which *E(spl)* represses the neural fate and, (iii) the cloning and functional analysis of the *gro* homologue from the housefly (*Musca domestica*). The aims and experimental approach to each of these projects are outlined below.

(i) An analysis of the interaction domains of the E(spl) and Groucho proteins.

In chapters 3 and 4 a mutational analysis of the carboxyl-terminal WRPW motif of the E(spl)-m8 protein which is necessary for interaction with the Gro protein is performed. Previous data has indicated that the WRPW motif of the Hairy-related proteins, including those of the *E(spl)-C*, is critical for function. Precise removal of the domain or mutations therein disrupt *in vivo* function (Wainwright and Ish-Horowicz, 1992) and interactions with the Gro protein (Paroush *et al.*, 1994). In this study a systematic approach is taken to determine the importance of individual amino acids within this domain. Mutagenesis of the motif is performed and the mutant derivatives are tested, firstly, for *in vivo* function in the developing fly (chapter 3) and secondly, for ability to interact with the Gro protein in the yeast two-hybrid system (chapter 4).

In chapter 5 the reciprocal experiment is performed in an attempt to map the domains within the Gro protein which are required for interaction with the E(spl)-m8 protein. In order to do this the reverse two-hybrid system, a variation on the yeast two-hybrid scheme, is used.

(ii) An analysis of the modes of repression of E(spl) during neural fate specification.

The potential modes by which the E(spl) proteins repress the neural fate have been outlined above. In chapter 6 an assay is performed in order to determine which of these modes of repression are likely to be occurring *in vivo* during the specification of the SOP cells in imaginal development. The assay is specifically designed to determine whether other modes of repression occur *in vivo* distinct from those involving direct transcriptional repression of the *scute* or *daughterless* genes occur *in vivo*. The assay is performed by ectopically co-expressing either the *scute* or *daughterless* gene with the *E(spl)-m8* gene and analysing the adult bristle phenotype. The crux of the assay relies on the fact that ectopic expression of the *scute* and *daughterless* transgene is under control of a heterologous promoter (the upstream activation element or UAS element from yeast) and as such is relieved from direct transcriptional regulation by the E(spl) proteins. Therefore modification of the UAS-scute or UAS-daughterless driven phenotype by co-expression with UAS-E(spl)-m8 would indicate the involvement of modes of repression other than transcriptional repression.

(iii) The cloning and functional analysis of the gro homologue from the housefly.

In chapter 7 the cloning and functional analysis of the *gro* gene from the housefly (*Musca domestica*) is described. This work has formed part of an on-going project in the laboratory with the express aim of cloning all the *E(spl)-C* genes from the housefly in order to provide a comparative tool for the evolutionary analysis of the locus. A preliminary functional analysis of *Musca gro* is performed by using the technique of RNA interference (RNAi).

Chapter 2

Materials and Methods

2.1. Fly stocks, husbandry and methods.

2.1.1. Stocks.

E(spl)-m8WRPW mutant transformant lines:

w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RH}$}29	Chromosome 2.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RH}$}40	Chromosome 3.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RH}$}44	Chromosome 2.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RH}$}64	Chromosome 3.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RG}$}28	Unmapped. Multiple insertion.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RG}$}57	Chromosome 3.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RG}$}75	Chromosome 2.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RG}$}94	Chromosome 2.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RD}$}22	Unmapped.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RD}$}50	Chromosome 2.
w^{1118}; P{$w^{+mW.hs}$=UAS-E(spl)-$m8^{RD}$}51	Chromosome 3.

Other UAS-responder lines:

w^{*}; P{$w^{=mW.hs}$=UAS-sc^{+t}}39	Chromosome 2.
w^{*}; P{$w^{=mW.hs}$=UAS-E(spl)-$m8^{+t}$}4	Chromosome 3.
w^{*}; P{$w^{=mW.hs}$=UAS-da^{+t}}*	Chromosome 3.

P[GAL4] driver lines:

69B w*; P{$w^{+mW.hs}$=GawB}69B

GAL4 Protein in embryonic epidermis and imaginal discs. Chromosome 3.

T80 w*; P{$w^{+mW.hs}$=GawB}T80/CyO

Ubiquitous GAL4 protein in third instar imaginal discs. Chromosome 2.

ptc-559.1 w*; P{$w^{+mW.hs}$=GawB}ptc-559.1

GAL4 Protein in *ptc* gene pattern. Chromosome 2.

c591 w*; P{$w^{+mW.hs}$=GawB}

Near ubiquitous GAL4 protein in third instar wing disc. Chromosome 2.

Balancer stocks:

w*; Sco/CyO; MKRS/Tm6b

Other stocks:

p{ry$^{+t7.2}$=lArB}neur^{A101} ry^{506}/TM6B Tb^{1}

2.1.2. Drosophila fly culture.

All stocks were maintained on oatmeal medium at 25°C in vials unless otherwise stated. Brewers yeast media was used when crosses between single flies or small numbers of flies were performed. A grape juice medium was used for collecting *Drosophila* embryos.

Oatmeal medium.

6.8g agar, 130g oatmeal, 40ml black treacle, per litre of water and 0.11% v/v Nipagen M (4-hydroxybenzoic acid methylester, 20% w/v) to inhibit fungal growth. The surface of the medium was supplemented with a small amount of live yeast paste when set.

Brewers yeast medium.

9g agar, 46.3g sucrose, 46.3g dried yeast, per litre of water and 0.2% v/v Nipagen M (4-hydroxybenzoic acid methylester, 20% w/v) to inhibit fungal growth. The surface of the medium was supplemented with a small amount of dried yeast when set.

Grape juice media.

20g agar, 500ml grape juice, 500ml water, 10ml phosphoric/propionic acid mix. The surface of the medium was supplemented by a small amount of live yeast paste when set.

Alternative grape juice media.

19.8g agarose, 52.2g glucose, 26g sucrose, 7g dried yeast, 9% (v/v) red grape juice per litre of water and 0.1% (v/v) Nipagen M (4-hydroxybenzoic acid methylester, 20% w/v).

2.1.3. P-element mediated transformation.

pUAST-E(spl)m8 WRPW mutant plasmids were prepared for microinjection using a Qiagen midi-prep column following manufacturers instructions. DNA was co-injected at a concentration of 750µg/ml, with the helper plasmid phs-π-Δ2,3wc (Misra and Rio, 1990) at a concentration of 75µg/ml in injection buffer (0.1mM Na phosphate buffer, pH 7.8, 5mM KCl) into w^{1118} embryos using standard procedures (Rubin & Spradling, 1982; Spradling, 1986). Briefly, needles for injection were prepared using glass capillaries (Narishige) and pulled on a Narishige PC-10 needle puller, following manufacturers instructions. DNA was centrifuged to pellet any suspended particles and back-filled into needles using Eppendorf

microloader pippete tips. Embryos were collected on grape juice agar plates at room temperature at 1 hourly intervals, dechorionated in 3% sodium hypochlorite, lined up on agarose plates, transferred to a glass slide and desiccated for approximately 10 minutes. Embryos were then covered with Voltalef 10S oil and viewed using a inverted stage microscope and injected at 18°C using a micromanipulator and Narishige IM-30 nitrogen pump. Approximately 500 embryos were injected with each construct and allowed to hatch on grape juice agar plates at 18°C. Individual G_0 adults were mated with w^{1118} flies and G_1 progeny screened for red/orange eyes. The chromosomal location of insertions were mapped by crossing flies heterozygous for the insertion to the balancer stock *w; Sco/CyO; MKRS/Tm6b*. The progeny from this cross were then mated to w^{1118} and the location of insertions determined relative to the segregation of w^+ and the *CyO/MKRS* chromosomes.

2.1.4. Lethal phase analysis.

Flies were placed in plastic beakers inverted over grape juice agar plates and allowed to lay eggs overnight at 25°C. The development of the eggs was then monitored.

2.1.5. Wing dissection and mounting.

Wings form adult flies were dissected in 70% EtOH and mounted in canada balsam (Sigma).

2.1.6. Musca fly husbandry.

Adult *Musca* were maintained in plastic tanks (approximate dimensions 50cm x 30cm x 40cm) at 25°C and fed on sucrose and dried milk (provided separately in petri dishes). A large bottle of water, plugged with paper towels to form a wick, was provided as the water source. Egg laying begins 4-5 days after eclosion. For egg laying and larval culture, plastic tubs (15cm x 10cm x 10cm) containing natural bran mixed with fresh milk to a 'wet-paste' consistency and containing a pinch of dried yeast and approximately 1ml of 20% Nipagen M (4-hydroxybenzoic acid methylester) were used. After approximately 2-4 days (ensuring, during this period, that the egg laying vessel did not dry out) the egg laying tub was removed and placed in a fresh tank. The larvae will complete their development in the bran mixture, and will either pupate in the mixture or escape and pupate on the floor of the tank. Care should be taken during this period that the bran mixture does not dry; milk or water can be added (but do not add too much otherwise the larvae will drown). The complete life cycle takes approximately 13 days at 25°C.

Musca egg collection plates.

Petri dishes containing *Drosophila* brewers yeast media were supplemented with a chunk of Purrfect® rabbit and chicken catfood.

2.2. *in situ* hybridisation, immunohistochemistry, enhancer trap staining.

2.2.1. DIG-labelled RNA probe synthesis.

A linear template molecule was generated by restriction digest of pBluescript SK^- containing the appropriate sequence (using an enzyme which creates a 5' overhang). The digest was subsequently incubated with proteinaseK at 37°C for 30 minutes, phenol-chloroform extracted, EtOH precipitated and resuspended in RNase-free water at a concentration of 0.5mg/ml. Probe synthesis was carried out in a 20µl final volume using 1µg of template DNA, 1x RNA transcription buffer, 40 units of T7 RNA polymerase, 1x DIG RNA labelling mix (Digoxigenin-UTP) (Boehringer Mannheim), 40 units of RNasin (Promega), and incubated at 37°C for 1-2 hours. All procedures were carried out using RNase-free plasticware. To the probe synthesis reaction 70µg of tRNA and 0.4M LiCl was added and the final volume adjusted to 100µl. The probe was then precipitated by with 2.5 volumes EtOH, and washed in 70% EtOH. The precipitated probe was resuspended in 100µl of RNase-free water.

2.2.2. in situ hybridisation - wing imaginal discs.

Wing imaginal discs were dissected from third instar larvae in PBS (phosphate-buffered saline), fixed in 4% paraformaldehyde/PTW (PBS + 0.3% Tween), washed in PTW and hybridised overnight at 70°C in 400µl of hybridisation buffer (50% v/v deionised formamide, 5x SSC, 1mg tRNA, 500µg heparin, 0.1% tween 20, pH adjusted to 4.5 with 1M citric acid) containing 1µl of the heat denatured RNA probe. The excess probe was removed with several washes with PTW. The hybridised discs were then incubated with pre-absorbed anti-DIG antibody in PTW (1:2000 final concentration) for 1 hour at room temperature. Staining was performed in the dark in 1ml of staining buffer (0.1M NaCl, 50mM $MgCl_2$, 0.1M Tris-HCl pH9.5, 0.1% Tween 20) with 4.5µl NBT and 3.5µl BCIP (Boehringer Mannheim). Discs were dehydrated in an EtOH series and mounted on slides in methacrylate (JB-4, Polysciences).

2.2.3. in situ hybridisation - Drosophila and Musca embryos.

Embryos were collected overnight, harvested and dechorionated for 2 minutes in 3% bleach. Embryos were fixed in 10% paraformaldehyde (PBS)/heptane on rotor at room temperature for 20 minutes. Chorion removal was carried out by osmotic and temperature shock in MeOH at 70°C with vigorous shaking. Embryos were rehydrated through a MeOH/PBS series, subjected to a brief 5-10 minute secondary fixation in paraformaldehyde, washed 3x in PTW (PBS + 0.3% Tween) and hybridised overnight at 70°C in 1ml of hybridisation buffer (50% v/v deionised formamide, 5x SSC, 1mg tRNA, 500µg heparin, 0.1% tween 20, pH adjusted to 4.5 with 1M citric acid) containing 1-4µl of the heat denatured RNA probe. The excess probe was removed with several washes with PTW. The hybridised

discs were then incubated with pre-absorbed anti-DIG antibody in PTW (1:2000 final concentration) for 1 hour at room temperature. Staining was performed in the dark in 1ml of staining buffer (0.1M NaCl, 50mM $MgCl_2$, 0.1M Tris-HCl pH9.5, 0.1% Tween 20) with 4.5µl NBT and 3.5µl BCIP (Boehringer Mannheim). Embryos were dehydrated in an EtOH series and mounted on slides in methacrylate (JB-4, Polysciences).

2.2.4. α-HRP antibody staining.

Embryos were collected overnight, harvested, dechorionated in 3% bleach and fixed in 4% paraformaldehyde (PBS)/heptane on rotor at room temperature for 20 minutes. Vitelline membrane removal was carried out by osmotic shock with MeOH and vigorous shaking. Embryos were rehydrated through a MeOH/PBS series, subjected to a brief 5-10 minute secondary fixation in paraformaldehyde, washed 3x in PBT (PBS + 0.3% Triton), and blocked in PBT + 5% goat serum for at least 30 minutes. The embryos were then incubated with rabbit α-HRP antibody (Jackson) in PBT (1:400) at 4°C overnight. The excess antibody was removed with several washes in PBT (PBS + 0.1% Triton). The embryos were then incubated with pre-absorbed goat α-rabbit secondary antibody (Jackson) in PBT (1:1000) for 1-2 hours at room temperature. Prior to staining embryos were pre-incubated for 10 minutes in 0.5mg/ml DAB (PBS). Staining was performed in 0.5mg/ml DAB (PBS) + 0.03% H_20_2 for 1-10 minutes. The staining reaction was stopped by several rinses in PBS. Embryos were dehydrated in an EtOH series and mounted on slides in methacrylate (JB-4, Polysciences).

2.2.5. A101 enhancer trap line staining.

Third instar imaginal discs were dissected and accumulated on ice in PBS. The discs were then fixed 10 minutes at room temperature in 1% glutaraldehyde/PBS and then washed x3 5-10 minutes in PBS. Staining was performed in staining solution (5mM $K_4[Fe^{II}(CN)_6]$, 5mM $K_3[Fe^{III}(CN)_6]$, 0.3% Triton, 0.2% (w/v) Xgal in PBS) for a minimum of 2 hours. Stained discs were dehydrated in an EtOH series and mounted on slides in methacrylate (JB-4, Polysciences).

2.3. Site directed mutagenesis.

The site directed mutagenesis was carried out using the Altered Sites™ *in vitro* mutagenesis system (Promega) following manufacturers instructions. In summary, pALTER (plasmid with inactivated ampicillin resistance gene) containing the wild-type *E(spl)m8* coding sequence was denatured, annealed with the ampicillin repair oligonucleotide and one of the mutagenic oligonucleotides (m8RHb, m8RG, m8RD). The primers were extended by DNA synthesis using T4 DNA polymerase, followed by DNA ligation using T4 DNA ligase. This

DNA was transformed into ES1301 *mutS* (repair defective strain) and grown overnight in Luria broth in the presence of 65μg/ml ampicillin. The overnight culture was miniprepped and the plasmid DNA was transformed into DH5α competent cells. Ampicillin resistant clones were miniprepped and the DNA was subject to restriction analysis using diagnostic restriction enzymes (*Nco*I, *Apa*I). Positive clones were sequenced using the Xm8 primer to confirm the presence of the mutation.

2.4. Bacteria.

2.4.1. Strains.

ES1301 *mutS*	*lacZ*53 *mutS*201::Tn*5* *thyA*36 *rha*-5 *metB*1 *deoC* IN(*rrnD-rrnE)*
DH5α	F$^-$ ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR* *recA*1 *endA*1 *hsdR*17(r_K^-, m_K^+) *phoA* *sup*E44 λ$^-$ *thi*-1 *gyrA*96 *relA*1

2.4.2. Bacteria growth and storage.

Bacterial colonies were grown on inverted plates, or in liquid culture with vigorous shaking, at 37°C, with appropriate antibiotic selection when necessary.

Culture media.

L-broth: 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, per litre of water and adjusted to pH 7.0 with NaOH.

L-agar: As above with the addition of Bacto-agar (Difco) to 1.5% (w/v).

Storage.

Bacteria were stored as glycerol stocks (1.5ml glycerol added to 0.85ml of culture) at -70°C.

2.4.3. Antibiotics and indicators.

When selection of bacteria in culture was necessary ampicillin was added at a final concentration of 65μg/ml (50mg/ml stock solution in sterile distilled water) to broth or agar. In order to detect recombinant clones, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and IPTG (isopropyl-β-D-thiogalactopyranoside) were added to molten agar (50°C) to a final concentration of 40μg/ml and 0.5mM respectively.

2.4.4. Transformation of E. coli.

Preparation of competent cells.

DH5α competent cells were prepared according to Nishimera *et al.*, 1990. Briefly, a 50ml culture of L-broth supplemented with 10mM $MgSO_4.7H_2O$ and 0.2% glucose was inoculated with 0.5ml of an overnight culture of DH5α and grown with aeration at 37°C until cells had entered the logarithmic growth phase (OD_{600} = 0.4). The cells were cooled on ice for 10 minutes, and pelleted at 1500g for 10 minutes at 4°C. The cell pellet was then resuspended gently in 0.5ml of cold L-broth supplemented with 10mM $MgSO_4.7H_2O$ and 0.2% glucose, and then 2.5ml storage solution (36% glycerol, 12% PEG (MW7500), 12mM $MgSO_4.7H_2O$ added to L-broth and sterilised by filtration) was added. The competent cells were stored at -80°C until use.

Transformation.

Routinely, 3µl of ligation reactions (10µl) was added to 100µl of competent cells, the mixture was placed on ice for 30 minutes, heat shocked at 42°C for 1 minute and then cooled on ice for 2 minutes. 0.4ml of L-broth was added to each transformation and incubated at 37°C for 1 hour to allow expression of antibiotic resistance. 200µl was spread onto pre-dried L-agar plates containing the appropriate antibiotics and indicators, and incubated overnight at 37°C.

2.4.5. Plasmid DNA isolation.

Small scale plasmid isolation was carried out using either the silica binding method of Carter & Milton (1993) or using Qiagen spin mini-prep columns following the manufacturers instructions. Large scale plasmid isolation was carried out using Qiagen midi- or maxi-prep columns following manufacturers instructions.

2.5. Yeast.

2.5.1. Yeast strains.

EGY48 *MATα trp1 his3 ura3 ::(LexAop)6-LEU2*

2.5.2. Yeast growth and storage.

Yeast strains were grown on inverted plates, or in liquid culture with vigorous shaking, at 30°C.

Culture media.

YPD media: 10g yeast extract 20g peptone, 0.1g NaOH, 100ml glucose per litre of water.

Drop-out media:	6.7g yeast nitrogenous base (YNB), 2g drop-out powder (DOP) lacking appropriate nutrients, 0.1g NaOH, 20 ml sugar source (glucose or galactose/faffinose).
X-gal plates	6.7g YNB, 1.5g DOP, 20g agar, 0.1g NaOH, 20ml sugar source, 100ml 10x BU salts, X-gal to a final concentration of 80µg/ml.

Storage.

Yeast strains were stored in 15% glycerol at -80°C. To inoculate vials for storage, 500µl of late-log phase culture was added to vials containing 30% sterile glycerol, frozen on dry ice and transferred to the -80 °C freezer. The cells were revived by scraping some of the cells off the frozen surface onto a plate containing the appropriate media.

2.5.3. Yeast transformation.

A 20ml culture of YPD or appropriate drop out media was inoculated with a single yeast colony and grown with shaking overnight at 30°C to approximately mid logarithmic phase. The cells were harvested, washed once in 25ml sterile water, and once in 1ml 0.1M lithium acetate, transferred to an eppendorf and then resuspended in a final volume of 500µl 0.1M lithium acetate. 50µl aliquots of the cell suspension were transferred to a fresh eppendorf, pelleted and the supernatant was removed. The following reagents were added to the cell pellet (in the order stated): 240µl 50% PEG 4000, 36µl 1M lithium acetate, 25µl 2mg/ml sonicated single stranded salmon sperm DNA, and 50µl water containing 1µg of plasmid DNA. The reagents were mixed by vortexing vigorously for at least 1 minute followed by incubation at 30°C for 30 minutes. The transformation mix was then heat shocked at 42°C for 20-25 minutes. The cells were pelleted at 7000rpm for 15 seconds, gently resuspended in sterile water, then plated onto drop out media plates containing glucose and incubated at 30°C for 2-3 days until colonies of the appropriate size had grown.

2.5.4. Plasmid rescue from yeast.

A 5ml culture containing the appropriate drop out media was inoculated with a single yeast colony and grown overnight with shaking at 30°C. the cells were harvested, washed once in 1ml sterile water, transferred to an eppendorf, pelleted and resuspended in 800µl βME buffer (50mM sodium phosphate buffer pH7.5, 0.9M sorbitol, 1µl of β-mecaptoethanol per 20ml). Zymolase (Sigma) was added at a final concentration of 0.3mg/ml and incubated at 37°C for 45 minutes, after which the enzyme was heat inactivated at 70°C for 20 minutes. 5M potassium acetate was added to a final volume of 1M and placed on ice for 45 minutes. The precipitate was pelleted, the supernatant was transferred to a fresh eppendorf and the plasmid DNA was precipitated with isopropanol, and washed in 70% EtOH. The pellet was redissolved in 20µl TE. Larger quantity, higher quality DNA was generated by transforming a 2µl aliquot of this preparation into *E.coli*. Plasmid preparations from *E.coli* were then performed using Qiagen columns following the manufacturers instructions.

2.5.5. β-galactosidase assay in liquid culture.

A 5ml culture of appropriate drop out media and sugar source was inoculated with a single yeast colony and grown overnight with shaking at 30°C to mid- to late-log phase (OD_{600} = 0.4 to 0.7). The cells were harvested, resuspended in 5ml of Z buffer and OD_{600} was measured. The following two reaction tubes for each sample were set up: (i) 100µl cells with 900µl Z buffer (60mM $Na_2HPO_4.7H_2O$, 40mM $NaH_2PO_4.H_20$, 10mM KCl, 1mM $MgSO_4.7H_2O$, 50mM β-mecaptoethanol) and (ii) 50µl cells with 950µl Z buffer. To permeabalize the cells 10µl of 0.1% SDS and 50µl of chloroform were added to each tube and then vortexed for 15 seconds. The cells were equilibrated in a 30°C waterbath for 15 minutes and then 0.2ml of 4mg/ml ONPG (*o*-nitrophenyl-β-D-galactoside)was added and the tubes returned to the waterbath. The reactions were stopped by the addition of 450µl 1M sodium carbonate when a medium yellow colour developed. The cells were pelleted and the OD_{420} and OD_{550} of the supernatant was measured. Units of β-gal activity were calculated using the following equation:

$$U = \frac{1000 \times [(OD_{420}) - (OD_{550} \times 1.75)]}{(t) \times (v) \times (OD_{600})}$$

where t = time of reaction (mins)

v = volume of culture used in the assay

OD_{600} = cell density at the start of the assay

OD_{420} = combination of absorbence by *o*- nitrophenol and light scattering by cell debris

OD_{550} = light scattering by cell debris

2.6. Reverse two-hybrid screen.

2.6.1. Mutagenic PCR.

Four separate reactions, each containing a limiting amount of one of the dNTPs were performed. The mutagenic PCR reactions were carried out in a 100µl volume containing 10ng of template DNA, dNTP concentration was either 20µM for the limiting dNTP or 200µM for the remaining nucleotides, 200µM dITP, 1x standard PCR buffer, primer m9I and primer m9II at a concentration of 0.5µM each and 5 units of *Taq*-polymerase. Reactions were performed in an MJ Research PTC-200 peltier thermal cycler. The following PCR conditions were used for 30 cycles: 94°C for 30 seconds; 48°C for 1 minute; 72°C for 2 minutes. The products from each of the four reactions were then pooled.

Control non-mutagenic amplicons were generated by performing PCR under non-mutagenic conditions: 10ng template DNA, 200μM of all four dNTPs 1x standard PCR buffer, primer m9I and primer m9II at a concentration of 0.5μM each and 5 units of *Pfu* polymerase.

2.6.2. The screen.

Serial transformation of the yeast strain EGY48 was performed, firstly with the *lacZ* reporter plasmid pSH18-34 and then with pEG202-E(spl)-m8 to generate the host strain used in the screen. The pJG4-5-Gro vector was linearised with *Mlu*I and the ends repaired with Klenow and dephosphorylated. 1μg of linearized vector was transformed together with a five fold molar excess of fragment from either the mutagenic PCR or the non-mutagenic control PCR into the host yeast strain. The transformation procedure was identical to that previously described (section 2.5.3). To perform the screen colony lifts were taken using nylon membranes (Hybond-N) and transferred to X-gal indicator plates containing 2% galactose and 1% raffinose and allowed to grow for 24 hours at 30°C. White colonies were picked and plasmid rescue was performed to obtain the pJG4-5-Gro vector. The vector was then retransformed back into the host strain and a secondary screen was performed. Sequence analysis of the pJG4-5-Gro vector was conducted with the m9V and m9VI sequencing primers.

2.7. General molecular biology.

2.7.1. Fly genomic DNA isolation.

Rapid genomic DNA isolation.

High molecular weight chromosomal DNA was prepared from 20 flies by the method described by Hamilton *et al.*, (1991) with modifications. Flies were homogenised in a 1.5 ml Eppendorf microcentrifuge tube with an Eppendorf micropestle in 500μl of 5% sucrose, 80mM NaCl, 0.1M Tris-HCl, pH 8.5, 0.5% SDS (w/v), 50mM EDTA. The homogenate was stored at -20°C until frozen then incubated at 70°C for 30 minutes. KOAc was added to a final concentration of 160mM and placed on ice for 30 minutes. After centrifugation to remove the precipitate, the aqueous phase was collected and extracted twice with an equal vol of phenol/chloroform (1:1), and once with chloroform. DNA was precipitated with 0.7 vol of isopropanol and collected by centrifugation. The DNA pellet was washed with 70% EtOH, dried and resuspended in 50μl of TE (1mM EDTA, 10mM Tris-HCl, pH 8.0) containing 10μg/ml of RNase A. The yield was typically 15μg of DNA/20 flies.

2.7.2. Quantification of DNA.

Quantification of the amount of DNA in a sample was performed by taking readings in a spectrophotomoter at a wavelength of 260nm. An OD_{260} = 1 corresponds to ~50μg/ml for double-stranded DNA. Readings were also taken at a wavelength of 280nm, the ratio

between the two readings (OD_{260}/OD_{280}) provided an estimate of the purity of the DNA. An OD ratio of 1.8 indicates a pure preparation of DNA. Quantification of small amounts of DNA were estimated by comparing the fluorescent yield of the sample run out on a agarose gel with that of a series of standards.

2.7.3. Restriction enzyme digests.

DNA was digested with restriction enzymes supplied by GibcoBRL with the appropriate buffer recommended and provided by the manufacturer. At least 2 units of enzyme/μg of DNA was used and incubations were carried out at 37°C. Plasmid DNA was digested for ~1 hour and genomic DNA for ~5 hours.

2.7.4. DNA ligation.

Ligations were carried out in a final volume of 10μl in 1X ligation buffer (50mM Tris-HCl: pH 7.6, 10mM $MgCl_2$, 10mM DTT), 0.5mM ATP (for ligation of cohesive termini) or 0.2mM (for ligation of blunt-ended termini) and 1 unit of T4 DNA ligase (GibcoBRL) with the appropriate volumes of insert and vector DNA (in general a 3:1 molar ratio of insert to vector was used). In addition, for blunt-ended ligations, 5% PEG 8000 was added. Ligations were incubated overnight at 16°C.

2.7.5. Vector phosphorylation.

Phosphorylation reactions were carried out in a final volume of 20μl, with 1 unit shrimp alkaline phosphatase (USB) and 1x SAP buffer (20mM Tris-HCl, pH8.0, 10mM $MgCl_2$) and incubated for 1-2 hours at 37°C, followed by incubation for 15 minutes at 65°C to inactivate the enzyme.

2.7.6. Dideoxy sequencing.

Sequencing of double-stranded DNA was performed by the dideoxy chain-termination method (Sanger *et al.*, 1977) using the Sequenase Version 2.0 kit (U.S. Biochemical Corporation) with T7 DNA polymerase and S^{35}, following instructions provided by the manufacturer. Products of the DNA sequencing reactions were separated on 6% polyacrylamide gels (National Diagnostics).

2.7.7. Gel electrophoresis of nucleic acids.

Loading buffer.

Before DNA samples were loaded onto gels, 0.1 vol of 10x loading buffer (25% (w/v) Ficoll 400, 0.4% (w/v) bromophenol blue, 0.4% (w/v) xylene cyanol) was added to each.

Agarose gel electrophoresis.

DNA was electrophoresed on agarose gels in 0.5 x TBE (45mM Tris-borate, pH 8.3, 1mM EDTA). A range of agarose concentrations [0.8-1.5% (w/v)] were used depending on the sizes of fragments to be resolved (Sambrook *et al.*, 1989). Gels were run in electrophoresis buffer (0.5x TBE, 0.5µg/ml EtBr). Applied voltages varied between 2 and 10 V/cm, depending on the time of running. 200ng of either 1Kb ladder, ϕx174 or λ*Hin*DIII (GibcoBRL) were used as DNA size markers.

Low melting point agarose gels.

DNA fragments were separated on 1.2% (w/v) LMP (low melting point agarose) agarose gels in 0.5 x TBE (45mM Tris-borate, pH 8.3, 1mM EDTA) and were excised in a small slice of gel. Purification of the DNA fragments from the gel slice was performed using the Qiagen Qiaquick gel extraction kit following the manufacturers instructions.

Denaturing polyacrylamide gel electrophoresis.

Products of DNA sequencing reactions were separated on denaturing polyacrylamide gels prepared with 6% (w/v) acrylamide (N,N'-methylenebisacrylamide, 19:1), 7M urea, in 1 X TBE. Polymerisation was catalysed by the addition of 80µl of 10% (w/v) ammonium persulfate and 4µl of TEMED (N,N,N'N'-tetramethylenediamine) per 10ml of gel. Polymerised gels were pre-run at ~80W until they reached 50°C . Before loading, samples were denatured for 2 min at 75°C and gels were run for 2-5 hours. Gels were dried for 2 hr at 80°C under vacuum onto Whatman 3MM paper and then exposed to X-ray film overnight at room temperature without intensifying screens.

Visualisation of gels.

DNA was visualised by UV induced fluorescence of EtBr (0.5µg/ml) on a short wave (254nm) transilluminator.

2.7.8. Labelling of nucleic acids.

Gel-purified DNA fragments were labelled by random priming (Feinberg and Vogelstein, 1983), using the "Ready-to-Go" kit (Pharmacia) following the instructions of the manufacturer. Probes were separated from unincorporated radionucleotides by Sephadex G50 (Pharmacia) chromatography. Probes generally had a specific activity of 10^8-10^9 cpm/µg of DNA.

2.7.9. Southern blotting.

DNA was electrophorised on an agarose gel and transferred to nylon membranes (Hybond-N), by capillary transfer (Sambrook *et al.*, 1989) and fixed to the membrane by UV treatment as instructed by the manufacturer (Amersham UK). Filters were pre-hybridised for a minimum of 4 hr at 65°C in church buffer (0.5M Sodium phosphate buffer, 1% w/v

BSA, 1μM EDTA, 7% w/v SDS). Hybridisation with a denatured radioactive DNA probe was carried out at 65°C for a minimum of 12 hr in fresh church buffer. After hybridisation, the filters were washed in 2x SSPE, 0.1% SDS at 65°C for 15-20 minutes, 1x SSPE at 65°C for 15-20 minutes and then (in some cases) in 0.1X SSPE, 0.1% SDS at 65°C for 15-20 minutes.

2.7.10. Autoradiography.

Autoradiography of radioactive filters was carried out at -80°C with intensifying screens with exposure to Fuji NIF RX X-ray film. Films were developed using a DuPont Cronex CX-130 film processor.

2.7.11. Preparation of single stranded high molecular weight carrier DNA.

Salmon sperm DNA (Sodium salt -Sigma) was dissolved in TE pH 8.0 at a final concentration of 10mg/ml by stirring overnight at 4°C. The DNA solution was sonicated to achieve a size distribution of between 2kb and 15kb with a mean size of approximately 7kb. Estimation of the size distribution was carried out on a 0.8% agarose gel. The sonicated DNA solution was extracted twice with phenol/chloroform and once with chloroform, ethanol precipitated and resuspended at a final concentration of 2mg/ml in TE pH 8.0. The DNA was boiled in a waterbath for 10 minutes, quenched on ice and stored in aliquots at -20°C. Immediately prior to use an aliquot was thawed, boiled for 10 minutes and placed on ice until required.

2.8. Screening of bacteriophage libraries.

Screening of a λZAP *Musca domestica* embryonic cDNA library was performed as described in Sambrook *et al.*, (1989). Host cells were prepared by inoculating 100ml of L-broth containing 0.2% maltose with 100μl of an overnight culture of XL1-blue, which were grown at 37°C with vigorous aeration until cells had entered logarithmic growth phase (OD_{600} of 0.4-0.6). Cells were pelleted by brief centrifugation and resuspended in 0.2 culture volumes of 10mM $MgSO_4$. Phage particles were mixed with 2ml of host cells and incubated at 37°C for 20-30 mins to allow adhesion of phage to cells. This was mixed with 30ml of warm BBL-top overlay and poured onto a dried 20 x 20cm, pre-warmed L-broth plate. The plate was inverted once set and incubated for 8-13 hours at 37°C. Approximately 3×10^5 recombinant phage were screened. Duplicate filter lifts were taken, which were then pre-hybridised for 5 hr at 65°C in Church buffer. Filters were hybridised with labelled probe in hybridisation solution at 65°C for 16 hr, and washed with 2x SSC at room temperature for 15 min, then 1x SSC, 0.1% SDS at 65°C for 15 min with a final wash of

0.1x SSC, 0.1% SDS at 65°C for 5 min. One positively hybridising phage was isolated and purified.

2.9. The polymerase chain reaction (PCR).

PCR reactions were carried out in a volume of 10-100µl containing approximately 100-200ng of template DNA in 1X buffer (50mM KCl, 10mM Tris-HCl (pH 8.3), 1.5mM $MgCl_2$, 0.01% (v/v) gelatin), with the four dNTPs at a final concentration of 200µM each, primers at a concentration of between 0.33µM-0.5µM each and 1-10 units of *Taq* polymerase (Advanced Biotechnologies). Reactions were performed in a MJ Research PTC-200 peltier thermal cycler. Standard cycling parameters used were as follows for 30 cycles: denaturation of 30 seconds at 94°C, annealing of 1 minute at 55-65°C, and extension of 1 minute at 72°C.

2.9.1. Purification of PCR products.

When purified PCR products were required, PCR reactions were separated on 1.2% (w/v) LMP (low melting point) agarose gel in 0.5x TBE. DNA fragments were visualised by staining with EtBr (0.5µg/ml) and excised in a small slice of agarose gel. Purification of PCR products from the gel slice was performed using the Qiagen Qiaquick gel extraction kit following the manufacturers instructions.

2.9.2. Cloning of PCR products.

PCR products were cloned into pBluescript II SK^- by taking advantage of the terminal transferase activity of *Taq* polymerase (Clark, 1988). 5µg of pBluescript II SK^- were digested with *Eco*RV, purified using Qiagen gel extraction columns, and incubated with *Taq* polymerase (1.6 unit/20µl volume) using standard PCR buffer in the presence of 2mM dTTP for 2 hours at 72°C. The reaction was then purified using Qiagen gel extraction columns. DNA ligations were carried out in a final volume of 10µl, using 20-100µg of T-tailed pBluescript II vector, with a 4 to 6-fold excess of insert DNA to vector.

2.9.3. Degenerate PCR.

Degenerated PCR reactions were generally carried out in a 20µl volume containing approximately 100-200ng of template DNA in 1X buffer (50mM KCl, 10mM Tris-HCl (pH 8.3), 1.5mM $MgCl_2$, 0.01% (v/v) gelatin), the four dNTPs at a final concentration of 200µM each, the degenerate primers at a concentration of between 2-5µM each and 2 units of *Taq* polymerase (Promega). Reactions were performed in a MJ Research PTC-200 peltier thermal cycler. Standard cycling parameters used were as follows for 30 cycles:

denaturation of 1 minute at 94°C, annealing of 1 minute at 40-55°C, and extension of 1 minute at 72°C.

Mutagenic PCR (see above 2.6.1)

2.10 Primers.

2.10.1.Oligoneucleotides.
Olidonucleotide synthesis.
Oligonucleotides were synthesised by Bioline.

Quantification of oligonucleotides.
Concentration of an oligonucleotide solution was determined by using its molar extinction coefficient. This was calculated by summing the products of the molar extinction coefficients of the four deoxynucleotides multiplied by the number of times each of those nucleotides occurred within the sequence. Molar concentrations were thus calculated from the relationship: [DNA] = A_{260} / [(pG x 12010) + (qA x 15200) + (rT x 8400) + (sC x 7050)] for an oligonucleotide of composition GpAqTrCs.

Table 2.1. Primer sequences used in this study.

The following table lists the primers used in this work. All primers are orientated 5' to 3'. 'P' denotes a 5' phosphate group. The degenerate code is as follows: AC = M, AG = R, AT = W, GC = S, CT = Y, GT = K, AGC = V, ACT = H, AGT = D, GCT = B, ACGT = N.

Primer name	Sequence (5' to 3')	Chapter
m8RHb	P CCCTTGTGGCACCCATGGTAAAAC	3
m8RG	P CTTGTGGGGCCCCTGGT	3
m8RD	P GCCCTTGTGGGACCCATGGTAAAAAC	3
AmpR	P GTTGCCATTGCTGCAGGCATCGTGGTG	3
Xm8	GCTCTAGATGCGTTGTTGG	3
LEX1	CGTCAGCAGAGCTTCACCATT	2
m9I	TGGAATCCCACTCACCGC	5
m9II	GTGTTGATCTGACGGGCG	5
m9V	CAAGAGGATGAAGGCGAG	5
m9IV	GTGTACACGTACTTTGTGG	5

Mugro 1	AYGTIATGTAYTAYGARATG	7
Mugro 2	GTRCGDKCIGTYTADTTRTG	7
Mugro 3	TGYGTNAARGTNTGGGAYAT	7
Mugro 4	TGGYTRGCYGTNGGYATGGA	7
Mugro 4	TCCATRCCNACRGCYARCCA	7
Mugro 6	TANACKGTRGCYTTYTTRTC	7
chico	AGCTGACTAATTGGACTC	7
T3	TAACCCTCACTAAAGGGA	-
T7	TAATACGACTCACTATAGG	-

2.11. RNA interference.

2.11.1. Generation of dsRNA.

pBluescript II SK$^-$ containing the 770bp *Musca gro* fragment was used to generate the dsRNA. To generate the template the plasmid was linerised with either *Eco*RI (sense) or *Xho*I (antisense), incubated with protenase K at 37°C for 30 minutes, phenol-chloroform extracted, EtOH precipitated and resuspended in RNase-free water. The sense and antisense transcription reactions were performed separately *in vitro* in a final volume of 20µl using 1µg of template DNA, 1x RNA transcription buffer, 40 units of either T7 (sense) or T3 (antisense) RNA polymerase, 1mM rNTPs (Boehringer Mannheim), 40 units of RNasin (Promega), and incubated at 37°C for 1-2 hours. RNA was then phenol-chloroform extracted, EtOH precipitated and resuspended in 100µl RNase-free TE. Annealing was carried out by mixing equimolar ratios of each strand, boiling for 1 minute in a large beaker of water, removing the beaker from the heat and cooling to room temperature over a period of approximately 18 hours. A small aliquot was checked on a agarose gel to determine that annealing had occurred. dsRNA was stored as 'single use' aliquots under ethanol at -80 °C. When required an aliquot of the dsRNA was precipitated by centrifugation, washed in 70% ethanol and resuspended in injection buffer at an approximate concentration of 3µM. Immediately prior to use the dsRNA was backloaded into RNase-free glass capillary needles. All procedures were carried out using RNase-free plasticware.

2.11.2. Quantification of the dsRNA.

dsRNA was quantified by running a small aliquot on an agarose gel and comparing the fluorescent yield of the sample with that of Øx174 DNA marker of known concentration.

2.11.3. Injection procedure.

The optimal site of injection was determined by injecting syncytial blastoderm *Musca* embryos with injection buffer alone and then determining the percentage of survivors. The

optimal site of injection was determined to be an anterior region on the dorsal side of the embryo. *Musca* eggs were collected at approximately two hourly intervals on egg laying plates. Eggs were harvested from the plate, dechorionated for approximately 30 seconds in 1.2% bleach and washed firstly with 0.7% NaCl plus 0.02% Triton and then with 0.7% NaCl alone. The eggs were then lined up directly onto a glass slide, desiccated for 5-10 minutes and covered with halocarbon oil. Embryos were injected at the syncytial blastoderm stage in a position on the anterior dorsal side of the embryo. Embryos which had progressed beyond the syncytial blastoderm stage were not injected and were subsequently removed. The slides containing the injected embryos were then allowed to develop at 18°C in humid conditions.

Chapter 3

Mutational analysis of the E(spl)-m8 WRPW motif

INTRODUCTION

3.1. The WRPW motif of the Hairy-related proteins is required for transcriptional repression and for interaction with the Groucho protein.

The Hairy-related protein family includes members encoded by the *Drosophila* loci *hairy*, *deadpan* and bHLH genes of the *Enhancer of split complex* (*E(spl)-C*), and also proteins encoded by the mammalian *Hes* genes (*hairy* and *E(spl)* homologues). These proteins are characterised by a number of highly conserved domains: (i) a basic Helix-Loop-Helix (bHLH) region involved in DNA binding and protein dimerization; the Orange domain, that may form two putative helices and which could be responsible for determining specificity amongst the members of the family (Dawson *et al.*, 1995); and a carboxyl-terminal tryptophan-arginine-proline-tryptophan (WRPW) motif. The importance of the WRPW motif was initially shown by Wainwright and Ish-Horowicz (1992) when they analysed *hairy* alleles and found two independently isolated mutations which both resulted in proline to leucine substitutions within the motif. Interestingly, the WRPW is present in all bHLH repressors (i.e. the Hairy-related proteins) and absent from all bHLH activators (i.e. the proneural proteins). Taken together, these observations suggested that the WRPW motif was an important functional domain and hinted that it may be involved, in some way, in the repression function of proteins possessing it.

The first indication of the functional role of the motif came from a yeast two-hybrid screen performed by Paroush *et al.* (1994) in an effort to select for proteins that interacted with the *Drosophila* Hairy protein. A fusion protein containing Hairy fused to the LexA DNA binding domain was used as the 'bait' to fish for activator-tagged interacting proteins (see Chapter 4 for a description of the yeast two-hybrid system). A cDNA clone which interacted with Hairy was isolated and was found to encode a carboxyl-terminal portion of the Groucho (Gro) protein, corresponding to amino acids 251 to 719. Subsequent analysis demonstrated that the full length Gro protein interacted with Hairy and also with a number of other Hairy-related proteins including E(spl)-mδ, -m3, -m5, -m7, -m8 and Deadpan (Dpn). The domains within the Hairy-related proteins required to mediate interaction with Gro were then mapped

by mutational analysis. Hairy derivatives with missense mutations in the bHLH domain retained interaction with Gro, as did Hairy and E(spl)-m7 derivatives lacking the bHLH domain entirely, suggesting that the interaction domain was independent of this region and must map elsewhere in the protein. In contrast to changes in the bHLH domain, precise removal of the WRPW motif from both the Hairy and E(spl)-m7 proteins abolished interaction, revealing that the functional role of the WRPW motif in these proteins may involve mediating interaction with the Gro protein.

The authors were also able to demonstrate genetic interactions between *hairy* and *gro* during segmentation by showing that reductions in maternal *gro* dosage enhanced the penetrance of the segmentation phenotype produced by weak *hairy* alleles. In addition, maternal *gro* was shown to be required for segmentation, neurogenesis and sex determination, three developmental processes regulated by Hairy, E(spl) and Dpn respectively. Collectively, these data provide evidence that Hairy-related proteins and Gro form biologically active complexes. Furthermore, the mutational analysis suggests that the WRPW motif of these proteins is the domain with which interactions with Gro are mediated. Based on these observations Paroush *et al.* (1994) suggested that the Gro protein may act as a co-repressor which is recruited to target gene promoters via interaction with WRPW of the DNA-bound Hairy-related proteins.

Corroborative evidence for this hypothesis came from a series of experiments conducted by Fisher *et al.* (1996). In the first experiment, a repression assay was performed in *Drosophila* cell culture to determine whether a number of the Hairy-related proteins and deletion variants were able to repress a reporter gene construct. It should be noted that, the cells used in this experiment, *Drosophila* Schneider cells, express *gro* endogenously. In the assay hybrid proteins containing the full length Hairy and Hes-1 proteins fused to a heterologous DNA binding domain were able to efficiently repress reporter gene transcription. Equivalent levels of repression were shown by variants with amino terminal deletions, whereas truncated proteins with carboxyl-terminal deletions including the WRPW motif, showed no or markedly reduced levels of repression. Remarkably, the WRPW motif alone fused to a

heterologous DNA binding domain demonstrated repressor activity nearly equivalent to that of full length Hairy. Equivalent results were also obtained in a similar assay conducted in human HeLa cells, demonstrating that the WRPW motif also functions as a transcriptional repression domain in mammalian cells. These results indicated that the WRPW motif was necessary and sufficient for transcriptional repression in both insect and mammalian cells.

In agreement with Paroush *et al.* (1994), Fisher *et al.* (1996) also mapped the Gro interaction domain of the Hairy-related proteins to the WRPW motif. Only those moieties with an intact WRPW motif were able to interact with Gro. Once again a remarkable result was obtained in that the fusion protein containing just the WRPW motif was able to mediate wild-type levels of interaction with Gro. Additionally, the WRPW motif alone was shown to be able to interact with the Gro mammalian homologues the TLE proteins. These results showed that WRPW was necessary (as previously shown by Paroush *et al.* 1994) and sufficient to interact with *Drosophila* Gro and the mammalian TLE proteins in yeast and *in vitro*. The results added additional support to the hypothesis proposed by Paroush *et al.* that WRPW is the repression domain of the Hairy-related proteins and that repression is mediated by recruitment of the Gro protein.

In a third study, Giebel and Campos-Ortega (1997) further demonstrate a requirement for the WRPW motif for E(spl)-m8 function by performing an *in vivo* ectopic expression assay in the developing fly. In this experiment, two mutant E(spl)-m8 proteins, one lacking the WRPW motif entirely and a second with a proline to leucine substitution (corresponding to the defect found in the null allele of the *hairy* gene) were both unable to suppress development of SOP cells when ectopically expressed *in vivo*. This experiment demonstrates, in association with the data from the *hairy* alleles (Wainwright and Ish-Horowicz, 1992), that the WRPW motif of Hairy-related proteins is required *in vivo* for normal function.

In this study, single amino acid substitutions within the WRPW motif of one of the E(spl) proteins, E(spl)-m8, are made and the effects of these mutations are observed *in vivo* in flies

(this Chapter) and in the yeast two-hybrid system (Chapter 4). While the above studies show that complete removal of the motif abolishes function of the protein, this study observes the effects of more subtle alterations within the motif, since WRPW variants with single amino acid substitutions may still retain function. It was hoped that this would provide insight into the flexibility of the WRPW peptide sequence and would indicate which amino acids are invariantly required.

3.2. Manipulation of gene expression in *Drosophila*.

In this chapter a number of WRPW mutation derivatives are tested for function *in vivo* in the developing fly. The assay is based on a comparison between the phenotype generated by ectopic expression of the wild-type E(spl)-m8 protein and that generated by ectopic expression of the mutant derivative proteins (the assay is described in section 3.3). The assay therefore relies on the ability to drive ectopic expression from a transgene during development. A number of methods have been employed to manipulate gene expression in *Drosophila*. The first is to drive expression of a gene from a heat shock promoter which has the advantage of permitting temporal control of induced expression (for example, Ish-Horowicz and Pinchin, 1987). Several disadvantages are that expression is ubiquitous, basal levels of expression are sometimes observed, and the heat shock itself can induce phenocopies. A second method is to drive gene expression from a defined tissue specific promoter (for example, Parkhurst *et al.*, 1990). This provides spatial and temporal control of expression but is limited by the availability of previously cloned and characterised promoters that direct expression in the desired pattern. Additionally, if the expression of the target gene is deleterious to the fly it is difficult or impossible to generate stable transformant lines using this method.

A third method, the UAS-P[GAL4] system (Brand and Perrimon, 1993) permits temporal and spatial control of transgene expression and is not hindered by any of the problems associated with the methods described above. For these reasons the UAS-P[GAL4] system is used as a means of targeted gene expression in this study. The GAL4 protein is a yeast

transcriptional activator which can drive expression of target genes from promoters containing GAL4 binding sites (GAL4 Upstream Activation Sequences or UAS). In this system the transcriptional activator, the GAL4 protein, and the target gene (under UAS control) are separated in two distinct transgenic fly lines; in one line the activator protein is present but without its target and in the other line the target gene is silent without the activator. Expression of the target gene is achieved by crossing the two lines, bringing the GAL4 activator and the UAS target gene together in the same cells. Because the UAS sites can be placed upstream of any coding sequence, expression of any cloned gene can be performed using this system. A library of P[GAL4] expressing lines was generated by adaptation of the enhancer trap technique (O'Kane and Gehring, 1987): a P-element construct containing the GAL4 coding sequence fused to a promoter able to respond to neighbouring transcriptional regulatory elements was randomly integrated into the genome. This generated a large number of fly lines expressing the GAL4 protein in a wide range of patterns in embryos, larvae and adults, depending on the genomic site of integration. Visualisation of GAL4 expression patterns in these lines was observed by crossing each line to a UAS-*lac*Z line and staining for β-galactosidase activity. The UAS-P[GAL4] system has the advantage that it allows spatial and temporal control of transgene expression by careful choice of a fly line from the P[GAL4] expressing library.

3.3. The ectopic expression assay.

It has previously been shown that ectopic expression of two wild-type *E(spl)* genes, *E(spl)-m5* and *-m8*, using the UAS-P[GAL4] system, produces a mutant phenotype (Tata and Hartley, 1995). The phenotype caused by ectopic expression is dependent upon the developmental stage at which ectopic expression is induced and reflects the dual function of *E(spl)* in cell fate specification. Ectopic expression in third instar imaginal discs results in a bristle-loss phenotype resulting from the suppression of sensory organ precursor (SOP) cell formation. Ectopic expression later in development, during early pupal development (8-24h after puparium formation), result in a variable aberrant bristle phenotype which reflect perturbations in accessory cell determination and differentiation. These phenotypes include:

double sockets with no shaft, split bristle shafts, multiple shafts in single sockets and stunted or deformed bristle shafts.

In this study, the bristle-loss phenotype caused by ectopic expression of the wild-type *E(spl)-m8* gene in third instar wing discs is used as a point of comparison for the phenotype produced by ectopic expression of the *E(spl)-m8* WRPW mutant derivative genes. In order to test the neurogenic function of the WRPW mutant derivatives *in vivo,* the mutants proteins are expressed ectopically in various patterns in the larval wing disc and assayed for their ability to produce the bristle-loss phenotype associated with expression of the wild-type E(spl)-m8 protein. It was reasoned that, if the protein was still functional despite possessing a mutation in the WRPW motif then ectopic expression would produce a bristle-loss phenotype identical to that resulting from ectopic expression of the wild-type protein. An intermediate phenotype would indicate partial function and an inability to produce the bristle-loss phenotype would indicate that protein function had been completely abolished.

RESULTS

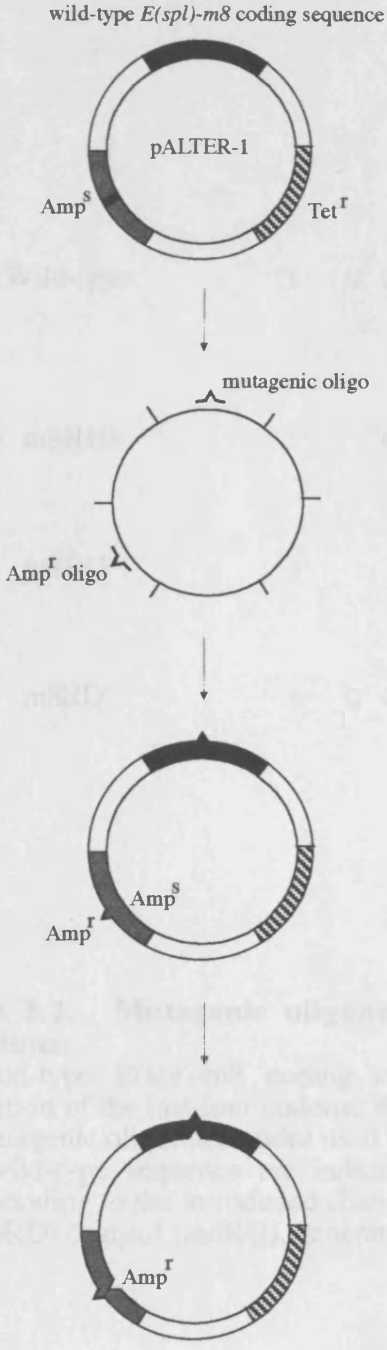
3.4. Generating the WRPW mutations.

Within the WRPW motif the arginine residue was selected as the amino acid which would be mutated. There was no *a priori* reason for this choice, or any evidence to suggest that arginine was more important than the other residues for WRPW function. Three substitutions were made at this position: arginine to histidine; arginine to glycine, and; arginine to aspartic acid. These changes represent an increasing difference from the original residue with respect to charge; Arginine (R) and histidine (H) have positively charged side chains (representing a conservative change), glycine (G) has an uncharged side chain, and the side chain of aspartic acid (D) is negatively charged.

In order to generate the WRPW derivatives, site directed mutagenesis approach was used to introduce base changes in the arginine codon of the wild-type coding sequence. Site directed mutagenesis technology involves hybridisation of a mismatched oligonucleotide to single stranded DNA (ssDNA) of the target sequence followed by second strand synthesis. All mutations made in this study were performed using the Promega Altered Sites II *in vitro* mutagenesis system. In this system the target sequence is cloned into the vector pALTER-1, a vector which also contains an inactivated gene for ampicillin resistance, to generate the template molecule used to create the mutation. In the mutagenesis reaction, an ampicillin repair oligonucleotide and a mutagenic oligonucleotide anneal to the same strand of the ssDNA template. Subsequent synthesis and ligation of the mutant strand links the two regions resulting in restoration of ampicillin resistance and introduction of the desired mutation (figure 3.1).

An *Eco*RI-*Xba*I fragment of the wild-type *E(spl)-m8* coding sequence was cloned into the *Eco*RI-*Xba*I sites of the pALTER-1 vector, to generate pALTER-m8. Three mutagenic oligonucleotides were designed from the wild-type *E(spl)-m8* sequence containing base substitutions in the arginine codon corresponding to each of the mutations (figure 3.2). Novel restriction sites were also incorporated into the oligonucleotides to generate markers to facilitate screening of potential mutants. In the case of m8RG the single base change in the arginine codon introduces a novel *Apa*I site. In the m8RHb and m8RD oligonucleotides additional silent base changes were introduced into the proline codon to generate new *Nco*I sites.

For each mutant, denatured pALTER-m8 was co-annealed with one of the mutant oligonucleotides and the ampicillin resistance repair oligonucleotide. Synthesis and ligation of the mutant strand linked the two oligonucleotides. The mutants were then isolated by the selection for the restoration of ampicillin resistance. Plasmid DNA from ampicillin resistant colonies was checked by cutting with the introduced novel restriction site. Positive clones were then sequenced to confirm the presence of the mutation (figure 3.3).



1. Wild-type *E(spl)-m8* coding sequence cloned into pALTER-1.

2. Denature pALTER-m8 to give ssDNA.

3. Co-anneal mutagenic oligo and ampicillin resistance repair oligo to ssDNA.

4. Synthesize and ligate the mutant strand.

5. Transform into repair minus strain of *E.coli* (*mutS*). Prepare miniprep DNA.

6. Transform into DH5α. Select for mutation by selecting for ampicillin resistance.

7. Screen for mutations by restriction analysis and direct sequencing.

Figure 3.1. Site directed mutagenesis.
The figure shows the site directed mutagenesis stratergy used to generate the *E(spl)-m8* WRPW mutant derivatives investigated in this study.

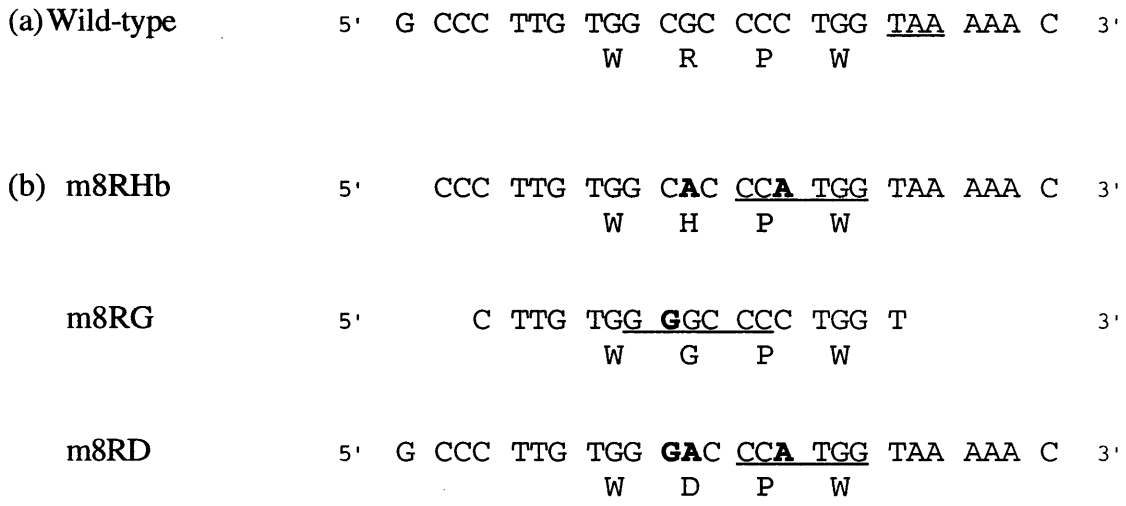


Figure 3.2. Mutagenic oligonucleotides used in the generation of the WRPW mutations.
(a) Wild-type E(spl)-m8 coding strand encoding the carboxyl-terminus of the protein. Translation of the last four codons, WRPW, is given below. The stop codon is underlined. (b) The mutagenic oligonucleotides used to generate the WRPW mutations. Nucleotide differences from wild-type sequence are indicated in bold type and the altered amino acid sequence corresponding to the introduced changes is given below. Novel restriction sites, *Nco*I (m8RHb and m8RD) or *Apa*I (m8RG), generated as a result of the nucleotide changes are underlined.

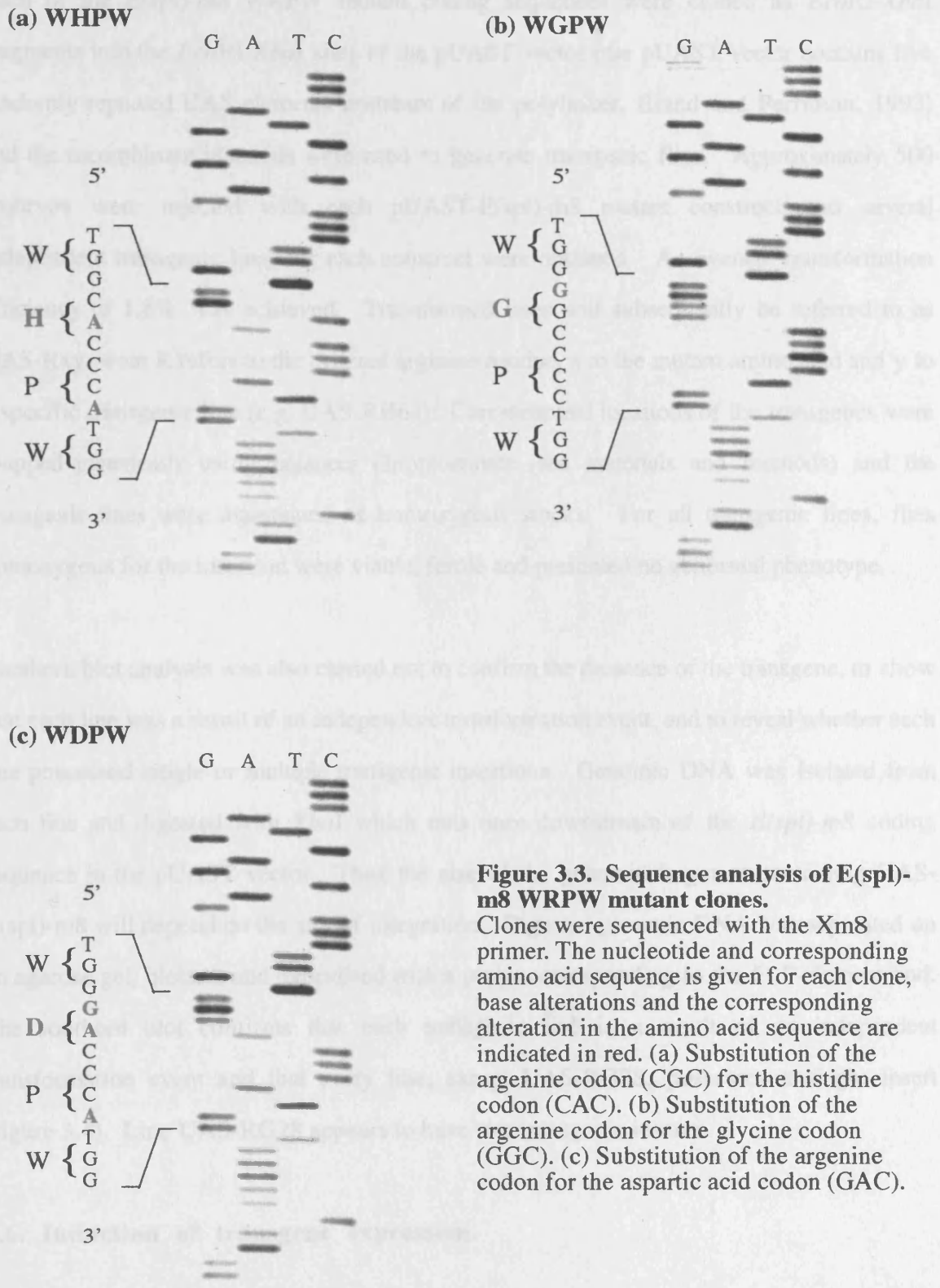


Figure 3.3. Sequence analysis of E(spl)-m8 WRPW mutant clones.
Clones were sequenced with the Xm8 primer. The nucleotide and corresponding amino acid sequence is given for each clone, base alterations and the corresponding alteration in the amino acid sequence are indicated in red. (a) Substitution of the argenine codon (CGC) for the histidine codon (CAC). (b) Substitution of the argenine codon for the glycine codon (GGC). (c) Substitution of the argenine codon for the aspartic acid codon (GAC).

3.5. Generating the transgenic flies.

Each of the *E(spl)-m8* WRPW mutant coding sequences were cloned as *Eco*RI-*Xba*I fragments into the *Eco*RI-*Xba*I sites of the pUAST vector (the pUAST vector contains five tandomly repeated UAS elements upstream of the polylinker, Brand and Perrimon, 1993) and the recombinant plasmids were used to generate transgenic flies. Approximately 500 embryos were injected with each pUAST-E(spl)-m8 mutant construct and several independent transgenic lines for each construct were obtained. An average transformation efficiency of 1.8% was achieved. Transformed lines will subsequently be referred to as UAS-Rxy, were R refers to the original arginine residue, x to the mutant amino acid and y to a specific transgenic line (e.g. UAS-RH64). Chromosomal locations of the transgenes were mapped genetically using balancer chromosomes (see materials and methods) and the transgenic lines were maintained as homozygous stocks. For all transgenic lines, flies homozygous for the insertion were viable, fertile and presented no abnormal phenotype.

Southern blot analysis was also carried out to confirm the presence of the transgene, to show that each line was a result of an independent transformation event, and to reveal whether each line possessed single or multiple transgenic insertions. Genomic DNA was isolated from each line and digested with *Xba*I which cuts once downstream of the *E(spl)-m8* coding sequence in the pUAST vector. Thus the size of the genomic fragment containing UAS-E(spl)-m8 will depend on the site of integration. Digested genomic DNA was separated on an agarose gel, blotted, and hybridised with a probe corresponding to the 5' P-element end. The southern blot confirms that each transgenic line is a result of an independent transformation event and that every line, except UAS-RG28, possesses a single insert (figure 3.4). Line UAS-RG28 appears to have two transgenic inserts.

3.6. Induction of transgene expression.

Expression of the UAS-transgene is only activated by crossing the transgenic lines to fly strains which express the GAL4 protein. Four such P[GAL4] expressing fly strains or 'driver lines' are used in this study, these are referred to as 69B, T80, ptc-559.1 and c591.

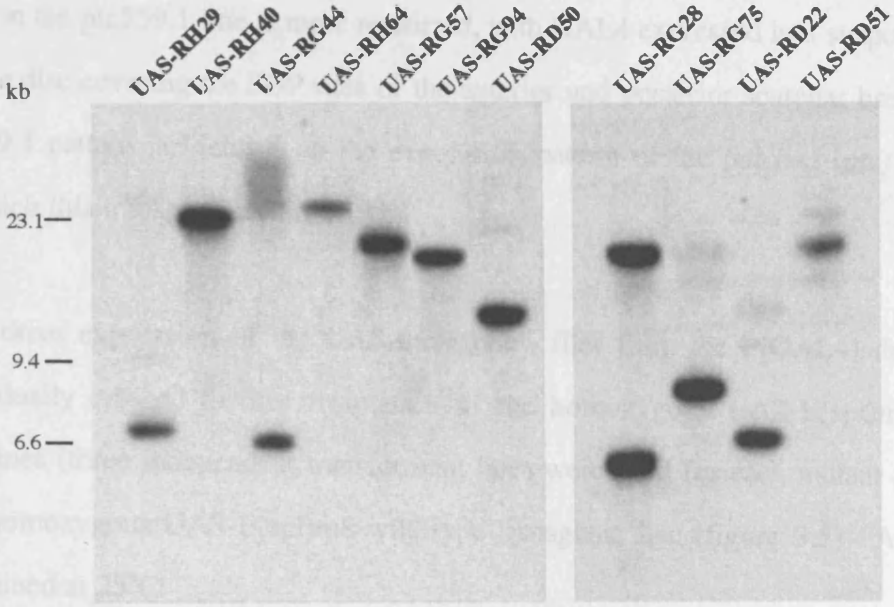


Figure 3.4. Southern blot analysis of genomic DNA from transgenic fly lines.
Genomic DNA from homozygous flies was digested with *Xba*I, probed with a 5' P-element probe, followed by high stringency washing (0.1% SSPE, 0.1% SDS, 65°C 20 minutes).

The GAL4 expression patterns of these lines are given in table 3.1. These patterns are the descriptions which accompany the flies in the Bloomington stock centre and represent known GAL4 expression domains. It is possible that expression occurs in other tissues and at other times of development. These lines were chosen because the spatial and temporal profile of GAL4 expression concurs with the formation of SOP cells. The lines 69B, T80 and c591 have ubiquitous or near ubiquitous GAL4 expression patterns in third instar wing discs and therefore expression covers all or most sites of the bristle SOPs. The pattern of expression in the ptc559.1 line is more restricted, with GAL4 expressed in a stripe down the centre of the disc covering the SOP sites of the anterior and posterior scutellar bristles only. The ptc-559.1 pattern is identical to the expression pattern of the *patched* (*ptc*) gene, the locus of which this transgene has inserted.

In order to drive expression of the UAS-transgenes, flies from the P[GAL4] driver lines were individually crossed to flies from each of the homozygous UAS-E(spl)m8 mutant transgenic lines (three independent transformant lines were used for each mutant construct) and to the homozygous UAS-E(spl)m8 wild-type transgenic line (figure 3.5) All crosses were maintained at 25°C.

In order to ensure the transgenes were being expressed in the correct spatial patterns, at the correct time of development and at sufficiently high levels, *in situ* hybridisation was performed on third instar wing discs from the F_1 progeny of crosses shown in figure 3.5. For the probe, wild-type *E(spl)-m8* in pBluescript II SK$^-$ was linearised with *Xba*I to generate the template and a DIG-labelled RNA probe was synthesised *in vitro* using T7 polymerase. Wing imaginal discs were dissected from third instar F_1 progeny from the cross shown in figure 3.5b and hybridised overnight with the *E(spl)-m8* DIG-labelled RNA probe. It should be noted that 50% of the dissected discs derive from individuals containing the *CyO* balancer chromosome and not the P[GAL4] chromosome; in these discs transgene expression is not induced and the expression of the endogenous wild-type gene can be seen. These discs served as an internal control for the specificity of the probe.

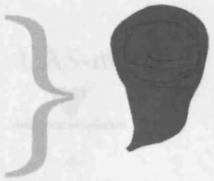
p[GAL4] driver line	GAL4 protein expression pattern	Wing disc expression pattern
c591	Ubiquitous in wing disc	
T80	Ubiquitous in third instar imaginal discs	
69B	Embryonic epidermis and imaginal discs	
ptc-559.1	Identical to the patched gene - stripe down the centre of the wing and eye disc	

Table 3.1. p[GAL4] driver lines used in the study.
List of the p[GAL4] driver lines used in this study and their expression patterns as stated in the Bloomington stock centre. Pattern of wing disc expression is also given (right).

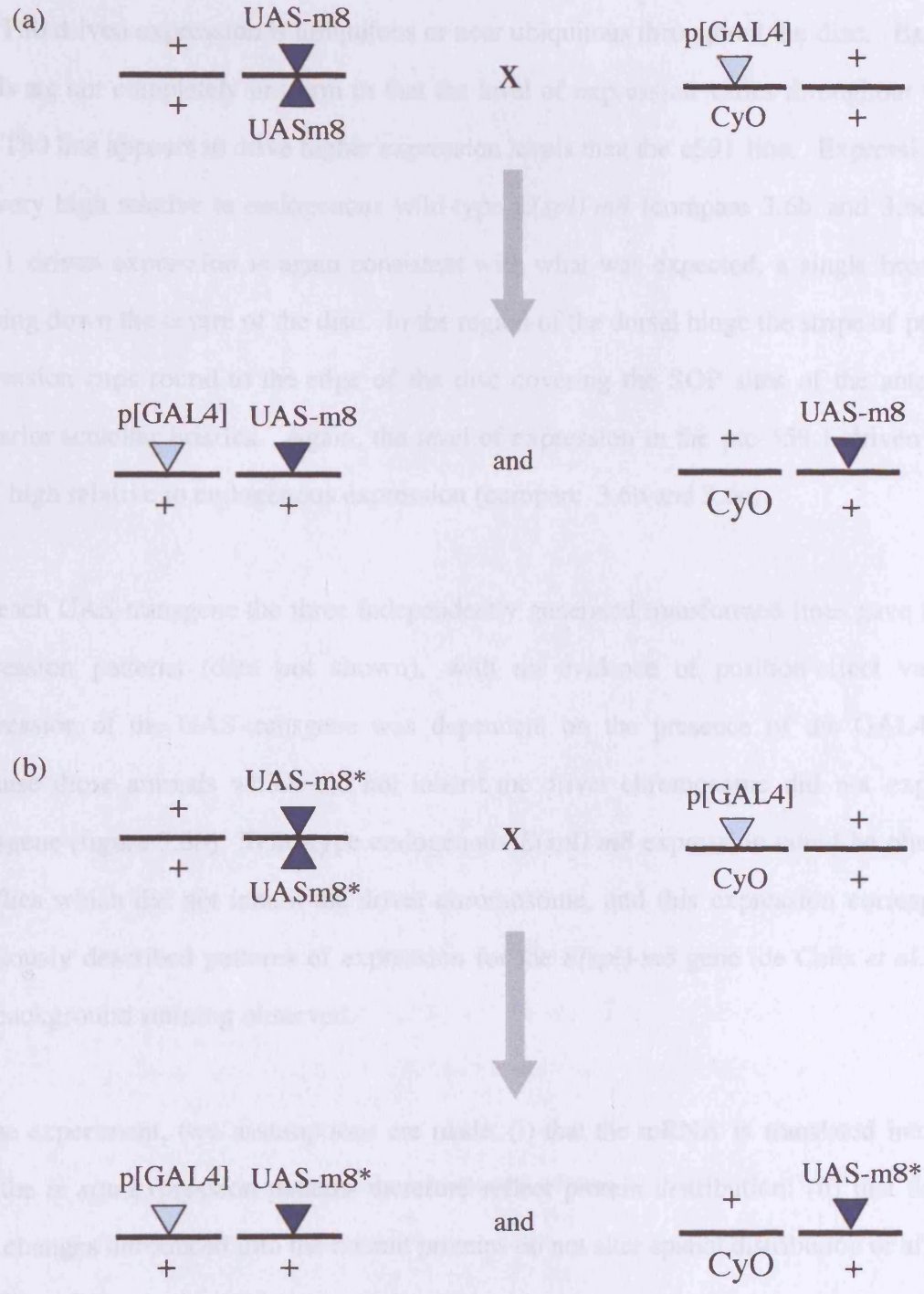


Figure 3.5. Genetic cross of UAS-responder lines to p[GAL4] driver lines. (a) Fly lines homozygous for the wild-type UAS-E(spl)-m8 transgene or, (b) the UAS-E(spl)-m8 mutant derivative transgenes (UAS-m8*) were crossed to balanced p[GAL4] driver lines. All crosses were maintained at 25°C.

The results of the *in situ* hybridisation are shown in figure 3.6. All discs were stained for an equal amount of time (approximately 40 minutes), allowing for direct visual comparison. The UAS-transgenes were expressed in the anticipated spatial and temporal patterns. c591 and T80 driven expression is ubiquitous or near ubiquitous throughout the disc. Expression levels are not completely uniform in that the level of expression varies throughout the disc. The T80 line appears to drive higher expression levels than the c591 line. Expression levels are very high relative to endogenous wild-type *E(spl)-m8* (compare 3.6b and 3.6d). ptc-559.1 driven expression is again consistent with what was expected, a single broad stripe running down the centre of the disc. In the region of the dorsal hinge the stripe of ptc-driven expression cups round to the edge of the disc covering the SOP sites of the anterior and posterior scutellar bristles. Again, the level of expression in the ptc-559.1 driven discs is very high relative to endogenous expression (compare 3.6b and 3.6c).

For each UAS-transgene the three independently generated transformed lines gave the same expression patterns (data not shown), with no evidence of position-effect variability. Expression of the UAS-transgene was dependent on the presence of the GAL4 protein because those animals which did not inherit the driver chromosome did not express the transgene (figure 3.6b). Wild-type endogenous *E(spl)-m8* expression could be observed in the flies which did not inherit the driver chromosome, and this expression corresponds to previously described patterns of expression for the *E(spl)-m8* gene (de Celis *et al.*, 1996). No background staining observed.

In the experiment, two assumptions are made: (i) that the mRNA is translated into protein and the *in situ* expression patterns therefore reflect protein distribution; (ii) that the amino acid changes introduced into the mutant proteins do not alter spatial distribution or affect their stability relative to the wild-type protein.

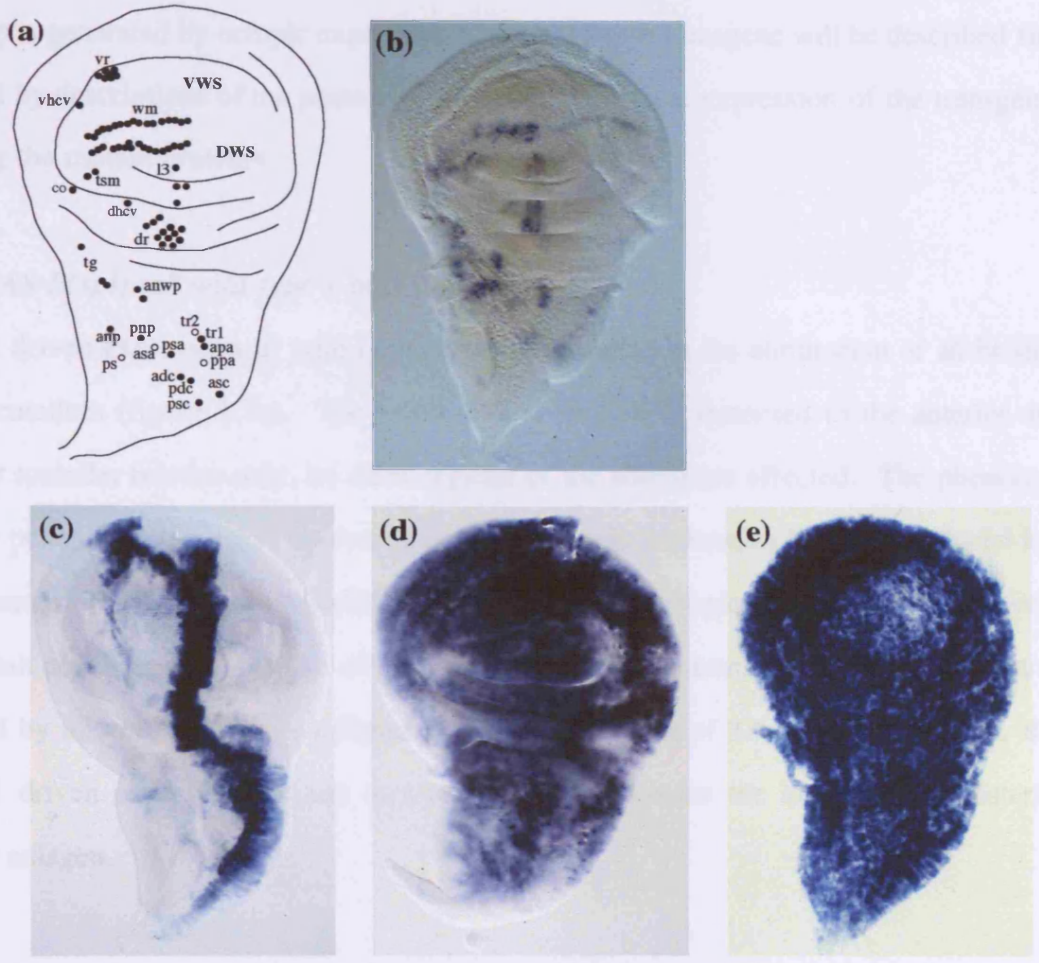


Figure 3.6. RNA expression pattern of E(spl)-m8 mutant derivatives in the wing disc.
(a) Schematic diagram of the third instar wing imaginal disc showing sensorgan fate map (refer to figure 1.2 for full description of abbreviations). (b, c, d,e) Third instar wing imaginal discs hybridised with a wild-type E(spl)-m8 DIG-labelled RNA probe. (b) Wild-type disc showing endogenous E(spl)-m8 expression. (c) Disc showing ectopic expression of E(spl)-m8RH64 driven by the ptc-559.1 driver line. (d) Disc showing ectopic expression of E(spl)-m8RH64 driven by the c591 driver line. (e) Disc showing ectopic expression of E(spl)-m8RH64 driven by the T80 driver line.

3.7. Functional analysis of *E(spl)-m8* mutant derivatives.

The crosses shown in figure 3.5 were performed and the results from these crosses are described below. In each case 50-100 adult progeny from each cross were examined. The phenotypes generated by ectopic expression of the wild-type transgene will be described first followed by descriptions of the phenotype produced by ectopic expression of the transgenes encoding the mutant proteins.

3.7.1. UAS-E(spl)-m8 wild-type x ptc559.1

ptc559.1 driven expression of wild-type *E(spl)-m8* resulted in the elimination of all bristles on the scutellum (figure 3.7a). The bristle-loss phenotype is restricted to the anterior and posterior scutellar bristles only, no other regions of the notum are affected. The phenotype is highly penetrant since every fly scored in which ectopic expression had been induced had missing anterior and posterior scutellar bristles (n = 87). The region of bristle-loss observed in the adult corresponds to regions of ectopic expression of the transgene in the wing disc as observed by *in situ* hybridisation (figure 3.6c). Comparison of 3.6a and 3.6b shows that ptc559.1 driven expression occurs in a region which includes the anterior and posterior scutellar anlagen.

3.7.2. UAS-E(spl)-m8 wild-type x c591

c591 driven expression of wild-type *E(spl)-m8* produced an extensive bristle loss phenotype (n = 61). The exact pattern of bristle loss was variable. In the most extreme case (figure 3.7b) most of the macrochaetae and microchaetae were absent, although less penetrant phenotypes were also observed which were typified by a reduction in the number of macrochaetae coupled with a less extensive loss of microchaetae (data not shown). The bristles most sensitive to c591 driven expression were the anterior and posterior scutellar bristles as these bristles were always absent after ectopic expression of the wild-type transgene. The pattern of bristle-loss observed in the adult fly corresponds to the pattern of ectopic expression of the transgene in the wing disc as observed by *in situ* hybridisation (figure 3.6d). c591 drives expression over a large proportion of the wing disc, including

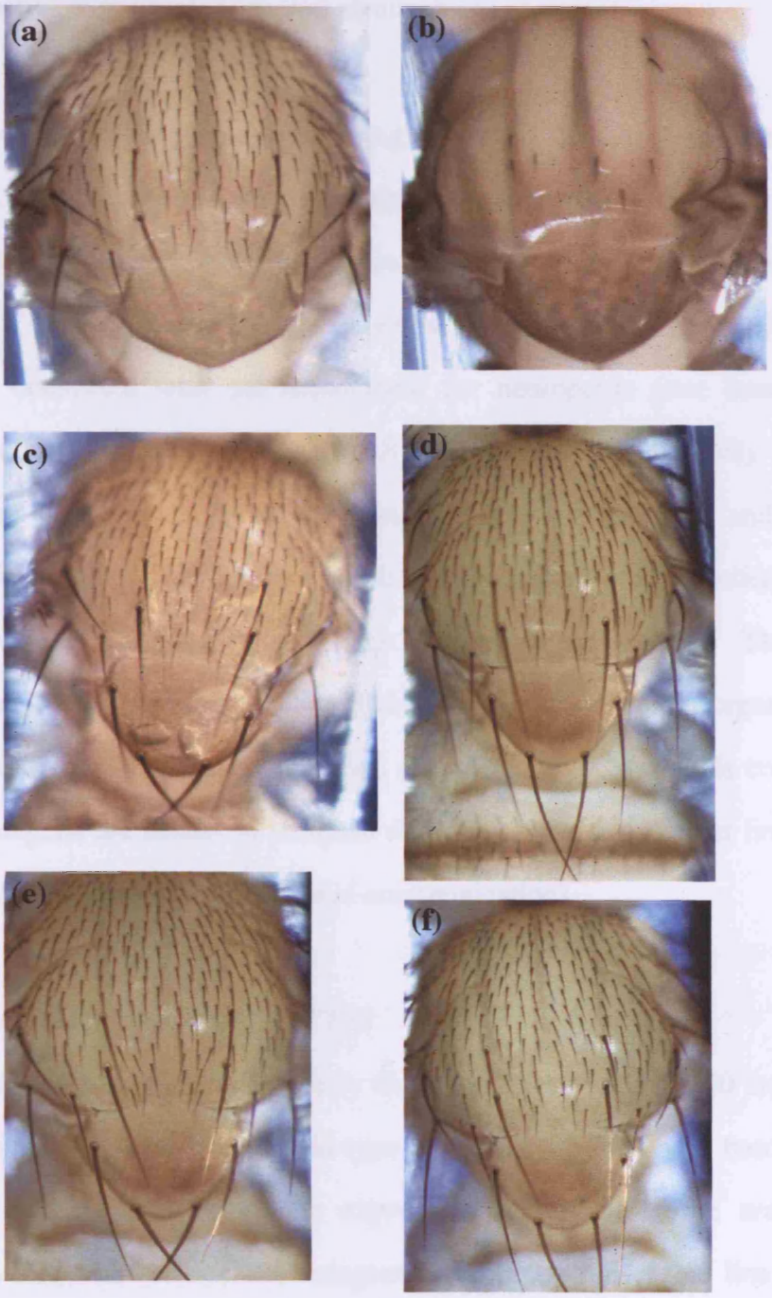


Figure 3.7. Ectopic expression of E(spl)-m8 and E(spl)-m8 WRPW derivatives.
Nota of adult flies showing the result of ectopic expression of either wild-type UAS-E(spl)-m8 or UAS-E(spl)-m8RH64 with a variety of p[GAL4] driver lines: (a) wild-typeUAS-E(spl)-m8 driven with ptc-559.1, (b) wild-typeUAS-E(spl)-m8 driven with c591, (c) UAS-E(spl)-m8RH64 driven with ptc-559.1, (d) UAS-E(spl)-m8RH64 driven with c591, (e) UAS-E(spl)-m8RH64 driven with T80, (f) UAS-E(spl)-m8RH64 driven with 69B.

regions corresponding to the anlagen of several of the bristles. The reason for the variability of the bristle-loss phenotype is difficult to explain because *in situ* hybridisation patterns observed in a number of individuals appeared identical.

Ectopic expression of the wild-type transgene using the c591 driver line produced two other mutant phenotypes, (i) disruptions in wing veination and (ii) abnormal locomotor behaviour of the flies. These two additional phenotypes were present in every individual. In the wings the distal portions of the fourth and fifth wing veins were gapped or absent (figure 3.8). This phenotype is consistent with the requirement for neurogenic gene function in the selection of vein versus non-vein cells (Garcia-Bellido and de Celis, 1992). Abnormal locomotor behaviour was manifest by a staggering movement of the flies and inability to walk in a straight line. This inability to co-ordinate movement often resulted in a large proportion of individuals perishing after falling into their food source. The abnormal locomotor behaviour is likely to be a result of disturbances in chordotonal organ formation. In support of this, in the A101 enhancer trap line, *lacZ* staining of SOP cells corresponding to the chordotonal organs are absent in imaginal discs when the c591 driver line is used to drive wild-type UAS-E(spl)-m8 (Tata, personal communication).

3.7.3. UAS-E(spl)-m8 wild-type x 69B and T80

Wild-type UAS-E(spl)-m8 expression driven by the driver lines 69B and T80 are embryonic lethal. After ectopic expression of the wild-type transgene no embryos hatch. This is probably a result of extensive embryonic expression which, in some way disrupts, embryogenesis. In line with this the Bloomington stock centre report that line 69B drives expression in the embryonic epidermis (table 3.1).

3.7.4. UAS-E(spl)-m8 WRPW mutant derivatives x ptc-559.1 and c591

ptc-559.1 and c591 driven ectopic expression of the UAS-m8RH, UAS-m8RG and UAS-m8RD transgenes had no phenotypic effect in that ectopic expression of each of the mutant transgenes produced flies which were phenotypically wild-type (n = 612 for ptc-559.1 and n = 568 for c591, collected data from each of a number of independent transformant lines for

Figure 3.8. Ectopic expression of wild-type E(spl)-m8 results in a decrease in vein material.
Ectopic expression of wild-type E(spl)-m8 using the c591 driver line, which drives expression in the wing pouch of the wing imaginal disc, results in a reduction of wing vein material. This phenotype is manifest by either the complete absence or gapping of the distal portion of the IV and V wing veins.

each mutant transgene) (figure 3.7c and d). This result is in contrast to the bristle-loss phenotype produced by ectopic expression of the wild-type transgene. In addition, the wing veination and behavioural mutant phenotypes which resulted from ectopic expression of the wild-type transgene using the c591 driver line were not observed after expression of the three *E(spl)-m8* mutant derivatives.

3.7.5. UAS-E(spl)-m8 WRPW mutant derivatives x 69B and T80

69B and T80 driven expression of the UAS-m8RH, UAS-m8RG and UAS-m8RD transgenes resulted in viable progeny that were phenotypically wild-type (n = 438 for 69B and n = 450 for T80, collected data from each of a number of independent transformant lines for each mutant transgene) (figure 3.7e and f). This contrasts with the lethality associated with T80 and 69B driven ectopic expression of the wild-type *E(spl)-m8* transgene. This indicates that the activity of the wild-type transgene which caused embryonic lethality is not present in each of the three mutant transgenes.

In summary, the phenotype produced by ectopic expression of the mutant WRPW E(spl)-m8 proteins deviates from the phenotype produced by ectopic expression of the wild-type E(spl)-m8 protein. Ectopic expression of the wild-type transgene results in either a bristle-loss phenotype or lethality, depending upon expressivity. In the case of the WRPW mutant transgenes neither of these phenotypes were observed. Since ectopic expression of the mutant proteins has no phenotypic effect it suggests that the neurogenic activity of these proteins has been either severely reduced or completely abolished.

3.8. Dominant negative effects of the mutant derivative proteins.

The functional analysis described above was conducted using flies with one copy of the UAS-transgene and one copy of the P[GAL4] driver chromosome. As mentioned above, the conclusion drawn from these assays is that the proteins mutant for WRPW were either devoid of activity or their activity had been severely reduced. To determine whether the mutant proteins possessed residual activity, the level of ectopic expression was elevated to

increase the amount of protein in the cells. It was reasoned that if enough protein was present it might partially restore the ability to induce the bristle-loss phenotype. In order to increase the amount of the protein, fly lines were constructed that were homozygous for both the UAS-transgene and the P[GAL4] driver chromosome. Three lines were generated: one homozygous for the ptc-559.1 driver construct and the UAS-E(spl)-m8RD51 responder construct, one homozygous for the c591 driver construct and the UAS-E(spl)-m8RD51 responder construct, and one homozygous for the c591 driver construct and the UAS-E(spl)-m8RH64 responder construct. It was assumed that these homozygous lines had approximately double the amount of the mutant protein compared to lines with just one UAS-responder and one P[GAL4] driver chromosome.

Analysis of the double homozygous lines showed that the majority of flies were wild-type with respect to bristle phenotype, demonstrating that elevated levels of the ectopically expressed mutant proteins is unable to restore the bristle-loss phenotype and therefore suggesting that the WRPW mutations completely abolish the neurogenic function of the protein. Interestingly, approximately 20% of individuals from each line possess extra bristles on the notum. This phenotype was manifest by a single extra macrochaete at a position of an existing bristle or at a site relatively nearby (figure 3.9a). Bilateral bristle-gains were observed at a lower frequency (figure 3.9a). The morphology of the ectopic bristles was generally wild-type, however some bristles were of a size intermediate between that of a macrochaete and a microchaete. The bristle-gain phenotype did, in some instances, manifest as an increase in the density of the microchaetae lawn with extra rows of microcheatae appearing between the normal rows (figure 3.9b). The average number of notal microcheatae in these flies was 349 for females (n = 5) and 337 for males (n = 5), this contrasts with the wild-type number of microcheatae which corresponds to approximately 210 for females (n = 2) and 182 for males (n = 2). Additionally, it was noticed that individuals with extra bristles on the notum also possessed ectopic bristles on the wing blade and, occasionally, ectopic vein tissue (figure 3.9c and d). The ectopic vein tissue phenotype is the opposite phenotype to the vein gapping observed by ectopic expression of wild-type *E(spl)-m8* (figure 3.8). It is possible that the bristle-gain and the ectopic vein tissue

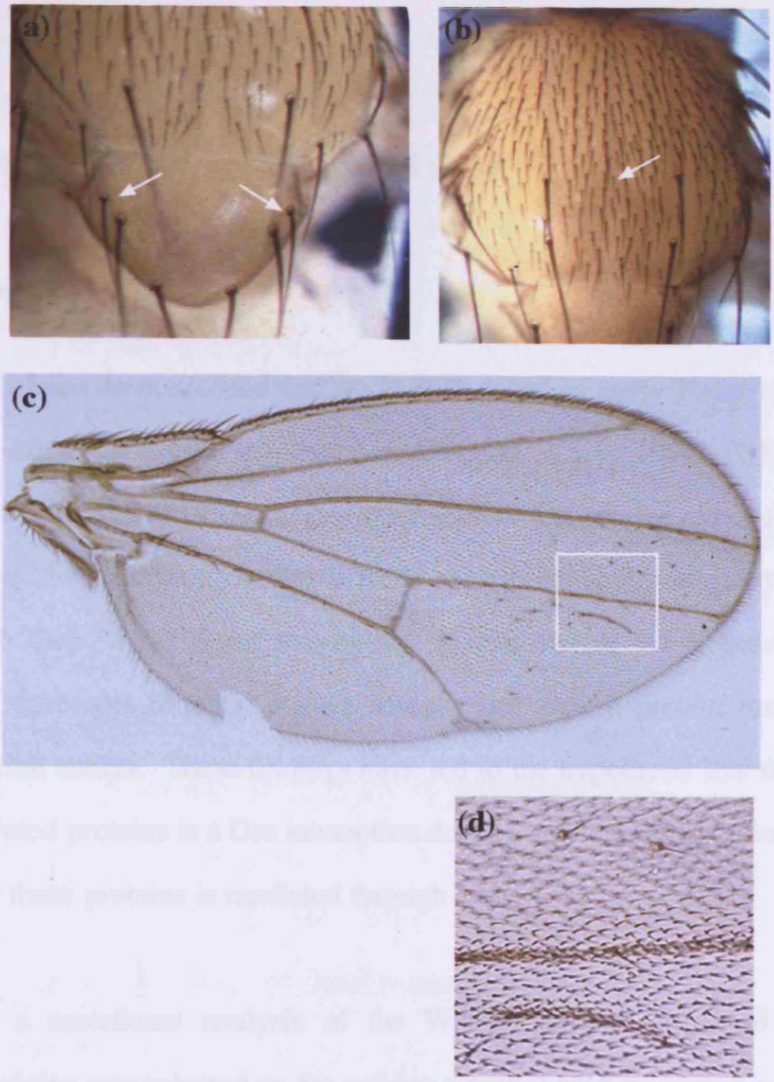


Figure 3.9. Ectopic expression of the E(spl)-m8 WRPW mutant derivatives result in a dominant negative phenotype.
Increasing the level of expression of the E(spl)-m8 WRPW mutant derivatives by generating lines homozygous for both the UAS-responder and the p[GAL4] driver constructs result in a number of dominant negative phenotypes. (a) adult notum of a fly homozygous for the UAS-m8RD51 responder and ptc-559.1 driver constructs shows a bilateral gain of macrochaetae on the scutellum at a position anterior to the anterior scutellar bristle (arrows). These additional bristles are smaller than the wild-type bristles on the scutellum. (b) adult notum of a fly homozygous for the UAS-m8RD51 responder and c591 driver constructs shows an increase in the density of the microchaetae lawn. Additional rows of small microchaete bristles are apparent between the normal rows. An additional row is found on the dorsal midline (arrow) which is not usually observed in the wild-type nota. (c) and (d) wing from a fly homozygous for the UAS-m8RD51 responder and c591 driver constructs demonstrating additional wing vein material and ectopic macrochaetae on the wing blade. (d) higher magnification of boxed region in (c).

phenotypes are dominant negative effects resulting from the interference of endogenous wild-type protein activity by the ectopically expressed mutant protein.

DISCUSSION

3.9. Discussion.

It has previously been demonstrated that the WRPW motif of many Hairy-related proteins is necessary and sufficient for protein function (Paroush *et al.*, 1994; Fisher *et al.*, 1996; Giebel and Campos-Ortega, 1997). These studies show that precise removal of WRPW and specific amino acid substitutions within WRPW abolish the ability of Hairy-related proteins to interact with Gro in the yeast two-hybrid system and *in vitro*, prevent the proteins functioning as repressors in reporter gene assays, and abolish protein function in *in vivo* ectopic expression assays. These findings have led to the hypothesis that the WRPW motif of the Hairy-related proteins is a Gro interaction domain and that transcriptional repression of target genes by these proteins is mediated through Gro.

In this study, a mutational analysis of the WRPW of the *E(spl)-m8* gene has been performed. Arginine was selected as the residue which would be altered and three separate mutations were generated at this position; arginine to histidine, arginine to glycine, and arginine to aspartic acid. The function of the mutant proteins was assayed *in vivo* in flies by comparison of the effects generated by ectopic expression of the mutant protein with the bristle-loss phenotype generated by ectopic expression of the wild-type E(spl)-m8 protein.

Ectopic expression of either the wild-type or mutant *E(spl)-m8* transgenes was induced in various spatial patterns at the pheno-critical period of SOP formation. Ectopic expression of the wild-type protein generated a bristle-loss phenotype similar to the phenotype obtained in a previous study (Tata and Hartley, 1995) and concordant with the proposed role of *E(spl)-m8* as a repressor of the neural fate. In contrast, however, ectopic expression of each of the

three mutant constructs (from a single copy of the transgene) did not result in a bristle-loss phenotype despite the presence of high levels of mutant transcript as determined by *in situ* hybridisation. The inability of the mutant proteins to induce the bristle-loss phenotype normally associated with ectopic expression of the wild-type protein suggests that the *in vivo* neurogenic function of the E(spl)-m8 protein is abolished when the arginine residue is substituted for either histidine, glycine or aspartic acid.

Previous studies highlighted the requirement for the WRPW motif by showing that truncated proteins lacking the whole motif were non-functional (Paroush *et al.,* 1994; Fisher *et al.*, 1996). The data presented here extends these findings by showing that more subtle alterations, such as amino acid substitutions, also render the protein non-functional with respect to the repression of the neural fate. The arginine to histidine amino acid substitution is a conservative change and yet the change abolishes the function of the protein, suggesting that this residue is critical for protein function. A more exhaustive mutation analysis would be required to determine whether there is any flexibility in the amino acid sequence of the motif as a whole, but the results presented in this study coupled with the specific WRPW mutations associated with various *hairy* alleles (Wainwright and Ish-Horowicz, 1992) indicate that any alterations within the motif would abolish the *in vivo* neurogenic function of the Hairy-related proteins.

If the level of the mutant protein is increased by driving expression from two copies of the transgene a novel phenotype is produced in approximately 20% of the flies. This phenotype is manifest in the presence of extra macrochaetae in the notum and wing or increases in the density of the microchaetae lawn on the notum and in the growth of extra vein tissue in the wing. These phenotypes probably represent weak dominant negative effects as they are the converse phenotype to that produced by ectopic expression of the wild-type *E(spl)-m8* transgene. Two other *E(spl)-m8* mutant derivatives which display weak dominant negative effects have been reported elsewhere (Welshons, 1956; Giebel and Campos-Ortega, 1997), one of these includes the original *E(spl)* mutatant allele, $E(spl)^{D}$ (Welshons, 1956). Interestingly these two mutant derivatives completely lack the WRPW motif. These data,

along with the evidence described here, suggest that removal or alteration of the WRPW motif of the E(spl) protein interfers with the function of the endogenous protein. There are two possible explanations for these dominant negative effects. Firstly, a mutant protein homodimer could bind the N-box sequences of *E(spl)-m8* target genes, this complex, which itself cannot repress target genes because it does not contain a functional WRPW motif and therefore cannot recruit Gro, hinders access to endogenous wild-type homodimers preventing them from repressing target genes. A second possibility is that optimal recruitment of Gro by E(spl)-m8 dimers may require two functional WRPW motifs; heterodimers containing one wild-type molecule and one mutant molecule, and thus possessing only one wild-type WRPW, may only be able to recruit Gro with reduced efficiency. It is probable that dimeric complexes between wild-type and mutant proteins would be more prevalent than complexes between two wild-type proteins because levels of expression of mutant protein are considerably higher than endogenous wild-type expression. The net effect of either of these two hypotheses would be a decrease in repressive efficiency of endogenous wild-type E(spl)-m8 proteins causing the observed bristle-gain and wing veination phenotypes.

In order to try to establish the molecular basis for the loss of function observed for the E(spl)-m8 mutant derivative proteins, an analysis of the interaction between these proteins and the Gro protein is performed in the yeast two-hybrid system and is described in the following chapter.

Chapter 4

Analysis of protein interactions between Groucho and the E(spl) mutant derivatives

INTRODUCTION

4.1. Molecular interactions between E(spl) and Groucho.

In the previous chapter a mutational analysis of E(spl)-m8 revealed that substitution of the arginine residue in the WRPW motif was sufficient to abolish the function of the protein when assayed *in vivo* in the developing fly. Given the extent of conservation of the motif in all members of the Hairy-related protein family it is perhaps not surprising that mutations within the motif result in disturbances in function. The data obtained in chapter 3 is consistent with the findings of Wainwright and Ish-Horowicz (1992) who showed that a number of mutations in the WRPW motif disrupt *in vivo* function of the structurally and functionally related Hairy protein. These data indicate that WRPW motif of Hairy and related proteins is required for the normal function of the proteins, but the studies do not however, determine the molecular basis for the requirement of the WRPW motif.

The first clues about the function of the WRPW motif came from a yeast two-hybrid library screen performed by Paroush *et al.* (1994) who used the Hairy protein as the bait to fish for activation domain-tagged interacting proteins. One of the interacting clones identified in this screen encoded a peptide corresponding to the C-terminal region of the Gro protein. Gro was subsequently shown to interact with a number of other Hairy-related proteins, including the E(spl) proteins and Deadpan. Further analysis revealed that the WRPW motif of these proteins was necessary to facilitate this interaction as variants which lacked this motif did not interact with Gro when assayed either in yeast or *in vitro*. A remarkable result was obtained in a subsequent study by Fisher *et al.* (1996) who demonstrated that the WRPW motif was not only necessary to facilitate interaction with Gro but that WRPW alone was sufficient for this interaction. They were also able to show that WRPW alone was sufficient to allow interaction with mammalian Gro homologs, the TLE proteins, providing evidence that the function of the domain, in terms of interaction with Gro, is evolutionary conserved. In line with these results Grbavec and Stifani (1996) demonstrated that interaction between human TLE1 and HES-1 was dependant upon the WRPW motif of HES-1. The combined results

of these studies provide very strong evidence that the WRPW motif of the Hairy-related proteins is the domain through which interactions with the Gro protein occur.

Fisher *et al.* (1996) were also able to demonstrate that WRPW was necessary and sufficient to mediate repression in a repression assay performed in cultured *Drosophila* cells. The combined data of this repression assay and the interaction assays described above supported the hypothesis initially proposed by Paroush *et al.* (1994) that the function of Hairy and related proteins is mediated, at least in part, by Gro. In the corepression hypothesis Paroush *et al.* (1994) propose that Hairy or the Hairy-related proteins bind DNA upstream of target genes and recruit the Gro protein via interaction with the WRPW motif, once Gro has been recruited to the DNA it then represses transcription of the target gene by a mechanism which is not yet fully understood (see section 1.10).

In light of the findings by Paroush *et al.* (1994) and Fisher *et al.* (1996) it is proposed here that the E(spl)-m8 WRPW derivatives described in the previous chapter are nonfunctional when assayed *in vivo* because they are defective in their ability to interact with the Gro protein. In other words, the mutant E(spl)-m8 protein is present in the proneural clusters of the wing imaginal disc cells at the time of cell fate commitment but as a consequence of an inability to interact with endogenous Gro it is unable to repress the target genes and is therefore rendered nonfunctional.

To investigate this hypothesis, interaction between each of the E(spl)-m8 mutant WRPW derivative proteins and Gro was assayed using the yeast two-hybrid system. The two-hybrid system allows a relative measure of the strength of interaction between two proteins to be determined and therefore provides an excellent tool with which to compare the strengths of interaction of wild-type E(spl)-m8 with Gro, on the one hand, to each of the E(spl)-m8 mutant derivatives with Gro on the other.

4.2. The yeast two-hybrid system, a historical perspective.

The yeast two-hybrid system is a methodology for analysing or detecting interactions between two proteins (Bartel and Fields, 1997). The advantage of studying protein-protein interactions using this system compared to *in vitro* analyses is that the assay is performed in a cellular environment, the yeast cell, therefore maximising the likelihood of maintaining the structural integrity of the proteins under study. The system can be used to investigate the interaction between two known proteins (an investigation of this kind is performed in this chapter) or used to screen a library to detect proteins that interact with a given protein.

The yeast two-hybrid system was conceived by Stanley Fields in 1987. The initial idea for the system was based on a number of discoveries about the modular domain structure of eukaryotic transcription factors which came to light around that time. It was found that many transcription factors had at least two distinct functional domains, a domain required for binding to specific DNA sequences and a separate domain required to activate or repress transcription (Keegan *et al.*, 1986; Hope and Struhl, 1986). The modular nature of transcription factors was first demonstrated by domain swap experiments carried out between unrelated transcription factors. Brent and Ptashne (1985) showed that a hybrid transcriptional activator could be generated by the fusion of the DNA-binding domain of the *Escherichia coli* LexA transcriptional repressor to the activation domain of the yeast GAL4 protein. The resulting hybrid protein could activate transcription of genes containing LexA binding sites in their promoters. Similarly, the fusion of a number of heterologous activation domains to the yeast GAL4 DNA binding domain could also activate transcription of genes containing GAL4-binding sites (Chasman *et al.*, 1989; Ma and Ptashne, 1987; Sadowski *et al.*, 1988), demonstrating that modularity was a general feature of at least a number of transcription factors. A crucial corollary to the modular nature of transcriptional activators which was pertinent to the development of the two-hybrid system was the discovery that the DNA-binding domain and the activation domain need not be covalently linked to each other for activation to occur (Ma and Ptashne, 1988; Triezenberg *et al.*, 1988). Ma and Ptashne (1988) elegantly demonstrated that covalent attachment was not required in an experiment

using the GAL80 protein, a repressor protein which interacts with GAL4 and lacks a naturally occurring activation domain. They showed that an artificial GAL80-AD fusion could restore, *in trans*, the ability of a mutant GAL4 protein lacking a functional activation domain to activate transcription.

The first trial of the system was performed by Fields and Song (1989) using two yeast proteins, Snf1 and Snf4, which had previously been shown to interact *in vitro*. Snf1 was expressed as a fusion to the GAL4 DNA-binding domain and Snf4 was expressed as a fusion to the GAL4 activation domain. Interaction between the two proteins was demonstrated by activation of a reporter gene bearing GAL4 binding sites, thereby demonstrating the validity of the two-hybrid system.

The success of this experiment led Fields and Song to suggest that the yeast two-hybrid system could be used to clone cDNAs encoding proteins that interact with a given known protein. To perform such an experiment a known protein is expressed fused to a DNA-binding domain and a cDNA library is expressed so that proteins encoded by the cDNA are fused to an activation domain (activation-tagged). Transcription of a reporter gene is only activated in yeast cells containing activation-tagged cDNA-encoded proteins that interact with the known protein. The first library screen was performed using the yeast protein Sir4p fused to the DNA-binding domain of GAL4 as the bait to fish for activation domain-tagged interacting proteins (Chien *et al.*, 1991). Because Sir4p was known to form homodimers it was anticipated that Sir4p itself would be obtained in the screen. Sir4p was detected as an interacting protein as expected, demonstrating that a library approach was feasible.

Since its development the two-hybrid system has become an important tool in molecular biology for the study of protein-protein interactions. There have been many studies which have used the system to assay interactions between two known proteins. In addition, and potentially the most powerful application of the system, has been the ability to screen cDNA libraries to detect novel protein-protein interactions. Many successful screens have been performed, a relevant example being the screen performed by Paroush *et al.*, 1994. More

recently, several permutations and variations on the two-hybrid theme have been developed, these include: (i) variations to study interactions between other macromolecules, such as the one-hybrid system to detect DNA-protein interactions (Chong and Mandel, 1997) or the three-hybrid system to detect RNA-protein interactions (Zhang *et al.*, 1997), (ii) the reverse two-hybrid system utilized to map specific interaction surfaces within a protein (Vidal *et al.*, 1996), (iii) mammalian based systems using mammalian cells as hosts for the assay, providing a more natural environment in which to study interactions between mammalian proteins (Tsou Tsan *et al.*, 1997), and (iv) the application of the two-hybrid system on a genome-wide scale for analysis of global protein interactions encoded by the whole genome of an organism (for example, the T7 bacteriophage, Bartel *et al.*, 1996).

4.3. The yeast two-hybrid system, an overview.

A brief description of how the system works is given below using protein X and protein Y as examples of two proteins whose interaction is under study (figure 4.1). The proteins X and Y can represent any two proteins, because in theory, the interaction between any two proteins can be studied using the two-hybrid system. A more comprehensive description of the particular system used in this study (Gyuris *et al.*, 1993) is given in section 4.4.

The yeast two-hybrid system consists of the following components. Two yeast vectors into which cDNAs corresponding to the coding sequences of the proteins under study can be inserted (e.g. proteins X and Y). One of the yeast vectors, often referred to as the 'bait' vector, is used to express a fusion protein linking protein X to a DNA-binding domain (GAL4 or LexA binding domain). The other vector, often referred to as the 'prey' vector, produces a fusion protein linking protein Y with an activation domain (usually the B42 activation domain derived from *E.coli* or VP16 activation domain derived from the Herpes simplex virus protein). In a situation where the two-hybrid system is used to perform a library screen, a library of prey vectors direct expression of activation-tagged cDNA encoded proteins. Another component of the system is the reporter gene, the most commonly used reporters are *lacZ* and/or nutritional markers (e.g. *LEU2*). The reporter genes contain the

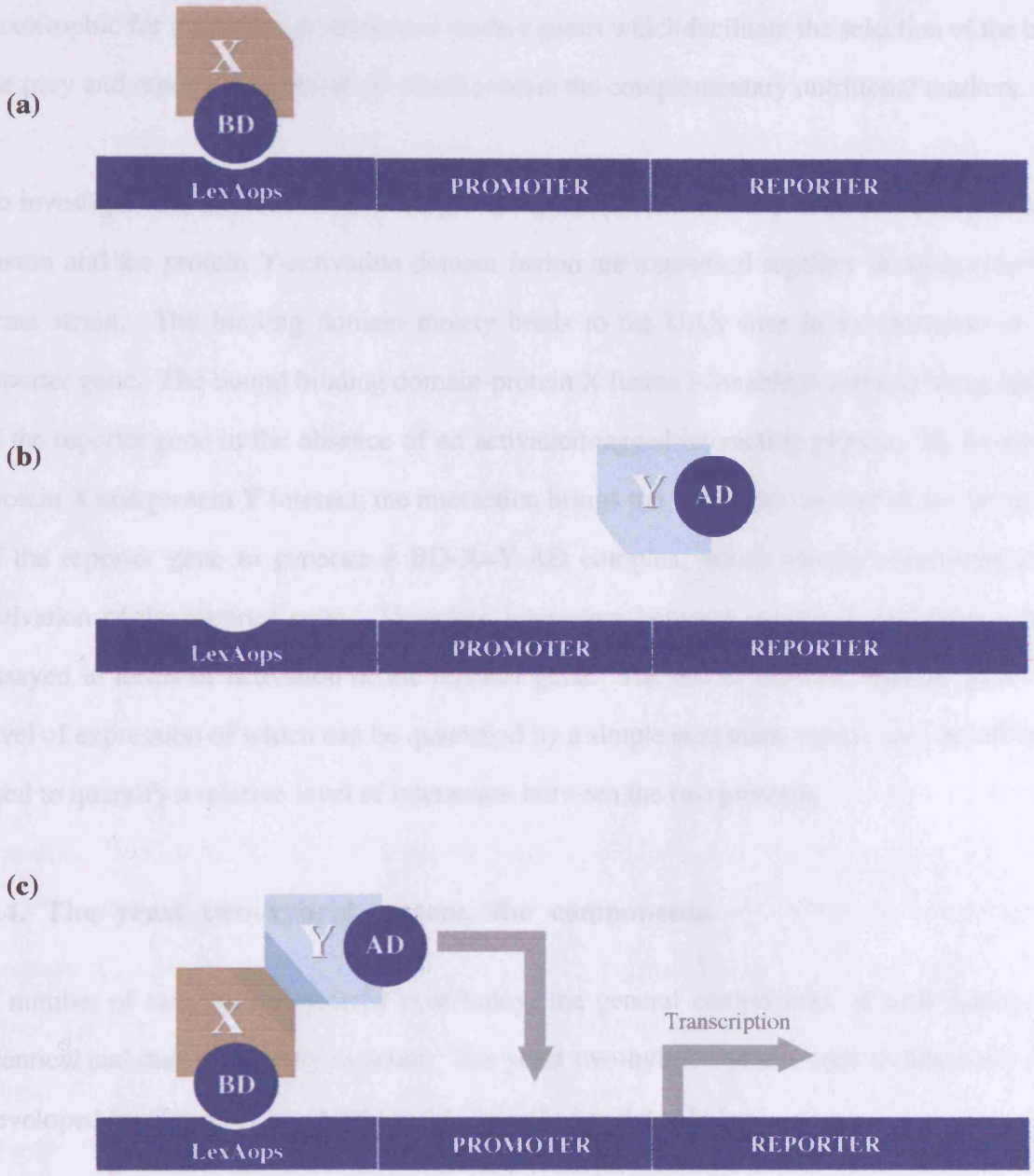


Figure 4.1. The yeast two-hybrid system.
(a) Protein X is expressed as hybrid protein fused to a DNA binding domain (e.g. LexA binding domain). The fusion protein binds the UAS in the promoter of the reporter gene but does not activate transcription of the reporter gene because it does not contain an activation domain.
(b) Protein Y is expressed as a hybrid protein fused to a DNA binding domain. The fusion protein is unable to activate transcription of the reporter gene because, in the absence of a DNA binding domain, it is not targeted to the reporter gene promoter.
(c) If protein X and protein Y interact a functional transcription factor is reconstituted at the promoter of the reporter gene and interaction is detected and/or quantified by the activity of the reporter gene.

binding sites corresponding to the cognate DNA-binding protein of the bait vector. The final component is a host yeast strain in which the assay is performed. The yeast strain used is auxotrophic for a number of nutritional marker genes which facilitate the selection of the bait, the prey and reporter vectors all of which contain the complementary nutritional markers.

To investigate the interaction between protein X and protein Y, the protein X-binding domain fusion and the protein Y-activation domain fusion are expressed together in an appropriate yeast strain. The binding domain moiety binds to the UAS sites in the promoter of the reporter gene. The bound binding domain-protein X fusion is unable to activate transcription of the reporter gene in the absence of an activation-tagged interacting protein. If, however, protein X and protein Y interact, the interaction brings the activation domain to the promoter of the reporter gene to generate a BD-X=Y-AD complex, which results in transcriptional activation of the reporter gene. Therefore interaction between protein X and protein Y is assayed in terms of activation of the reporter gene. The use of the *lacZ* reporter gene, the level of expression of which can be quantified by a simple enzymatic assay, can therefore be used to quantify a relative level of interaction between the two proteins.

4.4. The yeast two-hybrid system, the components.

A number of two-hybrid systems exist today, the general components of each system are identical and they differ only in detail. The yeast two-hybrid system used in this study was developed by Gyuris *et al.* (1996) and is described in detail below.

The bait vector (pEG202).

The pEG202 vector is used to express a hybrid protein containing one of the proteins under study fused to a DNA binding domain. The DNA binding domain utilised in this system is the *E.coli* LexA transcriptional repressor. The binding domain of this repressor recognises and binds specific DNA sequences referred to as LexA operators (LexAops), which are positioned upstream of the reporter gene in the system. The pEG202 vector contains the coding sequence corresponding to amino acids 1-202 of the LexA repressor (encoding the

DNA binding region) followed immediately downstream by a polylinker cloning site into which the coding sequence of a protein under study can be inserted. The vector therefore produces a hybrid protein containing the LexA binding domain at the amino terminus and the peptide sequences corresponding to the protein of interest at the carboxy-terminus. Flanking the LexA coding sequence is the alcohol dehydrogenase (ADH1) promoter which drives constitutive expression of the hybrid protein. Although the hybrid protein does not contain a nuclear localisation signal, expression levels from the ADH1 promoter are sufficient to provide occupancy of the LexAops upstream of the reporter gene. The vector contains the yeast 2μm origin of replication and is therefore maintained autonomously at 20-100 copies per cell. The vector also contains the *HIS3* selectable marker which permits selection and maintenance in a host yeast strain carrying a *his3* mutation (see yeast strain, below).

The prey vector (pJG4-5).

The pJG4-5 vector is used to express a hybrid protein containing the second protein under study fused to a DNA activation domain. The activation domain used in this system is the B42 acid blob activation domain from *E.coli.* The pJG4-5 vector contains the coding sequence for the SV40 nuclear localisation signal, the B42 activation domain sequence and the HA epitope tag from the influenza virus hemagglutinin gene. Immediately downstream from the HA epitope tag is a polylinker containing the *Eco*RI and *Xho*I restriction sites into which the coding sequence of the second protein under study can be inserted. The hybrid protein expressed from this vector contains the B42 activation domain at the amino terminus and the peptide sequence corresponding to the protein of interest at the carboxy terminus. Expression of the fusion protein is under control of the conditional GAL1 promoter, which is induced in the presence of galactose and repressed in the presence of glucose. The vector also contains the 2μm origin of replication and additionally contains the *TRP1* selectable marker which permits selection and maintenance in a host yeast strain carrying a *trp1* mutation.

The reporter genes.

This system utilises two reporter genes, a *lacZ* reporter gene carried on a plasmid and a chromosomally integrated *LEU2* reporter gene. The *lacZ* reporter provides a relative measure of the 'strength' of interaction between two proteins. β-galactosidase (β-gal) activity, which can be quantified in a simple *in vitro* assay, is directionally proportional to the amount of *lacZ* transcription which, in turn, is proportional to the stability of the complex formed between the two hybrid proteins on the promoter of the reporter gene. Therefore the 'strength' of interaction between two proteins can be quantified in units of β-gal activity. The *lacZ* reporter gene resides on the pSH18-34 plasmid which contains the LexA operators fused to the *lacZ* coding sequence. The plasmid also contains the 2μm origin of replication and the *URA3* selectable marker. The system utilises a second reporter gene, the *LEU2* reporter gene. This reporter has been constructed by the replacement of the *LEU2* regulatory regions at the endogenous *LEU2* locus by LexA regulatory regions to create LexAops-*LEU2*. Homologous recombination was used to integrate this reporter construct into the chromosomal location of the endogenous *LEU2* gene (see host strain, below).

Host yeast strain (EGY48).

The genotype of the EGY48 strain is: *MATα trp1 his3 ura3::(LexAop)6-LEU2*. As mentioned above, the endogenous *LEU2* locus of the EGY48 yeast strain has been replaced by a *LEU2* reporter gene. EGY48 will not grow in the absence of leucine unless LexAops-*LEU2* is transcribed. Additionally the strain also contains deficiencies in three genes which are required to allow selection of the plasmids described above: a *his3* mutation, which is complemented by the *HIS3* marker gene on pEG202; a *trp1* mutation, which is complemented by the *TRP1* marker gene on pJG4-5; and a *ura3* mutation, which is complemented by the *URA3* marker gene on pSH18-34.

RESULTS

4.5. Vector construction.

The *E(spl)-m8* and *E(spl)-m8* mutant derivatives coding sequences were excised from pBluescript II KS as *Bgl*II-*Not*I fragments and inserted in frame into the *Bam*HI-*Not*I sites of the pEG202 vector. The *Bgl*II site in E(spl)-m8 is 23bp downstream from the first methionine codon, therefore these clones encode E(spl) proteins which are missing the first eight amino-terminal residues. The correct junction between the LexA coding sequence and the *E(spl)-m8* sequences was confirmed by DNA sequencing using the LEX1 primer. The pJG4-5-Gro clone was a gift from Ze'ev Paroush (Paroush *et al.*, 1994).

4.6. Testing the suitability of the bait vector for two-hybrid analysis.

Before a two-hybrid analysis is performed two control experiments are required to determine that the LexA fusion protein (i) does not have the capacity to activate transcription of the reporter genes itself, and (ii) is able to enter the yeast nucleus and bind the LexA operators upstream of the reporter genes.

(i) Verification that LexA-E(spl)-m8 and the LexA-E(spl)-m8 mutant derivatives do not activate transcription of the reporter genes.

In this control experiment the LexA fusion proteins are expressed alone and the level of activation of the *lacZ* and *LEU2* reporter genes is assayed. LexA fusion proteins which activate the reporter genes in the absence of the activation domain-tagged fusion protein are unsuitable for two-hybrid analysis. Two control strains which express different fusion proteins are used for side-by-side comparison with the strains expressing the experimental LexA fusion proteins. The pRFHM1 plasmid is used as a negative control for activation of the reporter gene. This plasmid encodes a hybrid protein corresponding to the LexA DNA binding domain fused to a carboxy-terminal region of the *Drosophila* Bicoid protein that does not contain an activation domain and therefore does not activate transcription of the reporter genes. EGY48 cells that contain the pRFHM1 and pSH18-34 vectors do not grow on media

lacking leucine and remain white on media containing X-gal. The pSH17-4 plasmid is used as a positive control for activation of the reporter genes. It encodes a hybrid protein constructed from the LexA DNA binding domain fused to the activation domain of the GAL4 protein which strongly activates transcription of the reporter genes. EGY48 cells containing pSH17-4 and pSH18-34 are able to grow on media lacking leucine and turn blue on media containing X-gal. LexA fusion proteins behaving like the fusion encoded by the pSH17-4 plasmid would be wholly unsuitable for two-hybrid analysis.

Serial transformations of the yeast strain EGY48 were performed, firstly with the *lacZ* reporter plasmid pSH18-34 and then with either pEG202-E(spl)-m8 or the pEG202-E(spl)-m8 mutant derivatives or the control plasmids pRFHM1 or pSH17-34 to generate the following strains:

EGY48: pSH18-34; pEG202-E(spl)-m8
EGY48: pSH18-34; pEG202-E(spl)-m8RH
EGY48: pSH18-34; pEG202-E(spl)-m8RG
EGY48: pSH18-34; pEG202-E(spl)-m8RD
EGY48: pSH18-34; pRFHM1
EGY48: pSH18-34; pSH17-34

In order to test for activation of the *LEU2* reporter gene each of the above strains were grown to mid logarithmic phase in liquid culture under conditions which selected for the maintenance of the plasmids. Ten- and one hundred-fold dilutions were made from the cultures and each culture, along with the dilutions were spotted onto plates either containing or lacking leucine. The plates were incubated at 30°C and the growth rate of each of the spots was monitored over the subsequent three days.

Growth rate of each strain was identical when grown in the presence of leucine. The growth rate of the strains containing the experimental plasmids were equal to the strains containing the control plasmids indicating that the LexA-E(spl)-m8 and E(spl)-m8 mutant derivative fusion proteins were non-detrimental to growth when expressed in yeast. In the absence of

leucine, in contrast, the only strain which sustained growth was the strain containing the pSH17-34 plasmid, no growth was observed for any of the other strains after three days incubation at 30°C. The strain containing pSH17-34 grew under these conditions, as expected, because the plasmid encodes a LexA fusion with an activation domain which activates the *LEU2* reporter gene and therefore supports growth in the absence of exogenous leucine. The pRFHM1 plasmid, which encodes a non-activating LexA fusion protein and which served as a negative control for reporter gene activation, did not support growth in the absence of leucine as expected. Similarly, the strains containing LexA-E(spl)-m8 and E(spl)-m8 mutant derivative fusion proteins were also unable to grow in the absence of leucine indicating that these fusions do not activate the transcription of the reporter gene.

In order to test for activation of the *lac*Z reporter gene the strains listed above were then streaked onto plates containing X-gal. The strain containing the pSH17-34 vector turned blue in a matter of a few hours whereas all the other strains remained white after several days. These results are equivalent to those obtained for the activation of the *LEU2* reporter, in that the only LexA fusion able to activate reporter gene transcription was that encoded by the pSH17-34 vector.

The results from this experiment show that the LexA-E(spl)-m8 fusion and the LexA-E(spl)-m8 derivative fusions do not result in spurious activation of the reporter gene in the absence of an activation-tagged interacting protein, thereby confirming their suitability for two-hybrid analysis.

(ii) Demonstration that the LexA fusion proteins enter the nucleus and bind LexA operators.

In order to determine whether the LexA fusion proteins are able to enter the nucleus and bind the LexA binding sites in the promoter of the reporter genes, an assay known as the repression or blocking assay is performed. The assay is a measure of the ability of a particular LexA fusion protein to block transcription of a *lac*Z reporter gene by virtue of occupancy at LexA operator sites within the promoter of the reporter. The *lac*Z reporter plasmid used in the blocking assay, pJK101, contains the *lac*Z coding sequence fused to the

GAL1 promoter. pJK101 contains most of the GAL1 upstream activating sequences (UASg) but, in addition, it contains two LexA operators positioned between the UASg and the TATA box. Transcriptionally inert LexA fusion proteins that enter the nucleus and bind the LexA operator in pJK101 block *lacZ* transcription. The previously described plasmid pRFHM1 is used as the positive control for LexA operator occupancy. It encodes a transcriptionally inert LexA fusion protein (see above) which enters the nucleus, binds to the LexA operator in pJK101 and blocks *lacZ* transcription. As a negative control the plasmid pRS423 is used. This plasmid does not produce any fusion protein products and therefore transcription is not blocked in this assay. The assay is performed under two different growth conditions (i) in the presence of glucose, and (ii) in the presence of galactose/raffinose. In the presence of glucose the level of *lacZ* transcription from the pJK101 plasmid is lower due to the repressive effect glucose molecules have on the UASg. Performing the assay with glucose as the sugar source is therefore a more sensitive assay to detect LexA fusions which exhibit weak blocking.

Serial transformation of the yeast strain EGY48 was performed, firstly with pJK101 and then with either pEG202-E(spl)-m8, the pEG202-E(spl)-m8 mutant derivatives or the control plasmids PRFHM1 or pRS423 to generate the following strains:

EGY48: pJK101; pEG202-E(spl)-m8
EGY48: pJK101; pEG202-E(spl)-m8RH
EGY48: pJK101; pEG202-E(spl)-m8RG
EGY48: pJK101; pEG202-E(spl)-m8RD
EGY48: pJK101; pRFHM1
EGY48: pJK101; pRS423

Three separate transformants from each of the above strains were streaked side by side onto the following plates, (a) galactose/raffinose containing X-gal and (b) glucose containing X-gal. The plates were incubated at 30°C and the colour of each of the streaks was monitored over the following three days.

In the absence of glucose the yeast strain containing pRS423 turned blue overnight, while all the other strains remained white. After three days incubation the strain containing pSR423 had turned deep blue and all the other strains had turned very light blue. The intensity of colouration of the strains containing the experimental plasmids and the strain containing the pRFHM1 control plasmid was identical. In the presence of glucose the strain containing pRS423 turned blue overnight but the intensity of colouration was lower than on the plate containing galactose/raffinose, the remaining strains remained white even after 3 days.

The strain containing the *lacZ* reporter and the pRS423 plasmid demonstrated that transcription of the *lacZ* reporter occurred in the absence of a LexA fusion protein. In the presence of a LexA fusion protein encoded by the pRFHM1 vector, known to bind the LexA operator in the promoter of the pJK101 *lacZ* reporter gene, transcription was markedly reduced. There was no difference in the intensity of coloration between the strain carrying the pRFHM1 control vector and strains carrying each of the experimental LexA fusion proteins. This indicated that the experimental LexA fusion proteins were able to enter the nucleus and bind the LexA operator in the *lacZ* reporter gene as efficiently as the fusion protein encoded by the pRFHM1 vector.

The two control experiments described above show that all LexA-E(spl) fusion proteins tested fulfil the criteria required of a bait protein for two-hybrid analysis, in that all of the fusions enter the nucleus and bind LexA operators but do not cause spurious activation of reporter genes in the absence of an activation-tagged interacting protein.

It has not been formally shown by western blotting that the LexA-E(spl) fusion proteins are made and that the level of production of the wild-type and each of the three derivative LexA-E(spl) fusions are approximately the same. However, the results of the blocking assay provide evidence that the fusion proteins are made and suggest that the stability of the fusions proteins are similar in that they behave identically in this assay.

4.7. Wild-type E(spl)-m8 interacts with Gro in the yeast two-hybrid system.

Before analysing the level of interaction of the mutant E(spl)-m8 proteins with Gro it was first necessary to demonstrate interaction between the wild-type E(spl)-m8 protein and Gro and to show that the level of interaction was (i) similar to previously published data (Paroush *et al.*, 1994) and (ii) reproducible, to provide a reference level of interaction to which the interaction of the mutant proteins could be compared.

The level of interaction between the two hybrid proteins is quantified by performing a β-galactosidase (β-gal) liquid assay to determine the amount of transcription form the *lac*Z reporter gene. The appropriate yeast strain (carrying the *lac*Z reporter plasmid and the plasmids encoding the two hybrid proteins) is grown in liquid culture, the yeast cells are then permeabilized, and the chromogenic substrate *o*-nitrophenyl-β-D-galactoside (ONPG) is added in excess. After incubation at 30°C the reaction is stopped and product formation (*o*-nitrophenol) is determined spectrophotometrically. β-gal activity is expressed in β-gal units which are calculated according to the equation provided in the materials and methods.

The assay is performed under two growth conditions (i) in the presence of galactose and (ii) in the presence of glucose. As the activation-tagged fusion protein is under control of the conditional GAL1 promoter (section 4.4) by performing the assay under the two different growth conditions it is possible to demonstrate that activation of the reporter gene is dependant on expression of the activation-tagged fusion protein.

Serial transformations of the yeast strain EGY48 were performed with the *lac*Z reporter plasmid pSH18-34, pEG202-E(spl)-m8 and either pJG4-5-Gro or pJG4-5 to generate the following two yeast strains:

EGY48: pSH18-34; pEG202-E(spl)-m8(wild-type); pJG4-5-Gro
EGY48: pSH18-34; pEG202-E(spl)-m8(wild-type); pJG4-5

The strain containing an 'empty' pJG4-5 plasmid, which encodes just the activation domain moiety, is used to control for spurious interactions which might occur between the LexA-E(spl)-m8 fusion protein and the activation domain encoded by the pJG4-5 vector.

Three independent colonies from each of the strains listed above were grown to mid logarithmic phase in liquid media selecting for the presence of the plasmids and in the presence of either glucose or galactose/raffinose. Aliquots of the cultures were permeabilized and the level of β-gal activity was assayed.

The results show that wild-type E(spl)-m8 protein interacts with the Gro protein. The mean level of interaction was 517 (SD ± 34) β-gal units which is approximately 1.5 times higher than the value obtained in a previous study (Paroush *et al.*, 1994). The level of interaction was reproducible, the standard deviation was ± 34 β-gal units from the mean. Activation of the *lac*Z reporter gene was dependant upon the presence of the activation domain-tagged Gro fusion protein: only cultures grown in galactose, in which expression of the activation domain-tagged Gro fusion protein is induced, demonstrated reporter gene activity whereas cultures grown in the presence of glucose, where expression of the activation domain fusion protein is not induced, did not result in substantial activity of the reporter gene (2.9 β-gal units, SD ± 0.2). Yeast strains expressing LexA-E(spl)-m8 and the activation domain moiety alone (not fused to Gro) did not exhibit reporter gene activity indicating that spurious interactions between E(spl)-m8 and the activation domain moiety encoded by the 'empty' pJG4-5 vector did not occur (data not shown).

The assay was also performed on glucose or galactose/raffinose plates containing X-gal and equivalent results were obtained (data not shown). The strains were steaked onto X-gal indicator plates and their colour was monitored over several days. Strains expressing both the LexA-E(spl)-m8 fusion and the activation domain-tagged Gro fusion turned blue overnight on galactose/raffinose plates, in contrast an equivalent strain grown in the presence of glucose remained white after several days. The control strain, expressing LexA-E(spl)-

m8 along with the activation domain alone, did not turn blue either in the presence of galactose/raffinose or glucose.

4.8. The E(spl)-m8 WRPW mutants derivatives are defective in their ability to interact with Gro.

Having demonstrated interaction between wild-type E(spl)-m8 and Gro and shown that the level of interaction was reproducible, interaction between the E(spl)-m8 mutant derivatives proteins and Gro were then tested.

Serial transformations of the yeast strain EGY48 were performed with the *lacZ* reporter plasmid pSH18-34, pEG202-E(spl)-m8 (wild-type) or the pEG202-E(spl)-m8 mutant derivatives and either pJG4-5-Gro or pJG4-5 to generate the following yeast strains:

EGY48: pSH18-34; pEG202-E(spl)-m8 (wild-type); pJG4-5-Gro
EGY48: pSH18-34; pEG202-E(spl)-m8RH; pJG4-5-Gro
EGY48: pSH18-34; pEG202-E(spl)-m8RG; pJG4-5-Gro
EGY48: pSH18-34; pEG202-E(spl)-m8RD; pJG4-5-Gro
EGY48: pSH18-34; pEG202-E(spl)-m8 (wild-type); pJG4-5
EGY48: pSH18-34; pEG202-E(spl)-m8RH; pJG4-5
EGY48: pSH18-34; pEG202-E(spl)-m8RG; pJG4-5
EGY48: pSH18-34; pEG202-E(spl)-m8RD; pJG4-5

The strains were then assayed as previously described.

The level of β-gal activity in galactose/raffinose and glucose conditions are presented in figure 4.2 and table 4.1. The level of β-gal activity presented in this figure represent the mean level of activity measured for six independent yeast colonies (three colonies derived from each of two separate transformations).

In the presence of galactose, the level of interaction between wild-type E(spl)-m8 and Gro, 573 (SD ± 32) β-gal units, is equivalent to the data obtained in section 4.7. In contrast,

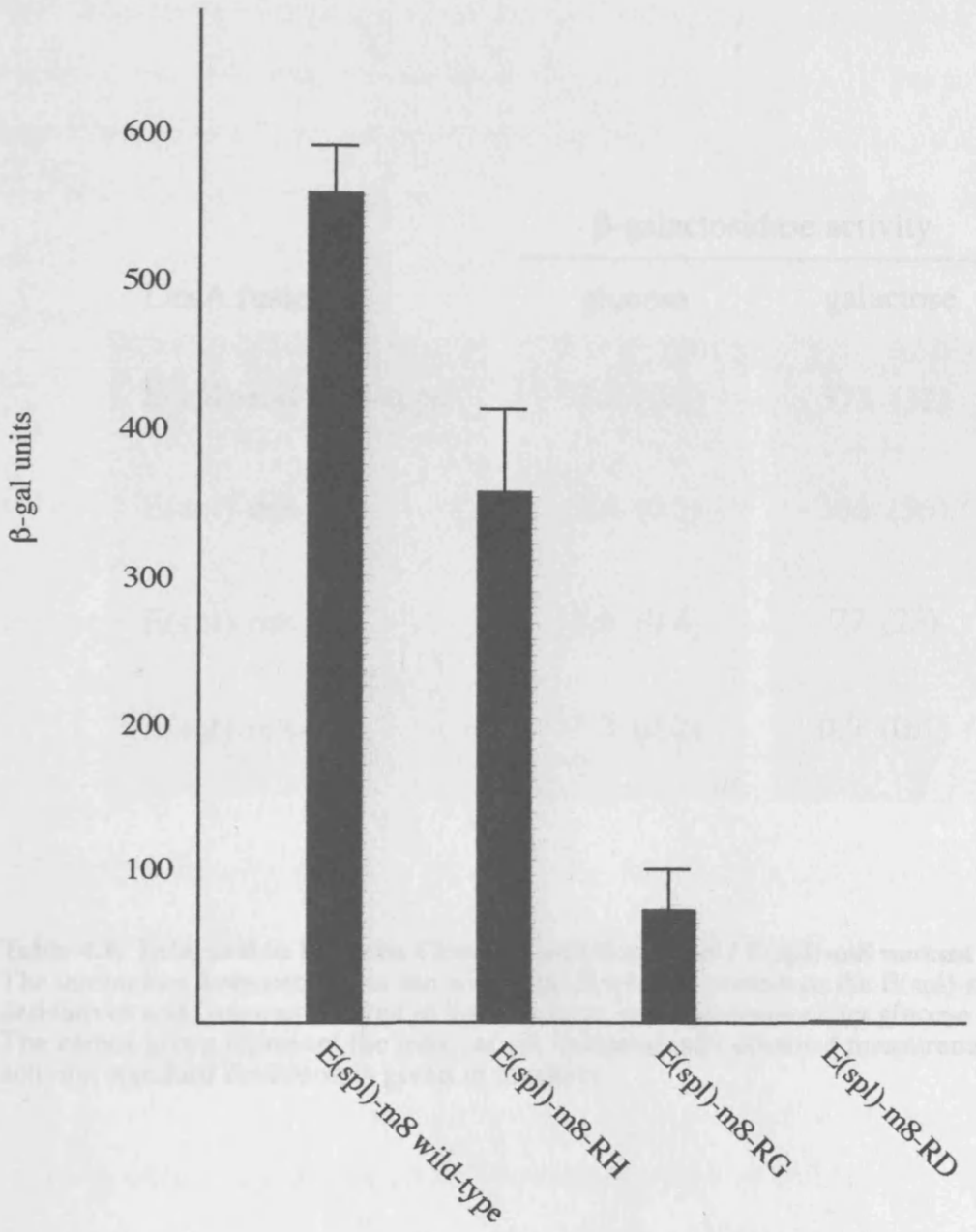


Figure 4.2. Interaction between Groucho and E(spl)-m8 and the E(spl)-m8 mutant derivatives.
The interaction between either the wild-type E(spl)-m8 protein or the E(spl)-m8 mutant derivatives and Gro was assayed in liquid culture. The values given represent the mean of six independently obtained measurements of β-gal activity. Error bars represent standard deviation from the mean.

	β-galactosidase activity	
LexA fusion	glucose	galactose
E(spl)-m8 wild-type	2.2 (0.2)	573 (32)
E(spl)-m8-RH	2.4 (0.3)	366 (56)
E(spl)-m8-RG	3.8 (0.4)	77 (28)
E(spl)-m8-RD	3.3 (0.2)	0.9 (0.1)

Table 4.1. Interaction between Groucho and E(spl)-m8 / E(spl)-m8 mutant derivatives. The interaction between either the wild-type E(spl)-m8 protein or the E(spl)-m8 mutant derivatives and Gro was assayed in liquid culture in the presence either glucose or galactose. The values given represent the mean of six independently obtained measurements of β-gal activity, standard deviation is given in brackets.

however, the level of interaction for each of the mutant E(spl)-m8 proteins with Gro were lower relative to the level observed with the wild-type E(spl)-m8 protein. In all cases a reduction in the strength of interaction was observed relative to wild-type, the extent of this reduction was, however, variable between each of the mutants. The strength of the interaction between E(spl)-m8RH and Gro was 366 (SD ± 56) β-gal units corresponding to a level of interaction of 64% relative to wild-type. The strength of interaction between E(spl)-m8RG and Gro was 77 (SD ± 28) β-gal units corresponding to a level of 13% relative to wild-type. The interaction between E(spl)-m8RD and Gro was 0.9 (SD ± 0.15) β-gal units, which corresponds to a complete loss of interaction.

The level of β-gal activity in all cultures grown in glucose media was extremely low, as expected in the absence of the Gro fusion, with the highest level observed being 3.8 (SD ± 0.36) β-gal units (for E(spl)-m8RG) indicating that *lacZ* reporter gene activity was dependant upon the expression of the activation-tagged Gro fusion protein. In addition, no *lacZ* reporter gene activity was observed when any of the LexA E(spl) fusion proteins were co-expressed with activation-domain moiety alone (not fused to Gro) indicating that none of the LexA fusion proteins exhibited spurious interactions with the activation-domain moiety encoded by the pJG4-5 vector (data not shown).

The assay was also performed on glucose or galactose/raffinose plates containing X-gal and equivalent results were obtained (data not shown). Strains containing both the E(spl)-m8 and the activation domain-tagged Gro fusion turned blue overnight on galactose/raffinose plates as expected. The strain containing the E(spl)-m8RH binding domain fusion and the strain containing the E(spl)-m8RG binding domain fusion did turn blue but only after 2-3 days and the intensity of the colouration did not reach the level of that observed for their wild-type counterpart. The strain containing the E(spl)-m8RD binding domain fusion remained white even after four days incubation. All strains grown in the presence of glucose remained white after several days.

DISCUSSION

4.9. Discussion.

In the previous chapter three E(spl)-m8 derivatives were generated which contained single amino acid substitutions at the position of the arginine residue in the WRPW motif. An analysis conducted *in vivo* in the developing fly demonstrated that these mutant derivatives were no longer functional. In this chapter the molecular basis of this loss of function has been investigated by analysing the ability of these E(spl)-m8 derivatives to interact with Gro protein.

The role of *gro* as a transcriptional corepressor for a subset of transcriptional regulators, including the E(spl) proteins, is now widely accepted (for review see Fisher and Caudy, 1998). Previous data has indicated that the WRPW motif of the E(spl) proteins is the domain through which interaction with the Gro protein takes place (Paroush *et al.*, 1994; Fisher *et al.*, 1996; Grbavec and Stifani, 1996). It is likely that all E(spl) function, in *Drosophila* at least, is mediated by Gro because null *gro* mutations are at least as strong as mutations which remove the whole of the E(spl) complex (Schrons *et al.*, 1992).

For these reasons it was suggested in the introduction to this chapter that the E(spl)-m8 mutant derivatives described in the previous chapter are either no longer able to interact with Gro or have a reduced affinity for interaction, with the result that, when expressed *in vivo*, they do not interact with endogenous Gro at a level sufficient to mediate function. In order to investigate this hypothesis a yeast two-hybrid analysis has been carried out to assay the interaction between the mutant E(spl)-m8 proteins and Gro.

It was first necessary to determine the level of interaction between the wild-type E(spl)-m8 protein and the Gro protein. The level of interaction obtained was 517 (34) β-gal units, which was approximately 1.5 times higher than the interaction value obtained by Paroush *et al.* (1994). This discrepancy in values may reflect minor differences in the LexA-E(spl)-m8

hybrid protein produced resulting from differential cloning strategies or from differences in the assessment of β-gal activity. In this study, the cloning strategy used to generate LexA-E(spl)-m8 wild-type and each of the mutant derivatives was identical, and because the β-gal activity of the wild-type and the mutant hybrid proteins were measured side-by-side, differences in the values of interaction must result as a consequence of the mutation introduced into the WRPW motif. For these reasons the discrepancy in values obtained in this study compared with the previous study was not of concern.

When the mutant E(spl)-m8 proteins were assayed for interaction with Gro it was discovered that all three mutations disrupted this interaction. Substitution of the arginine residue to histidine or to glycine resulted in a reduction in the level of interaction to approximately 64% or 13% respectively relative to the wild-type E(spl)-m8 protein, whereas substitution to aspartic acid completely abolished interaction. This data demonstrates that the arginine residue in WRPW is important to mediate stable wild-type interaction with the Gro protein. Moreover, it was noted that there was a positive correlation between the strength of interaction and the similarity, with respect to charge, that the substituted amino acid had to the original arginine residue; arginine and histidine are both positively charged whereas glycine is neutral and aspartic acid is negatively charged. This therefore indicates, that one important feature of the arginine residue in the WRPW motif is its positive charge. On the basis of this it could be predicted that a negatively charged residue exists at a position on the interaction surface of the Gro protein with which arginine forms an electrostatic bond. It is clear however that the nature of the charge is not the only important feature of the arginine residue, as substitution of arginine to histidine, another positively charged residue, does not result in wild-type levels of interaction. This suggests that other features of arginine in WRPW are important for wild-type interaction with Gro. These features could include, for example, the size of the residue or its hydrophobicity. Generation of additional substitutions at this position in which the nature of the charge is kept constant but other variables such as size or hydrophobicity are manipulated would help define other features required for wild-type levels of interaction with Gro.

The hypothesis put forward at the begining of this chapter was that E(spl)-m8 mutant proteins are non-functional when expressed *in vivo* because of their inability to interact with endogenous Gro. The two-hybrid analysis performed in this chapter has shown that all three of the E(spl)-m8 mutants either reduce or abolish the level of interaction with the Gro protein when assayed in the yeast cell. If molecular interactions observed in the yeast nucleus can be extrapolated to interactions occurring in the *Drosophila* nucleus then the two-hybrid results described here provide evidence that loss of *in vivo* function is indeed a consequence of reduced or abolished interaction with the Gro protein in the *Drosophila* nucleus. Of particular interest in this respect is the arginine to histidine substitution which demonstrated interaction with Gro in the two-hybrid system at approximately 65% efficiency compared to wild-type but, when tested *in vivo*, the protein was unable to support function. This data indicates that a reduction in the stability of interaction to 65% efficiency is below a threshold required for *in vivo* function and therefore suggests that a relatively highly stable complex between E(spl) and Gro is required *in vivo,* at the target gene promoter, in order for repression to occur.

There are, however, a number of caveats relating to direct extrapolation of the two-hybrid read-out to interactions occurring within the *Drosophila* cell. Although the two-hybrid assay is conducted in a cellular environment it is still an artificial environment. Firstly, the levels of the two proteins under study are likely to be far greater in the yeast cell compared to the *Drosophila* cell because expression of the hybrid proteins in yeast is driven from multicopy plasmids by the highly active ADH1 and GAL1 promoters. A second consideration is that in the two-hybrid system the proteins under study are assayed for interaction as fusion proteins and therefore the level of interaction obtained may not be a true reflection of the strength of interaction between the native proteins in the *Drosophila* cell. For these reasons direct extrapolation from the yeast cell to the *Drosophila* nucleus can only be made on a tentative basis.

Two other E(spl)-m8 derivatives have been generated, by colleagues in the laboratory, by substitution of residues in the WRPW motif; in one of these derivatives the last tryptophan

residue is replaced with phenylalanine (WRPF), and in the other the two central residues, arginine and proline, have been reversed (WPRW) (Chester and Jean-Ettiene personal communication). The former represents the substitution of one highly hydrophobic residue for another and in the latter the whole WRPW motif is inverted with respect to the rest of the protein. The interaction of these derivatives with Gro has been assayed in the two-hybrid system. It was discovered that substitution of tryptophan to phenylalanine resulted in a reduction of interaction to a level approximately 60% relative to the wild-type E(spl)-m8 protein. Whereas the inverted arginine-proline change resulted in complete loss of interaction, similar to what was observed for the arginine to aspartic acid substitution described in this chapter. These two derivatives have not been assayed *in vivo* but the data obtained from the WRPW mutant proteins described in this and the previous chapter would predict that they would not function when expressed *in vivo* in the fly.

It is possible that there is an absolute requirement for each of the four amino acids of the WRPW motif in order to achieve wild-type levels of interaction with the Gro protein. This is difficult to validate unless a whole battery of WRPW derivatives are generated in which substitutions of all possible combinations are represented. However, given the data presented in this chapter in combination with the highly conserved nature of the motif it seems likely that all four residues in WRPW are required to mediate stable interaction with Gro.

In this chapter a two-hybrid analysis has been carried out to assay the interaction between three E(spl)-m8 mutant derivatives and Gro. All three of the mutant E(spl) proteins tested interacted with Gro less stably than the wild-type E(spl) protein, the extent of the of the reduction in stability ranged from 60% to 0% interaction relative to wild-type and reflected the charged nature of the residue at the arginine position. These data provide an explanation for the loss of function observed in the *in vivo* assay performed in the previous chapter, and

provide additional evidence that the functional role of the WRPW motif is to mediate interaction with Gro.

Chapter 5

Mapping interaction domains in the Groucho protein

INTRODUCTION

5.1. The reverse two-hybrid system.

Previous studies by Paroush *et al.* (1994) and Fisher *et al.* (1996) in combination with the data presented in chapter 4 of this work provide evidence that the WRPW motif of the Hairy-related proteins is necessary and sufficient for interaction with the Gro protein. In contrast however, the reciprocal region or regions in the Gro protein necessary to mediate the interaction have not yet been determined in detail. Paroush *et al.* (1994) demonstrated that a truncated version of Gro between amino acids 251-414, corresponding to a central region of the protein, interacted with a number of the E(spl) proteins thereby suggesting that the domain of Gro required for interaction with the E(spl) proteins lies somewhere in this region. In this chapter an attempt is made to map the domain within the Gro protein which is required for interaction with the E(spl)-m8 protein.

The development of the yeast two-hybrid system to analyse interactions between known proteins and to isolate novel interacting partners for a protein of interest has allowed many functional protein-protein interactions to be determined. Once an interaction between two proteins has been identified however, the structural components in each of the proteins which mediate the interaction is often sought in an effort to help understand the mechanism and specificity of the interaction. Traditional methods to determine interaction domains within a protein often rely on biochemistry or require time- and labour-consuming cloning techniques and often lack efficient selection procedures to identify the interaction. More recently a new technique, the 'reverse' two-hybrid system, has been developed which circumvents these problems and has been successfully used to determine interaction domains in a number of proteins (see below).

The reverse two-hybrid system is based largely on the technology of the conventional two-hybrid system discussed in detail in chapter 4 but, as the name suggests, the concept of the system has been turned on its head so that protein-protein dissociation events, as opposed to

protein-protein association events, are selected for. The principle of the assay can be described in three steps (figure 5.1). The first of these is a mutagenesis step in which the coding region of the protein under study (protein Y) is randomly mutagenised using a PCR based approach. The mutagenesis can be carried out across the whole coding region or can be targeted to a particular region within it, depending on whether an interaction domain within the protein has previously been identified. In the second step this library of mutagenised fragments is used to replace the corresponding wild-type sequence contained in the yeast vector expressing protein Y fused to the transcriptional activation domain, which is performed in yeast by means of homologous recombination. To do this the vector is linearised in the target region by restriction digest and subsequently co-transformed into yeast with the mutagenic PCR fragment. In the recombination event the wild-type sequence within the vector is replaced with a mutant sequence resulting in the reconstitution of the recombinant clone. In the final step an interaction assay is performed to select for dissociation events. As the yeast strain used for the co-transformation, described above, expresses the other protein partner, protein X, which is targeted by a DNA binding domain to the promoter of the *lac*Z reporter gene, interaction between the proteins can be determined by monitoring β-galactosidase activity on X-gal indicator plates. A dissociation event between protein X and protein Y leads to the of loss of *lac*Z reporter gene activity and is recognised as a white yeast colony. In contrast, colonies expressing a wild-type Y protein or a mutant protein which has an intact interaction surface still bind protein X and lead to activation of *lac*Z, thus giving rise to blue colonies when stained for β-galactosidase activity.

The first reverse two-hybrid screen was performed by Li and Fields (1993) to determine mutations in the tumour suppresser p53 that disrupt interaction with the simian virus 40 large T antigen. Association between p53 and the large T antigen was demonstrated in a two-hybrid system using p53 fused to the GAL4 binding domain (p53-BD) and the large T antigen fused to the GAL4 activation domain. To isolate dissociation mutations, p53 was randomly mutagenised by PCR to generate a library of mutant p53-BD fusions and the library was screened for fusions which produced pale blue or white colonies when co-expressed with the large T antigen-AD fusions. From the screen 34 distinct mutations were

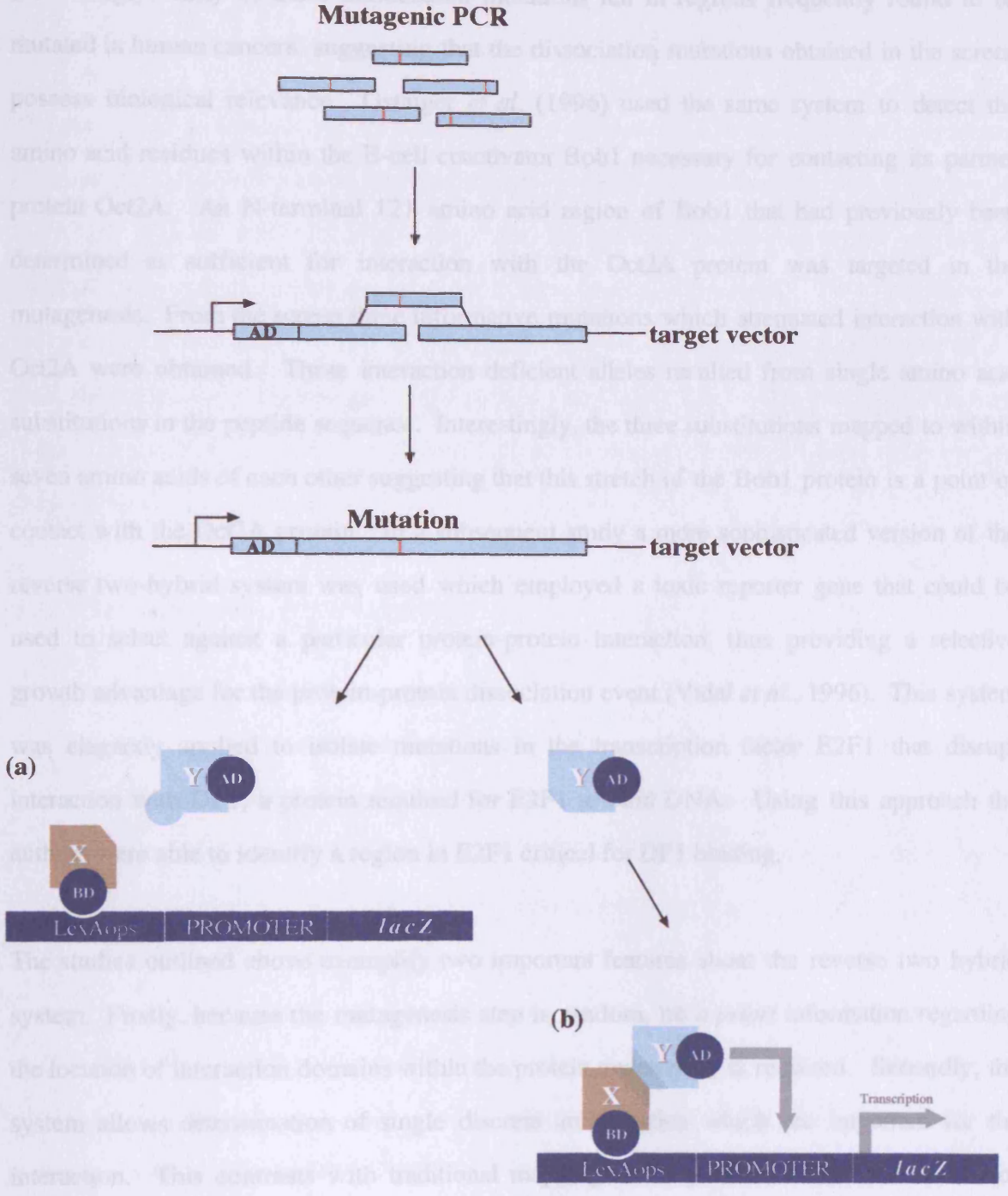


Figure 5. 1. The reverse two-hybrid technique.
The coding sequence of interacting protein Y is randomly mutagenised using a PCR based approach. The mutated PCR fragments are then co-transformed with a linearised target vector that expresses interacting protein Y fused to a transcriptional activation domain. Homologous recombination facilitates substitution of the wild-type coding sequence with the corresponding mutated PCR fragments. An interaction assay is then performed and dissociation events are revealed by loss of reporter gene activity (a). Colonies expressing wild-type protein Y or a mutant protein with an intact interaction surface still interact with protein X and lead to activation of the reporter gene (b).

obtained in p53 which caused disruptions in interaction with the large T antigen. Interestingly, many of these dissociation mutations fell in regions frequently found to be mutated in human cancers, suggesting that the dissociation mutations obtained in the screen possess biological relevance. Gstaiger *et al.* (1996) used the same system to detect the amino acid residues within the B-cell coactivator Bob1 necessary for contacting its partner protein Oct2A. An N-terminal 121 amino acid region of Bob1 that had previously been determined as sufficient for interaction with the Oct2A protein was targeted in the mutagenesis. From the screen three informative mutations which attenuated interaction with Oct2A were obtained. These interaction deficient alleles resulted from single amino acid substitutions in the peptide sequence. Interestingly, the three substitutions mapped to within seven amino acids of each other suggesting that this stretch of the Bob1 protein is a point of contact with the Oct2A protein. In a subsequent study a more sophisticated version of the reverse two-hybrid system was used which employed a toxic reporter gene that could be used to select against a particular protein-protein interaction, thus providing a selective growth advantage for the protein-protein dissociation event (Vidal *et al.*, 1996). This system was elegantly applied to isolate mutations in the transcription factor E2F1 that disrupt interaction with DP1, a protein required for E2F1 to bind DNA. Using this approach the authors were able to identify a region in E2F1 critical for DP1 binding.

The studies outlined above exemplify two important features about the reverse two hybrid system. Firstly, because the mutagenesis step is random, no *a priori* information regarding the location of interaction domains within the protein under study is required. Secondly, the system allows determination of single discrete amino acids which are important for the interaction. This contrasts with traditional mapping techniques where the resolution for mapping interaction domains often spans hundreds of residues.

5.2. The target region in the reverse-two hybrid screen.

The results of Paroush *et al.* (1994) indicate that a central region of the Gro protein spanning amino acids 251-414 is sufficient to mediate wild-type levels of interaction with a number of the E(spl) proteins including E(spl)-m8. For this reason the 251-414 region of the Gro molecule was chosen as the target in the reverse two-hybrid screen. Two primers were designed (m9I and m9II) to use in the mutagenic PCR corresponding to sequence which flank the region encoding this part of the protein (see figure 5.3a). The template DNA used in all PCR reactions described below was a clone containing the whole of the wild-type *gro* coding sequence in pBluescript II.

5.3. The mutagenic PCR protocol.

A mutagenic PCR protocol was sought which would, on average, result in a single random base substitution per molecule of PCR product. Mutation rate in excess of one per molecule was not desirable because the introduction of multiple mutations would complicate interpretation of the results. The mutagenic PCR protocol adopted was developed by Spee *et al.* (1993) and permits the adjustment of the mutation frequency to the size of the target DNA molecule. The method is based on the following principles: (i) one of the four dNTPs is present in limiting amounts in each of four separate PCR reactions, (ii) under these conditions it is possible that misincorporation of one of the other dNTPs is favoured and, (iii) misincorporation might be stimulated when dITP is present (inosine is a base analog that will base pair with the other bases) which, in the next cycle, would result in the incorporation of any of the other nucleotides as a complement to dITP. Spee *et al.* (1993) determined that a reaction containing 20μM of one dNTP (with the remaining dNTPs at a standard concentration of 200μM) in association with 200μM dITP and using standard PCR buffer resulted in a mutation frequency of 27.1 mutations per 10^4 bp. Under these conditions amplification of a 550bp fragment, which is the approximate size of Gro amplicon

used in this study, is estimated to give a mutation frequency of approximately 1.5 mutations per molecule.

The mutagenic PCR was performed as four separate reactions, each containing a limiting amount of one of the dNTPs. The reactions contained the following components: 20 fmols of template DNA, dNTP concentration was 20μM for the limiting dNTP and 200μM for the remaining nucleotides, 200μM dITP, standard PCR buffer, primer m9I and primer m9II at a concentration of 0.5μM each and 5 units of *Taq*-polymerase. The following PCR conditions were chosen for 30 cycles: 94°C for 30 seconds; 48°C for 1 minute; 72°C for 2 minutes. The products from each of the four reactions were then pooled. In addition, a separate PCR was performed under non-mutagenic conditions (200μM of all four dNTPs using the proof reading enzyme *Pfu*) to use as a control.

In order to check that random single base substitutions of approximately the correct frequency were obtained from the mutagenic PCR, aliquots of the products from each of the four limiting dNTP reactions were cloned separately into pBluescript II. Five individual clones derived from each PCR reaction were isolated and approximately 200bp from each was sequenced (figure 5.2). A total of seven base substitutions were obtained from the 20 clones sequenced (approximately 4000 bases were sequenced). This corresponds to a mutation frequency of approximately 1.75 x 10^{-3} per basepair. The results show that the mutagenic PCR protocol fits the criteria required for the experiment. Firstly, the mutation rate corresponds to 0.96 per 550bp molecule, secondly the only mutation events observed were base substitutions (no insertions or deletions were found) and finally, the mutations detected were randomly distributed throughout the molecule.

5.4. The reverse two-hybrid screen.

Serial transformation of the yeast strain EGY48 was performed, first with the *lacZ* reporter plasmid pSH18-34 and then with pEG202-E(spl)-m8 to generate the host strain into which the library of mutant Gro-AD fusion vectors would be introduced. Yeast vector pJG4-5-Gro,

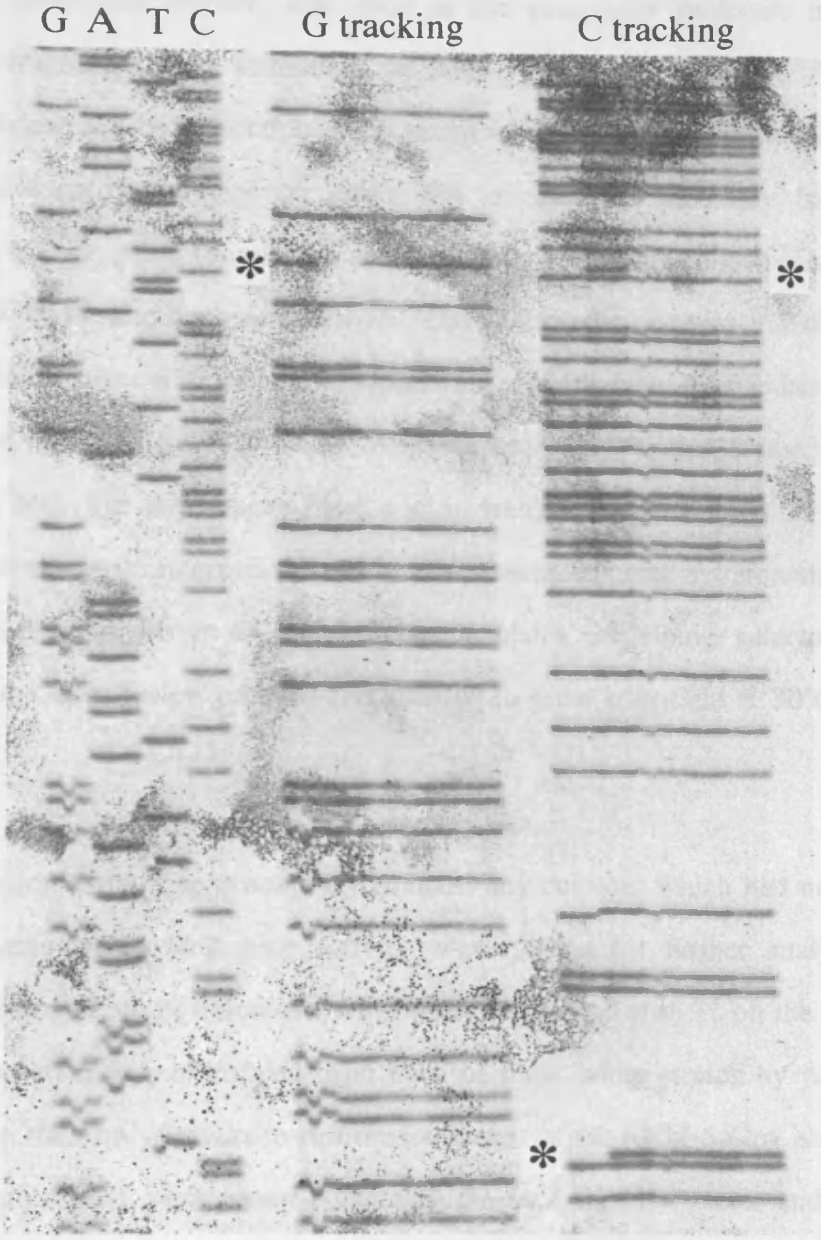


Figure 5.2. Determination of mutation frequency of the mutagenic PCR protocol.
Products from the four mutagenic PCR reactions (limiting-dATP, -dCTP, -dGTP and -dTTP) were cloned separately into pBluescript II and sequenced. The gel shows wild-type gro sequence (left) and sequencing data from five clones derived from a PCR reaction containing limiting amounts of dCTP. G tracking (middle) and C Tracking (right) were performed on clones derived from PCR reactions containing limiting amounts of dCTP and dGTP, A tracking and T tracking were performed on clones derived from PCR reactions containing limiting amounts of dATP and dTTP. Asterisks indicate base substitution events.

which expresses a hybrid protein containing the entire wild-type Gro protein fused to a transcriptional activation domain, was used as the progenitor molecule into which the mutated PCR fragments were introduced by homologous recombination. To prepare the vector it was linearised by restriction digest using *Mlu*I, which cuts once within the target region and leaves at least 40bp on either side of the gap necessary for homologous recombination to take place (Sambrook *et al.*, 1989), the ends of the molecule were then repaired with Klenow and dephosphorylated. The linearised clone was transformed into the host yeast strain together with a five fold molar excess of fragment from either the mutagenic PCR or the non-mutagenic control PCR. Approximately 20,000 transformed colonies were obtained from both the experimental and control transformation. In order to perform a screen for protein-protein interaction, colony lifts were taken using nylon membranes and the membranes were transferred to X-gal indicator plates containing galactose (to induce expression of the Gro fusion protein) and allowed to grow overnight at 30°C to assay for *lac*Z activity.

After an incubation time of approximately 24 hours any colonies which had not turned blue, i.e. did not demonstrate *lac*Z gene activity, were picked for further analysis. On the experimental plates 47 white colonies were obtained compared with 22 on the control plates. The pJG4-5-Gro vector was isolated from each of these white strains by plasmid rescue. The first step in the analysis was to re-transform each of the pJG4-5-Gro isolates from the initial screen into a host yeast strain containing the *lac*Z reporter vector and the pEG202-E(spl)-m8 vector and repeat the assay for *lac*Z activity to determine whether or not the dissociation event observed in the first screen was reproducible. This second screen for *lac*Z activity revealed that a large proportion of the original isolates did not occur as a consequence of a true dissociation event because, upon retransformation, many of the strains now turned blue on X-gal indicator plates. In the second screen twelve of the original experimental pJG4-5-Gro vectors produced colonies which lacked reporter gene activity in comparison to only one of the control pJG4-5-Gro vectors. These twelve experimental pJG4-5-Gro clones were then subjected to sequence analysis.

Sequence analysis was performed using the sequencing primers m9V and m9VI which flanked the region targeted in the mutagenesis (figure 5.3a). The results of the sequence analysis are given in table 5.1 and depicted in figure 5.3b. Eight out of the twelve Gro dissociation clones have a completely wild-type sequence in the region analysed. One clone, GM^2, had a single T to C base substitution at position 1188 on the coding strand corresponding to a synonymous change in third position of a serine codon. The remaining three dissociation clones, GM^1, GM^{12} and GM^{25} all possess nucleotide changes which alter the peptide sequence. The GM^{12} clone has two nucleotide substitutions, a T to G substitution at position 1075 changing the serine residue at codon position 359 into an alanine, and a T to C substitution at position 1190 changing the phenylalanine residue at codon position 397 into a serine. The GM^{25} clone contains a single nucleotide deletion at nucleotide position 850 causing a frameshift in the reading frame. The GM^{25} clone therefore encodes a 387 residue product containing 284 N-terminal amino acids of wild-type sequence followed by 103 amino acids of nonsense sequence. The GM^{25} clone additionally contains an A to G substitution at nucleotide position 997. The GM^1 clone possesses a number of nucleotide changes, an A to C substitution at nucleotide position 1198 leads to an asparagine to histidine amino acid substitution at codon position 400. GM^1 also contains a number of synonymous nulceotide substitutions which interestingly cluster around the region of the m9II primer, an A to G change at position 1257, a G to A change at position 1263 and a G to A change a position 1287. The proximity of these three changes to the end of the PCR fragment suggests that these mutations have resulted as a consequence of the homologous recombination process itself. Although all of the nucleotide substitutions around the m9II primer in this clone are silent third base position changes which do not translate to peptide sequence changes there was still concern that other changes existed outside the region sequenced. To address this concern a *Mlu*I-*Ssp*I fragment (figure 5.3a) was subcloned into the *Eco*RV site of pBluescript II SK^- and sequencing was carried out using the T3 and T7 universal primers. Sequence analysis of the subcloned fragment revealed that the concerns expressed above were justified because a single base deletion was found at nucleotide position 1318. This deletion results in a truncated product of 484 amino acids containing the N-terminal 439 residues of wild-type sequence followed by 45 residues of nonsense

pJG4-5-Gro clone	Position / type of mutation
GM [1]	1198 substitution 1257 substitution 1263 substitution 1287 substitution 1318 deletion
GM [2]	1188 substitution
GM [5]	w-t
GM [9]	w-t
GM [11]	w-t
GM [12]	1075 substitution 1190 substitution
GM [14]	w-t
GM [15]	w-t
GM [25]	850 deletion 997 substitution
GM [34]	w-t
GM [39]	w-t
GM [42]	w-t

Table 5.1. Sequence analysis of the non-interacting Gro clones.
Summary of the sequence data obtained for the twelve non-interacting pJG4-5-Gro clones. The sequence data was obtained using the primers m9V and m9IV. The nucleotide position and type of mutation found for each clone is given. Clones in which mutations were not discovered in the region sequenced are designated wild-type (w-t).

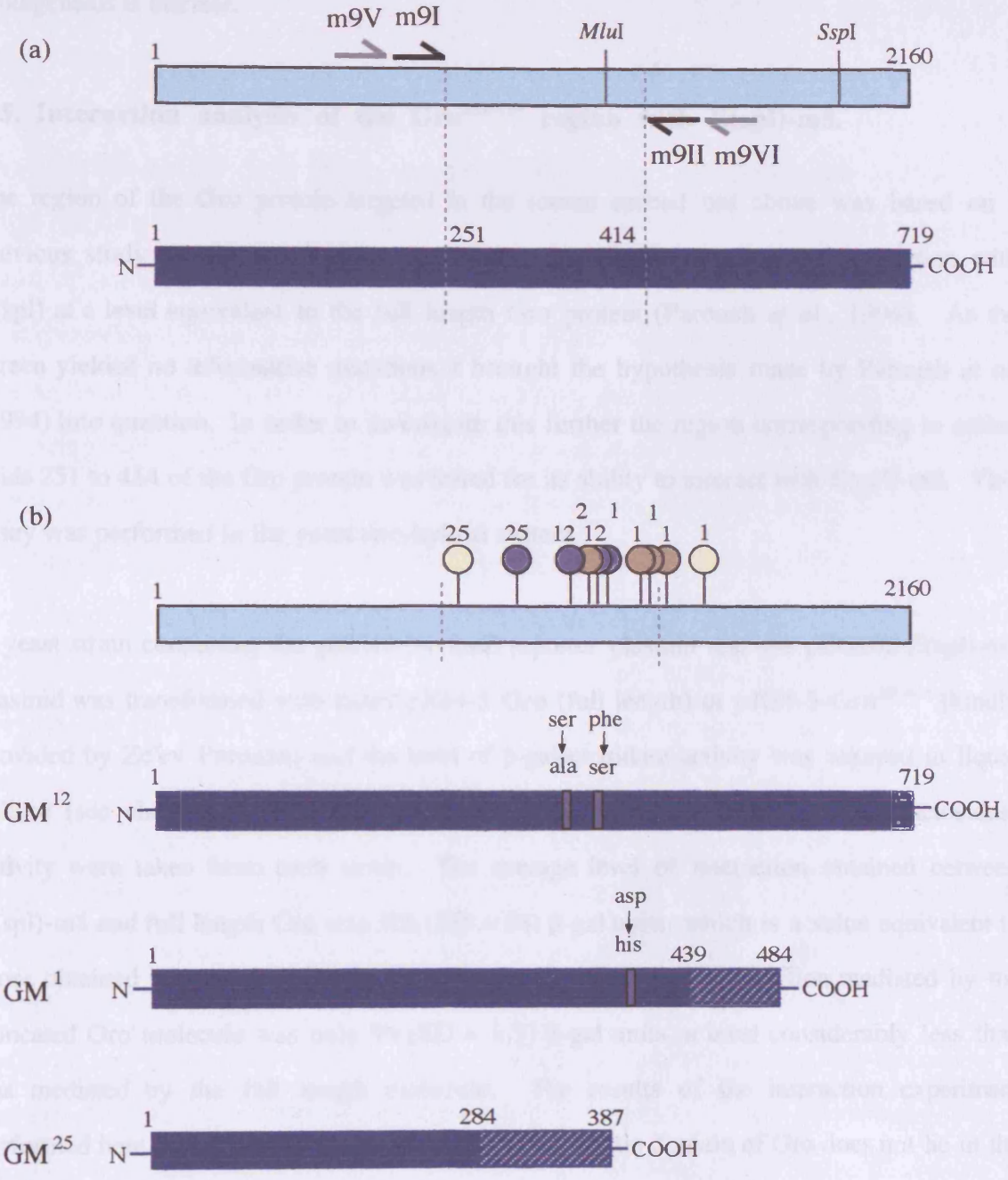


Figure 5.3. Reverse two-hybrid mutagenesis.
(a) Representation of the *Drosophila gro* coding sequence (upper) and protein primary sequence (lower). The m9I and m9II primers (black) were used in the mutagenic PCR to amplify the region of *gro* which encodes amino acids 251-414. The m9V and m9VI primers (grey) are the sequencing primers used in the sequence analysis.
(b) Representation of the mutations obtained in the screen. The lollipops designate nucleotide changes in the coding sequence: deletions are shown in yellow, nonsynonymous changes in purple and synonymous changes are shown in brown. The predicted protein products from clones GM^{12}, GM^{1} and GM^{25} are given. Amino acid substitutions are shown in purple and nonsense peptide sequence resulting from a frameshift is indicated with stippling.

sequence. The reason why a mutation was obtained outside the region targeted in the mutagenesis is unclear.

5.5. Interaction analysis of the Gro$^{251-414}$ region with E(spl)-m8.

The region of the Gro protein targeted in the screen carried out above was based on a previous study which reported that this central portion of Gro mediated interaction with E(spl) at a level equivalent to the full length Gro protein (Paroush *et al.*, 1994). As the screen yielded no informative mutations it brought the hypothesis made by Paroush *et al.* (1994) into question. In order to investigate this further the region corresponding to amino acids 251 to 414 of the Gro protein was tested for its ability to interact with E(spl)-m8. This assay was performed in the yeast two-hybrid system.

A yeast strain containing the pSH18-34 *lac*Z reporter plasmid and the pEG202-E(spl)-m8 plasmid was transformed with either pJG4-5-Gro (full length) or pJG4-5-Gro$^{251-414}$ (kindly provided by Ze'ev Paroush) and the level of β-galactosidase activity was assayed in liquid culture (see chapter 2 for details). Four independent measurements of β-galactosidase activity were taken from each strain. The average level of interaction obtained between E(spl)-m8 and full length Gro was 526 (SD = 88) β-gal units, which is a value equivalent to those obtained in chapter 2. In contrast, however, the level of interaction mediated by the truncated Gro molecule was only 90 (SD = 8.7) β-gal units, a level considerably less than that mediated by the full length molecule. The results of the interaction experiment performed here suggest that the main E(spl)-m8 interaction domain of Gro does not lie in the region spanning amino acids 251-414. This result does not correspond to the result obtained by Paroush *et al.* (1994) even though their analysis was performed using exactly the same Gro fusion molecule. The reason for the discrepancy in results between this study and those of the previous study is unknown.

DISCUSSION

5.6. Discussion.

A reverse two-hybrid screen was performed using the E(spl)-m8-LexA fusion and the a full length Gro-AD fusion both of which were described previously in chapter 2. A central region of the Gro protein between amino acids 251 and 414 was targeted in the screen based on data from a previous study which reported that this region was sufficient to give wild-type levels of interaction with E(spl)-m8, thereby suggesting that the interaction domain lay exclusively within this portion of the protein (Paroush *et al.*, 1994). To isolate dissociating mutants the *gro* coding sequence encoding the 251-414 region was randomly mutagenised using a PCR based method to generate a library of mutant Gro-AD fusions, the library was then screened for fusions that produced white colonies when co-expressed with the E(spl)-m8-LexA fusion. 47 non-interacting clones were obtained in the initial screen, which was subsequently whittled down to only 12 non-interacting clones after the pJG4-5-Gro vector was rescued, re-transformed and subjected to a secondary screen. Sequence analysis revealed that 8 out of 12 of these clones were wild-type in the region targeted in the screen, one additional clone was also obtained which possessed a single synonymous nucleotide substitution which did not alter the amino acid sequence. The remaining three clones however all contained nucleotide changes which altered the amino acid sequence. GM^{12} contained two base substitutions one at position 1075 and one at 1190 resulting in a serine to alanine substitution at codon position 359 and a phenylalanine to serine substitution at codon position 397 respectively. Clones GM^{1} and GM^{25} contained single basepair deletions at nucleotide positions 850 for GM^{25} and 1318 for GM^{1}. These deletions result in frameshifts producing truncated proteins containing 284 amino acids of correct sequence followed by 103 amino acids of nonsense sequence for GM^{25} and 439 amino acids of correct sequence followed by 45 amino acids of nonsense sequence for GM^{1}. GM^{1} additionally contained a number of synonymous nucleotide substitutions in a region close to the m9II primer and a substitution at position 1198 leading to a asparagine to histidine amino acid change in codon 400.

In order to define an interaction domain using data generated from a mutagenesis screen of this kind requires that a large number of separate clones be obtained containing single amino acid substitutions which cluster to a region of the protein. None of the clones isolated in this study, with the possible exception of GM^{12} (see below), satisfy these criteria. The clones GM^{25} and GM^{1} both contained frameshift mutations resulting in proteins containing nonsense peptide sequence at a point after codons 283 and 439 respectively. The observed loss of interaction caused by the mutation in GM^{25} is unambiguously a consequence of the loss of the peptide sequence C-terminal to codon 283. It could be argued for the GM^{1} clone however, given that this clone encodes 439 wild-type amino acids and contains the region determined by Paroush *et al.* (1994) as sufficient for interaction with the E(spl) proteins, that loss of interaction resulting from the mutation in GM^{1} is not a result of the loss of peptide sequence C-terminal to codon 439 but a result of the asparagine to histidine residue substitution in codon 400. This claim however is unlikely especially in light of the data obtained in this study for the analysis of the interaction between $Gro^{251-414}$ and E(spl)-m8. A number of other intriguing features were noted about the GM^{1} clone. Firstly, the base deletion resulting in the frameshift did not occur in the region targeted in the mutagenesis but at a location over 30bp downstream from it and secondly, three base substitutions were observed within a 31bp stretch of nucleotides situated around the m9II primer. It is unclear why there is such a high incidence of mutations around this site and indeed, whether the clustering of the mutations is in any way significant. It is also unclear why a mutation is observed outside the target region. One possible explanation for these occurrences is that they are a consequence of the homologous recombination process itself.

The GM^{12} clone remains a mystery in that either one or both of the base substitutions discovered in the clone could represent the true dissociation mutation. It is also possible that the GM^{12} clone contains other undetermined mutations elsewhere in the sequence which result in the disruption of interaction. To determine whether both mutations or just one of them is the cause of the dissociation event it would be necessary to isolate them and perform

the interaction assay individually for both mutations. Additionally, it would also be necessary to determine that mutations do not exist elsewhere in the Gro molecule.

Sequence analysis of the target region of the remaining nine clones isolated revealed they encoded wild-type peptide sequences. The reason why these clones appear to be wild-type and yet were unable to interact with E(spl)-m8 is not known. One possible explanation is that they contain mutations elsewhere in the coding sequence which result in the production of a non-interacting Gro protein. This is not an unprecedented claim given that mutations outside the target region have been observed in the GM^1 clone. In addition, one non-interacting clone was obtained in control transformations using DNA which had not been mutagenised (section 5.4). The reason for this is unknown, but it may be a consequence of the background mutation rate.

There are a number of reasons why this reverse two-hybrid approach has been unsuccessful, including: (i) technical difficulties encountered with the screening procedure itself; (ii) the possibility that an interaction domain does not exist in the region targeted in the screen. These reasons are discussed in detail below.

Two main technical problems were encountered which hindered the screen. One problem related to the procedure for selection of dissociation events which relied on screening for *lacZ* reporter gene activity and selecting colonies which had no activity when assayed on a plate containing X-gal. This selection procedure was not powerful enough. A total of 47 colonies demonstrating loss of reporter gene activity were obtained in the initial screen but after plasmid rescue of the pJG4-5-Gro vector and re-transformation a large proportion of colonies exhibited wild-type levels of β-galactosidase activity. The reason for the difference in the behaviour of the strains between the primary and secondary screens is unclear. One possibility is that loss of either the pSH18-34 reporter plasmid or the pEG202-E(spl)-m8 plasmid may have occurred, in which case loss of reporter gene activity could be explained by the loss of the E(spl)-m8 fusion protein or the loss of the *lacZ* reporter gene itself. This would explain the observation that, upon re-transformation, reporter gene activity is

restored, but it is difficult to understand how the auxotrophic yeast strain is able to survive and grow on selective media in the absence of the nutritional marker present on each of the yeast vectors. It is also possible that white colonies were obtained due to non-uniformity of X-gal in the indicator plates or because the membranes on which the yeast was growing did not contact the media on the plate.

The second technical problem encountered was an inability to distinguish biologically relevant from biologically irrelevant mutations, such as frameshift mutations, without having to perform sequence analysis. Only three clones (GM^1, GM^{12} and GM^{25}) were isolated which exhibited a loss of reporter gene activity and possessed mutations which could account for this loss in activity. These clones demonstrate that the selection procedure, although inefficient, does work in some cases. Two of these clones however (GM^1 and GM^{25}) were biologically uninformative as they were both a result of a frameshift mutation which only became apparent after sequence analysis and, in the case of GM^1, a subcloning step was additionally required. As the sequence analysis is a rate limiting step in the protocol the elimination of clones which do not provide useful data is desirable.

More sophisticated versions of the reverse two-hybrid system have now been developed and the modifications introduced in these new versions remedy some of the technical problems described above. One technical modification allows a dissociation event to be selected for on the basis that the dissociation event itself provides a selective growth advantage. This has been accomplished by the introduction of a reporter gene whose product can be toxic to the growing yeast cells, therefore the wild-type BD-X=Y-AD interaction results in the death of all cells in which it occurs and, therefore, only those cells in which the interaction has been lost are able to grow. Two toxic marker genes that are used for negative selection are *URA3* and *CYH2* (Vidal *et al.*, 1996; Leanna and Hannink, 1996). The *URA3* gene product is essential for uracil biosynthesis, but additionally, it can also catalyse the transformation of 5-flouroorotic acid (5-FOA) into a toxic compound. Therefore a *URA3* marker gene allows both negative and positive growth selection, on a medium containing 5-FOA or on a medium lacking uracil respectively. Similarly the *CYH2* gene product confers sensitivity to

cycloheximide and therefore a two-hybrid *CYH2* marker gene can be used as a negative growth selector in the presence of cycloheximide. The second problem encountered in the screen performed in this study related to the inability to distinguish biologically relevant mutations from mutations causing frameshifts in the coding sequence. A clever way to distinguish between biologically relevant mutations and those which result in frameshifts has now also been incorporated into the system. This involves the addition of a scorable marker protein whose activity can easily be assayed, such as β-galactosidase or green fluorescent protein (GFP) to the C-terminus of the fusion protein to generate, for example, AD-Y-GFP. Interaction-defective alleles are then screened in the standard way described above but combined with a β-galactosidase or fluorescence assay directly on the yeast colonies to distinguish full length dissociation mutations from nonsense mutations resulting from truncated proteins.

Aside from the technical difficulties encountered with the screening procedure itself a third, more fundamental, problem was one of experimental design. As previously discussed, the region within *gro* targeted in the mutagenesis was based wholly on the data obtained by Paroush *et al.* (1994) who reported that a portion of the protein between amino acid 251 and 414 was able to mediate interaction with a number of E(spl) proteins, including E(spl)-m8, at a level equivalent to the full length Gro protein. These results suggest that the region required to mediate interaction with the E(spl) proteins lay exclusively in this part of the Gro protein. After the screen had been performed and no biologically informative interaction deficient alleles were found, the pJG4-5-Gro$^{251\text{-}414}$ construct was obtained from Paroush *et al.* (1994) and tested for interaction with wild-type E(spl)-m8. In contrast to the data obtained in the previous study it was shown here that the level of interaction fostered by the truncated Gro variant was considerably lower than that fostered by the full length Gro protein. The level of association between the truncated Gro variant and E(spl)-m8 was approximately 83% less than the level of the full length Gro molecule. This data indicates that this part of the protein may have a minor role in the mediation of interaction with E(spl)-m8, but suggets that the main interaction domain maps outside of this region. If the 251-414 amino acid region of Gro plays only a minor role in the mediation of interaction with E(spl)

then it is not surprising that no interaction-deficient alleles were obtained in the screen. It could be argued that since this region has a minor involvement in mediating interaction, that the screen should have revealed some mutations which disrupt rather than abolish interaction with E(spl)-m8. This however is unlikely because, in an effort to avoid ambiguity, only those colonies which had completely lost *lacZ* activity were selected for further analysis and therefore those mutations which only partially reduced interaction would have been missed by this selection procedure.

If the E(spl) interaction domain in the Gro protein is not in the region initially proposed by Paroush *et al.* (1994) then where does it reside? One strong candidate region must be the C-terminal WD40-repeat domain, an ancient motif present in many proteins and previously implicated in the mediation of a number of protein-protein interactions (van der Vroon and Ploegh, 1992; Neer *et al.*, 1994). The WD40-repeat domain of Gro has recently been shown to be required for interaction with the Hairy protein, in that truncations of Gro missing the final WD40-repeat demonstrate weak interactions with Hairy *in vitro* (Jiménez *et al.*, 1997). The authors additionally show however that sequences located in the amino-terminal half of the protein are also required to facilitate *in vitro* interaction with Hairy. The implication from these experiments is that there may be multiple domains in the Gro protein which are required for stable interaction with the Hairy and E(spl) proteins.

Despite the problems encountered in this study the reverse two-hybrid approach may still prove to be a valuable tool for mapping the interaction domains of Gro. Based on the findings of the work described in this chapter, a subsequent screen to determine the interaction domains of Gro using the reverse two-hybrid system should take into account the following recommendations. The first recommendation is that the whole of the Gro molecule be targeted in the screen. The reason for this is that the locations of interaction domains within Gro are not well defined and there is some evidence that multiple interaction domains may exist in disparate regions of the protein (Jiménez *et al.*, 1997). Secondly, the use of a counterselectable marker such as *URA3* or *CYH2* should be used to provide a more powerful selection procedure for dissociation events. To maximise the sensitivity of the

screen, the minimal number of binding sites in the promoter of the *URA3* or *CYH2* reporter genes and the minimal concentration of either 5-FOA or cycloheximide that still allows selection against the BD-E(spl)-m8=Gro-AD interaction should first be determined. This would allow the detection of even small changes in reporter gene expression that would result from weakly dissociating mutations. A sensitised screen such as this would be extremely valuable if a number of interaction domains exist in Gro which each contribute to the stability of interaction. Thirdly, incorporation of a C-terminal marker protein, such as β-galactosidase or GFP, should be introduced which would facilitate the distinction of interaction-defective alleles in the context of a stable, full length protein from nonsense peptides resulting from frameshift mutations. This would allow the latter to be efficiently and rapidly eliminated from the analysis. One other rate limiting step in the protocol was the isolation of the pJG4-5-Gro vector after the interaction screen has been performed. The standard method to do this is to perform a crude plasmid preparation from yeast and use this to subsequently transform bacteria to obtain larger quantities of better quality DNA. All three vectors in the system used in this study contain the bacterial Amp^r selection marker which hindered the selection of the desired vector when transformed into *E.coli.* To alleviate this problem a unique bacterial selection marker gene such as Kan^r should be incorporated to enable rapid isolation of the desired vector. The final recommendation is to use automated sequencing in preference to manual sequencing which would greatly increase the number of clones which could be analysed and would make the prospect of screening the whole Gro molecule less daunting.

Additionally, a separate screen could be performed using the E(spl)-m8 WRPW derivatives described previously in chapter 4 to determine second site suppressor compensatory mutations in the Gro protein which would restore interaction back to wild-type levels. A screen of this kind could be performed as a complement to the one described above.

Chapter 6

Determining the mode of E(spl)-m8 mediated repression during neural fate commitment

INTRODUCTION

6.1. Peripheral nervous system development, the role of the bHLH genes.

The allocation and subsequent implementation of cell fate is a critical feature in the development of all metazoan organisms. One system in which cell fate allocation has been intensively studied is the allocation of neural fate in the development of the peripheral nervous system (PNS) in the *Drosophila* embryo and imaginal discs. The larva and adult fly both posses a large number of external sensory organs most of which perform a mechano- or chemosensory function to allow the animal to sense and react to the external world. The adult fly, for example, contains over 1000 sensory bristles. Many of these bristles are located in a highly stereotypical pattern, so much so in fact, that individual bristles can be uniquely recognised on the basis of their location. It is not surprising therefore, to find that specification of the neural fate during the development of the PNS is highly regulated by a large number of positively and negatively acting factors.

Sensory organ formation in both the larva and adult occurs progressively in a defined temporal sequence of steps during development, these steps include: the positioning of the sense organ in response to local cues, the singling out of the precursor cell through a cell-cell communication process, and the allocation of fates within the lineage through a combination of intrinsic and extrinsic determinants (reviewed by Vervoort *et al.*, 1997). There are multiple levels of both positive and negative control at each step. A summary of the steps during the development of a sense organ is provided in section 1.3.2 and figure 1.1.

A number of the key players involved in the development of the PNS have been identified and their role during neural development has been elucidated. Many of these genes encode products which contain the basic helix-loop-helix (bHLH) domain that mediates both homo- and hetero-dimeric interactions. This domain is found in genes which activate the neural fate, such as the proneural genes of the *AS-C*, and repress the neural fate such as the *E(spl)* genes. It has come to light that the bHLH genes involved in *Drosophila* neurogenesis are

also involved in cell fate choices in other developmental contexts such as muscle precursor determination, malpighian tubule development and sex determination (Carmena *et al.*, 1995; Corbin *et al.*, 1991; Hoch *et al.*, 1994; Parkhurst *et al.*, 1990). Furthermore, the vertebrate homologues of these bHLH genes have also been implicated in cell fate determination during neurogenesis (Guillemot *et al.*, 1993), cardiac muscle development (Srivastava *et al.*, 1995), haematopoiesis (Porcher *et al.*, 1996), mesodermal cell determination (Burgess *et al.*, 1995) and skeletal development (Cserjesi *et al.*, 1995).

The involvement of the same activator and repressor bHLH genes in multiple developmental processes has led to the proposal that the bHLH genes from a 'functional gene cassette' that acts to carry out similar regulatory functions in different contexts (Jan and Jan, 1993). It is not currently known exactly why the bHLH genes have such versatility in function and, although some of these bHLH genes were cloned over ten years ago, the characterisation of their action and the levels at which regulation takes place is still far from complete. For these reasons the work in this chapter attempts to determine the modes by which one of these bHLH proteins, the E(spl)-m8 protein, negatively regulates the neural fate. Cell fate specification in the developing PNS is used as a model to address this problem.

6.2. The role of E(spl) is to repress the neural fate.

The E(spl) proteins are expressed as a result of the activation of the Notch signalling pathway during allocation of SOP fate and again during specification of the progeny of the SOP. The functional role of the E(spl) proteins in both of these steps is to restrict neural development by repressing the neural fate, thereby ensuring a proportion of the cells are available to adopt the alternative, ectodermal fate. This role is exemplified by loss of function (LOF) and gain of function (GOF) genetic analyses. LOF analysis of individual genes from the *E(spl)* locus is difficult because functional redundancy exists between the seven genes of the locus. Tata and Hartley (1995) however used the technique of mitotic recombination to produce clones of cells homozygous for large deletions of the locus. The

results of this experiment show that within the mutant clones derepression of the neural fate occurs and as a consequence supernumerary sensory bristle neurons develop.

GOF analysis has been performed making use of the UAS-P[GAL4] system described in chapter 3. Tata and Hartley (1995) showed that ectopic expression of either *E(spl)-m5* or *E(spl)-m8* during development of the adult microchaete resulted in either complete absence of the sense organ or resulted in sense organs with aberrant cuticular structures. The difference in phenotype reflected a difference in the timing of ectopic expression; induction during early pupal development affected allocation of the SOP whereas later induction affected specification of the SOP daughter and granddaughter cells. Both phenotypes are, however, consistent with a role for E(spl) in the repression of the neural fate; early ectopic expression represses neural precursor formation and later ectopic expression represses the neuron fate (see figure 1.1).

It is clear from these analyses that the role of *E(spl)* is to restrict neural development by repressing the neural fate, but how exactly is this achieved? Several hypotheses pertaining to the mechanisms of *E(spl)* action have been put forward based on the presence in the E(spl) proteins of a number of evolutionary conserved domains. These domains include: a basic region, a Helix-Loop-Helix region, the Orange domain and the WRPW motif.

The basic region is a DNA binding domain which recognises and contacts a specific DNA sequence known as the N-box. This ability to bind DNA is consistent with a role for E(spl) as a transcription factor, and direct binding to DNA is invoked in models of transcriptional repression in association with the corepressor protein Groucho (Gro) (see below). The HLH domain mediates homo- and hetero-dimer formation with other molecules possessing an HLH domain. Homo- and hetero-dimer formation between the E(spl) proteins is a prerequisite for DNA binding. There is also evidence that the E(spl) proteins form dimers with other HLH containing proteins including some of the proneural proteins such as Acheate, Scute and Daughterless (Gigliani *et al.*, 1996; Alifragis *et al.*, 1997). It has been argued that dimer formation between E(spl) and the proneural proteins can sequester

proneural protein activity and thereby inhibit the neural fate (see below). A third conserved feature of the E(spl) proteins is a region downstream of the bHLH domain which is suggested to form two amphipathic helices (helix III and helix IV) (Knust *et al.*, 1992), and which has been called the Orange domain (Dawson *et al.*, 1995). Based on data demonstrating differences in the behaviour between the Orange domain of the Hairy and the E(spl) proteins, it has been suggested that this domain may contribute to functional diversity between the proteins of the family (Dawson *et al.*, 1995). A final feature of the E(spl) proteins is the C-terminal WRPW motif. As described previously in chapters 3 and 4, the WRPW motif is the domain through which interactions with the transcriptional corepressor protein Gro occur (Paroush *et al.*, 1994; Fisher *et al.*, 1996; this study). The accepted hypothesis is that E(spl) and Gro complex via WRPW, the complex binds DNA and the target genes are transcriptionally repressed by Gro in some, as yet, uncharacterised way (see below).

6.3. Models of E(spl)-mediated repression of the neural fate.

A number of mechanisms by which E(spl) proteins repress the neural fate have been proposed. Each of these models is described below and experimental evidence for each model is given. In some cases the evidence can be confusing and sometimes contradictory. It should be noted that the models described are by no means mutually exclusive.

(i) Transcriptional repression of the proneural genes by an E(spl)-Gro repression complex.

A transcriptional repression complex is formed by the association of two E(spl) molecules (dimerising via the HLH domain) and a Gro molecule recruited via interaction with the E(spl) WRPW motif(s). It is possible that other, as yet unidentified, transcriptional corepressors are also recruited to this complex. The basic domain in the E(spl) proteins then binds a specific sequence, the N-box, located upstream of the target genes, the proneural genes. It is unclear at present whether the complex binds DNA fully assembled or whether assembly occurs on the DNA once the E(spl) dimer has bound. The Gro protein (maybe in association with other unidentified factors) is the repression domain of this complex and represses

transcription of the proneural genes by an, as yet, unidentified mechanism (see section 1.10). This model of repression is shown in figure 6.1a.

There is considerable evidence for this mode of repression. It is well established that *gro* function is required for normal neural development as removal of *gro* activity results in extreme neural hypertrophy (Preiss *et al.*, 1988; Schrons *et al.*, 1992). Additionally, physical interaction between E(spl) and Gro (Paroush *et al.*, 1994; Fisher *et al.*, 1996) indicate that the role of Gro is probably executed in association with the E(spl) proteins. Transcriptional repression by Gro has also been demonstrated *in vivo* (Fisher *et al.*, 1996); the mechanism by which this occurs is presently unclear, but some data suggest that changes in the chromatin configuration may be involved (Palaparti *et al.*, 1997). Finally, it has been shown that E(spl) dimers bind to the N-box motif CACGCG, *in vitro* (Oellers *et al.*, 1994), a sequence which is found upstream of the proneural genes *ac* and *sc*, therefore indicating that E(spl) proteins have the potential to bind these sites *in vivo* and influence transcription of the *ac* and *sc* genes.

(ii) Sequestration of proneural protein activity by direct E(spl) binding.

A second potential mechanism that is independent of DNA-binding involves direct interference with activator bHLH proteins. This mechanism is analogous to that previously described for the *Drosophila* HLH repressor Extramacrochaete (Emc) and the homologous mammalian counterpart Id, which form non-DNA-binding heterodimers with the activator proteins thereby titrating their activity (Van Doren *et al.*, 1991; Van Doren *et al.*, 1992; Martinez *et al.*, 1993). It has been shown that Emc dimerises with both Ac and Sc and that dimer formation with Emc prevents Ac and Sc from binding DNA *in vitro* (Cabrera *et al.*, 1994; Ellis *et al.*, 1994; Garrell and Modolell, 1990; Van Doren *et al.*, 1991). As the *ac* and *sc* genes autoregulate and transregulate one another this mechanism of repression also indirectly leads to transcriptional down-regulation of these genes (Van Doren *et al.*, 1992). It is proposed that the E(spl) proteins, which also contain a HLH domain, are capable of regulating proneural activity in this way. As the E(spl) proteins and the Ac and Sc proteins

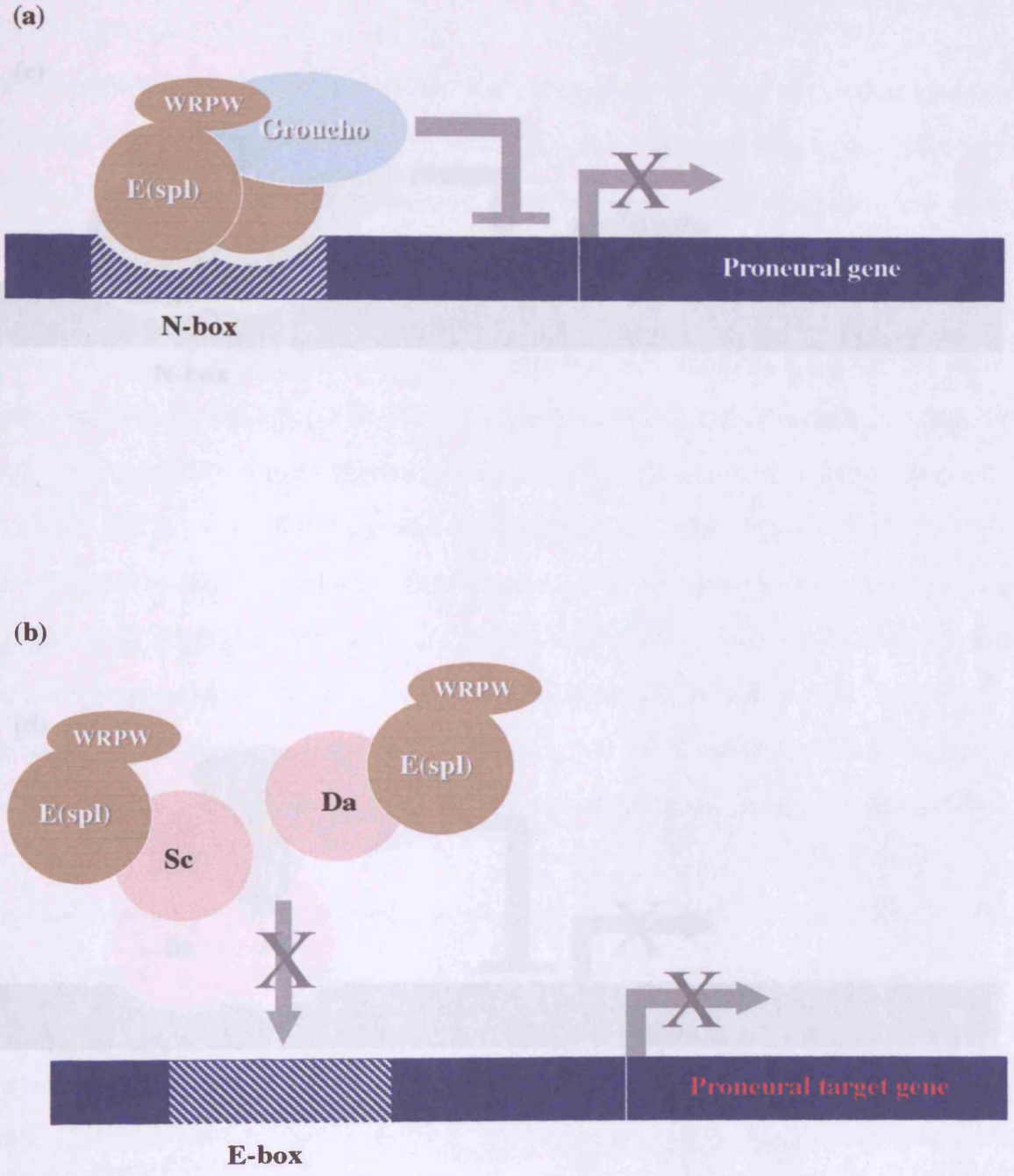


Figure 6.1. Potential modes of E(spl)-mediated repression of the neural fate.

contd.

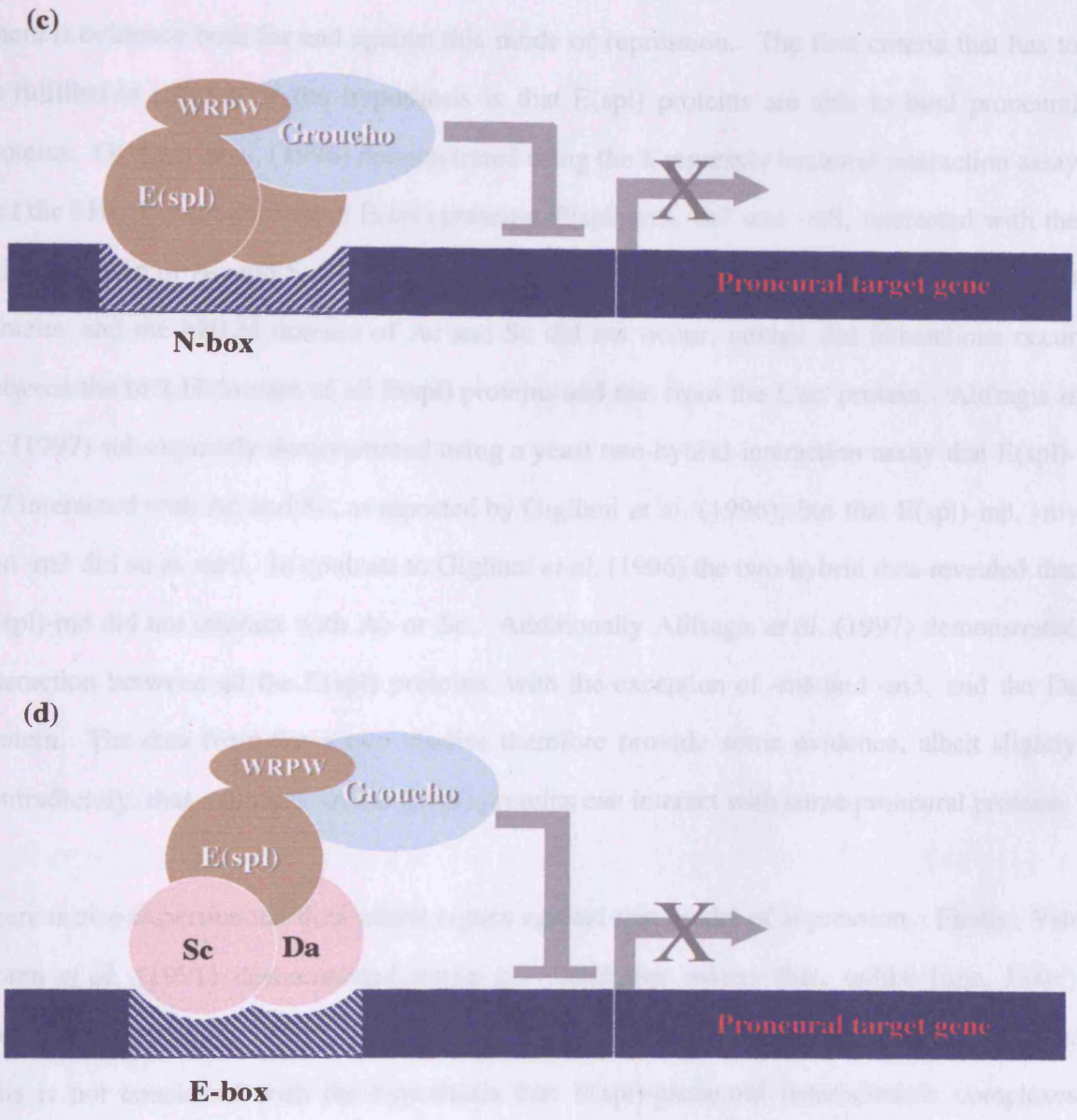


Figure 6.1. Potential modes of E(spl)-mediated repression of the neural fate.
Figure shows four potential modes of repression by which E(spl) could repress the neural fate. (a) An E(spl)-Gro repression complex binds N-box sequences upstream of the proneural genes and transcriptionally represses them. (b) E(spl) protein physically interacts with the proneural protein and titrates their activity. (c) An E(spl)-Gro repression complex binds N-box sequences upstream of the proneural target genes and transcriptionally represses them. (d) The E(spl)-Gro repression complex is recruited to the promoter of target proneural target genes by physical interaction with DNA-bound proneural proteins.

are often co-ordinately expressed there exists the potential for this mechanism of repression. This model of repression is depicted in figure 6.1b

There is evidence both for and against this mode of repression. The first criteria that has to be fulfilled in support of the hypothesis is that E(spl) proteins are able to bind proneural proteins. Gigliani *et al.* (1996) demonstrated using the λ-repressor bacterial interaction assay that the bHLH domain of three E(spl) proteins, E(spl)-m5, -m7 and -m8, interacted with the bHLH domain of Ac and Sc. Interaction between the bHLH domain of the remaining E(spl) proteins and the bHLH domain of Ac and Sc did not occur, neither did interactions occur between the bHLH domain of all E(spl) proteins and that from the L'sc protein. Alifragis *et al.* (1997) subsequently demonstrated using a yeast two-hybrid interaction assay that E(spl)-m7 interacted with Ac and Sc, as reported by Gigliani *et al.* (1996), but that E(spl)-mβ, -mγ and -m3 did so as well. In contrast to Gigliani *et al.* (1996) the two-hybrid data revealed that E(spl)-m8 did not interact with Ac or Sc. Additionally Alifragis *et al.* (1997) demonstrated interaction between all the E(spl) proteins, with the exception of -mδ and -m3, and the Da protein. The data from these two studies therefore provide some evidence, albeit slightly contradictory, that a number of the E(spl) proteins can interact with some proneural proteins.

There is also experimental data which argues against this model of repression. Firstly, Van Doren *et al.* (1991) demonstrated using gel retardation assays that, unlike Emc, E(spl) proteins do not inhibit the binding of Da/Ac or Sc heterodimers to E-box DNA sequences. This is not consistent with the hypothesis that E(spl)-proneural heterodimeric complexes titrate proneural activity by preventing DNA-binding and transcriptional activation. Secondly, an *in vivo* expression study conducted by Giebel and Campos-Ortega (1997) showed that an E(spl)-m8 derivative that lacked the bHLH domain completely still possessed residual activity suggesting that repression of the neural fate is mediated, at least in part, by a mechanism independent of interactions involving the HLH domain. Further evidence against this particular model of repression is presented in chapter 3 of this study. It will be recalled that the expression of E(spl)-m8 derivatives containing a single amino acid substitution in the WRPW motif are unable to suppress bristle development. If repression is mediated by

titration of the proneural protein activity via interactions through the HLH domain then it would be expected that these E(spl)-m8 derivatives, since they contain only a single amino acid substitution in an unrelated domain, would possess at least some residual bristle-suppression activity. As they do not the data suggests that a repression mechanism of this kind does not operate *in vivo*.

(iii) Transcriptional repression of the proneural target genes.

It is believed that the proneural proteins activate the transcription of a number of target genes, that are likely to include other regulatory genes involved in the implementation of the neural developmental program and genes involved in neural differentiation itself. It is possible that transcriptional repression by E(spl) could occur, not just at the level of the proneural genes themselves as described above, but also at the level of these proneural target genes. In a similar manner to the transcriptional repression of the proneural genes, this model invokes assembly of the E(spl)-Gro repression complex at N-box sequences upstream of the proneural target genes resulting in transcriptional repression of these genes.

There is presently little evidence for transcriptional repression of proneural target genes by E(spl) proteins. This lack of evidence is largely a consequence of the paucity of information regarding the targets of the proneural proteins. One characterised proneural target however is the neural precursor gene *asense* (*ase*). The *ase* gene is expressed soon after the SOP is determined and its expression is dependent on the proneural genes. In addition to E-box sequences, the sites at which the proneural proteins bind and which are required for wild-type expression of *ase*, the promoter of the *ase* gene also contains N-box target sites for the E(spl) proteins (González *et al.*, 1989). The possession of N-box sequences therefore makes *ase* a potential target for E(spl) repression. Additionally, it has been recently reported that the sequence motif that the E(spl) proteins bind is more similar to the E-box than previously thought and it has been shown that E(spl) can compete with proneural proteins for the same DNA target sites as *in vitro* (Jennings *et al.*, 1999). These experiments suggest that E(spl) may compete for the same sites in target gene promoters *in vivo* and repress, rather than acivate, their transcription.

Circumstantial evidence for this mode of regulation comes form an experiment performed using a chimeric protein containing E(spl)-m7 fused to the VP16 transcriptional activation domain (Jiménez and Ish-Horowicz, 1997). The chimeric protein retains E(spl)-m7 binding specificity but activates, rather than represses, transcription of its target genes. Ectopic expression of this chimeric protein in a wild-type backgroud during imaginal development results in ectopic bristle formation. This can be explained by ectopic transcriptional activation of the *ac* and *sc* genes by the E(spl)-m7^{Act} protein. However, ectopic expression in an *ac/sc* mutant background does, in some instances, result in the formation of ectopic bristles. This suggests that E(spl)-m7^{Act} can activate alternative targets, raising the possibility that E(spl) proteins also function downsteam of proneural target genes to regulate neural differentiation genes.

(iv) Transcriptional repression of the proneural target genes independent of DNA-binding.

A mode of transcriptional repression which does not involve direct binding of the E(spl) protein to the DNA has been proposed based on unpublished data from Delidakis *et al.* (pers. comm.). They show that an artificial proneural target gene consisting of a *lac*Z reporter transgene containing multiple upstream E-box sequences is expressed in the proneural cluster and in the SOP cells of the wing disc and is dependent upon endogenous *ac* and *sc*. Presumably Ac and Sc in association with Da, bind the E-box sequences in the promoter of the reporter gene and activate transcription of *lac*Z. Accordingly, ectopic expression of UAS-sc results in reporter gene activity in regions of the disc where *sc* is driven. In contrast, however, ectopic co-expression of UAS-E(spl) with UAS-sc results in either severe reduction or loss of reporter gene activity, demonstrating that E(spl) is able to repress Sc activated reporter gene transcription. Importantly, as the reporter gene construct does not contain the N-box sequences required for DNA-binding by E(spl) this mode of repression must be mediated by a mechanism independent of DNA-binding. In a second experiment a hybrid protein consisting of E(spl)-m7 fused to the VP16 transcriptional activation domain was used in the co-expression assay. In contrast to the previous experiment, co-expression of UAS-E(spl)-m7^{Act} and UAS-sc resulted in hyperactivation of the *lac*Z reporter gene

arguing that the UAS-E(spl)-m7^{Act} fusion associates with the reporter gene promoter in some way to activate transcription. In summary, the first experiment shows that E(spl) can repress *sc* target genes by a mechanism independent of direct DNA-binding, whilst the second experiment suggests this mechanism may involve indirect recruitment of E(spl) to the DNA, perhaps by association with DNA-bound Sc protein. The combined data from the two experiments suggest a model of transcriptional repression whereby an E(spl)-Gro repression complex is recruited to the promoters of the proneural target genes not via association with N-box DNA sequences but by association with the E-box-bound proneural proteins themselves. Supporting evidence for this hypothesis comes from data showing that an E(spl)-m8 derivative, unable to bind DNA because its basic domain has been neutralized, is still able to suppress the neural fate when assayed *in vivo* (Nakao and Campos-Ortega, 1996), implying that DNA binding is dispensable for function.

6.4. The ectopic co-expression assay.

It is generally accepted that regulation of neural fate by *E(spl)* occurs by direct transcriptional repression of at least some of the proneural genes. The evidence pertaining to the other mechanisms of regulation described above, in contrast, is relatively inconclusive and there is some debate as to whether or not these mechanisms actually occur *in vivo.* The focus of the work carried out in this chapter is to determine whether E(spl), specifically E(spl)-m8, exerts a regulatory effect in the development of the PNS at levels other than direct transcriptional repression of the proneural genes. In order to do this an *in vivo* co-expression assay was performed, which is described below in detail.

The assay makes use of the UAS-P[GAL4] system previously described in Chapter 3. It will be recalled that ectopic expression of *E(spl)* suppresses the development of bristle SOPs and results in adult flies with a reduction in the number of bristles (Tata and Hartley, 1995; Chapter 3, this work). In contrast, ectopic expression of *sc* or *da* in the imaginal wing disc induces ectopic SOP formation and results in the opposite phenotype, i.e. an adult fly with additional bristles (Hinz *et al.,* 1994). The mutant phenotypes caused by ectopic expression

of *E(spl)*, *sc* and *da* reflect the wild-type role of the genes during neural fate development: *sc* and *da* on the one hand promote the neural fate, whereas the *E(spl)* genes suppress it. Both phenotypes can be quantified by the extent of deviation from the wild-type bristle plan, either in terms of increases in bristle number, for *sc* or *da*, or decreases in bristle number, for *E(spl)-m8*. The assay itself is performed by co-expressing *E(spl)-m8* with either *sc* or *da* in an identical pattern during SOP development and then scoring the outcome in terms of the adult bristle phenotype. Ectopic expression of *E(spl)-m8*, *sc* and *da* is driven from transgenes which are under the control of the UAS enhancer element. Expression from each of these transgenes is induced in the presence of the GAL4 protein and, importantly for the purposes of the assay described here, are not subject to the transcriptional control associated with their endogenous counterparts.

The assay is a measure of the degree to which co-expression with *E(spl)-m8* modifies the phenotype associated with ectopic expression of UAS-sc or UAS-da. In this way, the phenotypic consequence of co-expressing either *sc* or *da* with *E(spl)-m8* provides some indication as to whether regulation by *E(spl)-m8* occurs at levels other than transcriptional repression of *sc* and *da*. Since the ectopic E(spl)-m8 protein cannot transcriptionally repress the transgene-driven expression of *sc* or *da*, any affect on the extra bristle phenotype generated by expression of UAS-sc or UAS-da can only be a result of alternative, post-transcriptional mechanisms of regulation by E(spl)-m8. If the UAS-sc/UAS-da extra bristle phenotype is modified by ectopic expression of UAS-E(spl)-m8 then this would indicate that E(spl)-m8 does have some regulatory effect on neural development aside from repressing the transcription of *sc* and *da*. Alternatively, an inability of ectopically driven *E(spl)-m8* to modify the phenotype would indicate that regulation by *E(spl)-m8* is mediated only at the level of transcriptional repression. The assay is designed to determine whether one or more of these other E(spl)-mediated mechanisms of repression occur in addition to transcriptional repression. It is important to note that the assay can only provide evidence for or against the existence of these other mechanisms, it is unable to distinguish between them.

The analysis carried out in this chapter is performed for two of the four proneural genes, namely *sc* and *da*, the reasons why these two particular proneural genes have been chosen are given below. The aim of the work carried out is to uncover mechanisms of *E(spl)-m8* mediated regulation other than direct transcriptional repression of the proneural genes. As described in 6.3 one of these mechanisms could involve sequestration of proneural activity by the formation of inactive E(spl)-proneural heterodimers. To maximise the likelihood of observing such a mechanism *in vivo* proneural proteins previously demonstrating interactions with E(spl)-m8 were chosen for use in the assay described here. Gigliani *et al.* (1996) demonstrated that both the Ac and Sc bHLH domains interact with the E(spl)-m8 bHLH domain but that of L'sc did not. For these reasons *l'sc* was discounted and *sc* was chosen as a representative for *ac/sc*. In a separate study, Da demonstrated strong levels of interaction with E(spl)-m8 in the yeast two-hybrid system (Alifragis *et al.*, 1997) and was therefore also chosen for use in the assay.

As described above the adult bristle phenotype serves as an easily scored phenotypic marker for a cell fate choice occurring during neural development. Additionally, the A101 enhancer trap line is used as a marker of the SOP cells in the imaginal disc. The A101 line was generated by O'Kane and Gehring (1987) in a *lacZ* enhancer trap screen and represents a transgenic insert at the *neuralized* locus, a gene which is specifically expressed in the SOP cells shortly after they are determined. A101 expression is the earliest known marker of SOP fate and has been utilised extensively as a tool to map the spatial and temporal origin of SOPs within the imaginal discs. The A101 line is used in this study to visualise commitment to the sensory organ fate in the third instar wing imaginal disc.

RESULTS

6.5. Construction of the fly lines and determination of suitable P[GAL4] driver lines.

The UAS-responder fly lines used in the co-expression assays described here were either obtained from other fly groups or were generated previously as part of this study (Chapter 3). A list of these lines and their genomic location is provided in Materials and Methods (Chapter 2). All lines are homozygous viable and are maintained as homozygous stocks.

Two P[GAL4] driver lines, ptc-559.1 and c591, are used for the induction of the UAS-responder expression. The GAL4 expression pattern in these flies have been described previously in Chapter 3. The choice of these driver lines was based on a number of criteria. Firstly, they induce expression in the correct temporal and spatial pattern required for the assay, i.e. during allocation of the SOP cells of the wing disc (see Chapter 3). A second consideration was related to viability after P[GAL4]-driven expression of each of the UAS constructs. The experimental analysis was to be carried out using the adult bristle phenotype and therefore viability to adulthood was required. For this reason the ptc-559.1 was chosen on the recommendation of U. Hinz who had previously shown that ptc-559.1-driven expression of UAS-da produced viable adults (in contrast to many other P[GAL4] lines). A pilot experiment assessing the suitability of the P[GAL4] lines ptc-559.1 and c591 lines in this respect is outlined below. The final criterion was that the P[GAL4] insert in the driver lines had to have a chromosomal location other than on the third chromosome. This requirement was simply to ease construction of the fly stocks, as the UAS-E(spl)-m8 and UAS-da transgenes are both located on the third chromosome.

In order to assess the suitability of the ptc-559.1 and c591 driver lines with respect to adult viability a pilot experiment was carried out in which both P[GAL4] lines were crossed to the UAS-da and UAS-sc (the remaining UAS-responder lines have already been tested in chapter 3). As temperature can have an effect on the penetrance of the phenotype associated

with P[GAL4] mediated expression the crosses were performed at both 18°C and at 25°C. The results of this pilot experiment are presented in table 6.1.

Ectopic expression of UAS-sc with either ptc-559.1 or c591 driver lines at both 18°C and 25°C results in viable adults that eclose normally. This indicates that both driver lines are suitable for use with the UAS-sc responder line. A temperature dependent effect is observed when the same driver lines are used to express UAS-da. Ectopic expression of UAS-da using both driver lines, when crosses are maintained at 18°C, results in viable adults that eclose normally. When grown at 25°C, however, ptc-559.1-driven expression results in lethality during the first larval instar and c591-driven expression results in flies that complete development to adulthood but do not eclose from the pupal case. Therefore, by performing crosses at 18°C the ptc-559.1 and c591 driver lines are suitable for use with the UAS-da responder line. The c591 line can also be used in crosses maintained at 25°C as bristle phenotypes can be scored after dissection of the imago from the pupal case. The difference in the phenotype between the higher and lower temperatures probably reflects an increase in the expression of the UAS-da construct at 25°C due to higher efficiency of the GAL4 protein at this temperature which is closer to the physiological optimum temperature for yeast (30°C).

Construction of the stocks with multiple transgenic inserts was hindered by the fact that most of the transgenes are marked with the w^+ marker gene. It was therefore necessary to balance the desired chromosomes over marked balancer chromosomes which then permitted the selection of the desired chromosome by negative selection for the balancing chromosome. An example of the construction of a stock is given in figure 6.2, the same scheme was used to generate all the others stocks used in the study.

6.6. The phenotypic consequence of ectopic expression of UAS-E(spl)-m8, UAS-sc or UAS-da.

To determine and quantify the adult bristle phenotypes resulting from ectopic expression of UAS-E(spl)-m8, the UAS-E(spl)-m8 mutant derivatives, UAS-sc and UAS-da they were

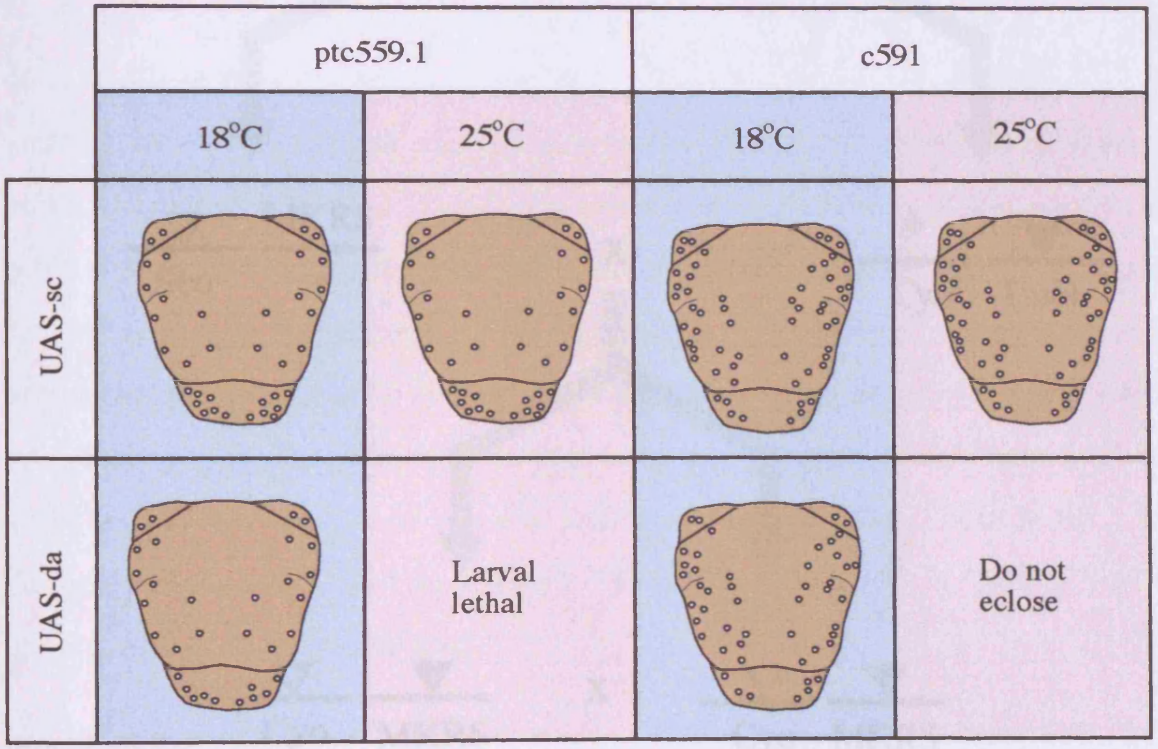


Table 6.1. p[GAL4] driven ectopic expression of UAS-sc and UAS-da.
To determine the suitability of the p[GAL4] driver lines for use in the co-expression assay, UAS-sc and UAS-da fly lines were crossed to either ptc-559.1 or c591 driver lines and maintained at 18°C or 25°C. The figure shows the approximate position and number of ectopic macrochaete (open circles).

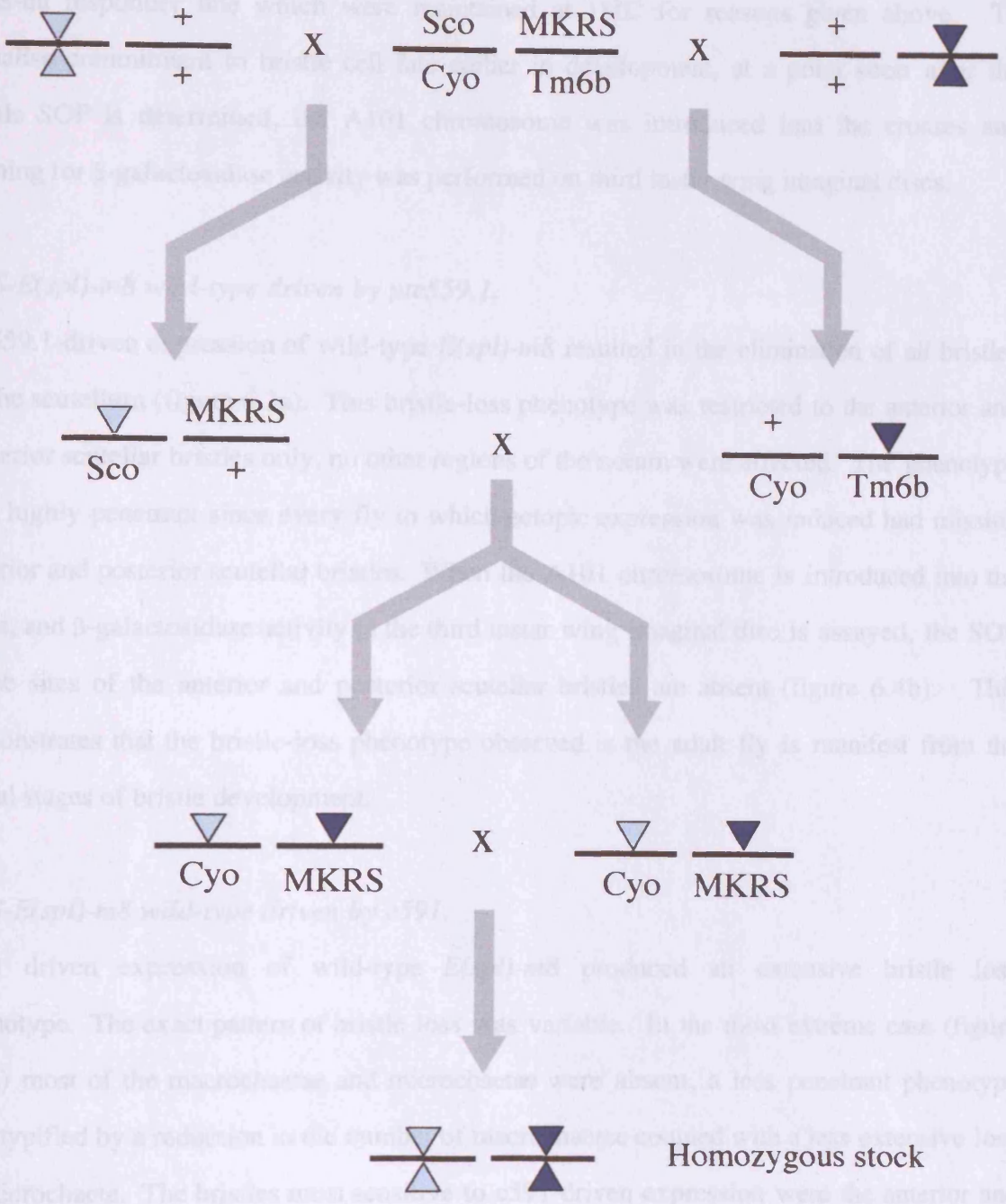


Figure 6.2. Mating scheme used to generate the fly stocks used in the study.
The 2nd and 3rd chromosomes are represented by horizontal bars, the light blue triangle represents a transgene on the 2nd chromosome and the dark blue triangle a transgenic insert on the 3rd chromosome. The transgenic inset on the 2nd and 3rd chromosomes were balanced over Sco and TM6b respectively and selected in the subsequent cross by selection against these balancing chromosomes. Individuals heterozygous for both transgenes and containing the CyO and MKRS balancer chromosomes were then selfed and progeny homozygous for both the transgenes were selected. The flies were subsequently maintained as homozygous stocks.

individually crossed to the P[GAL4] driver lines ptc-559.1 and c591 and the adult bristle phenotype was scored. All crosses were maintained at 25°C except those involving the UAS-da responder line which were maintained at 18°C for reasons given above. To visualise commitment to bristle cell fate earlier in development, at a point soon after the bristle SOP is determined, the A101 chromosome was introduced into the crosses and staining for β-galactosidase activity was performed on third instar wing imaginal discs.

UAS-E(spl)-m8 wild-type driven by ptc559.1.

ptc559.1-driven expression of wild-type *E(spl)-m8* resulted in the elimination of all bristles on the scutellum (figure 6.3a). This bristle-loss phenotype was restricted to the anterior and posterior scutellar bristles only, no other regions of the notum were affected. The phenotype was highly penetrant since every fly in which ectopic expression was induced had missing anterior and posterior scutellar bristles. When the A101 chromosome is introduced into the cross, and β-galactosidase activity in the third instar wing imaginal disc is assayed, the SOP at the sites of the anterior and posterior scutellar bristles are absent (figure 6.4b). This demonstrates that the bristle-loss phenotype observed in the adult fly is manifest from the initial stages of bristle development.

UAS-E(spl)-m8 wild-type driven by c591.

c591 driven expression of wild-type *E(spl)-m8* produced an extensive bristle loss phenotype. The exact pattern of bristle loss was variable. In the most extreme case (figure 6.3b) most of the macrochaetae and microchaetae were absent, a less penetrant phenotype was typified by a reduction in the number of macrochaetae coupled with a less extensive loss of microchaete. The bristles most sensitive to c591-driven expression were the anterior and posterior scutellar bristles as these bristles were always absent after ectopic expression of the wild-type transgene. When the A101 chromosome is introduced into this cross, and β-galactosidase activity in the third instar wing imaginal disc is assayed, a reduction in the number of SOP sites are again observed. Due to the extensive expression of E(spl)-m8 using this particular driver line a large number of the SOP sites are absent including most SOPs corresponding to the notum bristles (figure 6.4f).

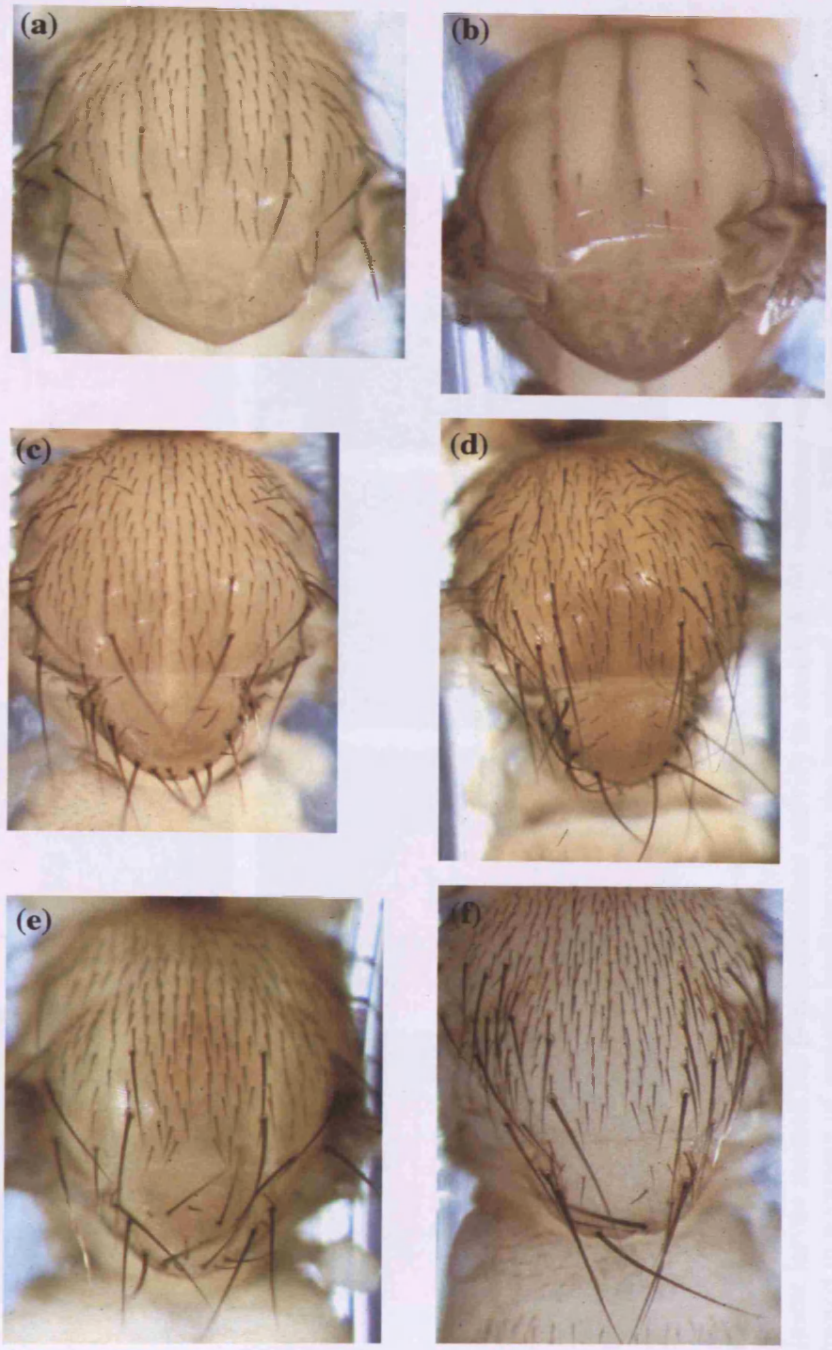


Figure 6.3. The consequence of ectopic expression of *E(spl)-m8*, *scute* and *daughterless* on bristle development.
Adult nota showing the effects of ectopic expression of UAS-E(spl)-m8, UAS-sc and UAS-da on bristle development. Left pannel shows ptc-559.1 driven ectopic expression of (a) wild-type UAS-E(spl)-m8, (c) UAS-sc, (e) UAS-da. The right pannel shows c591 driven ectopic expression of (b) wild-type E(spl)-m8, (d) UAS-sc, (f) UAS-da.

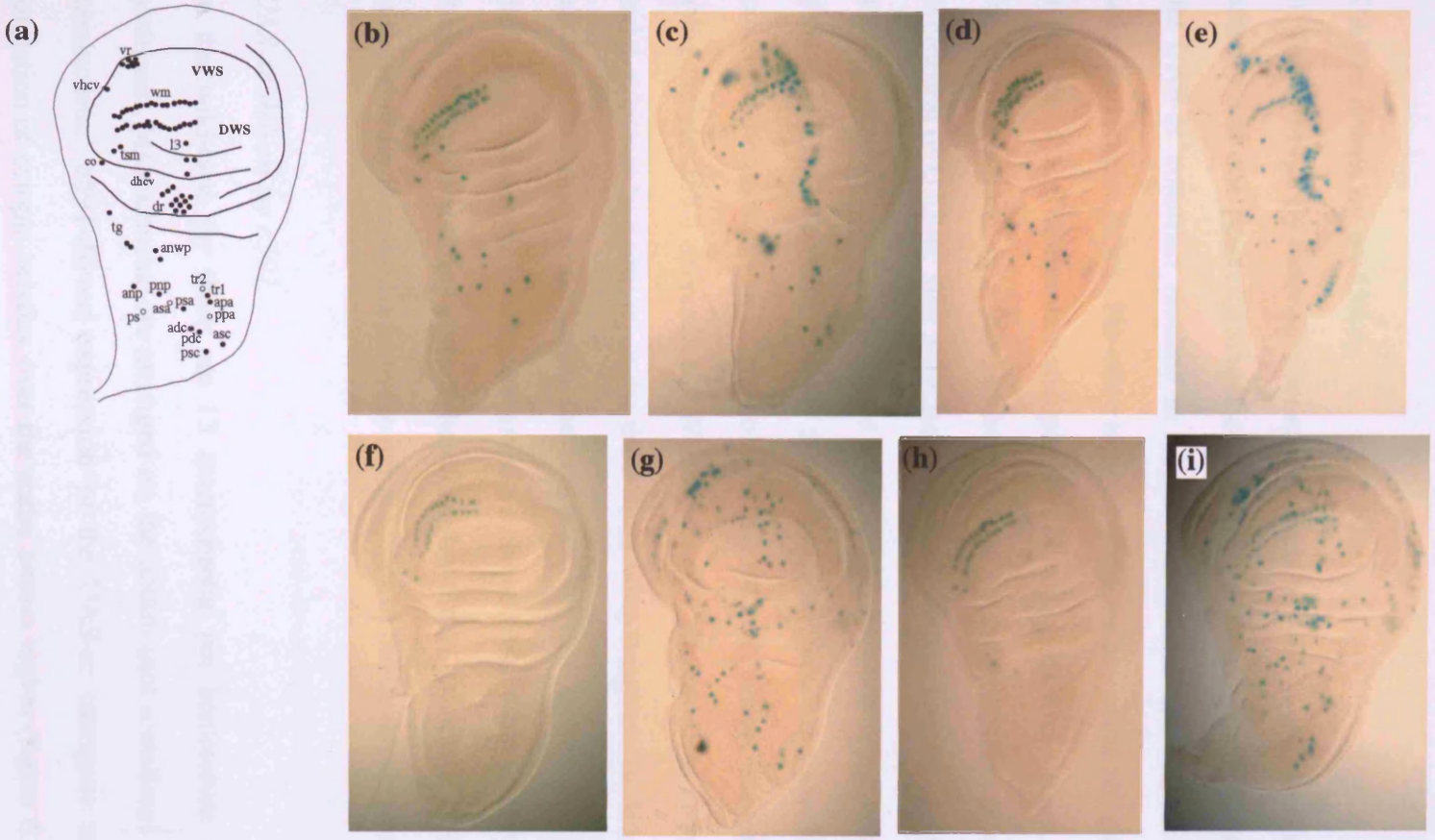


Figure 6.4. The consequence of ectopic expression of E(spl)-m8 and scute on SOP formation.
(a) Schematic figure showing adult sensory structure fate map (see figure 1.2 for full description of abbreviations). (b to i) wing imaginal discs from third instar larvae stained for β-galactosidase activity to reveal A101 enhancer trap expression. A101 enhancer trap expression was determined in a variety of genetic backgrounds: the top pannel shows ptc-559.1 driven ectopic expression of (b) wild-type UAS-E(spl)-m8, (c) UAS-sc, (d) wild-type UAS-E(spl)-m8 and UAS-sc, (e) UAS-E(spl)-m8RH64 and UAS-sc. The lower pannel shows c591 driven ectopic expression of (f) wild-type UAS-E(spl)-m8, (g) UAS-sc, (h) wild-type UAS-E(spl)-m8 and UAS-sc, (i) UAS-E(spl)-m8RH64 and UAS-sc.

P[GAL4] driven expression of the E(spl)-m8 mutant derivatives.

Expression of the UAS-E(spl)-m8-RH, -RG and -RD mutant derivatives using the ptc559.1 and the c591 driver lines does not result in a bristle-loss phenotype (data shown in chapter 3).

UAS-sc driven by ptc559.1.

In the wild-type fly there are four macrochaete bristles on the scutellum corresponding to one anterior and one posterior scutellar bristle per heminotum. ptc-559.1-driven expression of the UAS-sc construct resulted in ectopic macro- and microchaete bristle formation in the scutellum (figure 6.3c). No other regions of the notum where affected by driving expression using this line. The number and position of ectopic bristles varied, and ranged from 11-15 for the macrochaete (n = 90) and from 5-21 for the microchaete (n = 89). ptc-559.1-driven expression of UAS-sc also generates a small number of bristles of intermediate size along the anterior-posterior boundary of the wing blade, between wing veins III and IV (data not shown). It is important to note that the regions where ectopic bristles form correspond exactly to the domains of bristle loss associated with ectopic expression of UAS-E(spl)-m8 using the ptc-559.1 driver line. When the A101 chromosome is introduced into the cross, and β-galactosidase activity in the third instar wing imaginal disc is assayed, additional SOPs are observed in a stripe down the centre of the disc (figure 6.4c). This stripe of SOPs prefigures the ectopic bristles observed on the scutellum and on the wing blade of the adult flies. The A101 staining pattern therefore shows that the ectopic bristle phenotype observed in the adult fly is manifest from the initial stages of bristle development.

UAS-sc driven by c591.

In the wild-type fly there are 13 macrochaetae per heminotum which occupy invariant positions. Microchaetae are arranged on the notum (not scutellum) in five regular rows per heminotum. c591-driven expression of the UAS-sc transgene resulted in the extensive formation of ectopic bristles over the entire notum region (figure 6.3d). The exact number and position of the ectopic bristles varied with the number of macrochaete ranging from 62 to 72 (n = 39). Ectopic expression of the UAS-sc transgene produced three other mutant

phenotypes: extra bristles of various sizes on the wing blade (data not shown), failure of fusion of the dorsal and ventral wing blades resulting in a 'inflated wing phenotype' (data not shown) and uncoordinated locomotor behaviour. These additional phenotypes were observed in all flies scored. Once again, regions of ectopic bristle formation induced using c591 driver line corresponded exactly to the domains of bristle loss generated when UAS-E(spl)-m8 is expressed using the same driver line (see above). When the A101 chromosome is introduced into the cross, and β-galactosidase activity in the third instar wing imaginal disc is assayed, additional SOPs are observed. As a consequence of the extensive expression of UAS-sc when driven by c591 many ectopic SOP arise throughout the whole disc (figure 6.4g). These SOPs approximately prefigure the ectopic bristles positions observed throughout the notum of the adult fly.

UAS-da driven by ptc559.1.

ptc-559.1 driven expression of the UAS-da construct resulted in ectopic macro- and micro-cheate bristle formation in the scutellum (figure 6.3e). No other regions of the notum where affected by driving expression using this line. The number and position of ectopic bristles within the scutellum varied, and ranged from 10-13 for the macrochaete (n = 21) and from 8-11 for the microchaete (n = 21). ptc-559.1-driven expression of UAS-da, similar to that of ptc-559.1-driven UAS-sc, also results in a small number of ectopic bristles of intermediate size along the anterior-posterior boundary of the wing blade, between wing veins III and IV (data not shown). It is important to note that the regions where ectopic bristles form correspond exactly to the domains of bristle loss associated with ectopic expression of UAS-E(spl)-m8 using the ptc-559.1 driver line.

UAS-da driven by c591.

c591-driven expression of the UAS-da transgene resulted in the extensive formation of ectopic bristles over the entire notum region (figure 6.3f). The exact number and position of ectopic bristles varied, with the number of macrochaete ranging from 30 to 41 (n = 18). Once again, regions of ectopic bristle formation induced by using c591 driver line

corresponded exactly to the domains of bristle loss generated when UAS-E(spl)-m8 is expressed using the same driver (see above).

6.7. Ectopic co-expression of UAS-E(spl)-m8 and UAS-sc / UAS-da.

Co-expression of UAS-E(spl)-m8 with UAS-sc.

UAS-E(spl)-m8 and UAS-sc expression, when driven together using the ptc-559.1 line, resulted in the elimination of all bristles on the scutellum (figure 6.5a). The loss of the scutellar bristles occurred in every fly scored in which co-expression of the two transgenes was induced (n=52). Correspondingly, when the A101 chromosome was introduced into the cross and β-galactosidase activity in the third instar wing imaginal disc was assayed, the SOP cells at the sites of the anterior and posterior scutellar bristles were absent (figure 6.4d). Equivalent results were obtained when the c591 line was used to drive expression of both transgenes together, in that co-expression resulted in an extensive bristle loss phenotype (figure 6.5b) which was identical to the phenotype resulting from expression of the UAS-E(spl)-m8 transgene alone (n=32). The A101 expression pattern from these crosses reveal that the absence of adult bristles is prefigured by an absence of the corresponding SOP (figure 6.4h). Therefore, ectopic co-expression of UAS-E(spl)-m8 and UAS-sc together produces an identical phenotype to that observed when UAS-E(spl)-m8 is expressed by itself. *in situ* hybridisation using an antisense probe to the *sc* transcript revealed that the UAS-sc transgene was expressed at levels equivalent to that observed when the transgene is expressed alone (figure 6.6), indicating, as predicted, that mRNA expression from the UAS-sc transgene is not regulated by the presence of the product from the UAS-E(spl)-m8 transgene

As described previously, ectopic expression of *E(spl)-m8* during SOP formation causes a bristle loss phenotype whereas ectopic expression of *sc* results in the opposite phenotype manifest by an increase in the number of bristles. Inducing ectopic expression of *E(spl)-m8* and *sc* together in the same pattern, as performed above, resulted in the bristle loss phenotype associated with E(spl)-m8 expression indicating that the E(spl)-m8 protein was

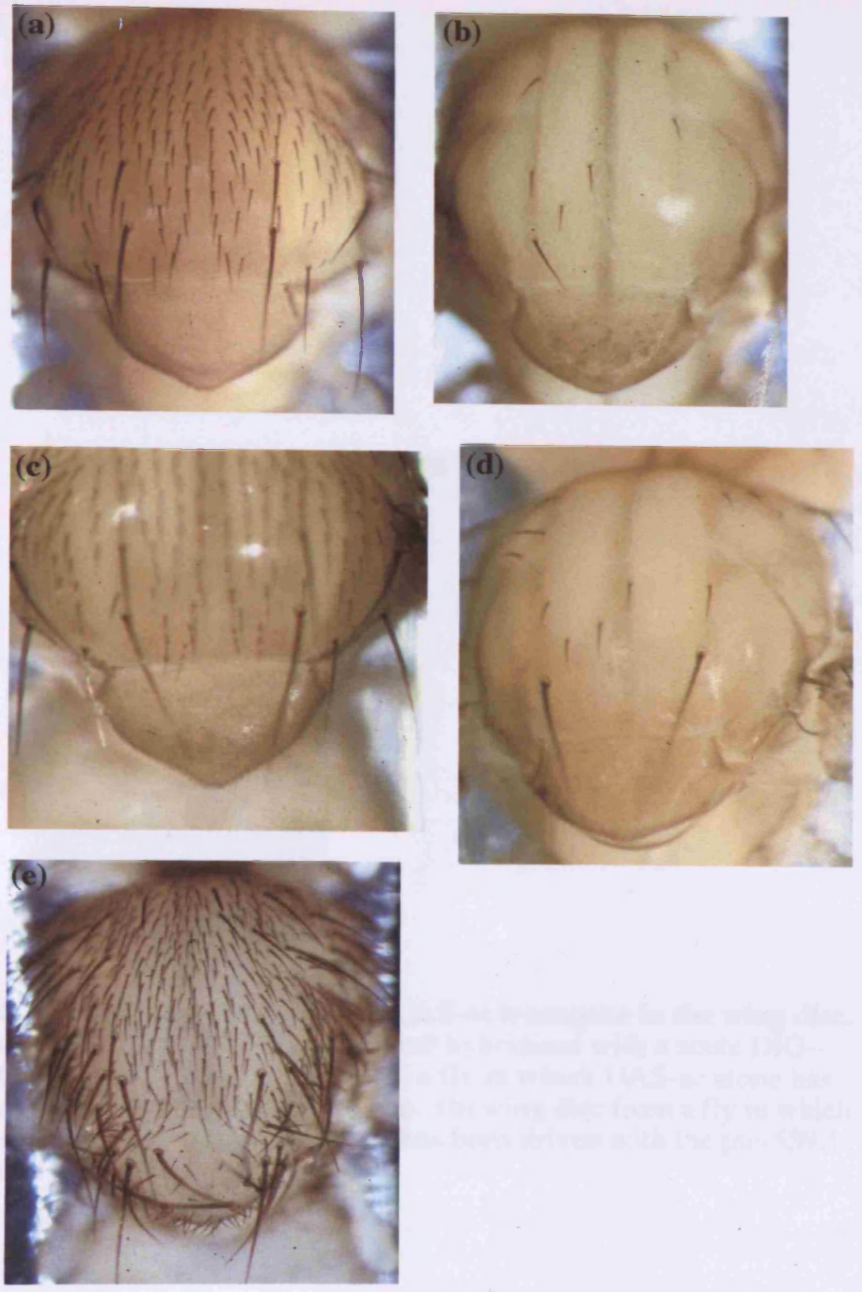


Figure 6.5. The consequence of ectopic co-expression of *E(spl)-m8* with *scute* and *daughterless* on bristle development.
Adult nota showing the effects of ectopic co-expression of UAS-E(spl)-m8 with UAS-sc and UAS-da on bristle development. ptc-559.1 driven (a) and c591 driven (b) co-expression of wild-type UAS-E(spl)-m8 and UAS-sc result in a bristle loss phenotype indistinquishable from ectopic expression from UAS-E(spl)-m8 alone (figure 6.3a and b). ptc-559.1 driven (c) and c591 driven (d) co-expression of wild-type UAS-E(spl)-m8 and UAS-da similarly result in a bristle loss phenotype indistinquishable from ectopic expression from UAS-E(spl)-m8 alone (figure 6.3a and b). (e) c591 driven co-expression of UAS-E(spl)-m8RD51 and UAS-sc results in an ectopic bristle phenotype similar to the phenotype produced by ectopic expression of UAS-sc alone.

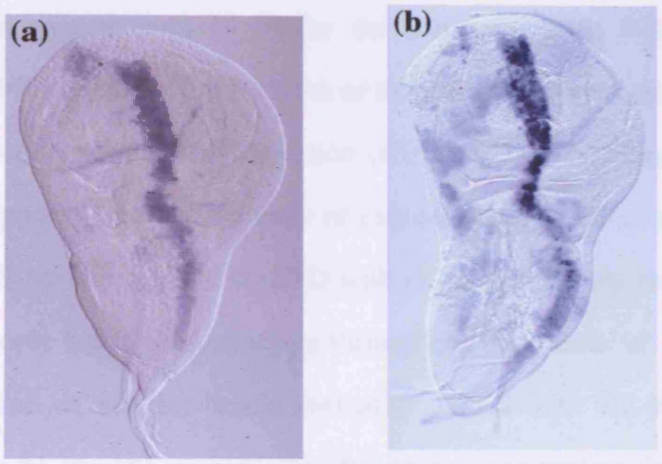


Figure 6.6. Ectopic expression of the UAS-sc transgene in the wing disc. Third instar wing discs were dissected and hybridised with a scute DIG-labelled RNA probe. (a) wing disc from a fly in which UAS-sc alone has been driven with the ptc-559.1 driver line. (b) wing disc from a fly in which UAS-sc and wild-type UAS-E(spl)-m8 has been driven with the ptc-559.1 driver line.

repressing the neural fate by post-transcriptional regulation of Sc activity. To observe whether the WRPW motif of E(spl)-m8 was necessary for this repression the co-expression analysis was repeated using the *E(spl)-m8* WRPW mutant derivatives.

Co-expression of UAS-E(spl)-m8 WRPW mutant derivatives and UAS-sc.

ptc-559.1-driven co-expression of UAS-m8RH or UAS-m8RD with UAS-sc resulted in ectopic macro- and microchaetae formation on the scutellum. This phenotype was similar to that obtained after ectopic expression of UAS-sc alone. The number and position of ectopic bristles varied and ranged from 11-20 for the macrochaete (n = 168) and 6-24 for the microchaete (n = 140) (pooled data for each of the mutant derivatives). Co-expression using this driver line also resulted in the formation of ectopic bristles along the anterior-posterior wing boundary, a phenotype symptomatic of expression of UAS-sc alone. c591-driven co-expression of UAS-m8RH, or UAS-m8RD with UAS-sc similarly resulted in an increase in the number of ectopic bristles at locations throughout the whole of the notum. The exact number and position of ectopic bristle formation varied with the number of macrochaete ranging from 62 to 81 (n=38) (pooled data for each of the mutant derivatives (figure 6.5e). This phenotype is similar to that obtained with ectopic expression of UAS-sc alone. However, the number of ectopic macrochaetae on the head (data not shown) and the density of microchaetae on the notum was greater when UAS-sc was co-expressed with the E(spl) mutant deivatives than when expressed alone (compare the microchaete density in figure 6.3d to figure 6.5e). When the A101 chromosome is introduced into these crosses and β-galactosidase activity in the third instar wing imaginal disc is assayed ectopic SOP sites are observed either in a stripe down the centre of the disc for the ptc559.1 driver line or throughout the disc for the c591 driver line (figure 6.4e and 6.4i).

Co-expression using the c591 driver line additionally caused ectopic bristle formation in the wing blade, an 'inflated-wing' phenotype and abnormal locomotor behaviour, three phenotypes observed after ectopic expression of UAS-sc alone. Therefore ectopic co-expression of either UAS-m8RH or UAS-m8RD with UAS-sc produced a similar phenotype observed when UAS-sc is expressed alone.

Co-expression of UAS-E(spl)-m8 and UAS-da.

UAS-E(spl)-m8 and UAS-da expression, when driven together using the ptc-559.1 line, resulted in the elimination of all bristles on the scutellum (figure 6.5c). The loss of the scutellar bristles occurred in every fly scored in which co-expression of the two transgenes was performed (n=12). Equivalent results were obtained when the c591 line was used to drive expression of both transgenes together (figure 6.5d), in that co-expression resulted in an extensive bristle loss phenotype which was identical to the phenotype resulting from expression of the UAS-E(spl)-m8 transgene alone (n=9). Thus, ectopic co-expression of UAS-E(spl)-m8 and UAS-da together produced a phenotype identical to that observed when UAS-E(spl)-m8 is expressed alone. Co-expression of UAS-E(spl)-m8 and UAS-da resulted in a high degree of lethality during the early larval stages, the reasons for this are unclear.

6.8. Interactions between E(spl)-m8 / E(spl)-m8 derivatives and Sc / Da.

In order to study the molecular interactions between the E(spl)-m8 protein and the Sc / Da proteins and to address whether the inability to suppress the neural fate observed for the E(spl)-m8 WRPW mutant derivatives in the assay described above was a consequence of a change in ability to interact with Sc / Da proteins a yeast two-hybrid analysis was carried out (see Chapter 4 for a detailed description of the yeast two-hybrid system). The pEG202 constructs encoding the E(spl)-m8 and the E(spl)-m8 mutant derivative proteins as fusions to the LexA DNA binding domain have been described previously in Chapter 4. The constructs encoding the Sc and Da fusion proteins were a gift from Christos Delidakis (Alifragis *et al.*, 1997). These constructs, VP16-Sc and VP16-Da, encode either Sc or Da fused to the VP16 transcriptional activation domain derived from the Herpes simplex virus protein which is a very strong transcriptional activator relative to the B42 activation domain encoded by the pJG4-5 construct used previously in Chapter 4.

Serial transformation of the yeast strain EGY48 were performed to generate the following strains:

EGY48: pSH18-34; pEG202-E(spl)-m8 (wild-type); pVP16-Sc
EGY48: pSH18-34; pEG202-E(spl)-m8 (wild-type); pVP16-Da
EGY48: pSH18-34; pEG202-E(spl)-m8RH; pVP16-Sc
EGY48: pSH18-34; pEG202-E(spl)-m8RH; pVP16-Da
EGY48: pSH18-34; pEG202-E(spl)-m8RG; pVP16-Sc
EGY48: pSH18-34; pEG202-E(spl)-m8RG; pVP16-Da
EGY48: pSH18-34; pEG202-E(spl)-m8RD; pVP16-Sc
EGY48: pSH18-34; pEG202-E(spl)-m8RD; pVP16-Da

Three independent colonies from each of these strains were grown to mid logarithmic phase in glucose liquid media under the selection for the maintenance of the plasmids. Aliquots of the cultures were permeabilized and the level of β-galactosidase (β-gal) activity was assayed.

The level of β-gal activity obtained for E(spl)-m8 or the E(spl)-m8 derivatives with Sc is given in table 6.2. The results show that the wild-type E(spl)-m8 protein does not interact with the Sc protein. This data corresponds to previous data obtained by Alifragis *et al.* (1997) who also demonstrated that interaction between wild-type E(spl)-m8 and Sc does not occur in yeast. The data however contrasts with that obtained by Gigliani *et al.* (1996) who showed, using the λ-repressor bacterial interaction assay, that interaction did occur between the bHLH domains of the two proteins. Similarly, the three E(spl)-m8 mutant derivatives also demonstrated an inability to interact with the Sc protein.

The level of β-gal activity obtained between the E(spl)-m8 or the E(spl)-m8 derivative proteins with the Da protein is given in table 6.2 and depicted in figure 6.7. The wild-type E(spl)-m8 protein interacts very strongly with the Da protein. The mean level of interaction from three independent colonies was 1308 β-gal units (SD = 101). This data again corresponds to the previous study by Alifragis *et al.* (1997) who showed that E(spl)-m8 interacts with Da when tested in yeast, although the level of interaction obtained in this study was higher (2000 β-gal units) than that obtained here. When the three E(spl)-m8 mutant derivatives were tested for interaction with Da a surprising result was obtained. The level of interaction between the E(spl)-m8RH derivative and Da was 1293 (SD = 143) which is identical to the level of interaction observed between Da and the wild-type E(spl)-m8 protein.

	β-galactosidase activity	
	Scute	Daughterless
E(spl)-m8 wild-type	2.9 (0.2)	1308 (101)
E(spl)-m8-RH	3.6 (0.3)	1293 (143)
E(spl)-m8-RG	3.3 (0.3)	20 (10)
E(spl)-m8-RD	2.9 (0.2)	3.2 (1.0)

Figure 6.2. Interaction between E(spl)-m8 / E(spl)-m8 mutant derivatives with Scute / Daughterless.
The interaction between either the wild-type E(spl)-m8 or the E(spl)-m8 mutant proteins with either Scute or Daughterless was assayed in liquid culture. The values given represent the mean of three independently obtained measurements of β-gal activity, standard deviation is given in brackets.

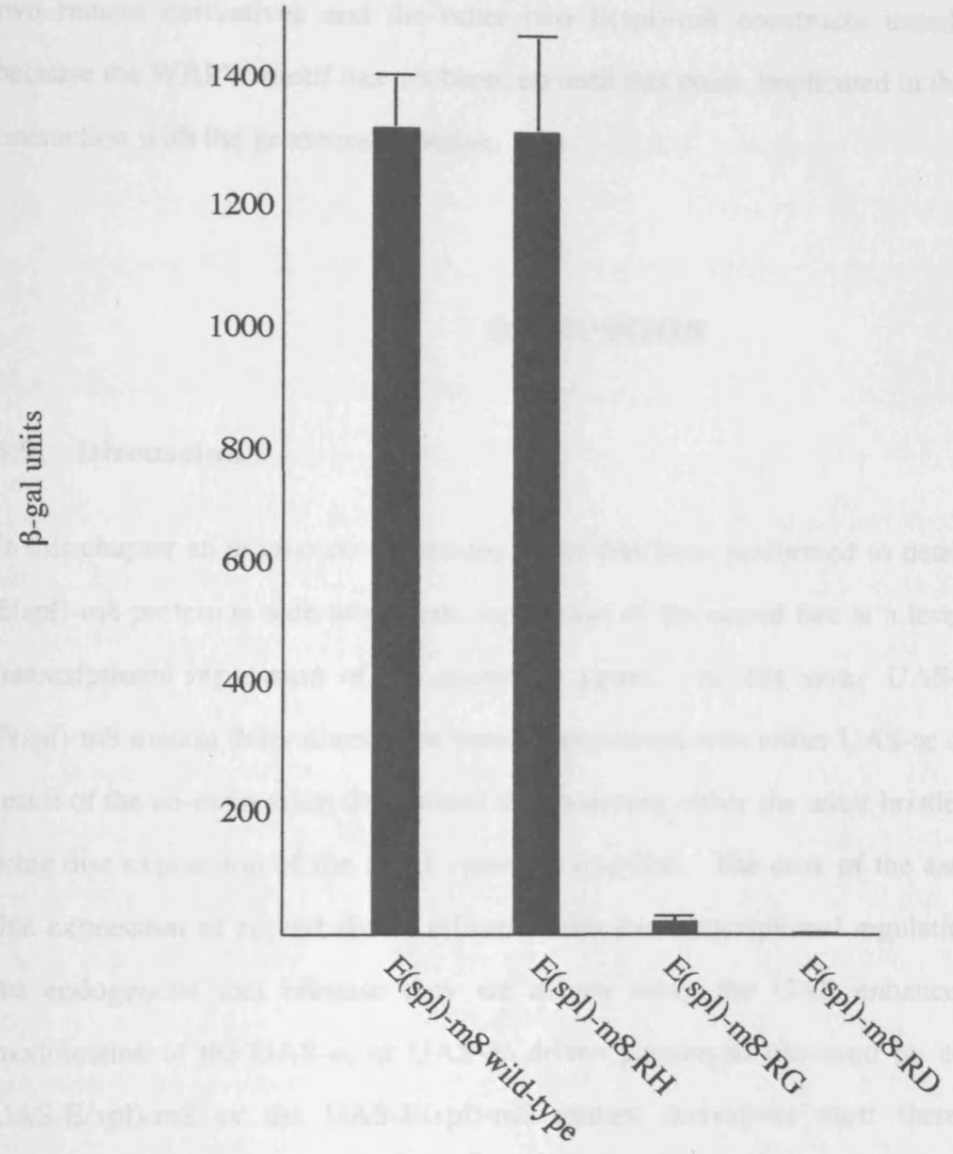


Figure 6.7. Interaction between E(spl)-m8 / E(spl)-m8 mutant derivatives and Daughterless.
The interaction between either the wild-type E(spl)-m8 protein or the E(spl)-m8 mutant proteins and Daughterless was assayed in liquid culture. The values given represent the mean of three independently obtained measurements of β-gal activity. Error bars represent standard deviation from the mean.

The other two E(spl)-m8 derivatives did not however interact with Da. The mean value of interaction obtained for E(spl)-m8RG and E(spl)-m8RD with Da were 20 β-gal units (SD = 10) and 3.2 β-gal units (SD = 1) respectively. The difference in behaviour between these two mutant derivatives and the other two E(spl)-m8 constructs tested was unexpected because the WRPW motif has not been, up until this point, implicated in the mediation of the interaction with the proneural proteins.

DISCUSSION

6.9. Discussion.

In this chapter an *in vivo* co-expression assay has been performed to determine whether the E(spl)-m8 protein is able to mediate repression of the neural fate at a level other than direct transcriptional repression of the proneural genes. In this assay UAS-E(spl)-m8 or the E(spl)-m8 mutant derivatives have been co-expressed with either UAS-sc or UAS-da and the result of the co-expression determined by observing either the adult bristle phenotype or the wing disc expression of the A101 enhancer trap line. The crux of the assay lies in the fact that expression of *sc* and *da* are relieved from the transcriptional regulation associated with the endogenous loci because they are driven using the UAS enhancer element. Any modification of the UAS-sc or UAS-da driven phenotype observed by co-expression with UAS-E(spl)-m8 or the UAS-E(spl)-m8 mutant derivatives must therefore result as a consequence of regulation by E(spl)-m8 at a level other than direct transcriptional repression of these genes. In addition a yeast two-hybrid analysis has been conducted to determine the molecular interactions between the E(spl)-m8 / E(spl)-m8 WRPW mutant derivatives and the Sc and Da proteins in an effort to distinguish between the different potential mechanisms of repression employed by E(spl)-m8.

The results show that ectopic co-expression of UAS-sc or UAS-da with UAS-E(spl)-m8 produces a phenotype identical to that produced by expression of UAS-E(spl)-m8 alone. In

other words, UAS-E(spl)-m8 is able to completely repress the extra bristle phenotype associated with expression of UAS-sc and UAS-da. A similar result is obtained at the level of A101 expression in the wing disc, in that the SOP cells corresponding to the bristles lost in the adult fly are also absent, indicating that the bristle-loss phenotype is manifest from the initial stages of development. The results strongly argue that E(spl)-m8 can represses the neural fate at levels other than just transcriptional repression of the proneural genes. Complete penetrance of the E(spl)-m8 bristle-loss phenotype in the co-expression assay further suggest that these other mechanism(s) are very efficient in repressing the neural fate.

The E(spl)-m8 WRPW mutant derivatives were also tested in the co-expression assay to determine their ability to suppress the UAS-sc driven phenotype. Co-expression of the UAS-E(spl)-m8 WRPW derivatives with UAS-sc resulted in a phenotype similar to that of ectopic expression of UAS-sc alone demonstrating that the WRPW motif of E(spl)-m8 is essential for repression of the UAS-sc driven phenotype.

As the ectopically driven E(spl)-m8 protein is able to suppress the UAS-sc and UAS-da driven phenotype in the absence of the ability to directly repress the expression of either transgene, it indicates that the protein mediates repression at levels other than direct transcriptional regulation of the proneural genes. Furthermore, demonstration that suppressor activity is lost by the substitution of the arginine residue in the WRPW motif to either histidine or aspartic acid suggests that that this residue within the WRPW motif is essential for this mode of repression. The results obtained in this study agree with data obtained previously in which *E(spl)-m8* and mutant derivatives thereof were co-expressed with the proneural gene *l'sc* (Giebel and Campos-Ortega, 1997). Giebel and Campos-Ortega (1997) found that ectopic co-expression of the wild-type UAS-E(spl)-m8 and UAS-l'sc in the wing disc during bristle SOP allocation led to a suppression of the UAS-l'sc phenotype, a result identical to the result obtained here using the *sc* proneural gene. Moreover, the study showed that precise removal of the WRPW motif or a substitution of the proline to a leucine residue in the WRPW motif abolished the suppressor function of the protein. These

observations, similar to those described here, indicate that the WRPW motif is essential for suppression of the UAS-l'sc driven phenotype.

It will be recalled that, in addition to direct transcriptional repression of the proneural genes three other potential models of repression were described: (i) sequestration of proneural activity by direct E(spl) binding, (ii) transcriptional repression of the proneural target genes, and (iii) transcriptional repression of the proneural target genes by a mechanism independent of DNA binding. The results from this and the previous study by Giebel and Campos-Ortega (1997) are consistent with one or more of these mechanisms being involved in E(spl) mediated repression of the neural fate. Given the data described above for the E(spl)-m8 WRPW mutant derivatives, coupled with the known function of this motif as a Gro interaction domain, a tentative prediction can be made that the mode of repression by E(spl)-m8 involves transcriptional repression of the proneural target genes mediated by interaction with the Gro protein rather than the titration of proneural activity. A caveat to this prediction is that the WRPW motif is required to facilitate interaction with the proneural proteins. Another line of evidence against a model of repression involving sequestration of proneural activity by HLH interaction with the E(spl) proteins comes from data obtained in the study by Giebel and Campos-Ortega (1997) using a second E(spl)-m8 derivative lacking the bHLH domain entirely. This derivative was still able to suppress the neural fate, albeit with reduced efficiency when compared to the wild-type E(spl) protein, suggesting that repression of the neural fate is mediated, at least in part, by a mechanism independent of interactions involving the HLH domain. A caveat to this interpretation is that interaction between the E(spl) and proneural proteins may not be mediated by the HLH domain,but by another region within the protein, the Orange domain may be important in this respect (see below).

A yeast two-hybrid analysis was also performed in this study to examine the molecular interactions between the E(spl)-m8 protein and the Sc or Da proteins and, to address the caveat outlined above, determine whether the WRPW mutant derivatives were unable to suppress the neural fate as a consequence of a change in ability to interact with the either the Sc or Da proteins. In agreement with a previous yeast two-hybrid study (Alifragis *et al.*,

1997) it was found that wild-type E(spl)-m8 exhibited strong interaction with Da but did not interact with Sc. When the E(spl)-m8 WRPW mutant derivatives were analysed for interaction with Da an interesting result was obtained. The E(spl)-m8RH derivative, similar to the wild-type E(spl)-m8 protein, demonstrated very high levels of interaction with Da. In contrast, the other two mutant derivatives, E(spl)-m8-RG and -RD, did not interact with Da. The two-hybrid results from the three E(spl)-m8 derivatives suggest that the arginine residue in WRPW is important, although not absolutely essential, for mediating interaction with Da. This is an interesting result in itself as the WRPW motif has not until this point, been implicated in the mediation of the interaction with any of the proneural proteins. In terms of the co-expression assay however the ability to interact with Da is not an issue because the E(spl)-m8-RH derivative, despite being able to interact with Da in a manner identical to the wild-type protein in the two-hybrid system, was unable to suppress the UAS-sc driven phenotype. The most parsimonious explanation for the loss of suppressor activity is that the arginine substitution in WRPW results in a disturbance of function which is not connected with ability to interact with Da. As described above one such function could be interaction with the Gro protein.

The data from this study and the previous studies described above therefore point towards a mode of repression involving transcriptional repression of the proneural target genes. If we accept that E(spl) repression is mediated at the transcriptional level of the proneural target genes then we have a choice between two models, one model which involves direct binding of the E(spl) protein to N-box sequences in the DNA and one which involves the recruitment of E(spl) to the promoter of the target genes by interaction with the DNA bound proneural proteins. In both models transcriptional repression is mediated by Gro which binds E(spl) via the WRPW motif. Distinguishing between these two different models is difficult. There is evidence however which suggests that E(spl) is able to suppress the neural fate in the absence of binding to DNA. Firstly, an E(spl) derivative has been generated in which the DNA-binding domain, the basic domain, has been neutralised (Nakao and Campos-Ortega, 1996). It was demonstrated that this derivative is unable to bind DNA *in vitro* as a consequence of the change made in the basic domain but is still able to suppress the neural

fate when tested *in vivo*. Secondly, the E(spl)-m8 derivative lacking the bHLH domain in the Giebel and Campos-Ortega study still demonstrates an ability to suppress the UAS-l'sc phenotype despite the absence of the basic domain. Another line of evidence supporting a mode of repression invoking a non-DNA-binding mechanism comes from the unpublished work of Delidakis *et al.* (pers. comm.). They have shown that an artificial proneural target gene consisting of a *lacZ* reporter gene connected to a number of E-box sequences, the sites at which the proneural proteins bind, is negatively regulated *in vivo* by E(spl) even though this construct does not contain E(spl) N-box binding sites. Regulation of this reporter gene was shown to take place on the DNA because an E(spl) fusion protein containing the VP16 activation domain led to an increase in activation of the reporter gene. These data suggest that E(spl) can be recruited to the DNA via interaction with one or more of the proneural proteins and subsequently transcriptionally repress the downstream gene, possibly in association with Gro. It has recently been reported, however, that the optimal E(spl) DNA binding motif may include sequences more similar to E-boxes than previously thought (Jennings *et al.*, 1999). In light of these findings, it will be necessary to determine whether E(spl) can bind the artificial proneural target gene *in vitro,* before the hypothesis proposed by Delidakis *et al.* can be accepted.

As the data obtained by Giebel and Campos-Ortega (1997) implies that the HLH domain is dispensable for suppressor function, the model proposed by Delidakis *et al.* would need to invoke a different proneural interaction domain in the E(spl)-m8 protein. In line with this, the Orange domain has been shown to be required *in vivo* for the repression function of Hairy and E(spl) in contexts related to sex determination and neurogenesis respectively (Dawson *et al.*, 1995; Giebel and Campos-Ortega, 1997). Although not formally tested, this data has led to the suggestion that the Orange domain in Hairy and E(spl) may be involved in mediating interaction with the proneural proteins. The yeast two-hybrid data obtained in this study would predict that Da, but not Sc, would be a target for interaction with E(spl)-m8. Furthermore, data from previous studies (Giebel and Campos-Ortega, 1997; Gigliani *et al.*, 1996) would suggest that the Ac protein and the L'sc protein are also unlikely targets for E(spl)-m8. The six remaining *Drosophila* E(spl) proteins, however, have different

interaction profiles with the proneural proteins suggesting that a mode of repression of this kind may be mediated by the other E(spl) proteins binding to other proneural partners. These differences could contribute to the functional specificity of the E(spl) proteins.

The observations made in this study, within the context of the GAL4-driven ectopic expression assay, strongly suggest that E(spl)-m8 mediated repression of the neural fate also occurs at levels other than direct transcriptional regulation of the proneural genes. It has been shown here, in agreement with Giebel and Campos-Ortega (1997), that the WRPW motif of E(spl)-m8 is essential to mediate this mode of repression indicating that Gro may also be involved and suggesting a model involving transcriptional repression of proneural target genes.

Chapter 7

Cloning and functional analysis of the *Musca domestica groucho* gene

INTRODUCTION

7.1. The *groucho* gene in *Drosophila* and other species.

In this chapter the cloning and characterisation of the *groucho* (*gro*) gene from the housefly (*Musca domestica*) is described and a preliminary functional analysis is performed. Using a degenerate PCR approach a region of the gene has been cloned which demonstrates a high degree of sequence identity to *Drosophila gro* at both the DNA and amino acid level. *in situ* hybridisation to *gro* revealed that the expression pattern in the *Musca* embryo closely resembles the pattern of expression in the *Drosophila* embryo. Finally, using the technique of RNA interference it has been shown that *gro* activity is required during neural development in the housefly embryo.

The *gro* gene in *Drosophila* was initially identified on the basis of its physical proximity to, and genetic interactions with, members of the E(spl)-C (Hartley *et al.*, 1988). Since then a number of *gro* genes have been identified in worms (Pflugrad *et al.*, 1997; Farida *et al.*, 1997), fish (Wylbeck and Campos-Ortega, 1997), rat (Schmidt and Sladek, 1993), mouse (Miyasaka *et al.*, 1993; Koop *et al.*, 1996) and human (the TLE family genes) (Stifani *et al.*, 1992) and shown to encode proteins which share a similar primary sequence structure to *Drosophila* Gro. Sequence similarity encompasses a series of carboxyl-terminal WD40-repeats, a highly conserved amino terminus and a variable region which links the two domains (figure 1.6).

gro plays a pleiotropic role during development in *Drosophila* and embryos lacking maternally contributed *gro* activity die during late embryogenesis as a consequence of defects in a number of developmental processes, including sex determination, segmentation, neurogenesis, dorsovental and terminal fate specification (Paroush *et al.*, 1994; Dubnicoff *et al.*, 1997; Paroush *et al.*, 1997). The requirement for *gro* throughout embryogenesis is reflected in the pattern of expression of the gene. In the early embryo of the syncytial and cellular blastoderm stage, maternally contributed message is detected ubiquitously at high

levels in the embryo which, as embryogenesis progresses, becomes more and more confined to the developing CNS (Hartley *et al.*, 1988).

A molecular role for Gro as a transcriptional co-repressor for the E(spl) bHLH proteins was initially hypothesised by Tata and Hartley (1993) based on sequence similarities between Gro and Tup1, a yeast transcriptional co-repressor containing WD40 repeats (see main introduction). Evidence for this hypothesis was provided by Paroush *et al.* (1994) who demonstrated genetic interactions between *gro* and a number of Hairy-related genes including the *E(spl)* genes, physical interactions between Gro and the products of the genes form these loci, and the requirement for *gro* in a number of developmental processes regulated by the Hairy-related proteins. Subsequent data obtained by Fisher *et al.* (1996) corroborated this hypothesis by showing that Gro and the mammalian homologues, the TLE proteins, were required for the repression activity of the Hairy-related proteins. It soon became clear however that Gro was also involved in transcriptional repression elicited by a subset of other, structurally unrelated transcription factors including, amongst others, Engrailed, Dorsal, the Runt domain family of proteins and Pangolin (*Drosophila* Tcf) (Jiménez *et al.*, 1997; Dubnicoff *et al.*, 1997; Aronson *et al.*, 1997; Cavallo *et al.*, 1998). The general theme that is emerging therefore is that the Gro family proteins are transcriptional co-repressors for a subset of genes which are involved in developmental regulation.

The housefly (*Musca domestica*) belongs, like *Drosophila*, to the Dipteran family and the evolutionary distance between the two species is approximately 100 million years (Beverley and Wilson, 1984). The characterisation of the housefly *E(spl)* genes is underway to use as a comparative tool for the evolutionary analysis of the complex between the two fly species (Duncanson, Glittenberg and Tata, unpublished data). To this end a number of E(spl) homologues from the housefly have already been isolated, including *E(spl)-mβ*, *E(spl)-mδ* and a gene which is very similar but distinct from *E(spl)-mβ*, which has been named *E(spl)-mβ-like*. Sequence analyses show that these genes demonstrate a high degree of similarity to their *Drosophila* counterparts especially in the regions of the bHLH and WRPW domains. A preliminary functional analysis for the *mβ*-like gene has been performed in *Drosophila*: the

gene has been introduced into the *Drosophila* genome by germline transformation and a functional assay was performed by driving ectopic expression and observing the phenotypic consequence (Duncanson, pers. comm.). Ectopic expression of the *Musca E(spl)-mβ-like* gene using the c591 driver line (see chapter 3 for a description of the c591 driver line) results in a loss of bristles from the notum and a loss of wing vein material (data not shown), a phenotype identical to that generated by ectopic expression of the *Drosophila E(spl)-mβ* gene (Ligoxygakis *et al.*, 1999). This experiment indicates that the *Musca* homolog of *E(spl)-mβ-like* is functionally equivalent to the *Drosophila E(spl)-mβ* gene when tested in *Drosophila.* However, transgenic analysis of this kind can be time consuming because a number of transgenic lines have to be generated, and additionally, heterospecific analyses may not reflect the true function of the gene in the species from which it originated. With this in mind an analysis of gene function which can be performed in the species from which the gene was originally isolated would be of great benefit. Such a method is now available with the recent discovery of RNA interference (RNAi).

7.2. RNA interference (RNAi).

RNAi describes a simple and rapid method for inhibiting specific gene function by the introduction (usually by direct injection into the animal or developing animal) of a double-stranded RNA (dsRNA) species corresponding to the gene of interest. The presence of a dsRNA effectively and specifically inhibits the activity of the endogenous gene to which it is directed. The phenomena of RNAi was uncovered in *C. elegans*, where it was shown for a large number of genes that injection of dsRNA into the adult worm could inactivate gene function (Fire *et al.*, 1998). The potency with which dsRNA interferes with endogenous genes was shown to be considerably higher when compared to either the sense or antisense RNA strand alone. A number of other interesting observations were noted: the effect of the interference was shown to be dependent upon only a few molecules of dsRNA per cell thereby arguing against a model of stochiometric interference with endogenous transcripts and implicating a catalytic or amplification process; the effect was systemic irrespective of the site of injection, indicating that dsRNA or the effect initiated by it is transported across

cellular boundaries; and inheritance of gene interference was sometimes observed in the progeny of injected animals.

The mechanisms involved in RNA-mediated interference are still unclear, however there are several lines of evidence to suggest the response occurs at a post-transcriptional level. Two initial observations pointed towards a post-transcriptional mechanism, firstly dsRNA containing intron sequences did not cause RNAi, and secondly *in situ* hybridisation showed that nascent transcripts corresponding to genes targeted by RNAi are briefly detected in the nucleus but quickly degrade and therefore cytoplasmic transcripts never accumulate (Fire *et al.*, 1998). More conclusive evidence comes from a study of the *lin-15* locus in worms. The *lin-15* locus is an operon coding for the two mature transcripts *lin-15a* and *lin-15b*; a single mutation in either *lin-15a* or *lin-15b* alone does not result in a mutant phenotype, however a mutation in both genes results in a worm which develops multiple vulva. If RNAi acted to prevent transcription then RNA corresponding to either *lin-15a* or *lin-15b* should produce the mutant phenotype because this locus produces a single primary transcript. Experiments have shown that injection of dsRNA directed to *lin-15a* or *lin-15b* alone does not result in a multi-vulval phenotype, in contrast, co-injection with both dsRNA species does result in the mutant phenotype (Mongomery *et al.*, 1998), suggesting that RNAi acts at a stage sometime after the mature transcripts are produced.

Traditionally the mechanism of delivery of dsRNA has been microinjection, however new delivery techniques such as feeding the worms on bacteria expressing dsRNA specific for a gene has been shown to be effective in inducing gene-specific phenotypes in the worms and their progeny (Timmons and Fire, 1998). Simple, large scale delivery techniques such as this have allowed genetic screens to be performed, and recently a number of genes necessary for implementation of the RNAi response have been determined. Tabara *et al.*(1999) identified four RNA interference-deficient mutants (*rde*). Interestingly, two of these loci, *rde-2* and *rde-3*, appear to be involved in the suppression of transposon activity and a third locus, *rde-1*, encodes a protein belonging to a large family, one member of which, the *sting* gene in *Drosophila,* is needed to silence the repetative germline specific *Stellate* locus

(Schmidt *et al.*, 1999). In addition, another group have reported that a large number of genes isolated in a screen for mutations which activate transposon activity are also resistant to RNAi (Ketting *et al.*, 1999). One of the mutations isolated in this second screen, the *mut-7* gene, encodes a protein with similarity to RNaseD. All these data suggest that the physiological function of RNAi may be to suppress transposon activity and indicate that RNAi might work by dsRNA directed, enzmatic RNA degradation.

More recently, this technology has been successfully transferred to *Drosophila* (Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999), *Trypanosoma* (Ngo *et al.*, 1998), *Planeria* (Sánchez Alvarado and Newmark, 1999), as well as to the plants *Nicotiana tabaccum* and *Oryza sativa* (Waterhouse *et al.*, 1998) and suggests that the technique may also become an effective tool for analysing gene function in organisms other than just*C. elegans*. Although the technique is at present widely used in *C. elegans* and is fast becoming an important technique in *Drosophila*, possibly the greatest application of RNAi will be as a tool to study gene function in organisms were more traditional genetic manipulations are not possible. For these reasons the technique of RNAi was used for the work carried out in this chapter, firstly to determine whether the technique of RNAi could be extended to analyse gene function in the housefly and secondly, and more importantly, in an attempt discover the functional role of the *gro* gene in this species of fly. The technique and experimental manipulations devised for RNAi analysis in *Musca* will be described below in detail.

RESULTS

7.3. A single copy a *gro* gene is present in the *Musca* genome.

Southern blot analysis was used to determine whether a *gro* homologue was present in the *Musca* genome. The analysis was performed using a number of restriction enzymes to help determine the copy number. Genomic DNA was isolated from the 'Cooper' strain and digested with either *Bam*HI, *Eco*RI, *Hin*dIII, *Xba*I or *Xho*I. The digested genomic DNA was separated on an agarose gel, blotted, and hybridised with a probe generated from a PCR

fragment corresponding to the complete coding region of the *Drosophila gro* gene. The filters were hybridised overnight at 65°C, and then washed at 65°C for 20 minutes in each of the following: 2x SSPE + 0.1% SDS, 1x SSPE + 0.1% SDS and again in 1x SSPE + 0.1% SDS. The results for the Southern blot analysis are given in figure 7.1. A single hybridising band is present in lanes were genomic DNA was digested with *Bam*HI, *Eco*RI, *Xba*I or *Xho*I indicating that a sequence with similarity to the *Drosophila gro* gene does indeed exist in the *Musca* genome. The presence of a single hybridising band in each of these lanes additionally suggests that the homologue is represented as a unique sequence within the genome. In contrast, DNA digested with *Hin*dIII produced two hybridising bands: a larger, less strongly hybridising band of approximately 8kb and a smaller, more strongly hybridising band of approximately 3.5kb. There are a number of possible explanations for the discrepancy between results obtained for the different restriction enzymes: the presence of two bands in the *Hin*dIII digest could be taken as evidence that there are two *gro* homologues in *Musca*, it could indicate that *Hin*dIII cleaves the gene into two fragments both of which hybridise with the probe, or alternatively, and probably more likely given the difference in intensity of the two hybridising bands, it could be a result of an incomplete *Hin*dIII digest with the larger band representing a partially digested DNA fragment and the smaller band representing the fully digested fragment.

7.4. Cloning the *Musca gro* homologue.

Having demonstrated by Southern blot analysis that a *gro*-like sequence exists in the *Musca* genome an attempt was then made to clone the gene. Several *gro* homologues have been cloned to date from organisms ranging from worms through to humans, and when alignments of the amino acid sequence are performed areas of high conservation are revealed at various regions along the length of the protein. The availability of such a resource made degenerate PCR an attractive approach to cloning the *gro* homologue from *Musca*. Another consideration which prompted the use of this approach was that only a portion of the coding sequence was required to perform the RNAi analysis. Additionally, obtaining a fragment of

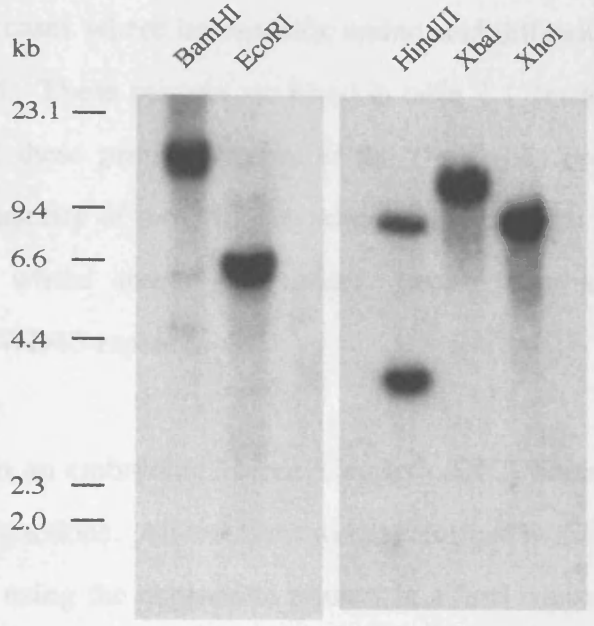


Figure 7.1. Southern blot analysis of Musca genomic DNA. Genomic DNA was extracted from the *Musca domestica* 'Cooper' strain, digested with *Bam*HI, *Eco*RI, *Hin*dIII, *Xba*I or *Xho*I, separated on a 0.8% agarose gel, blotted and hybridised with a ^{32}P-labelled probe generated from the complete coding sequence of the *Drosophila gro* gene. The filters were washed at 65°C for 20 minutes in 1x SSPE, 0.1% SDS and exposed overnight at -80°C.

Musca gro using a PCR approach would then greatly facilitate subsequent efforts to screen a cDNA library to isolate a full length cDNA clone.

An alignment of ten amino acid sequences from multiple species (fly, mouse, rat and human) (Tata and Hartley, 1993) was used as the basis to design a number of degenerate primer sets. The *Musca domestica* codon usage was figured into the design of the primers where appropriate and in cases where interspecific amino acid differences occured the *Drosophila* sequence was used. These primers are listed in table 2.1 (materials and methods section) and the position of these primers relative to the *Drosophila* coding sequence is shown in figure 7.2. The majority of the 'Mugro' primers cluster to the 3' end of the gene, as this encodes a region where amino acid identity between species is greatest (the region corresponds to the WD40-repeat motif).

cDNA isolated from an embryonic *Musca* 'Cooper' cDNA library was used as the template DNA for the PCR reactions. All reactions were performed in a 20µl volume using 2µl of the template DNA and using the degenerate primers at a final concentration of 2µM each. The following PCR conditions were chosen for 30 cycles: 1 minute denaturation at 94°C; 1 minute annealing at 50°C; 1 minute elongation at 72°C.

PCR reactions were set up using the following pair-wise combinations of each primer: Mugro1 + Mugro5, Mugro1 + Mugro6, Mugro3 + Mugro5, Mugro3 + Mugro6, and Mugro4 + Mugro6. In addition, a negative control (without template DNA) for each primer set was performed (data not shown) and, in order to detect products resulting from spurious amplification by a single primer alone, reactions containing just one of each of the primers were also performed. The results from this experiment are presented in figure 7.3. Selection of potentially positive products was carried out based on size (as estimated from the size of product expected from amplification from *Drosophila* cDNA) and that all associated control reactions were negative. A product resulting from amplification using primer set Mugro3 + Mugro6 was of approximately the correct size (770bp) when compared to the *Drosophila* sequence, the control reaction without DNA was negative and amplfication

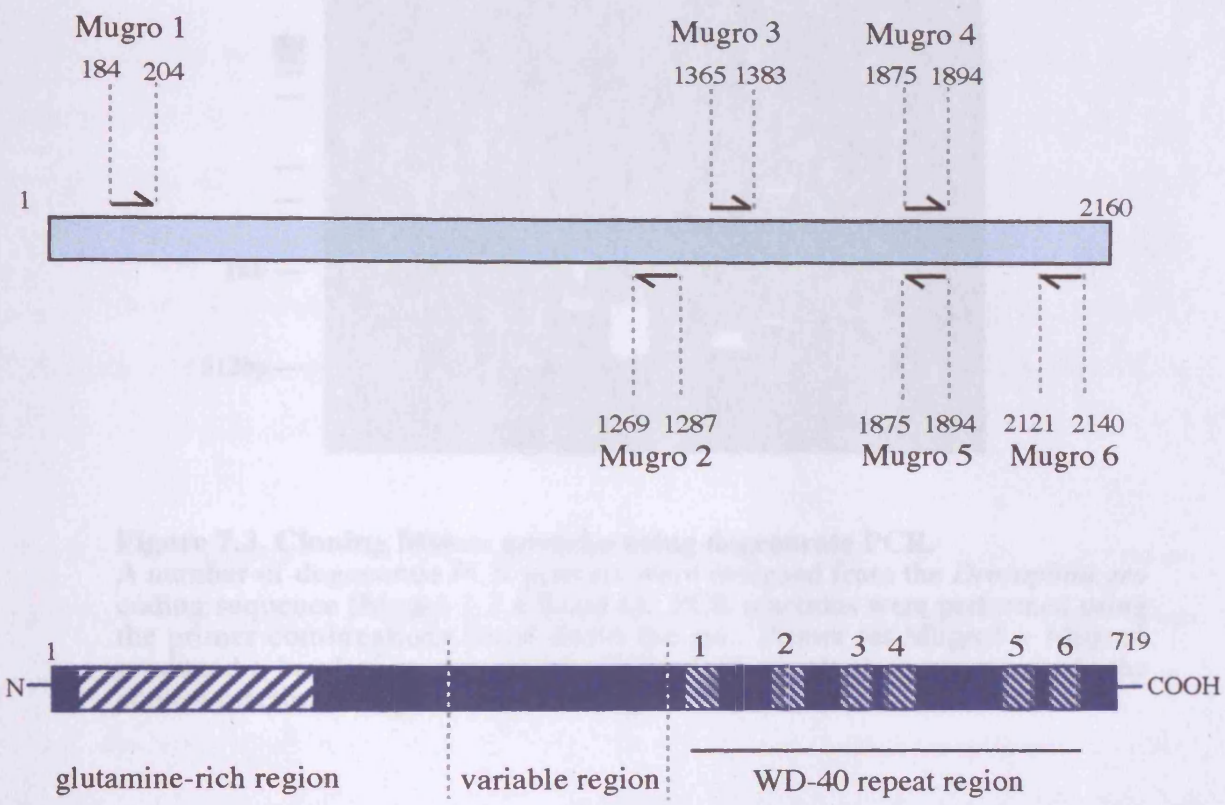


Figure 7.2. *Musca groucho* degenerate PCR primers.
An alignment of ten amino acid sequences from multiple species (Tata and Hartley, 1993) was used to design six degenerate primers (Mugro 1 to Mugro 6). Shown is the *Drosophila* coding sequence (upper) and the position of the six primers on the sequence. The *Drosophila* primary protein sequence is shown below and the positions of the glutamine-rich domain, the variable region and the WD40 repeats are given. The majority of the Mugro primers map to the 3' end of the gene because this encodes a region where amino acid identity between species is the greatest.

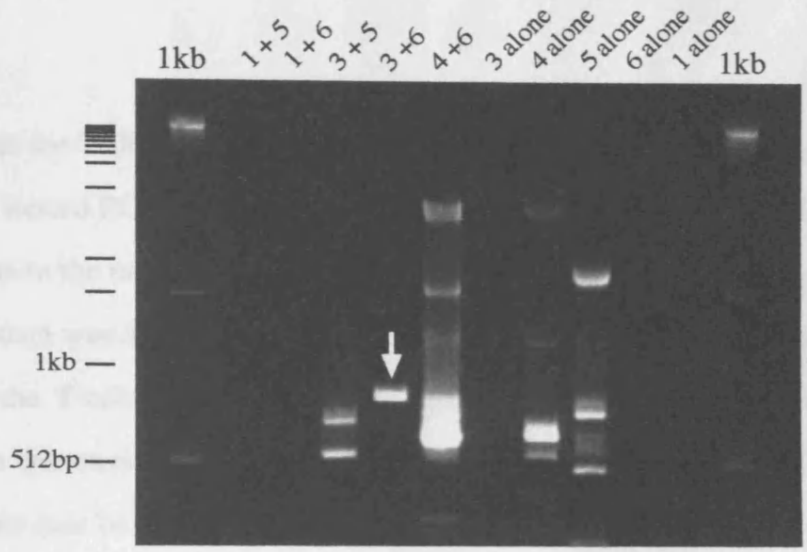


Figure 7.3. Cloning Musca groucho using degenerate PCR.
A number of degenerate PCR primers were designed from the *Drosophila gro* coding sequence (Mugro 1,3,4,5 and 6). PCR reactions were performed using the primer combinations listed above the gel. Primer set Mugro3 + Mugro6 produced a band of approximately 770bp (arrow) which is not present in the control reactions Mugro3 alone or Mugro6 alone.

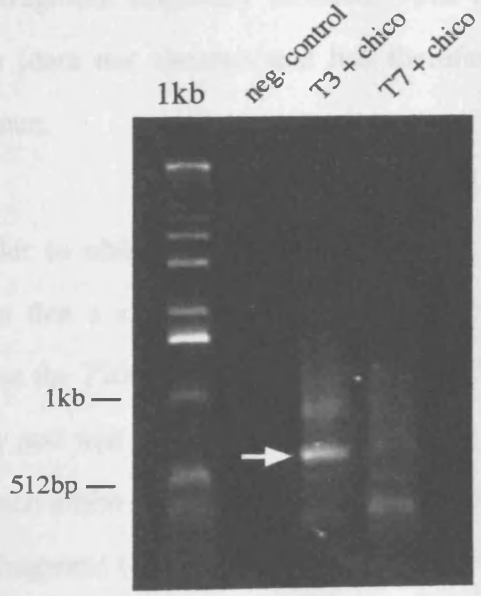


Figure 7.4. PCR cloning of the 5' sequence of *Musca gro*.
The chico primer was designed from the 5' sequence of the 770bp fragment obtained in the degenerate PCR. Priming sites for the T3 and T7 universal primers are present in the vector in which the library is cloned. PCR reactions were performed using the primer combinations listed above the gel. Primer set T3 + chico produced a band of approximately 700bp (arrow).

with Mugro3 or Mugro6, when tested individually, did not result in spurious products. None of the other primer combinations yielded products which fitted the criteria set out above.

To validate that the 770bp product did correspond to *gro*, the band was isolated and used as a template for nested PCR using primer sets Mugro4 + Mugro6 and Mugro3 + Mugro5. The two primer sets in the nested PCR both produced bands of the expected size confirming that the 770bp product was indeed *gro* (data not shown). The 770bp product was subsequently cloned using the T-tailed pBluescript vector (see materials and methods) and sequenced. Sequence data demonstrated that the 770bp PCR product corresponded to a portion of the *Musca gro* gene (see below).

The 770bp fragment of *Musca gro* was used to generate a ^{32}P-labelled probe to screen an embryonic *Musca* λZap cDNA library. A single clone was isolated from the library which contained an insert of approximately 3kb. Sequence analysis however demonstrated that the clone was rearranged and contained only part of the *gro* sequence along with another unidentified sequence. The region of the clone corresponding to *gro* was smaller than the PCR fragment originally isolated. The clone isolated includes the entire 3' untranslated region (data not shown) and has therefore allowed determination of the entire 3' coding sequence.

In order to obtain further 5' sequence a different approach was therefore taken. It was known that a clone larger than the one isolated from the screen did exist in the library because the 770bp fragment obtained by PCR was derived from a pooled preparation of that library and was larger than the rearranged *gro* clone isolated from the screen. For the new approach a new reverse primer (chico) was designed from the sequence at the 5' end of the PCR fragment (table 2.1, materials and methods section) and used in combination with either Mugro1 or the T3 or T7 universal primers (priming sites for T3 and T7 are present in the vector in which the library is cloned) to amplify cDNA extracted from the library. Using this approach, a product of approximately 700bp was obtained (figure 7.4). This fragment was

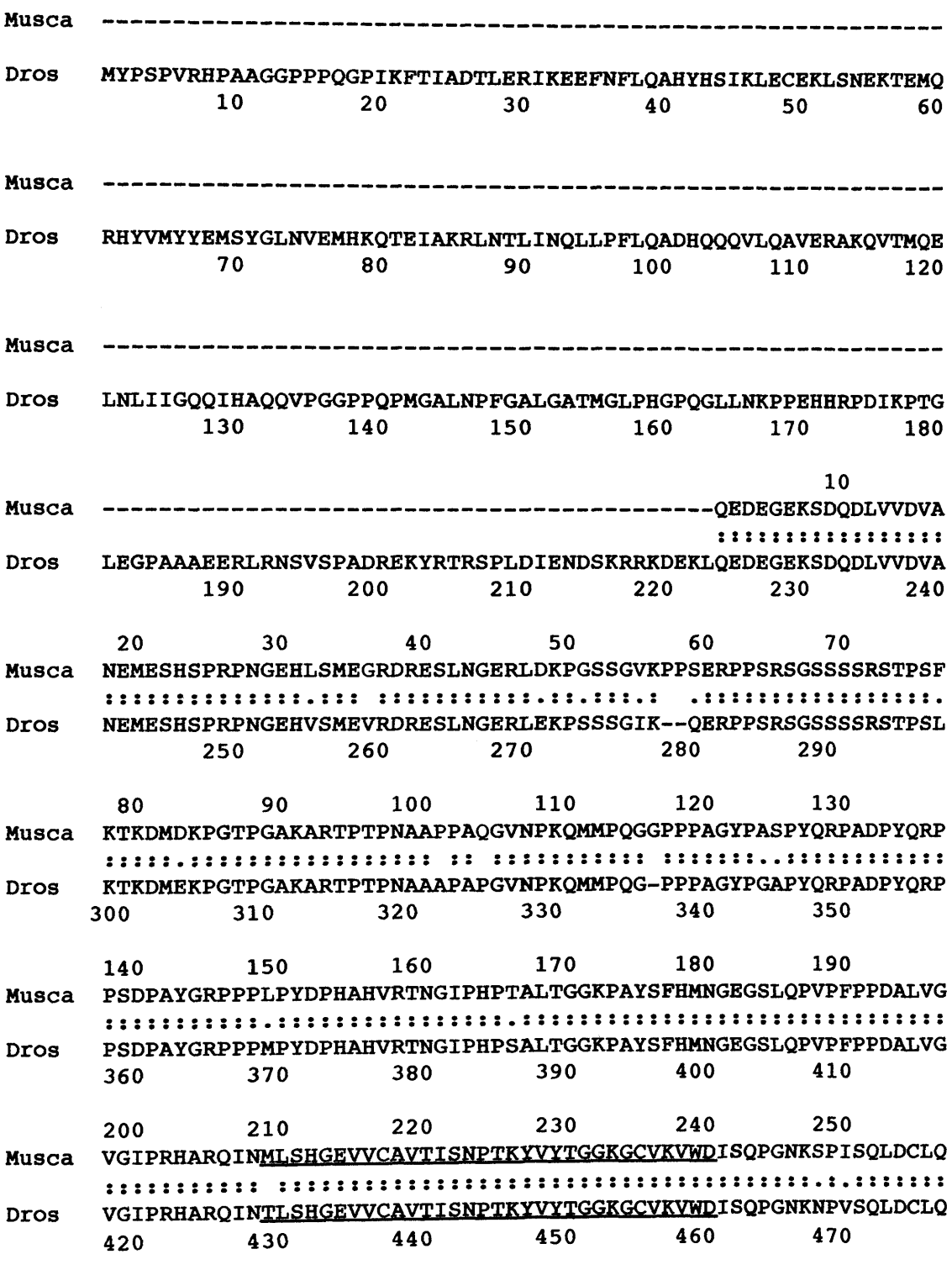
then cloned using T-tailed pBluescript vector and sequenced. Sequence data demonstrated that the 700bp PCR product corresponded to a portion of the *Musca gro* gene.

A total of 1497bp of *Musca gro* coding sequence has been obtained using the two PCR-based approaches described above. This fragment corresponds to the 3' end of the *gro* gene and includes the region encoding the six carboxy-terminal WD40 repeats and a portion of the variable domain. At the level of the DNA sequence the 1497bp fragment from *Musca* shares 53% identity with the corresponding region of *Drosophila gro* (see appendix 3). The predicted primary protein sequence encoded by the 1497bp fragment is highly similar to that of the corresponding region of the *Drosophila* protein, the approximate overall identity is 95% (figure 7.5). Comparison of the sequence to corresponding regions of other Gro family members such as the *C. elegans* UNC-37 and human TLE1 proteins reveal that the highest level of conservation occurs in the WD40 repeat domain (59% for *C. elegans* and 89% for humans), whereas the region outside this domain is more variable (25% *for C. elegans* and 44% for humans) (figure 7.6).

7.5. RNA expression pattern of the *gro* gene in *Musca* embryos.

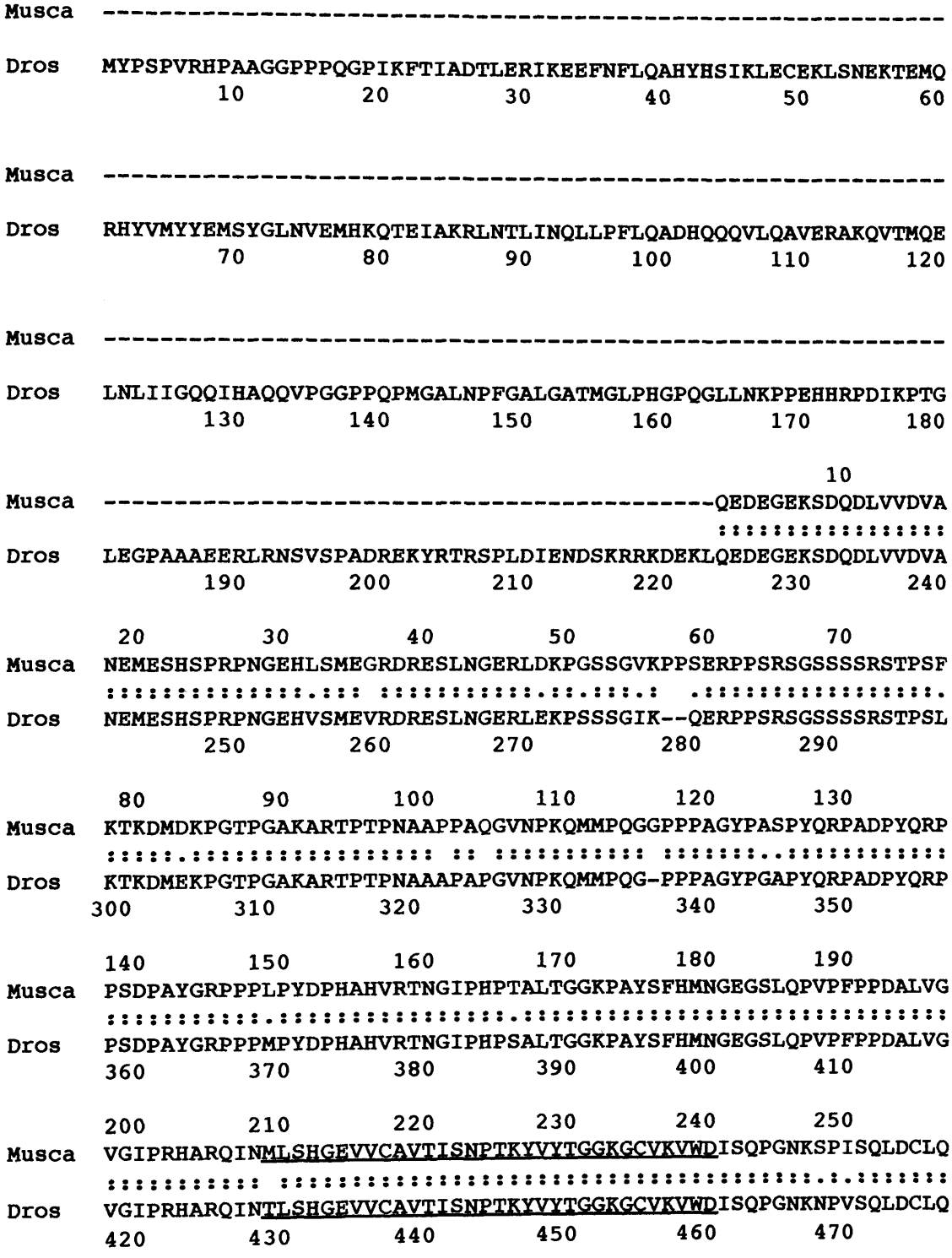
The first step in the analysis of the function of the *Musca gro* gene was to determine its expression pattern by *in situ* hybridisation to *Musca* embryos. The 770bp *Musca gro* clone in pBluescript II SK$^-$ was linearised with *Xho*I to generate a template for the antisense probe and with *Eco*RI to generate a template for the sense probe. DIG-labelled RNA probes were then synthesised *in vitro* using T7 and T3 polymerase for the antisense and sense probe respectively. Embryos were collected overnight on egg-laying plates supplemented with a chunk of cat food (Purrfect® rabbit and chicken flavour), harvested and subsequently manipulated in a manner identical to *Drosophila* embryos. The results of the *in situ* hybridisation are presented in figure 7.7. Very high levels of ubiquitous *gro* expression is detected throughout syncytial and blastoderm embryos which probably corresponds to maternally contributed message. In later embryos of the extended germband stage the expression level is not as high as in the early embryo and is no longer uniformly distributed,

Figure 7.5. ***Musca*** **v** ***Drosophila*** **Gro protein alignment.**



contd.

Figure 7.5. *Musca* v *Drosophila* Gro protein alignment.



contd.

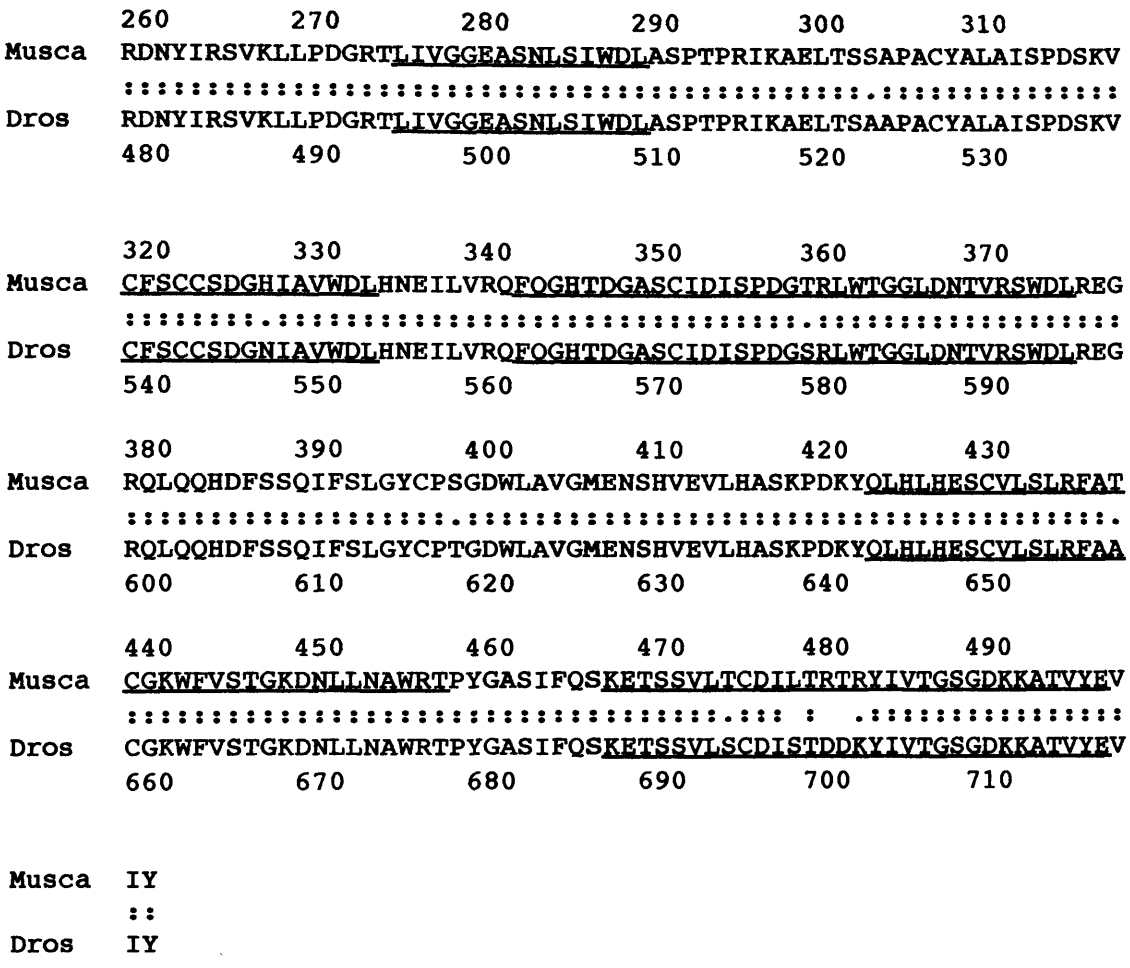


Figure 7.5. *Musca* v *Drosophila* Gro protein alignment.
Figure shows an aligment of the predicted protein sequence of *Musca* Gro against the *Drosophila* Gro protein. Identical residues are indicated in black, similar residues in blue and non-similar residues in red. The six WD40 repeats are underlined. The alignment was performed using FASTA.

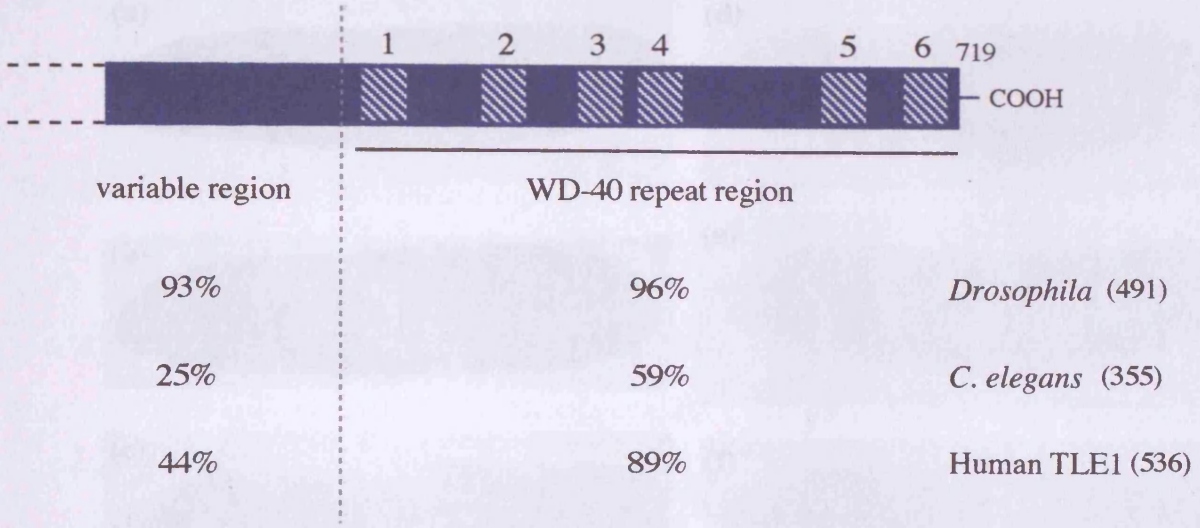


Figure 7.6. Schematic representation of the *Musca gro* protein and similarity with other Gro family members.
A schematic diagram of the product encoded 1497bp region of *Musca gro* obtained in this study is shown. The variable region and the conserved WD40-repeat region are delineated on the figure. The percentage identity of *Musca gro* to *Drosophila gro*, *C. elegans unc-37* and Human TLE1 is shown for the amino acid sequence for the two regions.

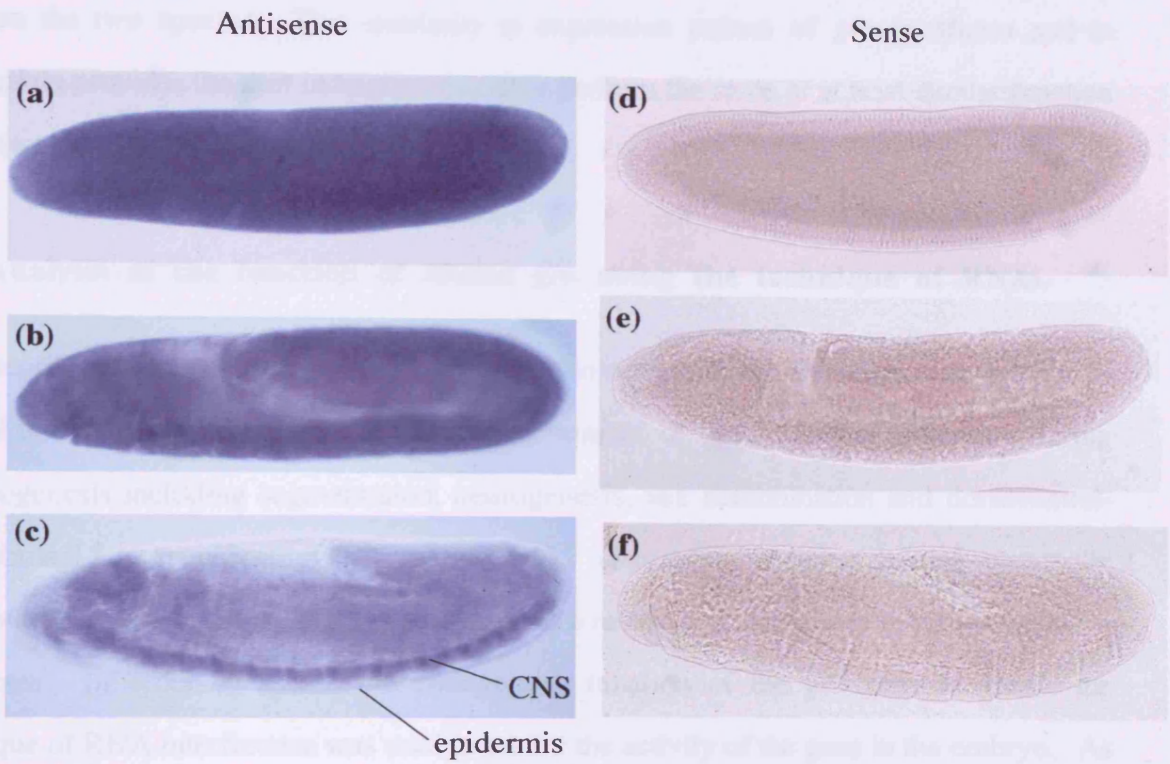


Figure 7.7. RNA expression pattern of *gro* in *Musca* embryos.
Musca embryos were hybridized overnight with an antisense (a,b,c) or sense (d,e,f) DIG-labelled RNA probe corresponding to the 770bp fragment of *Musca gro*. (a) Very high levels of ubiquitous *gro* mRNA is detected in the syncytial blastoderm embryo. (b) In the extended germ band embryo (equivalent to *Drosophila* stage 9) *gro* expression is lower relative to the early embryo and is no longer distributed uniformly. (c) By the retracted germband stage (equivalent to *Drosophila* stage 13) expression is confined to the differentiating CNS and at this point epidermal staining is no longer detected. (d,e,f) Staining is never observed at any stage of development with the DIG-labelled sense RNA probe. In all cases anterior is right and dorsal is up.

with the highest level of expression in the germband (which corresponds to the developing CNS). After retraction of the germband (equivalent to *Drosophila* stage 13) the transcript becomes confined to the differentiating CNS, and at this stage epidermal expression is no longer detected. No staining is observed at anytime during embryonic development with the sense probe. Comparison of the embryonic expression patterns of *gro* in *Musca* with those described for *Drosophila* (Hartley *et al.*, 1988) reveals that no substantial differences exist between the two species. This similarity in expression pattern of *gro* in *Musca* and in *Drosophila* provides the first indication that they perform the same or at least similar function in the two fly species.

7.6. Analysis of the function of *Musca gro* using the technique of RNAi.

In *Drosophila* the Gro protein is required as a co-repressor for the repression activity of several transcriptional regulators involved in a number of developmental processes during embryogenesis including segmentation, neurogenesis, sex determination and dorsoventral- and terminal-fate specification (see section 7.1). *Drosophila* embryos lacking maternally contributed *gro* die during the embryonic stage as a result of disturbances in all of the above processes. In order to assess the endogenous function of the *gro* gene in *Musca* the technique of RNA interference was used to inhibit the activity of the gene in the embryo. As RNAi affects both zygotic and maternal gene activity (Kennerdel and Carthew, 1998) it provides an excellent tool with which to remove all *gro* activity from *Musca* embryos, thus creating a true null phenotype. To determine the consequences of eliminating *gro* function using RNAi two different assays were used. The first assay simply involved scoring the number of embryos which hatched as larvae subsequent to injection with either dsRNA or with injection buffer alone (control embryos). The second assay made use of the anti-Horseradish Peroxase (α-HRP) antibody which recognises an epitope in both CNS of *Drosophila* and of the grasshopper (Jan and Jan, 1982) and, as shown in this study, also recognises an equivalent epitope in *Musca*. The α-HRP antibody has been previously utilised in *Drosophila* to reveal neural defects associated with mutations in a number of neurogenic loci including *gro* (Preiss *et al.*, 1988). The use of this antibody in the RNAi

experiment therefore provides a more specific assay to determine whether the product of the *gro* gene performs a neurogenic function in the *Musca* embryo similar to the role it plays in the *Drosophila* embryo.

7.6.1. α-HRP antibody staining of Musca embryos.

The α-HRP antibody labels all CNS neurons and the brain in both *Drosophila* and the grasshopper (Jan & Jan, 1982). This cross-specificity suggested that the antibody may also recognise an epitope in the *Musca* CNS and would therefore be a useful tool for the RNAi analysis. In order to test this an overnight collection of *Musca* embryos was stained with the α-HRP antibody. The results from this staining are given in figure 7.8 and show that the antibody specifically labels the CNS neurons and the brain in the housefly as it does in *Drosophila*. The staining pattern in the housefly is largely identical to the staining pattern observed in *Drosophila*.

7.6.2. Preparation of the double-stranded RNA.

The 770bp *Musca gro* clone in pBluescript II SK$^-$ was used to generate dsRNA. The plasmid was linearised with either *Eco*RI or *Xho*I to generate the template for the sense and antisense strands respectively. The sense and antisense transcription reactions were performed separately *in vitro* and annealing was then carried out by mixing equimolar amounts of each strand, boiling for one minute in a large beaker of water, after which the beaker was removed from the heat and allowed to cool to room temperature over a period of approximately 18 hours. After 18 hours a small aliquot was checked on a agarose gel to (i) determine that annealing had occurred (Figure 7.9) and (ii) to determine the approximate concentration of the dsRNA (see materials and methods).

7.6.3. The dsRNA injection procedure.

In order to determine the optimal site of injection a number of embryos were injected at various positions with buffer alone and the percentage survival after injection at each position was calculated (data not shown). From this experiment it was determined that an anterior region on the dorsal side of the embryo was the optimal site because injection in this region

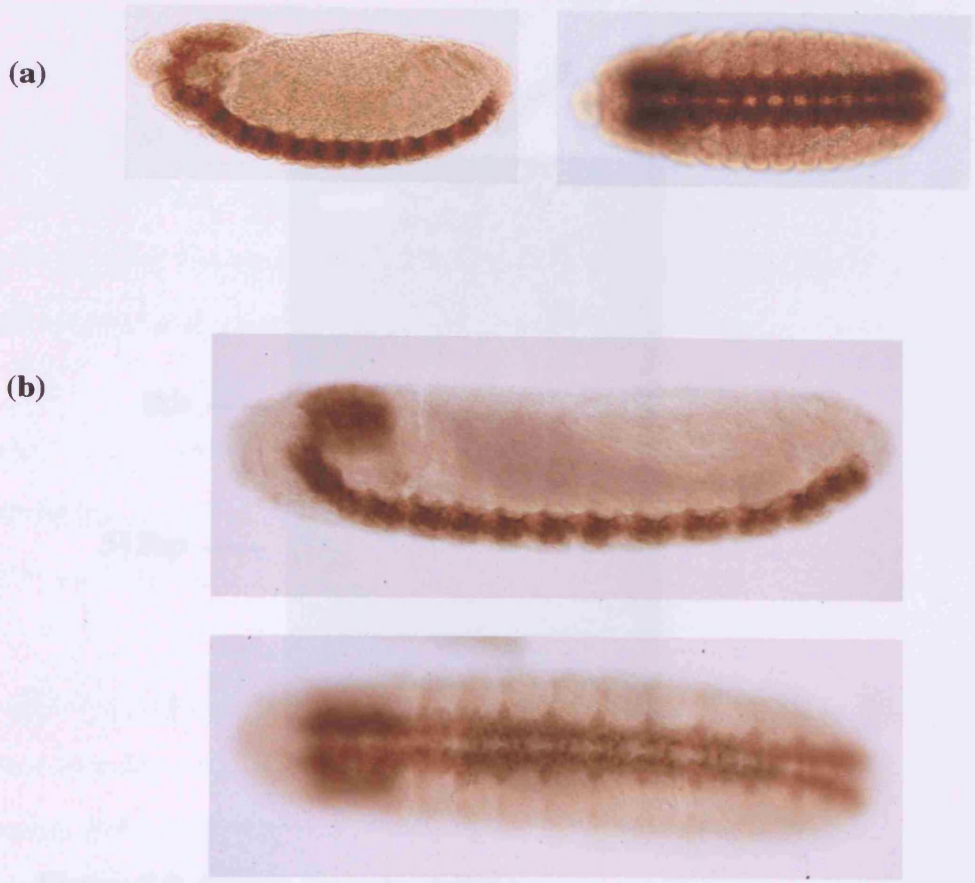


Figure 7.8. α-HRP staining patterns of stage 13 *Drosophila* and *Musca* embryos. *Drosophila* or *Musca* embryos were hybridized for 1-2 hours with the α-HRP antibody. (a) Lateral and ventral view of wild-type stage 13 *Drosophila* embryo. (b) Lateral and ventral view of wild-type retracted germband *Musca* embryo. The antibody specifically labels the CNS neurons and the brain in both *Drosophila* and *Musca*. The staining pattern in the two fly species is largely identical. In all cases anterior is right.

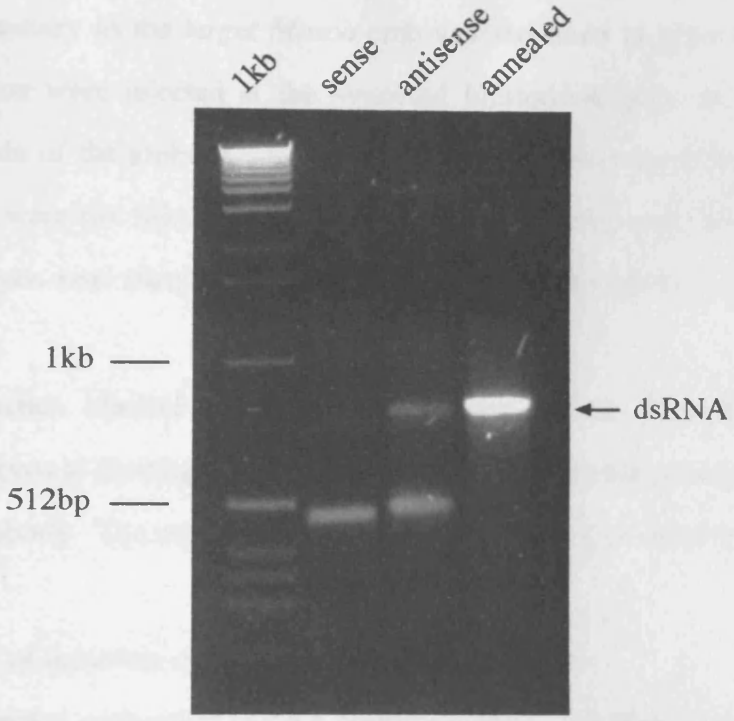


Figure 7.9. *Musca groucho* dsRNA.
A linearized pBluescript vector containing the 770bp fragment of *Musca gro* was used as the template to synthesize the RNA transcripts. Transcription reactions were performed separately *in vitro* using T7 and T3 polymerases to generate the sense and antisense transcripts respectively. Equimolar amounts of the transcripts were mixed, boiled for 1 minute in a large beaker of water and then allowed to cool to room temperature over a period of approximately 18 hours. A product migrating at the expected size (770bp) is observed after the annealing reaction is performed (arrow).

resulted in the lowest percentage lethality. Similar findings have been observed by a colleague (Salameh, pers. comm.).

Musca embryos were collected at approximately two hourly intervals on egg laying plates supplemented with a chunk of cat food. Embryos were harvested from the plate, dechorionated and aligned directly onto a glass slide. It was discovered that the use of sticky tape was not necessary as the larger *Musca* embryos remained in place on the slide when injected. Embryos were injected at the syncytial blastoderm stage in a position on the anterior dorsal side of the embryo. Embryos which had progressed beyond the syncytial blastoderm stage were not injected and were subsequently removed. The slides containing the injected embryos were then allowed to develop in humid conditions at 18°C.

As mentioned earlier, injected embryos were assayed in two different ways, either by allowing the embryos to develop and scoring the number of larvae which hatched or by use of the α-HRP antibody. The results from the two methods will be described below.

7.6.4. The effects of injection of gro dsRNA on survival.

Embryos were injected with either dsRNA corresponding to the *Musca gro* fragment or with injection buffer alone and then allowed to develop at 18°C for approximately 60 hours (embryonic development of *Musca* at 18°C is calculated to take approximately 50-60 hours, therefore all embryos that are going to hatch will have done so after 60 hours). The data from this experiment are given in table 7.1a. It is clear from the values presented in table 7.1a that injection of *gro* dsRNA results in a higher percentage of embryonic lethality than those embryos injected with buffer alone. Out of a total of 142 embryos injected with the dsRNA only 19 larvae were recovered corresponding to a survival rate of only 13%. This contrasts with the data obtained for the mock injected embryos (injection of buffer alone), where 128 embryos were injected and a total of 77 larvae were recovered corresponding to a survival rate of 60%. Injection with buffer alone does result in a significant degree of embryonic lethality (40%). The most likely explanation for this is that the process of microinjection itself causes trauma which ultimately results in termination of development.

		Buffer	dsRNA
(a)			
	Total injected embryos	128	142
	Number of survivors	77	19
	Survival rate (%)	60.2	13.3
(b)			
	Total injected embryos	160	165
	Number non-staining	71	76
	Number wild-type staining	89	12
	Number abnormal staining	0	77

Table 7.1. RNAi of *Musca groucho*.
Syncitial blastoderm *Musca* embryos were injected with either buffer (mock injected embryos) or with dsRNA. (a) The effects of the RNAi were assayed by comparing embryonic survival rate of embryos injected with dsRNA to the mock injected embryos. (b) The effects of the RNAi were assayed by staining with the αHRP antibody and comparing the staining pattern between embryos injected with dsRNA or buffer alone (lower panel).

From these results it is clear that injection of *gro* dsRNA into the *Musca* syncytial blastoderm embryo results in embryonic lethality.

7.6.5. The effects of injection of gro dsRNA on the α-HRP staining pattern.

Although the above experiment shows that introduction of *gro* dsRNA into *Musca* embryos results in embryonic lethality, as expected if the role of *gro* in *Musca* is equivalent to that in *Drosophila*, it does not provide any indication of specific roles for the gene in *Musca*. In order to investigate the possible involvement of *gro* during neurogenesis, *Musca* embryos were injected with dsRNA and then subsequently stained with the α-HRP antibody to reveal the morphology of the CNS. Injection of the embryos was carried out using either dsRNA or buffer alone in the manner described above and post-injected embryos were allowed to develop for a period of 20-24 hours at 18°C. It was determined that a period of between 20-24 hours development at 18°C is the equivalent to stage 13 in *Drosophila* development (Campos-Ortega and Hartenstein, 1985), at the point when the germband becomes fully retracted. After 20-24 hours the embryos were stained using the α-HRP antibody. Embryos injected with dsRNA or mock injected embryos were stained in an identical manner side-by-side in the same experiment.

The data from this experiment are given in table 7.1b and in figure 7.10. A total of 160 mock injected embryos were stained with the α-HRP antibody, of this number 71 embryos did not stain at all (44%). These non-staining embryos probably correspond to those which abort early in development as a result of trauma brought on by the microinjection itself and therefore do not start to undergo neural development and hence do not express the neural epitope recognised by α-HRP. The remaining 89 embryos (55%) all displayed wild-type staining patterns equivalent to the staining pattern observed with uninjected *Musca* embryos. Abnormal staining patterns were not observed in mock injected embryos. A total of 165 embryos were injected with dsRNA and stained with α-HRP, of these 76 embryos failed to stain (46%). Again, these non-staining embryos probably correspond to those which abort early in development. Of the remaining 89 embryos, 77/89 (46% of total injected) display abnormal phenotypes which fall into a number of different classes (figure 7.10). The most

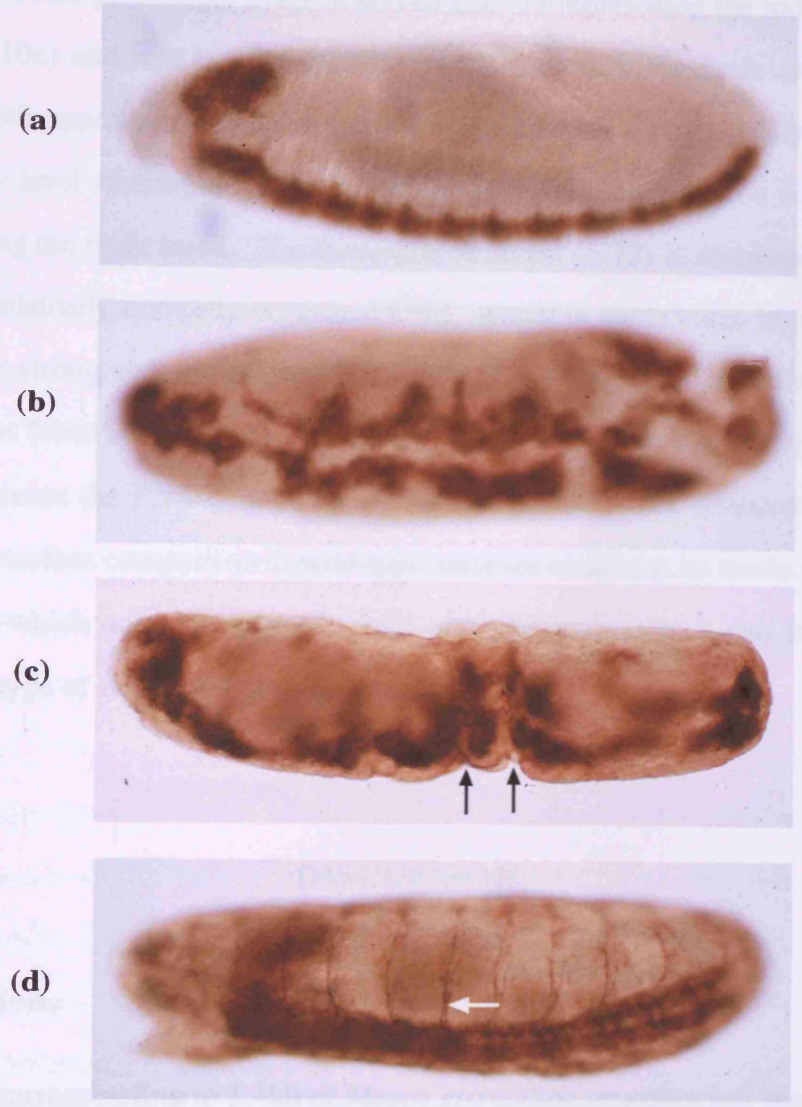


Figure 7.10. The effects of *gro* dsRNA on neurogenesis in *Musca*.
Syncytial blastoderm embryos were injected with either buffer alone (a) or *gro* dsRNA (b,c,d), allowed to develop for 20-24 hours at 18°C and hybridized with an α-HRP antibody. (a) Lateral view of a retracted germband embryo showing a normal CNS morphology. (b) Ventral view showing severe disruption and fragmentation of the CNS. (c) Lateral view showing strong hyperplasic defects in the brain and ventral nerve cord and low level staining throughout the embryo. Restrictions along the main trunk of the embryo are seen in some cases (arrows). (d) Lateral view showing staining of the PNS (arrow). In all cases anterior is right and dorsal is up.

frequent class (47/77) observed display disruptions in the CNS when viewed from the ventral surface, which is manifest by a mildly hypertrophied ventral nerve cord in addition to minor disruptions including gapping and fragmentation along its length (compare 7.8d with 7.10b). The second phenotypic class observed (22/77) represented the most severe defect seen (figure 7.10c) and was typified by hyperplasic defects in the brain and ventral nerve cord, which both appear disorganised when viewed laterally. Additionally these embryos often show low level staining throughout the whole of the embryo and a number also have restrictions along the main trunk. The third class of defect (8/77) is represented by embryos which have a relatively normally organised CNS, which in some cases is mildly enlarged, but which have strong staining of the PNS which is not usually observed with the α-HRP antibody (figure 7.10d). It is unclear whether the morphology of the PNS in these animals is abnormal because the PNS of wild-type *Musca* embryos is not revealed by the α-HRP antibody and therefore comparison to wild-type embryos could not be made. Twelve out of the 77 embryos which stained displayed a wild-type phenotype which was indistinguishable from the phenotype of embryos injected with buffer alone.

DISCUSSION

7.7. Discussion.

A partial clone corresponding to 1.4kb of *Musca gro* coding sequence has been cloned using a PCR approach. The portion of *gro* cloned encodes part of the variable region and the carboxl-terminal WD40 repeats. Comparisions between these two domains and the coresponding domains in other species reveals that the WD40 repeat domain is highly conserved, whereas the variable region, as its name suggests, is less well conserved. In order to make a comprehensive comparison between *Musca gro* and the other previously cloned homologues the full coding sequence is required, this is especially true if function is to be inferred from the primary sequence data. Cloning of the remainder of the molecule is presently underway using a number of techniques including 5' RACE and suppression PCR.

In addition to providing ability to conduct a complete comparative analysis, obtaining a full length cDNA would also enable a more extensive functional analysis to be performed. It has previously been demonstrated using yeast two-hybrid and *in vitro* analyses that *Drosophila* Gro and a number of Gro homologues physically interact with a range of transcriptional regulatory proteins such as the Hairy-related proteins, Engrailed, Dorsal and Runt. Similar analyses could be performed using *Musca* Gro to determine whether interactions with these proteins also occur. An obvious starting point for such an experiment would to be assay interaction with the *Musca* E(spl) homologues, a number of which have already been cloned (Duncanson and Tata, unpublished data). A second experiment which could be carried out (given the availability of a full length cDNA) is to determine whether or not a *Musca gro* transgene would be able to rescue lethality in *Drosophila* embryos lacking *gro*. Such an experiment would indicate whether the function of the *Musca gro* gene has been conserved in all developmental contexts, such as sex determination, segmentation and dorsoventral/terminal fate specification over an evolutionary distance of approximately 100 million years.

The pattern of expression of *gro* in *Musca* embryos was determined by *in situ* hybridisation and was shown to closely resemble the pattern of expression previously described for *Drosophila* embryos (Hartley *et al.*, 1988). During early development ubiquitous maternal expression is observed throughout the syncytial and cellular blastoderm embryo, then, as development continues, the transcript becomes progressively more confined to the CNS, and by the time the germband has retracted (equivalent to *Drosophila* stage 13) the transcript is found exclusively in the CNS. In the *Drosophila* embryo a similar expression pattern reflects the pleiotropic role *gro* plays in numerous developmental contexts, from a point very early in development during sex determination, through segmentation, to later stages when it is required for neurogenesis. The neural pattern of expression observed in later *Musca* embryos strongly implies that *gro* plays an equivalent role during *Musca* CNS development. The earlier patterns of *gro* expression in *Musca* are largely identical to the patterns in *Drosophila* so it is also tempting to predict that *gro* plays a role in sex determination and in the patterning of the early embryo. The latter prediction could be tested by using the

technique of RNAi to inhibit *gro* activity and assaying for segmentation defects by, for example, staining embryos using an antibody which recognises one of the segmentation genes, such as the segment polarity gene *engrailed* (Patel *et al.*, 1989).

RNAi describes a technique of genetic interference whereby introduction of dsRNA corresponding to a gene of interest is able to specifically disrupt the activity of the endogenous gene. RNAi was used for a functional analysis by observing the consequences of removing *gro* activity from the *Musca* embryo. It was shown that injection of *gro* dsRNA at the syncytial blastoderm stage resulted in an increase in embryonic lethality when compared to control embryos which had been injected with buffer alone. This result is consistent with the lethal phenotype associated with removal of maternal *gro* activity in *Drosophila* which results as a consequence of disruptions in embryonic patterning and in neurogenesis (Paroush *et al.*, 1994). One caveat with the RNAi experiment conducted in this study is that the lethal phenotype observed after injection of dsRNA may not be a specific consequence of introduction of *gro* dsRNA but may represent a more general response elicited by the presence of a foreign RNA species. A suitable control to test for non-specific effects caused by introduction of dsRNA would be to inject dsRNA corresponding to a benign target or to a target which does not exist in *Musca* such as *lacZ*.

Drosophila embryos which lack maternally contributed *gro* exhibit neural hypertrophy with varying degrees of penetrance (Priess *et al.*, 1988; Schrons *et al.*, 1992; Paroush *et al.*, 1994). To determine whether *Musca gro* plays a similar role in neural development the phenotype of *Musca* embryos injected with dsRNA was assayed using the α-HRP antibody. Injection of the dsRNA resulted in a number of neuronal defects ranging from an increased staining of the PNS to more severe disruptions in the CNS. The mutant phenotypes observed in these injected *Musca* embryos are not as extreme as the neural hypertrophy phenotype described for *Drosophila* embryos homozygous for the *gro* alleles $l(gro)^{X1}$ and $E(spl)^{BX22}$ (Preiss *et al.*, 1988). The reason for this is probably related to the fact that these *Drosophila gro* alleles result from extensive genomic deletions removing other genes including some of the *E(spl)* transcription units, thereby causing a phenotype more severe

than that associated with lesions in *gro* alone. In line with this, embryos homozygous for the gro^{E107} allele (point mutation within *gro*) result in a less severe neural hypertrophy (Preiss *et al.*, 1988) which is a phenotype more akin to the type of neural defect observed in some of the *Musca* embryos described in this study (figure 7.10b,c,d). Similarly, Schrons *et al.* (1992) describe a number of *gro* alleles which exhibit a range of embryonic phenotypes in homozygous individuals in which the maternal contribution is also compromised, these were classified into a phenotypic series from class 1 embryos exhibiting mild neural defects through to class 8 exhibiting severe neural hypertrophy of both the CNS and PNS. Using the criteria set out by Schrons *et al.* (1992) for scoring neurogenic defects, the phenotype of RNAi treated *Musca* embryos depicted in figure 7.10c closely resembles *Drosophila* embryos with class 4 defects, whereas the increase in PNS staining of those depicted in figure 7.10d is more reminiscent of those from class 6. Interestingly, Schrons *et al.* (1992) also note that some *gro* embryos exhibit fragmentation of the ventral nerve cord which, in some cases, is observed in association with expansion of the nerve cord in other regions. This is a phenotype which is also detected in *Musca* embryos injected with *gro* dsRNA (figure 7.10b). In summary, removal of *gro* activity from *Musca* embryos results in defects in the development of the nervous system. In some cases, these defects do not appear to be as severe as the phenotypes observed in *Drosophila* embryos lacking *gro*, but nevertheless the data suggest that *Musca gro* like *Drosophila gro* plays at least some role in nervous system development.

In *Drosophila*, mutations at the *gro* locus lead to a failure in regulation of cell fate specification in the neuroectoderm, which results in an abnormally large proportion of cells within this population adopting the neural fate rather than the alternative, epidermal fate. Neural hypertrophy is observed by staining with a marker of neural tissue (e.g. the α-HRP antibody, as performed above) whereas reduction in epidermal tissue can be assayed by observing the cuticle phenotype, which is either absent or severely reduced in *gro* mutant embryos. It would be interesting to observe the cuticle phenotype of *Musca* embryos treated with RNAi to determine whether the neural hypertrophy defects that have been observed in this study are associated with a corresponding loss in epidermal tissue.

In this chapter it is reported that a portion of the *Musca gro* gene has been cloned which exhibits a high degree of similarity to *Drosophila gro*. *in situ* hybridisation has shown that the developmental expression pattern of *gro* in *Musca* is identical to that previously described in *Drosophila* embryos suggesting that regulation and function of the gene is conserved. A preliminary functional analysis of the gene has also been conducted using the technique of RNAi, the results from which indicate that *gro* in *Musca* plays a role in the development of the nervous system, a function which is consistent with the known role of *gro* in *Drosophila*. In addition, successful transfer of RNAi technology to the housefly has been demonstrated thus adding weight to the argument that RNAi will become an important tool in deciphering gene function in other genetically intractable organisms.

Chapter 8

Summary and final conclusions

8.1. The function of E(spl)-m8 in repression of the neural fate.

The bHLH genes of the *E(spl)-C* and *groucho* are required for the allocation of appropriate cell fate in a number of developmental contexts in *Drosophila* and in other species. In an effort to better understand the specific function of these genes in development, the work carried out in this study was to investigate the role of these genes during neurogenesis in the fly.

The conservation of the WRPW motif in the Hairy-related proteins from several species, from flies to humans, implies that the motif performs some important functional role. Previous studies have highlighted a requirement for the motif *in vivo* (Wainwright and Ish-Horowicz, 1992; Giebel and Campos-Ortega, 1997) and for interaction with the Gro protein (Paroush *et al.*, 1994; Fisher *et al.*, 1996). These studies have led, to the now widely accepted hypothesis, that Gro is a transcriptional corepressor for the Hairy-related proteins. In this work, a mutational analysis of the WRPW motif in the E(spl)-m8 protein was performed by generating a number of derivatives containing single residue substitutions in the motif and analysing the function of these derivatives, *in vivo* in the developing fly (chapter 3) and for interaction with the Gro protein in the yeast two-hybrid system (chapter 4). It was hoped that subtle alterations such as these, would provide some indication about which residues are required for protein function.

When tested *in vivo* all three of mutant derivatives generated were completely non-functional. These results are in agreement with previous data which show that removal of the whole motif or single amino acid substitutions therein abolish *in vivo* function of Hairy (Wainwright and Ish-Horowicz, 1992) and E(spl)-m8 (Giebel and Campos-Ortega, 1997). The combined data from these studies and the data obtained here suggest that a number of residues within WRPW are critical for *in vivo* function. In order to determine whether there is a requirement for all four amino acids in the motif, a more exhaustive mutagenesis would have to be performed. However, the data from the studies described above, in association

with the highly conserved nature of the motif, would suggest that all four residues in the motif are required for function.

It was also noted that increasing the expression level of the mutant proteins, by generating flies homozygous for the mutant transgene and the p[GAL4] driver line, led to weak dominant negative effects. These effects were manifest by increases in the number of bristles on the notum and wing and the formation of ectopic vein tissue in the wing. The data indicate that the presence of the mutant protein interferes with the function of the endogenous protein in some way. Similarly, two other *E(spl)-m8* mutant derivatives, completely lacking the WRPW motif, have been reported elsewhere which display weak dominant negative effects (Welshons, 1956, Giebel and Campos-Ortega, 1997), one of which includes the original *E(spl)* mutant allele, $E(spl)^{D}$ (Welshons, 1956). In the discussion in chapter 3, it was speculated that interference may be a result of occupancy of the E(spl)-m8 target sites by the mutant proteins. Occupancy of these sites by the mutant proteins would not result in transcriptional repression of the target genes because the mutant proteins are unable to recruit Gro efficiently, but would hinder access to the wild-type endogenous proteins, thereby resulting in a net derepression of the target genes.

With the corepression hypothesis in mind, it was reasoned that the loss of function observed *in vivo* could result from a disturbance in interaction with the Gro protein. In order to examine this hypothesis, the interaction between the mutant derivatives and Gro was assayed and compared to the level of interaction between the wild-type E(spl)-m8 protein and Gro (chapter 4). It was discovered that the level of interaction between the WRPW mutant derivatives and Gro was altered in all three cases. The level of interaction ranged from 60% efficiency to 0% efficiency relative to the wild-type protein. Two other E(spl)-m8 WRPW derivatives have also been generated by colleagues in the laboratory, and similarly both mutations are found to disrupt interaction with Gro. These data offer a molecular basis for the observed loss of *in vivo* function: *in vivo* interaction between the mutant E(spl)-m8 proteins and endogenous Gro is either abolished or destabilised and as a consequence, effective recruitment of Gro to upstream regions of target genes does not occur. These target

genes are not therefore transcriptionally repressed. These results provide additional evidence that the WRPW motif is the Gro interaction domain and offer further support for the corepression hypothesis put forward by Paroush *et al.* (1994).

In the *in vivo* assay the function of the protein appeared to be completely abolished. If, therefore, we accept that loss of *in vivo* function is a sole consequence of disruption in ability to interact with Gro, we must also accept that all E(spl)-m8 function is mediated by Gro. Supporting evidence for this comes from the analysis by Schrons *et al.* (1992) showing that the neurogenic phenotype of *gro*⁻ embryos is at least as severe as that of *E(spl)-C* mutant embryos. Implicit in this suggestion is that repression of the neural fate is mediated exclusively at the level of transcriptional repression. In order to determine whether E(spl)-m8-mediated repression of the neural fate occurs at levels other than transcriptional repression of the proneural genes an *in vivo* ectopic co-expression assay was performed (chapter 6).

The ability of the E(spl)-m8 protein to suppress bristle development was determined in a co-expression assay with either the Sc or Da proteins. In both cases E(spl)-m8 was able to suppress bristle development even in the absence of ability to repress transgenic expression of UAS-sc or UAS-da. This data indicates that E(spl)-m8-mediated repression of the neural fate occurs at levels other than just transcriptional repression of the proneural genes. Furthermore, co-expression of the mutant E(spl)-m8 WRPW derivatives with either Sc of Da did not suppress bristle development indicating that this mode of repression, whatever it may be, is dependent on the WRPW motif. This therefore implicates a possible role for Gro.

As previously discussed in chapter 6, a mode of repression invoking sequestration of proneural protein activity by E(spl) mediated through bHLH interaction seems unlikely given that repression in the assay described above requires the WRPW motif: which is not necessary for interactions between these two classes of proteins (Gigliani *et al.*, 1996). A more likely explanation is that repression by E(spl)-m8 is mediated at the level of transcriptional repression of the proneural target genes. During neural development the

proneural genes are likely to be involved in the transcription of genes required for adopting the neural fate (which will probably include additional regulatory genes involved in the implementation of the neural fate) and neural differentiation genes. It is possible that E(spl) may also be involved in the direct transcriptional repression of these downstream genes. There are two lines of evidence suggesting that this might be the case. Firstly, Jiménez and Ish-Horowicz (1997) have shown that a chimeric protein containing E(spl)-m7 fused to a transcriptional activation domain (E(spl)-m7^{Act}) is able to induce certain leg bristles in the absence of the *ac* and *sc* genes. Usually, in the absence of *ac* and *sc*, no bristles develop, therefore in this context it is likely that E(spl)-m7^{Act} is acting on genes with functions downstream of the proneural proteins to cause neural differentiation. This suggests that the E(spl) proteins may directly regulate the same targets as the proneural proteins, but repress rather than activate their transcription. Secondly, recent work by Jennings *et al.* (1999) have demonstrated that the E(spl) and proneural proteins recognise and bind the same DNA target sites and show that the two classes of protein compete for the same sites *in vitro*. This raises the possibility that E(spl) and the proneural proteins compete for identical sites *in vivo*. Until the genes which are regulated directly by the proneural and E(spl) proteins have been identified it will not be possible to define the precise modes of action of E(spl). Identification and characterisation of these target genes will probably constitute one of the more fruitful and enlightening areas of research over the next few years.

At this juncture we shall turn to another potential mode of repression, that of regulation by RNA:RNA duplex formation. Lai and Posakony (1998) have proposed that RNA:RNA duplex formation occurs between two novel classes of conserved sequence motif, the GY box (GYB) and the proneural box (PB), found in the 3' UTR of *E(spl)* transcripts and proneural transcripts respectively. The GYB (GUCUUCC) which is found in the 3'UTR of a number of genes including *Bearded* (*Brd*), *hairy*, and three of the genes of the *E(spl)-C* (*m3*, *m4* and *m5*) is exactly complementary to the central seven nucleotides of the PB (GGAAGAC) which is found once in *ac*, *l'sc* and *ato* along with a second variant in both *l'sc* and *ato*. Lai and Posakony (1998) propose that RNA:RNA duplex formation between GYB- and PB-bearing transcripts may constitute a mechanism of translational repression

(ultimately resulting in a reduction of proneural activity). An example of this type of regulation has been described in *C. elegans* where small antisense RNAs encoded by the *lin-4* locus bind complementary sequences in the 3'UTR of *lin-14* and *lin-28* transcripts and block their translation (Slack and Ruvkun, 1997). Although there is no substantial evidence that this mechanism of regulation occurs in *Drosophila in vivo*, the opportunity to form such complex clearly exists, as transcripts from proneural genes and their regulators very frequently accumulate in coincident patterns. This hypothesis is supported by experimental evidence demonstrating that duplex formation between PB- and GYB-containing transcripts does occur *in vitro* and that this association is PB and GYB dependant.

The main reason for mentioning this here, is to emphasise (i) that further complexities exist in the already complex network of regulatory interactions that control cell fate decisions in neurogenesis, and (ii) that others may await discovery. The inherent complexity of the regulatory interactions operating during neurogenesis in *Drosophila* may bring into question the merits of this system as a model for studying regulation of cell fate. This may well be a fair criticism. Alternatively, it could be argued, that the vast amounts of time, work and thought invested in this system has revealed complexities which remain, as yet, undiscovered in other less well studied, seemingly less complex systems. The answer to this will be borne out in time. We should not be surprised, however, if the latter is closer to the truth.

8.2. The *Musca groucho* gene.

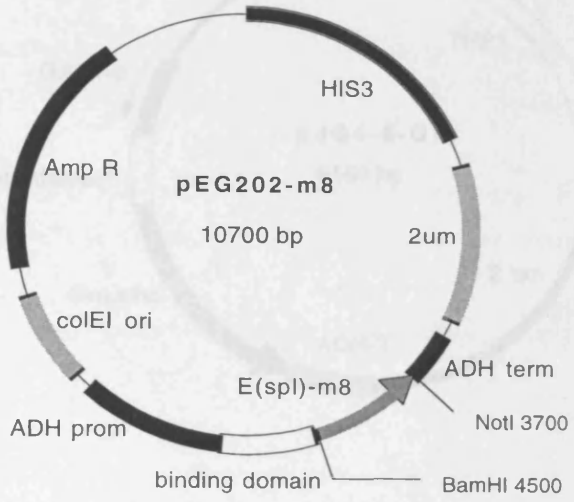
In chapter 7 of this study, cloning and characterisation of a portion of the *Musca gro* gene was described. The *Musca* Gro protein is highly similar to the *Drosophila* protein and the region of greatest identity is the conserved WD40 repeat region (96%). *in situ* hybridisation was performed to determine the expression of the *gro* gene in *Musca* embryos. The results from this experiment demonstrate that the gene is expressed in an identical pattern to the *gro* gene in *Drosophila*; expression is ubiquitous in early embryos, but later in development it is refined and becomes closely associated with the differentiating CNS. The expression pattern

in later embryos suggests that *Musca gro*, like *Drosophila gro*, may be involved in neural development. Furthermore, the pattern of expression in the blastoderm embryo indicates that *gro* may have other roles earlier in development such as during sex determination or segmentation. A preliminary analysis of gene function was performed by using the technique of RNAi to remove *gro* activity from *Musca* embryos. The effects of removal of *gro* activity on embryonic neural development was analysed by staining embryos with an antibody specific to the CNS. This analysis revealed that removal of *gro* activity produced neurogenic phenotypes similar to the phenotypes described for *Drosophila* embryos mutant for certain *gro* alleles (Schrons *et al.*, 1992) indicating that *Musca gro* is involved in neural development in the housefly.

The work described in chapter 7 has additionally demonstrated successful transfer of RNAi technology to the housefly, increasing the list of species where RNAi has been shown to work, and showing that the technique is an important tool in determining gene function in genetically intractable organisms.

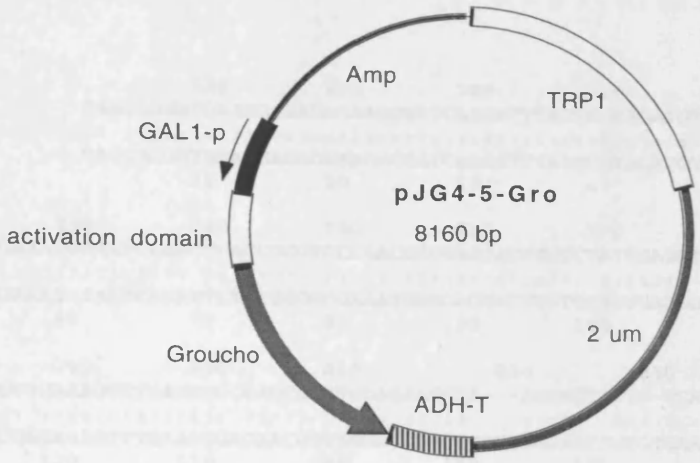
Appendices

Appendix 1. Yeast vector pEG202-m8.



pEG202-E(spl)-m8 was used in the yeast two-hybrid system to express a fusion protein containingthe LexA DNA-binding domain and either wild-type E(spl)-m8 or one of each of the three E(spl) WRPW mutant derivative proteins. The plasmid was generated by inserting a *Bgl*II – *Not*I fragment into the *Bam*HI – *Not*I sites in the vector.

Appendix 2. Yeast Vector pJG4-5-Gro.



pJG4-5-Gro was used in the yeast two-hybrid system to express a fusion protein containing the B42 activation domain and the full length *Drosophila* Groucho protein. The plasmid was a gift from Z. Paroush (Paroush *et al.*, 1994).

Appendix 3. *Musca* v *Drosophila* DNA.

Drosophila gro 1513 nt
Musca gro 1521 nt
52.9% identity

```
                      680       690       700       710       720
Drosophila      CAAGAGGATGAAGGCGAGAAAAGCGATCAAGATTTAGTCGTAGATGTTGCA
                :: ::::::::::: :::::::::::::::::::::::::::::::::::
Musca           CAGGAGGATGAAGGAGAGAAAAGCGATCAAGATTTAGTCGTAGATGTTGCA
                        10        20        30        40        50

             730       740       750       760       770       780
Dros    AATGAAATGGAATCCCACTCACCGCGTCCCAACGGCGAGCACGTGTCTATGGAGGTGCGC
        :::::::::::::: :: :: :: :: :: :: ::::::::: :::: ::::::::  :
Musca   AATGAAATGGAATCACATTCTCCTCGGCCAAATGGCGAGCACTTGTCCATGGAGGGTAGA
             60        70        80        90       100       110

             790       800       810           820       830
Dros    GATCGGGAAAGCTTGAATGGCGAGCGCCTGGAGAAGCCA---AGCAGTAGTG-GCATCAA
        :: :: ::::: :::::::: :: ::  :::: :: ::    :::::  ::: : : : :
Musca   GACCGAGAAAGTTTGAATGGAGAACGGTTGGATAAACCGGGTAGCAGCGGTGTGAAACCA
            120       130       140       150       160       170

         840       850       860       870       880       890
Dros    GCAG--GAACGGCCGCCCTCACGCTCCGGCTCCAGTTCGTCACGTTCCACACCCAGCCTC
        ::   ::  : :: :: :::::::: ::::: :: :::::::: :: :: :::::: :
Musca   CCATCAGAGAGACCACCTTCACGCTCAGGCTCAAGCTCGTCACGATCAACGCCCAGCTTT
            180       190       200       210       220       230

           900       910       920       930       940       950
Dros    AAGACAAAAGATATGGAAAAGCCGGGTACACCGGGCGCCAAGGCACGCACACCGACACCG
        ::::::::::::::: ::::: ::::: :: :: :: :: :: :::::::: :: :: ::
Musca   AAGACAAAAGATATGGATAAGCCTGGTACTCCAGGTGCAAAAGCACGCACTCCAACGCCC
            240       250       260       270       280       290

           960       970       980       990      1000      1010
Dros    AACGCCGCTGCTCCGGCGCCAGGCGTTAATCCTAAACAAATGATGCCGCAGGGA---CCA
        ::::: ::  : :: :: :  :::::::::: ::::::::::::::::: :::   ::
Musca   AACGCAGCGCCGCCAGCACAGGGCGTTAATCCCAAACAAATGATGCCGCAAGGAGGTCCT
            300       310       320       330       340       350

          1020      1030      1040      1050      1060      1070
Dros    CCGCCAGCCGGATATCCGGGTGCACCGTATCAAAGGCCGGCCGATCCCTACCAGCGTCCA
        ::::: :: :: ::::: :   : :: :::::  : :: ::::: :: :::::: : :::
Musca   CCGCCGGCTGGCTATCCAGCGTCCCCATATCAGCGACCAGCCGACCCATACCAGAGACCA
            360       370       380       390       400       410

          1080      1090      1100      1110      1120      1130
Dros    CCGTCAGATCCAGCCTATGGACGACCGCCACCAATGCCGTACGATCCGCACGCCCATGTG
        :: :: ::::: :: :: ::: : :: :: ::  : :: :::::::: :: :: :: ::
Musca   CCCTCCGATCCGGCTTACGGAAGGCCACCGCCTCTACCATACGATCCACATGCGCACGTT
            420       430       440       450       460       470

          1140      1150      1160      1170      1180      1190
Dros    CGAACCAATGGCATTCCACATCCCTCGGCCCTAACGGGTGGAAAGCCTGCATACTCTTTC
        ::::::::::::::::::::  : :: :: :: :::::::::::::: :: :: :::
Musca   AGAACCAATGGCATTCCACATCCGACTGCGCTTACCGGTGGAAAGCCTGCTTATTCCTTC
            480       490       500       510       520       530

          1200      1210      1220      1230      1240      1250
Dros    CATATGAACGGCGAGGGTAGTCTACAACCGGTGCCGTTCCCGCCGGACGCGTTGGTGGGT
        :::::::: :: :: :::::::::::::: :: :: ::::: :: ::::: :::::::::
Musca   CATATGAATGGTGAAGGTAGTCTACAACCCGTTCCCTTCCCTCCAGACGCCTTGGTGGGT
            540       550       560       570       580       590
```

```
          1260       1270       1280       1290       1300       1310
Dros      GTTGGAATTCCGCGGCACGCCCGTCAGATCAACACGCTGTCGCATGGAGAGGTCGTCTGT
          :: ::::: ::  : :: :: :: :: :: :: :: : :::: :::::::: :: :: ::
Musca     GTCGGAATACCCAGACATGCACGCCAAATAAACATGTTGTCACATGGAGAAGTTGTTTGC
          600        610        620        630        640        650

          1320       1330       1340       1350       1360       1370
Dros      GCGGTAACCATCTCTAATCCCACAAAGTACGTGTACACGGGCGGCAAGGGCTGCGTCAAG
          :: :: ::::: :: :: :: :::::::: :: :::::::: :::::::: :::::::: ::
Musca     GCAGTCACCATTTCAAACCCCACAAAATATGTGTACACCGGCGGCAAAGGCTGCGTGAAA
          660        670        680        690        700        710

          1380       1390       1400       1410       1420       1430
Dros      GTATGGGACATCTCGCAACCGGGCAACAAGAATCCAGTTAGCCAGCTGGATTGTCTGCAG
          :: ::::: :: ::::: :: :: :: :: :::: :::: :::: :::::::::::  ::::
Musca     GTGTGGGATATATCGCAGCCTGGTAATAAGAGTCCAATTAGTCAGCTGGATTGCTTGCAA
          720        730        740        750        760        770

          1440       1450       1460       1470       1480       1490
Dros      CGCGACAACTACATCCGCTCGGTAAAGCTGCTGCCCGACGGCCGTACGCTGATCGTGGGC
          :::::::::: :: :: :: ::::: :: :::::::: ::  : ::  : :: :: :: ::
Musca     CGCGACAACTATATTCGATCAGTAAAACTACTGCCCGATGGAAGAACATTAATTGTAGGT
          780        790        800        810        820        830

          1500       1510       1520       1530       1540       1550
Dros      GGTGAGGCGTCCAACCTGTCCATCTGGGATCTGGCCAGTCCGACGCCTCGCATAAAGGCG
          :: :::::::: :: :: ::::: :::::::: ::::: :: :: ::  : :: :: :: ::
Musca     GGAGAGGCGTCAAATCTATCCATTTGGGATCTTGCCAGCCCCACTCCCAGAATTAAAGCT
          840        850        860        870        880        890

          1560       1570       1580       1590       1600       1610
Dros      GAACTAACATCGGCGGCGCCCGCCTGCTACGCTCTGGCCATTAGCCCTGACTCGAAGGTG
          ::: :::::::  : :: :: :: :: :: :: :: ::::: ::::: :: :: ::::::::
Musca     GAATTAACATCATCTGCCCCAGCTTGTTATGCCCTGGCAATTAGTCCGGATTCGAAGGTT
          900        910        920        930        940        950

          1620       1630       1640       1650       1660       1670
Dros      TGCTTCTCGTGCTGCAGCGACGGCAATATCGCTGTGTGGGACCTGCACAACGAGATCCTG
          :::::::::: :::::::: ::: : :: ::::: :::::: :::::::::: ::  ::
Musca     TGCTTCTCGTGTTGCAGCGATGGCCACATAGCTGTTTGGGACTTGCACAACGAAATATTG
          960        970        980        990        1000       1010

          1680       1690       1700       1710       1720       1730
Dros      GTGCGCCAGTTCCAGGGCCATACCGACGGCGCTTCATGCATCGACATCAGTCCGGATGGC
          :: ::::: ::::: :: ::::: :: :: :: :: :: :: ::::: :: :: :: ::
Musca     GTACGCCAATTCCAAGGTCATACAGATGGTGCGTCTTGTATTGACATTAGCCCAGACGGT
          1020       1030       1040       1050       1060       1070

          1740       1750       1760       1770       1780       1790
Dros      TCCAGGCTGTGGACGGGCGGCTTAGACAACACGGTGCGCTCCTGGGATCTGCGCGAGGGT
           :: :  ::::::: :: ::  : ::::: :: :: :: :: :::::  :::: :: ::
Musca     ACCCGATTGTGGACCGGTGGTCTTGACAATACCGTTCGTTCATGGGACTTGCGAGAAGGC
          1080       1090       1100       1110       1120       1130

          1800       1810       1820       1830       1840       1850
Dros      CGCCAGCTGCAACAGCACGACTTCAGCTCTCAAATATTCTCGCTCGGCTACTGTCCCACA
          :::::: : ::::: :: :: ::::: :: :: :: :::::  : :: :: :::::: ::
Musca     CGCCAGTTACAACAACATGATTTCAGTTCACAGATTTTCTCATTGGGATATTGTCCCTCA
          1140       1150       1160       1170       1180       1190

          1860       1870       1880       1890       1900       1910
Dros      GGCGACTGGCTGGCTGTGGGTATGGAGAACTCGCATGTGGAGGTGCTGCACGCATCGAAA
          :: :::::::::: :: :::::::: :: :: ::::: ::::: :::::::::: :::
Musca     GGAGACTGGCTGGCAGTTGGTATGGAAAATTCACATGTCGAGGTTCTGCACGCATCAAAA
          1200       1210       1220       1230       1240       1250

          1920       1930       1940       1950       1960       1970
Dros      CCGGACAAGTATCAACTGCATCTGCACGAGAGCTGCGTTCTGTCGCTGCGCTTTGCCGCC
          :::::::: :: ::: : ::: ::::::: ::::: ::  ::::  :::: :::::: ::
Musca     CCGGACAAATACCAATTACATTTGCACGAAAGCTGTGTGTTGTCATTGCGTTTTGCCACC
          1260       1270       1280       1290       1300       1310
```

```
          1980       1990       2000       2010       2020       2030
Dros   TGCGGCAAATGGTTCGTTTCCACCGGCAAAGACAACCTGCTTAACGCATGGCGAACACCT
       :: :::::::::: :: ::::: ::::: :: ::: : :: :: :: ::: ::::::::
Musca  TGTGGCAAATGGTTTGTGTCCACAGGCAAGGATAACTTACTCAATGCTTGGAGAACACCT
          1320       1330       1340       1350       1360       1370

          2040       2050       2060       2070       2080       2090
Dros   TACGGTGCCAGCATATTCCAGTCGAAGGAAACATCATCCGTACTTAGCTGCGACATATCA
       :: :: :: :::::::::::::: :::::::::: :: :::::::  ::::::::  :
Musca  TATGGAGCAAGCATATTCCAGTCAAAGGAAACATCCTCAGTACTTACTTGCGACATACTA
          1380       1390       1400       1410       1420       1430

          2100       2110       2120       2130       2140       2150
Dros   ACTGACGACAAATACATTGTGACGGGTTCGGGCGATAAGAAGGCTACTGTCTACGAAGTT
       ::       ::  ::::: :::::::: :: :: :: :: ::::: :: :: ::::::::
Musca  ACCCGGACCAGGTACATCGTGACGGGATCCGGAGACAAAAAGGCAACAGTTTACGAAGTA
          1440       1450       1460       1470       1480       1490

          2160
Dros   ATTTATTAA
       :: :: :::
Musca  ATCTACTAA
```

Bibliography

Alifragis, P., Poortinga, G., Parkhurst, S.M. and Delidakis, C. (1997) A network of interacting transcriptional regulators involved in *Drosophila* neural fate specification revealed by the yeast two-hybrid system. *Proc. Natl. Acad. Sci. USA* **94:** 1309913104.

Arnosti, D. N., Gray, S., Barolo, S., Zhou, J. and Levine, M (1996) The gap protein knirps mediates both quenching and direct repression in the *Drosophila* embryo. *EMBO J.* **15:** 3659-3666.

Aronson, B. D., Fisher, A.L., Blechman, K., Caudy, M. and Gergen, J.P. (1997) Groucho-dependent and -independent repression activities of Runt domain proteins. *Mol. Cell. Biol.* **17:** 5581-5587.

Artavanis-Tsakonas, S., Matsuno, K. and Fortini, M.E. (1995) Notch signaling. *Science* **268:** 225-232.

Ashraf, S. I., and Ip, Y.T. (1998) Transcriptional control: Repression by local chromatin modification. *Curr. Biol.* **8:** 683-686.

Axelrod, J. D., Matsuno, K., Artavanis-Tsakonas, S. and Perrimon, N. (1996) Interaction between Wingless and Notch signaling pathways mediated by Dishevelled. *Science* **271:** 1826-1832.

Ayer, D. E., Lawrence, Q.A. and Eisenman, R.N. (1995) Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell* **80:** 767-776.

Bailey, A. M., and Posakony, J.W. (1995) Suppressor of Hairless directly activates transcription of *Enhancer of split* Complex genes in response to Notch receptor activity. *Genes and Dev.* **9:** 2609-2622.

Balcells, L. I., Modolell, J. and Ruiz-Gomez, M. (1988) A unitary basis for different *Hairy-wing* mutations of *Drosophila melanogaster. EMBO J.* **7:** 3899-3906.

Barolo, S., and Levine, M. (1997) Hairy mediates dominant repression in the *Drosophila* embryo. *EMBO J.* **16:** 2883-2891.

Bartel, P. L., Roecklein, J.A., SenGupta, D. and Fields, S. (1996) A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat. Genet.* **12:** 72-77.

Bartel, P. L., and Fields S. (1997). The Yeast Two-hybrid System. (NY.: Oxford University Press.).

Bate, M., Rushton, E. and Frasch, M. (1993) A dual requirement for neurogenic genes in *Drosophila* myogenesis. *Development* Suppl. pp. 149-161.

Bate, M. (1990) The embryonic development of larval muscles in *Drosophila*. *Development* **110:** 791-804.

Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L. and Weintraub, H. (1990) The protein Id: a negative regulator of helix-loop-helix binding proteins. *Cell* **61:** 49-59.

Beverley, S. M., and Wilson, A.C. (1984) Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. *J. Mol. Evol.* **21:** 1-13.

Blaumueller, C. M., Qi, H.L., Zagouras, P. and Artavanis-Tsakonas, S. (1997) Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane. *Cell* **90:** 281-291.

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

Brand, M., Jarman, A.P., Jan. L.Y. and Jan, Y.N. (1993) *asense* is a *Drosophila* neural precursor gene and is capable of initiating sense organ formation. *Development* **119:** 1-17.

Brand, M., and Campos-Ortega, J.A. (1989) Two groups of interrelated genes regulate early neurogenesis in *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.* **197:** 457-470.

Bray, S. (1998) Notch signalling in *Drosophila*: three ways to use a pathway. *Cell Dev. Biol.* **9:** 591-597.

Brent, R., and Ptashne, M. (1985) A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43:** 729-736.

Burgess, R., Cserjesi, P., Ligon, K.L. and Olson, E.N. (1995) Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. *Dev. Biol.* **168:** 296-306.

Cabrera, C. V., Martinez-Arias, A. and Bate, M. (1987) The expression of three members of the *achaete-scute* gene complex correlates with neuroblast segregation in *Drosophila*. *Cell* **50:** 425-433.

Cabrera, C. V. (1994) Regulation of scute function by *extramacrochaete* in vitro and in vivo. *Development* **120:** 3595-3603.

Cabrera, C. V., and Alonso, M.C. (1991) Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10:** 2965-2973.

Campos-Ortega, J. A., and Hartenstein, V. (1985). The embryonic development of *Drosophila melanogaster*. (Berlin: Springer Verlag).

Carmena, A., Bate, M. and Jimenez, F. (1995) *lethal of scute*, a proneural gene, participates in the specification of muscle progenitors during Drosophila embryogenesis. *Genes and Dev.* **9:** 2373-2383.

Carter, M. J., and Milton, I.D. (1993) An inexpensive and simple method for DNA purifications on silica particles. *Nucleic Acids Res.* **21:** 1044.

Cavallo, R. A., Cox, R.T., Moline, M.M., Roose, J., Polevoy, G.A., Clevers, H., Peifer, M. and Bejsovec, A. (1998) *Drosophila* Tcf and Groucho interact to repress Wingless signalling activity. *Nature* **395:** 604-608.

Chasman, D. I., Leatherwood, J., Carey, M., Ptashne, M. and Kornberg, R.D. (1989) Activation of yeast polymerase II transcription by herpes virus VP16 and GAL4 derivatives *in vitro*. *Mol. Cell. Biol.* **9:** 4746-4749.

Chen, G., Nguyen, P.H. and Courey, A.J. (1998) A role for Groucho tetramerization in transcriptional repression. *Mol. Cell. Biol.* **18:** 7259-7268.

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

Chitnis, A., Henrigue, D., Lewis, J., Ish-Horowicz, D. and Kintner, C. (1995) Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene *Delta*. *Nature* **375:** 761-766.

Chong, J. A., and Mandel, G. (1997). Isolation of DNA-binding proteins using a one-hybrid genetic screen. In The Yeast Two-hybrid system., P. L. Bartel, and Fields, S., ed. (NY).

Choudhury, B. K., Kim, J., Kung, H.F. and Li, S.S. (1997) Cloning and developmental expression of Xenopus cDNAs encoding the Enhancer of split groucho and related proteins. *Gene* **195:** 41-48.

Corbin, V. M., A.M., Abmayr, S.M., Neel, V., Alcamo, E., Maniatis, T. and Young, M.W. (1991) A role for the *Drosophila* neurogenic gene in mesoderm differentiation. *Cell* **67:** 311-323.

Cserjesi, P., Brown, D., Ligon, K.L., Lyons, G.E., Copeland, N.G., Gilbert, D.J., Jenkins, N.A. and Olson, E.N. (1995) Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development* **121:** 1099-1110.

Cubas, P., and Modolell, J. (1992) The *extramacrochaetae* gene provides information for sensory organ patterning. *EMBO J* **11:** 3385-3393.

Cubas, P., de Celis, J.F., Campuzano, S. and Modolell, J. (1991) Proneural clusters of *achaete-scute* expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes Dev.* **5:** 996-1008.

Culi, J., and Modolell, J. (1998) Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated by *Notch* signaling. *Genes Dev.* **12:** 2036-2047.

Dambly-Chaudiere, G., A., Jan, L.Y. and Jan. Y.N. (1988) The determination of sense organs in *Drosophila*: Interaction of *scute* with *daughterless*. *Roux's Arch. Dev. Biol.* **197:** 419-423.

Dawson, S. R., Turner, D.L., Weintraub, H. and Parkhurst, S.M. (1995) Specificity for the Hairy/Enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. *Mol. Cell. Biol.* **15:** 6923-6931.

de Celis, J. F., Mari-Beffa, M. and Garcia-Bellido, A. (1991) Function of trans-acting genes of the *achaete-scute* complex in sensory organ patterning in the mesonotum of *Drosophila*. *Roux's Arch. Dev. Biol.* **200:** 64-76.

de Celis, J. F., de Celis, J., Ligoxygakis, P., Preiss, A., Delidakis, C. and Bray, S.J. (1996) Functional relationships between *Notch*, *Su(H)* and the bHLH genes of the *E(spl)* complex: the *E(spl)* genes mediate only a subset of *Notch* activities during imaginal development. *Development* **122:** 2719-2728.

de la Concha, A., Dietrich, U., Weigel, D. and Campos-Ortega, J.A. (1988) Functional interactions of neurogenic genes of *Drosophila melanogaster*. *Genetics* **118:** 499-508.

De la Pompa, J. L., Wakeham, A., Correia, K.M., Samper, E., Brown, S., Aguilera, R.J., Nakano, T., Honjo, T., Mak, T.W., Rossant, J. and Conlon, R.A. (1997) Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124:** 1139-1148.

Delidakis, C., and Artavanis-Tsakonas, S. (1992) The Enhancer of split [*E(spl)*] locus of *Drosophila* encodes seven independent helix-loop-helix proteins. *Proc. Natl. Acad. Sci. USA* **89:** 8731-8735.

Delidakis, C., Preiss, A., Hartley, D.A. and Artavanis-Tsakonas, S. (1991) Two genetically and molecularly distinct functions involved in early neurogenesis reside within the *Enhancer of split* locus of *Drosophila melanogaster*. *Genetics* **129:** 803-823.

Diederich, R. J., Matsuno, K., Hing, H. and Artavanis-Tsakonas, S. (1994) Cytosolic interaction between deltex and Notch ankyrin repeats implicates deltex in the Notch signalling pathway. *Development* **120:** 473-481.

Doe, C. Q., and Goodman, C.S. (1985) Early events in insect neurogenesis. II. The role of cell interactions and cell lineages in the determination of neuronal precursor cells. *Dev. Biol.* **111:** 206-219.

Dominguez, M., and Campuzano, S. (1993) *asense*, a member of the *Drosophila achaete-scute* complex, is a pronueral and neural differentiation gene. *EMBO J.* **12:** 2049-2060.

Dubnicoff, T., Valentine, S.A., Chen, G., Shi, T., Lengyel, J.A., Paroush, Z. and Courey, A.J. (1997) Conversion of dorsal from an activator to a repressor by the global corepressor Groucho. *Genes Dev.* **11:** 2952-2957.

Ellis, H. M. (1994) Embryonic expression and function of the *Drosophila* helix-loop-helix gene *extramacrochaetae*. *Mech. Dev.* **47:** 65-72.

Farida, S. S., Tsoi, S.C.-M. and Li, S.-L. (1997) cDNA cloning and genomic organization of Enhancer of split groucho gene from nematode *Caenorhabditis elegans*. *Biochem. Mol. Biol. Int.* **43:** 327-337.

Fehon, R., Kooh, P., Rebay, I., Regan, C., Xu, T., Mustavich, M. 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:** 523-534.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabelling DNA retriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.

Fields, S., and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340:** 245-246.

Finley, R. L. J., and Brent, R. (1994). In Gene Probes: A Practical Approach., D. Hames, and Glover, D., ed. (Oxford: Oxford University Press).

Fire, A., Xu, S., Montgomer, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabdities elegans*. *Nature* **391:** 806-810.

Fisher, A., and Caudy, M. (1998) The function of Hairy-related bHLH repressor proteins in cell fate decisions. *BioEssays* **20:** 298-306.

Fisher, A. L., Ohsako, S. and Caudy, M. (1996) The WRPW motif of the Hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino acid transcription repression and protein-protein interaction domain. *Mol. Cell. Biol.* **16:** 2670-2677.

Fortini, M. E., and Artavanis-Tsakonas, S. (1994) The Suppressor of Hairless protein participates in Notch receptor signalling. *Cell* **79:** 273-282.

Garcia-Alonso, L. A., and Garcia-Bellido, A, (1986) Genetic analysis of the *Hairy-wing* mutations. *Roux's Arch. Dev. Biol.* **195:** 259-264.

Garcia-Bellido, A. (1979) Genetic analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics*, **91:** 491-520.

Garcia-Bellido, A., and Santamaria, P. (1978) Developmental analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics* **88:** 469-486.

Garcia-Bellido, A., and de Celis, J.F. (1992) Developmental genetics of the veination pattern of *Drosophila*. *Annu. Rev. Genet.* **26:** 277-304.

Garrell, J., and Modolell, J. (1990) The *Drosophila* extramacrochaete locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* **61:** 39-48.

Ghysen, A., and Dambly-Chaudiere, C. (1989) Genesis of the *Drosophila* peripheral nervous system. *Trends. Genet.* **5:** 251-255.

Giebel, B., and Campos-Ortega, J.A. (1997) Functional dissection of the *Drosophila* Enhancer of split protein, a suppressor of neurogenesis. *Proc. Natl. Acad. Sci. USA* **94:** 6250-6254.

Giebel, B., Stüttem, I., Hinz, U. and Campos-Ortega, J.A. (1997) Lethal of Scute requires overexpression of Daughterless to elicit ectopic neuronal development during embryogenesis in *Drosophila*. *Mech. Dev.* **63:** 75-87.

Gigliani, F., Longo, F., Gaddini, L. and Battaglia, P.A. (1996) Interactions among the bHLH domains of the proteins encoded by the *Enhancer of split* and *achaete-scute* gene complexes of *Drosophila*. *Mol. Gen. Genet.* **251:** 628-634.

Gomez-Skarmeta, J. L., de Corral, R.D., de la Calle-Mustienes, E., Ferre-Marco, D. and Modolell, J. (1996) *araucan* and *caupolican*, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell* **85:** 95-105.

Gomez-Skarmeta, J. L., Rodriguez, I., Martinez, C., Culf, J., Ferres-Marco, M.D., Beamonte, D. and Modolell, J. (1995) Cis-regulation of *achaete* and *scute*: shared enhancer-like elements drive their coexpression in proneural clusters of the imaginal discs. *Genes Dev.* **9:** 1869-1882.

Gonzalez, F., Romani, S., Cubas, P., Modolell, J. and Campuzano, S. (1989) Molecular analysis of the *asense* gene, a member of the *achaete-scute* complex of *Drosophila melanogaster*, and its novel role in optic lobe development. *EMBO J.* **8:** 3553-3562.

Gray, S., Szymanski, P. and Levine, M. (1994) Short-range repression permits multiple enhancers to function autonomously within a complex promoter. *Genes Dev.* **8:** 1829-1838.

Grbavec, D., and Stifani, S. (1996) Molecular interaction between TLE1 and the carboxyl-terminal domain of HES-1 containing the WRPW motif. *Biochem. Biophys. Res. Commun.* **223:** 701-705.

Gstaiger, M., Knoepfel, L., Georgiev, O., Schaffner, W. and Hovens, C.M. (1995) A B-cell coactivator of octamer-binding transcription factors. *Nature* **373:** 360-362.

Gstaiger, M., Georgiev, O. and Schaffner, W. (1996) Fine mapping of protein interaction surfaces with a PCR-based mutagenesis screen in yeast. *Trends Genet.* **12:** 393-394.

Guo, M., Jan, L.Y. and Jan, Y.N. (1996) Control of daughter cell fates during asymmetric division: interaction of Numb and Notch. *Neuron* **17:** 27-41.

Gyuris, J., Golemis, E., Chertkov, H. and Brent, R. (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75:** 791-803.

Häder, T., Wainwright, D., Shandala, T., Saint, R., Taubert, H, Brönner, G and Jäckle, H. (2000) Receptor tyrosine kinase signaling regulates different modes of Groucho-dependent control of Dorsal. *Curr. Biol.* **10:** 51-54.

Haenlin, M., Kunisch, M., Kramatschek, B. and Campos-Ortega, J.A. (1994) Genomic regions regulating early embryonic expression of the *Drosophila* neurogenic gene *Delta*. *Mech. Dev.* **47:** 99-110.

Hamilton, B. A., Palazzolo, M..J., and Chang, J.H. (1991) Large scale screen for transposon insertions into cloned genes. *Proc. Natl. Acad. Sci. USA* **88:** 2731-2735.

Han, K., Levine, M.S. and Manley, J.L. (1989) Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56:** 573-583.

Hartenstein, V. (1988) Development of *Drosophila* larval sensory organs: spatiotemporal pattern of sensory neurons, peripheral axonal pathways and sensilla differentiation. *Development* **102:** 869-886.

Hartley, D. A., Preiss, A. and Artavanis-Tsakonas, S. (1988) A deduced gene product from the *Drosophila* neurogenic locus, *Enhancer of split*, shows homology to mammalian G-protein b subunit. *Cell* **55:** 785-795.

Heitzler, P., and Simpson, P. (1993) Altered epidermal growth factor-like sequences provide evidence for a role of *Notch* as a receptor in cell fate decisions. *Development* **117:** 1113-1123.

Heitzler, P., and Simpson, P. (1991) The choice of cell fate in the epidermis of *Drosophila*. *Cell* **64:** 1083-1092.

Heitzler, P., Bourouis, M., Ruel, L., Carteret, C. and Simpson, P. (1996b) Genes of the *Enhancer of split* and *achaete-scute* complexes are required for the regulatory loop between *Notch* and *Delta* during lateral signalling in *Drosophila*. *Development* **122:** 161-171.

Heitzler, P., Haenlin, M., Ramain, P., Calleja, M. and Simpson, P. (1996a) A genetic analysis of *pannier*, a gene necessary for viability of dorsal tissues and bristle positioning in *Drosophila*. *Genetics* **143:** 1271-1286.

Hinz, U., Giebel, B. and Campos-Ortega, J.A. (1994) The basic helix-loop-helix domain of *Drosophila* lethal of scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* **76:** 77-87.

Hoch, M., Broadie, K., Jäckle, H. and Skaer, H. (1994) Neurogenesis in the malpighian tubules of *Drosophila*. *Development* **120:** 3439-3450.

Hope, I. A., and Struhl, K. (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46:** 885-894.

Irvine, K. D., and Vogt, T.F. (1997) Dorsal-ventral signaling in limb development. *Curr. Opin. Cell. Biol.* **9:** 867-876.

Ishibashi, M., Moriyoshi, K., Sasai, Y., Shiota, K., Nakanishi, S. and Kageyama, R. (1994) Persistent expression of helix-loop-helix factor HES-1 prevents mammalian neural differentiation in the central nervous system. *EMBO J.* **13:** 1799-1805.

Ishibashi, M., Ang, S.L., Shiota, K., Nakanishi, S., Kageyama, R. and Guillemot, F. (1995) Targeted disruption of mammalian *hairy* and *Enhancer of split* homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis and severe neural tube defects. *Genes Dev.* **9:** 3136-3148.

Jan, L. Y., and Jan, Y.N. (1982) Antibodies to horseradish peroxidase as specific neuronal markers in *Drosophila* and in grasshopper embryos. *Proc. Natl. Acad. Sci. USA* **79:** 2700-2704.

Jan, Y. N., and Jan, L.Y. (1993b) HLH proteins, fly neurogenesis, and vertebrate myogenesis. *Cell* **75:** 827-830.

Jan, Y. N., and Jan, L.Y. (1993a). The peripheral nervous system. In The development of *Drosophila melanogaster.*, M. Bate, and Martinez-Arias,A., ed. (Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press), pp. 1207-1244.

Jarman, A. P., Grau, Y., Jan, L.Y. and Jan, Y.N. (1993) *atonal* is a proneural gene that directs chordotonal organ formation in the *Drosophila* peripheral nervous system. *Cell* **73:** 1307-1321.

Jarman, A. P., Grell, E.H., Ackerman, L., Jan, L.Y. and Jan, Y.N. (1994) *atonal* is the proneural gene for *Drosophila* photoreceptors. *Nature* **369:** 398-400.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E.H., Kopan, R. and Israel, A. (1995) Signalling downstream of activated mammalian Notch. *Nature* **377:** 355-358.

Jennings, B., Preiss, A., Delidakis, C and Bray, S.J. (1994) The Notch signalling pathway is required for *Enhancer of split* bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development* **120:** 3537-3548.

Jennings, B., de Celis, J., Delidakis, C., Preiss, A. and Bray, S.J. (1995) Role of *Notch* and *achaete-scute* complex in the expression of *Enhancer of split* bHLH proteins. *Development* **121:** 3745-3752.

Jennings, B. H., Tyler, D.M. and Bray. S.J. (1999) Target specificities of *Drosophila* Enhancer of split basic helix-loop-helix proteins. *Mol. Cell. Biol.* **19:** 4600-4610.

Jimenez, G., and Ish-Horowicz, D. (1997) A chimeric Enhancer-of-split transcriptional activator drives neural development and *achaete-scute* expression. *Mol. Cell. Biol.* **17:** 4355-4362.

Jimenez, G., Paroush, Z. and Ish-Horowicz, D. (1997) Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. *Genes Dev.* **11:** 3072-3082.

Johnson, A. D. (1995) The price of repression. *Cell* **81:** 655-658.

Keegan, L., Gill, G. and Ptashne, M. (1986) Separation of DNA-binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231:** 699-704.

Keleher, C. A., Redd, M.J., Schultz, J., Carlson, M. and Johnson, A.D. (1992) Ssn6-Tup1 is a general repressor of transcription in yeast. *Cell* **68:** 709-719.

Kennerdell, J. R., and Carthew, R.W. (1998) Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in Wingless pathway. *Cell* **95:** 1017-1026.

Ketting, R. F., Haverkam, T.H.A., van Luenen, H.G.A.M. and Plasterk, R.H.A. (1999) *mut-7* of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner Syndrome Helicase and RNaseD. *Cell* **99:** 1330141.

Kidd, S., Kelley, M.R. and Young, M.W. (1986) Sequence of the *Notch* locus of *Drosophila melanogaster*: relationship of the encoded protein to mammalian clotting and growth factors. *Mol. Cell. Biol.* **6:** 3094-3108.

Klämbt, C., Knust, E., Tietze, K. and Campos-Ortega, J.A. (1989) Closely related transcripts encoded by the neurogenic gene complex *Enhancer of split* of *Drosophila melanogaster*. *EMBO J.* **8:** 203-210.

Knust, E., Tietz, K. and Campos-Ortega, J.A. (1987) Molecular analysis of the neurogenic locus *Enhancer of split* of *Drosophila melanogaster*. *EMBO J.* **6:** 4113-4123.

Knust, E., Schrons, H., Grawe, F. and Campos-Ortega, J.A. (1992) Seven genes of the *Enhancer of split* complex of *Drosophila melanogaster* encode helix-loop-helix proteins. *Genetics* **132:** 505-518.

Koop, K. E., MacDonald, L.M. and Lobe, C.G. (1996) Transcripts of *Grg4*, a murine *groucho*-related gene, are detected in tissues adjacent to other murine neurogeneic gene homologs during embryonic development. *Mech. Dev.* **59:** 73-87.

Kopan, R., Chroeter, E.H., Weintraub, H. and Nye, J.S. (1996) Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. *Proc. Natl. Acad. Sci. USA* **93:** 1683-1688.

Kunisch, M., Haenlin, M. and Campos-Ortega, J.A. (1994) Lateral inhibition mediated by the *Drosophila* neurogenic gene *Delta* is enhanced by proneural proteins. *Proc. Natl. Acad. Sci. USA* **91:** 10139-10143.

Lai, E. C., and Posakony, J.W. (1998) Regulation of *Drosophila* neurogenesis by RNA:RNA duplexes? *Cell* **93:** 1103-1104.

Lawrence, P.A. (1992) The Making of a Fly. (Oxford: Blackwell Scientific Publications).

Lawrence, P.A., and Green, S.M. (1979) Cell lineage in the developing retina of *Drosophila*. *Dev. Biol.* **71:** 142-152.

Leanna, C. A., and Hannink, M. (1996) The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. *Nucleic Acids Res.* **24:** 3341-3347.

Lecourtois, M., and Schweisguth, F. (1998) Indirect evidence for *Delta*-dependent intracellular processing of Notch in Drosophila emryos. *Curr. Biol.* **8:** 771-774.

Lecourtois, M., and Schweisguth, F. (1995) The neurogenic Suppressor of Hairless DNA-binding protein mediates the transcriptional activation of the *Enhancer of split* complex genes triggered by Notch signaling. *Genes Dev.* **9:** 2598-2608.

Lehmann, R., Jimenez, F., Dietrich, U. and Campos-Ortega, J.A. (1983) On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.* **192:** 62-74.

Lehmann, R., Jimenez, F., Dietrich, U. and Campos-Ortega, J.A. (1992) On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.* **192:** 62-74.

Leiber, T., Kidd, S., Alcamo, E., Corbin, V. and Young, M.W. (1993) Antineurogenic phenotypes induced by truncated Notch proteins indicate a role in signal transduction and may point to a novel function for Notch in nuclei. *Genes Dev.* **7:** 1949-1965.

Leyns, L., Dambly-Chaudiere, C. and Ghysen, A. (1989) Two different sets of cis elements regulate *scute* to establish two different sensory patterns. *Roux's Arch. Dev. Biol.* **198:** 227-232.

Li, B., and Fields, S. (1993) Identification of mutations in p53 that affect its binding to sv40 large t-antigen by using the yeast two-hybrid system. *FASEB* **7:** 957-963.

Ligoxygakis, P., Bray, S.J., Apidianakis, Y. and Delidakis, C. (1999) Ectopic expression of individual *E(spl)* genes has differential effects on different cell fate decisions and underscores the biphasic requirement for Notch activity in wing margin establishment in *Drosophila*. *Development* **126:** 2205-2214.

Ligoxygakis, P., Yu, S.Y., Delidakis, C. and Baker, N.E. (1998) A subset of Notch functions during *Drosophila* eye development require *Su(H)* and the *E(spl)* gene complex. *Development* **125:** 2893-2900.

Lindsley, D. L., and Grell, E.H. (1968). *Genetic Variations of Drosophila melanogaster.*: Carnegie Inst. Wash. Publ.), pp. 627.

Lindsley, D. L., and Zimm, G.G. (1992). The Genome of *Drosophila melanogaster*. (San Diego, CA: Academic Press, Inc.).

Logeat, F., Bessia, C., Brou, C., LeBail, O., Jarriault, S., Seidah, N.G. and Israël, A. (1998) The Notch1 receptor is cleaved constitutively by a furin-like convertase. *Proc. Natl. Acad. Sci. USA* **95:** 8108-8112.

Ma, J., and Ptashne, M. (1988) Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55:** 443-446.

Ma, J., and Ptashne, M. (1987) A new class of yeast transcriptional activators. *Cell* **51:** 113-119.

Mannervik, M., Nibu, Y., Zhang, H. and Levine, M. (1999) Transcriptional coregulators in development. *Science* **284:** 606-609.

Martin-Bermudo, M.D., Martinez, C. Rodriquez, A. and Jiminez, F. (1991) Distribution and function of the *lethal of scute* gene productduring early neurogenesis in *Drosophila*. *Development* **113:** 1-17.

Martin-Bermudo, M. D., Carmena, A. and Jimenez, F. (1995) Neurogenic genes control gene expression at the transcriptional level in early neurogenesis and in mesectoderm specification. *Development* **121:** 219-224.

Martinez, C., and Modolell, J. (1991) Cross-regulatory interactions between the proneural *achaete* and *scute* genes of *Drosophila*. *Science* **251:** 1485-1487.

Martinez, C., Modolell, J. and Garrell, J. (1993) Regulation of the proneural gene *achaete* by helix-loop-helix proteins. *Mol. Cell. Biol.* **13:** 3514-3521.

Matsuno, K., Diederich, R.J., Go, M.J., Blaumueller, C.M. and Artavanis-Tsakonas, S. (1995) Deltex acts as a positive regulator of Notch signaling through interactions with the Notch ankyrin repeats. *Development* **121:** 2633-2644.

Misquitta, L., and Paterson, B.M. (1999) Targeted disruption of gene function in *Drosophila* by RNA interference (RNA-i): A role for *nautilus* in embryonic somatic muscle formation. *Proc. Natl. Acad. Sci. USA* **96:** 1451-1456.

Misra, S., and Rio, D.C. (1990) Cytotype control of *Drosophila* P element transposition: the 66 kd protein is a repressor of transposase activity. *Cell* **62:** 269-284.

Miyasaka, H., Choudhury, B.K., Hou, E.W. and Li, S.S. (1993) Molecular cloning and expression of mouse and human cDNA encoding AES and ESG proteins with strong similarity to *Drosophila* enhancer of split groucho protein. *Biochem.* **216:** 343-352.

Mohr, M. (1923) *f. indukt. Abst. Verebl.* **32.**

Montgomery, M. K., Xu, S. and Fire, A. (1998) RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **95:** 15502-15507.

Morata, G., and Ripoll, P. (1975) Minutes: mutants of *Drosophila* autonomously affecting cell division rate. *Dev. Biol.* **427:** 211-221.

Muller, H. J. (1935) A viable two-gene deficiency phenotypically resembling the corresponding hypomorphic mutations. *J. Heredity* **26:** 469-478.

Murre, C., Schonleber McCaw, P. and Baltimore, D. (1989) A new DNA binding and dimerization motif in Immunoglobulin enhancer binding. *Cell* **56:** 777-783.

Nakao, K., and Campos-Ortega, J.A. (1996) Persistent expression of genes of the *Enhancer of split* complex suppresses neural development in *Drosophila*. *Neuron* **16:** 275-286.

Neer, E. J., Schmidt, C.J., Nambudripad, R. and Smith, T.F. (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371:** 297-300.

Ngo, H., Tschudi, C., Gull, K. and Ullu, E. (1998) Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci.* **95:** 14687-14692.

Nishimura, A., Morita, M., Nishimura, Y. and Sugino, Y. (1990) A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucleic Acids Res.* **18:** 6169.

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

Oellers, N., Dehio, M. and Knust, E. (1994) bHLH proteins encoded by the *Enhancer of split* complex of *Drosophila* negatively interfere with transcriptional activation mediated by proneural genes. *Mol. Gen. Genet.* **244:** 465-473.

Ohsako, S., Hyer, J., Panganiban, G., Oliver, I. and Caudy, M. (1994) *hairy* function as a DNA-binding helix-loop-helix repressor of *Drosophila* sensory organ formation. *Genes Dev.* **8:** 2743-2755.

Palaparti, A., Baratz, A and Stifani, S. (1997) The Groucho/transducin-like Enhancer of split transcriptional repressors interact with the genetically defined amino-terminal silencing domain of histone H3. *J. Biol. Chem.* **272:** 26604-26610.

Panin, V. N., Papayannopoulos, V., Wilson, R. and Irvine, K.D. (1997) Fringe modulates Notch ligand interactions. *Nature* **387:** 908-912.

Parkhurst, S. M., Bopp, D. and Ish-Horowicz, D. (1990) X:A ratio is transduced by helix-loop-helix proteins in *Drosophila*. *Cell* **63:** 1179-1191.

Paroush, Z., Finley Jr., R.L., Kidd, T., Wainwright, S.M., Ingham, P.W., Brent, R. and Ish-Horowicz, D. (1994) Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with Hairy-related bHLH proteins. *Cell* **79:** 805-815.

Paroush, Z., Wainwright, S.M. and Ish-Horowicz, D. (1997) Torso signalling regulates terminal patterning in *Drosophila* by antagonising Groucho-mediated repression. *Development* **124:** 3827-3834.

Patel, N. H., Martinblanco, E., Coleman, K.G., Poole, S.J., Ellis, M.C., Kornberg, T.B. and Goodman, C.S. (1989) Expression of Engrailed proteins in Arthropods, Annelids, and Chordates. *Cell* **58:** 955-968.

Pflugrad, A., Meir, J.Y., Barnes, T.M. and Miller, D.M. (1997) The Groucho-like transcription factor UNC-37 functions with the neural specificity gene *unc-4* to govern motor neuron identity in *C. elegans*. *Development* **124:** 1699-1709.

Porcher, C. S., W., Rockwell, K. Fujiwara, Y., Alt, F.W. and Orkin, S.H. (1996) The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **86:** 47-57.

Poulson, D. F. (1937) Chromosomal deficiencies and embryonic development of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **23:** 133-137.

Preiss, A., Hartley, D.A. and Artavanis-Tsakonas, S. (1988) The molecular genetics of *Enhancer of split*, a gene required for embryonic neural development in *Drosophila*. *EMBO J.* **7:** 3917-3927.

Prokop, A., and Technau, G.M. (1991) The origin of postembryonic neuroblasts in the ventral nerve cord of *Drosophila melanogaster*. *Development* **111:** 79-88.

Qi, H., Rand, M.D., Wu, X., Sestan, N., Wang, W., RAkic, P., Xu, T. and Artavanis-Tsakonas, S. (1999) Processing of the Notch ligand Delta by the metalloprotease Kuzbanian. *Science* **283:** 91-93.

Ramain, P., Heitzler, P., Haenlin, M. and Simpson, P. (1993) *pannier*, a negative regulator of *achaete* and *scute* in *Drosophila* encodes a zinc finger protein with homology to the vertebrate transcription factor GATA-1. *Development* **119:** 1277-1291.

Rebay, I., Fleming, R.J., Fehon, R.G., Chervas, P. and Artavanis-Tsakonas, S. (1991) Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell* **67:** 687-699.

Reinke, R., and Zipursky, S.L. (1988) Cell-cell interaction in the *Drosophila* retina: the *bride of sevenless* gene is required in photoreceptor cell R8 for R7 development. *Cell.* **55:** 321-330.

Rodriguez, I., Hernandez. R., Modolell, J. and Ruiz-Gomez, M. (1990) Competence to develop sensory organs is temporally and spatially regulated in *Drosophila* epidermal primordia. *EMBO J.* **9:** 3583-3592.

Romani, S., Campuzano, S., Macagno, E.R. and Modolell, J. (1989) Expression of *achaete* and *scute* genes in *Drosophila* imaginal discs and their function in sensory organ development. *Genes Dev.* **3:** 997-1007.

Rooke, J., Pan, D., Xu, T. and Rubin, G.M (1996) KUZ, a conserved metalloprotease-disintegrin protein with two roles in *Drosophila* neurogenesis. *Science* **273:** 1227-1231.

Royet, J., Bouwmeester, T. and Cohen. S.M. (1998) *Notchless* encodes a novel WD40-repeat-containing protein that modulates Notch signaling activity. *EMBO J.* **24:** 7351-7359.

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

Ruiz-Gomez, M., and Modolell, J. (1987) Deletion analysis of the *achaete-scute* locus of *D. melanogaster*. *Genes Dev.* **1:** 1238-1246.

Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M. (1988) GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335:** 563-564.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning: A laboratory manual., Second Edition (Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.).

Sánchez Alvarado, A., and Newmark, P.A. (1999) Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl. Acad. Sci. USA* **96:** 5049-5054.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R. and Nakanishi, S. (1992) Two mammalian helix-loop-helix factors structurally related to *Drosophla hairy* and *Enhancer of split*. *Genes Dev.* **6:** 2620-2634.

Schmidt, A., Palumbo, G., Bozzetti, M.P., Tritto, P., Pimpinelli, S. and Schafer, U. (1999) Genetic and molecular characterization of *sting*, a gene involved in crystal formation and meiotic drive in the male germ line of *Drosophila melanogaster*. *Genetics* **151:** 749-760.

Schmidt, C. J., and Sladek, T.E. (1993) A rat homolog of the *Drosophila* enhancer of split (groucho) locus lacking WD-40 repeats. *J. Biol. Chem.* **268:** 25681-25686.

Schroeter, E. H., Kisslinger, J.A. and Kopan, R. (1998) Notch-1 signalling requires ligand-induces proteolytic release of intracellular domain. *Nature* **393:** 382-386.

Schrons, H., Knust, E. and Campos-Ortega, J.A. (1992) The *Enhancer of split* complex and adjacent genes in the 96F region of *Drosophila melanogaster* are required for segregation of neural and epidermal progenitor cells. *Genetics* **132:** 481-503.

Schweisguth, F., and Posakony, J.W. (1994) Antagonistic activities of *Supressor of Hairless* and *Hairless* control alternative cell fates in the *Drosophila* adult epidermis. *Development* **120:** 1433-1441.

Schweisguth, F., and Posakony, J.W. (1992) *Suppressor of Hairless*, the *Drosophila* homolog of the mouse recombination signal-binding protein gene, control sensory organ cell fates. *Cell* **69:** 1199-1212.

Schweisguth, F. (1995) *Supressor of Hairless* is required for signal reception during lateral inhibition in the *Drosophila* pupal notum. *Development* **121:** 1875-1884.

Sharief, F. S., Tsoi, S.C. and Li, S.S. (1997) cDNA cloning and genomic organization of enhancer of split groucho gene from nematode *Caenorhabditis elegans*. *Biochem. Mol. Biol. Int.* **43:** 327-337.

Simpson, P., and Morata, G. (1981) Differential mitotic rates and patterns of growth in compartments in the *Drosophila* wing. *Dev. Biol.* **85:** 299-308.

Simpson, P., and Carteret, C. (1990) Proneural clusters: equivalence groups in the epithelium of *Drosophila*. *Development* **110:** 927-932.

Singson, A., Leviten, M.W., Bang, A.G., Hua, X.A. and Posakony, J.W. (1994) Direct downstream targets of proneural activators in the imaginal disc include genes involved in lateral inhibitory signaling. *Genes Dev.* **8:** 2058-2071.

Skaer, H. (1989) Cell division of malpighian tubule development in *D. melanogaste*r is regulated by a single tip cell. *Nature* **342:** 556-569.

Skeath, J. B., and Doe, C.Q. (1996) The *achaete-scute* complex proneural genes contribute to neural precursor specification in the *Drosophila* CNS. *Curr. Biol.* **6:** 1146-1152.

Skeath, J. B., and Carroll, S.B. (1991) Regulation of *achaete-scute* gene expression and sensory organ pattern formation in the *Drosophila* wing. *Genes Dev.* **5:** 984-995.

Skeath, J. B., and Carroll, S.B. (1992) Regulation of proneural gene expression and cell fate during neuroblast segregation in the *Drosophila* embryo. *Development* **114:** 939-946.

Slack, F., and Ruvkun, G. (1997) Temporal pattern formation by heterochronic genes. *Annu. Rev. Genet.* **31:** 611-634.

Spee, J. H., de Vos, W.M. and Kuipers, O.P. (1993) Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Res.* **21:** 777-778.

Spradling, A. (1986). In *Drosophila,* a practical approach., D. B. Roberts, ed. (Oxford: IRL Press).

Srivastava, D. C., P. and Olson, E.N. (1995) A subclass of bHLH proteins required for cardiac morphogenesis. *Science* **270:** 1995-1999.

Stern, C. (1954) Two or three bristles. *Am. Sci.* **42:** 213-247.

Stifani, S., Blaumueller, C.M., Redhead, N.J., Hill., R.E. and Artavanis-Tsakonas, S. (1992) Human homologs of a *Drosophila* Enhancer of split gene product define a novel family of nuclear proteins. *Nature Genet.* **2:** 119-127.

Struhl, G., Fitzgerald, K. and Greenwald, I. (1993) Intrinsic activity of the Lin-12 and Notch intracellular domains *in vivo*. *Cell* **74:** 331-345.

Struhl, G., and Adachi, A. (1998) Nuclear access and action of Notch *in vivo*. *Cell* **93:** 649-660.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., Mello, C.C. (1999) The *rde-1* Gene, RNA interference, and Transposon Silencing in *C. elegans*. *Cell* **99:** 123-132.

Takebayashi, K. A., C., Nakanishi, S. and Kageama, R. (1995) Structure and promoter analysis of the gene encoding the mouse helix-loop-helix factor Hes-5. Identification of the neural precursor cell-specific promoter element. *J. Biol. Chem.* **270:** 1342-1349.

Tata, F., and Hartley, D.A. (1995) Inhibition of cell fate in *Drosophila* by *Enhancer of split* genes. *Mech. Dev.* **51:** 305-315.

Tata, F., and Hartley, D.A. (1993) The role of the *Enhancer of split* complex during cell fate determination in *Drosophila*. *Development* Suppl. 139-148.

Technau, G. M., and Campos-Ortega, J.A. (1986) Lineage analysis of transplanted individual cells in embryos of *Drosophila melanogaster*. II. Commitment and proliferative capabilities of epidermal and neural cell progenitors. *Roux's Arch. Dev. Biol.* **195:** 445-454.

Tietze, K., Oellers, N. and Knust, E. (1992) Enhancer of split^{D}, a dominant mutation of *Drosophila* and its use in the study of functional domains of a helix-loop-helix protein. *Proc. Natl. Acad. Sci. USA* **89:** 6152-6156.

Timmons, L., and Fire, A. (1998) Specific interference by ingested dsRNA. *Nature* **395:** 854.

Triezenberg, S. J., Kingsbury, R.C. and McKnight, S.L. (1988) Functional dissection of VP16, the trans-activator of herpes-simplex virus immediate early gene-expression. *Genes Dev.* **2:** 718-729.

Truman, J. W., Taylor, B.J. and Awad, T.A. (1993). Formation of the adult nervous system. In The development of *Drosophila melanogaster*., M. a. M.-A. Bate, A., ed. (Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.), pp. 1245-1275.

Truman, J. W., and Bate, M. (1988) Spatial and temporal patterns of neurogenesis in the CNS of *Drosophila melanogaster. Dev. Biol.* **125:** 146-157.

Tsou Tsan, J., Wei Wang, Z., Jin, Y., Hwang, L-Y., Bash, R.O., and Baer, R. (1997). Mammalian Cells as Hosts for Two-hybrid Studies of Protein-Protein Interaction. In The Yeast Two-hybrid system., P. L. Bartel, and Fields, S., ed. (NY.: Oxford University Press.), pp. 217-232.

Vaessin, H., Caudy, M., Bier, E., Jan, L.Y. and Jan. Y.N. (1990) Role of helix-loop-helix proteins in *Drosophila* neurogenesis. *Cold Spring Harbor Symp. Quant. Biol.* **55:** 239-245.

van der Voorn, L., and Ploegh, H.L. (1992) The WD-40 repeat. *FEBS Lett.* **2:** 131-134.

Van Doren, M., Ellis, H.M. and Posakony, J.W. (1991) The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development* **113:** 245-255.

Van Doren, M., Bailey, A.M., Esnayra, J. Ede, K. and Posakony, J.W. (1994) Negative regulation of proneural gene activity: Hairy is a direct repressor of *achaete. Genes. Dev.* **8:** 2729-2742.

Van Doren, M., Powell, P.A., Pasternak, D., Singson, A. and Posakony, J.W. (1992) Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *extramacrochaetae. Genes Dev.* **6:** 2592-2605.

Vässin, H., Brand, M., Jan, L.Y. and Jan. Y.N. (1994) *daughterless* is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in *Drosophila* embryos. *Development* **120:** 935-945.

Vässin, H., Vielmetter, J. and Campos-Ortega, J.A. (1985) Genetic interactions in early neurogenesis of *Drosophila melanogaster. J. of Neurogenet.* **2:** 291-308.

Vervoort, M., Dambly-Chaudiere, C. and Ghysen, A. (1997) Cell fate determination in *Drosophila. Curr. Opin. Neuro.* **7:** 21-28.

Vidal, M., Brachmann, R.K., Fattaey, A., Harlow, E. and Boeke, J.D. (1996) Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc. Natl. Acad. Sci. USA* **93:** 10315-10320.

Wainwright, S. M., and Ish-Horowicz, D. (1992) Point mutations in the *Drosophila hairy* gene demonstrate *in vivo* requirements for basic, helix-loop-helix, and WRPW domains. *Mol. Cell. Biol.* **12:** 2475-2483.

Wang, S., Younger-Shepherd, S., Jan, L.Y. and Jan, Y.N. (1997) Only a subset of binary cell fate decisions mediated by numb/Notch signalling in *Drosophila* sensory organ lineage requires *Suppressor of Hairless*. *Development* **124:** 4435-4446.

Waterhouse, P. M., Graham, M.W., and Wang, M.B. (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. USA* **95:** 13959-13964.

Weintraub, S. J., Prater, C.A. and Dean, D.C. (1992) Retinoblastoma protein switches the E2F site from positive to negative element. *Nature* **358:** 259-261.

Welshons, W. J. (1956) Dosage experiments with *split* mutants in the presence of an Enhancer of *split*. *Dros. Inf. Serv.* **30:** 157-158.

Wettstein, D. A., Turner, D.L. and Kintner, C. (1997) The *Xenopus* homolog of *Drosophila Supressor of Hairless* mediates Notch signaling during primary neurogenesis. *Development* **124:** 693-702.

Wharton, K. A., Johansen, K.M., Xu, T. and Artavanis-Tsakonas, S. (1985) Nucleotide sequence from the neurogenic locus *Notch* implies a gene roduct that shares homology with proteins containing EGF-like repeats. *Cell* **43:** 567-581.

Wigglesworth, V. B. (1940) Local and general factors in the development of 'pattern' in *Rhodnius prolixus* (Hemiptera). *J. Exp. Biol.* **17:** 180-200.

Wilcox, M., Mitchison, G.J. and Smith, R.J. (1973) Pattern formation in the blue-green algae Anabaena. I. Basic mechanisms. *J. Cell Sci.* **12:** 707-723.

Wolk, C. P., and Quine, M.P. (1975) Formation of one-dimensional patterns by stochastic processes and by filamentous blue-green algae. *Dev. Biol.* **46:** 370-382.

Wylbeck, C., and Campos-Ortega, J.A. (1997) Two zebrafish homologues of the *Drosoophila* neurogenic gene *groucho* and their pattern of transcription during early embryogenesis. *Dev. Genes Evol.* **207:** 156-166.

Xu, T., Rebay, I., Fleming, R.J., Nelson Scottgale, T. and Artavanis-Tsakonas, S. (1990) The *Notch* locus and the genetic circuitry involved in early *Drosophila* neurogenesis. *Genes Dev.* **4:** 464-475.

Yew, P. R., Liu, X. and Berk, A.J. (1994) Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53. *Genes Dev.* **8:** 190-202.

Zhang, B., Gallegos, M., Puoti, A., Fields, S., Kimble, J. and Wickens, M. (1997) A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390:** 477-484.

Ziemer, A., Tietze, K., Knust, E. and Campos-Ortega, J.A. (1988) Genetic analysis of *Enhancer of split*, a locus involved in neurogenesis in *Drosophila melanogaster*. *Genetics* **119:** 63-74.