Analysis of *slmo*, a gene required for normal motor function in *Drosophila melanogaster*

A thesis submitted for the degree of Doctor of Philosophy

University of Warwick

By
Christopher T. Dee

Department of Biological Sciences
University of Warwick

May 2004

Table of Contents

List of Figures

List of Tables

Acknowledgments

Many thanks go to all of the following people. Kevin Moffat for his always selfless and attentive supervision, and for his practical contribution to the yeast-2-hybrid screen. Richard Baines, who performed all electrophysiology, provided essential help with dissections, and who's thoughtful advice on this project was always greatly appreciated. Che-Hsiung Liu (my mentor for two years), Chris Mee, Ed Pym, Wendy Moore, Pete Hawkins and Nara Muraro for help, advice and entertainment value. Finally, thanks to my Mum and Dad, and especially Crystal for often necessary support and encouragement.

Some of the data in this thesis has been published in:

Carhan, A., Reeve, S., Dee, C.T., Baines, R.A. and Moffat, K.G. (2003) Mutation in *slowmo* causes defects in *Drosophila* larval locomotor behaviour. *Invert Neurosci.* [Epub ahead of print]

Declaration

I hereby declare that all of the work reported in this thesis is my own unless stated otherwise in the text or figure legend. None of this work has been previously submitted for any other degree or at any other institution. All sources of information used in the preparation of this thesis are indicated by reference.

C.T. Dee

Summary

The stereotyped motor behaviour of the *Drosophila* larva provides a useful model for the development and function of neural circuitry. However, the cellular and molecular basis of this behaviour is poorly understood. Reported in this thesis is an analysis of *slowmo* (*slmo*), a gene previously shown to be expressed in the developing embryonic central nervous system of *Drosophila*. Null mutants of *slmo* are able to hatch, but the resulting larvae exhibit a progressive phenotype of defective locomotor activity. An enhancer trap, P(GAL4)c682, which is inserted in the *slmo* locus, reports expression in a subset of neurons within the embryonic, larval and adult central nervous systems. Inactivation of marked neurons with tetanus toxin light chain leads to severely impaired motor function, but not total paralysis. Affected embryos are capable of some sporadic movements, but are unable to hatch and normal peristaltic contraction waves are largely absent. The *slmo* gene is shown to encode a member of a novel family of conserved proteins of unknown function. GFP fusions of Slmo are shown to localise to the mitochondria in a cultured cell line. Two novel *slmo* related genes, termed *prelil* (*prel*) and *real-time* (*retm*), are identified in *Drosophila*. The *prel* gene is expressed ubiquitously during embryonic development, and disruption of the gene by a P-insertion results in lethality during larval development. *retm* encodes a member of a novel subclass of larger Slmo related proteins which contain the conserved CRAL-TRIO domain thought to be involved in the transport of small hydrophobic ligands. GFP fusions of both Prel and Retm are also associated with the mitochondria in cell culture, suggesting this might be true all proteins of this family. Using a yeast-2-hybrid approach, the identification of candidate Slmo interacting proteins is described, providing a basis for future work on the function of Slmo.

List of Common Abbreviations

A – adenine

ATP – adenosine triphosphate

b – base

BDGP – Berkeley *Drosophila* genome project

BLAST – basic local alignment search tool

bp – base pair

C – cytosine

°C- degrees Celsius

cDNA – complementary deoxyribonucleic acid

CNS – central nervous system

CPG – central pattern generator

CS – Canton-S

CyO – Curly of Oster

dH_20 – distilled water

DIG - Digoxygenin

DNA –deoxyribonucleic acid

DNase – deoxyribonuclease

dNTP – deoxynucleotide triphosphate

EDTA – ethyenediaminetetraacetate

EST – expressed sequence tag

G – guanine

g –gram

GFP – green fluorescent protein

GSC – germline stem cell

hrs – hours

his – histidine

kb – kilobase

kDa – kiloDalton

l – litre

L1, L2, L3 – first, second or third instar larva

LB – Luria Bertani broth

leu – leucine

M – Molar

μg – microgramme

mins - minutes

μl – microlitre

ml – millilitre

mM – millimolar

mRNA – messenger ribonucleic acid

MSF1' – mitochondrial sorting factor 1'

mV – millivolts

ng – nanogram

nm - nanometre

OD – optical density

ORF – open reading frame

PAGE – polyacrylamide gel electrophoresis

PBS – phosphate buffered saline

PBT – phosphate buffered saline plus 0.1% (w/v) Tween-20

PBTx – phosphate buffered saline plus 0.1% (w/v) Triton X-100

PCR – polymerase chain reaction

PH- post-hatching

PNS – peripheral nervous system

prel – prelil

PRELI – protein of relevant evolutionary and lymphoid interest

retm – real-time

RNA – ribonucleic acid

RNase – ribonuclease

rpm – revolutions per minute

RT-PCR – reverse transcriptase polymerase chain reaction

SDS – sodium dodecyl sulphate

secs - seconds

slmo – slowmo

SSC – somatic stem cell

T – thymine

TBE – tris-borate-ethylenediamine buffer

TeTxLC – tetanus toxin light chain

Tris – tris (hydroxymethyl) aminomethane

trp - tryptophan

UAS – upstream activating sequence

ura – uracil

v/v – volume / volume

VNC – ventral nerve cord

WCS – *white* Canton-S

w/v – weight/ volume

YNB – yeast nitrogen base (yeast "dropout" media)

YPD – yeast extract peptone dextrose medium (yeast rich medium)

X-gal – 5-bromo-4-chloro-3-indoyl-β-D-galactopyroside

Chapter 1: Introduction

Chapter 1: Introduction

1.1 *Drosophila* as a model organism

The nervous system represents the most complex organ system of animals, and is composed of a vast range of neurons and glia, which interact in a precise and intricate manner. The neurons extend axons, which must form the appropriate synaptic connections in a highly complex neural network. This network provides the basis for critical functions such as the control of behaviour, learning and memory. The *Drosophila melanogaster* nervous system has long been established as an excellent model system in which to study the development and function of the nervous system. Its relative simplicity, combined with an amenability to powerful genetic, molecular, anatomical and electrophysiological techniques make it possible to dissect the emergence and function of nervous system structure, as well as the roles of the individual circuits, cells, and molecules within it.

Much research has focussed on the adult fly, where the isolation and characterisation of mutations which modify behaviour has provided insights into the molecular and cellular basis for a variety of adult behaviours such as circadian rhythms (Hall, 1998), olfaction (Carlson, 1996), courtship (Hall, 1994; Yamamoto *et al.*, 1997), hearing (Eberl, 1999), learning (Carew, 1996; Davis, 1996; Dubnau and Tully, 1998), vision (Zuker, 1992; Zuker, 1996) and locomotion (Strauss and Heisenberg, 1993). In addition, it has been possible to link certain behavioural tasks to particular structures of the brain. For example, it has been shown that the mushroom bodies are critical for complex behaviours such as olfactory learning and memory (Heisenberg, 2003; Roman and Davis, 2001), control of spontaneous walking activity (Martin *et al.*, 1998), and the mediation of context generalisation in

visual learning (Liu *et al.*, 1999; Waddell and Quinn, 2001). Similarly, the central complex has been shown to be involved in the higher control and modulation of locomotor behaviour (Martin *et al.*, 1999; Strauss, 2002), and it is now well established that the "jump response" escape behaviour is mediated by the giant fibre system (Allen *et al.*, 1998; Allen *et al.*, 2000).

The developing *Drosophila* embryo on the other hand, has proven invaluable as a model for the establishment of neuronal circuits, in this case, of the larval nervous system. During embryogenesis, the machinery underlying larval behaviour is assembled, and circuits are generated and begin to function. In particular, there has been rapid progress in unravelling the mechanisms involved in early development of the nervous system, such as neurogenesis, axon growth and guidance, target recognition and the formation of synaptic terminals (Bhat, 1998; Goodman, 1993; Prokop, 1999; Tessier-Lavigne and Goodman, 1996). More recently, some progress has been made on the later stages of nervous system development when the basic connectivity is established, and the component neurons develop their electrical properties while maturing as part of a functional circuit (Baines, 2003; Baines *et al.*, 1999; Baines *et al.*, 2002; Baines *et al.*, 2001).

1.2 Larval locomotor behaviour

Drosophila larvae crawl by performing rhythmic, longitudinal waves of cyclic muscle contraction and extensions of the body known as telescopic peristalsis (Berrigan, 1995). When larvae move spontaneously across a plain non-nutritive substrate, periods of linear locomotion are interrupted by phases of head swinging and digging, followed by a change in direction. Larval locomotion therefore represents a simple, stereotyped behaviour comprised of episodes of co-ordinated

motor output and decision making (Wang *et al.*, 1997). However, the basic machinery which drives the peristalsis performed by larvae is primarily established during the later stages of embryonic development, and underlies the vigorous movements of the embryo prior to and during the hatching process (13-16 hours after laying; hatching occurs at approximately 21 hours) (Kaliss, 1939; Suster and Bate, 2002). Repeated waves of forward and backward peristalsis begin to occur within the vitelline membrane, and are supplemented by swinging and wriggling movements of the head. Once the egg shell is punctured, repeated forward peristaltic movements drive the new larva out of its vitelline and chorionic membranes (Siekhaus and Fuller, 1999).

An important achievement in neurobiology would be to elucidate both the development, and regulation, of the functional properties of defined neuronal circuits that underlie particular behaviours. The relatively simple, coordinated motor output that drives larval locomotion provides a compelling model, in which a range of powerful techniques can be employed (see below). In particular, it may be possible to identify exactly which cells are involved in driving this behaviour, which molecules are required to regulate the functional development of these cells, and which genes are involved in modulating the output of such a circuit in order to meet the requirements of a specific behaviour.

1.3 Dissecting the circuitry

In the last few years, some studies have begun to examine the nature of the circuitry that drives larval locomotion in *Drosophila*. In general, rhythmic motor activities are based on the output of specialised neural networks called central pattern generators (CPGs) (Marder and Bucher, 2001). In most systems, these represent

circuits of pre-motor interneurons that are capable of generating patterned, rhythmic discharges in motor neurons. The defining quality of a CPG is that it is capable of producing this rhythmic activity independently (in the absence of extrinsic timing information), and indeed, many isolated central networks will generate fictive motor patterns as *in vitro* preparations (Marder and Bucher, 2001). It is this property that allows these circuits to form the basis of a wide range of repetitive behaviours, such as feeding in crustaceans (Nusbaum and Beenhakker, 2002) and molluscs (Jing and Weiss, 2002), swimming in *Xenopus* (Sun and Dale, 1998) and zebrafish (Drapeau *et al.*, 2002; Saint-Amant and Drapeau, 2001), and respiration in both mammals (Bellingham, 1998) and crabs (Dicaprio, 1997; Dicaprio *et al.*, 1997). In insects, CPGs located in the ventral nerve cord ganglia, have been shown to drive walking behaviour of adult locusts (Ryckebusch and Laurent, 1994), larval feeding in *Drosophila* (Gorczyca *et al.*, 1991) and larval locomotion in the hawkmoth (Johnston *et al.*, 1999; Johnston and Levine, 1996). Electrophysiological evidence supports the notion that this is also true for the rhythmic peristalsis in *Drosophila* larvae (Baines *et al.*, 1999; Baines *et al.*, 2002; Baines *et al.*, 2001; Cattaert and Birman, 2001). Despite this, the development of CPGs is still poorly understood (Fenelon *et al.*, 1998; Marder and Bucher, 2001).

Although a CPG is capable of generating the basic rhythmic output that drives such behaviours in isolation, in reality movement depends on interactions of the CPG with sensory feedback and descending commands from the brain (Dickenson, 2000; Marder and Bucher, 2001). These act to activate, block and modulate the output of these circuits according to the requirements of the particular behaviour. In addition, multiple CPGs must interact with each other in order to coordinate muscle innervation in the appropriate pattern. In this case, the muscles of

different segments must contract in the correct temporal-spatial pattern for telescopic peristalsis. Furthermore, any CPG must interact, directly or indirectly, with the appropriate motoneurons. The importance of these extra layers of complexity was elegantly illustrated by the elimination of all sensory input from a developing *Drosophila* embryo (Suster and Bate, 2002). Under these circumstances, the CPG driving locomotory peristalsis is still able to develop and produce coordinated rhythmic output. Despite this, the resultant larvae exhibited severely abnormal locomotor activity. These larvae are able to produce coordinated peristaltic contraction waves, but sensory feedback is required to ensure the predominant direction of peristaltic waves is from posterior to anterior, and therefore, to integrate the rhythmic output into an actual pattern of locomotion. It has also been shown that the motoneurons that excite locomotory muscles require synaptic excitation in order to acquire their normal electrical properties (Baines *et al.*, 2001), the implication being that the precise functional development of individual components of the circuit is not autonomous, but dependent on the integrated circuit as a whole. The overall circuitry that governs larval locomotion therefore carries a considerable amount of subtlety, but still represents a more easily interpretable system than that of a higher cognitive behaviour such as learning.

The isolation of mutants has proven widely informative in establishing the biochemical basis of behaviour and neural development in *Drosophila* (1.1). Mutant screens in the vertebrate system of zebrafish have also demonstrated the potential of a forward genetic approach to understanding the nature of locomotor circuits (Bate, 1999; Drapeau *et al.*, 2002; Granato *et al.*, 1996). Additional advantages of *Drosophila* are the availability of neuronal markers, in the form of enhancer traps (O'Kane and Gehring, 1987), and the ability to spatially target the expression of

transgenes using the GAL4-UAS system (Brand and Perrimon, 1993). In this thesis, I have applied these approaches in an attempt to identify cells and molecules involved in regulating the development, maturation and function of the larval locomotor circuit.

1.4 P-element mutagenesis

Many of the classical genetic screens in *Drosophila* involved the use of chemical mutagenesis, using ethyl methane sulphonate (EMS) or ethyl nitroso urea (Nusslein-Volhard and Wieschaus, 1980; Wieschaus *et al.*, 1984). This method is highly efficient at inducing mutations, but the changes made to the DNA are subtle, often of a single nucleotide. Mutations identified by phenotype are therefore difficult to isolate and clone. Transposable elements also provide a potent means of disrupting gene function (Kidwell, 1986). The ability of these mobile fragments to excise from and reinsert into the genome creates the possibility that a genetic lesion will be created at the insertion site. This approach greatly facilitates cloning as the identified transposon leaves a molecular tag in or near the interrupted gene (Cooley *et al.*, 1988). Over the last couple of decades, extensive research has focussed on the P-family of transposable elements, which have been developed to form the basis of a range of powerful genetic manipulations.

P-elements are naturally occurring transposons that move via a non-replicative mechanism without the need for an RNA intermediate. A full length P-element contains 31bp terminal inverted repeats and 11bp subterminal inverted repeats, which are critical for efficient transposition (O'Hare and Rubin, 1983). Transposition also requires the action of a transposase, which can be encoded from within the P-element, or from an extrinsic source (another P-element) (Engels, 1984;

Kidwell, 1986). Importantly, P-elements exhibit a high degree of mobility, and are able to retain this mobility despite drastic modifications to their internal sequences. Experimentally modified P-elements can therefore be used to manufacture stable transgenic strains (Fujioka *et al.*, 2000). In addition, movement of the P-element can be controlled by limiting the availability of the transposase (Engels, 1984; Kidwell, 1986). As result, it has been possible to mobilise engineered P-elements in large genetic screens, and isolate vast numbers of stable mutant lines. In particular, this technique has been successfully applied by the Berkley *Drosophila* genome project (BDGP) as part of the gene disruption project (Spradling *et al.*, 1999) (http://www.fruitfly.org/p_disrupt/index.html), which has now isolated corresponding mutants of more than 25% of essential genes. In addition, a large database of P-element insertion lines has also been generated commercially (http://genexel.com/eng/htm/genisys.htm).

1.5 The enhancer trap

The enhancer trap technique initially developed by O'Kane and Gehring (1987), offers an alternative approach to the traditional P-element mutagenesis screen, and has been employed to identify and clone many neurally important genes. This method is based on engineered P-elements which contain a reporter gene under the control of a weak constitutive promoter. Significant levels of reporter gene expression require that the P-element become inserted into the *Drosophila* genome close to an endogenous enhancer of gene expression (Bellen *et al.*, 1989). The pattern of reporter expression should therefore reflect part or all of that of the endogenous gene associated with the regulatory element. In this way it is possible to generate large numbers of enhancer trap insertion lines and screen for genes by

expression pattern, without requiring the selection of a phenotype (Wilson *et al.*, 1989). A researcher interested in the larval locomotor circuit might therefore look for enhancer traps expressed in the embryonic and larval ventral nerve cord. In addition, the insertion site may have disrupted either the coding sequence of the corresponding gene, or an important regulatory sequence (see section 1.4). Where this is not the case, it is often possible to generate deletions of the gene via imprecise excision of the P-element, or selection of male recombination events (Kidwell, 1986; Preston *et al.*, 1996).

Many enhancer traps employ the *E.coli lacZ* gene as an easily assayable reporter. However, the system is far more versatile when the reporter used is GAL4, a yeast transcriptional activator which is functional in *Drosophila* (Brand and Perrimon, 1993). GAL4 will drive the expression of any gene engineered under the control of the "upstream activating sequence" (UAS). Rather than visualising GAL4 directly, it can be used to drive a variety of secondary reporters with different advantageous properties, such as UAS-*lacZ* or UAS-Green Fluorescent Protein (GFP). With the pattern of expression established, this system can be used to drive a range of UAS transgenes, allowing *in vivo* manipulation of the GAL4 expressing cells. It is this property that has made the GAL4-UAS system a particularly useful tool.

1.6 Applications of the GAL4 system in exploring neural circuitry

The spatial control of transgene expression provided by the GAL4-UAS system has proven one of the most productive experimental techniques applied to the *Drosophila* system (Brand and Perrimon, 1993). The versatility of this genetic tool provides an elegant method for elucidating the nature of poorly defined neural

circuits. Initially, this system facilitates the anatomical characterisation of specific neuronal circuitry by targeted expression of a variety of different reporter genes. In addition, it is possible to control the expression of toxic gene products, which will disrupt the normal function of defined groups of cells in an otherwise normally functioning fly. Indeed, relatively inaccessible cells can be disabled *in vivo* as long as there is an appropriately expressing GAL4 line. In this way, it is possible to examine the role of particular cells and cell-cell interactions, an important step in understanding the cellular basis of behaviour.

The GAL4 system provides a means with which to label and visualise neural pathways, whilst avoiding the technical limitations associated with direct dye-filling of small *Drosophila* neurons with tracers such as DiI or Lucifer yellow (Bossing and Technau, 1994; Sink and Whitington, 1991). Various UAS-*tau* transgenes have been used to visualise the axons of specific neurons (Ito *et al.*, 1997). By simply crossing certain GAL4 lines with transgenic lines bearing UAS-*tau*, the microtubule based Tau protein can be introduced into particular neurons where it is actively distributed into all cellular processes. Tau protein can then be detected directly with an antibody (Ito *et al.*, 1997), but has also been expressed as fusions with GFP and *lacZ* (Hidalgo *et al.*, 1995; Murray *et al.*, 1998). More recently UAS-mCD8GFP has been widely used for the same purpose, in this case allowing the targeted expression of a fusion protein which anchors brightly fluorescent GFP to the cell membrane (Lee and Luo, 1999). The plant lectin wheat germ agglutinin (WGA), a protein that has been shown to transfer across synapses, has also been expressed in *Drosophila* neurons under UAS control as a method for identifying the pre-synaptic targets of GAL4 expressing neurons (Tabuchi *et al.*, 2000).

In some cases the GAL4 system has been used to express toxins such as the A chains of ricin (RTA) (Allen *et al.*, 2002; Hidalgo *et al.*, 1995) and diptheria (DTA) (Lin *et al.*, 1995), as an effective method with which to selectively ablate particular populations of cells and examine the developmental consequences. Alternatively, the gene products of genes such as *ced3*, *ice*, *head involution defective* (*hid*) and *reaper* (*rpr*) can be used to elicit cell death in *Drosophila* by driving them into apoptosis (Shigenaga *et al.*, 1997; Zhou *et al.*, 1997). This approach has proved useful in describing a number of developmental phenomena. For example, specific ablation of embryonic glia by targeted expression of a UAS-RTA transgene demonstrated the crucial role of these cells in the establishment and maintenance of axon tracts (Hidalgo *et al.*, 1995). Likewise, cell ablation by expression of both UAS-RTA and UAS-DTA has been used to examine the role of pioneer neurons in the developing embryonic CNS (Hidalgo and Brand, 1997). This technique has also been used to ablate brain structures thought to be mediating particular behaviours (McNabb *et al.*, 1997; Renn *et al.*, 1999). However, for behavioural studies, the usefulness of targeted cell ablation is often limited given that the death of a cell can affect the fate, development or function of surrounding neurons (Lin *et al.*, 1995).

The tetanus toxin light chain (TeTxLC) gene (Sweeney *et al.*, 1995) and a semi-dominant, temperature sensitive allele of the *Drosophila shibire* gene (shi^{ts1}) (Kitamoto, 2001) have both been brought under control of the GAL4 system as a means of silencing neuronal function, whilst avoiding the problems associated with cell ablation. Tetanus toxin belongs to the family of clostridial neurotoxins and in nature consists of two polypeptide chains. The heavy chain mediates the translocation of the light chain to the neuronal cytosol, where the light chain acts catalytically to inhibit synaptic transmission from the cell (Schiavo *et al.*, 2000).

Delivery of the isolated TeTxLC into neuronal cells (e.g. using the GAL4 system) is sufficient by itself to inhibit all exocytotic neurotransmitter release. The toxin operates by proteolytic cleavage of a small membrane protein called neuronal Synaptobrevin (n-Syb) (also known as the vesicle associated membrane protein or VAMP), an essential component of the machinery regulating exocytosis of synaptic vesicles (Sweeney *et al.*, 1995). TeTxLC will therefore silence the synaptic output of any neuron in which it is expressed, but its mode of action is also specific to neuronal cells. This property increases the number of GAL4 lines available to the experimenter, as any extraneous expression of the toxin in non-neuronal tissues, such as glia or muscle, can effectively be ignored.

Unsurprisingly, embryos that express TeTxLC throughout the nervous system are paralysed, and do not show the coordinated peristaltic contractions characteristic of the later stages of embryogenesis (Sweeney *et al.*, 1995). Importantly however, the embryonic nervous system assembles normally in the absence of evoked synaptic release. There are no apparent changes to the morphology of the nervous system, and synapses are able to form on muscles as normal. UAS-TeTxLC is therefore regarded as a method for selectively inactivating groups of neurons in an otherwise developmentally and functionally normal nervous system. In reality of course, inhibition of synaptic transmission from a neuron can have more subtle developmental consequences, as it will inevitably disrupt the activity dependant phase of development by which neurons develop their functional characteristics. In this way, the tetanus toxin offers a method by which to learn about this important aspect of neuronal development. For example, Baines *et al.* (1999) were able to demonstrate that targeted expression of TeTxLC in particular embryonic motoneurons not only caused the failure of synaptic transmission between these

neurons and their target muscles, but that these motoneurons themselves receive substantially less synaptic input. The implication of this result is that a mechanism of retrograde signalling is required from the postsynaptic neuron to modulate the functional properties of its presynaptic partner, and more broadly reinforces the notion that neuronal circuits require functional interplay between individual cells to refine their active properties.

TeTxLC has been employed widely in conjunction with the GAL4 system to analyse a broad range of behaviours in both adults and larvae, including studies of locomotor activity. In the adult brain, targeted expression of TeTxLC in groups of neurons within the mushroom bodies and the central complex has demonstrated the involvement of these structures in regulating locomotor activity. In the case of the mushroom bodies, disruption with the neurotoxin results in increased locomotor activity (Martin *et al.*, 1998), whilst inactivation of particular groups of cells within the central complex was found to causes more subtle defects in the maintenance and temporal patterning of locomotor activity (Martin *et al.*, 2002; Martin *et al.*, 1999). In the larval system, Suster *et al.* (2002) used UAS-TeTxLC to screen for GAL4 lines that express in groups of neurons important for the regulation of locomotion.

This toxigenetic method has made it possible to inhibit the activity of very discreet groups of neurons and examine their role in behaviour. However, the potential of this approach to drive toxin expression in particular cells and brain structures is limited by the available GAL4 lines. This problem is exacerbated by the fact that most GAL4 enhancer trap lines are expressed at various stages of development and that the toxic nature of TeTxLC will impair development at an early stage. UAS-*shi*^{ts1} was envisaged as a means by which to circumvent some of these problems, by inhibiting neuronal activity in a conditional manor (Kitamoto,

2001). The *Drosophila shi* gene encodes the protein dynamin, a vital component of the machinery required for synaptic vesicle recycling. The semi-dominant allele shi^{ts1} is defective in this process at restrictive temperatures (>29°C), and when over expressed, will cause rapid and reversible inhibition of synaptic transmission even in the presence of a wild type allele. The researcher therefore has temporal control over the effector gene, permitting the use of GAL4 lines in later stages of development that might have resulted in earlier lethality when expressing TeTxLC. More recently, attention has been devoted to the development of an inducible GAL4 system, in which it is the transcriptional activator itself that is conditionally active thereby allowing temporal control over the expression of any UAS transgene. The Gene-Switch GAL4 system combines the *S. cerevisae* GAL4 domain with the human progesterone receptor ligand binding domain, and the human p65 transcriptional activation domain (Roman *et al.*, 2001). RU486, an antiprogestin, will bind to the progesterone receptor ligand binding domain and activate the Gene-Switch molecule. The Gene-Switch GAL4 system can therefore be used to generate numerous lines under the control of various enhancer sequences (Roman and Davis, 2002), but restricting the availability of RU486 through specific feeding regimes can be used to control the activity of the GAL4.

1.7 Identification of P(GAL4)c682

P(GAL4)c682, was originally isolated as part of a screen for enhancer trap lines expressed in the adult nervous system (D. Shepherd, University of Southampton). In particular, expression had been described in the cell bodies and axons of the giant fibre system, as with other lines studied in the Moffat lab at the time (Allen *et al.*, 1998). Using UAS-*lacZ* as the reporter, Carhan (1999) was able to

confirm this, and also described expression in other cell bodies of the adult brain and ventral nerve cord. In addition, reporter expression was shown to be restricted to cell bodies in the developing ventral nerve cord of stage 16 embryos. This type of neural expression pattern strongly suggests that the enhancer trap marks a gene which functions in the nervous system. Indeed, following the cloning of the corresponding gene (section 1.8), *in situ* hybridisation using a Digoxygenin labelled DNA probe also showed embryonic expression of the endogenous transcript to be CNS specific. However, it was also noted that males homozygous for the insert were sterile, and it is this phenotype that has formed the basis of previous research by Carhan (1999) and Reeve (2002).

1.8 Molecular Cloning of *slmo*

The gene marked by the P(GAL4)c682 enhancer trap was originally named *kisir* (Turkish for sterile), but is now referred to as *slowmo* (*slmo*) on account of the striking sluggish behaviour exhibited by *slmo* null mutants (see chapter 3). The P(GAL4)c682 insertion was originally mapped to the 26B region of the left arm of the *Drosophila* second chromosome, and genomic DNA flanking the P-element was cloned by plasmid rescue (Bellen *et al.*, 1989; Carhan, 1999; Wilson *et al.*, 1989). It was then possible to isolate the corresponding cDNA by using the rescued sequence to screen existing cDNA (Stratagene) and cosmid libraries (Hoheisel *et al.*, 1991). Examination of the sequence revealed a gene with a simple two-exon structure, and an open reading frame contained entirely within the second exon (Figure 1.1). The EST database also indicates the existence of an additional transcript transcribed from an alternative promoter. The coding sequence is predicted to encode a 215 amino acid protein of 25kDa molecular weight, although later analysis by western blot has

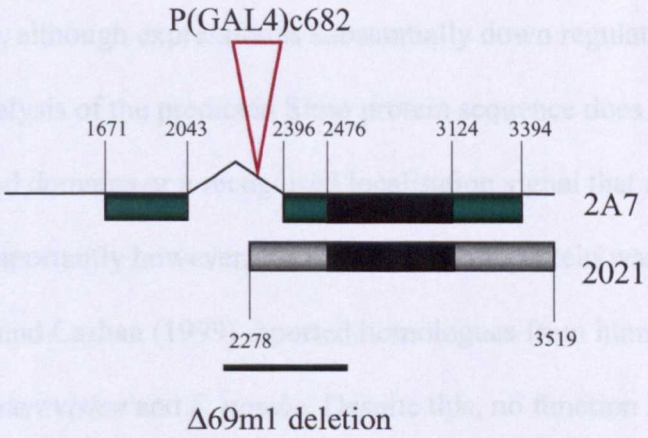


Figure 1.1: **Schematic representation of the *slmo* gene structure**. The *slmo* gene is expressed as two alternative transcripts which differ at the 5' ends. Both transcripts are mapped relative to each other as though transcribed from the same strand. All numbers refer to the 6kb genomic DNA fragment cloned by Carhan (1999) (Appendix 1). The **2A7** transcript (**green blocks**) was originally isolated from a cDNA library (Carhan, 1999), and was found to have a two exon structure. The enhancer trap **P(GAL4)c682** is inserted within the single intron. The second transcript **2021** (**grey blocks**) was identified from the EST database (Reeve, 2002). The ORF (**black blocks**) is identical in both transcripts and is predicted to encode a 215 amino acid protein of unknown function. The **Δ69m1** deletion, was generated by imprecise excision of the P-element, and represents a *slmo* null allele.

shown the protein to migrate on a polyacrylamide gel as a 30kDa molecule, presumably due to post-translational modification (Reeve, 2002). Sequencing out of the P-element revealed that the P(GAL4)c682 insertion site is within the single *slmo* intron, and further analysis of this allele (referred to as $slmo^{c682}$) by northern blot (Carhan, 1999) and western blot (Reeve, 2002) has shown that this does not represent a null allele, although expression is substantially down regulated.

Analysis of the predicted Slmo protein sequence does not reveal any characterised domains or a recognised localisation signal that might offer clues as to function. Importantly however, the predicted Slmo protein was found to be highly conserved, and Carhan (1999) reported homologues from human, mouse, chicken, *C. elegans*, *S. cerevisiae* and *S. pombe*. Despite this, no function has been assigned to any member of this protein family. The closest human homologue, CGI-107 (Lai *et al.*, 2000), exhibits 55% identity and 70% similarity to Slmo at the protein level, and has been identified by the BodyMap project as being expressed in the adult brain (Sese *et al.*, 2001). The chicken protein, px19 (Niu *et al.*, 1996), and the human protein PRELI (Guzman-Rojas *et al.*, 2000) are more distantly related, and are expressed in a variety of tissues including brain, lung, liver, kidney and B-lymphocytes. No other mutant phenotypes have been reported, although it is known that knockouts of homologous genes in *S. cerevisiae* (T. Endo pers. comm., Nagoya University) and *C. elegans* (http://www.wormbase.org, release WS115, date 15/01/04) are not lethal.

1.9 Sterility phenotype of $slmo^{c682}$ homozygotes

The *Drosophila* germline, both in the male (Gonczy *et al.*, 1997) and the female (King and Lin, 1999), are primarily of interest as models for studying stem

cell function. Stem cells have the unique ability to divide asymmetrically, regenerating the parental stem cell, whilst also producing daughter cells able to proliferate and differentiate into specific cell types. In this way, different stem cell populations form the basis of an organism, initially producing the cells of the developing embryo, then later generating and maintaining tissues and organs such as blood (Dzierzak *et al.*, 1998), skin (Toma *et al.*, 2001), nervous system (Doetsch *et al.*, 1999), and the germline (Lin, 2002). The *Drosophila* gonad provides a relatively simple, well-characterised organ in which both the actively dividing stem cells, and the surrounding support cells, are clearly defined. In addition, these cells can be studied and manipulated with the powerful molecular and genetic approaches available in *Drosophila* (Fuller, 1993; Lin, 2002).

Stem cell function is regulated and maintained in a contained micro-environment referred to as a niche. Here essential support cells are thought to send molecular signals to the stem cells, which act to repress differentiation of the parental stem cell and control proliferation and specification of the daughter cells (Spradling *et al.*, 2001). In males the niche is composed of a group of somatic hub cells at the apex of the tubular shaped testes (Figure 1.2). These remain in constant contact with both the germline stem cells (GSCs) and the pairs of somatic stem cells (SSCs) that flank each GSC (Lin, 2002). In this way the hub cells can communicate with the stem cell population via multiple signalling pathways, and maintain the integrity of the stem cell system. As the GSCs divide, the daughter progeny or gonialblasts are displaced away from the hub, and undergo four synchronous divisions with incomplete cytokinesis. The result is the generation of a cyst of 16 interconnected spermatogonia that remain enclosed by two somatic cyst cells (derived from the

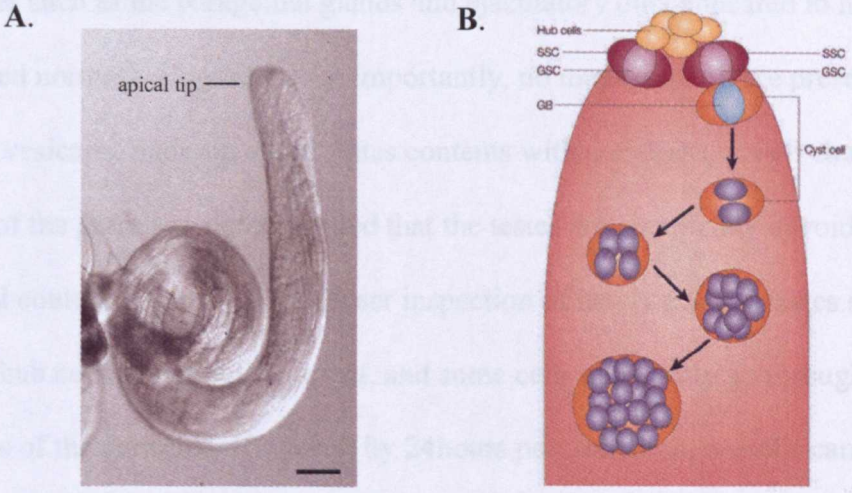


Figure 1.2: Germline and somatic stem cells of the *Drosophila* testis. **A**. The Drosophila testis is a coiled tubular structure. Scale bar is equal to 100μm. **B**. Schematic diagram representing the apical tip of the testis. Only 2 out of 5-9 germline stem cells (**GSC**s) are represented with their accompanying pairs of somatic stem cells (**SSC**s). Both GSCs and SSCs remain in constant contact with somatic hub cells at the apex where they divide to generate a gonialblast (**GB**) surrounded by two cyst cells. The GB undergoes four synchronous divisions with incomplete cytokinesis resulting in 16 interconnected spermatogonia enclosed by the two somatic cyst cells. Figure adapted from Lin (2002) and Carhan (1999).

SSCs). These must then undergo several further phases of division and differentiation to develop into mature sperm (Fuller, 1993).

Carhan (1999) showed that in males homozygous for the *slmo*^{c682} allele, this tightly controlled process is completely disrupted at an early stage. Dissection from 4-10 day old mutant males revealed severely shrivelled testes, while other genital structures such as the paragonial glands and ejaculatory duct appeared to have developed normally (Figure 1.3A). Importantly, no mature sperm are present in the seminal vesicalis. Staining of the testes contents with anti-Vasa, a well characterised marker of the germline, demonstrated that the testes are completely devoid of any germinal content (Figure 1.3C). Closer inspection of newly eclosed testes showed that the hub cells are initially present, and some cells apparently go through the early divisions of the germline. However, by 24hours post-eclosion, no cells can be identified and vacuolations appear, which become rapidly more pronounced over the next few days (Figure 1.3E, F) and may be the result of apoptosis (although this has never been formally tested). Indeed, degeneration of the germline is already apparent in the testes of 3rd instar larvae (Reeve, 2002).

Mutations in many genes can disrupt gametogenesis, largely because the germline is particularly sensitive to defects in general or metabolic functions. However, very few cause clear arrests in the early proliferation phase of spermatogenesis (Castrillon *et al.*, 1993; Fuller, 1993). It has been argued that *slmo* may be in some way involved in the cell-cell signalling events that constitute the stem cell niche (Carhan, 1999; Reeve, 2001). Certainly, it is true that expression reported by promoter constructs driving *lacZ* (Reeve, 2002), appears to be absent from the early germline. This line of thought is also supported by parallel work in the female germline. Carhan (1999) had classified female *slmo*^{c682} homozygotes as

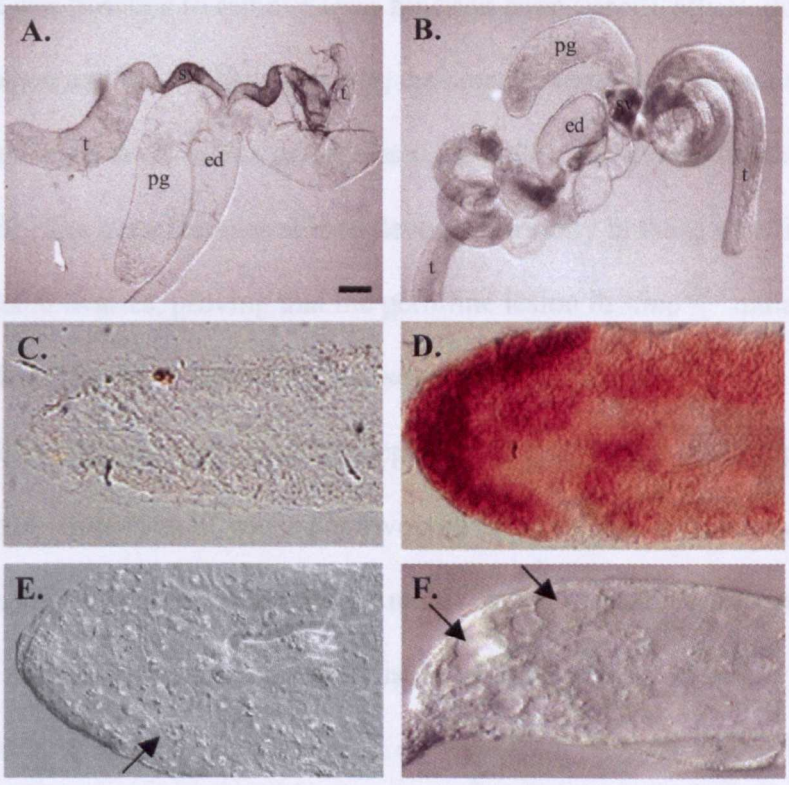


Figure 1.3: **Male sterility phenotype of *slmo*^{c682} homozygotes**. **A**. Testes dissected from *slmo*^{c682} homozygous males show abnormal morphology. Testes (**t**) appear shrivelled when compared to the coiled tubular shaped testes of the control wild type fly (**B**). Other genital structures such as the paragonial glands (**pg**), seminal vesicalis (**sv**) and the ejaculatory duct (**ed**) appear to have developed normally. **C**. Staining with the germline marker anti-Vasa demonstrates a lack of germinal content in the *slmo*^{c682} mutant as compared to the strong staining observed in wild type testes (**D**). **E**. Newly eclosed mutant testes contain some cellular architecture, and cell bodies can be seen (**black arrow**). **F**. In aged testes no cells can be identified, and severe vacuolations (**black arrows**) have developed by 5 days post-eclosion. Scale bar is **A**,**B**: 100μm, **C**-**F**: 10μm. Images taken from Carhan (1999).

having a severe reduction in fecundity, although not completely sterile. Since then Reeve (2002) has described a similar phenotype in the ovaries to that seen in the testes, with the majority of the female germline failing to divide beyond an early stage. Using the FLP-FRT system of site directed recombination (Xu and Rubin, 1993), it was possible to remove *slmo* function from specifically within the germline (Theodosiou and Xu, 1998). However, the somatic terminal filament and cap cells that form the stem cell niche in females (Lin, 2002), still expressed the wild type *slmo* gene. The germline was able to develop normally in these genetically manipulated ovaries, proving that the germline lesion in $slmo^{c682}$ mutant females results from abnormal function of the somatic cells.

However, it is important to keep in mind that similar sterility phenotypes can be caused by mutation of genes not involved in signalling. Alleles of *chickadee* (*chi*) and *diaphanous* (*dia*) result in similar male sterility phenotypes to that seen in $slmo^{c682}$ homozygotes, and both were isolated in screens for male sterile mutants (Castrillion *et al.*, 1993; Fuller, 1993). *chi* encodes a *Drosophila* profilin protein and is known to play an important role in the assembly of cytoskeletal actin filaments (Castrillion *et al.*, 1993; Verheyen and Cooley, 1994), although homologues have been implicated in signal transduction (Lassing and Lindberg, 1985; Goldschmidt-Clermont *et al.*, 1990). *dia* encodes a member of the Formin Homology family, and is critical for actin mediated cytokinesis (Castrillion and Wasserman, 1994; Afshar *et al.*, 2000). In both these cases the germline dies at an early stage, although the somatic cell lineages survive. The use of enhancer trap markers of hub and cyst cell lineages showed that somatic cells are also still present in $slmo^{c682}$ mutant testes (Carhan, 1999). Null alleles of both *chi* and *dia* result in lethality, as is the case for

slmo (see 1.10), whilst the male sterile alleles also disrupt development of the female germline.

1.10 Generation of a *slmo* null allele

The presence of a P-element in the *slmo* intron disrupts, but does not eliminate transcription of the gene. Carhan (1999) performed an excision screen, primarily in an attempt to generate precise excisions of the enhancer trap that returned to fertility, therefore proving the insertion to be the cause of the sterility phenotype. A total of 108 excision events were generated, as determined by reversion to a white eye phenotype (indicating removal of the mini-*white* gene, located as a marker within the P-element). As expected, the majority (85) were restored to fertility, whilst 17 excision lines remained sterile. More importantly, six homozygous lethal lines were also isolated. One of these, designated Δ69m1, was later proven by Southern blot analysis (Reeve, 2002) to represent an imprecise excision event, in which a 5' portion of the *slmo* coding sequence (including the ATG start codon) has been deleted (Figure 1.1). This null allele is referred to as $slmo^{\Delta 69m1}$.

Carhan (1999) reported that embryos homozygous for $slmo^{\Delta 69m1}$ hatch into first instar larvae, but that the mutant larvae are immobile and never reach the second instar stage. Subsequently, Reeve (2002) suggested that newly hatched mutant larvae are mobile, but progress towards paralysis in the following hours. A variety of defects have the potential to cause an immobility phenotype, for example, defective muscle development (Bate, 1993). Alternatively, lack of movement could be a secondary effect of a more global problem. However, given that embryonic expression of *slmo* (as reported by the enhancer trap and *in situ* hybridisation) is

specific to the nervous system, it is reasonable to propose that the defect in *slmo*$^{\Delta 69m1}$ mutants is specific to the nervous system. Staining of embryos with various markers of CNS morphology does not reveal any gross defects in nervous system development (Carhan, 1999). Likewise, examination of the peripheral nervous system (PNS) suggests that motoneuron axons have developed correctly and formed the appropriate synapses (Reeve, 2002).

1.11 Aims

1) Use of P(GAL4)c682 as a tool for examining neuronal circuitry

This enhancer trap had been shown to be expressed specifically, but not ubiquitously within the CNS. I chose to examine expression of the GAL4 line in detail, and explore possible use of the enhancer trap in identifying cells involved in neuronal circuitry.

2) To determine the potential role of *slmo* in the *Drosophila* nervous system

Previous research had clearly implicated an involvement of *slmo* in the development and/or function of the *Drosophila* nervous system. However, at the beginning of this project, this potentially very interesting relationship had been largely neglected. It was therefore desirable to perform a more detailed analysis of the null mutant phenotype and its possible causes.

3) Function of the Slmo protein

Very little is known about any proteins in the Slmo family, and it is important to gain as much insight into the functional properties of these proteins as possible. I chose to use three approaches: First, to establish the subcellular localisation of Slmo,

second, to search for interacting proteins, and third, to perform a functional analysis of the two *Drosophila* homologues I identified from the database.

Chapter 2: Materials and Methods

Chapter 2: Materials and Methods

2.1 Fly manipulations

2.1.1 Fly food

2.1.1.1 Standard fly food

Flies were routinely raised on media consisting of 94g sucrose, 18g yeast, 103g maize, 5.6g agar and 15mls of Nipagen (Tegosept M, p-hydroxy benzoic acid methyl ester, 10% w/v in ethanol) per litre of distilled water. Fly food was stirred well, allowed to simmer for approximately 15mins to kill live yeast, and then allowed to cool prior to the addition of the Nipagen. The food was then dispensed into 1/3 pint bottles or plastic vials, and allowed to set overnight. A few drops of a live yeast suspension was added to each vial or bottle and allowed to dry before use.

2.1.1.2 Apple agar medium

For the collection of embryos and larvae, flies were allowed to mate and lay on an apple juice based medium. Medium reagents (250mls apple juice, 25g sucrose, 25g agar) were made to 1litre in distilled water. Medium was then simmered until the ingredients had fully dissolved, and allowed to cool before the addition of 15mls Nipagen (10%w/v in ethanol). The medium could then be dispensed into 4.5cm Petri dishes and stored at 4°C. Apple plates were warmed to room temperature before use.

2.1.2 Fly stocks

All fly stocks were maintained on standard food in vials at 18°C.

Table 2.1: Fly stocks used

Stock	Details	Genotype	Source
CS	Canton-S (wild-type stock)	+; +; +	Martin Heisenberg, Wurzberg
WCS	Canton-S carrying *white* (w^{1118}) mutation	w^{1118}; +; +	Lab stock
yw	Canton-S carrying *yellow* (y^1) and *white* (w^{1118}) mutation	y^1w^{1118}; +; +	Bloomington stock centre
w; *CyO*/*Sp*	Balancer stock for mapping	*w*; *CyO*/*Sp*	Lab stock
w; *CyO*/*If*; TM6B/MKRS	Double-balancer stock for mapping	*w*; *CyO*/*If*; TM6B/MKRS	Lab stock
slmo^{c682}	Enhancer trap P(GAL4)c682	*w*; *slmo*^{c682}/ *CyO*$^{[kr\text{GAL4,UASGFP}]}$	Lab stock
slmo$^{\Delta 69m1}$	*slmo* null allele	w; *slmo*$^{\Delta 69m1}$/ *CyO*$^{[kr\text{GAL4,UASGFP}]}$	Lab stock
Df(2L)BSC5 /SM6a	Chromosome 2 deficient for the 26B region	Df(2L)BSC5, w[+mC]/SM6a	Bloomington stock centre
UAS-GFP	Expression of GFP	*w*; +; UAS-GFP	Bloomington stock centre
UAS -mCD8GFP	GFP directed to the cell membrane	*w*; *CyO*/ *If*; UAS-mCD8GFP	Sean Sweeney, UCSF, California
UAS-GFPn	GFP localised to the cell nucleus	*w*; +; UAS-GFPn	Bloomington stock centre
UAS-TeTxLC	Tetanus (G) toxin light chain	*w*; +; UAS-TeTxLC	Sean Sweeney, UCSF, California
UAS-TeTxLC -VIF	Inactive form of TeTxLC, negative control	*w*; +; UAS-TeTxLC-VIF	Sean Sweeney, UCSF, California
D42-GAL4, UAS-mitoGFP	GFP is expressed in most motoneurons and is directed exclusively to the mitochondria	*w*; +/+; P(*w*+)D42-GAL4; P{*w*+mC=UASmito-GFP3M2}, *e*	Bill Saxton, Indiana
GFP*slmo*^{2F1}	GFP-*slmo* fusion transgenic	*w*; *CyO*/*If*; pGFP*slmo*^{2F1}	*Drosophila* transgenesis
GFP*slmo*^{3F2}	"	*w*; *CyO*/*If*; pGFP*slmo*^{3F2}	"
GFP*slmo*^{3M1}	"	*w*; *CyO*/*If*; pGFP*slmo*^{3M1}/TM6B	"
GFP*slmo*^{4M1}	"	*w*; pGFP*slmo*^{4M1}/ CyO	"
GFP*slmo*^{4M2}	"	*w*; *CyO*/ *If*; pGFP*slmo*^{4M2}	"
GFP*slmo*^{4F2}	"	*w*; *CyO*/ *If*; pGFP*slmo*^{4F2}/TM6B	"
GFP*slmo*^{4F3}	"	*w*; *CyO*/ *If*; pGFP*slmo*^{4F3}	"
ChaGAL4; UASGFP	GFP expressed in cholinergic neurons	*w*; ChaGAL4; UASGFP	Richard Baines, Warwick
prel$^{l(2)k12402}$	P-insert allele of *prel* (CG8806)	*yw*; *prel*$^{l(2)k12402}$/*CyO*$^{[kr\text{GAL4,UASGFP}]}$	Bloomington stock centre
retm^{KG05639}	P-insert allele of *retm* (CG9528)	y^1;*retm*P{*y*[+mDint2]*w*[BR.E.BR] =SUPor-P}KG05639; *ry*[506]	Bloomington stock centre
*Dr*Δ2-3	Transposase source	*w*; *CyO*/ *Sp*; *Dr*Δ2-3/TM6B	Bloomington stock centre
*Sb*Δ2-3	Transposase source	*w*; *CyO*/ *Sp*; *Sb*Δ2-3/TM6B	Lab stock
retm$^{\Delta 69}$	*retm* (CG9528) deletion allele	*w*; *retm*$^{\Delta 69}$	Excision screen

2.1.3 Behavioural assay

Larvae, aged 0-2hrs and 18-20hrs post-hatching, were genotyped by selection for or against the CyO [*kruppel*GAL4 UASGFP] balancer chromosome (Casso *et al.*, 2000) and then placed on a fresh, room temperature agar plate and allowed 5mins to acclimatise. The number of full body peristaltic contraction waves completed within 3mins was then counted. For each larva, three independent recordings were taken, and then averaged to produce a single data point. A sample of 15 larvae was examined for each genotype. For some samples the non-parametric Kruskal-Wallis test was used to test the significance of the behavioural assay using SPSS 10.0 for Windows (standard version, SPSS Inc.).

2.2 DNA manipulations

2.2.1 Agarose gel electrophoresis

Agarose gels used for DNA analysis were prepared by dissolving 1% (w/v) agarose in 1X TBE buffer (89mM Tris-HCl, pH8, 89mM boric acid, 1mM EDTA), and adding 0.5μg/ml ethidium bromide (Sigma). DNA samples were mixed with DNA loading buffer (10mM Tris-HCl, 50mM EDTA, 10% Ficoll 400, 0.25% Bromophenol blue, 0.25% Xyline Cyanol FF, 0.4% Orange G) before loading. Gels were run at between 100-160V until DNA fragments had migrated the desired distance. A 1kb DNA ladder (Invitrogen) was used to compare size of samples. Gels were visualised on a GelDoc digital UV light box (Bio-Rad) and photographed.

2.2.2 Molecular cloning

2.2.2.1 Restriction digestions

DNA was routinely cut with various restriction enzymes according to the manufacturers instructions. Digestions were performed in 15-100µl reaction volumes containing 1X reaction buffer (appropriate for the particular enzyme) and <10% volume of restriction enzyme. Reactions were incubated at 37°C (or the optimal temperature for the relevant enzyme) for 1-2hrs. Digests were terminated by gel electrophoresis (2.2.1), and the resolved DNA fragments could be purified from the gel using the GeneCleanII kit (QBIOgene) according to the manufacturers instructions.

2.2.2.2 Dephosphorylation of DNA

For dephosphorylation, vector DNA digested for cloning (2.2.2.1) was incubated in a reaction mixture containing 0.2U alkaline phosphatase (calf intestine isolated, Roche), and the appropriate buffer for 15mins at 37°C, then 15mins at 56°C. A further 0.1U of the enzyme was then added, before incubating again for 15mins at 37°C then 15mins 56°C. DNA was purified from the enzyme using the GeneClean II kit (QBIOgene) according to the manufacturers instructions.

2.2.2.3 Ligation of DNA fragments into plasmid vectors

"Sticky end" ligations were performed using T4 DNA ligase (MBI Fermentas), with molar ratio of approximately 1:3 vector to insert. Reactions used between 50-200ng linearised and dephosphorylated vector DNA, 5 units T4 DNA ligase, 1X T4 ligase buffer and were performed in final volumes of 10 or 20µl. Reactions were incubated either 1hr at room temperature or over night at 4°C.

2.2.2.4 LB cultures and LB Agar plates

Bacterial cells used in cloning were grown and maintained in LB (Luria Bertani broth) and LB agar plates containing an appropriate selective antibiotic. Per litre of LB: 10g bacto-tryptone, 5g bacto-yeast extract, 5g NaCl, pH7.5, 15g bacto-agar (for LB agar plates only). Autoclaved LB/ LB agar plates were cooled to about 55°C before the addition of the antibiotic (ampicillin: 100μg/ml, kanamycin: 25μg/ml, chloramphenicol: 50μg/ml).

2.2.2.5 Preparation of competent cells for transformation

The *E. coli* strain DH5α (genotype: *sup*E44 Δ*lac*U169 (ϕ80 *lac*ZDM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1) was used routinely for amplification of plasmid DNA. To generate cells competent for transformation with cloned DNA, the following protocol was used (Hanahan, 1983).

A single colony from an LB agar plate was incubated in 5ml of LB medium (without antibiotic) at 37°C overnight with shaking (200rpm). The following day, 50μl of this culture were used inoculate a fresh 25mls of LB. Cells were then grown until the OD_{550} reached 0.3-0.4 (typically 3-5hours). Two fresh (pre-warmed to 37°C) 250ml LB cultures were then each inoculated with 12mls of the previous culture and grown for a further 3-4hrs until the OD_{550} had reached 0.4-0.5. The cultures were then cooled rapidly in ice water and the cells harvested by centrifugation at 2500rpm in a Beckman J2-21M/E centrifuge for 10mins at 4°C. The supernatant was discarded and the pellet resuspended in 100mls ice cold TFBI (100mM RbCl, 50mM $MnCl_2$, 10mM $CaCl_2$, 30mM CH_3COOK, 19% (w/v) glycerol, pH5.8 with acetic acid). Cells were again pelleted as before and resuspended in 10mls ice cold TFBII (10mM MOPS, 75mM $CaCl_2$, 10mM RbCl, 19% (w/v)

glycerol, adjusted to pH7 with NaOH). 200-400µl aliquots were then frozen on dry ice and stored at –70°C.

2.2.2.6 Transformation of competent cells

100µl aliquots of competent cells (2.2.2.5) were thawed on ice and incubated on ice with 10µl of ligation reaction (or 1µl of previously cloned plasmid) for 30-60mins. Cells were then heat shocked at 42°C for 2mins and immediately placed on ice for 15mins. 1ml of non-selective LB was added and the cells incubated at 37°C for 1hr. Cells were pelleted by briefly spinning at maximum speed in a microfuge, resuspended in 100µl of fresh LB, plated on the appropriate selective media (2.2.2.4) and grown overnight at 37°C.

2.2.2.7 Production of glycerol stocks

A single colony of transformed bacteria was grown overnight in 10ml LB containing the relevant antibiotic. Glycerol stocks were made by mixing 250µl of 30% glycerol with 250µl of the bacterial culture and freezing on dry ice. Glycerol stocks were stored at –70°C.

2.2.2.8 Preparation of plasmid DNA

High quality plasmid DNA was prepared using the QIA miniprep and maxiprep kits (Qiagen) following the manufacturers instructions. Cloned inserts and plasmid DNA were checked by digestion with the appropriate restriction enzymes and visualisation on an agarose gel (2.2.1).

2.2.2.9 Colony PCR

Where low ligation efficiency was expected, bacterial colonies were screened by PCR for the presence of a cloned insert. Single colonies were picked and suspended in 50µl dH_2O. 10µl of the cells were transferred to a fresh tube and the remaining cells boiled for 5mins, then pelleted by centrifuging at 13000rpm for 1min. 5µl of the supernatant was used as template for a PCR reaction (2.2.3).

2.2.3 Polymerase chain reaction (PCR)

2.2.3.1 Primer design

PCR primers were designed using the Primer Designer program (Scientific and Educational Software), and manufactured by TAGN or Invitrogen. Primers were made up as 50ng/µl working stocks in Tris-HCl (pH8) buffer.

2.2.3.2 Basic PCR

PCR was routinely carried out using Taq DNA polymerase (MBI Fermentas). Standard reactions contained 10-500ng template DNA, 50ng of each primer, 0.2mM dNTP mix (MBI Fermentas), 2.5mM $MgCl_2$, 10X reaction buffer (1X in final volume) and 1U of Taq DNA polymerase (5U/µl stock), made up to final volume of 50µl with dH_2O. A typical cycle of reactions were as follows:

Table 2.2: Typical PCR cycle

Stage 1: 1 cycle	**Stage 2: 35 cycles**	**Stage 3: 1 cycle**
step 1 94°C 1min	step 1 94°C 30secs	step 1 72°C 10mins
	step 2 55°C 30secs	
	step 3 72°C 1min	

2.2.3.3 *Pfu* and *Pwo* PCR

Pfu (MBI Fermentas) and *Pwo* (Roche) high fidelity DNA polymerases were used on different occasions for PCR based cloning of DNA fragments. These enzymes were used according to the manufacturer's instructions.

2.2.4 Isolation of genomic DNA from *Drosophila*

Drosophila genomic DNA was prepared using the DNeasy kit (Qiagen) following the manufacturers instructions (isolation of genomic DNA from animal tissues protocol).

Crude genomic DNA was prepared from single flies where a large number of samples were to be screened by PCR. Individual flies were placed in 50µl of squishing buffer (10mM Tris-HCl, pH8.2, 1mM EDTA, pH8.0, 25mM NaCl, 200µg/ml freshly added proteinase K), and disrupted with a pipette tip before incubating at 37°C for 20mins. Samples were then placed at 85°C for 2mins to inactivate the proteinase K, and centrifuged at 13,000rpm for 3mins. The supernatant was removed and stored at 4°C or –20°C. 1µl of genomic was used as template in a PCR.

2.2.5 DNA sequencing

Automatic sequencing was performed by the University of Warwick Molecular Biology Service, using Applied Biosystems Big Dye Terminator Version 3.1 Chemistry, and run on the 3100 Genetic Analyser.

2.2.6 *In vitro* mutagenesis

For the purposes of some cloning protocols, restriction sites were artificially introduced into cloned DNA fragments by *in vitro* mutagenesis. This was achieved using the QuikChange™ *in vitro* site directed mutagenesis kit (Stratagene) following the manufacturer's instructions.

2.3 RNA manipulations

2.3.1 Maintenance of an RNase free environment

When handling RNA, RNase free water was routinely used to prevent degradation. To produce RNase free water, sterile, deionised water was treated with 0.1% diethyl pyrocarbonate (DEPC), mixed well and left to stand overnight. The water was then autoclaved for 20mins to remove the DEPC.

2.3.2 Whole mount *in situ* hybridisation

All washes were performed in 1.5ml Eppendorf tubes on a rotator, and at room temperature unless otherwise stated.

2.3.2.1 Preparation of Digoxygenin (DIG) labelled RNA probe

Template DNA was prepared (2.2.2.8) and linearised with the appropriate restriction enzyme (2.2.2.1). 1µg of template was made up to 11µl with RNase free dH_20. To this was added, 2µl DIG labelling mix (Roche), 1µl of RNase inhibitor (MBI), 1µl of either T3, T7 or SP6 RNA polymerase (Gibco), and 4µl of the appropriate 5X transcription buffer. The reaction was then incubated for 2hrs at 37°C. 1µl was then saved for later analysis. To the remaining reaction mix 2µl of RNase free DNase was added and this incubated for 15mins at 37°C. Another 1µl

aliquot was removed for later analysis, and the DNase reaction was halted by the addition of 2µl EDTA (200mM, pH8).

Newly synthesised probe was precipitated by the addition of 80µl of dH_2O, 20µl of LiCl (4M) and 360µl of 100% ethanol. The probe was incubated at –20°C for 20mins then pelleted by centrifugation at 13,000rpm. The probe was then resuspended in 100µl of RNase free H_2O.

2.3.2.2 Acid hydrolysis of RNA probes

100µl of $NaHCO_3$ (0.2M) were added to the 100µl of riboprobe prepared as above, and the mixture heated for t mins as calculated by the following formula (Sambrook *et al.*, 1989):

$$t \text{ mins} = \frac{L_0 - L_f}{0.11\,(L_0)\,(L_f)}$$

where L_0 = initial length of probe in kb

Lf = desired length in kb (0.15kb for *Drosophila* embryos)

The probe was then precipitated again by adding 20µl of 5% acetic acid, 20µl of sodium acetate, 500µl of 100% ethanol, incubated at –20°C for 15mins the pelleted by centrifugation at 13,000rpm for 30mins. The pellet was then resuspended in 100µl of RNase free dH_2O. All probes were diluted both 1/100 and 1/50 in Hyb (50% (v/v) deionised formamide, 5X SSC, 0.1% (w/v) Tween-20, 2% (w/v) DIG blocking reagent (Roche)), and stored at –20°C as 100µl aliquots.

2.3.2.3 Collection of embryos

Canton S flies were allowed to lay onto apple agar plates (2.1.1.2), which were changed at 16 and 8hr intervals to ensure embryos were collected at a range of

developmental stages. Embryos were harvested with a fine paintbrush and dechorionated by washing in 50% (v/v) sodium hypochlorite (Fisher Scientific) for 2mins. Embryos were then washed and fixed for 20mins (on a rotator) in a two-phase solution containing 2mls 4% (w/v) paraformaldehyde (in PBS) and 2mls of n-heptane. The lower aqueous phase was then removed and 1ml of methanol added. Embryos were then devitellinised by vortexing at low speed for approximately 30secs. Devitellinised embryos (which sink to the bottom) could then be transferred to an Eppendorf, washed in methanol, and stored at –20°C until needed.

2.3.2.4 Fixation

Embryos were removed from –20°C and washed for 3mins in a 1:1 PBT: methanol mix, then twice for 3mins in PBT. Embryos were then fixed for 20mins in 4% paraformaldehyde (in PBT), washed twice for 5mins in PBT, and then treated for 5mins in proteinase K (5µg/ml in PBT) at room temperature. The proteinase reaction was stopped by rinsing once in PBT then two 5min washes in 2mg/ml glycine (in PBT).

2.3.2.5 Post-fixation

Embryos were fixed again for 20mins in 4% paraformaldehyde (in PBT) followed by two 5min washes in PBT.

2.3.2.6 Pre-hybridisation

Embryos were washed for 15mins in a 1:1 PBT: Hyb mix (2.3.2.2). Following this, embryos were incubated in Hyb solution in a 55°C water bath for

20mins, transferred to a 65°C water bath for 1hour, then back to 55°C for a further 20mins.

2.3.2.7 Hybridisation

The pre-hybridisation solution was removed and replaced with 100µl of riboprobe (diluted in Hyb, see section 2.3.2.2). Embryos were incubated with the riboprobe overnight in a 55°C water bath. The following day, the probe was replaced with fresh Hyb, and the embryos incubated at 60°C for 10mins. Embryos were then subjected to a series of four 10min washes (all at 60°C) in 4:1, 3:2, 2:3, 1:4 Hyb: NTB (NTB: 150mM NaCl, 100mM Tris-HCl (pH7.5), 0.1% (w/v) Tween-20, 0.2% (w/v) DIG blocking reagent), followed by a 10mins, 60°C wash in NTB.

2.3.2.8 Detection

Embryos were washed in 2% goat serum (in NTB) at 4°C for 4hrs on a rotator. This solution was then replaced with fresh NTB 2% goat serum to which 5µl of anti-DIG-AP Fab fragments (Roche) had been added, and left to wash at 4°C overnight. The following day, embryos were washed three times in PBT for 30mins at 4°C, followed by two 10min, 4°C washes in freshly prepared NMTT (0.1M NaCl, 0.1M Tris-HCl (pH9.5), 50mM $MgCl_2$, 0.1% (w/v) Tween-20), and final 10min wash in NMTT at room temperature. Embryos were then transferred to a watch glass and incubated in colouration solution (1ml NMTT, 8µl NBT/BCIP mix (Roche)) in the dark to allow the colouration reaction to proceed (typically 30mins to overnight). When complete, the reaction was stopped by washing in PBT.

2.3.2.9 Mounting

Stained embryos were dehydrated via an ethanol series, cleared with Xylene (Fisher) and mounted in DePeX mounting medium (BDH Laboratory Supplies).

2.3.3 Extraction of total RNA from *Drosophila*

Total RNAs from embryos, larvae, pupae and adult flies were isolated using the QIA RNeasy kit (Qiagen), following the manufacturer's instructions. RNA samples were visualised by agarose gel electrophoresis and quantified using an Ultraspec III spectrophotometer (Pharmacia Biotech).

2.3.4 RT-PCR

2.3.4.1 First strand cDNA synthesis

To prepare first strand cDNA, 100ng-5μg of RNA (equal for each sample) were added to 1μl of 1X hexanucleotide mix (Roche), and made up to a volume of 11μl with RNase free dH_2O. The mixture was then incubated at 65°C for 5mins, and chilled on ice for 2mins. 4μl of 5X M-MuLV reaction buffer (MBI Fermentas), 2μl of 10mM dNTP mix, and 20U RNase inhibitor were then added and made up to a final volume of 19μl with RNase free dH_2O. The sample was then incubated at 25°C for 5mins before the addition of 1μl (200U) M-MuLV reverse transcriptase (MBI Fermentas). The reaction mixture was next incubated at 25°C for 10mins followed by 1hour at 42°C. The reaction was then halted by incubation at 65°C for 20mins. Newly synthesised cDNAs were cooled on ice and stored at –20°C.

2.3.4.2 Amplification of first strand cDNA

A modified PCR protocol was used for optimal detection of specific transcripts. Standard reactions contained 1-2µg cDNA, 50ng of each primer, 0.02mM dNTP mix (MBI Fermentas), 1.25-2.5mM $MgCl_2$ (dependent on properties of primer pair, see below), 10X reaction buffer (1X in final volume) and 2.5U of Taq DNA polymerase (MBI Fermentas), made up to final volume of 50µl with dH_2O. The basic cycle of reactions presented in section 2.2.3.2 was followed, but was modified for each primer pair in order to optimise the reaction (see below).

Table 2.3: Primers for RT-PCR

Transcript	Primer Sequence	Stage2 Cycle No.	Annealing Temp.	$MgCl_2$ conc.
RpL32	5'-aagatgaccatccgcccagcatac-3' 5'-ctcgttcttcttgagacgcagg-3'	30	55°C	1.25mM
prel	5'-gcagaaggatggtagtactc-3' 5'-gagagctccttggttgttga-3'	35	52.3°C	2.5mM
retm	5'-ctggtgatgaaggcctatga-3' 5'-gtagcgacaccgctcgaaga-3'	35	54.4°C	2.5mM

2.3.4.3 Semi-quantitative analysis

"Semi-quantitative" RT-PCR was used to estimate relative levels of transcript expression. To allow for small variations in the initial quantity of template cDNA, *RpL32*, a well characterized housekeeping gene, was used as a control for quantification. Equal volumes (usually 5µl) of each of the samples to be compared were loaded and run on an agarose gel. For each sample, the corresponding *RpL32* PCR product (i.e. amplified from the same cDNA) was also run out on the same gel. All PCR products were quantified using GelDoc software (Bio-Rad). Arbitrary

values were generated for relative fluorescence, which were expressed as a percentage of the corresponding *RpL32* product.

2.4 Protein manipulations

2.4.1 Preparation of *Drosophila* protein

Twenty adult flies were collected and disrupted in 100µl of homogenization buffer (10mM Tris-HCl, pH8, 60mM NaCl, 10mM EDTA, pH8, 0.15mM spermidine, 0.5% (v/v) Triton X-100). The sample was then pelleted by centrifugation at 13,000rpm for 5mins and the supernatant transferred to a fresh tube. This was mixed with 20µl of 2X SDS loading buffer (250mM Tris-HCl, pH6.8, 2% SDS, 10% glycerol, 20mM DTT, 10% β-mercaptoethanol, 0.01% bromophenol blue), before boiling the sample for 5mins and storing at –20°C. Protein samples were re-boiled for 5mins before loading 5-10µl on an SDS-PAGE gel (2.4.4).

2.4.2 Preparation of *S. cerevisiae* protein

A 2.5ml culture of selective YPD medium (2.8.1) was inoculated with a single colony and grown overnight at 30°C. The following day, cells were pelleted by centrifugation at 2500rpm in a Centra MP4R centrifuge (International Equipment Company) for 10mins at room temperature, and the pellet resuspended in 1ml of 0.25M NaOH, 1% β-mercaptoethanol solution. The sample was then incubated on ice for 10mins prior to the addition of 160µl 50% (w/v) Tri-chloroacetic acid (TCA). The sample was mixed, placed back on ice for a further 10mins then centrifuged at 13,000rpm for 10mins. The pellet was resuspended in 1ml of ice cold acetone by vortexing vigorously, then centrifuged again at 13,000rpm for 10mins. The pellet was air dried, resuspended in 300µl of 1X SDS loading buffer, and boiled for 5mins

before storage at –20°C. Protein samples were re-boiled for 5mins before loading 25µl on an SDS-PAGE gel (2.4.4).

2.4.3 SDS-polyacrylamide gel (SDS-PAGE) analysis

SDS-PAGE gels were prepared using the Protein Mini-Gel kit (Bio-Rad) according to the manufacturers instructions. To prepare the resolving gel (12.5%), 4.28ml of 30% acryl-bis stock (29:1 w/w of acrylamide to bis-acrylamide, Fisher Scientific), were mixed with 2.85ml of Solution B (1.57M Tris-HCl, pH8.8, 0.4% (w/v) SDS), 2.87ml dH_2O, 150µl of 10% (w/v) APS (ammonium per-sulphate) and 15µl of TEMED (Sigma). To prepare the stacking gel, 1.3ml of acryl-bis stock were mixed with 2.5ml of solution C (0.5M Tris-HCl, pH6.8, 0.4% (w/v) SDS), 6.3ml dH2O, 150µl 10% APS and 15µl of TEMED.

Electrophoresis was carried out in 1X running buffer (3.03g/l Tris-HCl pH8.3, 14.4g/l glycine and 0.01% (w/v) SDS). Pre-stained size markers (MBI Fermentas) were loaded on all gels for comparison. Protein samples were run on SDS-PAGE at 100V until reaching the edge of the resolving gel, then at 130V for 1-2hrs.

2.4.4 Staining of SDS-PAGE gels

Proteins were visualized by staining with Coomassie blue staining solution (0.1% (w/v) Coomassie brilliant blue, 40% (v/v) methanol, 10% (v/v) acetic acid) for 1hr at room temperature with agitation. Staining solution was then removed and replaced with destaining solution (10% (v/v) methanol, 10% (v/v) acetic acid) for 2-4hrs. The gel was then left in fresh destaining solution overnight (without agitation) to completely remove background.

2.4.5 Western blotting

2.4.5.1 Transfer of protein to nitrocellulose membrane

Following electrophoresis the gel was soaked in western blot transfer buffer (39mM glycine, 48mM Tris base, 0.037% (w/v) SDS, 20% methanol). The gel was placed, with a nitrocellulose membrane, in a gel transfer kit (Bio-Rad) according to the manufacturers instructions, and the transfer carried out for 1hr at 100V.

2.4.5.2 Immunodetection of membrane bound antigens

To reduce non-specific binding of the antibody, the membrane was incubated for 2hrs at room temperature in 5% (w/v) powdered milk (Tesco) in PBS. The membrane was then incubated in the relevant primary antibody (diluted in 2% (w/v) powdered milk in PBS) overnight at 4°C.

Primary antibodies: Rabbit polyclonal anti-GFP 290 at 1: 2000 (AbCam), Rabbit polyclonal anti-LexA (Invitrogen) at 1:5000.

The following day, the membrane was given five 10min washes in PBT before incubating at room temperature in the secondary antibody (goat anti-rabbit IgG AP conjugate, Bio-Rad), diluted 1/2000 in 2% (w/v) powdered milk. The membrane was then washed a further five times for 10mins in PBT.

2.4.5.3 Colour development

The membrane was given two 10min washes in NMTT (100mM Tris-HCl, pH9.5, 100mM NaCl, 5mM $MgCl_2$, 0.1% (w/v) Tween-20), before transferring it to a Petri dish and treating it with colouration solution (80µl NBT/BCIP mix (Roche) in 10ml NMTT). The colour was then allowed to develop in the dark (10mins to

overnight). Colouration reaction was halted by washing twice in PBS for 15mins. The membrane was air dried on filter paper.

2.4.6 *In vitro* transcription - translation

Reactions were performed using the TNT Coupled Reticulocyte Lysate System (Promega). Double stranded plasmid DNA was transcribed and translated as described in manufacturer's instructions, in the presence of ^{35}S methionine. Translation product was then run on an SDS-PAGE gel (2.4.4). The gel was then exposed to X-ray film (Fuji Photo Film Co.) overnight to visualise the protein.

2.5 Whole mount antibody staining

2.5.1 Dissection of *Drosphila* CNS

Embryos were collected from overnight lays, and dechorionated in 50% sodium hypochlorite solution (Fisher Scientific). Early stage 16 embryos (Hartenstein, 1993), were identified by the presence of three disk-like contractions of the gut. Late stage 17 embryos (Hartenstein, 1993) were readily identifiable by the presence of inflated trachea. 1st instar larvae could be readily collected from aged laying plates. Dissections were carried out under either PBS or dissection saline (135mM NaCl, 5mM KCl, 2mM $MgCl_2$, 0.5mM $CaCl_2$, 5mM *N*-Tris[hydroxymethyl]methyl-2-aminoethanesulfonic acid (TES), and 36mM sucrose) using sharp tungsten dissecting needles. Embryos were removed from their vitelline membranes by hand. Devitellinised stage 16 embryos readily adhered to polylysine (Sigma)-coated coverslips. Stage 17 embryos and 1st instar larvae were fixed at their anterior and posterior ends to a Sylgard (DOW Corning)-coated coverslip by mouth pipetting cyanoacrylate glue (Histoacryl; Braun, Melsungen, Germany). Embryos

and larvae were then opened dorsally and stuck flat to the polylysine, or glued to the Sylgard. Gut and fat bodies were removed to fully expose the brain and ventral nerve cord.

Adult brains were dissected in PBT by pinning the fly to a Sylgard coated Petri dish and using fine forceps. The dissection was carried out essentially as described in Allen *et al.*, 1998.

2.5.2 Antibody staining

Dissected embryos were fixed on coverslips with 4% paraformaldehyde in PBS at room temperature for either 1hr (stage 16 embryos) or 30mins (stage 17 embryos). Embryos were then washed several times in PBTx (PBS + 0.1% (w/v) Triton X-100) before blocking for 1 hour in PBTx + 3%BSA (bovine serum albumen, Sigma). Primary antibodies were then incubated overnight at 4°C diluted in PAT (PBTx + 1%BSA). The following day, samples were washed 5 times in PAT, and incubated in secondary antibody (diluted in PAT) for 1hr at room temperature. Embryos were then washed several times in PBTx, and finally mounted in PBTx using nail varnish to seal the slides. Samples were visualised using a Leica confocal microscope, with LCS lite software.

Primary antibodies: Rabbit polyclonal anti-GFP 290 at 1:2000 (AbCam), mouse monoclonal anti-Eve 3C10 at 1:250 (Doe *et al.*, 1988), mouse monoclonal anti-Futsch 22C10 at 1:250 (Hummel *et al.*, 2000), mouse monoclonal anti-Acj6 9C52 at 1:2 (Certel *et al.*, 2000), mouse anti-Repo 8D12 at 1:300 (Developmental Studies Hybridoma Bank), rabbit polyclonal anti-Synaptotagmin at 1:500 (Littleton *et al.*, 1993).

Secondary antibodies: The Molecular Probes fluorescent secondaries Alexa Fluor 488 (goat anti-mouse and goat anti-rabbit), and Alexa Fluor 594 (goat anti-mouse and goat anti-rabbit) were used in this study.

2.6 Cell culture techniques

2.6.1 Maintenance

NIH 3T3 fibroblast cells were seeded in flasks and grown in a monolayer in growth medium (Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS), 2mM glutamine, 20units/ml penicillin and 150units/ml streptomycin). Cultures were kept at 37°C in a humidified atmosphere containing 5% CO_2 in air, and maintained in logarithmic growth by harvesting with trypsin solution (0.25% (v/v) trypsin/ 1mM EDTA in PBS) and re-plating before the cells reached confluence.

2.6.2 Liquid nitrogen stocks

Cells were detached from flasks by incubating in trypsin solution (2.6.1) and transferred to fresh media. This was then pelleted by centrifugation at 1200rpm for 5mins, and the pellet re-suspended in media containing 10% (v/v) DMSO and an additional 10% (v/v) FCS. Aliquots were left on ice for 30mins, and frozen to –80°C before finally being transferred to liquid nitrogen for long-term storage.

2.6.3 Transfection and staining

NIH-3T3 fibroblasts were cultured on coverslips in 23mm dishes in 1ml of growth medium (2.6.1). When the culture reached 50-80% confluence, the cells were transfected with 0.5μg of plasmid under growth conditions, for ~48hours, using

GeneJuice (Novagen) according to the manufacturers instructions. The medium was then replaced with fresh medium containing 25nM MitoTracker Red CMXRos (Molecular Probes, Incorporated) and incubated in growth conditions for 30mins. Cells were then washed with PBS and fixed in 4% paraformaldehyde for 10mins, washed three times in PBS then incubated in DAPI (1μg/ml in PBS) at room temp for 10mins. Cells were again washed in PBS then mounted in Vectorshield mounting medium (Vector Labs, Inc.).

2.6.4 Fluorescence microscopy

Cells were viewed under a Zeiss Axioskop fluorescence microscope equipped with FITC, Texas Red and DAPI filters. Images were captured digitally using the Quips Smartcapture (Vysis, Ltd.) software and processed using Adobe Photoshop.

2.6.5 Confocal microscopy

Cells were viewed as optical sections under a Leica confocal microscope.

2.7 Production of transgenic *Drosophila*

2.7.1 Preparation of DNA for injection

High quality transgene construct DNA was prepared using the QIA Maxiprep kit (Qiagen), and 1μg/μl mixed with 250ng/μl of helper plasmid (pπΔ2-3, Roberts, 1986). This plasmid mix was then ethanol precipitated and resuspended in injection buffer (Sambrook, 1989).

2.7.2 Fly transformation

yw embryos (2.1.2) were collected hourly from apple agar plates (2.1.1.2). Embryos were dechorionated by hand and injected as previously described (Roberts, 1986). Following injection embryos were incubated at 18°C for 16-48hrs. Hatched larvae were collected, transferred to standard fly food (2.1.1.1) and incubated at 25°C until adulthood.

2.7.3 Mapping of inserts

All inserts were mapped genetically to a particular chromosome, and subsequently maintained over an appropriate balancer chromosome (2.1.2). Where inserts were homozygous viable, stocks were maintained without a balancer chromosome.

2.8 Yeast manipulations

Yeast protocols were primarily adapted from the DupLEX-A yeast two-hybrid system (OriGene Technologies) manual.

2.8.1 Yeast media

The yeast strain EGY48 (genotype: **MAT**α *trp1 his3 ura3 leu2*::6 LexAop-*LEU2*) was cultured in YPD (rich medium) or on YPD plates. Per litre of YPD: 20g peptone, 10g yeast extract, 20g glucose, 20g agar (plates only). Transformed yeast strains were cultured and maintained in selective YNB (glu) dropout media or YNB (glu) plates. Per litre YNB(-ura-his-leu-trp): 1.7g yeast nitrogen base without amino acids (QBIOgene), 5g ammonium sulphate, 0.6g-his-leu-trp-ura dropout mix (QBIOgene), 20g glucose, 20g agar (plates only). The appropriate selective media

was produced by adding tryptophan (0.04mg/ml), uracil (0.02mg/ml), leucine (0.06mg/ml) and histidine (0.02mg/ml) as required.

When screening for activation of LexAop-*LEU2* and *lacZ* reporters, YNB (gal) selective media was used. This was prepared in the same way as standard YNB media, but used 20g galactose and 10g raffinose instead of the normal 20g glucose per litre. All media was autoclaved for 15mins before use. Per litre of X-gal indicator medium, 100ml was added of 7% (w/v) sodium phosphate (dibasic) and 3% (w/v) sodium phosphate (monobasic) solution (prepared using sterile water autoclaved), and 0.8ml of 100mg/ml X-gal (in N, N-dimethyl formamide).

2.8.2 Yeast Transformation

A large yeast colony was used to inoculate 10ml of the appropriate media (2.8.1) and incubated overnight at 30°C with shaking (200rpm). The overnight culture was then used to inoculate a 100ml culture to an OD_{600} of 0.2-0.3, and incubated at 30°C with shaking for 3hrs. The cells were then pelleted at 2500rpm in a bench top centrifuge (Centra MP4R, International Equipment Company) for 5mins at room temperature, and resuspended in sterile TE (0.01M Tris-HCl, pH7.5, 1mM EDTA, pH8). Cells were then pelleted again as before and resuspended in 1.5ml of sterile, freshly prepared TE/lithium acetate solution (0.1M lithium acetate in TE). 100µl of yeast suspension was then mixed with 1-5µg of plasmid DNA and 50µg of salmon testes carrier DNA (Sigma) prior to the addition of 600µl sterile PEG/LiAc solution (40%polyethylene glycol (PEG) 3350, 0.1M lithium acetate in TE). This was mixed by vortexing, and incubated at 30°C for 30mins before heat shocking at 42°C for 15mins and then chilling on ice for 2mins. Cells were then spun at 13,000rpm in a bench top microfuge for 5secs and the pellet resuspended in TE. The

cells could then be plated on the relevant selective media and incubated at 30°C for 3-4days to allow colonies to appear.

2.8.3 Large Scale Transformation

For transformation of the library, the transformation protocol described above was scaled up in size. A colony of an EGY48 strain containing the bait and *lac*Z reporter plasmids was used to inoculate 50ml of selective YNB and incubated overnight at 30°C with shaking (200rpm). The following day, 30ml of the overnight culture were used to inoculate 300ml of selective YNB and incubated at 30°C for 3hrs to an OD_{600} of 0.4-0.6. The cells were then centrifuged at 2500rpm in a Centra MP4R bench top centrifuge (International Equipment Company) for 5mins at room temperature, the pellet resuspended in sterile TE, recentrifuged as before and finally resuspended in 1.5ml TE/lithium acetate solution (2.8.2). Thirty 50µl aliquots of yeast suspension were each mixed with 6-10µg of plasmid DNA and 50µg of salmon testes carrier DNA (Sigma) prior to the addition of 300µl PEG/lithium acetate solution (2.8.2) to each tube. Samples were incubated at 30°C for 30mins, heat shocked at 42°C for 15mins, then chilled on ice for 2mins. Cells were next centrifuged at 13,000rpm in a bench top microfuge for 5secs and resuspended in 500µl of TE before plating on the relevant selective medium and incubating at 30°C for 3-4 days. Transformed colonies were harvested from the plates by washing them from the surface in TE with the aid of a sterile glass spreader. The resulting slurry was then pipetted into a fresh tube. Transformants were pelleted by centrifuging at 2500rpm in a bench top centrifuge (Centra MP4R, International Equipment Company) for 5mins, washed in TE, repelleted and then resuspended in an equal

volume of TE. The yeast suspension was then mixed with an equal volume of sterile 50% glycerol and stored at –70°C as 1ml aliquots.

2.9 Electron Microscopy

To expose the CNS, 1st instar larvae were dissected in dissection saline (2.5.1) as flat preparations by gluing to a Sylgard (DOW-Corning)-coated coverslip as described in 2.5.1. Dissected larvae were washed in PBS, fixed in EM fixative (3% glutaraldehyde, 4% formaldehyde, 0.1M PIPES in PBS) and removed from the Sylgard using a razor blade. Sample were processed for electron microscopy according to published methods (Baines *et al.*, 2002) and analysed using a JEOL-1200EX transmission electron microscope.

Chapter 3: Analysis of *slmo* Function in the *Drosophila* Central Nervous System

Chapter 3: Analysis of *slmo* Function in the *Drosophila* Central Nervous System

3.1 Introduction

The locomotor behaviour of *Drosophila* larvae represents a particularly robust and relatively simple motor behaviour, and provides a useful model with which to screen for mutations that perturb motor output. The isolation of mutations in *Drosophila* that cause paralytic phenotypes have led to the identification of many of the components required for neuronal signalling. For example, ion channel genes such as *para* (Loughney *et al.*, 1989; Suzuki *et al.*, 1971) and *cacophony* (Kawasaki *et al.*, 2000), and synaptic vesicle components such as *syntaxin* (Littleton *et al.*, 1998) and *synaptobrevin* (Deitcher *et al.*, 1998). Mutations that disrupt larval locomotion have also led to the identification of a variety of factors required for the development of a functional motor circuit. For example, the role of the POU transcription factor encoded by *acj6* in regulating neuronal connectivity and synapse formation (Certel *et al.*, 2000), and *hikaru genki*, which has been shown to encode a cell adhesion protein (Hoshino *et al.*, 1999; Hoshino *et al.*, 1996). Screens in larvae have also identified genes involved in related behaviours such as foraging and wandering (Shaver *et al.*, 2000). However, there is a paucity of genetic loci identified as being involved in the early stages of locomotor peristalsis, and very little is known of how the circuitry that underlies this behaviour develops or is regulated.

The availability of P(GAL4) enhancer trap strains expressed in specific subsets of neurons (Brand and Perrimon, 1993) has provided a means to begin unravelling the function of neurons, as behaviour can be analysed in parallel with genetic misexpression (Ferveur *et al.*, 1995; Hidalgo and Brand, 1997; Sweeney *et*

al., 1995). Suster *et al.* (2003) applied this genetic approach in an attempt to identify specific neurons required for the spatial and temporal pattern of larval locomotion. A small screen was performed of about 150 P(GAL4) enhancer traps by looking for abnormal larval behaviour after crossing each of these to a UAS-TeTxLC strain. A small number of GAL4 lines were isolated which generated defects specifically in the "decision making" phase of locomotory behaviour, such as abnormal turning and pausing. The ideal would be to characterise genes that correspond to interesting enhancer trap lines and identify molecules involved in regulating locomotor behaviour.

Previous research on the *slmo* locus had focused on the involvement of this gene in survival of the germline, but evidence had been reported that the gene is expressed in the nervous system, and that loss of *slmo* function results in defective motor control (Carhan, 1999; Reeve, 2002). The research presented in this chapter focuses on the role of *slmo* within the nervous system. The *slmo* mutant phenotype, and the corresponding enhancer trap line P(GAL4)c682, are consistent with this locus as a candidate for regulating the motor output required for normal locomotory peristalsis. Possible causes of the phenotype are explored and discussed.

3.2 Results

3.2.1 *slmo* mutant lethality and behavioural phenotype

Initial observations showed that while null mutant embryos were capable of hatching they were clearly compromised in their motor function. Whilst it was originally proposed that mutant 1st instars are completely paralysed (Carhan, 1999), closer inspection revealed that newly hatched larvae have some motor function (Reeve, 2002). Mutant larvae were tested for the ability to perform a simple "twist-

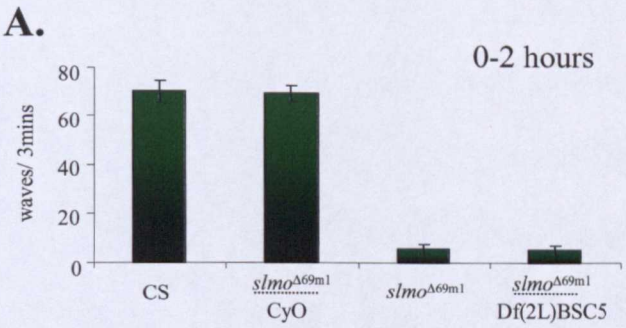
and-roll" righting reflex when placed in an inverted position (Reeve, 2002). This represents a relatively complex behavioural response requiring coordinated motor control, and had previously been used to show the requirement for the gene *turtle* (which encodes a neurally expressed member of the Ig superfamily) in executing complex coordinated movements (Bodily *et al.*, 2001). As with *turtle* null mutant larvae, it was found that newly hatched $slmo^{\Delta 69m1}$ homozygotes were able to perform the reflex, and right themselves from an inverted position, but were clearly defective in this behaviour, taking a substantially longer time to perform the reflex than wild-type animals. By the end of the 1st instar stage (~20hrs), mutants were totally unable to perform this behaviour and remained stranded in an inverted position.

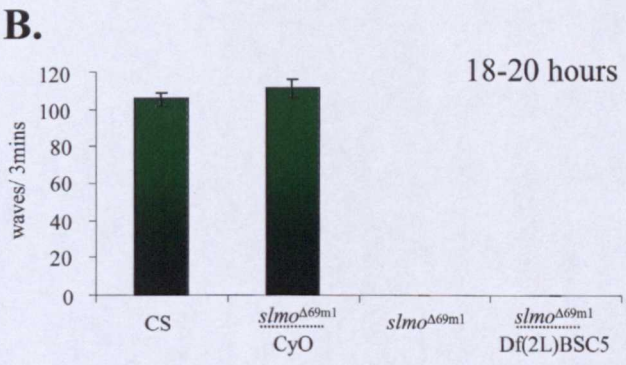
In the case of the *turtle* gene, mutant 3rd instar larvae were found to be defective in the righting reflex behaviour, and other coordinated movements such as response to tactile stimulation, turning behaviour during locomotion, and flying in adults, whilst being able to execute basal locomotory peristalsis in a manner comparable to wild-type (Bodily *et al.*, 2001). In this case the authors argued that the *turtle* gene fulfils a specific role in regulating complex coordinated behaviours, and that the inversion assay provides an example of such behaviour. In the case of *slmo*, this behavioural assay demonstrated that these mutants have defective motor function, and that this defect is progressive. However, a more thorough consideration of *slmo* mutant behaviour might suggest that the defect reported for this complex behaviour, was in this case secondary to a more general disruption of motor function.

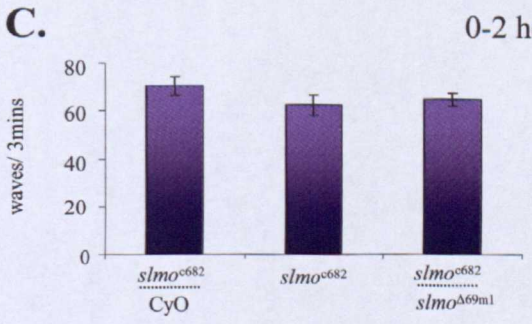
Observation suggested that the *slmo* null mutation, which causes lethality at an earlier stage than *turtle*, results in a much more severe impairment of motor control, with animals exhibiting severely sluggish locomotor behaviour immediately after hatching. Newly hatched larvae performed visibly slow and weak waves of

peristaltic contraction that drives larval locomotion, and incomplete contraction waves were often performed. It therefore follows that mutant larvae would perform badly in the righting reflex assay performed by Reeve (2002), as this movement requires coordinated bilateral motor control involving the innervation of body wall muscles in a specific pattern (Bodily *et al.*, 2001). Interestingly, *slmo* null mutants were still capable of moving their head and mouth parts during pauses. Given that *slmo* expression in the embryo had been reported previously to be specific to the CNS, it seemed reasonable to postulate that the locus might be involved in regulating the control of locomotor behaviour.

A detailed analysis of the locomotor defect is presented in Figure 3.1 (2.1.3). Mutant and heterozygous larvae were genotyped by selection for, or against the CyO [*kruppel*GAL4, UASGFP] balancer chromosome, which causes bright fluorescence of the Bolvic's organs (Casso *et al.*, 2000). At 0-2hrs post-hatching, $slmo^{\Delta 69m1}$ homozygotes exhibit an approximately 12-fold reduction in the number of peristaltic contraction waves executed in a 3min time period when compared to the wild type control line Canton S (CS) or the heterozygous siblings (Figure 3.1A). The same phenotype is also observed with larvae carrying the $slmo^{\Delta 69m1}$ mutation hemizygous with a deficiency for the chromosome region (Df(2L)BSC5). As expected, the phenotype becomes more severe with age, and by 18-20hrs post-hatching larvae homozygous or hemizygous for the $slmo^{\Delta 69m1}$ allele were unable to perform a full body peristaltic contraction wave (Figure 3.1B). However, aged mutant larvae were still able to move their head and mouth parts in response to tactile stimulation. In addition, raising the larvae on coloured food demonstrated that null mutants do not feed and grow (Figure 3.1E), and by 18-20hrs post-hatching have developed a flaccid







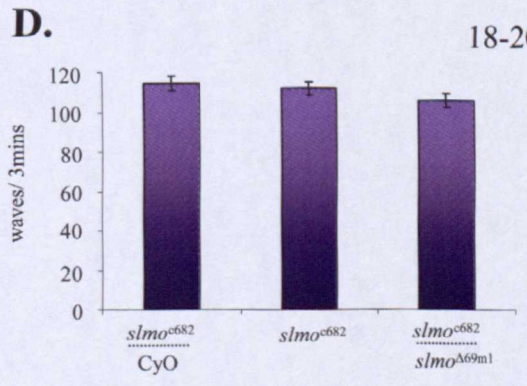
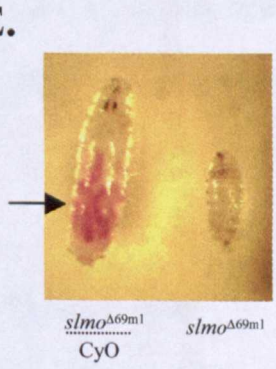


Figure 3.1: ***slmo* mutants exhibit defects in locomotor behaviour**. **A-D**. Locomotor activity was measured as the number of full body peristaltic contraction waves performed by L1 larvae in a 3min time period. For all lines in **A-D** n = 15. Error bars indicate standard error. **A.** At 0-2 hrs post-hatching (PH), $slmo^{\Delta 69m1}$ null mutants, and $slmo^{\Delta 69m1}$/Df(2L)BSC5 hemizygotes perform badly in the locomotor assay as compared to wild-type (Canton-S, CS) and $slmo^{\Delta 69m1}$/CyO controls. **B.** By 18-20 hrs PH, no peristaltic contraction waves are performed by $slmo^{\Delta 69m1}$ homozygotes or hemizygotes. **C-D.** Homozygotes for the hypermorphic allele $slmo^{c682}$ and $slmo^{c682}/slmo^{\Delta 69m1}$ transheterozygotes perform locomotor peristalsis at a rate comparable to wild-type at both 0-2hrs (**C.** P-value=0.305) and 18-20hrs (**D.** P-value=0.184) PH. **E.** $slmo^{\Delta 69m1}$ and $slmo^{\Delta 69m1}$/CyO larvae were fed coloured food and aged for 20 hrs. Control heterozygous siblings grow to more than double their original size, and the red-dyed food is clearly visible in the gut (**arrow**). In contrast, mutant larvae do not grow, and do not appear to have ingested the coloured food.



appearance. This raises the possibility that the lethality may ultimately be due to lack of food or hydration, rather than as a direct result of a defect in the nervous system.

To examine whether the phenotype is dosage dependent, the original *slmo*^{c682} allele was also examined. Northern and western blot analysis had previously shown the *slmo*^{c682} mutation, which carries the P(GAL4)c682 enhancer trap, to be a hypermorphic allele of *slmo* (Carhan, 1999; Reeve, 2002). Expression of the gene is greatly reduced, but homozygotes survive to adulthood with no apparent behavioural phenotypes. Over time, homozygous adults had become increasingly rare within the stock (which was maintained over a CyO balancer), suggesting that the line had acquired a lethal mutation on the *slmo*^{c682} chromosome. Before any behavioural analysis could be performed, single flies of the *slmo*^{c682} line were out-crossed to WCS flies over 7 generations to allow any linked mutations to be lost by recombination (the P-element insertion could be followed by its *white*+ marker), and then balanced over the CyO [*kruppel*GAL4, UASGFP] chromosome. As expected, the genetically "Cantonised" *slmo*^{c682} stock returned to a 2:1 ratio of curly wing to straight wing flies.

Locomotor activity was analysed as before for *slmo*^{c682} homozygotes, *slmo*^{c682}/*slmo*$^{\Delta 69m1}$ transheterozygotes and *slmo*^{c682}/CyO heterozygous siblings. No significant differences (2.1.3) were observed between control flies or either of the mutant genotypes (Figure 3.1C, D), suggesting that the *slmo* phenotype is not dosage dependent.

3.2.2 Expression analysis of the P(GAL4)c682 enhancer trap

The enhancer trap P(GAL4)c682 had previously been subjected to very limited analysis in the developing embryo. Using UAS-*lac*Z as a reporter, expression

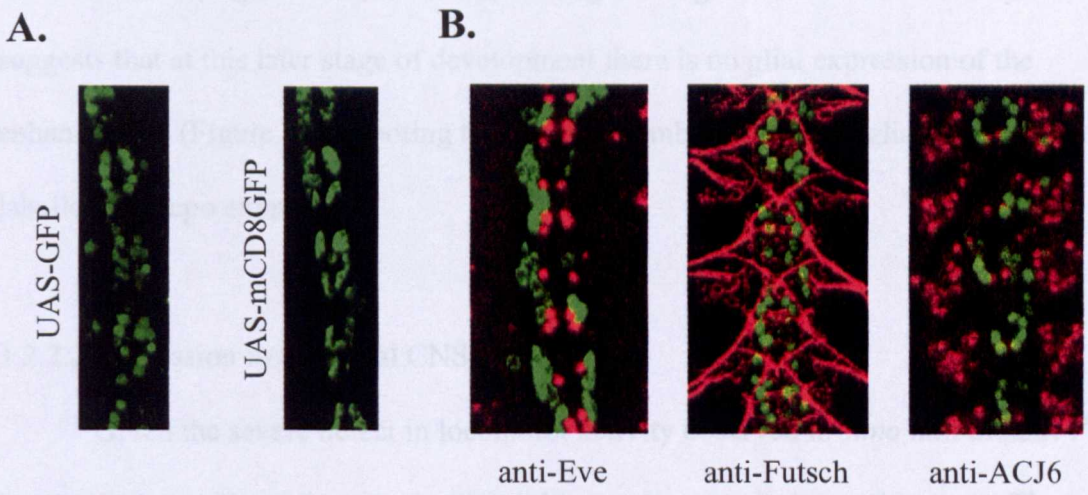
of the enhancer trap was reported in the ventral nerve cord (VNC) of stage 16 embryos (Carhan, 1999). No expression was observed outside the CNS, except in the salivary glands, which was dismissed as a common artefact of enhancer traps. Interestingly, reporter expression appeared to be restricted to a very limited group of cell bodies in the midline of the VNC. Following cloning of the gene, *in situ* hybridisation to *Drosophila* embryos using a Digoxygenin labelled DNA probe appeared to show the enhancer trap expression to be an accurate representation of the expression of the endogenous gene. The *slmo* gene is expressed only in the CNS and the salivary glands, and is not ubiquitous within the CNS. It was proposed that the cells expressing P(GAL4)c682 and *slmo* mRNA are identified as the longitudinal glia, based purely on morphology (Carhan, 1999). However, the glial identity of these cells was never proven, and no attempt was made to explore the expression of the enhancer trap at later stages of development, which might be more directly relevant to the mutant phenotype (which is first apparent after hatching). Attempts to use several antisera against Slmo have failed to work on whole mount preparations, which would accurately report the expression of the protein, although western blot analysis confirmed that Slmo is expressed at low levels in embryos and larvae (Carhan, 1999; Reeve, 2002). I attempted to address these issues by utilising the fluorescent reporters UAS-mCD8GFP (Lee and Luo, 1999), UAS-GFP (Yeh *et al.*, 1995) and UAS-GFPn (nuclear GFP; Martin-Blanco *et al.*, 2000) of enhancer trap expression, allowing for greater resolution of the expression pattern within the CNS.

3.2.2.1 Expression in the embryonic CNS

Expression of the enhancer trap was initially determined using the reporter UAS-mCD8GFP (Lee and Luo, 1999), which expresses a version of the green

fluorescent protein that becomes anchored to the cell membrane. This provides a brightly fluorescent reporter with the important advantage that neuronal projections can be visualised. In addition, UAS-GFP (Yeh *et al.*, 1995; GFP remains in the cytosol and nucleus) was also employed, but detected and visualised using immunofluorescence (2.5.2), allowing the cell bodies to be resolved more clearly. Importantly, the pattern of expression is similar with both reporters, suggesting that all the cells expressing GAL4 at significant levels are labelled. The embryonic CNS was viewed in filleted preparations using confocal microscopy (2.5.1).

As previously reported, the enhancer trap is not expressed until early in stage 16 of embryonic development, at around the time at which muscle contractions first become apparent (Broadie and Bate, 1993). Expression at this stage is found in a small group of 8 cell bodies per segment (Figure 3.2). To establish the identity of these cell bodies, CNS preparations were fixed and double labelled with the glial marker anti-Repo 8D12 (Developmental Studies Hybridoma Bank), and the neuronal markers anti-Evenskipped (Eve) 3C10 (Doe *et al.*, 1988), anti-Futsch 22C10 (Hummel *et al.*, 2000) and anti-Acj6 9C52 (Certel *et al.*, 2000) (Figure 3.2). At stage 16, Eve is expressed in the identified motoneurons aCC and RP2, plus the pCC interneurons (Doe, 1992). Futsch expression also labels motoneurons, and some interneurons at this stage including aCC, MP1, dMP2, vMP2, SP1, VUMs and their neuronal extensions (Grenningloh *et al.*, 1991), whilst Acj6 labels only interneurons thought to be involved in motor control (Certel *et al.*, 2000). The cell bodies marked by P(GAL4)c682 at stage 16 clearly co-localise with a subset of the midline glial cells labelled by anti-Repo, whilst not co-localising with any of the neuronal markers, confirming previous observations made by Carhan (1999).



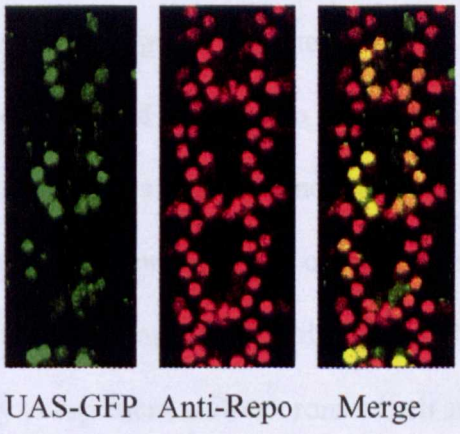


Figure 3.2: Expression of P(GAL4)c682 in stage 16 embryos. A. Expression in the ventral nerve cord as reported by both UAS-GFP and UAS-mCD8GFP. The enhancer trap is expressed in a small group of about 8 cell bodies per segment in the ventral midline. **B.** Enhancer trap expression reported by UAS-GFP (**green**), does not co-localise with neuronal markers anti-Eve, anti-Futsch, or anti-ACJ6 (**red**). **C.** The cell bodies marked by P(GAL4)c682 (**green**) co-localise with a subset of midline glia labelled by anti-Repo (**red**). All images were captured at 63X magnification and are orientated with anterior as upwards.

By stage 17, reporter expression has expanded and can be seen in a group of medially positioned neurons as well as a second set of neurons more peripheral within the CNS. The enhancer trap is also expressed in a limited subset of neurons in the brain lobes (Figure 3.3A). Double labelling the stage 17 CNS with anti-Repo suggests that at this later stage of development there is no glial expression of the enhancer trap (Figure 3.3B), noting that a small number of midline glial cells are not labelled by Repo expression.

3.2.2.2 Expression in the larval CNS

Given the severe defect in locomotor activity observed in *slmo* null mutant 1st instars, it seemed important to examine enhancer trap expression at this stage. The expression of P(GAL4)c682 was visualised in filleted preparations using UAS-mCD8GFP as the reporter and imaging with confocal microscopy. The reporter is expressed strongly in a limited group of midline cell bodies located in the dorsal half of the VNC, and weakly in a small group of lateral neurons most of which are located in ventral and intermediate positions (Figure 3.4). The reporter is expressed strongly in the neuropil and commissures, and neurons are seen to project along the segmental nerves. The dorsal position of the midline cells suggests that these are likely to represent motoneurons which are sending axons out to the muscles, whilst the lateral cell bodies may belong to interneurons projecting into the neuropil (Certel *et al.*, 2000; Odden *et al.*, 2002). In the brain lobes, the enhancer trap labels a restricted group of cell bodies in the cortex, which can be seen to project into the dorsal neuropil region (Younossi-Hartenstein *et al.*, 2003). In addition, strong expression is seen in the basoanterior compartment of the neuropil, which forms part

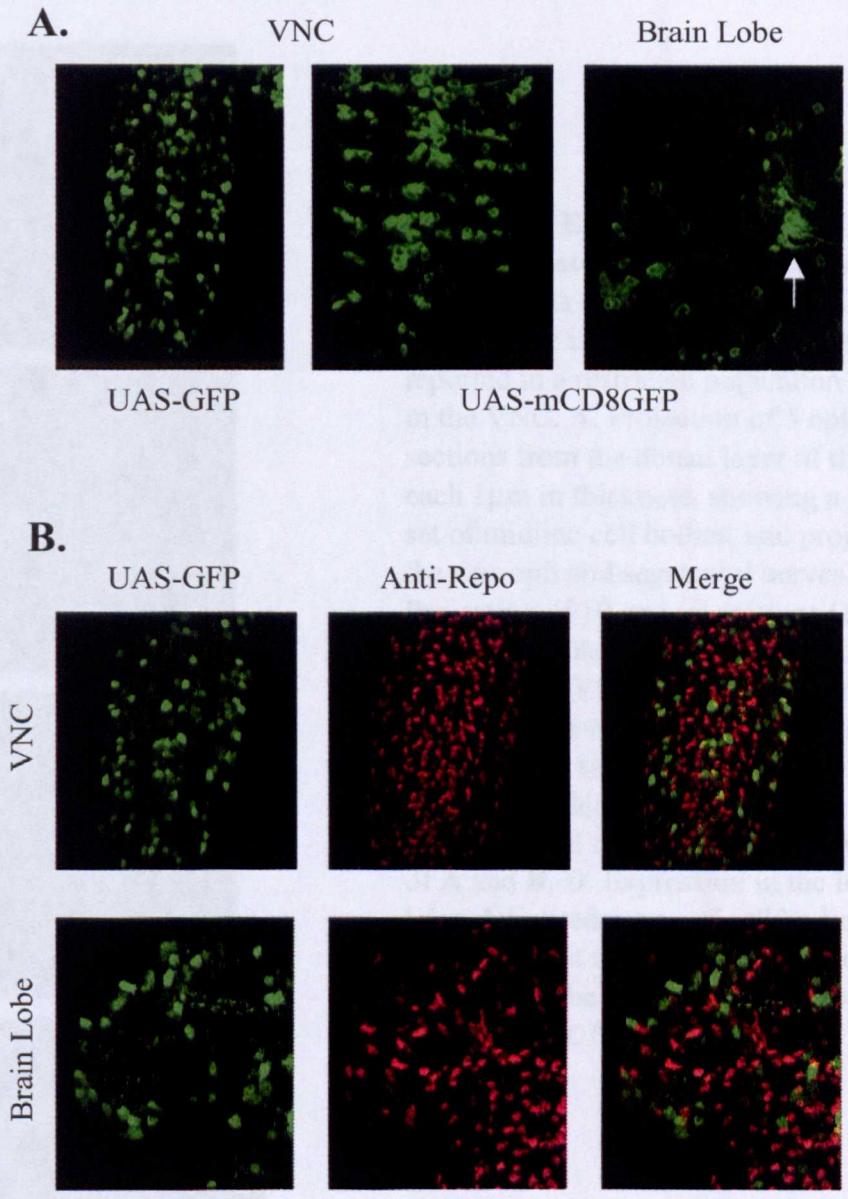


Figure 3.3: **Expression of the P(GAL4)c682 in stage 17 embryos**. **A**. Expression in the VNC and left brain lobe as reported by UAS-GFP and UAS-mCD8GFP. The enhancer trap is expressed in a population of neurons positioned medially as well as a second set of neurons more peripheral within the VNC. Expression is also reported in a subset of neurons in the cortex of the brain lobe, including a small group of strongly labelled cells projecting towards the VNC (**white arrow**). **B**. Enhancer trap expression reported by UAS-GFP (**green**) does not co-localise with the glial marker anti-Repo (**red**) at stage 17, indicating that all expression is neuronal at this stage. All images were captured at 63X magnification and are orientated with anterior as upwards.

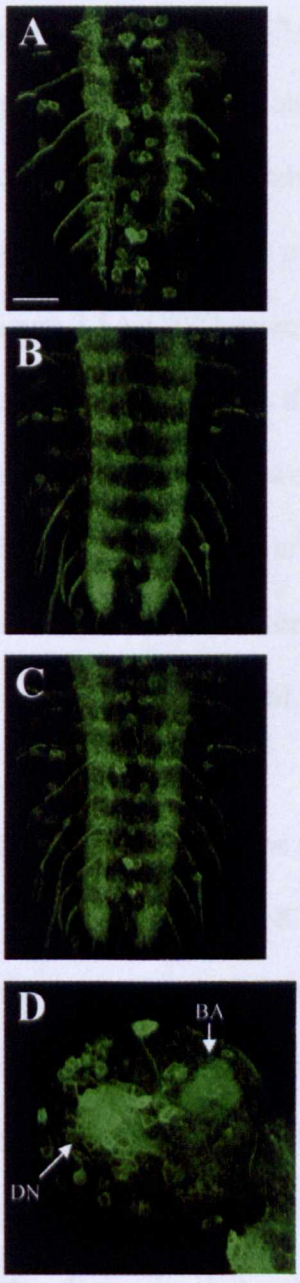


Figure 3.4: Expression of P(GAL4)c682 in the 1st instar larva. In all cases expression is reported by UAS-mCD8GFP, and anterior is up. **A-C**. Expression is reported in a restricted population of cells in the VNC. **A**. Projection of 5 optical sections from the dorsal layer of the VNC, each 1μm in thickness, showing a restricted set of midline cell bodies, and projections in the neuropil and segmental nerves. **B**. Projection of 10 optical sections (1μm thickness) from the intermediate and ventral layers of the VNC. Strong expression can be seen in the neuropil, commissural neurons, and segmental nerves. A group of weakly labelled cell bodies can be seen in the peripheral region of the CNS. **C**. Merge of **A** and **B**. **D**. Expression in the left brain lobe. A limited group of cell bodies in the cortex project into the dorsal neuropil (**DN**) region, and the basoanterior neuropil (**BA**) compartment. Scale bar is 25μm.

of a transitional region between the brain and VNC (Younossi-Hartenstein *et al.*, 2003).

The CNS was also dissected from 3rd instar larvae carrying the enhancer trap and the UAS-mCD8GFP reporter (Figure 3.5). At this stage of development a similar expression pattern was observed to that seen in 1st instars. A group of dorsally located cell bodies are labelled in the midline of the VNC, and very strong expression can again be seen in the neuropil, commissures and segmental nerves (Figure 3.5A-C). The neurons marked by the enhancer trap at this stage are presumably the same as those seen at the first instar stage, as the CNS undergoes little reorganisation between these stages (Younossi-Hartenstein *et al.*, 2003). In the brain lobes, expression is tightly restricted to a group of neurons, which again appear to correspond to those seen in the much smaller brain lobes of the first instar larva by projecting into the dorsal neuropil region (Figure 3.5D,E).

3.2.2.3 Expression in the adult CNS

The P(GAL4)c682 line was originally isolated for its expression in the adult giant fibres. To explore expression in the adult CNS, whole brains were dissected from flies carrying the enhancer trap and the UAS-mCD8GFP reporter (2.5.1). These were mounted in PBS and analysed by confocal microscopy (Figure 3.6A,B). Expression in the adult brain, although not ubiquitous, is seen to be widespread within the neuropil. An alternative reporter, UAS-GFPn, expresses GFP fused to a nuclear localisation signal (Martin-Blanco *et al.*, 2000), and was used to specifically label GAL4 expressing cell bodies. Expression is reported within a large number of cell bodies in the cortex of the adult brain.

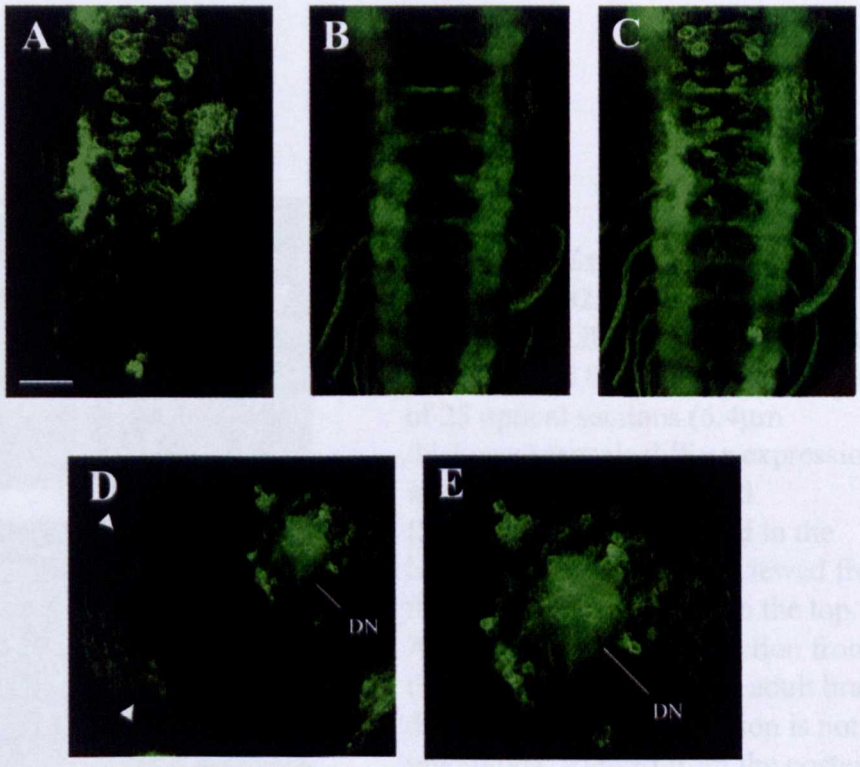


Figure 3.5: **Expression of P(GAL4)c682 in the 3rd instar larva**. In all cases expression is reported by UAS-mCD8GFP and anterior is up. **A**. Projection of 3 optical sections (1µm thickness) from the dorsal layer of the VNC. A small set of dorsally located cell bodies can be seen, which may correspond to those seen in the 1st instar VNC (see Figure 3.4A). **B**. A projection of 16 optical sections (1µm thickness) from the ventral and intermediate layers of the VNC shows expression in neuropil, commissural neurons, and segmental nerves. **C**. Merge of **A** and **B**. **D**. Expression in the left brain lobe may also correspond to that seen at the 1st instar stage (see Figure 3.4D). A small group of neurons can be seen to project into the dorsal neuropil region (**DN**). The perimeter of the brain lobe is indicated by **white arrowheads**. **E**. Close up of **D**. Scale bar is **A-D**: 50µm, **E**: 25µm.

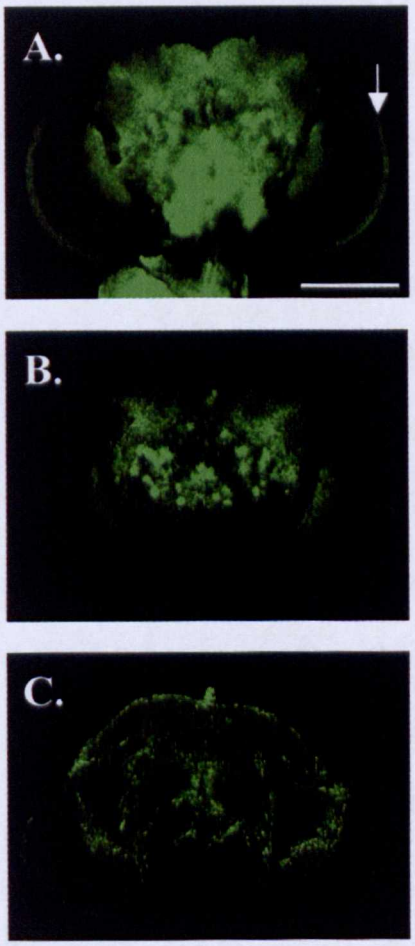


Figure 3.6: Expression of P(GAL4)c682 in the adult brain. A. UAS-mCD8GFP reports widespread expression in the adult brain. A merge of 25 optical sections (6.4µm thickness) reveals diffuse expression in the central neuropil, and fluorescence also observed in the lamina (**arrow**). Brain is viewed from the posterior with dorsal to the top. **B.** A single 6.4µm optical section from the posterior region of the adult brain demonstrates that expression is not ubiquitous. Viewed from the posterior with dorsal to the top. **C.** Merge of 25 optical sections (1.2µm thickness). UAS-GFPn reports expression of the enhancer trap in a large number of cell bodies within the cortex. Brain is viewed from the anterior with dorsal to the top. Scale bar is 75µm.

3.2.2 Inactivation of marked neurons with UAS-TeTxLC

To further characterise the subset of neurons marked by P(GAL4)c682, the neurons were inactivated using UAS-TeTxLC, which is sufficient to eliminate all evoked release of synaptic neurotransmitter (Sweeney *et al.*, 1995). Resultant embryos appeared to develop normally, but ultimately were unable to hatch. Closer inspection revealed that, like *slmo* mutant larvae, embryos expressing the toxin were not completely paralysed. The vigorous head swinging movements characteristic of the hatching process were still commonly observed, whilst peristaltic contraction waves were seen only very rarely. The phenotype generated by expression of TeTxLC in this population of neurons therefore appears to reflect that of *slmo* mutants, but with increased severity. Locomotor peristalsis is seemingly affected more severely than other movements such as head swinging.

To further investigate the neuronal basis of this phenotype, an electrophysiological analysis was performed in collaboration with Richard Baines (University of Warwick). In order to compare the activity of central synapses, a previously described preparation of the *Drosophila* embryo was used to record the synaptic drive in motoneurons (Baines and Bate, 1998; Baines *et al.*, 1999). The patterned synaptic excitation exhibited by this preparation is known to reflect the maximum rate of peristaltic contraction in an intact wandering larva (Baines and Bate, 1998). This is presumably because the damage necessarily done to the CNS in order to access and record from the motoneurons initiates a near maximal output of the underlying CPG circuitry. In a late stage 17 wild type embryo or larva, the centrally generated rhythmic pattern of motor output can clearly be seen as bursts of action potentials. Analysis of the synaptic drive to motoneurons in embryos expressing the toxin reveals the normal pattern of synaptic excitation to be severely

disrupted (Figure 3.7). Synaptic excitation is almost completely abolished, however, it is clear that the motoneurons do have some input, and some residual rhythmicity can still be observed. Inactivation of the neurons marked by the enhancer trap does appear to prevent the CPG from firing normally. The sporadic and weakened excitation that persists in the motoneurons indicates that the immediate presynaptic interneurons have retained the ability to release neurotransmitter, and therefore do not express GAL4. However, these interneurons have seemingly lost their own synaptic drive, suggesting that the enhancer trap labels neurons presynaptic to this subset of neurons.

3.2.4 Analysis of sectioned mutant CNS

Neurodegeneration is a pathological condition known to occur in diverse organisms, including humans and flies. Neurodegenerative phenotypes have often been ascribed to single gene defects (Muqit and Feany, 2002), and in model organisms such as *Drosophila*, a variety of mutations have been isolated that cause degeneration of the nervous system and manifest as behavioural phenotypes (Driscoll and Gerstbrein, 2003). In many cases, neuronal cell death is the consequence of mutations that cause aberrant neuronal signalling properties (Driscoll and Gerstbrein, 2003). For example, the mouse neurodegenerative mutants *Lurcher*, *tottering* and *weaver*, which were initially isolated on the basis of defective locomotor behaviour phenotypes, have now been shown to encode ion channels and neurotransmitter receptors (Fletcher *et al.*, 1996; Patil *et al.*, 1995; Zuo *et al.*, 1997). Likewise, mutation in *C. elegans* ion channel genes such as *deg-3*, *deg-1* and *mec-4*, have also been shown to cause neurodegeneration (Hall *et al.*, 1997; Yassin *et al.*, 2001). In *Drosophila*, mutations in *Adar*, which encodes an RNA editing enzyme that operates

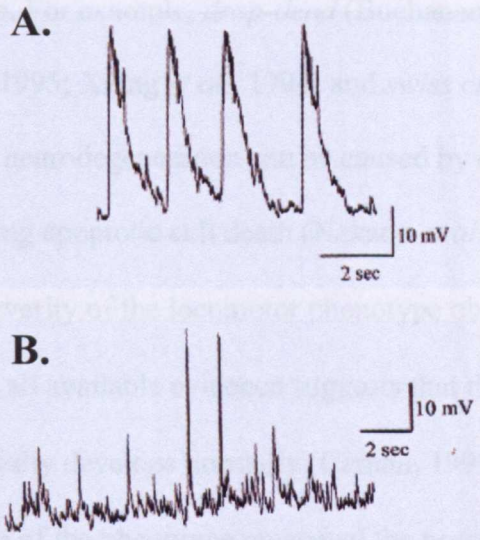


Figure 3.7: Synaptic excitation of motoneurons in TeTxLC expressing embryos. A. In a wild-type embryo, whole cell current clamp recordings from either aCC or RP2 motoneurons reveal the presence of rhythmic synaptic depolarisations that result from the activity of presynaptic cholinergic interneurons. **B.** Expression of UAS-TeTxLC, using the driver P(GAL4)c682, causes disruption of patterned synaptic excitation in the motoneurons aCC/RP2. The residual activity that can be recorded in these embryos suggests that the immediately presynaptic interneurons have maintained the ability to release neurotransmitter.

on transcripts encoding several ion channels (Palladino *et al.*, 2000), and *vacuous*, a mutation that results in hyperexcitability of young neurons (Palladino *et al.*, 2002), both result in extensive loss of neurons in the adult brain. In several cases, degeneration of neurons appears to result from defects in the glial cells required for neuronal maintenance. For example, *drop-dead* (Buchanan and Benzer, 1993), *repo* (Xiong and Montell, 1995; Xiong *et al.*, 1994) and *swiss cheese* (Kretzschmar *et al.*, 1997). Alternatively, neurodegeneration can be caused by direct interference with the mechanisms controlling apoptotic cell death (Nakano *et al.*, 2001).

Despite the severity of the locomotor phenotype observed in $slmo^{\Delta 69m1}$ mutants (Figure 3.1), all available evidence suggests that the nervous system of the $slmo^{\Delta 69m1}$ mutant initially develops normally (Carhan, 1999; Reeve, 2002). However, the progressive nature of the phenotype presented the possibility that neuronal circuitry was degenerating at a later stage. Certainly, the marked reduction of *slmo* expression caused by the original P(GAL4)c682 enhancer trap insertion had been shown to cause dramatic degeneration of the male germ line in a manner suggestive of apoptosis (Carhan, 1999; also see Figure 1.3). To address this possibility, *slmo* mutant larvae were sectioned and the CNS examined under the transmission electron microscope (TEM). This approach allowed the presence or absence of cellular degeneration within the CNS to be assessed directly, and also provides an opportunity to look for any obvious abnormalities in the structure of the mitochondria, to which the function of *slmo* has now been linked (see chapter 4).

In preparation for the TEM, $slmo^{\Delta 69m1}$ homozygous larvae were selected immediately after hatching and aged for 18-20hrs. Aged larvae were then dissected to expose the CNS, fixed, stained, and embedded in Araldite resin (2.9). Embryos were sectioned by Anton Page (University of Southampton). Ultrathin sections were

taken from the posterior region of the larval brain and viewed under the TEM (Figure 3.8). Heterozygous siblings were prepared in an identical manner as a negative control. Examination of several sections from multiple samples revealed no evidence of cellular degeneration. No vacuolisations can be seen in either the neuropil or cortex of mutant larvae. In addition, no obvious differences were observed in cell morphology, and the mitochondria also appeared indistinguishable from those of the heterozygous controls. Small vacuolisations could sometimes be identified in the cytoplasm of both neurons and glia, but these were also present in the control samples. The CNS of aged *slmo* null mutant larvae therefore appears to be intact, suggesting that cellular degeneration is not the cause of the motor phenotype.

3.2.5 Examination of axon transport in mutant larvae

The considerable distances that can separate a synapse from its cell body places special demands on the mechanisms of cytoplasmic motility. The axon itself supports little synthesis of proteins, making it necessary for newly synthesised components to be continually transported from the cytoplasm of the cell body to the presynaptic point of action (Martin *et al.*, 1999). In addition, this anterograde transport must be supported by a system of retrograde transport, which returns materials back to the cell body for recycling (Schnapp, 1997). Axonal transport is critical for the ability of a neuron to transmit signals to postsynaptic cells, and is therefore essential for both the development and maintenance of a functional neuronal circuit.

Intracellular transport requires the action of molecular motors that bind cargo and generate movement coupled to ATP hydrolysis along cytoskeletal filaments. Kinesin-I is a microtubule motor protein that facilitates the ATP dependent

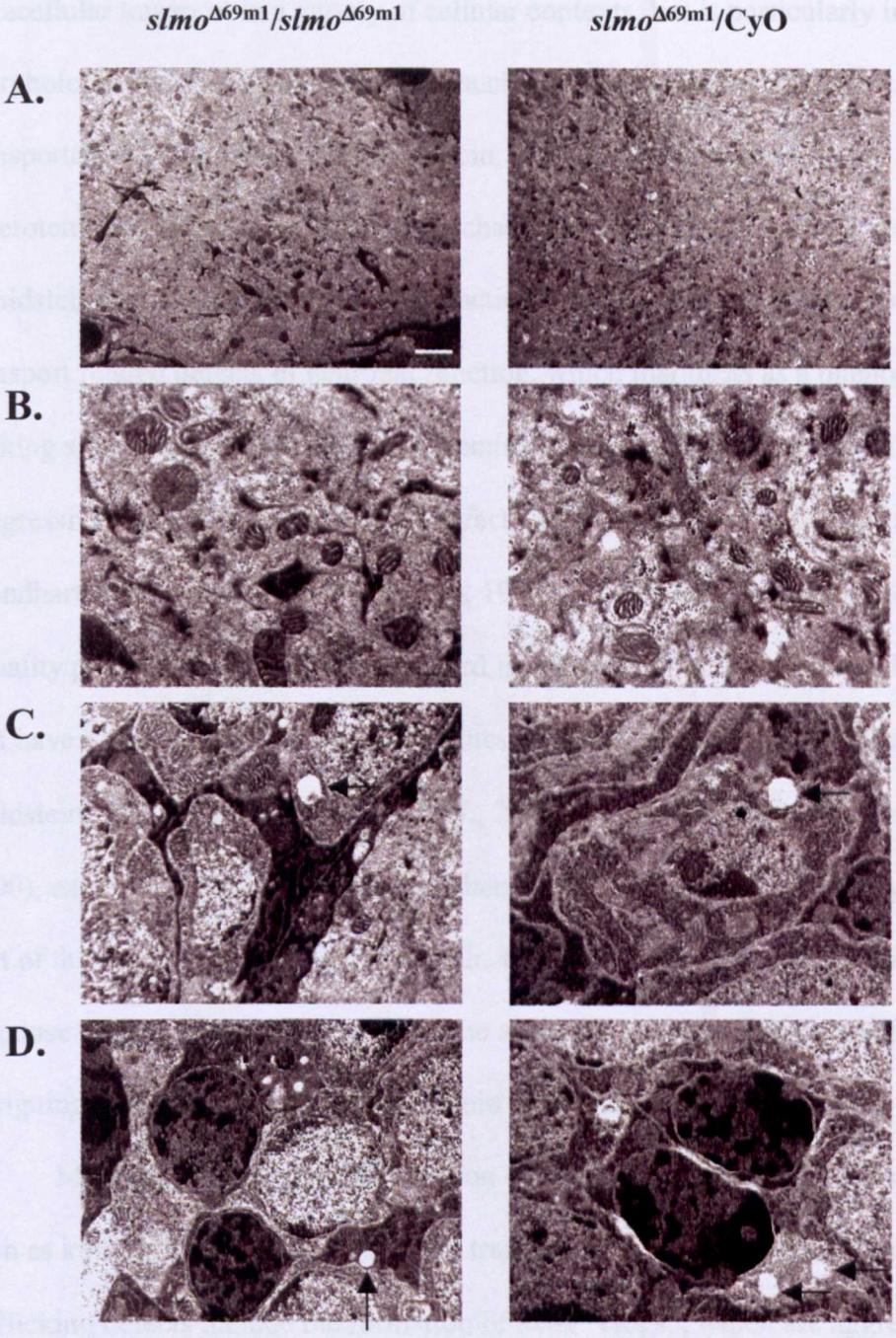


Figure 3.8: **Analysis of sectioned slmo$^{\Delta 69m1}$ mutant CNS**. Aged (18-20hrs) *slmo*$^{\Delta 69m1}$ homozygotes (**left column**) were compared to their heterozygous siblings (**right column**). Sections were taken from the posterior region of the larval brain lobes. **A**. Low magnification view of the neuropil showing a complete lack of cellular degeneration in the mutant, which is indistinguishable from the heterozygote controls. **B**. Mitochondria in the neuropil exhibit no obvious morphological defects. **C, D**. Small vacuolisations (**arrows**) can sometimes be identified in the cytoplasm of neuronal cell bodies (**C**) and glia (**D**), but these are also present in the control larvae. Scale bar is **A**: 2µm, **B**: 0.5µm, **C**: 0.75µm (**left column**), 0.5µm (**right column**), **D**: 0.75µm (**left column**), 0.5µm (**right column**).

movement of membrane bound vesicles and organelles. This protein is necessary for intracellular transport in a variety of cellular contexts, but is particularly important in morphologically specialised cell types, such as neurons, where cargos must be transported over long distances (Goldstein, 2001). Native kinesin-I is a heterotetramer composed of two heavy chains (KHC) and two light chains (KLC) (Goldstein and Philp, 1999). Loss of function of either of these subunits results in transport related defects of neuronal function, which manifests as a phenotype with striking similarities to that of *slmo*. Mutants of the *khc* or *klc* genes exhibit progressive lethargy and locomotory defects during the second instar larval stage (Gindhart *et al.*, 1998; Hurd and Saxton, 1996). This leads to paralysis and eventual lethality prior to the transition to the third instar stage. In addition, several proteins that have been shown to interact with kinesin-I, such as Appl (Gunawardena and Goldstein, 2001), Unc76 (Gindhart *et al.*, 2003) and Sunday driver (Bowman *et al.*, 2000), exhibit similar loss of function phenotypes. These factors are thought to form part of the mechanisms that allow kinesin-I to bind particular cargos and regulate the response of the system to the needs of the axon and synapse. This raised the intriguing possibility that the Slmo protein might also play a role in axon transport.

Mutations that disrupt the function of fast axonal transport motor complexes such as kinesin-I result in defects in the trafficking of axonal cargos. These trafficking defects include the formation of axon "clogs", which are aggregates of membrane bound transport cargos such as synaptic vesicle precursors, prelysosomal vesicles and mitochondria (Bowman *et al.*, 2000; Gindhart *et al.*, 1998). Axon clogs can be visualised in the segmental nerves of an affected larva, and the *slmo* enhancer trap reports strong expression in cells that send axonal projections along these fibres (Figure 3.9A). Given that the *slmo* null mutation causes behavioural defects similar

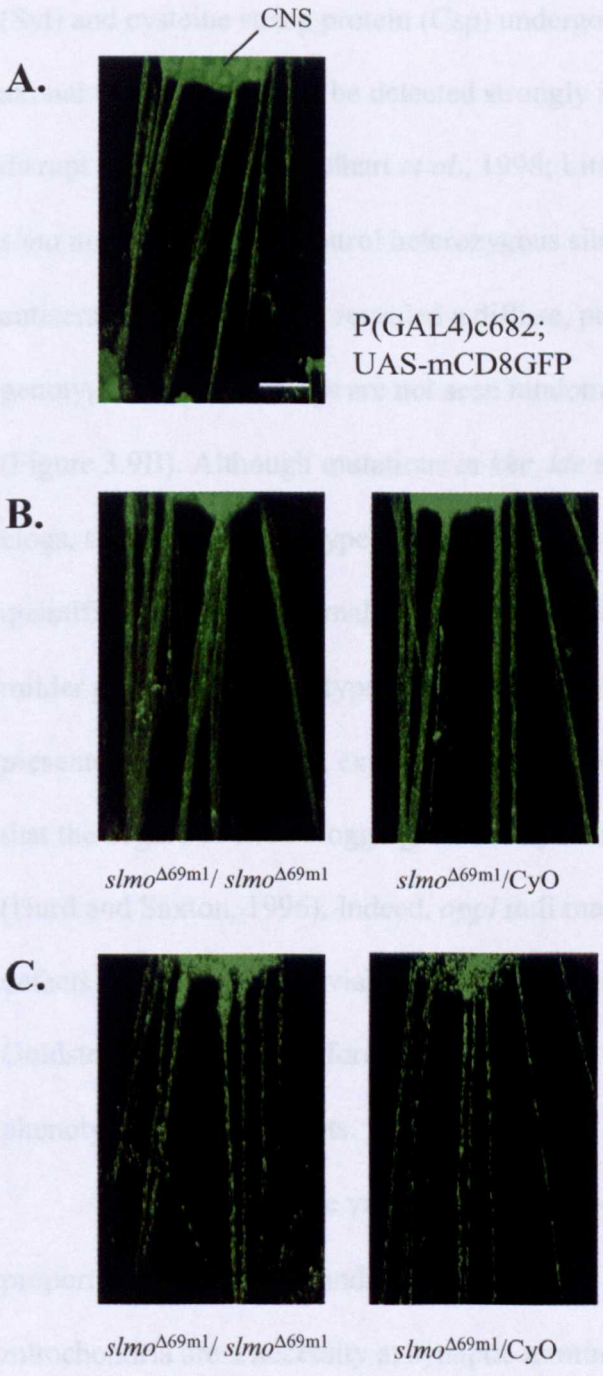


Figure 3.9: The *slmo* gene product is not required for normal axonal transport. Segmental nerves are viewed by confocal microscopy in neuromuscular preparations of aged (18-20hrs) 1st instar larvae. In all images anterior is upwards. **A**. UAS-mCD8GFP reports strong expression of the P(GAL4)c682 enhancer trap in the segmental nerves. **B**. The segmental nerves of *slmo* null mutant and control (heterozygous sibling) larvae were stained with the synaptic vesicle marker, anti-Syt. Diffuse, punctate staining is observed for both genotypes. No axon clogging is observed, indicating that axonal transport of synaptic vesicles is not affected by mutation in *slmo*. **C**. The mitochondria of *slmo* null mutant and control larvae are labelled by UAS-GFP driven by D42-GAL4. Distribution of the mitochondria in the segmental nerves is unaffected by slmo loss of function. Scale bar is 50μm.

to those observed when axonal microtubule motor function is disrupted, neuromuscular preparations of *slmo* mutant and control larvae were examined for disruption of axonal transport. Synaptic vesicle components such as synaptotagmin (Syt) and cysteine string protein (Csp) undergo high levels of microtubule based axonal transport, and can be detected strongly in the axon clogs of mutants that disrupt this function (Gindhart *et al.*, 1998; Littleton *et al.*, 1993). Aged 18-20hours *slmo* null mutants and control heterozygous siblings were immunostained with antisera to Syt (2.5). This revealed a diffuse, punctate staining pattern for both genotypes, and axon clogs are not seen randomly distributed in the nerve fibres (Figure 3.9B). Although mutations in *khc*, *klc* and *unc-76* result in many large axon clogs, the cellular phenotype of *appl* mutants is less dramatic, resulting instead in a quantifiable increase of small axon clogs (Gunawardena and Goldstein, 2001). A milder phenotype of this type cannot be formally discounted for *slmo* from the data presented here. However, experience with the various kinesin related genes suggests that the degree of axon clogging is directly related to the severity of the phenotype (Hurd and Saxton, 1996). Indeed, *appl* null mutants have detectable locomotory defects as larvae, but are viable and survive to adulthood (Gunawardena and Goldstein, 2001). It therefore seems very unlikely to apply to the very severe phenotype of *slmo* mutants.

Whilst all synaptic vesicles are exported to the nerve terminal, only a proportion of the mitochondria are transported down the axon. However, mitochondria are a necessity at synaptic terminals where they are not only required to meet the energetic demands of neuronal signalling, but are also thought to play an important role in buffering local Ca^{2+} concentrations (Berridge, 1998; Meldolesi, 2001). The kinesin associated protein Milton has been shown to play a role

specifically in the transport of mitochondria (Stowers *et al.*, 2002). Mutation causing *milton* loss of function results in axons and synaptic terminals that lack mitochondria, whilst transport of synaptic vesicles is found to be unaffected. As a result, Milton has been suggested as an adaptor protein, or part of an adaptor complex, that links the appropriate kinesin motor to mitochondria. Given the association of Slmo with the mitochondria (see chapter 4), it was important to test whether the *slmo* mutant phenotype might result from a mitochondria specific transport defect. To achieve this, the mitochondria were labelled using a GFP fusion protein that is targeted to the mitochodria by a well characterised localisation sequence. This reporter fusion, referred to as mitoGFP, was expressed under UAS control driven by D42-GAL4, which causes expression of the UAS construct in (almost) all motor neurons (A. Pilling, pers. comm.). The D42-GAL4 and UAS-mitoGFP transgenes were conveniently recombined together on the third chromosome resulting in permanently maintained expression (A. Pilling, pers. comm.). Viewed under confocal microscopy, mitochondria can be clearly seen distributed along the segmental nerves. When UAS-mitoGFP is expressed in a *slmo* mutant background, the distribution of the mitochondria is indistinguishable from wild type (Figure 3.9C). From these data, it is possible to eliminate defective axon transport as a possible cause for the *slmo* mutant phenotype.

3.3 Discussion

This chapter describes a mutant phenotype demonstrating the requirement for *slmo* in normal larval locomotor activity. Newly hatched *slmo* null mutants are mobile, but their ability to generate rhythmic locomotory peristaltic contraction waves is severely impaired. This phenotype subsequently progresses towards total

inactivity, although the ability to move the head and mouth parts is retained. In addition, I have reported an expression pattern that strongly suggests *slmo* operates primarily in the CNS throughout early development. Taken together, these data implicate *slmo* as a candidate for regulating the patterned motor output that drives basal locomotor activity in larvae.

A number of mutations defective in larval motor function have been isolated, and these are involved in a variety of developmental processes. For example, mutations in *acj6* result in larval locomotory phenotypes as a result of defects in axon pathfinding and synaptic connectivity of specific neurons, which is regulated by the POU transcription factor that the gene encodes (Certel *et al.*, 2000). Likewise, various mutant alleles of *hikaru genki* (*hig*) result in locomotory defects at both larval and adult stages, and in this case appears to encode a synaptic adhesion protein, as observed by abnormal synapse development in the mutant (Hoshino *et al.*, 1999; Hoshino *et al.*, 1996). Previously reported data on *slmo* suggest that this phenotype does not result from a gross defect in the development of the nervous system. Various preparations stained with markers of CNS morphology, combined with electrophysiological analysis, suggest that CNS morphology is normal, the neuromuscular junctions are intact, and all axonal processes are visible (Carhan, 1999; Reeve, 2002; R. Baines, pers. comm.). Although subtle abnormalities in central or peripheral nervous system development cannot be formally eliminated, the evidence available suggests that the basic circuitry required for larval locomotion is initially established. The disruption of motor output observed in *slmo* mutants would therefore appear to be the result of a defect at the functional stage of neuronal circuitry development or the modulation of an intact circuit's output. Alternatively, the circuit might have developed, but then degenerated at a later stage.

The data presented here addresses two possible hypotheses for the cause of the *slmo* phenotype. Firstly, examination of sections from the *slmo* mutant CNS established that there was no visible degeneration of the mutant CNS. In addition, examination of sections under the transmission electron microscope revealed no obvious abnormalities in the structure of neurons, glia or mitochondria (Figure 3.6). Secondly, analysis of neuromuscular preparations with markers of synaptic vesicles and mitochondria demonstrated that the phenotype was not caused by an impairment of the axonal transport machinery.

Given the highly restricted expression pattern initially observed in stage 16 embryos, I had initially hoped that a very discrete neuronal population could be identified that was critical in driving larval locomotion. This would then provide a starting point to begin unravelling the cellular basis of this patterned motor output. By stage 17 of embryogenesis, expression of the enhancer trap has expanded to a limited, but much larger population of cells, which are probably of heterogeneous neuronal cell types (Manseau *et al.*, 1997). Expression of TeTxLC in this neuronal subset results in a phenotype, which is superficially similar to that observed in *slmo* mutants, but with an increased severity that prevents the embryos from hatching. In this case the severe impairment of peristaltic movement is accompanied by a clear decrease in the synaptic drive to motoneurons. The fact that the motoneurons of toxin expressing embryos have some synaptic input indicates that expression is not found in the immediately presynaptic population of cholinergic interneurons (Baines *et al.*, 2002). This suggests that the enhancer trap does in fact mark a poorly defined population of neurons that includes interneurons presynaptic to those that directly synaptically drive the motoneurons. Initially it would be desirable to gain more

information on the identity of the neurons marked by P(GAL4)c682 at this stage and look for overlapping enhancer trap patterns that generate similar phenotypes.

Interestingly, after hatching, at the time when the *slmo* mutant phenotype first becomes apparent, reporter expression becomes markedly more refined. In the VNC, strong reporter expression marks a very discrete set of dorsally positioned neuronal cell bodies, which may represent motoneurons (Figure 3.4). Certainly, neuronal projections are clearly visible along the segmental nerves as well as in the neuropil. In the brain lobes, a small group of cell bodies are labelled in the cortex which project into specific regions defined as the dorsal neuropil and basoanterior neuropil compartments (Younossi-Hartenstein *et al.*, 2003). In addition, this concise pattern of expression appears to persist into later larval stages (Figure 3.5). Enhancer traps will often report incomplete or extraneous expression that does not accurately represent the corresponding endogenous gene. However, whilst the limitations of the enhancer trap as a reporter of endogenous gene expression should be acknowledged, the expression pattern described here is compelling with regards to the potential function of *slmo* in the nervous system. The implication is that *slmo* plays a critical role in modulating the functional output of only a small set of CNS neurons that are essential for a stereotyped motor output. It would be interesting to inactivate the neurons expressing GAL4 at the larval stages using a conditional effector gene such as UAS-*shibire*^{ts1} (Kitamoto, 2001), and look for a locomotor phenotype comparable to that caused by *slmo* loss of function.

Chapter 4: Subcellular Localisation of *Slmo*

Chapter 4: Subcellular Localisation of Slmo

4.1 Introduction

The severe behavioural phenotype and subsequent lethality caused by loss of *slmo* function suggests a vital role for the Slmo protein. However, very little is known about the biochemical function of Slmo, which contains no characterised domains and no defined localisation signals. Establishing the localisation of the protein within the cell is an important step towards establishing the cellular processes that involve Slmo. Several unsuccessful attempts have been made in our lab to raise an antibody to Slmo, which could be used to label the protein and determine its localisation at both the cellular and subcellular level (Carhan, 1999; Reeve, 2002). Most recently, Reeve (2002) was able to produce an antibody that could be used to detect Slmo on western blots, but did not work in whole mount preparations.

Although very little information was available about the nature of the Slmo protein, it was known that the *S. cerevisiae* homologue MSF1p' had been linked to the mitochondria. Researchers in Toshia Endo's lab (Nagoya University) had identified *MSF1'* while cloning the *YAP19* gene (Nakai *et al.*, 1993). Deletion of this gene is not lethal, but a myc-tagged fusion was reported to localise to the inner mitochondrial membrane. In addition, over-expression of the gene was reported to alter the location of a reporter directed to the mitochondrial intermembrane space (T. Endo, pers. comm.). Following this, it was shown that the antibody raised by Reeve (2002) was occasionally able to detect Slmo in the cytoplasm of the *Drosophila* Schneider 2 (S2) cell line. A punctate pattern was observed throughout the cytosol, as would be expected for the mitochondria (Reeve, 2002), although this staining was inconsistent and not reproducible.

An alternative approach is to engineer and express a version of the *Drosophila* Slmo protein which carries an easily detectable molecular tag. The green fluorescent protein (GFP) of *Aequorea victoria* is now used widely for this purpose, as its unique fluorescent properties provide a convenient way of visualising the target protein (Chalfie, 1998). GFP generates a strongly fluorescent signal in heterologous cell types, without the addition of any extrinsic factors. Importantly, it can be fused to other proteins without disrupting its fluorescent properties, and the resulting chimera will often retain the normal functions and localisations of the host protein. GFP fusions have now been successfully targeted to virtually every major organelle, including plasma membrane, nucleus, endoplasmic reticulum, Golgi apparatus, secretory vesicles, peroxisomes and the mitochondria (Chalfie, 1998). The range of applications of GFP have also been expanded by the isolation of mutant variants with altered spectral properties, such as the enhanced fluorescence of EGFP (enhanced GFP) and S65T, and the shifted excitation and emission spectra of BFP (blue), YFP (yellow), RFP (red) and CFP (cyan) (Chalfie, 1998).

This chapter documents attempts to gain further insight into the function of the Slmo protein by expressing GFP fusions in mammalian cell culture, and also in transgenic flies. There is considerable precedent for both these approaches (Rizzuto *et al.*, 1995; Wang and Hazelrigg, 1994). Although the Slmo protein is expressed in the *D. melanogaster* S2 cell line (Reeve, 2002), analysis of a *Drosophila* fusion protein in a well established mammalian fibroblast cell line has a number of important practical advantages. In particular, these cells are morphologically large and flat, providing an excellent system in which to study a cytoplasmic protein, and are also relatively easy to manipulate and transfect. GFP is often used in transgenic organisms to study complex phenomena such as the regulation of gene expression,

dynamics of organelles or protein sorting (e.g. chapter 3; Rizzuto *et al.*, 1995; Yeh *et al.*, 1995). The production of transgenic flies, which express GFP tagged Slmo under the control of the endogenous promoter and is able to rescue the mutant phenotype, would present a means to confidently assess the spatial and temporal expression of Slmo. In addition, a commercially available antibody could be used to label the protein in fixed *Drosophila* tissues and refine the localisation of the protein under the electron microscope. For example, if Slmo were indeed a mitochondrial protein, it would be possible to determine in which compartment(s) of the organelle the protein operates.

4.2 Results

4.2.1 Production of *slmo*-GFP and GFP-*slmo* fusion constructs

The expression of tagged protein constructs in mammalian cell lines is a commonly employed method for studying the relationship of the target protein with the cell. As a result, a variety of useful expression vectors are now commercially available for this purpose. Tagging a protein with GFP is often successful in generating a fusion that maintains all or some of the normal functions and localisation of the target protein. Certainly, it is not uncommon that the fusion of a GFP tag to the N-terminus of the target protein results in success, whilst the equivalent C-terminal fusion will result in failure, or vice versa (Chalfie, 1998). It has therefore become common practice to engineer both types of fusion protein and analyse the results in parallel. The vectors pEGFP-N1 and pEGFP-C1 (Clontech) provide a convenient means for the construction and expression of both N- and C-terminal GFP fusion proteins. The protein coding sequence of interest can readily be cloned in frame either upstream (EGFP-N1) or downstream (EGFP-C1) of an EGFP

coding sequence. When these constructs are transiently transfected into a cultured mammalian cell line, such as the NIH-3T3 fibroblast line used here, fusions are expressed under the control of the human cytomegalovirus (CMV-E1) promoter.

4.2.1.1 Cloning of *slmo* coding sequence into mammalian expression vectors

Three constructs, based on the mammalian expression vectors EGFP-N1 and EGFP-C1, were designed to express Slmo-GFP fusion proteins in mammalian cell culture. An N-terminal fusion of GFP (GFP-*slmo*), a C-terminal fusion (*slmo*-GFP), in which the *slmo* ORF is preceded by 20bps of endogenous 5'UTR, and a second C-terminal fusion (*slmo*-GFP') in which the *slmo* ORF is preceded by the vertebrate Kozak consensus sequence, GCCACC (Kozak, 2000). In all three cases PCR primers were designed to add both an *Eco*RI site upstream of the ORF, and a downstream *Kpn*I site to enable cloning into the expression vectors. For the construction of GFP-*slmo*, the primers used were 5'- aagcttgaattcg**ATG**AAAATCT GGACATCG-3' and 5'-aagcttggtaccGTCGACACCTACGTAATG-3'. For *slmo*-GFP, the primers were 5'-aagcttgaattcCGACAACGGCGCTTGCAT-3' and 5'-aagcttggtaccCCCGTAATG TGCATCGCCTT-3', and for *slmo*-GFP', the primers were 5'-aagcttgaattcGCCACC **ATG**AAAATCTGGACATCG-3' and 5'-aagcttggtaccCCCGTAATGTGCATCGCC TT-3' (*Eco*RI and *Kpn*I sites are underlined). For *slmo*-GFP and *slmo*-GFP' the primer pairs were designed to delete the endogenous *slmo* UAG stop codon.

With each primer pair the *slmo* coding sequence was amplified using a standard protocol for *Pwo* DNA polymerase (2.2.3.3), an enzyme with a high fidelity of synthesis. The *slmo* cDNA (referred to as 2A7) previously isolated by Carhan (1999), was used as the template for all three PCR reactions. Figure 4.1 shows the three amplified PCR products at the expected size for *slmo*. The PCR fragments were

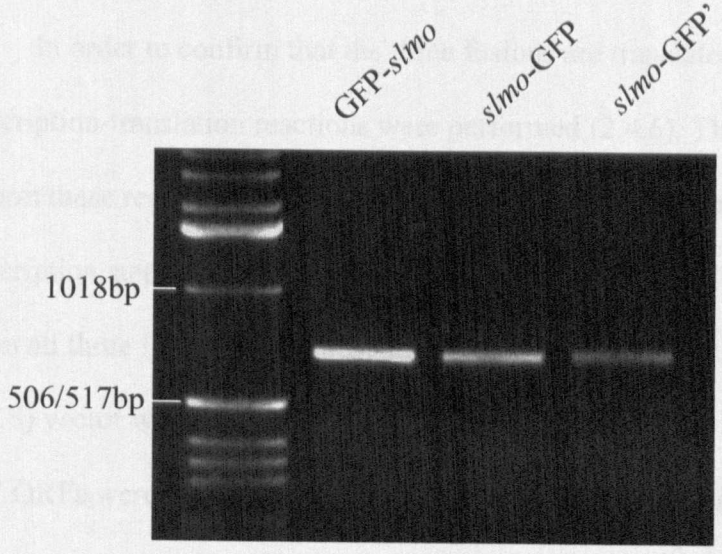


Figure 4.1: PCR amplification of the *slmo* coding region. Three different primer pairs were used to amplify and modify the *slmo* coding sequence for the construction of GFP-*slmo*, *slmo*-GFP and *slmo*-GFP'. Following amplification, all three products were resolved on an agarose gel, and single bands found to migrate between 1018bp and 506/517bp standard markers consistent with the expected size of 648bp.

then cloned (2.2.2) into corresponding restriction sites in the EGFP-C1 (for GFP-*slmo*) or EGFP–N1 (for *slmo*-GFP and *slmo*-GFP') vectors. Plasmid DNA was isolated (2.2.2.8) and the presence of the inserts confirmed by digestion with *EcoR*I and *Kpn*I restriction enzymes (Figure 4.2).

4.2.1.2 Confirmation of fusions by *in vitro* transcription/translation

In order to confirm that the three fusions are translated, *in vitro* coupled transcription-translation reactions were performed (2.4.6). The system used to perform these reactions requires standard T7, SP6 or T3 promoters for the transcription step, none of which are present in the EGFP-N1 or –C1 vectors. For this reason all three fusion coding sequences were subcloned into the pBluescriptIIKS (pBKS) vector which carries both T7 and T3 promoters. The *slmo*-GFP and *slmo*-GFP' ORFs were released from the EGFP-N1 vector by digestion with *Not*I and *Sac*I restriction enzymes, whilst GFP-*slmo* was excised using *Kpn*I and *Nhe*I. These fragments were then purified from an agarose gel and ligated into the corresponding restriction sites in pBKS (2.2.2). In the case of GFP-*slmo* the *Nhe*I site was ligated into the compatible *Xba*I site in the pBKS polylinker. Cloned plasmid DNA was digested with restriction enzymes to confirm the presence of the correctly sized inserts in all three constructs (Figure 4.3A). In each case the fusion coding fragments were orientated so that transcription can be initiated from the upstream T7 promoter (Figure 4.3B).

The *in vitro* transcription-translation reactions were performed in the presence of ^{35}S methionine and a fraction of each reaction was resolved by SDS-PAGE (2.4.4). An autoradiograph of this gel is shown in Figure 4.3C, and indicates that both *slmo*-GFP and *slmo*-GFP' are translated *in vitro*. The size of both these

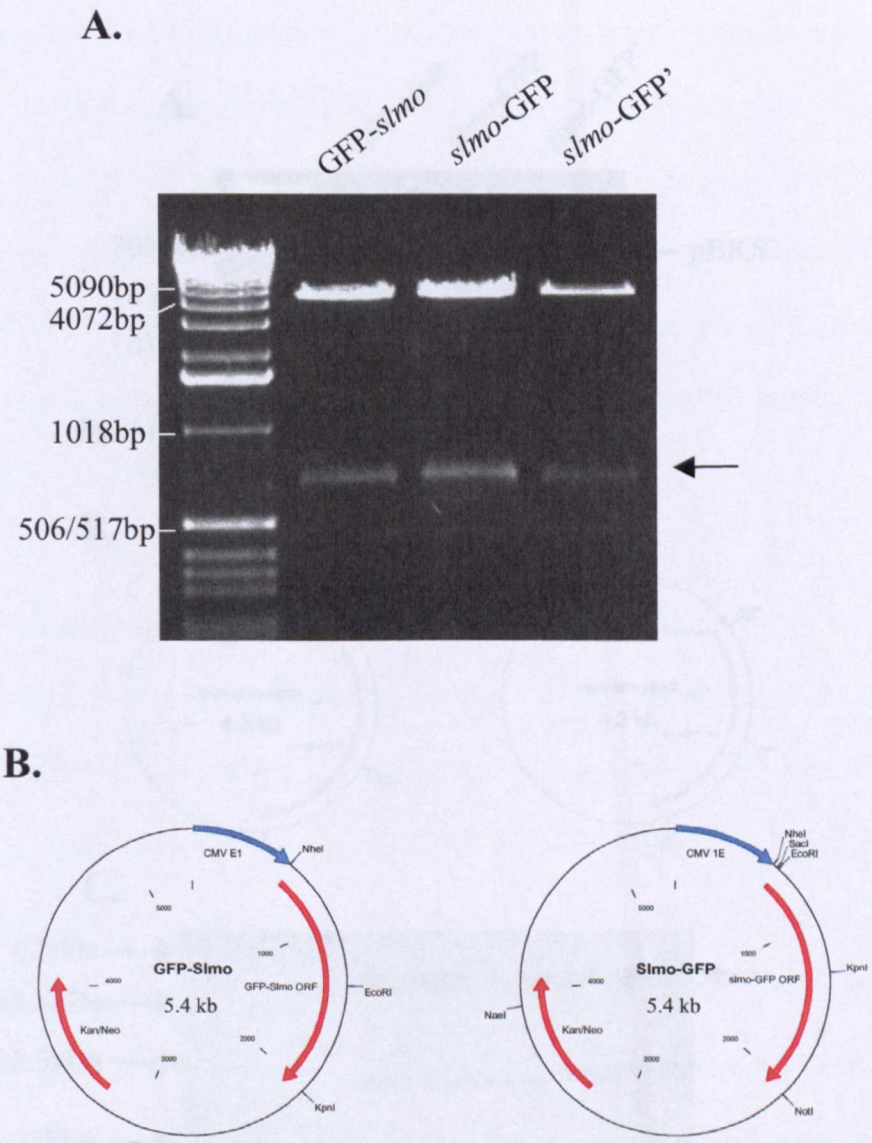


Figure 4.2: Production of GFP fusion constructs. Three constructs were engineered to express Slmo fused to GFP. The N-terminal fusion, GFP-*slmo*, was produced by cloning the *slmo* ORF into the pEGFP-C1 vector. The C-terminal fusions, *slmo*-GFP (*slmo* ORF is preceded by 20bps of endogenous 5'UTR), and *slmo*-GFP' (*slmo* ORF is preceded by the vertebrate Kozak consensus sequence, GCCACC) are based on pEGFP-N1. **A**. Gel showing diagnostic *Eco*RI and *Kpn*I double digests. The presence of the *slmo* insert (**arrow**) is confirmed in all three constructs. **B**. Schematic plasmid maps of GFP-*slmo* and *slmo*-GFP (identical to the *slmo*-GFP' map), showing relative positions of the *Eco*RI and *Kpn*I sites.

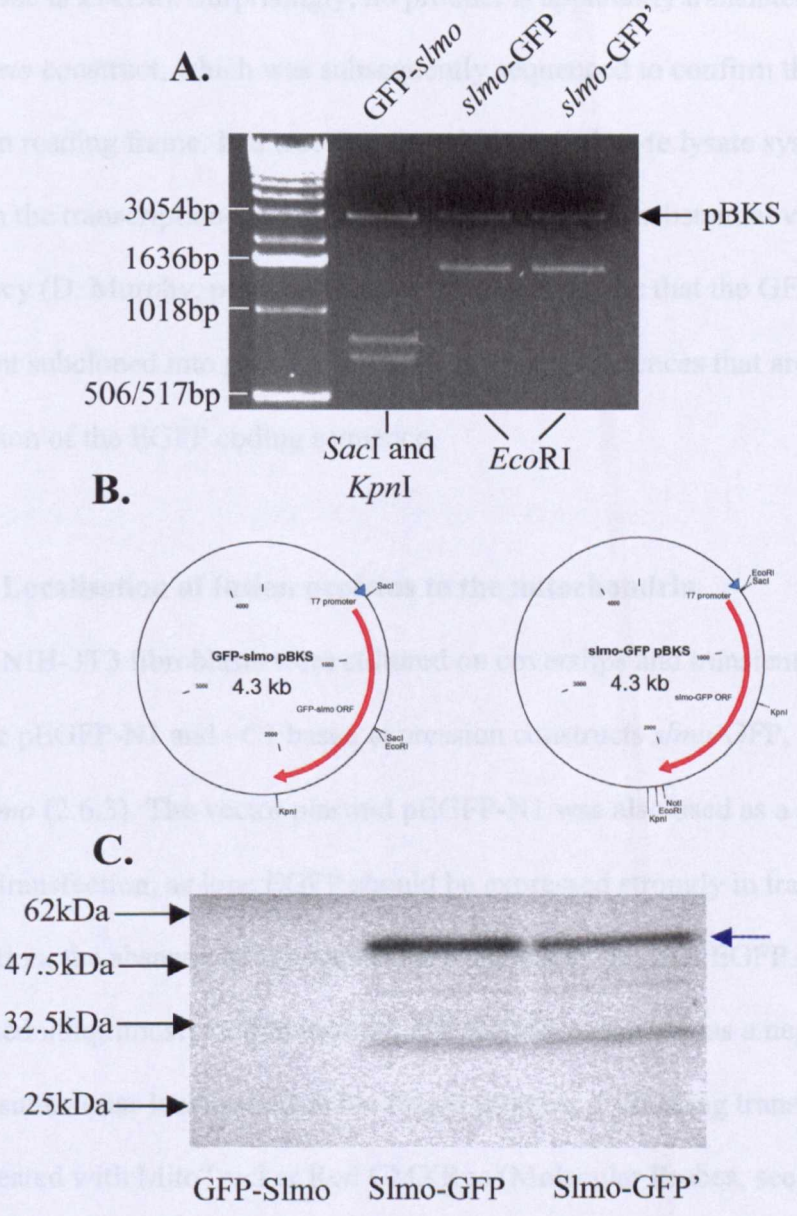


Figure 4.3: Expression of fusion proteins *in vitro*. A. The GFP-*slmo*, *slmo*-GFP and *slmo*-GFP' ORFs were subcloned into the vector pBluescript (pBKS). The presence of the inserts was confirmed by digestion with either *Sac*I and *Kpn*I (for GFP-*slmo*) or *Eco*RI (for *slmo*-GFP and *slmo*-GFP'). The GFP-slmo insert is released as two fragments of approximately 730bp and 650bp as *Sac*I cuts in between the EGFP and *slmo* coding sequences. The *slmo*-GFP and *slmo*-GFP' ORFs are both released as single fragments at approximately 1.4kb. **B.** Schematic plasmid maps showing the position of the GFP-*slmo* and *slmo*-GFP inserts (identical to *slmo*-GFP') relative to the relevant restriction sites and the T7 promoter. **C.** Autoradiograph of an SDS-PAGE gel showing the products of *in vitro* transcription-translation reactions using the above plasmids as templates. For Slmo-GFP and Slmo-GFP' product is detected at the predicted size of 52kDa (**blue arrow**). The GFP-Slmo clone is not translated in this system.

fusion proteins is consistent with the predicted size of 52kDa (Slmo alone is 25kDa, GFP alone is 27kDa). Surprisingly, no product is apparently translated from the GFP-*slmo* construct, which was subsequently sequenced to confirm the integrity of the open reading frame. It is true that the rabbit reticulocyte lysate system used to perform the transcription-translation reaction is prone to substantial variation in efficiency (D. Murphy, pers. comm.), or it may simply be that the GFP-*slmo* fragment subcloned into pBKS is missing upstream sequences that are important for translation of the EGFP coding sequence.

4.2.1.3 Localisation of fusion proteins to the mitochondria

NIH-3T3 fibroblasts were cultured on coverslips and transiently transfected with the pEGFP-N1 and –C1 based expression constructs *slmo*-GFP, *slmo*-GFP' and GFP-*slmo* (2.6.3). The vector plasmid pEGFP-N1 was also used as a positive control for the transfection, as lone EGFP should be expressed strongly in transfected cells. In addition, the absence of any localisation signals means that EGFP should be expressed ubiquitously within the cell, and therefore also acts as a negative control for the subcellular localisation of the fusion proteins. Following transfection, cells were treated with MitoTracker Red CMXRos (Molecular Probes, see methods 2.6.3); a red fluorescent die that stains the mitochondria of live cells and is resistant to fixation. The cells were then fixed briefly in paraformaldehyde before labelling the nucleus by DAPI staining (2.6.3). Cells were mounted on slides and analysed by fluorescence microscopy.

All three fusion proteins were expressed in cell culture and could be visualised by their green fluorescence, including GFP-Slmo, which was not successfully expressed *in vitro*. In each case, the tagged Slmo protein was found

exclusively in the cytoplasm in a punctate pattern. Consistent with observations in the *S. cerevisiae* system (T. Endo, pers. comm.), all three fusion proteins were seen to co-localise with the mitochondrial marker MitoTracker Red (Figure 4.4A-C), confirming that the *Drosophila* Slmo protein is indeed associated with the mitochondria. As expected, EGFP was found to be ubiquitous within the cell when expressed alone as a control, and does not co-localise with MitoTracker Red (Figure 4.4D).

4.2.2 Production of *Drosophila* GFP fusion transgenics

Transgenic *Drosophila* lines containing *slmo* based transgenes have been produced on two previous occasions. Carhan (1999) isolated a 2kb genomic *Pst*I fragment, which contains the entire *slmo* gene (Figure 4.5), and cloned this into the pCaSpeR-hs P-element vector. In this way transgenic lines were established that express *slmo* under the control of a heat shock inducible promoter. No overexpression phenotypes were observed, but the transgene was able to rescue the mutant phenotype even without the need for a heat shock regime, suggesting that the construct contains all the endogenous regulatory sequences required for normal *slmo* function. In this case, a promoter within the first exon and intron, which drives the 2021 transcript (see Figure 1.1), would therefore seem to be both necessary and sufficient for essential *slmo* expression.

As a method for studying the expression of *slmo*, Reeve (2002) chose to analyse the activity of the putative promoters. Arbitrarily chosen fragments containing presumptive promoter sequences (referred to as the "short" and "long" promoters, see Figure 4.5) were cloned into the pCaSpeR-AUG-βgal vector, which facilitates the analysis of promoters by allowing them to drive the expression of a

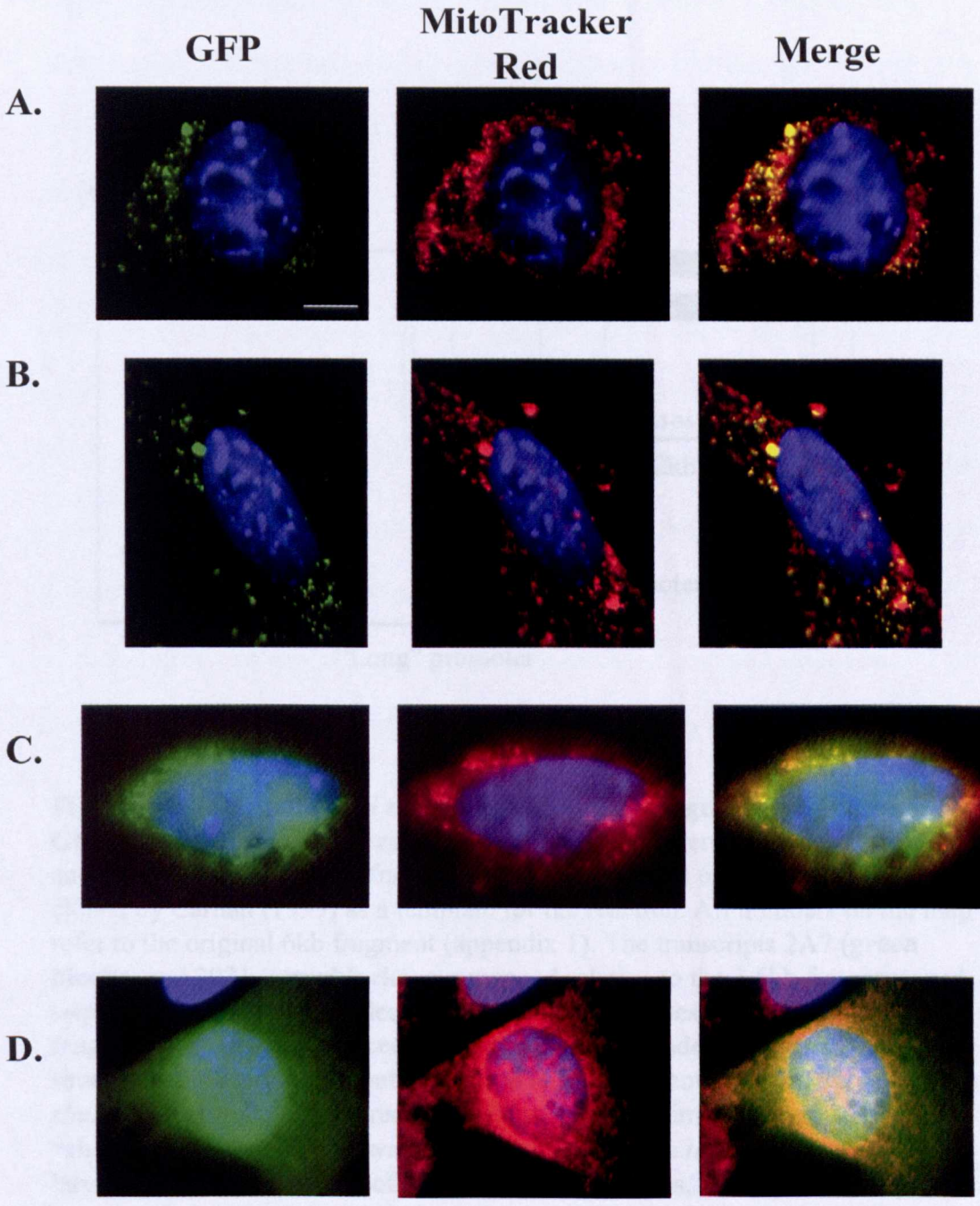


Figure 4.4: **Localisation of N- and C-terminal GFP-tagged Slmo in NIH-3T3 fibroblasts**. NIH-3T3 fibroblasts were transiently transfected with the expression constructs GFP-*slmo*, *slmo*-GFP and *slmo*-GFP'. After 24 hours, cells were fixed and labelled with the red fluorescent mitochondrial marker MitoTracker Red (**centre column**), and the nuclear marker DAPI (**blue**), before analysis by fluorescence microscopy at 100x magnification. The green fluorescent fusion proteins (**left column**) GFP-Slmo (**A**), Slmo-GFP (**B**) and Slmo-GFP' (**C**) are cytoplasmic, and co-localise with MitoTracker Red (**right column**). **D**. EGFP expressed alone is present in both the cytoplasm and the nucleus, and does not co-localise with MitoTracker Red. Scale bar: 5μm.

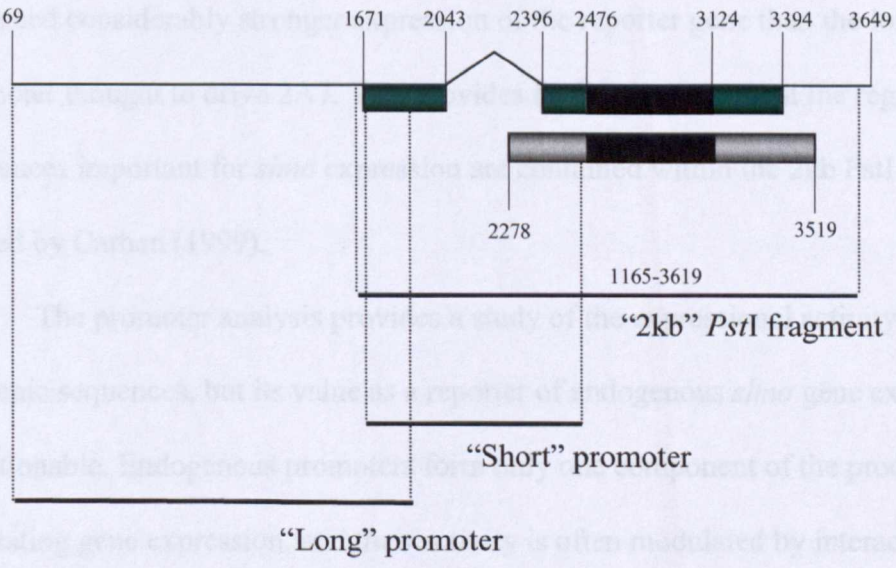


Figure 4.5: **Schematic map showing the genomic fragment used to generate GFP-*slmo* and *slmo*-GFP transgenes**. PCR primers were designed to amplify and clone a 3.6kb genomic fragment using the genomic 6kb *Bam*HI fragment cloned by Carhan (1999) as a template for the reaction. All numbers on the map refer to the original 6kb fragment (appendix 1). The transcripts 2A7 (**green blocks**) and 2021 (**grey blocks**) are mapped relative to the 3.6kb fragment, and sequences previously included in *Drosophila* transgenes. The **"2kb" *Pst*I fragment** had been introduced into transgenic flies under the control of a heat-shock inducible promoter, but was found to be sufficient to completely rescue the *slmo* null mutation without requiring a heat-shock regime (Carhan, 1999). The **"short" promoter** was shown to drive expression of a *lac*Z reporter gene in the larval nervous system and both larval and adult gonads, and is presumably responsible for regulating the 2021 transcript. The fragment labelled **"long" promoter** was included in a similar transgene and shown to have only weak promoter activity. Schematic is for reference and is not to scale.

lacZ reporter gene. As expected, reporter expression was observed from both constructs in the larval CNS, and the larval and adult gonads, although these constructs did not report any expression in the embryo. Interestingly, it was noted that the short promoter, which is thought to drive expression of the 2021 transcript, produced considerably stronger expression of the reporter gene than the long promoter thought to drive 2A7. This provides further evidence that the regulatory sequences important for *slmo* expression are contained within the 2kb PstI fragment cloned by Carhan (1999).

The promoter analysis provides a study of the expressional activity of isolated genomic sequences, but its value as a reporter of endogenous *slmo* gene expression is questionable. Endogenous promoters form only one component of the process of regulating gene expression, and their activity is often modulated by interaction with other regulatory sequences such as enhancer or suppressor elements. In addition, expression reported from these constructs does not take into account regulation in the post-transcriptional phase, and are therefore not a satisfactory substitute for a Slmo specific antibody. However, given that the putative promoter sequences have been shown to have activity (Reeve, 2002), and that only a portion of these sequences was required for rescue of the phenotype (Carhan, 1999) I reasoned that presumptive promoters could be employed to drive the expression of Slmo-GFP and GFP-Slmo fusion proteins. Ideally, these fusions would be functional *in vivo*, and would rescue the mutant phenotype. This approach would therefore have important advantages over the promoter constructs, as rescue of the mutant by a functional gene product proves that no important regulatory sequences are absent. In addition, the dynamics of the functional protein could be readily observed within the cell by GFP fluorescence or immunohistochemistry.

4.2.2.1 Construction of *slmo*-GFP and GFP-*slmo* transgenes

PCR primers were designed to amplify a 3.6kb genomic fragment containing the entire sequence which had previously been shown to rescue the *slmo* mutant (Carhan, 1999), and 1.6kb of additional upstream genomic sequence which had been shown to have weak promoter activity (Reeve, 2002; Figure 4.5). A previously cloned 6kb genomic *Bam*HI fragment (Carhan, 1999) was used as a template for the reaction. The primers were 5'-gcggccgcggatccGGCGTTCCGCTGGGGCTTCC-3' and 5'-gcggccgcggatccTTTATACACAGAATGCTTAA-3'. These primers add *Bam*HI sites (underlined) to each end of the PCR fragment, which were used to clone the fragment into the corresponding sites in the pBluescript vector (Figure 4.6A).

To generate N- and C-terminal GFP fusions, it was necessary to introduce restriction sites either immediately upstream or downstream of the *slmo* open reading frame to facilitate cloning of the GFP coding sequence in frame with the *slmo* coding sequence. This was achieved using the QuikChange™ *in vitro* site directed mutagenesis kit from Stratagene (2.2.6). A pair of oligonucleotide primers was designed, each complementary to opposite strands of the target vector, and containing the desired mutation. These were extended during thermal cycling by *PfuTurbo* DNA polymerase, which replicates both strands with high fidelity. Incorporation of the primers generates mutated copies of the plasmid containing staggered nicks. The parental DNA is removed by digestion with *Dpn*I, an endonuclease which specifically cuts methylated and hemimethylated DNA (such as the dam methylated template isolated from *E.coli* strains). The nicked, mutated plasmid DNA was then transformed into the supercompetant cell line XL1-Blue where the nicks were repaired. For the production of GFP-*slmo*, a *Bst*EII site was

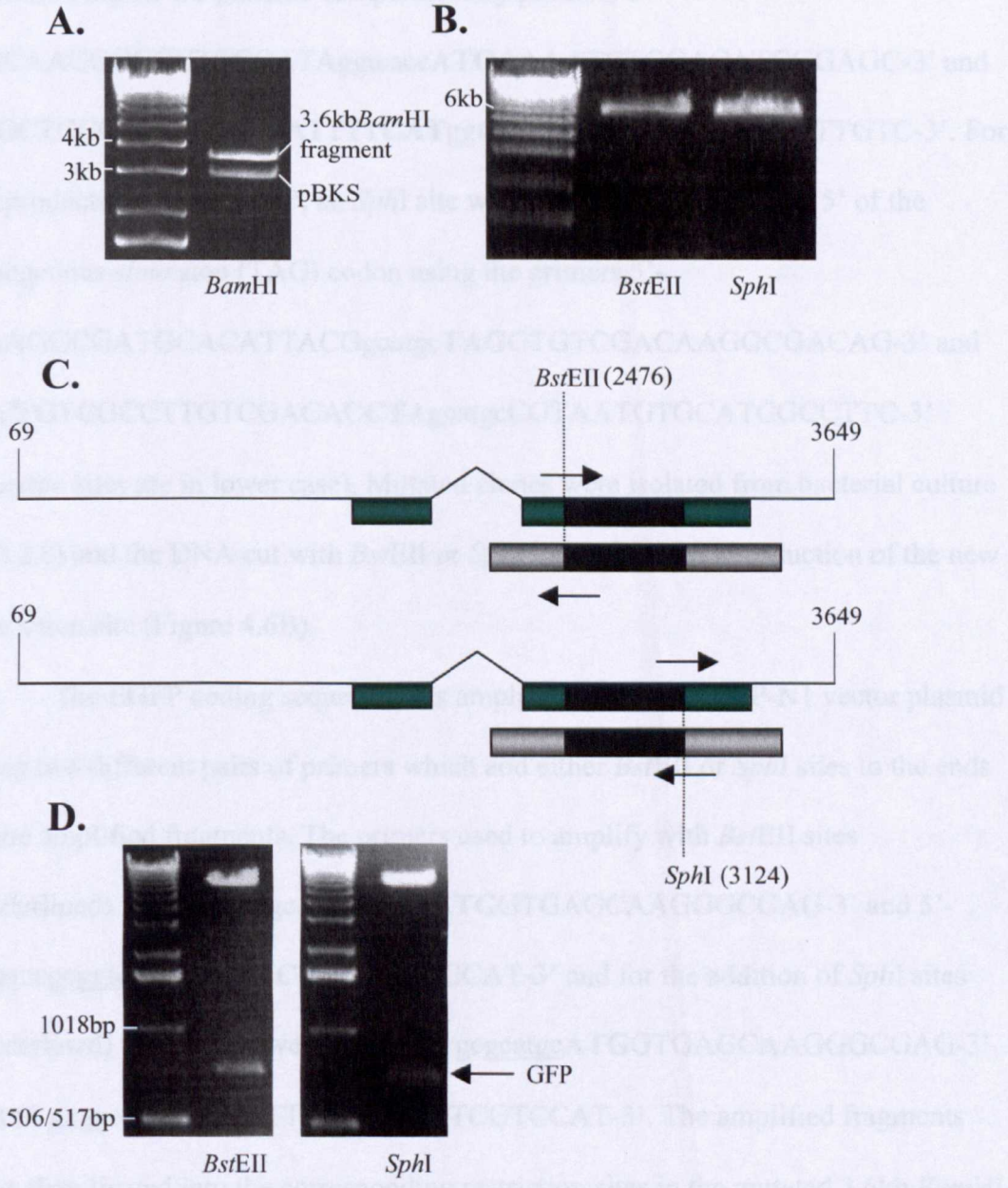


Figure 4.6: **Introduction of the EGFP coding sequence into the *slmo* gene**. **A**. A 3.6kb genomic DNA fragment, containing the slmo gene and putative promoter sequences, was cloned into the *Bam*HI sites of the vector pBKS. The presence of the 3.6kb insert was confirmed by digestion with *Bam*HI. **B,C**. To generate N- and C- terminal fusions site directed mutagenesis was used to introduce *Bst*EII or *Sph*I restriction sites either immediately upstream (for *Bst*EII) or immediately downstream (for *Sph*I) of the *slmo* ORF to facilitate the in frame cloning of the EGFP coding sequence. **B**. The introduction of the new restriction sites was confirmed by linearisation of the plasmids by digestion with either *Bst*EII or *Sph*I. **C**. Schematic showing the positions of the site directed mutagenesis primers (**black arrows**) and the introduced restriction sites to the original 3.6kb genomic fragment, the *slmo* gene (**green/grey blocks**), and the *slmo* ORF (**black blocks**). Numbers refer to the sequence in appendix 1. Schematic is for reference and is not to scale. **D**. The EGFP coding sequence was cloned into the *Bst*EII and *Sph*I sites. The presence of the insert was confirmed in both constructs by digestion with the relevant enzyme.

introduced immediately upstream of the endogenous *slmo* start (AUG) codon using the following HPLC purified complementary primers: 5'-GACAACGGCGCTTGCATAggtcacc**ATG**AAAATCTGGACATCGGAGC-3' and 5'-GCTCCGATGTCCAGATTTT**CAT**ggtcaccTATGCAAGCGCCGTTGTC-3'. For the production of *slmo*-GFP, an *Sph*I site was introduced immediately 5' of the endogenous *slmo* stop (TAG) codon using the primers 5'-GAAGGCGATGCACATTACGgcatgc**TAG**GTGTCGACAAGGCGACAG-3' and 5'-CTGTCGCCTTGTCGACAC**CTA**gcatgcCGTAATGTGCATCGCCTTC-3' (enzyme sites are in lower case). Mutated clones were isolated from bacterial culture (2.2.2.8) and the DNA cut with *Bst*EII or *Sph*I to confirm the introduction of the new restriction site (Figure 4.6B).

The EGFP coding sequence was amplified from the EGFP-N1 vector plasmid using two different pairs of primers which add either *Bst*EII or *Sph*I sites to the ends of the amplified fragments. The primers used to amplify with *Bst*EII sites (underlined) were 5'-gcggccgcggtcacc**ATG**GTGAGCAAGGGCGAG-3' and 5'-gcggccgcggtcacCTTGTACAGCTCGTCCAT-3' and for the addition of *Sph*I sites (underlined) the primers were 5'-gcggccgcgcatgc**ATG**GTGAGCAAGGGCGAG-3' and 5'-gcggccgcgcatgcCTTGTACAGCTCGTCCAT-3'. The amplified fragments were then ligated into the corresponding restriction sites in the mutated 3.6kb *Bam*HI fragment clones to generate GFP-*slmo* and *slmo*-GFP. Twelve individual *Bst*EII clones and eleven *Sph*I clones were isolated and screened for clones in which the EGFP insert was in the correct orientation (i.e. the same orientation as the *slmo* coding sequence) by sequencing across the linkers using the primers 5'-CCATCTATCGGAGCCACTGTCATC-3' (for sequencing across the *Bst*EII linker in GFP-*slmo*) and 5'-AGGACCTGCTCACCTCGACCATTA-3' (for sequencing

across the *Sph*I linker in *slmo*-GFP). The EGFP coding sequence was found to have inserted in the correct orientation in eight of the GFP-*slmo* (*Bst*EII) clones and nine of the *slmo*-GFP (*Sph*I) clones. Figure 4.6D shows examples of these GFP-*slmo* and *slmo*-GFP clones with the EGFP insert released from the plasmid by digestion with *Bst*EII or *Sph*I.

Sequencing of the clones revealed that an error had been introduced specifically in the *slmo*-GFP clones, presumably during the mutagenesis process. A single thymine residue had been deleted from position 3117 (appendix 1) near the end of the *slmo* coding sequence. This presented a problem, as any coding sequence downstream of this point mutation (including the entire EGFP coding sequence) will have been shifted out of frame. This was solved by performing *in vitro* mutagenesis (2.2.6) on a *slmo*-GFP clone for a second time using a complementary primer pair designed to reintroduce the deleted residue. In addition, the primers were purified by polyacrylamide gel electrophoresis (PAGE) rather than HPLC to increase the purity and hopefully the efficiency. The primers used were 5'-GAAGGCGATGCACAT**T**A CGgcatgc**ATG**GTG-3' and 5'-CAC**CAT**gcatgcCGT**A**ATGTGCATCCCTC-3' (the reintroduced residue is underlined in bold, the EGFP "start" codon is in bold, and the *Sph*I restriction site is in lower case). "Reverted" *slmo*-GFP clones were isolated, cut with *Sph*I to release the EGFP coding sequence and confirm their identity, and then resequenced across the linker to confirm the reintroduction of the missing thymine residue (and the absence of any additional errors).

In order to generate transgenic *Drosophila*, it was necessary to subclone the entire GFP-*slmo* and *slmo*-GFP gene constructs (including putative regulatory sequences) into a P-element transformation vector. The entire GFP-*slmo* and *slmo*-GFP inserts were excised from the pBluescript vector by double digestion with

*Bam*HI, which releases the insert from the vector and *Sca*I, which cuts within pBluescript and allows easier resolution of the fragments on an agarose gel. The inserts were excised and purified from the gel then ligated into the *Bam*HI site of pCaSpeR3 (Thummel, 1992) a P-element vector which contains the inverted repeat sequences, and selectable marker genes (*white+*) required for transformation, but does not contain any additional promoter or enhancer sequences (Figure 4.7). These constructs are referred to as pGFP-*slmo* and p*slmo*-GFP.

4.2.2.2 Generation of transgenic *Drosophila*

The plasmids pGFP-*slmo* and p*slmo*-GFP were each co-precipitated with the helper plasmid pπΔ2-3 and injected into *yw* embryos (2.7.1). The helper plasmid codes for the P-transposase enzyme needed for the integration of the transgene into the *Drosophila* genome. The helper plasmid itself is unable to integrate into the genome and is lost during cell division, whilst the integrated transgene becomes stable.

2403 *yw* embryos were injected with the plasmid pGFP-*slmo*, just 124 (5.2%) of which hatched into larvae and 34 (1.4%) eclosed as adults. 1734 *yw* embryos were injected with the plasmid p*slmo*-GFP, of which 50 (2.9%) hatched and only 8 (0.46%) eclosed as adults (2.7.2). The adults were crossed individually to flies of the genotype *w*; *CyO/Sp* and the progeny screened for the red-eyed phenotype that would indicate the fly was carrying the *white+* marker gene carried by the P-element. Many of the adults proved to be sterile (20% of those injected with pGFP-*slmo*, and 12.5% of those injected with p*slmo*-GFP) as is common following the injection process. Three of the adults injected with pGFP-*slmo* generated red-eyed progeny, and a total

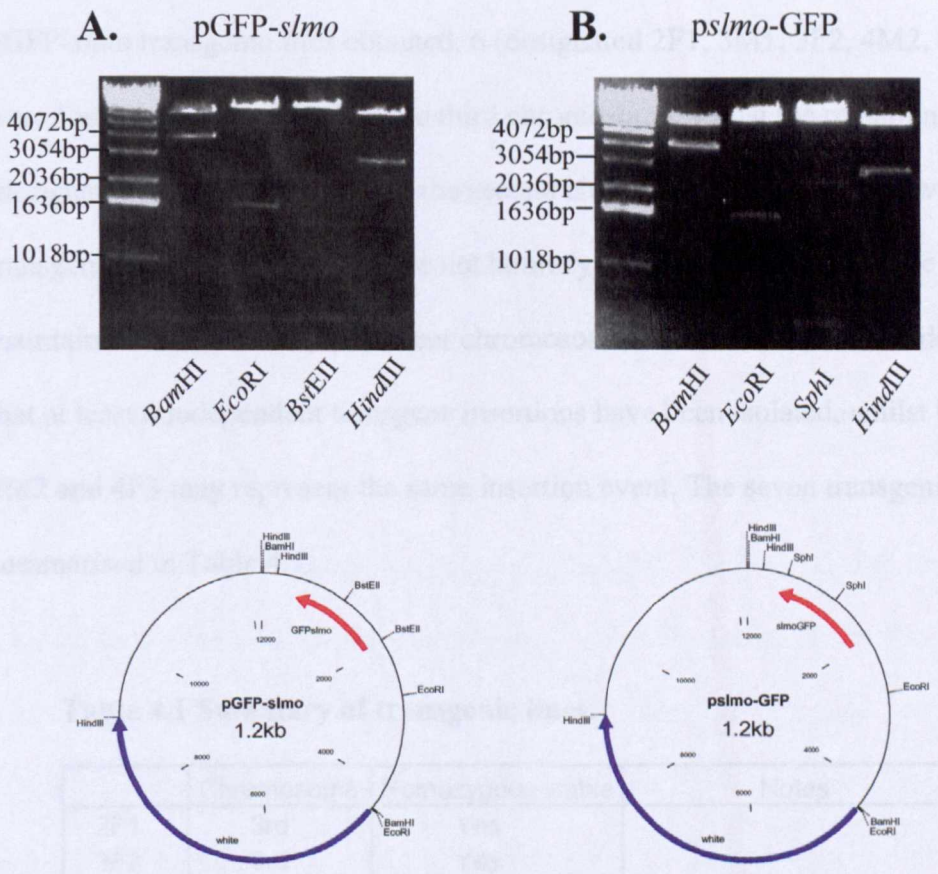


Figure 4.7: **Production of the GFP fusion transgenes**. The entire GFP-*slmo* and *slmo*-GFP gene constructs (see Figure 4.6) were subcloned into the P-element transformation vector pCaSpeR3 to generate the transgene constructs pGFP-*slmo* and p*slmo*-GFP. **A**. Constructs were cut diagnostically with the enzymes *Bam*HI (releases the entire gene from the pCaSpeR3 vector), *Eco*RI and *Hind*III (which cut at sites in the vector allowing orientation to be determined, and either *Bst*EII (for pGFP-slmo) or *Sph*I (for pslmo-GFP) to release the EGFP coding sequence. **B**. Schematic plasmid maps showing the position of the GFP-*slmo* and *slmo*-GFP insets (identical to *slmo*-GFP') relative to the relevant restriction sites.

of 7 transgenic individuals were obtained. No transgenic flies were obtained from the injection with p*slmo*-GFP.

For each transgenic individual, the transgene was genetically mapped by crossing to flies of the genotype *w*; *CyO*/*If*; TM6B/MKRS (2.7.3). Out of the 7 pGFP-*slmo* transgenic flies obtained, 6 (designated 2F1, 3M1, 3F2, 4M2, 4F2, 4F3) were found to have inserted on the third chromosome, whilst the remaining 1 (designated 4M1) was located on the second chromosome. In addition, two of the transgenic lines, 3M1 and 4F2, are not homozygous viable and have to be genetically maintained over the TM6B balancer chromosome. From this it could be determined that at least 6 independent transgene insertions have been isolated, whilst the lines 4M2 and 4F3 may represent the same insertion event. The seven transgenic lines are summarised in Table 4.1.

Table 4.1 Summary of transgenic lines.

	Chromosome	Homozygous viable	Notes
2F1	3rd	Yes	
3F2	3rd	Yes	
3M1	3rd	No	Maintained over TM6B balancer
4M1	2nd	No	Maintained over CyO balancer
4M2	3rd	Yes	Maybe same insertion as 4F3
4F2	3rd	No	Maintained over TM6B balancer
4F3	3rd	Yes	Maybe same insertion as 4M2

4.2.2.3 Analysis of transgenic lines

Having obtained stable transgenic lines it was initially important to confirm the presence of the insert in the genome by molecular analysis. To accomplish this, genomic DNA was prepared from the six third chromosome lines 2F1, 3M1/TM6B, 3F2, 4M2, 4F2/TM6B and 4F3 (2.2.4). The second chromosome line 4M1 proved to

be a very sick stock which was maintained with difficulty as a population of less than 15 flies, making it impractical to remove flies for genomic DNA extraction. Genomic DNA extracts were used as the template for PCR reactions using the primers 5'-gcggccgcgcatgc**ATG**GTGAGCAAGGGCGAG-3' and 5'-gcggccgcgcatgcCTTGTA CAGCTCGTCCAT-3' which amplify the entire EGFP coding sequence (without the "stop" codon), a sequence not found in the *yw* line originally injected). In all six lines, a fragment of the predicted size was amplified, thereby proving that these lines carry the pGFP-*slmo* transgene (Figure 4.8A).

To determine whether the fusion protein is expressed in these transgenic lines, total fly protein was extracted from each of the six 3rd chromosome lines (2.4.1), and a western blot analysis (2.4.5) was performed using a commercially produced rabbit polyclonal antibody to GFP (Ab290, AbCam). In addition, the western blot was also performed on protein extracted from line known to be expressing GFP strongly (Cha-GAL4; UAS-GFP, see Table 2.1) as a positive control, and the wild type line CS as a negative control (Figure 4.8B). No bands were detected at the expected size for the fusion protein (57kDa) in any of the six transgenic lines analysed, suggesting that the fusion protein is not expressed. Examination of dissected adult, embryonic and larval CNS, and both adult testes and ovaries revealed no fluorescence confirming that GFP-Slmo fusion protein, under the control of putative endogenous regulatory sequences, is not expressed in any of the transgenic lines produced here.

4.3 Discussion

The data presented in this chapter demonstrates an association of Slmo with the mitochondria. When expressed in a mammalian fibroblast cell line, both N- and

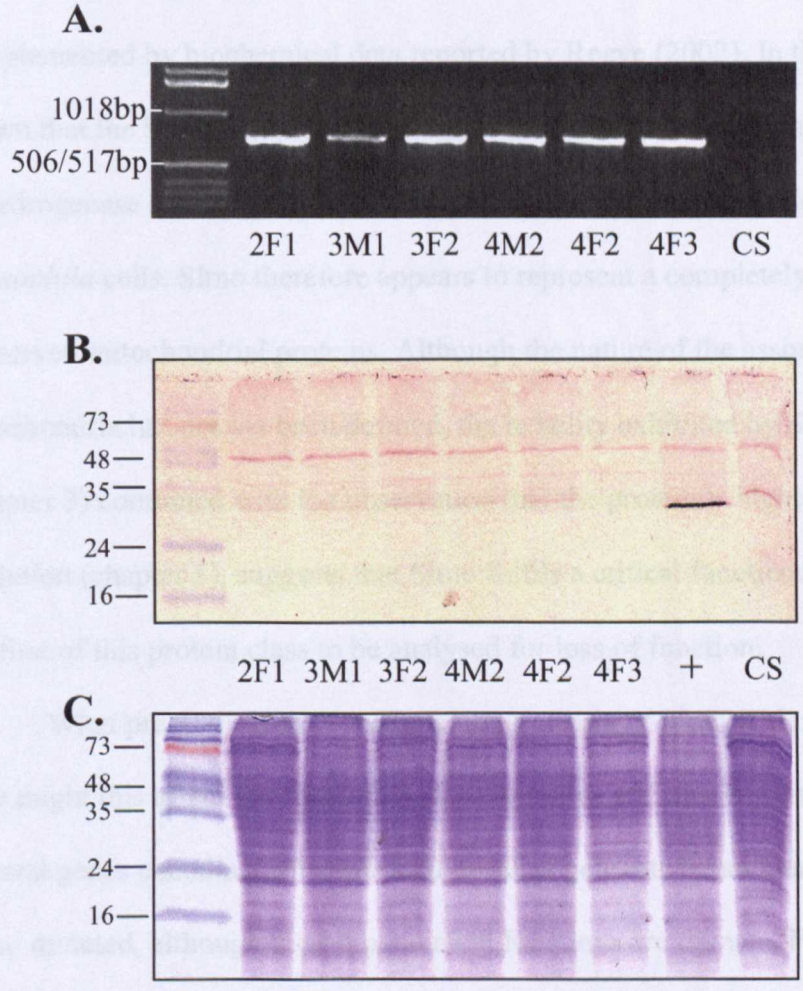


Figure 4.8: **Analysis of pGFP-*slmo* transgenic lines**. **A**. PCR for the detection of the EGFP coding sequence in *Drosophila* genomic DNA. Genomic DNA was extracted from six putative transgenic lines. In all six lines a fragment was amplified at the predicted size of 717bp confirming the presence of the transgene. No band was detected for the wild-type line Canton S (CS). **B**. Homogenates were prepared from the six transgenic lines, and subjected to western immunoblotting using a commercially available polyclonal antibody to GFP (Ab290, AbCam). A fly line known to be expressing GFP strongly (Cha-Gal4; UAS-GFP, **lane 7**), were included as a positive control, and a strong band is detected at approximately 27 kDa as expected for GFP. CS flies were also included as a negative control (**lane 8**). The fusion protein GFP-Slmo is predicted to run at 57kDa (GFP is 27kDa, Slmo runs a 30kDa, see Reeve, 2002). No band is detected at this size in any of the six transgenic lines. Bands at approximately 30, 36 and 50kDa are present in all lanes including the controls, and are presumably the result of non-specific binding of the antibody. **C**. Coomasie stained gel run in concert with **A** to show equivalence of loading.

C-terminal GFP fusions of Slmo co-localise with the mitochondrial marker MitoTracker Red. This finding is consistent with subcellular localisation reported for the yeast homologue MSF1p' (Huh *et al.*, 2003; T. Endo, pers. comm.), and is now complemented by biochemical data reported by Reeve (2002). In this case it was shown that the Slmo protein co-localises with the mitochondrial markers succinate dehydrogenase and cytochrome-C in a sucrose gradient fractionation of adult *Drosophila* cells. Slmo therefore appears to represent a completely novel class of conserved mitochondrial proteins. Although the nature of the association with the mitochondria has not yet been defined, the lethality exhibited by Slmo mutants (chapter 3) combined with the observation that the protein is highly conserved in evolution (chapter 5), suggests that Slmo fulfils a critical function. However, Slmo is the first of this protein class to be analysed for loss of function.

What precisely is the functional relationship of Slmo with mitochondria, and how might this affect the functional characteristics of Slmo expressing neurons? Several genes encoding mitochondrial proteins generate behavioural phenotypes when mutated, although their biochemical functions are diverse. For example, mutation of the *tamas* gene, which encodes the catalytic subunit of mitochondrial DNA polymerase, leads to impaired movement at the third instar larval stage and eventual lethality and is also required for normal development of the larval visual system (Iyengar *et al.*, 1999). *sluggish-A* encodes mitochondrial proline oxidase, an enzyme involved in glutamate biosynthesis and is expressed primarily in the embryonic CNS, although loss of function has only been shown to cause impaired locomotor behaviour in adult flies (Hayward *et al.*, 1993). Mutations of the *stress-sensitive B* gene, which encodes an inner mitochondrial membrane ATP/ADP antiporter, have also been reported to cause abnormal adult locomotor behaviour

(Homyk *et al.*, 1980; Zhang *et al.*, 1999), whilst mutants of *technical knockout*, a gene encoding the mitochondrial ribosomal protein S12, have various defects including deafness and "bang sensitivity" and larval lethality (Toivonen *et al.*, 2001). The disparity of the mutant phenotypes observed suggests that these are unlikely to be informative with regards to Slmo function.

Researchers in T. Endo's lab have suggested that the *S. cerevisiae* homologue MSF1p' is localised to the inner mitochondrial membrane, and that the protein might be involved in intercompartmental protein sorting (Nakai *et al.*, 1993). However, we currently have no data to infer this in *Drosophila*. It is for this reason that the failure of the GFP fusion transgenics is particularly unsatisfactory. The expression of functional GFP-tagged Slmo proteins under the control of endogenous regulatory elements would have provided the opportunity to localise the fusion to a particular mitochondrial compartment by immunogold labelling and electron microscopy. It is possible that expression of pGFP-*slmo* transgene has been silenced by the position of insertion within the genome, although at least five independent insertion lines have been analysed for expression of the fusion protein. Alternatively, the putative promoter sequences used in the construction of this transgene may not have been as complete as was originally proposed. Certainly, reporter expression driven by the promoter constructs described in Reeve (2002) was notably weak in tissues other than the germline. Carhan (1999) described rescue of the *slmo* mutant phenotype by a transgene containing only a portion of the sequences included in the pGFP-*slmo* construct (see above). Carhan's transgene had been engineered under the control of a heat shock inducible promoter, but was able to rescue both *slmo* mutant alleles without the need for a heat shock regime. This suggested that the pGFP-*slmo* transgene contained all the regulatory sequences required for *slmo* gene function. It is

possible the functional rescue observed for the heat shock transgene was the result of leaky expression from the heat shock inducible promoter at purportedly restrictive temperatures.

An alternative approach would be to generate transgenes that express GFP fusions of Slmo under the control of the GAL4-UAS system (Brand and Perrimon, 1993). Under this system, expression of the fusion proteins would not provide any direct information on the endogenous expression of Slmo protein, but significant questions on the functional requirement of the protein could be answered by testing which GAL4 lines are sufficient to rescue the *slmo* null mutation. For example, rescue of the locomotor phenotype induced by *elaV*-GAL4 (Lin and Goodman, 1994), a driver expressed only in neurons, would strongly support the hypothesis that *slmo* function is specifically neuronal during early larval development. Likewise, expression with drivers such as Cha-GAL4, which is expressed specifically in cholinergic neurons (Baines *et al.*, 2002), or TH-GAL4, which is expressed specifically in dopaminergic neurons (Friggi-Grelin *et al.*, 2003), might offer an insight into the particular neuronal populations that require Slmo.

Chapter 5: Bioinformatic Analysis of the Slmo Family

Chapter 5: Bioinformatic Analysis of the Slmo Family

5.1 Introduction

The observation that the *slmo* mutant phenotype is caused by the loss of a novel class of non-ubiquitous mitochondrial protein is of considerable cell biological interest. At the beginning of this project, it was clear that the Slmo protein was conserved, as similar sequences had been reported from a wide variety of eukaryotic organisms including human (Guzman-Rojas *et al.*, 2000; Lai *et al.*, 2000), chicken (Niu *et al.*, 1996), *S. cerevisiae* (Nakai *et al.*, 1993), *C. elegans* and *S. pombe* (Carhan, 1999). However, the extent of this evolutionary conservation had not been fully explored, and the rapidly expanding databases facilitated a more thorough analysis of this entirely new, and poorly characterised family of proteins.

Reported in this chapter are the results of an extensive database search and the identification of a highly conserved and widely spread family of Slmo related proteins. The relationships between the various members of this family are described and the existence of two predicted *slmo* related *Drosophila* genes is discussed.

5.2 Results

5.2.1 Conservation

An initial BLASTp (Altschul *et al.*, 1990) search of the Swiss-Prot database identified several Slmo related protein sequences, some of which had been described previously. These were the chick protein px19 (Niu *et al*, 1996), the human protein PRELI (Guzman-Rojas *et al.*, 2000), the human predicted protein CGI-107 (also

called CT45) (Lai *et al.*, 2000) and the *S. cerevisiae* protein MSF1p' (Nakai *et al.*, 1993). In addition, the predicted protein products of two further *Drosophila* genes, designated CG8806 and CG9528 were found to exhibit sequence similarity with Slmo. A total of three different *slmo* related genes were also identified in both *C. elegans* and *S. cerevisiae,* whilst the fission yeast *S. pombe* was found to contain only two such loci. A further tBLASTn search of the Genbank EST database identified similar sequences from a wide range of eukaryotic taxa including vertebrate model organisms such as the mouse *Mus musculus* and *Xenopus laevis*, parasitic worms (*Ancylostoma carninum* and *Strongyloides stercoralis*) and several plant species, with both angiosperms (*A. thaliana*) and gymnosperms (*Pinus taeda*) represented. A full list of the identified EST sequences is presented in appendix 2. A more recent search of the sequence databases, in the light of more complete genome sequencing projects for several model organisms, reveals the complete sequences of some of the proteins originally identified as translated ESTs. In addition, further members of the Slmo family can now be identified. In particular, it is now possible to find a total of four *slmo* related genes in both mouse and human. A list of identified Slmo related proteins/ genes from selected model organisms is presented in Table 5.1 including amino-acid sequence identities and similarities to the Slmo protein.

5.2.2 Phylogenetic analysis

The high degree of conservation of protein sequences from a broad range of species is illustrated by an amino acid sequence alignment generated using Clustal X (Thompson *et al.*, 1997) (Figure 5.1). The conserved domain appears to comprise an approximately 170 amino-acid region at the N-terminus of all the proteins identified. To further clarify the relationships between Slmo and the other members

Table 5.1A: Slmo related sequences identified in 2000

Organism	Gene/EST Name	Protein	SwissProt Accession	GenBank Accession	Identity (%)	Similarity (%)
D.melanogaster	*slmo*	Slmo	Q9V3U9			
C.elegans	F15D3.6	F15D3.6	O62173		42	58
H.sapiens	C20ORF45	CGI-107	Q9Y3B1		54	71
M.musculus	C20ORF45	CGI-107	Q9CYY7		53	70
X.laevis	BC047977			BC047977	49	69
S.cerevisiae	MSF1'	MSF1p'	P35200		34	54
S.cerevisiae	YDR185C		Q04006		33	51
S.pombe	SPBC36.10		O59707		37	55
D.melanogater	CG8806		Q95RE6		26	42
C.elegans	B0334.4	B0334.4	Q17476		24	41
H.sapiens	PRELI	PRELI	Q9Y255		25	46
M.musculus	Preli	Preli	Q8R107		25	46
X.laevis		Preli-like		AAH44981	22	44
Gallus gallus	Px19	Px19	Q90673		27	45
S.cerevisiae	YLR193C		Q05776		26	44
S.pombe	SPAP8A3.10		Q9UT07		29	48
D.melanogaster	CG9528		Q9VMD6		24	44
C.elegans	T23G5.2	T23G5.2	Q03606		26	43
A.thaliana	T19L5_30		Q9FYA3		28	49
L.esculentum (Tomato)	EST303596			AW218413	29	50
Pinus taeda	RTCNT1_73_B08			CF672720	31	47
Ceratopteris ricardii	Cri2_8_P09_SP6			BE643491	31	50

Table 5.1B: Slmo related sequences identified in 2004

Organism	Gene/EST Name	Protein	SwissProt Accession	GenBank Accession	Identity (%)	Similarity (%)
H.sapiens	KIAA0420	KIAA0420	O43304		22	47
M.musculus	SEC14L1	1200017E04 Rik	Q9DBQ0		22	46
H.sapien	MGC21644	MGC21644		MGC21644	21	43
M.musculus	C330008K14Rik	C330008K14 Rik	Q9CRD8		21	46
S.tuberosum (Potato)	EST504505			BG599610	27	48

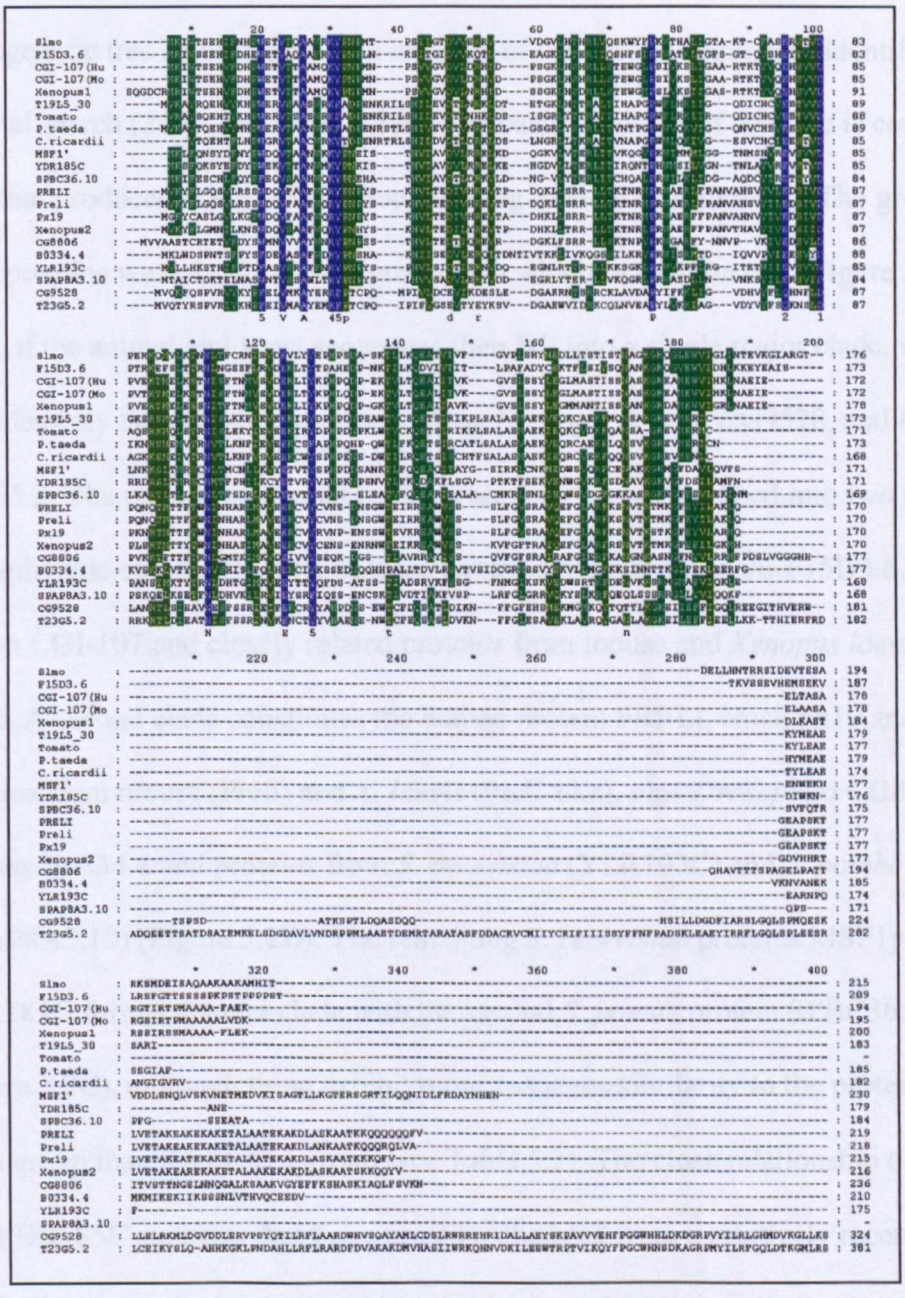


Figure 5.1: **Multiple sequence alignment of the Slmo protein family**. Full length (215 amino-acid) *Drosophila* Slmo was aligned to related proteins from selected model organisms using Clustal X. A conserved domain is identified that extends for approximately 170 amino-acids at the N-terminus. In this region a high degree of evolutionary conservation is observed between eukaryotic taxa as widely diverged as human (CGI-107, PRELI), *S. cerevisiae* (MSF1p', YDR185C, YLR193C) and *Arabadopsis thaliana* (T19L5_30). Universally conserved residues are highlighted in **blue**, widely conserved residues are in **green**.

of this protein family, a phylogenetic analysis was performed using the maximum likelihood (ML) program TreePuzzle (Schmidt *et al.*, 2002). Figure 5.2 shows the phylogenetic tree resulting from an analysis of the protein sequences identified in the original search (2000, see Table 5.1A). The basic topology of this tree is consistent with that produced using a Neighbour-Joining method (appendix 3). The group of plant proteins were used as an outgroup with which to root the tree (Figure 5.2A). Most of the animal and yeast sequences then fall into a single major clade, which excludes only the predicted products of the *Drosophila* gene CG9528, and *C. elegans* T23G5.2. The remaining animal proteins are then further grouped into two main monophyletic clades. The first of these contains Slmo, *C. elegans* F15D3.6, the human CGI-107 and closely related proteins from mouse and *Xenopus laevis* (Figure 5.2B). A second clade constitutes the human protein PRELI, chick px19 and similar proteins from mouse (Preli) and *X. laevis* (Preli-like), plus *Drosophila* CG8806, *C. elegans* B0334.4 and proteins from *S. cerevisiae* (YLR193C) and *S. pombe* (SPAP8A3.10) (Figure 5.2D). The remaining *S. cerevisiae* proteins MSF1p' and YDR185C form a distinct clade with the second *S. pombe* protein SPBC36.10 (Figure 5.2C), although these exhibit more sequence similarity to the proteins of the Slmo group than the PRELI group (see Table 5.1). The close relationship of MSF1p' and YDR185C suggests that these proteins diverged from a relatively recent gene duplication event.

The predicted proteins of the *Drosophila* CG9528 and *C. elegans* T23G5.2 form a separate group that appears to have diverged significantly from the other proteins (Figure 5.2E). In particular it was clear that this *C. elegans* protein was much larger, comprising 743 amino acids, than any other identified member of the family (the second largest identified protein being the 236 amino-acid product of

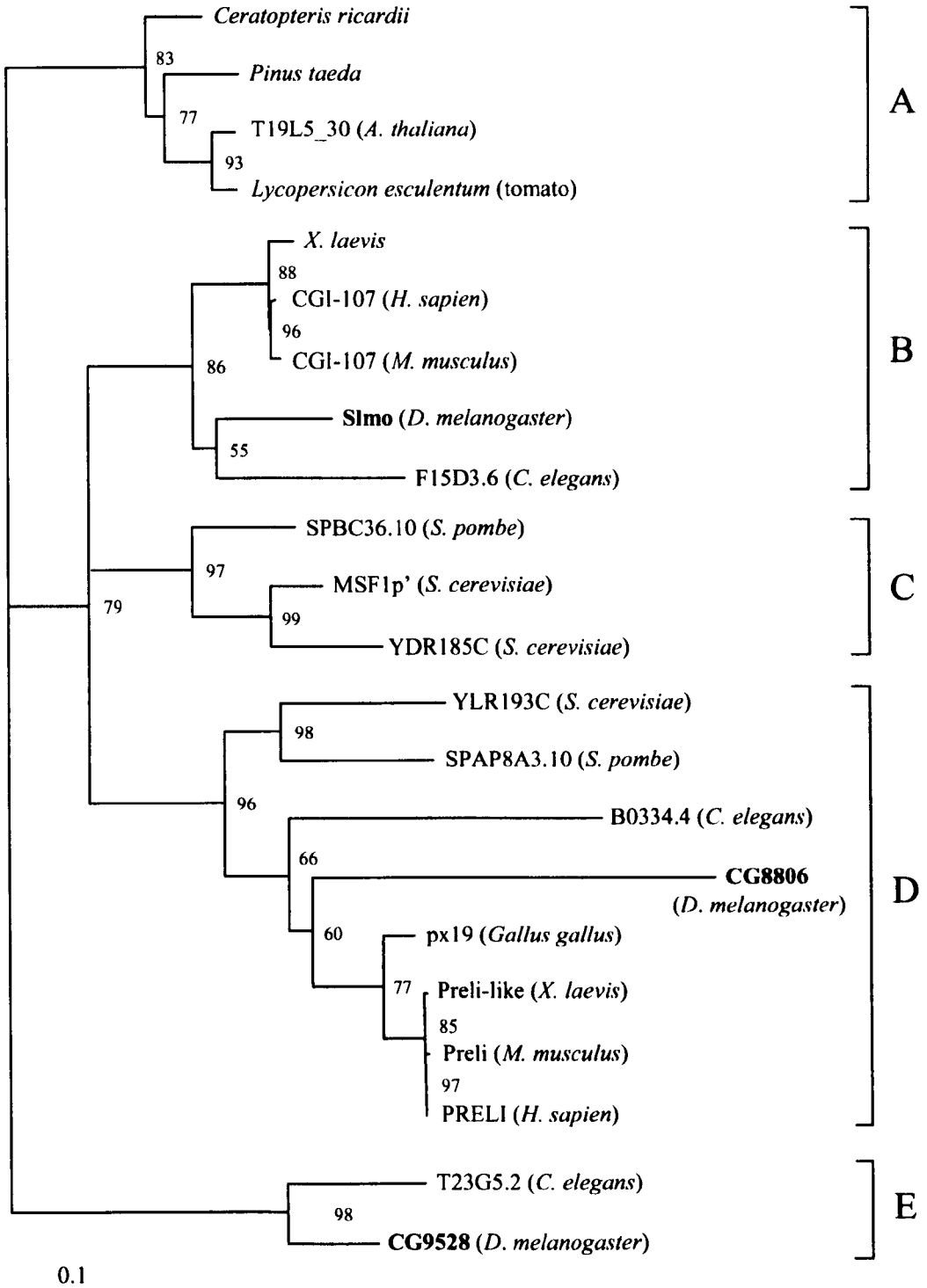
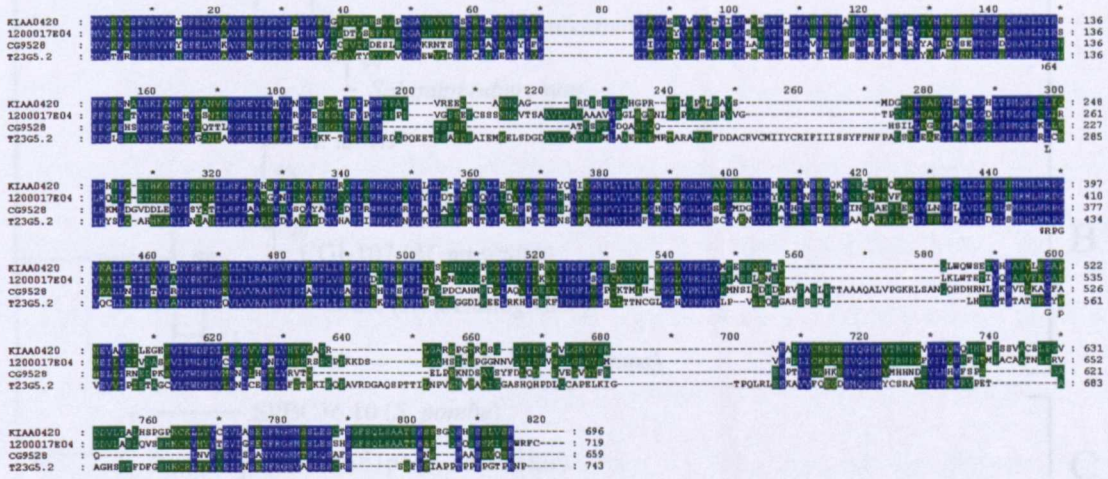


Figure 5.2: Phylogeny of the Slmo family of proteins (2000). Phylogeny was generated from a ClustalX protein sequence alignment using the maximum likelihood program TreePuzzle (Schmidt *et al.*, 2002), with 10,000 puzzling steps and a uniform rate of substitution. The phylogeny was rooted using the plant sequences as an outgroup and branching nodes are labelled with maximum likelihood support values. Proteins are grouped into five major clades corresponding to **A.** plant proteins **B.** sequences most closely related to Slmo, **C.** MSF1p' and closely related yeast proteins, **D.** sequences most closely related to the human PRELI protein, and **E.** predicted proteins that contain a C-terminal CRAL-TRIO domain.

CG8806). Analysis of the T23G5.2 protein sequence using the PFAM database of protein families (Bateman *et al.*, 2000) revealed the presence of a C-terminal CRAL-TRIO (cellular retinaldehyde binding protein/ triple function) domain (PFAM entry: PF00650), a conserved module thought to be involved in the binding and delivery of small lipophilic molecules. Closer inspection of the *Drosophila* genome sequence suggested CG9528 was also likely to contain such a domain, and that the predicted gene sequence was simply incomplete at the time, an observation that is confirmed by the current annotation of the *Drosophila* genome (http://fly.ebi.ac.uk). It is now possible to identify both human and mouse proteins carrying an N-terminal Slmo-like module (Slmo/MSF1p' domain), and a CRAL-TRIO domain (Table 5.1B). In addition, this group of proteins has also been shown to be associated with the GOLD domain, thought to be involved in mediating protein-protein interactions (Anantharaman and Aravind, 2002). The domain structure of these proteins is illustrated in Figure 5.3B. The CRAL-TRIO domain is also highly conserved as demonstrated by an amino acid sequence alignment of these proteins (Figure 5.3A).

Figure 5.4 shows a phylogenetic analysis of the Slmo protein family including protein sequences identified from a recent (2004) database search. The same basic branching pattern is observed, although a single branching node is missing which was previously seen to group most of the animal sequences together into a single major clade (Figure 5.2). It should be noted that the ML statistical confidence value applied to this node in Figure 5.2 was not very high (79). As expected, the CRAL-TRIO domain containing proteins identified in mouse and human form a single clade with those previously identified in *Drosophila* and *C. elegans* (Figure 5.4E). A fourth Slmo related protein can also be identified in both mouse and human and these proteins also fall together with the CRAL-TRIO domain

A.



B.

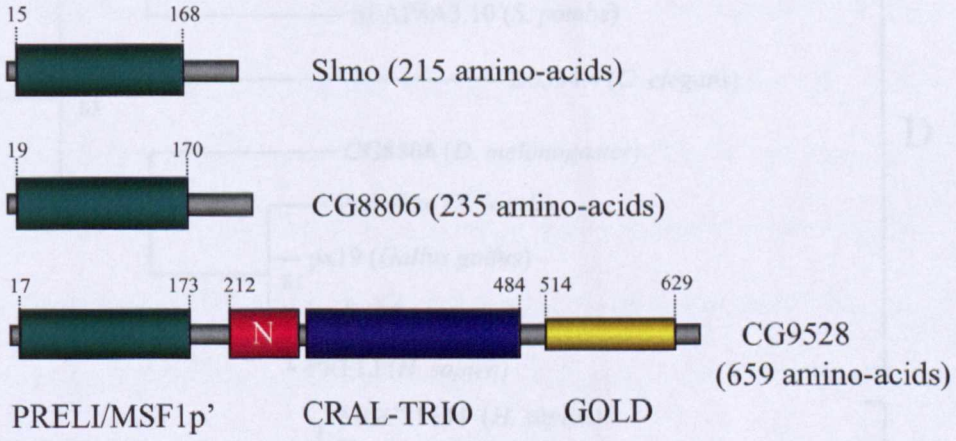


Figure 5.3: **Slmo related proteins containing the CRAL-TRIO domain**. Predicted proteins were identified in *Drosophila* (CG9528), *C. elegans* (T23G5.2), human (KIAA0420) and mouse (1200017E04) containing an N-terminal Slmo-like (PRELI/MSF1p') domain combined with the CRAL-TRIO domain and GOLD domains. **A**. Multiple sequence alignment of the CRAL-TRIO domain proteins using ClustalX. A high degree of conservation is observed along the full length of these proteins. Conserved residues are illustrated in **green**, whilst universally conserved residues are shown in **blue**. **B**. Schematic comparing the domain structures of the three *Drosophila* proteins Slmo, CG8806 and CG9528. As is the case for most proteins in the family, Slmo and CG8806 exist as stand alone PRELI/MSF1p' domains (**green**). The CRAL-TRIO domain (**blue**) exists in conjunction with the CRAL-TRIO-N domain (**red**) and is believed to be involved in the binding small hydrophobic ligands (Li *et al.*, 2000). The C-terminal GOLD domain is suggested to mediate protein-protein interactions (Anantharaman and Aravind, 2002). Limits of all domains were determined from the PFAM database except the GOLD domain which is described in Anantharaman and Aravind (2002).

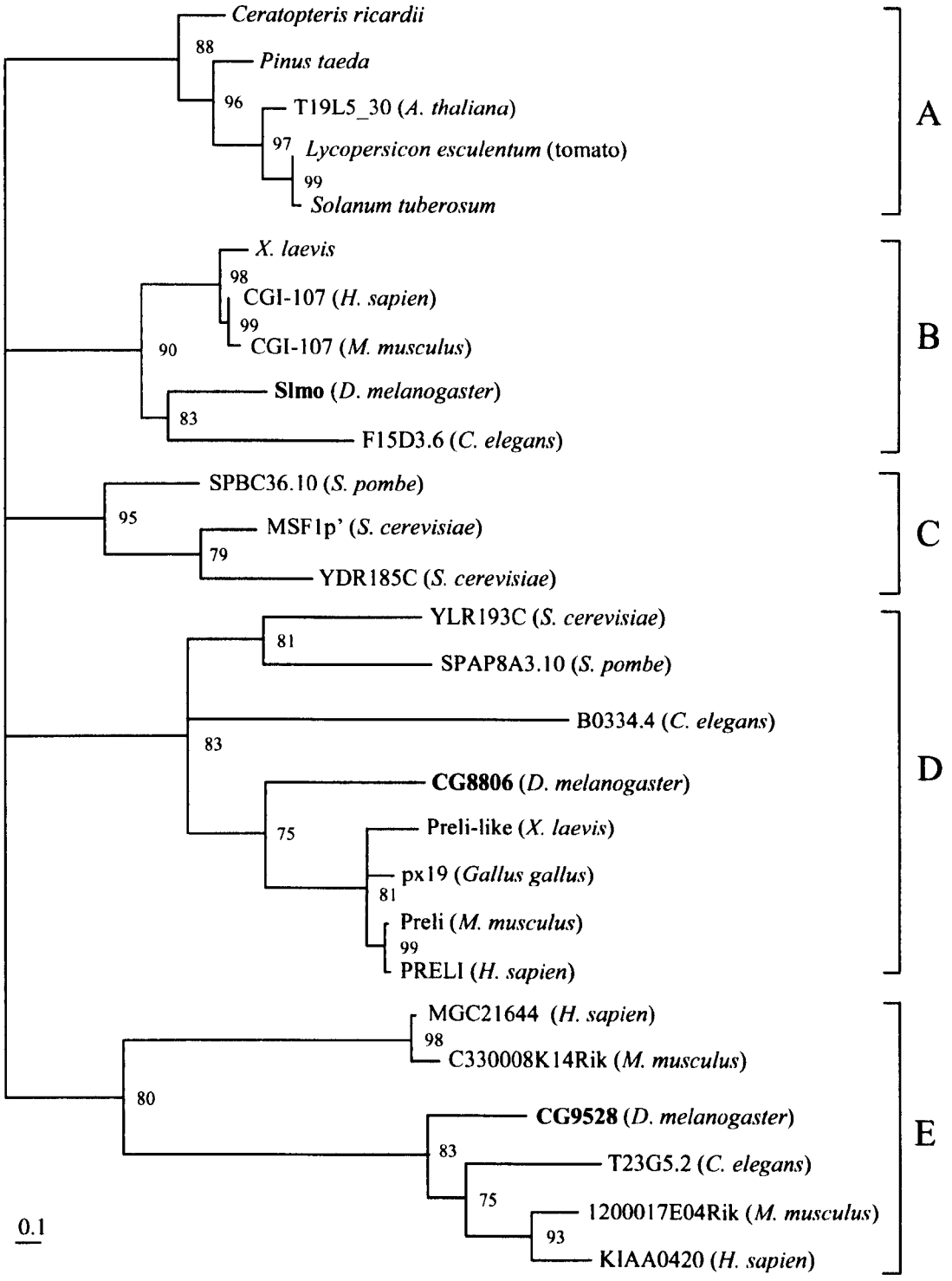


Figure 5.4: Phylogeny of the Slmo family of proteins (2004). Phylogeny was generated from a ClustalX protein sequence alignment using the maximum likelihood program TreePuzzle (Schmidt *et al.*, 2002), with 10,000 puzzling steps and a uniform rate of substitution. The phylogeny was rooted using the plant sequences as an outgroup and branching nodes are labelled with maximum likelihood support values. Proteins are grouped into five major clades corresponding to those described in Figure 5.2 (**A-E**). Clade **E** includes two additional sequences from human and mouse containing the CRAL-TRIO motif (1200017E04Rik and KIAA0420) and two less closely related sequences that do not appear to contain this domain (C330008K14Rik and MGC21644).

group in a single larger clade (Figure 5.4E), although neither of these proteins appear to contain either the CRAL-TRIO or GOLD domains.

5.3 Discussion

The identification of Slmo related proteins from organisms as diverse as yeast, plants, worms, flies and mammals may suggest that these proteins are a requirement in all eukaryotic organisms. Whilst the biochemical function of these proteins remains unknown, such widespread conservation would point to an ancient and vital role that has yet to be established. In animal species, the existence of multiple *slmo*-like genes appears to be the norm, with the invertebrates *Drosophila* and *C. elegans* containing three such genes, whilst mouse and human are found to possess at least four. Phylogenetic analysis indicates that the animal proteins of this family form three major groups. The first subgroup is composed of proteins apparently directly homologous to Slmo, which includes the CGI-107 proteins from human and mouse, F15D3.6 from *C. elegans* and may also include the previously identified yeast protein MSF1p'. It is currently unclear whether these proteins are related as true functional homologues, although the human CGI-107 is reportedly expressed in the adult central nervous system (BodyMap project; Sese *et al.*, 2001). Knockouts of these genes in yeast (T. Endo pers. comm., Nagoya University) and *C. elegans* (http://www.wormbase.org, release WS115, date 15/01/04) do not result in lethality. However, it is interesting to note that *C. elegans* lacking F15D3.6 expression exhibit a post-embryonic "slow growth" defect.

The second subgroup is comprised of apparent homologues of the previously reported human gene *PRELI* (Guzman-Rojas *et al.*, 2000), and includes the closely related *px19* from chick (Niu *et al.*, 1996), and the *Drosophila* computed gene

designated CG8806. These proteins have not been assigned to any particular biological role, although expression of *PRELI* and *px19* has been reported in a variety of tissues. In particular, expression of the *px19* gene was reported to be strong in the major sites of haematopoiesis, the blood islands of the area opaca during early development, and the blood cells of the liver during later embryonic development (Niu *et al.*, 1996). In the human foetus, *PRELI* is expressed predominantly in the liver, which also represents the major foetal haematopoietic organ, and the most important site for B-cell lymphopoiesis (Guzman-Rojas *et al.*, 2000). This has been interpreted as evidence of an important role for PRELI in these stem cell systems, however it should be noted that this protein is highly expressed in a variety of adult tissues including brain, lung, spleen and kidney and may be involved in the postnatal maturation of these organs. Previous reports on px19 and PRELI noted the presence of the LEA (Late Embryogenesis Abundant) motif (Guzman-Rojas *et al.*, 2000; Niu *et al.*, 1996), a consensus sequence previously reported only in plants (Puupponen-Pimia *et al.*, 1993). However, this motif is not found in any of the other family members including those proteins identified from plants and its functional significance is unclear.

Perhaps the most significant result of this analysis is the discovery of a third subgroup in which the Slmo/MSF1p' domain is not found as a "stand alone" protein, but rather as the N-terminal portion of a larger protein featuring a functionally characterised protein module known as the CRAL-TRIO domain. This conserved motif has been shown to constitute a hydrophobic lipid binding pocket, and is common to several proteins with established roles in mediating the transfer of small hydrophobic molecules through the aqueous phase. These include proteins involved in a variety of biological processes. For example, CRALBP (cellular retinaldehyde

binding protein) is a critical component of visual cycle, where it is responsible for the transport of the physiological ligands 11-cis-retinol and 11-cis-retinaldehyde and modulates interaction of these retinoids with visual cycle enzymes (Crabb *et al.*, 1998). Another example is the α-tocopherol transfer protein (α-TTP), a mammalian protein shown to mediate the recognition, binding and transport of vitamin E (α-tocopherol) (Arita *et al.*, 1995). The mammalian supernatant protein factor (SPF) on the other hand, has been linked to the metabolism of another hydrophobic substrate, squalene, where it has been suggested to regulate the availability of this molecule to the enzymes of the sterol biosynthetic pathway and ultimately the maintenance of cholesterol homeostasis (Shibata *et al.*, 2001).

The yeast phosphatidylinositol transfer protein, SEC14p, an example of an isolated CRAL-TRIO domain, mediates the binding and transfer of either phosphatidylinositol (PI) or phosphotidylcholine (PC), an activity which is essential for many cellular processes (Li *et al.*, 2000). By controlling the availability of these molecules, the transfer protein forms an integral part of the machinery regulating downstream signalling and membrane transport events. In *S. cerevisiae*, the protein operates specifically in the Golgi apparatus, and plays an essential role in regulating the efficient production of exocytotic vesicles (Li *et al.*, 2000). However, homologous *SEC14* genes can control distinct physiological processes, even though the underlying biochemical function is likely to be the same. For example, in the budding yeast species *Yarrowia lipolytica*, SEC14p is not required in the Golgi, but instead plays a role at the plasma membrane related to the dimorphic transition from the yeast to the mycelial form (Lopez *et al.*, 1994). In addition, proteins with similar activity in higher eukaryotes have been shown to play a variety of complex roles

including signal transduction and membrane transport (Cockcroft, 1998; Cockcroft and De Matteis, 2001).

Significantly, all of the three major subgroups of the Slmo family are represented by a *Drosophila* gene, these being *slmo*, CG8806 and CG9528. The two computed genes, which have been named *prelil* (PRELI-like) and *real-time* (see chapter 6) respectively, offer an opportunity to further explore the possible functions of the proteins of this new family.

Chapter 6: Analysis of the *slmo* Related *Drosophila* Genes *prelil* and *real-time*

Chapter 6: Analysis of the *slmo* Related *Drosophila* Genes *prelil* and *real-time*

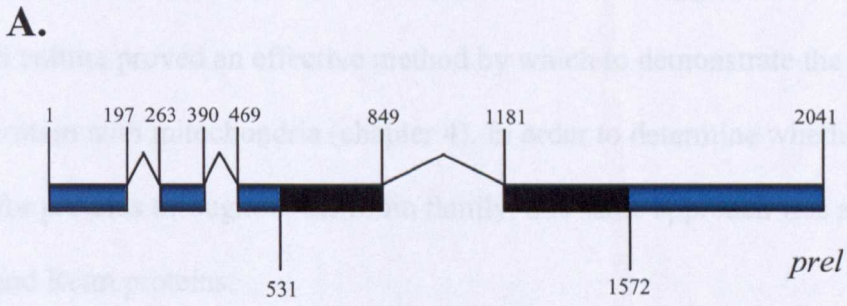
6.1 Introduction

Phylogenetic analysis demonstrated the existence of three diverged subgroups within the Slmo protein family, which may be functionally diverged. The three related *Drosophila* genes, *slmo*, *prelil* (*prel*) and *real-time* (*retm*) are each representative of one of these subgroups, and therefore provide an opportunity to gain insight into the function of the various factors within this new class of proteins. The gene structures of *prel* (CG8806), which maps to the 45D5 region of chromosome 2, and *retm* (CG9528), which maps to 26D4-5 of chromosome 2, are presented in Figure 6.1. Neither *prel* nor *retm* had been reported previously in any molecular or biochemical studies, though the presence of a CRAL-TRIO domain within the Retm protein suggests an undefined role for this factor in the transport of small hydrophobic molecules (Allen-Baume *et al.*, 2002). This chapter presents the first analysis of *prel* and *retm* function, including localisation, expression and the consequences of loss-of-function.

6.2 Results

6.2.1 Subcellular localisation of Prel and Retm

The knowledge that both Slmo and the *S. cerevisiae* MSF1p' are mitochondrial proteins, suggests that this is likely to be true of closely related homologous proteins such as vertebrate CGI-107 and *S. pombe* SPBC36.10. However, an association with either the mitochondria, or any other cellular compartment had never been demonstrated for any of the more distantly related



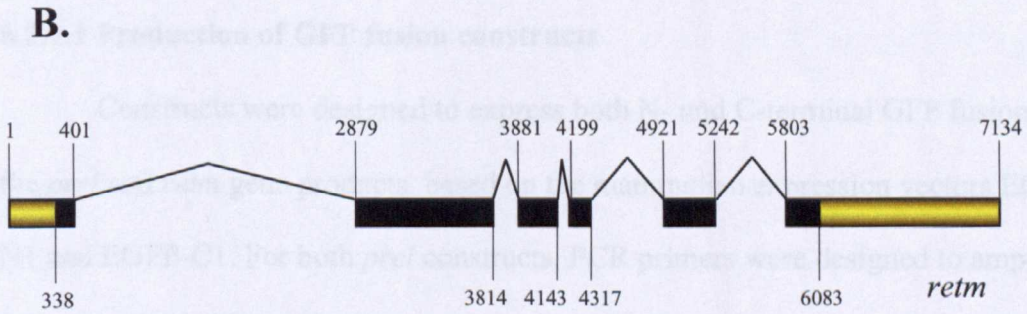


Figure 6.1: **Schematic representation of the *prel* and *retm* gene structures**. The intron exon structure of the *prel* (**A**) and *retm* (**B**) genes is presented based on gene region and mRNA sequences obtained from the GadFly database [gene IDs: FBgn0033413 (*prel*) and FBgn0031814 (*retm*)]. Exons are represented by **coloured blocks**. ORFs are represented by **black blocks**. All numbers refer to the gene region sequences presented in appendix 4 (*prel*) and appendix 5 (*retm*). Schematic is for reference and is not to scale.

proteins. This was particularly pertinent in the case of those proteins shown to contain the CRAL-TRIO and GOLD domains, which are both known to be associated with the Golgi apparatus in a variety of other proteins (Anantharaman and Aravind, 2002; Li *et al.*, 2000). The expression of GFP-tagged Slmo fusion proteins in cell culture proved an effective method by which to demonstrate the association of this protein with mitochondria (chapter 4). In order to determine whether this was the case for proteins throughout the Slmo family, this same approach was applied to the Prel and Retm proteins.

6.2.1.1 Production of GFP fusion constructs

Constructs were designed to express both N- and C-terminal GFP fusions of the *prel* and *retm* gene products, based on the mammalian expression vectors EGFP-N1 and EGFP-C1. For both *prel* constructs, PCR primers were designed to amplify the *prel* ORF with the addition of an upstream *Eco*RI site and a downstream *Kpn*I site to facilitate cloning into the expression vector. For the two *retm* constructs, the primers were designed to include an upstream *Xho*I site and a downstream *Kpn*I site. In the case of the C-terminal fusions *prel*-GFP and *retm*-GFP, the primers were designed to introduce the vertebrate Kozak consensus sequence (GCCACC) (Kozak, 2000) immediately upstream of the ORF. For the construction of GFP-*prel*, the primers were 5'-gcaagcttgaattcg**ATG**GTCGTGGCAGCATCCAC-5' and 5'-gcaagcttggtacc**TCA**GTTCTTGACAGAAAACA-3', and for *prel*-GFP, the primers were 5'-aagcttgaattcGCCACC**ATG**GTCGTGGCAGCATCCAC-3' and 5'-gcaagcttggtaccccGTTCTTGACAGAAAACAGTT-3'. For GFP-*retm* the primers were 5'-gcaagcttctcgagct**ATG**GTGCAAAAATTCCAGTC-3' and 5'-gcaagcttggtacc**TCA**T CGACTCTGCACCGAGC-3' and for *retm*-GFP the primers

were 5'-aagcttctcgaGCCACC**ATG**GTGCAAAAATTCCAGTC-3' and 5'-gcaagcttggtacccTCGACTCTGCACCGAGCTGG-3' (*Eco*RI, *Xho*I and *Kpn*I sites are underlined, start and stop codons are in bold). For *prel*-GFP and *retm*-GFP the primers were designed to delete the endogenous UGA stop codons.

The *prel* and *retm* coding sequences were then amplified with the above primer pairs using the high fidelity enzyme *Pfu* DNA polymerase (2.2.3.3). As templates for these PCR reactions, full length cDNAs were obtained for *prel* (LD36516, GenBank Acc. AI456484) and *retm* (GH05975, GenBank Acc. AI517941) as IMAGE clones (www.hgmp.mrc.ac.uk). These amplified fragments were found to run at the expected sizes for *prel* and *retm* on an agarose gel (Figure 6.2), and were cloned into to the relevant restriction sites in the EGFP-C1 (for GFP-*prel* and GFP-*retm*) or EGFP-N1 (for *prel*-GFP and *retm*-GFP) vectors. Plasmid clones were isolated, and the presence of the inserts confirmed by double digestion with *Kpn*I and either *Eco*RI (for GFP-*prel* and *prel*-GFP) or *Xho*I (for GFP-*retm* and *retm*-GFP) restriction enzymes (Figure 6.3). All constructs were sequenced to confirm the integrity of the fusions.

6.2.1.2 Expression of GFP fusions in cell culture

NIH-3T3 fibroblasts were cultured on coverslips and transiently transfected with each of the four expression constructs GFP-*prel*, *prel*-GFP, GFP-*retm* and *retm*-GFP (2.6.3), with the expression vector pEGFP-N1 used as a positive control for the transfection. After transfection, cells were treated with the red fluorescent mitochondrial marker MitoTracker Red CMXRos (Molecular Probes, see 2.6.3). Cells were then fixed briefly in paraformaldehyde, mounted on slides and analysed by confocal microscopy (2.6.5). Fusion of GFP to the N-terminus of Prel (GFP-Prel)

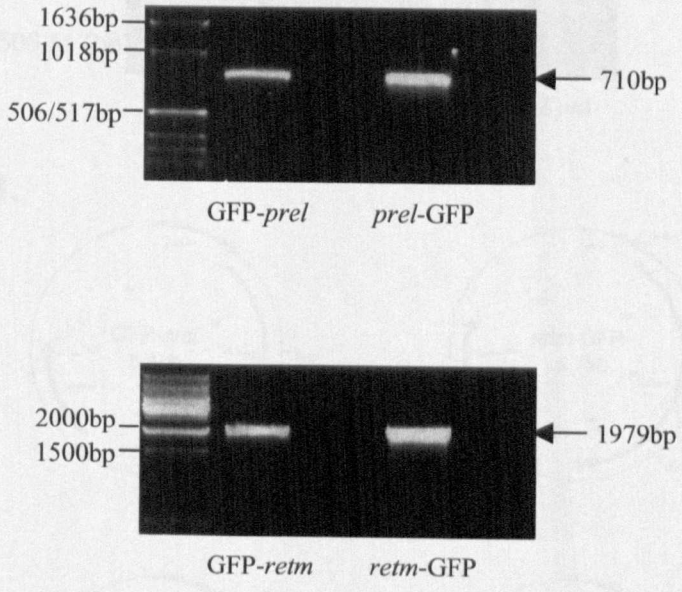


Figure 6.2: **PCR amplification of the *prel* and *retm* coding regions**. The *prel* and *retm* coding sequences were each amplified using two different primer pairs for the construction of GFP-*prel*, *prel*-GFP, GFP-*retm* and *retm*-GFP. Following amplification all for products were resolved on an agarose gel, and single bands were found to migrate at the expected sizes for both *prel* (710bps) and *retm* (1979bps).

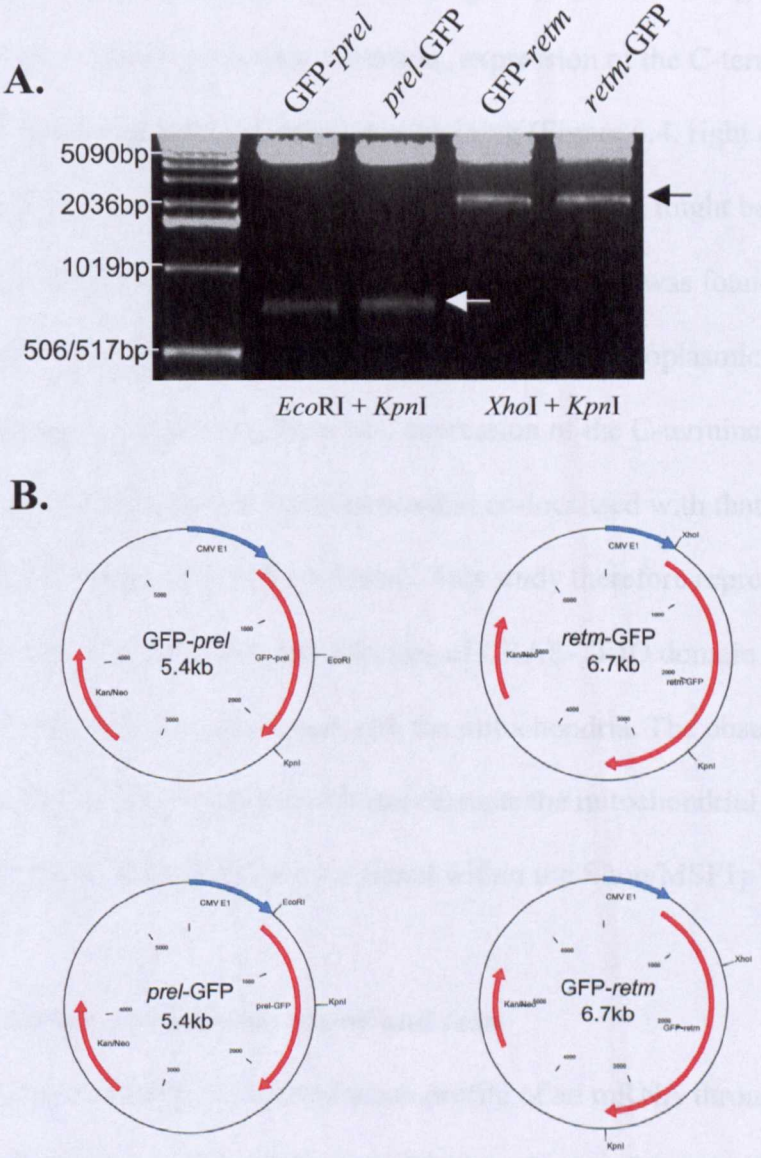


Figure 6.3: **Production of Prel and Retm GFP fusion constructs**. Constructs were engineered to express GFP fused to the N- and C- termini of Prel and Retm. The N-terminal fusion constructs GFP-*prel* and GFP-*retm*, were produced by cloning the *prel* and *retm* ORFs into the pEGFP-C1 vector. The C-terminal fusions *prel*-GFP and *retm*-GFP utilised the vector pEGFP-N1. **A**. All four constructs were cut diagnostically with restriction enzymes *Kpn*I and either *Eco*RI (for GFP-*prel* and *prel*-GFP) or *Xho*I (for GFP-*retm* and *retm*-GFP) and resolved on an agarose gel. The presence of the insert is confirmed in all constructs at the expected sizes for the *prel* (710bps, **white arrow**) and *retm* (1979bps, **black arrow**). **B**. Schematic plasmid maps of the four constructs showing the relative positions of the relevant restrictions sites.

resulted in a pattern of fluorescence that coincided with that observed upon staining the cells with MitoTracker Red CMXRos (Figure 6.4, left column), indicating that Prel is found in the mitochondria. However, expression of the C-terminal fusion, Prel-GFP resulted in diffuse cytoplasmic staining (Figure 6.4, right column), suggesting that an important mitochondrial targeting signal might be located at the C-terminus of the protein. In the case of Retm, the reverse was found to be true with the N-terminal fusion, GFP-Retm showing only diffuse cytoplasmic staining (Figure 6.5, left column). Importantly however, expression of the C-terminal fusion, Retm-GFP, resulted in a pattern of fluorescence that co-localised with that of MitoTracker Red CMXRos (Figure 6.5, right column). This study therefore represents the first direct evidence that the Retm-like subclass of CRAL-TRIO domain proteins (see chapter 5) is likely to be associated with the mitochondria. The observation that fusion of GFP to the N-terminal of Retm disrupts the mitochondrial localisation suggests that this is determined by a signal within the Slmo/MSF1p' domain.

6.2.2 Endogenous expression of *prel* and *retm*

An examination of the expression profile of an mRNA through a series of developmental stages can provide an insight into the possible role of the encoded protein. The *slmo* gene is known to be expressed throughout development, although the transcript appears to be restricted to the CNS during embryonic development (Carhan, 1999), and analysis of the enhancer trap suggests this is also the case during larval stages (see chapter 3). An obvious question is whether the related genes *prel* and *retm* are also associated with the developing nervous system, or perhaps fulfil a similar biochemical function in different tissues or developmental stages. To establish the expression of these genes during embryonic development, Digoxygenin

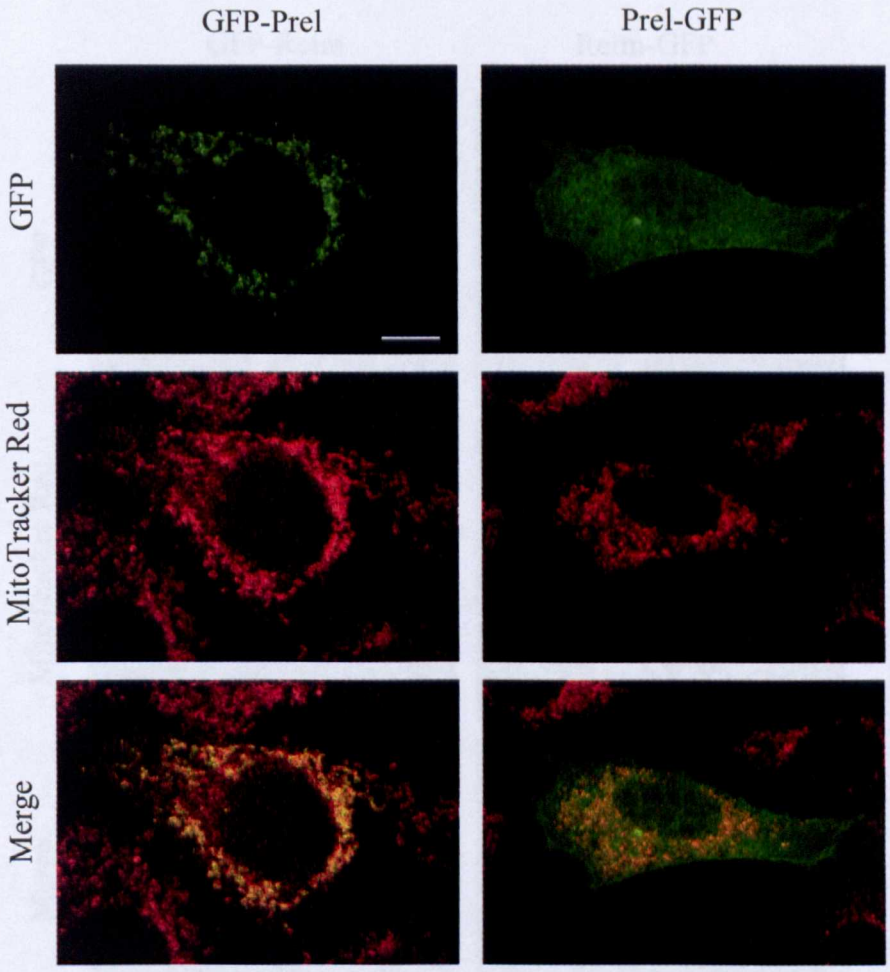


Figure 6.4: **Localisation of N- and C-terminal GFP-tagged Prel in NIH-3T3 fibroblasts**. NIH-3T3 fibroblasts were transiently transfected with the expression constructs GFP-*prel* and *prel*-GFP. Cells were also stained with the red fluorescent mitochondrial marker MitoTracker Red CMXRos and analysed by confocal microscopy. Expression of the GFP-Prel fusion results in a pattern of fluorescence (**green**) that co-localises with that of the mitochondrial marker (**red**), indicating that Prel is a mitochondrial protein. Expression of the C-terminal fusion, Prel-GFP, results only in diffuse cytoplasmic staining. Scale bars represent 5μm.

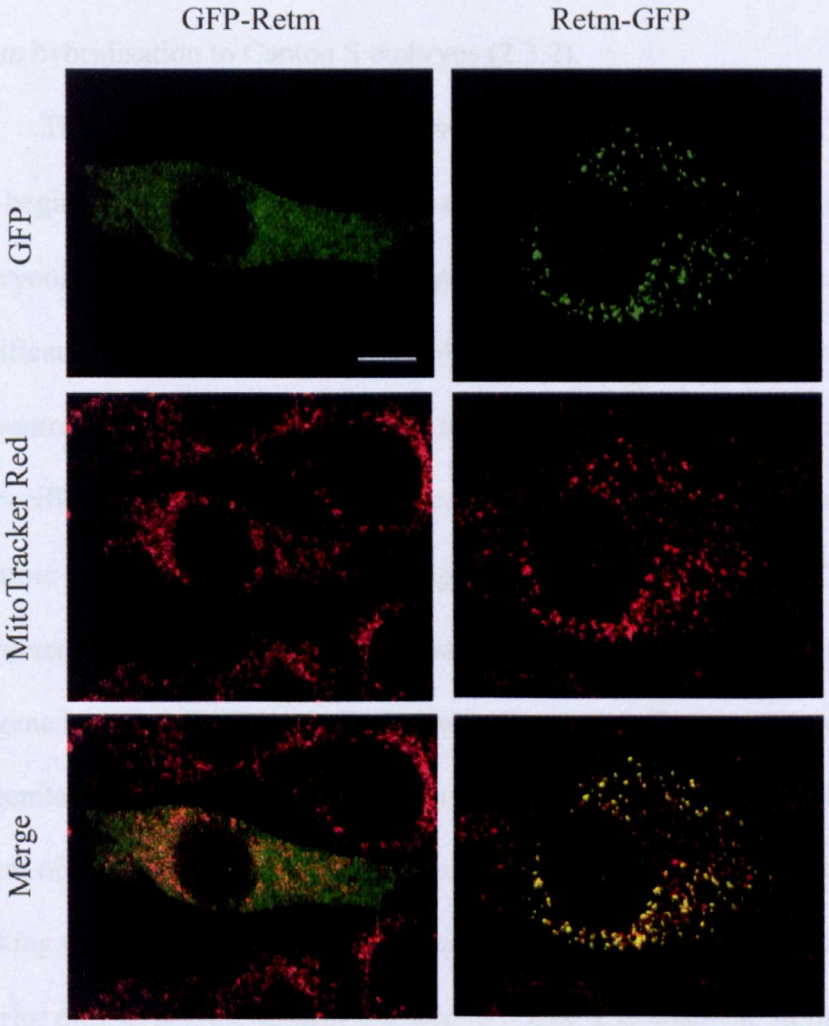


Figure 6.5: **Localisation of N- and C-terminal GFP-tagged Retm in NIH-3T3 fibroblasts**. NIH-3T3 fibroblasts were transiently transfected with the expression constructs GFP-*retm* and *retm*-GFP.The mitochondria were labelled by staining the cells with the red fluorescent dye MitoTracker Red CMXRos. Expression of the N-terminal fusion protein, GFP-Retm (**left column**), results in a diffuse pattern of fluorescence in the cytoplasm (**green**). Expression of the C-terminal fusion, Retm-GFP (**right column**), results in a pattern of fluorescence (**green**) that co-localises with the **red** fluorescence of the mitochondrial marker. Scale bars represent 5μm.

(DIG) labelled riboprobes, specific to the *prel* and *retm* transcripts, were manufactured as described in 2.3.2, using the full length cDNA IMAGE clones (see 6.2.1.1) as templates. The expression of *prel* and *retm* mRNA was then detected by *in situ* hybridisation to Canton S embryos (2.3.2).

The embryonic expression of *prel* is presented in Figure 6.6. Transcription of *prel* begins at the earliest stages of development, and is ubiquitous throughout embryonic development. At stage 16, *prel* expression appears to be upregulated significantly in the developing brain lobes, although, interestingly, this is not true of the ventral nerve cord. In addition, an increased level of expression can also be seen in specific regions of the midgut. Expression of the *retm* gene is considerably more dynamic (Figure 6.7). The *retm* message is first transcribed at stage 7, where it is expressed in the early germ band. By stage 11, when the germ band is fully extended, the gene is expressed strongly in the ventral neuroectoderm, which represents the progenitor of the central nervous system (Goodman, 1993). At this time, a small region of particularly strong staining is visible at the tip of the extended germ band marking the proctodeum (the primordial hindgut) and expression is also visible in the anterior midgut and the foregut primordial region known as the stomodeum (Skaer, 1993). Expression then persists in both the emerging gut and nervous system during the next three developmental stages (12-14). At stage 15, *retm* is no longer expressed in the newly formed midgut, but staining continues in the hindgut as it continues to develop by projecting towards the antero-dorsal region of the embryo and is also evident in differentiating components of the foregut, which might include the stomatogastric nervous system (Hartenstein, 1993). Interestingly, *retm* expression is detected strongly throughout the developing CNS and the intensity of this staining is maintained through stage 16, whilst the expression in regions of the gut is rapidly

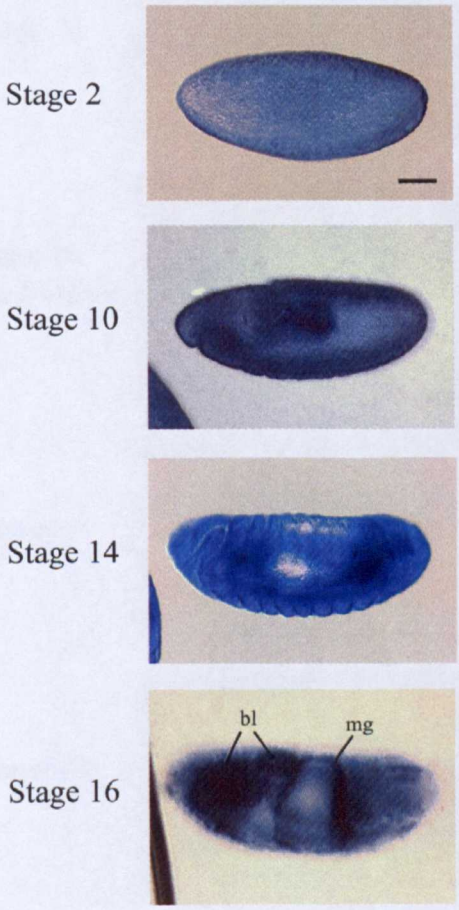


Figure 6.6: **Expression of *prel* during embryonic development**. Expression of *prel* mRNA was detected by *in situ* hybridisation using a DIG-labelled RNA probe. *prel* is expressed ubiquitously throughout embryonic development. At stage 16 expression is seen to be upregulated in specific regions of the midgut (**mg**), and in the developing brain lobes (**bl**) (although not in the ventral nerve cord). A sense control probe detects no staining at any stage (not shown). All embryos are orientated with anterior to the left and lateral view. Scale bar is 100μm.

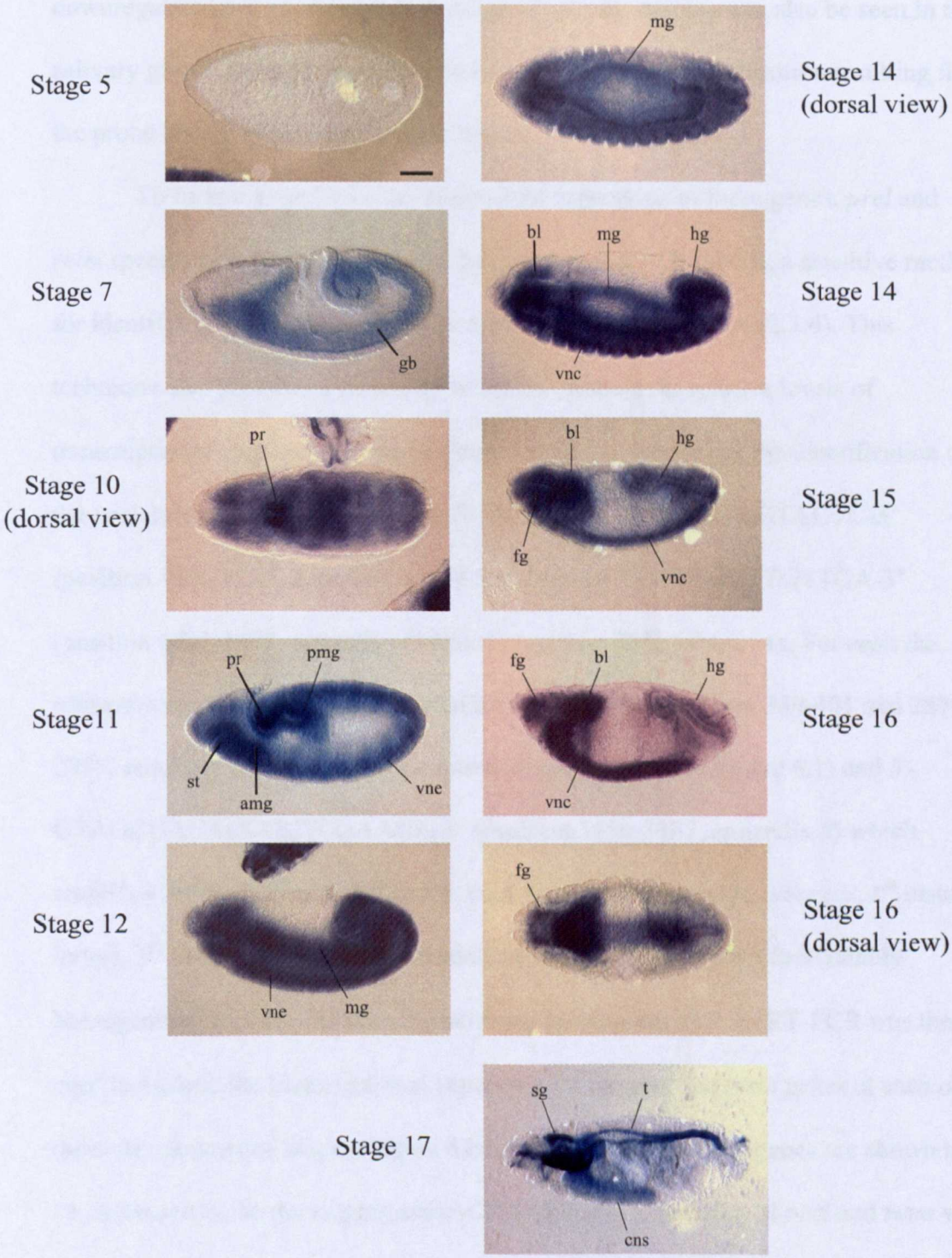


Figure 6.7: **Expression of *retm* during embryonic development**. Expression of *retm* mRNA was detected using a DIG-labelled riboprobe. *retm* is expressed in the developing gut and central nervous system. (**amg**) anterior midgut; (**bl**) brain lobes; (**cns**) central nervous system; (**fg**) foregut; (**gb**) germ band; (**hg**) hindgut; (**mg**) midgut; (**pmg**) posterior midgut; (**pr**) proctodeum; (**sg**) salivary glands; (**st**) stomodeum; (**t**) trachea; (**vnc**) ventral nerve cord; (**vne**) ventral neuroectoderm. A sense control probe detects no staining at any stage except stage 17, when non-specific staining is seen in the salivary glands and trachea, but not the CNS. All embryos are orientated with anterior to the left and lateral view unless otherwise stated. Scale bar is 100μm.

downregulated. Expression continues in the CNS late into stage 17, but is also downregulated prior to hatching. At stage 17, strong staining can also be seen in the salivary glands and trachea, a common artefact of *in situ* hybridisation resulting from the probe becoming trapped in these tissues.

To further explore the developmental expression of these genes, *prel* and *retm* specific primers were designed for the purposes of RT-PCR, a sensitive method for identifying the presence or absence of a particular transcript (2.3.4). This technique also provides a means by which to estimate the relative levels of transcriptional expression at various developmental stages. For the identification of the *prel* transcript, the primers were 5'-GCAGAAGGATGGTAGTACTC-3' (position 1206-1225, appendix 4) and 5'-GAGAGCTCCTTGGTTGTTGA-3' (position 1491-1472, appendix 4) which amplify a 285bp fragment. For *retm* the primers were 5'-CTGGTGATGAAGGCCTATGA-3' (positions 389-401 and 2879-2885, sequence is split by the first intron in appendix 5, see Figure 6.1) and 5'-GTAGCGACACCGCTCGAAGA-3' (position 3156-3137, appendix 5) which amplify a 290bp fragment. Using the wild-type strain Canton S, embryos, 1st instar larvae, 3rd instar larvae, pupae and adult males and females were individually homogenised and total RNA extracted from the samples (2.3.3). RT-PCR was then used to analyse the transcriptional expression of the *prel* and *retm* genes at each of these developmental stages (Figure 6.8A,B). As expected, both genes are shown to be expressed in the developing embryo. In addition, expression of *prel* and *retm* was detected during all the subsequent developmental stages from 1st instar larvae through to adulthood.

For estimation of relative levels of expression between stages, primers specific to the ubiquitously expressed housekeeping gene *RpL32* (2.3.4.2), which

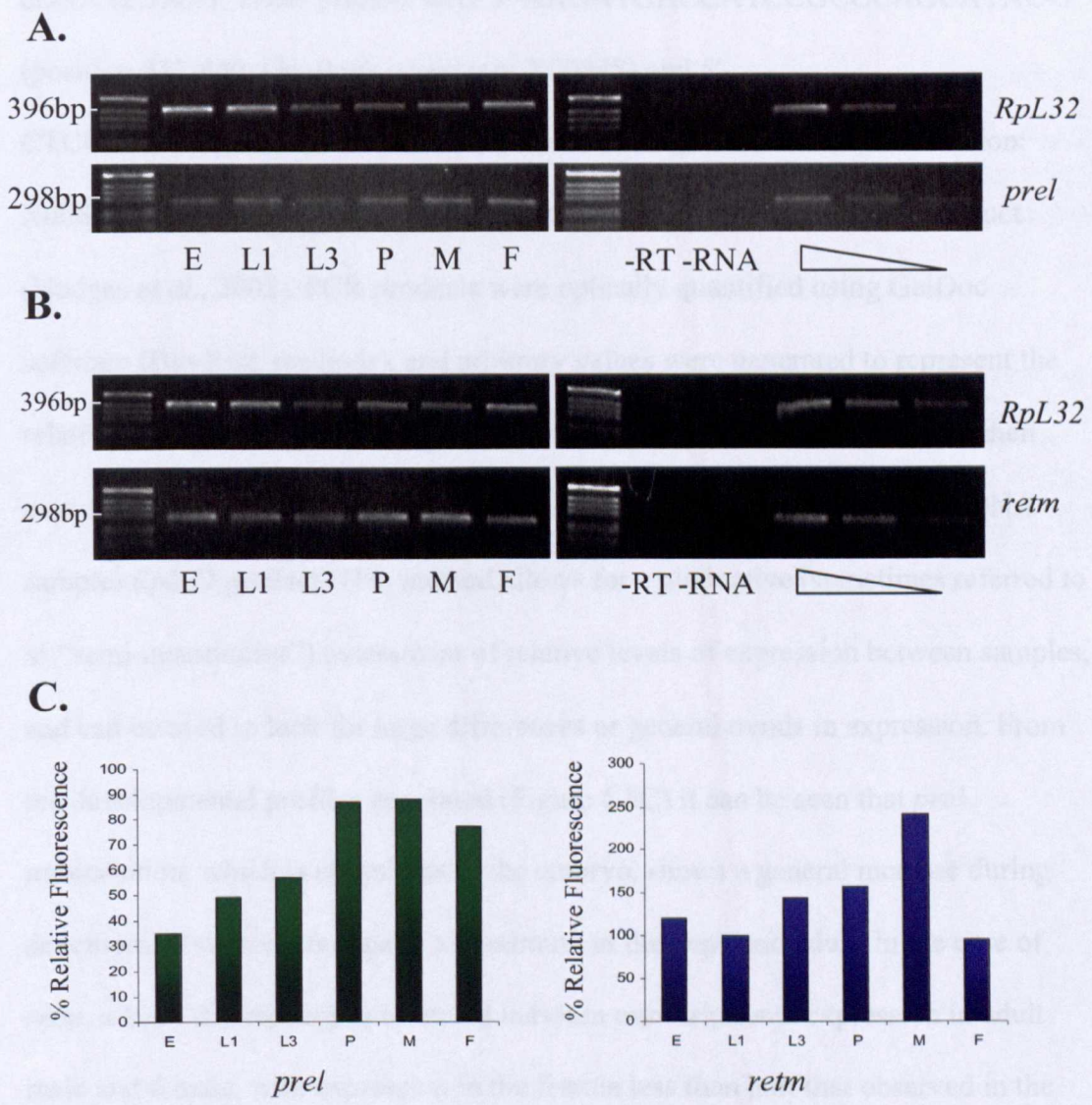


Figure 6.8: **Developmental expression of *prel* and *retm***. Semi-quantitative RT-PCR was performed on total RNA extracted from embryos (**E**), 1st instar larvae (**L1**), 3rd instar larvae (**L3**), pupae (**P**), adult males (**M**) and adult females (**F**). The housekeeping gene *RpL32* was used as a control recognised as a 456bp band. Each PCR set includes 2 negative controls to show there is no contamination with exogenous DNA (-RT and -RNA), and 3 linearity controls, in which the cDNA template is diluted 1/2, 1/4 and 1/8 (labelled by **triangle**), to demonstrate PCRs were in exponential phase. RT-PCR using primers specific to *prel* (**A**), *retm* (**B**) demonstrate expression of both these genes at all developmental stages. **C**. Developmental expression profiles of *prel* and *retm*. All PCR products were optically quantified using GelDoc software (Bio-Rad), and the arbitrary "relative fluorescence" values generated are expressed as a percentage of the corresponding *RpL32* product.

encodes the L32 ribosomal protein (O'Connell and Rosbash, 1984), were used for control PCR reactions to allow for small variations in the initial quantity of template cDNA (2.3.4.3). These primers were 5'-AAGATGACCATCCGCCCAGCATAC-3' (position 417-440, GenBank accession: X00848) and 5'-CTCGTTCTTCTTGAGACGCAGG-3' (position 876-855, GenBank accession: X00848) and had been shown previously to recognise a specific 459bp product (Hodges *et al.*, 2002). PCR products were optically quantified using GelDoc software (Bio-Rad, methods), and arbitrary values were generated to represent the relative fluorescence of each band. These relative fluorescence values were then expressed as a percentage of the corresponding (amplified from the same cDNA sample) *RpL32* product. This method allows for a qualitative (sometimes referred to as "semi-quantitative") assessment of relative levels of expression between samples, and can be used to look for large differences or general trends in expression. From the developmental profiles generated (Figure 6.8C) it can be seen that *prel* transcription, which is ubiquitous in the embryo, shows a general increase during development with expression at a maximum in the pupa and adult. In the case of *retm*, a large discrepancy is observed between transcriptional expression in adult male and female, with expression in the female less than half that observed in the male. This might suggest a male specific function for *retm*.

6.2.3 Characterisation of the P-element insertion lines l(2)k12402 and KG05639

Analysis of the expression pattern of a gene provides valuable insight into its function. The ubiquitous expression suggests a general requirement for *prel* in all cell types, whilst the more restricted distribution of the *retm* transcript would appear to indicate a more specific role in the development of the gut and central nervous

system. An essential step in establishing the functional requirement of any gene is an analysis of a loss of function mutant. In *Drosophila*, many large scale screens have been carried out to search for mutations disrupting gene function, notably by the BDGP gene disruption project (Spradling *et al.*, 1999; Spradling *et al.*, 1995) which has employed P-element mutagenesis to disrupt large numbers of genes. As a result, fly strains carrying single P-element insertions are readily available for many of the genes identified by the genome project. This is fortunately true for both *prel* and *retm*.

The P-inserts designated l(2)k12402 and KG05639 were reported by the database (http://fly.ebi.ac.uk) to map to the *prel* and *retm* genes respectively. These alleles, herein referred to as $prel^{l(2)k12402}$ and $retm^{KG05639}$, were obtained as independent strains from the Bloomington stock centre. Flanking sequence recovered by inverse PCR is available for both P-elements [GenBank accession numbers AQ026044 (l(2)k12402) and BH759618 (KG05639)] and was used to map the precise position of both insertions (Figure 6.9). In the case of both $prel^{l(2)k12402}$ and $retm^{KG05639}$, the P-elements are inserted within the transcribed region of the gene. The l(2)k12402 P-element has inserted in the *prel* 5' UTR, 266bp upstream of the translational start codon (Figure 6.9A). In addition, the $prel^{l(2)k12402}$ allele, which was genetically maintained over a *Cy*O balancer, was found to be homozygous lethal, suggesting that the insertion has disrupted the function of this gene. Prior to any analysis of the mutant line, it was important to first prove that the l(2)k12402 insertion is indeed the cause of the lethality. For this reason the line was subjected to an excision screen. The chromosome carrying the l(2)k12402 insertion was brought together with a transposase source located on the third chromosome (Figure 6.10). Fifty-seven crosses were set up and produced seven independent excision events.

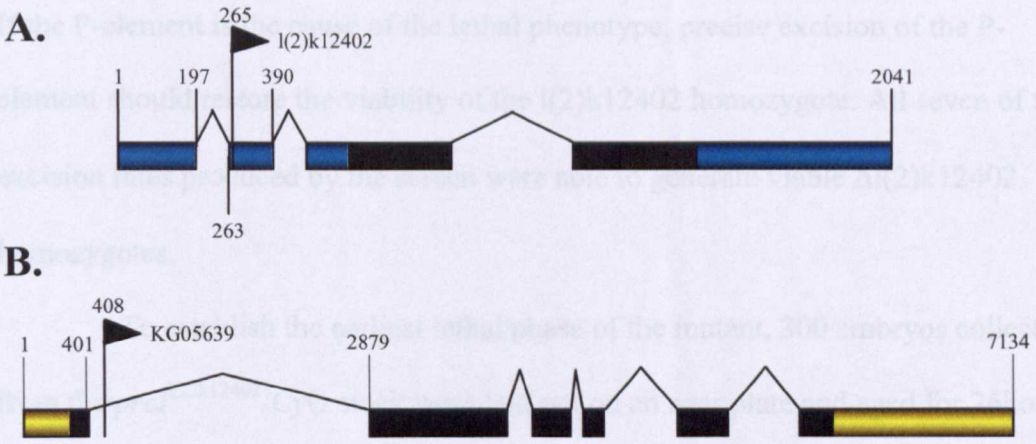


Figure 6.9: **Position of the l(2)k12402 and KG05639 P-elements**. **A**. The l(2)k12402 insertion (**black triangle**) is located within the second exon of the *prel* gene, 266bps upstream of the translational start codon. **B**. The KG05639 insertion is located within the first intron of the *retm* gene, 7bps from the end of the first exon. Coding sequences are represented by **black blocks**. All numbers refer to the sequences in appendix 4 (*prel*) or appendix 5 (*retm*). Schematic is for reference and is not to scale.

The system of nomenclature used for excisions is based simply on the vial number the excision line was produced from and the sex of the original fly. For example, line Δ5F1 represents the first excision event produced as a female from vial number 5. The excision lines were Δ5F1, Δ9M1, Δ13M1, Δ14F2, Δ34F1, Δ46M1 and Δ52M1. If the P-element is the cause of the lethal phenotype, precise excision of the P-element should restore the viability of the l(2)k12402 homozygote. All seven of the excision lines produced by the screen were able to generate viable Δl(2)k12402 homozygotes.

To establish the earliest lethal phase of the mutant, 300 embryos collected from the *prel*$^{l(2)k12402}$/*Cy*O stock were laid out on an agar plate and aged for 26hours. Of these, 97% hatched indicating that the *prel*$^{l(2)k12402}$ homozygotes do not die during embryonic development. To allow the homozygotes to be selected from heterozygous siblings, the *prel*$^{l(2)k12402}$ allele was balanced over the *Cy*O [*kruppel*GAL4 UASGFP] balancer chromosome, which causes bright fluorescence of the larval Bolvic's organs (Casso *et al.*, 2000). Fifty homozygous 1st instar larvae were selected and aged on apple agar medium (2.1.1.2). By 26hrs post-hatching, all of the larvae progressed to the 2nd instar stage, as identified by the morphology of the mouthparts (Ashburner, 1989), and dug downwards into the food. None of the mutant larvae ever progressed to the wandering 3rd instar or pupal stages, suggesting that lethality occurs during the 2nd instar or early 3rd instar stages. Upon closer observation, it was clear that *prel*$^{l(2)k12402}$ 1st instars do not exhibit an impaired locomotor phenotype similar to that described in *slmo* mutants (see section 3.2.1). In addition, mutant larvae are able to grow and feed, although growth appears to have slowed in the mutant (Figure 6.11).

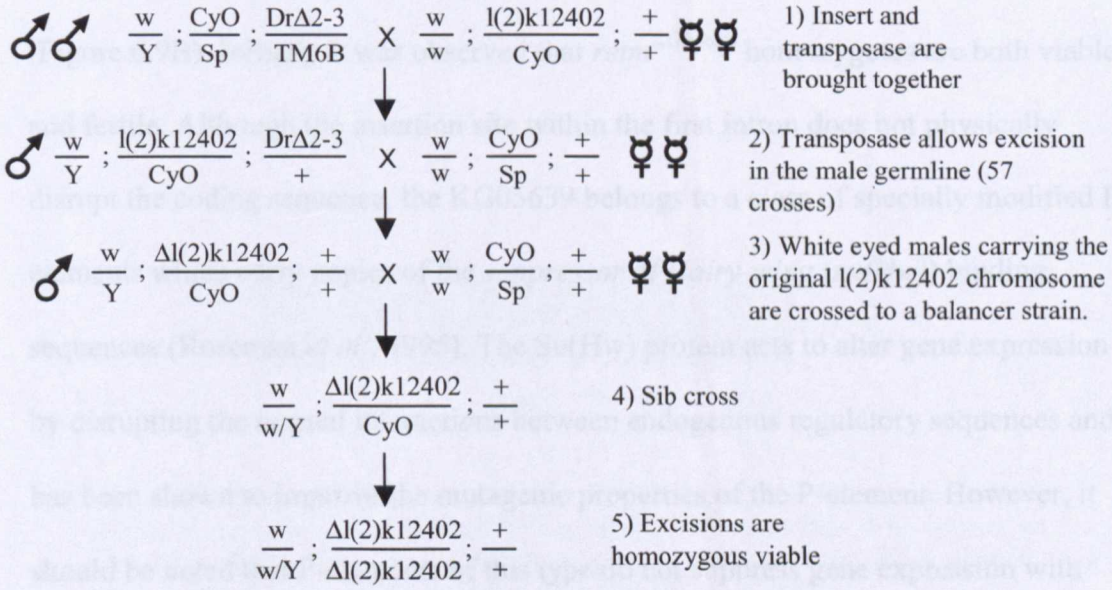


Figure 6.10: **Crossing scheme for l(2)k12402 excision screen**. Figure shows the excision of the lethal P-element insertion l(2)k12402 in the presence of a transposase source (DrΔ2-3) located on the third chromosome. The generation of 7 independent, homozygous viable excision lines is shown in step 5, and demonstrates the P-element is the cause of the lethality.

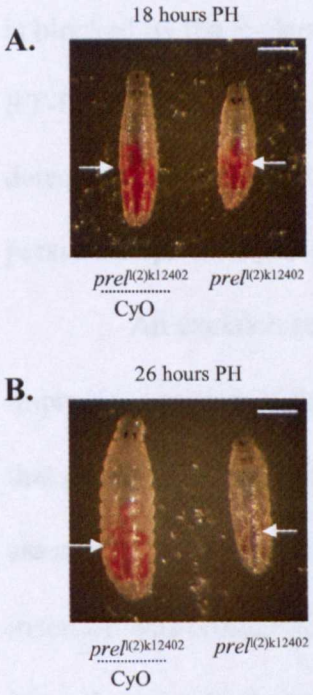


Figure 6.11: **Growth of *prel*$^{l(2)k12402}$ homozygous larvae**. Newly hatched *prel*$^{l(2)k12402}$ and *prel*$^{l(2)k12402}$/CyO larvae were fed coloured food and aged for 18 and 26 hour time periods. **A**. By 18 hours post-hatching (PH), shortly before the transition to the 2nd instar stage, 1st instar larvae have undergone a substantial period of feeding and growth during which they more than double in size. At this time mutant homozygotes exhibit a slight reduction in size when compared to the control heterozygous siblings, suggestive of slowed rate of growth. **B**. By 26 hours PH, mutant and control 2nd instars exhibit a marked difference in size. At both 18 and 26hours, the red-dyed food is visible in the gut (**arrows**) indicating that mutant larvae are able to feed. Scale bars are equal to 0.5mm.

The KG05639 P-element has also inserted into the transcribed region of the *retm* gene, although in this case the insertion site is located within the first intron (Figure 6.9B). Initially it was observed that *retm*^{KG05639} homozygotes are both viable and fertile. Although the insertion site within the first intron does not physically disrupt the coding sequence, the KG05639 belongs to a class of specially modified P-elements which carry copies of the *suppressor of Hairy-wing* (*su(Hw)*) binding sequences (Roseman *et al.*, 1995). The Su(Hw) protein acts to alter gene expression by disrupting the normal interactions between endogenous regulatory sequences and has been shown to improve the mutagenic properties of the P-element. However, it should be noted that P-elements of this type do not suppress gene expression with 100% efficiency, a notion supported by previous experience in our lab (Liu, 2002). In addition, it is well established that P-elements located in intronic regions are often spliced out of the resulting transcript and therefore do not fully disrupt expression of the protein (Horowitz and Berg, 1995). To establish whether expression of the gene is blocked by the P-element, *retm*^{KG05639} homozygotes were subjected to analysis by RT-PCR using *retm* specific primers (Figure 6.12, see 2.3.4). A weak band was detected at the expected size of 290bps, suggesting that transcription of the *retm* gene persists despite the presence of the P-element.

An excision screen was carried out on the *retm*^{KG05639} line in the hope that imprecise excision of the P-element would generate deletions of the flanking DNA that physically disrupt the *retm* coding sequence. Excisions were generated essentially as described in Figure 6.10. The chromosome carrying the KG05639 insertion was brought together with a transposase source (in this case SbΔ2-3) located on the third chromosome. A total of 200 crosses were set up and 95 independent excision lines were produced (Table 6.1). 13 excision lines were found

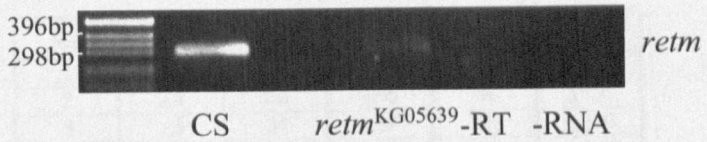


Figure 6.12: Expression of *retm* in the *retm*^{KG05639} mutant. RT-PCR was performed on total RNA extracted from adult *retm*^{KG05639} homozygotes. A weak band is detected at the expected size of 290bps, indicating that the retm gene is still transcribed in the presence of the P-insert. Canton S (CS) flies were used as a positive control.

Table 6.1: List of KG05639 excision strains.

A. The viable excisions

No.	Strain	No.	Strain	No.	Strain
1	Δ1	29	Δ64	57	Δ136
2	Δ4	30	Δ65	58	Δ138
3	Δ5	31	Δ67	59	Δ140
4	Δ6	32	Δ69	60	Δ145
5	Δ7	33	Δ70	61	Δ146
6	Δ8	34	Δ72	62	Δ149
7	Δ9	35	Δ75	63	Δ150
8	Δ12	36	Δ78	64	Δ151
9	Δ15	37	Δ87	65	Δ152
10	Δ16	38	Δ89	66	Δ155
11	Δ17	39	Δ92	67	Δ156
12	Δ20	40	Δ100	68	Δ158
13	Δ22	41	Δ105	69	Δ160
14	Δ23	42	Δ107	70	Δ163
15	Δ28	43	Δ110	71	Δ165
16	Δ33	44	Δ111	72	Δ168
17	Δ41	45	Δ114	73	Δ170
18	Δ43	46	Δ116	74	Δ171
19	Δ45	47	Δ117	75	Δ176
20	Δ50	48	Δ120	76	Δ177
21	Δ52	49	Δ122	77	Δ180
22	Δ53	50	Δ124	78	Δ183
23	Δ55	51	Δ125	79	Δ186
24	Δ56	52	Δ127	80	Δ187
25	Δ58	53	Δ129	81	Δ189
26	Δ59	54	Δ131	82	Δ196
27	Δ61	55	Δ134		
28	Δ63	56	Δ139		

B. The lethal excisions

No.	Strain	No.	Strain
1	Δ10	8	Δ143
2	Δ57	9	Δ164
3	Δ66	10	Δ175
4	Δ81	11	Δ191
5	Δ83	12	Δ194
6	Δ84	13	Δ197
7	Δ119		

to be lethal, although these were later found to genetically compliment each other, indicating that the lethality of these strains resulted from independent events unrelated to the *retm* locus. To screen the remaining 82 lines for deletion of the *retm* coding sequence, PCR primers were designed to amplify fragments from exon 1 and exon 2 (Figure 6.13A) and reactions performed on genomic DNA extracted from single adult flies (2.2.4). For exon 1 the primers were 5'-ACGCTAGGAGCCAACGGATA-3' and 5'-CTTCATCACCAGCTCAAAAG-3'. For exon 2 the primers were 5'-AGGCCTATGAGCGGCGGTTT-3' and 5'-ACATGCGTGATGCCTTCCTC-3'. A single excision line, designated $retm^{\Delta 69}$ was found to carry a deletion which includes part of exon 2 and disrupts the *retm* coding sequence (Figure 6.13C). In addition, primers were designed to amplify fragments from the predicted genes CG16947 and CG9527 that directly flank *retm* on chromosome 2 to demonstrate that the deletion did not represent a large deficiency removing multiple genes. For CG16947 the primers were 5'-ATGAAGAACCGCCAGCGAAC-3' (position 1-20, FlyBase gene region: FBgn0031816) and 5'-GTGCGCTGGATGATACTTGC-3' (position 284-303, FlyBase gene region: FBgn0031816), whilst for CG9527 the primers were 5'-TCGCGCCTAGCATTCGAAGC-3' (position 4-23, FlyBase gene region: FBgn0031813) and 5'-AAGGTCCGCTACGGAAGTCG-3' (position 308-327, FlyBase gene region: FBgn0031813). The gene was named *real-time* due to the initial observation that mutation does not result in a larval slow motion phenotype comparable to that seen in *slmo* mutants. The fact that $retm^{\Delta 69}$ homozygotes survive to adulthood suggests that *retm* is likely not an essential gene. Loss of *retm* function may result in more subtle defects which have yet to be determined, although preliminary analysis suggests that adult $retm^{\Delta 69}$ homozygotes perform normally in

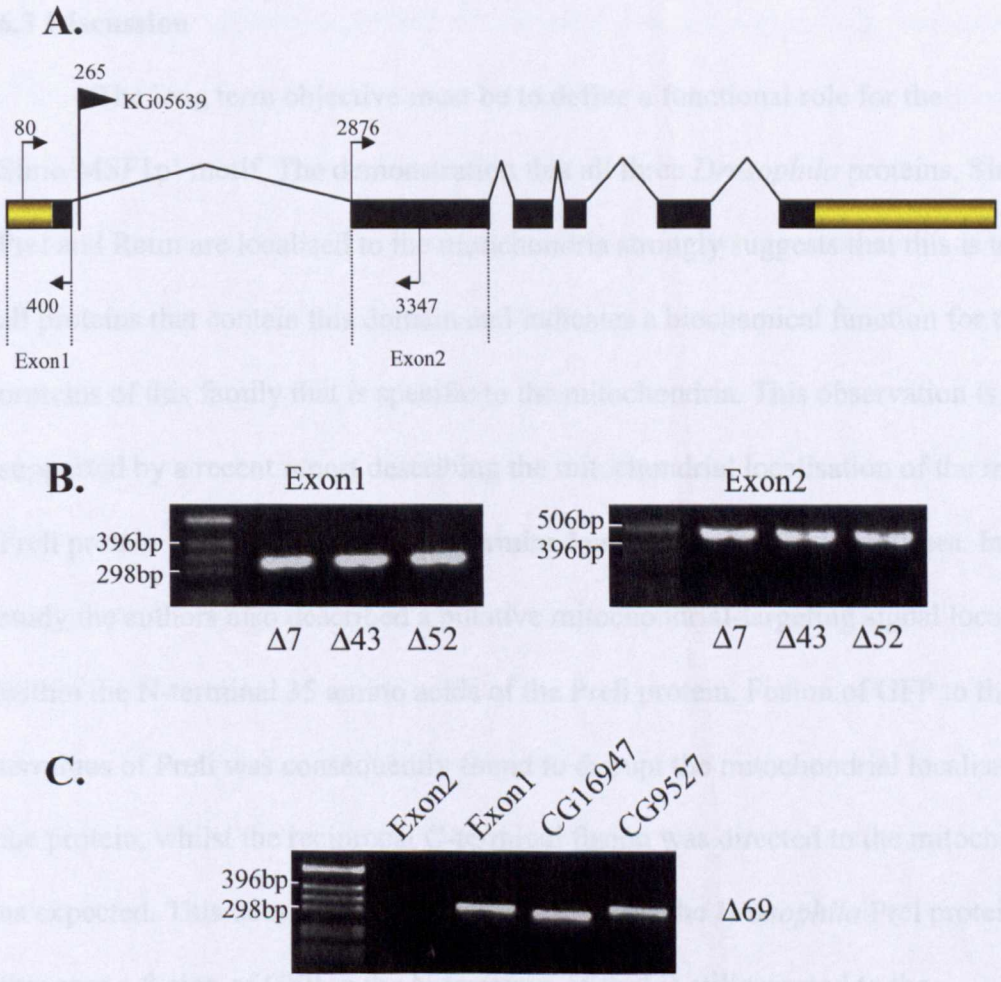


Figure 6.13: **Identification of a *retm* deletion allele**. **A**. Schematic to show the relative positions of PCR primers designed to amplify fragments from exon1 and exon2. The original position of the KG05639 P-element is indicated (**black triangle**). The protein coding sequence is indicated by **black blocks**. All numbers refer to the *retm* gene region sequence obtained from FlyBase (appendix 5) Schematic is for reference and is not to scale. **B**. The primers recognise bands of 320bps (exon1) and 471bps (exon2) in the putative precise excision lines Δ7, Δ43 and Δ52. **C**. In the excision line Δ69, no PCR product is recognised for exon 2 indicating that part of this exon has been deleted by imprecise excision of the P-element. The consequent disruption of the coding sequence means that the Δ69 excision represents an additional *retm* mutant allele. Exon 1 and the neighbouring genes CG16947 and CG9527 are still intact.

assays of locomotor activity, and the ability to right themselves from an inverted position (appendix 6).

6.3 Discussion

The long term objective must be to define a functional role for the Slmo/MSF1p' motif. The demonstration that all three *Drosophila* proteins, Slmo, Prel and Retm are localised to the mitochondria strongly suggests that this is true of all proteins that contain this domain and indicates a biochemical function for the proteins of this family that is specific to the mitochondria. This observation is now supported by a recent report describing the mitochondrial localisation of the mouse Preli protein (Fox *et al.*, 2004) as determined using a Preli specific antisera. In this study the authors also described a putative mitochondrial-targeting signal located within the N-terminal 35 amino acids of the Preli protein. Fusion of GFP to the N-terminus of Preli was consequently found to disrupt the mitochondrial localisation of the protein, whilst the reciprocal C-terminal fusion was directed to the mitochondria as expected. This does not appear to be the case for the *Drosophila* Prel protein. In this case a fusion of GFP to the N-terminus of Prel is still targeted to the mitochondria, whilst a fusion to the C-terminus is found to be diffuse within the cytoplasm (Figure 6.4). The various targeting signals that are used to direct newly synthesised proteins to the mitochondria are diverse in nature, and these can be located at the N- or C-terminus, or distributed internally within the protein (Truscott *et al.*, 2003). These observations would suggest that the signals used to target the various proteins of the Slmo/MSF1p' family to the mitochondria are not universal. Certainly, the N-terminal 35 amino acid region of the mammalian Preli proteins does not appear to be particularly well conserved in the fly (see Figure 5.1). This notion is

also supported by the finding that fusions to either terminus of the Slmo protein are successfully directed to the mitochondria (see chapter 4) and that Retm is directed to the mitochondria only by a C-terminal fusion (Figure 6.5).

The *prel* gene is expressed ubiquitously during embryonic development (Figure 6.6) and the transcript is present during all later developmental stages (Figure 6.8A), suggestive of a general requirement for the Prel protein in the mitochondria of all cells. This is in contrast to the restricted expression observed with both *slmo* and *retm*, but consistent with the reported expression of the human *PRELI* gene in a wide variety of tissues including brain, liver, lung, spleen and kidney (Guzman-Rojas *et al.*, 2000). Certainly, the lethality resulting from *prel* loss of function indicates that the function of this gene is essential. The survival of *prel* mutants to larval stages may simply be the result of maternal contribution. The PRELI subgroup of Slmo/MSF1p' proteins may ultimately come to be regarded as the archetypal members of the family, fulfilling an as yet to be defined, but generally required biochemical function within the mitochondria, whilst those proteins more closely related to Slmo or Retm have diverged to fulfil more specialised roles in specific cell types.

Whilst the ubiquitous transcription of *prel* would appear to indicate a general requirement for the gene product, the spatially dynamic expression of *retm* appears to suggest a specific developmental function. Expression of *retm* is limited to the developing gut and CNS, but is also temporally restricted to developmentally crucial stages for these organs (Figure 6.7). For example, strong expression in the primordial midgut is rapidly down regulated prior to stage 15, by when the presumptive midgut has closed both ventrally and dorsally and is largely established. Expression persists in the hindgut as it continues to develop by projecting towards the antero-dorsal

region of the embryo, and is itself only down regulated during stage 16. Likewise, the *retm* transcript is detected strongly in the CNS at stages 15 and 16, during periods of major developmental reorganisation (Goodman, 1993), but is clearly reduced by the end of stage 17 immediately prior to hatching.

Like both *prel* and *slmo* (Reeve, 2002), *retm* expression persists through later stages of development through to adult. Despite this, the viability of flies homozygous for a *retm* deletion allele, $retm^{\Delta 69}$, indicates that this gene does not perform an essential function in *Drosophila*. The fact that *retm* is conserved between flies and mammals (see chapter 5) might suggest that loss of function has more subtle consequences that confer a selective disadvantage, or that the Retm subclass of CRAL-TRIO domain proteins play important roles in other organisms.

Chapter 7: Yeast-2-Hybrid Screen

Chapter 7: Yeast-2-Hybrid Screen

7.1 Introduction

Specific interactions between proteins form the basis of many essential biological processes. Consequently, considerable effort has been made to identify those proteins that bind to proteins of interest. It also follows that a major step in elucidating the function of an uncharacterised protein is to understand the interactions it makes with previously identified proteins. The advent of large scale sequencing projects has identified large numbers of novel proteins. For many of these new proteins, like those of the Slmo family, the sequence alone sheds little or no light on their function.

The yeast-2-hybrid approach provides a sensitive and versatile method for identifying proteins that interact with a protein of interest, and can be applied to most, if not all, proteins for which the coding sequence is known. The system takes advantage of the fact that many eukaryotic transcription factors are modular, with distinct and separable functional domains directing the binding to specific DNA sequences and the activation of transcription (Brent and Ptashne, 1985). Neither is capable of activating transcription alone, but the DNA binding and activation domains of different transcription factors can often be interchanged to form functional hybrid proteins. For example, when the transcriptional activation domain of GAL4 is combined with the DNA binding domain of the bacterial repressor LexA, the resulting chimera is a functional transcriptional activator in yeast (Brent and Ptashne, 1985). Importantly, the DNA binding and activation domains do not have to be covalently bonded to each other in order for transcription to be activated. Fields and Song (1989) first demonstrated the usefulness of this property by showing that

the activation of transcription in yeast could be used to assay the interaction between two proteins if one of them was fused to a DNA-binding domain and the other to an activation domain. When both fusions are expressed together in yeast, a positive interaction results in the reconstitution of a functional transcription factor. This results in the expression of one or more reporter genes, which ultimately generate a phenotypic signal (Figure 7.1). In this way, a protein of interest fused to a DNA-binding domain can be used to hunt for possible interactions by screening large numbers of activation domain fusion proteins expressed from a cDNA library (Chien *et al.*, 1991). Selection of yeast which are positive in the 2-hybrid assay therefore provides a means by which to isolate and identify those interactors by rescue of the corresponding cDNAs.

Based on this concept, a number of yeast-2-hybrid systems have been developed independently, all of which utilise the DNA-binding domains of either the yeast GAL4 or the *E. coli* protein LexA (Toby and Golemis, 2001). This approach has been used widely, and has often proven successful in ascribing functions to previously uncharacterised molecules (Vidal and Legrain, 1999). Indeed, the 2-hybrid system has much wider applications than this and has been further developed for the proteomic dissection of entire protein networks (Giot *et al.*, 2003). This chapter describes the use of a LexA based 2-hybrid system, the interaction trap, to look for proteins that interact with the Slmo protein with the purpose of gaining some insight into the biochemical role of this molecule.

7.2 Results

A major factor in the success of a yeast-2-hybrid screen is the availability of a high quality cDNA library. Given what we knew of the expression and function of

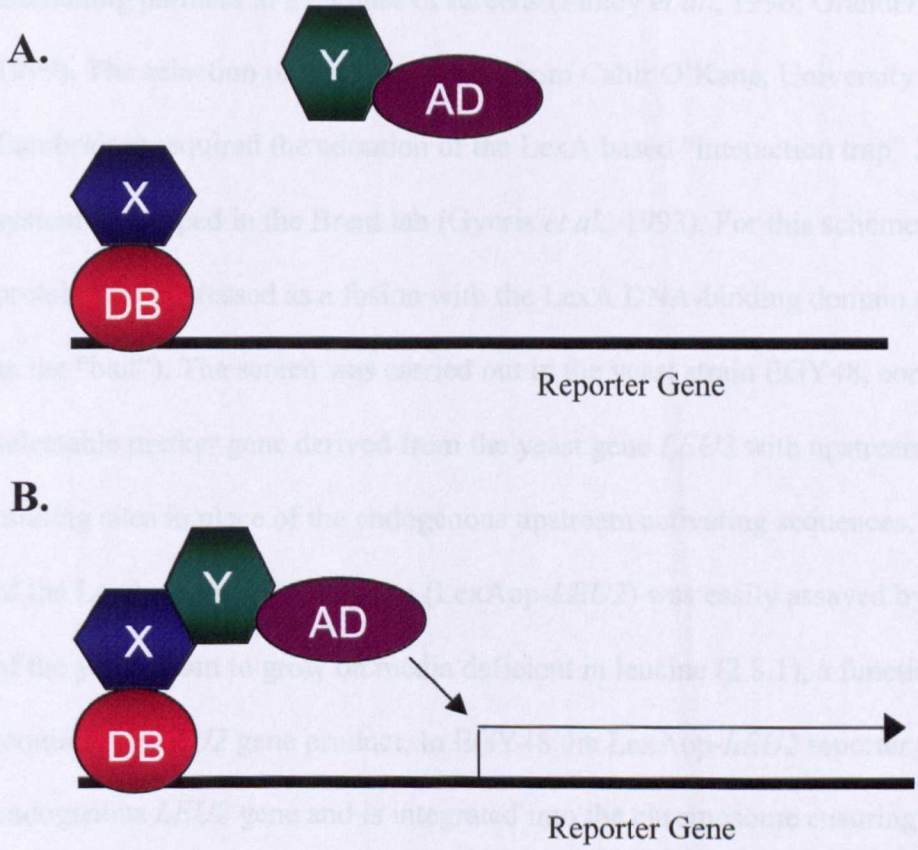


Figure 7.1: Schematic representation of the yeast-2-hybrid system assay. An interaction between two proteins is detected by the activation of one or more reporter genes. **A**. Protein of interest **X** is fused to a DNA-binding domain (**DB**) that binds specifically to sequences in the promoter of the reporter gene. Protein **Y** is fused to a transcriptional activation domain (**AD**) that recruits the transcription machinery. **B**. Expression of the reporter gene will only occur if proteins **X** and **Y** interact.

Slmo, either an embryonic or ovarian library was acceptable for the purposes of this screen. The *Drosophila* ovary library RFLY3 was originally manufactured in Roger Brent's lab (Finley *et al.*, 1996) and had been used successfully to identify interacting partners in a number of screens (Finley *et al.*, 1996; Granderath *et al.*, 1999). The selection of this library (gift from Cahir O'Kane, University of Cambridge) required the adoption of the LexA based "interaction trap" 2-hybrid system developed in the Brent lab (Gyuris *et al.*, 1993). For this scheme, the Slmo protein was expressed as a fusion with the LexA DNA-binding domain (referred to as the "bait"). The screen was carried out in the yeast strain EGY48, containing a selectable marker gene derived from the yeast gene *LEU2* with upstream LexA binding sites in place of the endogenous upstream activating sequences. Expression of the LexA operator *LEU2* gene (LexAop-*LEU2*) was easily assayed by the ability of the yeast strain to grow on media deficient in leucine (2.8.1), a function which requires the *LEU2* gene product. In EGY48 the LexAop-*LEU2* reporter replaces the endogenous *LEU2* gene and is integrated into the chromosome ensuring it is genetically stable. The system also incorporates a second, less sensitive LexA activated reporter based on the *lac*Z gene. This provides a secondary assay of activation that will eliminate many of the false positives that arise due to *cis* and *trans*-acting mutations that cause activation of the *LEU2* reporter. To prepare the selection strain for the interaction trap, EGY48 was transformed (2.8.2) with the *lac*Z reporter plasmid pSH18-34, which carries the *URA3* selectable marker that permits growth on media lacking uracil (2.8.1). The library cDNAs are contained in the "prey" vector pJG4-5, which directs the expression of the cDNA-encoded protein fused at the N-terminal to a moiety containing the SV40 nuclear localisation signal,

and the transcriptional activation domain B42. The plasmid also carries the selectable marker *TRP1*, which permits growth in the absence of tryptophan.

7.2.1 Construction of LexA-*slmo* bait plasmid

The purpose of using the interaction trap is to identify any protein that interacts with any part of Slmo. In addition, the lack of any characterised domains within the predicted Slmo sequence provided no basis for the exclusion or inclusion of any region of the protein. Therefore, for the purposes of this investigation, the entire Slmo coding sequence was included in the bait construct. The bait vector, pEG202, contains the entire coding sequence of the *E. coli* LexA protein (including the DNA binding domain and the dimerisation domain), and will maintain strong and constitutive expression of the bait as regulated by the yeast alcohol dehydrogenase 1 (*ADH1*) promoter. The LexA coding sequence is followed by a polylinker, which facilitates cloning of the coding sequence of interest to generate in frame fusions to LexA. In addition, the plasmid contains the *HIS3* selectable marker, which allows yeast transformants to grow on media lacking histidine (2.8.1), and the *E. coli* ampicillin resistance gene (*amp*) for selection of the plasmid during cloning. The LexA protein does not contain a yeast nuclear localisation sequence, but almost all fusion proteins become equally partitioned between the cytoplasm and the nucleus and will therefore bind LexA operators in the yeast nucleus (Golemis and Brent, 1992; Toby and Golemis, 2001).

For the construction of the bait plasmid, PCR primers were designed to amplify the *slmo* coding sequence adding both an upstream *Eco*RI restriction site and a downstream *Xho*I site to facilitate cloning of the fragment into the pEG202 polylinker. The primers were 5'-gcggccgcgaattc**ATG**AAAATCTGGACATCGGA-

3' and 5'-gcggccgcctcgag**CTA**CGTAATGTGCATCGC-3' (restriction sites are underlined, "start" and "stop" codons are in bold). The sequence was amplified with *Pfu* polymerase (2.2.3.3), using the *slmo* 2A7 cDNA (Carhan, 1999; see Figure 1.1) as a template for the reaction (Figure 7.2A). The PCR fragment was then cloned into the *Eco*RI and *Xho*I sites of the pEG202 vector and the presence of the insert confirmed by digestion with the corresponding enzymes (Figure 7.2B). The production of an in frame LexA-*slmo* fusion construct was confirmed by sequencing.

7.2.2 Testing the bait for autoactivation of reporters

Some bait proteins can activate transcription of the reporter genes alone. To be useful in an interactor screen, it is important that the bait fusion does not activate transcription independently. Screens involving proteins that have transcriptional activity require only portions of the protein to be cloned into the bait vector. It is important not to have a high background of false positive colonies arising from activation of the *lacZ* or LexAop-*LEU2* reporter genes by the bait alone. To test whether the bait autoactivates transcription of the reporters, an EGY48 strain containing the reporter plasmid pSH18-34 was transformed (2.8.2) with the LexA-*slmo* bait construct. In parallel with this, the reporter strain was also transformed with two *HIS3*+ control plasmids. The first of these, pSH17-4 directs the expression of a LexA-GAL4 fusion that activates transcription strongly and acts as a positive control for the activation assay. The second, pRFHM1 expresses LexA fused with a transcriptionally inert fragment of the *Drosophila* Bicoid homeoprotein, and therefore acts as a negative control. Transformed colonies were then patched onto X-gal indicator media (2.8.1) then incubated at 30°C for 3 days. As expected, the positive control yeast strain containing pSH17-4 went dark blue, indicating strong

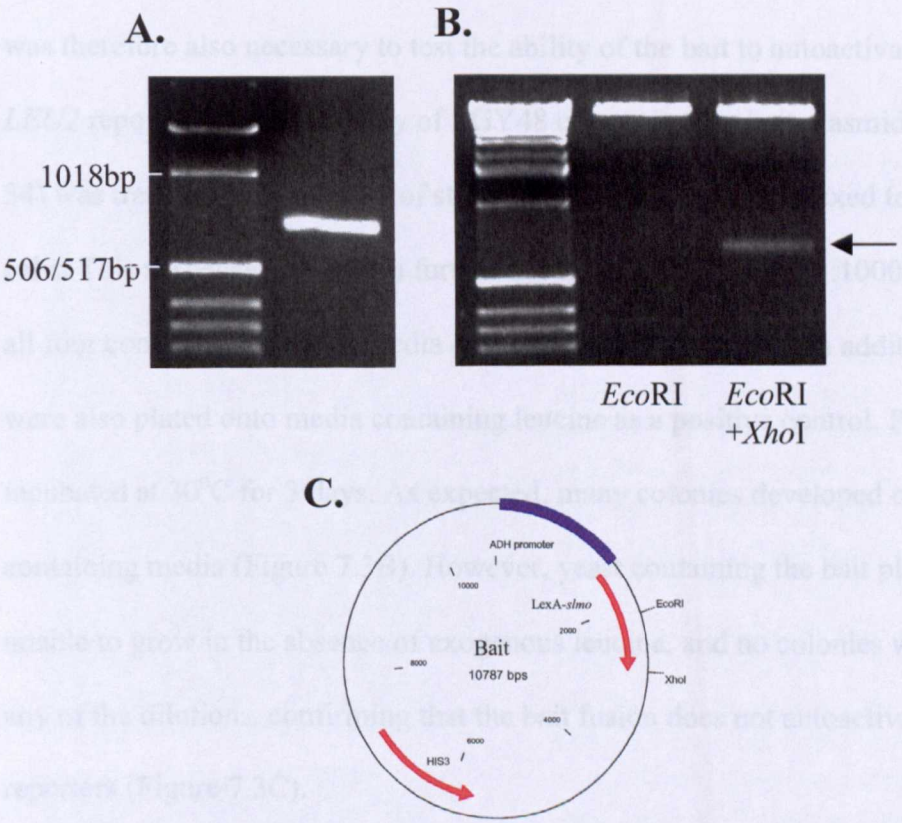


Figure 7.2: **Construction of the bait plasmid**. **A**. PCR amplification of *slmo* coding region. The product was resolved on an agarose gel, and a single band found to migrate between 1018bp and 506/517bp standard markers consistent with the expected size of 648bp. This fragment was therefore cloned into the bait vector pEG202. **B**. Gel showing diagnostic *Eco*RI single digest and *Eco*RI/ *Xho*I double digests, confirming the presence of the insert (**arrow**). **C**. Schematic plasmid map of the bait construct showing the relative positions of the LexA-*slmo* ORF, the ADH promoter and the *HIS3* gene.

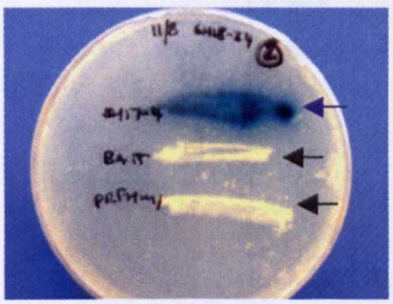
activation of the *lacZ* reporter gene. In contrast, neither the bait strain nor the negative control responded in this assay (Figure 7.3A).

For some baits, the LexAop-*LEU2* reporter of EGY48 is more sensitive than the *lacZ* reporter on pSH18-34 (DupLEX-A User Manual, OriGene Technologies). It was therefore also necessary to test the ability of the bait to autoactivate the LexAop-*LEU2* reporter. A single colony of EGY48 containing the bait plasmid (and pSH18-34) was transferred into 0.5ml of sterile distilled water and vortexed to disperse the cells. This suspension was then further diluted 1:10, 1:100 and 1:1000 before plating all four concentrations onto media deficient in leucine (2.8.1). In addition, dilutions were also plated onto media containing leucine as a positive control. Plates were then incubated at 30°C for 3 days. As expected, many colonies developed on leucine containing media (Figure 7.3B). However, yeast containing the bait plasmid were unable to grow in the absence of exogenous leucine, and no colonies were visible for any of the dilutions, confirming that the bait fusion does not autoactivate either of the reporters (Figure 7.3C).

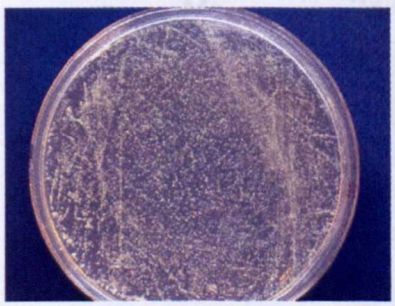
7.2.3 Test that the bait fusion is expressed in the yeast system

Before the bait construct could be employed in a large scale screen, it was important to confirm that the bait fusion protein is expressed in the yeast host. A single colony of the bait strain (EGY48 transformed with pSH18-34 and the bait construct) was used to inoculate an overnight culture. The culture was then lysed (2.4.2) and the sample run on an SDS-PAGE gel (2.4.3). As controls, yeast transformed with the pEG202 vector alone, and the pSH18-34 reporter plasmid alone were prepared in the same way, and a western blot was carried out on all three samples using an anti-LexA antibody (Figure 7.4). A band was detected at

A.



B.



C.

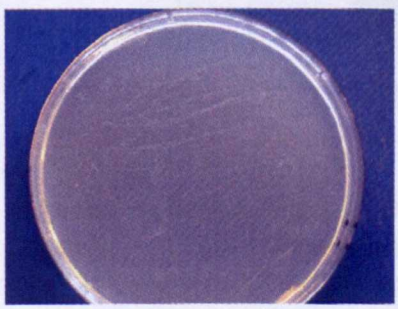


Figure 7.3: Test to determine whether the bait can autoactivate reporters. A. An EGY48 strain carrying the *lacZ* reporter plasmid pSH18-34 was transformed individually with the *slmo* bait plasmid, the positive control plasmid pSH17-4 and the negative control plasmid pRFHM1 before plating on selective media containing X-gal. The Lex-GAL4 fusion protein expressed from pSH17-4 is sufficient by itself to cause activation of the *lacZ* reporter and yeast carrying the plasmid are dark blue (**blue arrow**). Yeast carrying the bait plasmid or pRFHM1 are not blue, indicating that the LexA-Slmo fusion does not autoactivate the lacZ reporter (**black arrows**). **B,C.** The EGY48 strain containing the bait plasmid was plated as serial dilutions (only 1:1000 dilution is shown) on media containing leucine (**B**) and deficient in leucine (**C**). The bait expressing strain is unable to grow on leucine deficient media indicating that the LexA-Slmo fusion does not autoactivate the *LEU2* reporter.

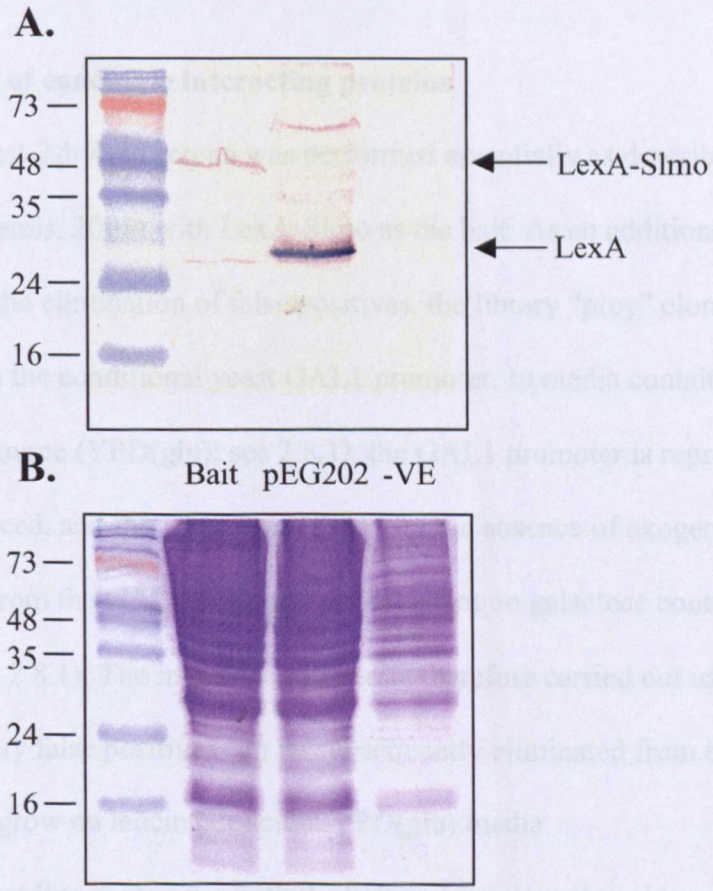


Figure 7.4: **Expression of the LexA-Slmo bait fusion in yeast**. **A**. Western blot using an anti-LexA antibody. The samples run on the gel are EGY48 transformed with the reporter plasmid (pSH18-34) and either the bait construct expressing the bait fusion (**LexA-Slmo**), the bait vector expressing LexA alone (**pEG202**), or no additional plasmid as a negative control (**-VE**). A band is detected at the expected size for a LexA-Slmo fusion protein (approximately 50kDa). This band is not found in either of the control lanes, indicating that the bait protein is expressed in the yeast system at an easily detectable level. **B**. Coomasie stained gel run in concert with **A** to show equivalence of loading.

approximately 50kDa, the size expected for a LexA-Slmo fusion protein (LexA~20kDa, Slmo runs at ~30kDa), confirming that there is a good level of expression of bait fusion. Importantly, this band was not present in the control lanes.

7.2.4 Isolation of candidate interacting proteins

The yeast-2-hybrid screen was performed essentially as described previously (Toby and Golemis, 2001) with LexA-Slmo as the bait. As an additional layer of stringency for the elimination of false positives, the library "prey" clones are expressed from the conditional yeast GAL1 promoter. In media containing glucose as the carbon source (YPD(glu); see 2.8.1), the GAL1 promoter is repressed, no prey fusion is produced, and the yeast cannot grow in the absence of exogenous leucine. Transcription from the GAL1 promoter is dependent on galactose containing media (YPD(gal); see 2.8.1). The interaction screen is therefore carried out using YPD(gal) media, and many false positives can be subsequently eliminated from the screen by their ability to grow on leucine deficient YPD(glu) media.

The yeast line containing both the bait and reporter plasmids was transformed with 200μg (2.8.3) of the RFLY3 cDNA library, plated on selective medium (YPD(glu)-ura-his-trp) and incubated at 30°C for 4 days. Many transformants were obtained (approximately 100,000), which were then harvested from the medium and 1ml aliquots stored as 50% glycerol stocks at –70°C (2.8.3). In order to ensure that all the library transformants are screened, it is necessary to screen at least 7 times the number of transformants initially obtained (DupLEX-A user manual, OriGene Technologies). Before screening the transformants for potential positive interactors, it was therefore important to first titre the number of viable cells, or colony forming units (cfu) present in each frozen aliquot of yeast transformants. A single aliquot was

thawed and diluted 1:10 in selective (YPD(glu)-ura-his-trp) medium, and incubated at 30°C for 4hrs. The transformants were then further diluted to 1:100, 1:1000 and 1:10,000 and all four dilutions plated on selective medium, and incubated for 3 days at 30°C. 100µl of the 1:10,000 dilution yielded approximately 2000 individual colonies, which is equal to $2x10^8$cfu/ml.

To screen the library for potential interacting proteins $1x10^6$cfu were plated onto selective media deficient in leucine (YPD(gal)-ura-his-trp-leu) and incubated at 30°C for 3 days. Many colonies were obtained of which 200 colonies were picked and patched onto fresh selective plates. 65 of these colonies did not grow on the selective medium a second time allowing them to be immediately eliminated as artefacts of the screen. To eliminate false positives from the remaining 135 candidates, these were re-patched onto selective X-gal indicator plates (2.8.1) to look for activation of the *lacZ* reporter. A total of 97 colonies were found to produce β-galactosidase and were blue in colour. Examples of the selection plates employed in the screen are shown in Figure 7.5. Colonies that activate the *lacZ* reporter do so to different levels, and as a result there is marked variation in the strength of blue colour exhibited by these yeast strains. In theory this is an indicator of the strength of the interaction. Candidates from the screen were therefore arbitrarily classified as weak, medium, strong or very strong activators (Figure 7.5). All candidates were then patched onto selective glucose (YPD(glu)-ura-his-trp-leu) medium to eliminate those false positives in which reporter activation is not conditional on the presence of galactose. Of these, 43 grew on the glucose medium and were eliminated as false positives, whilst 19 showed very weak growth and were counted as potential positives. The remaining 35 were completely unable to grow, indicating that activation of LexAop-*LEU2* reporter is dependent on activation of the library GAL1

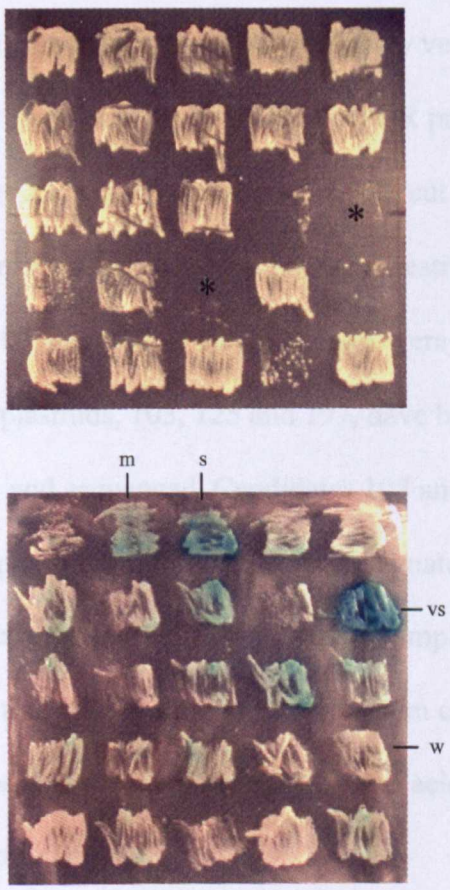


Figure 7.5: **Selection of candidate Slmo interactions**. To screen for potential interactions, EGY48 transformed with the bait plasmid (LexA-Slmo), *lac*Z reporter plasmid (pSH18-34) and RFLY3 cDNA library, was plated on galactose media deficient in leucine (2.8.1). Colonies which grew on this medium are potential positives that activate the LexAop-*LEU2* reporter. **A**. 200 potential positive colonies were selected and patched onto fresh selective medium. Shown is a single YPD(gal)ura-his-trp-leu selection plate (2.8.1) containing 25 colonies. Of the total 200 candidates, **65** were unable to grow when replated (**black asterisks**), indicating they were false positives. **B**. The remaining 135 candidates were plated on X-gal indicator plates (2.8.1) to screen for activation of the *lac*Z reporter. A single selection plate is shown containing 25 colonies. A total of **97** candidates were positive in this assay as indicated by their blue colour. Candidates were arbitrarily classified as weak (**w**), medium (**m**), strong (**s**) or very strong (**vs**) activators of the *lac*Z reporter.

promoter. The results of the yeast-2-hybrid screen are summarised in Figure 7.6 and Table 7.1.

To initially establish the nature of these candidate Slmo interactors, Kevin Moffat (University of Warwick) has rescued and sequenced several of these 35 primary candidates. 15 candidates were selected for further analysis (103, 109, 114, 122, 125, 126, 134, 139, 140, 150, 153, 154, 155, 196 and 197; see Table 7.1) Initially, primers specific to the library vector were used to amplify the relevant cDNA fragments by PCR. The 15 PCR products were then cut with the restriction enzyme *Rsa*I, a 4bp cutter that should cut often within each fragment to identify similar restriction patterns. From the restriction pattern it was determined that all but two (109 and 197) of the candidates were identical (K. Moffat, pers. comm.). Three of the plasmids, 103, 125 and 197, have been rescued from the candidate yeast strains and sequenced. Candidates 103 and 125 are both identified as cDNAs corresponding to the same gene, designated CG40218. This gene is predicted to encode a 241 amino acid protein of completely unknown function and has no characterised domains. The cDNA from candidate 197 is identified as CG6259, a gene predicted to encode a 226 amino acid protein with similarity to the yeast Snf7p protein.

7.3 Discussion

This chapter describes the application of a yeast-2-hybrid screen to search for proteins that interact with Slmo, since their identity might impart further information on the function of Slmo. Described here is the isolation of 35 candidate interactions as yeast strains positive in the yeast-2-hybrid screen. Preliminary analysis of these has identified two potential interacting partners of Slmo, encoded by the predicted

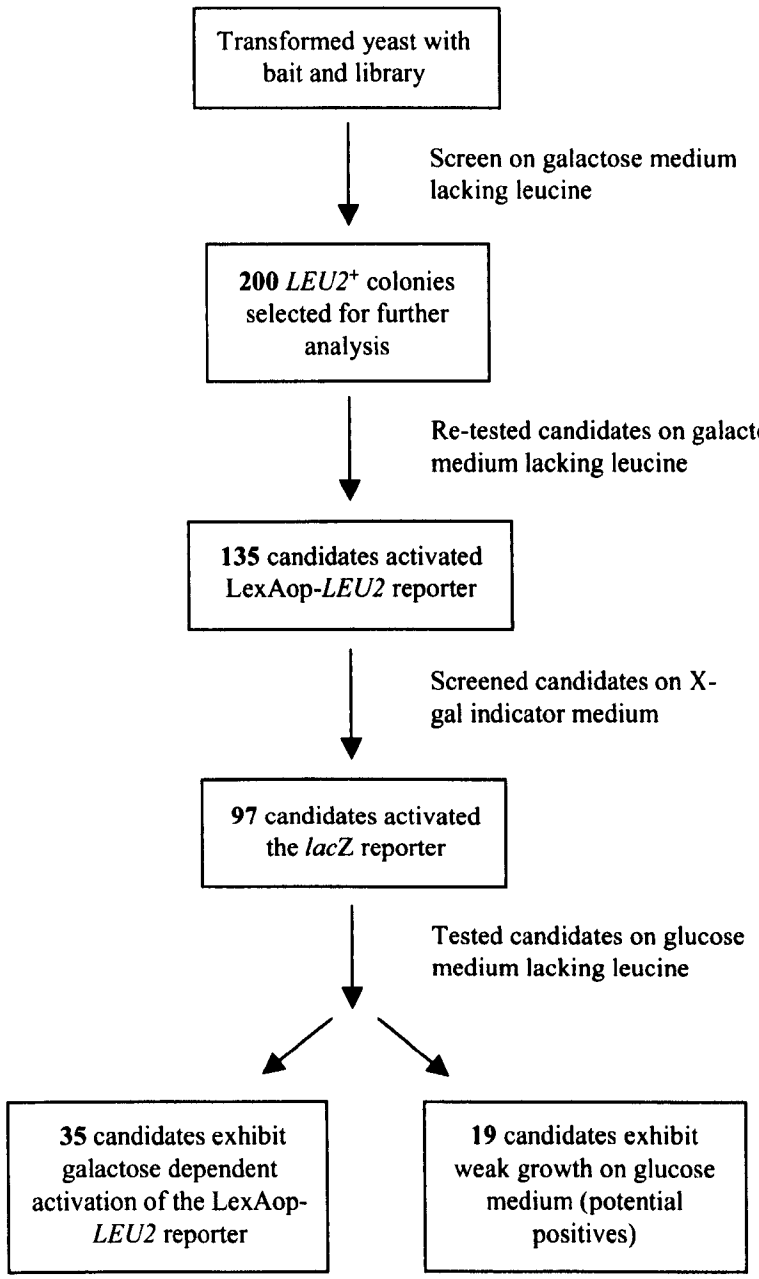


Figure 7.6: Flowchart to summarise the interaction trap screen performed using LexA-Slmo as the bait.

Table 7.1: Candidates identified from the yeast-2-hybrid screen.

(✔= positive, ×=negative, ✔×=potential positive)

Colony No.	Activate *LEU2*	Activate *lacZ*					Galactose Dependent
		None	Weak	Medium	Strong	Very Strong	
1	✔		✔				×
2	✔		✔				×
3	✔		✔				×
4	✔	×					
5	×						
6	×						
7	×						
8	✔		✔				×
9	✔		✔				×
10	×						
11	✔		✔				×
12	✔		✔				×
13	✔		✔				×
14	✔	×					
15	✔	×					
16	×						
17	✔	×					
18	✔	×					
19	✔	×					
20	✔		✔				×
21	✔		✔				×
22	✔		✔				×
23	✔		✔				×
24	✔		✔				×
25	×						
26	×						
27	✔		✔				×
28	✔		✔				×
29	✔		✔				×
30	✔		✔				×
31	✔	×					
32	✔		✔				×
33	✔		✔				×
34	×						
35	×						
36	×						
37	×						
38	✔	×					
39	✔		✔				×
40	✔		✔				×

Table 7.1 (cont.)

Colony No.	Activate *LEU2*	Activate *lacZ*					Galactose Dependent
		None	Weak	Medium	Strong	Very Strong	
41	×						
42	✔		✔				×
43	✔		✔				×
44	×						
45	✔		✔				×
46	✔		✔				×
47	✔		✔				×
48	×						
49	✔	×					
50	✔	×					
51	✔	×					
52	×						
53	×						
54	✔		✔				×
55	×						
56	✔	×					
57	✔	×					
58	×						
59	✔	×					
60	✔		✔				×
61	×						
62	✔		✔				×
63	✔		✔				×
64	✔		✔				×
65	✔	×					
66	×						
67	✔	×					
68	×						
69	×						
70	✔	×					
71	×						
72	×						
73	✔	×					
74	×						
75	×						
76	×						
77	×						
78	×						
79	×						
80	✔		✔				×

Table 7.1 (cont.)

Colony No.	Activate *LEU2*	Activate *lacZ*					Galactose Dependent
		None	Weak	Medium	Strong	Very Strong	
81	✔	×					
82	×						
83	×						
84	×						
85	✔		✔				×
86	×						
87	×						
88	×						
89	✔	×					
90	✔		✔				×
91	✔	×					
92	×						
93	×						
94	✔		✔				×
95	✔	×					
96	✔	×					
97	✔		✔				×
98	✔		✔				×
99	✔	×					
100	✔		✔				×
101	✔				✔		✔×
102	×						
103	✔			✔			✔
104	✔			✔			✔
105	✔			✔			✔
106	×						
107	×						
108	×						
109	✔		✔				✔
110	✔				✔		✔
111	✔				✔		✔×
112	×						
113	×						
114	✔		✔				✔
115	×						
116	×						
117	✔	×					
118	✔				✔		✔×
119	×						
120	×						

Table 7.1 (cont.)

Colony No.	Activate *LEU2*	Activate *lacZ*					Galactose Dependent
		None	Weak	Medium	Strong	Very Strong	
121	×						
122	✔		✔				✔
123	×						
124	×						
125	✔		✔				✔
126	✔		✔				✔
127	✔				✔		✔×
128	✔	×					
129	✔				✔		✔
130	✔					✔	✔×
131	×						
132	✔		✔				×
133	×						
134	✔		✔				✔
135	✔			✔			✔×
136	✔				✔		✔×
137	×						
138	✔				✔		✔
139	✔		✔				✔
140	✔		✔				✔
141	✔			✔			✔×
142	✔		✔				✔×
143	✔			✔			✔
144	✔	×					
145	✔	×					
146	✔			✔			×
147	✔			✔			✔
148	✔			✔			✔×
149	×						
150	✔			✔			✔
151	✔	×					
152	×						
153	✔		✔				✔
154	✔			✔			✔
155	✔		✔				✔
156	✔			✔			✔×
157	✔				✔		✔×
158	×						
159	✔			✔			×
160	×						

Table 7.1 (cont.)

Colony No.	Activate *LEU2*	Activate *lacZ*					Galactose Dependent
		None	Weak	Medium	Strong	Very Strong	
161	✔			✔			✔×
162	✔			✔			✔
163	✔				✔		✔×
164	✔		✔				✔
165	×						
166	✔					✔	✔×
167	✔		✔				✔
168	×						
169	✔				✔		✔×
170	✔				✔		✔×
171	×						
172	×						
173	✔			✔			×
174	✔		✔				✔
175	×						
176	✔				✔		✔×
177	✔	×					
178	×						
179	✔			✔			✔
180	✔		✔				×
181	✔	×					
182	✔			✔			✔×
183	✔				✔		✔
184	✔	×					
185	✔	×					
186	✔		✔				✔
187	×						
188	✔			✔			✔
189	✔	×					
190	✔		✔				✔
191	✔	×					
192	✔	×					
193	✔	×					
194	✔		✔				✔
195	✔			✔			✔
196	✔		✔				✔
197	✔		✔				✔
198	✔			✔			✔
199	✔			✔			✔
200	✔	×					

genes CG40218 and CG6259. Importantly, neither of these proteins are known as common artefacts of yeast-2-hybrid screens, unlike a number of critical mitochondrial proteins (e.g. cytochrome C, ATP-synthetase) that commonly arise as false positives (see Golemis homepage: http://www.fccc.edu/research/labs/golemis). Currently, the CG40218 and CG6259 proteins can only be regarded as candidate Slmo interacting proteins. To confidently report a true *in vivo* interaction with Slmo, it will be necessary to confirm the interaction biochemically, typically by immunoprecipitation and ultimately demonstrate some biological relevance (e.g. similar mutant phenotype). However, the isolation of these candidates, and others that have yet to be identified, provides a basis for future work on the Slmo protein.

The CG40218 protein is relatively small (241 amino acids), and completely uncharacterised with no ascribed or implied function. The PFAM database of protein families (Bateman *et al.*, 2000) identifies an 80 amino acid C-terminal domain of unknown function that is present in proteins from a diverse range of species (PFAM-B: 10149). These include mammalian proteins, BCNT, which was isolated from the bovine brain (Takahashi *et al.*, 1998), and CP27, which has been shown to be expressed in various tissues including mouse brain, heart, lung, liver, salivary glands, teeth and bones, and has been suggested to play a role in mouse embryogenesis (Diekwisch *et al.*, 1999). In addition, predicted proteins can be identified in *C. elegans* (F39H11.1 protein; SwissProt entry: P90867), *Arabidopsis* (BCNT-like protein; SwissProt entry: Q8L8Q4) and *S. cerevisiae* (YB81; SwissProt entry: P38326). The potential interaction with the CG40218 protein does not initially appear to be informative with regards to the biochemical function of Slmo. However, the apparent identification of this protein in 13 of the positive yeast strains, isolated individually from the screen, make this a strong candidate for a genuine interaction.

The CG6259 gene is also uncharacterised, but the predicted protein has similarity to a family of small charged coiled-coil factors typified by the yeast proteins Snf7p (sucrose non-fermenting) and Vps20p (vacuolar protein sorting) (Kranz *et al.*, 2001). Many members of this family (of which there are at least seven in both *Drosophila* and human) remain uncharacterised, but are highly conserved and, as with Slmo, are found in organisms as diverse as *Drosophila*, human, *S. cerevisiae* and *Arabidopsis* (Stauffer *et al.*, 2001; PFAM entry PF03357). In yeast, Snf7p (also known as Vps32) and Vps20p are associated with the late endosomal membrane, and operate together as part of an oligomeric complex called ESCRT-III (endosomal sorting complex required for transport). Here they play an essential role in protein sorting for the multivesicular body (MVB) pathway to the vacuolar lumen (Babst *et al.*, 2002). These proteins are therefore involved in the critical decision between recycling and degradation of internalised plasma membrane proteins (Lemmon and Traub, 2000). Mutation of *snf7* or *vps20* cause severe disruption of endosome to vacuole trafficking, which ultimately manifests as a defect in selective glucose metabolism, presumably due to altered turnover of a glucose sensor (Kranz *et al.*, 2001; Tu *et al.*, 1993). A similar function has been demonstrated for the mammalian Snf7p homologues CHMP4a and CHMP4b (charged multivesicular body protein), overexpression of which alters MVB trafficking of both EGF (epidermal growth factor) and cholesterol to the lysosome (Katoh *et al.*, 2003; Peck *et al.*, 2004).

If the CG6259 protein is involved in the same physiological process as Snf7p then it would be very unlikely to represent a true interactor of Slmo. However, given that the Slmo yeast homologue MSF1p' has been implicated in protein sorting at the mitochondria (Nakai *et al.*, 1993; T. Endo, pers. comm.), a candidate interacting

protein related to factors involved in protein sorting elsewhere in the cell is of considerable interest. Preliminary phylogenetic analysis (PFAM entry PF03357) suggests that CG6259 does not represent a genuine homologue of Snf7p, and there appears to be a substantial degree of functional divergence between the various members of this family. For example, the mammalian CHMP1 is associated with the nuclear matrix and has been shown to play a role in gene regulation by affecting chromatin structure (Stauffer *et al.*, 2001), whilst the closely related human gene BC-2 has also been shown to be upregulated in the nuclear matrix of breast adenocarcinoma (Stauffer *et al.*, 2001). Could the CG6259 protein be involved in protein sorting in the mitochondria? At present there is no data to support this hypothesis. Bioinformatic programs such as PSORTII, iPSORT (http://psort.nibb.ac.jp) and TARGETP (Emanuelsson *et al.*, 2000) do not predict a mitochondrial localisation for CG6259 with any confidence, but also do not eliminate the possibility. Clearly, it would be desirable to determine the subcellular localisation of this protein.

Chapter 8: General Discussion

Chapter 8: General Discussion

8.1 Introduction

This thesis describes an analysis of the *Drosophila* gene *slmo* (also known as *kisir*), and the related genes *prel* and *retm*. The project began with an interest in an enhancer trap line, P(GAL4)c682, expressed specifically, but not ubiquitously, within the CNS during embryonic development. Mutation of the corresponding gene, *slmo*, results in severe larval motor defects and eventual lethality. In the course of this thesis I have described Slmo to be a mitochondrial localised protein and a representative of a novel family of conserved proteins, found in organisms as diverse as flies, worms, yeast, mammals and plants. A discussion of experimental results is provided at the end of each of each of the results chapters 3-7. This section aims to draw together the major conclusions and address some of the under lying themes of this research.

8.2 The role of *slmo* in the CNS

A functional neuronal circuit has two basic requirements. Firstly, the component neurons must form the correct array of synaptic connections which mediate the correct pattern of communication between neurons. This allows them to integrate into a dynamic unit capable of generating functional output (Goodman, 1993). Secondly, the circuit must be able to modulate this functional output according to the specific requirements of a particular behaviour (Marder and Bucher, 2001). In the case of Slmo, the protein does not appear to be required for the establishment of neuronal connections within the CNS as all available evidence suggests strongly that the circuitry of *slmo* mutants is intact (Carhan, 1999; Reeve,

2002; see chapter 3). However, *slmo* is required for normal larval locomotor behaviour in *Drosophila* (3.2.1). Although initially able to hatch, 1st instars homozygous for a *slmo* null allele ($slmo^{\Delta 69m1}$) are severely impaired in their ability to generate the rhythmic peristaltic contraction waves that drive this relatively simple stereotyped behaviour. This phenotype subsequently progresses towards total inactivity, although the ability to move the head and mouth parts is retained. The severely compromised locomotor activity observed in $slmo^{\Delta 69m1}$ mutant larvae suggests abnormal functional output of the motor circuitry that drives this behaviour. In addition, available data on the expression of the *slmo* gene suggests that this phenotype of compromised motor output is caused by alteration of the function of a limited subset of cells within the CNS. In the developing embryonic CNS, both the P(GAL4)c682 enhancer trap (3.2.2.1) and *in situ* hybridisation to the *slmo* transcript (Carhan, 1999), report expression to be exclusive to a subset of cells within the CNS. During larval development, when the *slmo* mutant phenotype is first evident, the enhancer trap reports expression only in discrete populations of neurons within the VNC and larval brain lobes.

The precise mechanism by which Slmo might affect neuronal function is currently unknown. Various lines of evidence suggest that the motor circuitry that drives locomotor behaviour develops normally in the $slmo^{\Delta 69m1}$ mutant (chapter 3; Carhan, 1999; Reeve, 2002; R. Baines, pers. comm.). Despite this, subtle abnormalities in the development of the nervous system cannot be formally eliminated. For example, defective axon pathfinding of a very small subset of neurons would be difficult to establish. If the assembly of the nervous system is indeed unaffected by *slmo* loss of function, then this presents the possibility that Slmo is a factor important in modulating the functional activity of a specific

subpopulations of neurons. However, it should be noted that the reported expression of Slmo in early CNS glia (chapter 3; Carhan, 1999) and the adult gonads (Reeve, 2002) indicate that the function of this protein is not exclusively neuronal. The potential biochemical function of Slmo within neurons and other cell types is therefore of considerable interest. The establishment that Slmo function is associated with the mitochondria (chapter 4) suggests the cause of the phenotype might be viewed as a cell biological defect with severe consequences for neuronal function.

8.3 The association of Slmo with the mitochondria

Mitochondria are of course crucial in all cell types and have an integral involvement in many cellular processes, including energy production, regulation of intracellular calcium levels (Meldolesi, 2001; Vanden Berghe *et al.*, 2002), and the execution of apoptotic cell death (Claveria and Torres, 2003). The latter possibility was addressed in chapter 3 (3.2.4), and *slmo* does not appear to be required for neuronal viability. Indeed, the lack of any obvious developmental phenotype in *slmo* mutants is not inconsistent with the observation that even in the total absence of mitochondria from all neuronal projections, the synaptic terminals exhibit surprisingly normal ultrastructure and no degeneration (Stowers *et al.*, 2002). The implication is that other sources of ATP, such as anaerobic respiration or diffusion from the cell body, must compensate to some extent.

All cells depend on the mitochondria for the production of energy in the useable form of ATP. The electrical properties of neurons have a particularly high metabolic demand which needs to be serviced (Laughlin, 2001; Laughlin *et al.*, 1998). If the consequences of loss of Slmo function included a reduction in the availability of ATP this could adversely affect any number of ATP dependent

processes underlying the specialised function of neurons. These metabolic costs include the need to package neurotransmitter into synaptic vesicles, release of the neurotransmitter and the reuptake of neurotransmitter and metabolites (Laughlin, 2001; Laughlin *et al.*, 1998). In particular, are the metabolic demands of regulating the ionic contents of the cytoplasm in relation to the extracellular space. This is of course, essential to the ability to generate electrical signals, and the ion pumps that mediate this function are energetically expensive (Laughlin, 2001; Laughlin *et al.*, 1998). It follows that the overall demand for ATP varies in proportion to the level of neuronal activity, and it is necessary for the mitochondria to respond to the local needs of oxidative metabolism by changes to their activity (Bindokas *et al.*, 1998). It is possible that Slmo is involved somehow in regulating metabolic output in specific cells or tissues. Certainly, examples do exist of non-ubiquitous factors that affect activity in this way, such as the mammalian mitochondrial uncoupling factor BMCP1, which has been shown to modify mitochondrial respiratory efficiency in particular neurons within the human brain (Kim-Han *et al.*, 2001).

Another possibility is that loss of Slmo function in some way compromises the regulation of intracellular Ca^{2+} homeostasis, perturbations of which could affect the fitness of many cell types but would have particularly severe consequences for neuronal signalling and excitability (Meldolesi, 2001). Disruption of mitochondrial function might firstly indirectly affect this process by diminishing the extrusion of Ca^{2+} across the cell membrane as the Na^{2+} gradient diminishes due lack of ATP. In addition, mitochondria are known to work in partnership with the endoplasmic reticulum in buffering Ca^{2+} levels by rapidly and efficiently taking up Ca^{2+} when concentrations are high and therefore help reduce the concentration following a Ca^{2+} spike (Berridge, 1998; Meldolesi, 2001). The ER is the major organelle involved in

regulating calcium levels but the mitochondria act to increase the speed and efficiency of the response in localised compartments of the cell. This function is particularly important in crucial areas such as the synaptic terminal, where the volume of the cytoplasm is very low and regulating the intracellular ion concentration is more challenging (Berridge, 1998). Mitochondria are therefore closely physically associated with the ER at these locations and the relationship is to some extent reciprocal, as the mitochondria require a supply of Ca^{2+} ions for the process of oxidative metabolism. The tight regulation and buffering of calcium is crucial for neuronal excitability, neurotransmitter release, the regulation of gene transcription and synaptic plasticity (Berridge, 1998; Zucker, 1999). It would be interesting to analyse both Ca^{2+} levels and metabolic activity in Slmo expressing cells, but currently there is no data available that might suggest the precise biochemical action of Slmo within these processes.

8.4 A role in mitochondrial protein sorting?

Mitochondria contain many proteins, only a few of which are encoded by the mitochondrial genome. The vast majority of these factors, like Slmo, are nuclear encoded, and manufactured in the cytosol in parallel with proteins earmarked for other locations within the cell (Neupert and Brunner, 2002). This necessitates a highly organised system of protein trafficking that ensures the coordinated import of these proteins into the mitochondrion (Truscott *et al.*, 2003). Proteins destined for the mitochondria must be efficiently recognised and delivered to the organelle by a system of chaperone proteins (Hoogenraad *et al.*, 2002). In addition, imported proteins must be accurately sorted and directed to the functionally distinctive

subcompartments of the mitochondria, these being the outer membrane, the inner membrane, the intermembrane space, and the matrix (Truscott *et al.*, 2003).

The signal required for accurate targeting and sorting appears to be contained within the protein itself. This is traditionally viewed in terms of an N-terminal presequence that is ultimately removed, but in many cases, particularly proteins of the outer and inner membranes, the targeting and sorting information is contained within the mature protein (Truscott *et al.*, 2003). A number of specialised translocase complexes are responsible for the recognition of these targeting signals and facilitating the transfer of proteins across the mitochondrial membranes. Passage through the outer membrane is mediated by the TOM (translocase of the outer membrane) complex, which is essentially comprised of protein import receptors and an import channel (Truscott *et al.*, 2003; Wiedemann *et al.*, 2004). Proteins must then be sorted, by a number of mechanisms, either to the outer membrane, the intermembrane space or to the TIM (translocase of the inner membrane) complexes, these being TIM23, which mediates translocation into the matrix, or TIM22, which mediates insertion into the inner membrane (Truscott *et al.*, 2003; Wiedemann *et al.*, 2004). In addition, these major complexes are assisted by an array of molecular chaperones and processing enzymes. For example, a family of factors have been identified in yeast, referred to as the "small Tims", that are involved in mediating sorting and transport of inner membrane proteins across the intermembrane space (Koehler, 2004). These can operate in the aqueous chamber (e.g. Tim10), or be associated with the outer surface of the inner membrane (e.g. Tim12) (Koehler, 2004). A variety of other proteins have also been identified by Toshia Endo's lab with potential roles in mitochondrial protein sorting (Kanamori *et al.*, 1997).

Toshia Endo and colleagues were responsible for the identification of the yeast Slmo homologue, MSF1p' (Nakai *et al.*, 1993), a myc-tagged fusion of which was reported to be located at the inner mitochondrial membrane. Importantly, over-expression of the gene was reported to alter the location of a reporter directed to the intermembrane space (T. endo, pers. comm.). This observation has been interpreted as an indication that MSF1p' (mitochondrial sorting factor 1') is involved in the process of intracompartmental protein sorting (Nakai *et al.*, 1993). Might this be the function of the Slmo protein? Some circumstantial evidence in support of this view is provided by the identification, from a yeast-2-hybrid screen, of a member of the Snf7p family of charged coiled-coil proteins (Kranz *et al.*, 2001; Peck *et al.*, 2004; Stauffer *et al.*, 2001) as a potential interactor of the Slmo protein (see chapter 7). The protein in question, encoded by the predicted gene CG6259 (FlyBase entry: FBgn0036740), is itself uncharacterised, but similar proteins in yeast and mammals have essential roles in protein sorting systems outside the mitochondria, specifically, for the multivesicular body pathway between the endosome and the vacuole/ lysosome (Babst *et al.*, 2002; Peck *et al.*, 2004). Currently, most members of this protein family, of which there are a least six members present in both humans and flies, are identified from the databases only as hypothetical or predicted proteins. However, it is likely that the various members of the Snf7p family exhibit a significant degree of functional divergence, as indicated by the observation that two mammalian members of the Snf7p family, CHMP1 and BC-2, have functions in the cell nucleus (Stauffer *et al.*, 2001). Although no proteins of this type have yet been linked to the mitochondria, the identification of a potential Slmo interacting protein related to factors involved in protein sorting is of considerable interest and merits further research.

Interestingly, in humans, loss of function of the *DDP1* (deafness-dystonia peptide 1) gene, which encodes a protein with homology to the "small Tim" family of yeast proteins, has been identified as causing Mohr-Tranebjaer syndrome, an X-linked degenerative disease characterised by a number of progressive neurological symptoms including deafness, blindness, dystonia and mental retardation (Bauer and Neupert, 2001). This severe phenotype is thought to result from defective biogenesis of the TIM23 translocase, essential for protein import at the inner membrane, including many components of oxidative phosphorylation system. Perhaps the phenotypic consequences of *slmo* loss of function result from lowered abundance, possibly at the inner membrane, of some mitochondrial proteins important for neuronal function.

8.5 The function of Prel and Retm

Slmo exhibits a high degree of evolutionary conservation, and represents a novel family of proteins present in diverse eukaryotic taxa including *C. elegans*, *S. cerevisiae* (including MSF1p'), *Arabidopsis* and human (chapter 5). In animal and yeast species, multiple Slmo related factors can be identified. For example, *Drosophila* contains three such proteins, encoded by the *slmo*, *prel* and *retm* genes. An interesting question is whether the Slmo/MSF1p' domain fulfils a common biochemical role in all the proteins of the family. Although all three of Slmo, Prel and Retm (and related proteins in other organisms) have now been linked to the mitochondria (6.2.1.2; Huh *et al.*, 2003; Fox *et al.*, 2004), the phylogenetic relationships between these factors suggest that a degree of functional divergence is likely (5.2.2). Even so, the Slmo/MSF1p' domain may fulfil similar biochemical roles in different cellular contexts, or take part in alternative cellular processes.

The *slmo* gene appears to encode a factor with an essential role within the mitochondria, but the requirement for this function is restricted to particular tissues at specific developmental stages. In particular, research in both this thesis and previous reports has focussed on the embryonic and larval CNS (chapter 3), and in the adult germline (Carhan, 1999). In the case of the *prel* gene, expression appears to be widespread, and is indeed ubiquitous during embryonic development (6.2.2). This might suggest a general requirement for the Prel protein all tissues. Certainly, the lethality (during larval development) exhibited by a *prel* mutant indicates that the function of this gene is essential in *Drosophila* (6.2.3). However, the requirement for Prel is not necessarily uniform. In the developing embryo, *in situ* hybridisation suggests that transcription is upregulated specifically in the developing brain lobes and certain regions of the midgut (6.2.2). This might suggest that the requirement for Prel function is altered in different tissues at different developmental stages, a view supported by the reported expression of the homologous human PRELI gene (Guzman-Rojas *et al.*, 2000). The PRELI transcript is expressed in a wide range of human tissues including liver, brain, kidney, lung and spleen. In the developing foetus, expression is specifically high in the liver, whilst mRNA levels in other tissues are comparatively low. Conversely, at the adult stage, PRELI expression in the liver is substantially reduced, whilst the transcript is detected at high levels in spleen, brain, lung and kidney (Guzman-Rojas *et al.*, 2000). However, as with Slmo, the cellular processes affected by the Prel protein are at this stage unknown.

P-insert and deletion alleles of *retm* do not result in lethality (2.2.3), indicating that the function of this gene is not essential in *Drosophila*. However, the *retm* gene product remains interesting as a representative of a novel subclass of proteins containing the Slmo/MSF1p' domain in conjunction a functionally

characterised motif known as the CRAL-TRIO domain (5.2.2). This conserved domain has been shown in a number of proteins to mediate the binding and transport of small hydrophobic ligands. In particular, the domain is defined by its homology to the yeast phosphatidylinositol transfer protein (PITP), SEC14p. In *S. cerevisiae*, this essential protein mediates the binding and transfer of phosphatidylinositol or phosphatidylcholine, and by regulating the availability of these molecules, plays an integral role in a variety of signal transduction and membrane transport events (Li *et al.*, 2000). However, not all SEC14p related proteins are PITPs, and in higher eukaryotes, the CRAL-TRIO domain has been shown to bind alternative ligands such as retinol, α-tocopherol and squalene (Arita *et al.*, 1995; Crabb *et al.*, 1998; Shibata *et al.*, 2001). In the case of Retm, the ligand has not yet been identified.

8.6 Implications for human disease

The importance of the mitochondria to all living cells is well established. This view is accentuated by the more recent discovery that mitochondrial dysfunction is involved in a variety of degenerative diseases, ageing and cancer (Schon and Manfredi, 2003; Wallace, 1999). In many cases, the mechanisms underlying such conditions have been attributed to disruption of nuclear encoded proteins. The identification and characterisation of novel mitochondrial proteins is therefore significant not only for our understanding of normal cellular function, but may also have implications for understanding a variety of human disorders. In some cases, relevant mutations and comparable phenotypes have been identified in model systems such as *Drosophila*. For example, mutation of the *technical knockout* (*tko*) gene, which encodes the mitochondrial ribosomal S12 protein, has been shown to cause mechanosensory failure resulting in behavioural phenotypes such as sensitivity

to mechanical stress and impaired response to sound stimulus (Toivonen *et al.*, 2001). Such phenotypes would appear to parallel mitochondria related sensorineural deafness in human, itself a consequence of mechanosensory failure of the receptor cells of the inner ear (Jacobs, 1997). The *sluggish-A* gene was initially identified as a mutant causing locomotory phenotypes in adult *Drosophila*, and was found to encode mitochondrial proline oxidase (Hayward *et al.*, 1993). Mammalian homologues have subsequently been linked to human neurological disorders (Campbell *et al.*, 1997; Gogos *et al.*, 1999). The *scully* gene encodes the *Drosophila* homologue of the vertebrate type II L-3-hydroxylacyl-CoA dehydrogenase (HADH), a mitochondrial enzyme essential for β-oxidation of fatty acids (Torroja *et al.*, 1998). Mutation of *scully* results in abnormal photoreceptor development and sterility phenotypes suggested to reflect human pathologies caused by β-oxidation disorders. An excellent example is the *Drosophila parkin* gene, a homologue of the human *parkin* gene linked to an early onset form of Parkinson's disease. In this case, analysis of *parkin* loss of function has provided a valuable insight into the mechanisms of mitochondrial dysfunction that underlie this condition (Greene *et al.*, 2003). The identification of a novel family of mitochondrial associated proteins may therefore be of general relevance. Further analysis of the function of Slmo, Prel and Retm, and the relationship to mammalian homologues is likely to prove interesting.

References

Afshar, K., Stuart, B. and Wasserman, S. A. (2000). Functional analysis of the Drosophila diaphanous FH protein in early embryonic development. *Development* **127**, 1887-1897.

Allen, M. J., Drummond, J. A. and Moffat, K. G. (1998). Development of the giant fiber neuron of Drosophila melanogaster. *J Comp Neurol* **397**, 519-531.

Allen, M. J., O'Kane, C. J. and Moffat, K. G. (2002). Cell ablation using wild-type and cold-sensitive ricin-A chain in Drosophila embryonic mesoderm. *Genesis* **34**, 132-134.

Allen, M. J., Shan, X. and Murphey, R. K. (2000). A role for Drosophila Drac1 in neurite outgrowth and synaptogenesis in the giant fiber system. *Mol Cell Neurosci* **16**, 754-765.

Allen-Baume, V., Segui, B. and Cockcroft, S. (2002). Current thoughts on the phosphatidylinositol transfer protein family. *FEBS Lett* **531**, 74-80.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

Anantharaman, V. and Aravind, L. (2002). The GOLD domain, a novel protein module involved in Golgi function and secretion. *Genome Biol* **3**, research0023.

Arita, M., Sato, Y., Miyata, A., Tanabe, T., Takahashi, E., Kayden, H. J., Arai, H. and Inoue, K. (1995). Human alpha-tocopherol transfer protein: cDNA cloning, expression and chromosomal localization. *Biochem J* **306 (Pt 2)**, 437-443.

Ashburner, M. (1989). Drosophila, a laboratory handbook: Cold Spring Harbor Laboratory Press.

Babst, M., Katzmann, D. J., Estepa-Sabal, E. J., Meerloo, T. and Emr, S. D. (2002). Escrt-III: an endosome-associated heterooligomeric protein complex required for mvb sorting. *Dev Cell* **3**, 271-282.

Baines, R. A. (2003). Postsynaptic protein kinase A reduces neuronal excitability in response to increased synaptic excitation in the Drosophila CNS. *J Neurosci* **23**, 8664-8672.

Baines, R. A. and Bate, M. (1998). Electrophysiological development of central neurons in the Drosophila embryo. *J Neurosci* **18**, 4673-4683.

Baines, R. A., Robinson, S. G., Fujioka, M., Jaynes, J. B. and Bate, M. (1999). Postsynaptic expression of tetanus toxin light chain blocks synaptogenesis in Drosophila. *Curr Biol* **9**, 1267-1270.

Baines, R. A., Seugnet, L., Thompson, A., Salvaterra, P. M. and Bate, M. (2002). Regulation of synaptic connectivity: levels of Fasciclin II influence synaptic growth in the Drosophila CNS. *J Neurosci* **22**, 6587-6595.

Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T. and Bate, M. (2001). Altered electrical properties in Drosophila neurons developing without synaptic transmission. *J Neurosci* **21**, 1523-1531.

Bate, M. (1993). The mesoderm and its derivatives. In *The development of Drosophila melanogaster*, vol. 2 (ed. M. a. M. A. Bate, A), pp. 1013-1090. Cold Spring Harbor: Cold Spring Harbor press.

Bate, M. (1999). Development of motor behaviour. *Curr Opin Neurobiol* **9**, 670-5.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L. and Sonnhammer, E. L. (2000). The Pfam protein families database. *Nucleic Acids Res* **28**, 263-266.

Bauer, M. F. and Neupert, W. (2001). Import of proteins into mitochondria: a novel pathomechanism for progressive neurodegeneration. *J Inherit Metab Dis* **24**, 166-180.

Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. and Gehring, W. J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. *Genes Dev* **3**, 1288-1300.

Bellingham, M. C. (1998). Driving respiration: the respiratory central pattern generator. *Clin Exp Pharmacol Physiol* **25**, 847-856.

Berridge, M. J. (1998). Neuronal calcium signaling. *Neuron* **21**, 13-26.

Berrigan, D. and Pepin., DJ. (1995). How Maggots Move: Allometry and Kinematics of Crawling in Larval Diptera. *Journal of Insect Physiology* **41**, 329-337.

Bhat, K. M. (1998). Cell-cell signaling during neurogenesis: some answers and many questions. *Int J Dev Biol* **42**, 127-139.

Bindokas, V. P., Lee, C. C., Colmers, W. F. and Miller, R. J. (1998). Changes in mitochondrial function resulting from synaptic activity in the rat hippocampal slice. *J Neurosci* **18**, 4570-4587.

Bodily, K. D., Morrison, C. M., Renden, R. B. and Broadie, K. (2001). A novel member of the Ig superfamily, turtle, is a CNS-specific protein required for coordinated motor control. *J Neurosci* **21**, 3113-3125.

Bossing, T. and Technau, G. M. (1994). The fate of the CNS midline progenitors in Drosophila as revealed by a new method for single cell labelling. *Development* **120**, 1895-1906.

Bowman, A. B., Kamal, A., Ritchings, B. W., Philp, A. V., McGrail, M., Gindhart, J. G. and Goldstein, L. S. (2000). Kinesin-dependent axonal transport is mediated by the sunday driver (SYD) protein. *Cell* **103**, 583-594.

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

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

Broadie, K. and Bate, M. (1993). Activity-dependent development of the neuromuscular synapse during Drosophila embryogenesis. *Neuron* **11**, 607-619.

Buchanan, R. L. and Benzer, S. (1993). Defective glia in the Drosophila brain degeneration mutant drop-dead. *Neuron* **10**, 839-850.

Campbell, H. D., Webb, G. C. and Young, I. G. (1997). A human homologue of the Drosophila melanogaster sluggish-A (proline oxidase) gene maps to 22q11.2, and is a candidate gene for type-I hyperprolinaemia. *Hum Genet* **101**, 69-74.

Carew, T. J. (1996). Molecular enhancement of memory formation. *Neuron* **16**, 5-8.

Carhan, A. (1999). Analysis of the kisir mutation in Drosophila melanogaster. In *Biological Sciences*, (ed. Coventry: University of Warwick.

Carlson, J. R. (1996). Olfaction in Drosophila: from odor to behavior. *Trends Genet* **12**, 175-180.

Casso, D., Ramirez-Weber, F. and Kornberg, T. B. (2000). GFP-tagged balancer chromosomes for Drosophila melanogaster. *Mech Dev* **91**, 451-454.

Castrillon, D. H., Gonczy, P., Alexander, S., Rawson, R., Eberhart, C. G., Viswanathan, S., DiNardo, S. and Wasserman, S. A. (1993). Toward a molecular genetic analysis of spermatogenesis in Drosophila melanogaster: characterization of male-sterile mutants generated by single P element mutagenesis. *Genetics* **135**, 489-505.

Castrillon, D. H. and Wasserman, S. A. (1994). Diaphanous is required for cytokinesis in Drosophila and shares domains of similarity with the products of the limb deformity gene. *Development* **120**, 3367-77.

Cattaert, D. and Birman, S. (2001). Blockade of the central generator of locomotor rhythm by noncompetitive NMDA receptor antagonists in Drosophila larvae. *J Neurobiol* **48**, 58-73.

Certel, S. J., Clyne, P. J., Carlson, J. R. and Johnson, W. A. (2000). Regulation of central neuron synaptic targeting by the Drosophila POU protein, Acj6. *Development* **127**, 2395-2405.

Chalfie, M. and Kain, S. (1998). Green Fluorescent Protein : Properties, Applications, and Protocols: Wiley-Liss.

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 U S A* **88**, 9578-9582.

Claveria, C. and Torres, M. (2003). Mitochondrial apoptotic pathways induced by Drosophila programmed cell death regulators. *Biochem Biophys Res Commun* **304**, 531-537.

Cockcroft, S. (1998). Phosphatidylinositol transfer proteins: a requirement in signal transduction and vesicle traffic. *Bioessays* **20**, 423-432.

Cockcroft, S. and De Matteis, M. A. (2001). Inositol lipids as spatial regulators of membrane traffic. *J Membr Biol* **180**, 187-194.

Cooley, L., Kelley, R. and Spradling, A. (1988). Insertional mutagenesis of the Drosophila genome with single P elements. *Science* **239**, 1121-1128.

Crabb, J. W., Nie, Z., Chen, Y., Hulmes, J. D., West, K. A., Kapron, J. T., Ruuska, S. E., Noy, N. and Saari, J. C. (1998). Cellular retinaldehyde-binding protein ligand interactions. Gln-210 and Lys-221 are in the retinoid binding pocket. *J Biol Chem* **273**, 20712-20720.

Davis, R. L. (1996). Physiology and biochemistry of Drosophila learning mutants. *Physiol Rev* **76**, 299-317.

Deitcher, D. L., Ueda, A., Stewart, B. A., Burgess, R. W., Kidokoro, Y. and Schwarz, T. L. (1998). Distinct requirements for evoked and spontaneous release of neurotransmitter are revealed by mutations in the Drosophila gene neuronal-synaptobrevin. *J Neurosci* **18**, 2028-2039.

Dicaprio, R. (1997). Plateau potentials in motor neurons in the ventilatory system of the crab. *J Exp Biol* **200**, 1725-1736.

Dicaprio, R., Jordan, G. and Hampton, T. (1997). Maintenance of motor pattern phase relationships in the ventilatory system of the crab. *J Exp Biol* **200**, 963-974.

Dickenson, M., Farley, CT, Full, RJ, Koehl, MA, Kram, R, Lehman, S. (2000). How animals move: an integrative view. *Science* **288**, 100-106.

Diekwisch, T. G., Marches, F., Williams, A. and Luan, X. (1999). Cloning, gene expression, and characterization of CP27, a novel gene in mouse embryogenesis. *Gene* **235**, 19-30.

Doe, C. Q. (1992). Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system. *Development* **116**, 855-863.

Doe, C. Q., Smouse, D. and Goodman, C. S. (1988). Control of neuronal fate by the Drosophila segmentation gene even-skipped. *Nature* **333**, 376-378.

Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M. and Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* **97**, 703-716.

Drapeau, P., Saint-Amant, L., Buss, R. R., Chong, M., McDearmid, J. R. and Brustein, E. (2002). Development of the locomotor network in zebrafish. *Prog Neurobiol* **68**, 85-111.

Driscoll, M. and Gerstbrein, B. (2003). Dying for a cause: invertebrate genetics takes on human neurodegeneration. *Nat Rev Genet* **4**, 181-194.

Dubnau, J. and Tully, T. (1998). Gene discovery in Drosophila: new insights for learning and memory. *Annu Rev Neurosci* **21**, 407-444.

Dzierzak, E., Medvinsky, A. and de Bruijn, M. (1998). Qualitative and quantitative aspects of haematopoietic cell development in the mammalian embryo. *Immunol Today* **19**, 228-236.

Eberl, D. F. (1999). Feeling the vibes: chordotonal mechanisms in insect hearing. *Curr Opin Neurobiol* **9**, 389-393.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**, 1005-1016.

Engels, W. R. (1984). A trans-acting product needed for P factor transposition in Drosophila. *Science* **226**, 1194-1196.

Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164-166.

Fenelon, V. S., Casasnovas, B., Simmers, J. and Meyrand, P. (1998). Development of rhythmic pattern generators. *Curr Opin Neurobiol* **8**, 705-709.

Ferveur, J. F., Stortkuhl, K. F., Stocker, R. F. and Greenspan, R. J. (1995). Genetic feminization of brain structures and changed sexual orientation in male Drosophila. *Science* **267**, 902-905.

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

Finley, R. L., Jr., Thomas, B. J., Zipursky, S. L. and Brent, R. (1996). Isolation of Drosophila cyclin D, a protein expressed in the morphogenetic furrow before entry into S phase. *Proc Natl Acad Sci U S A* **93**, 3011-3015.

Fletcher, C. F., Lutz, C. M., O'Sullivan, T. N., Shaughnessy, J. D., Jr., Hawkes, R., Frankel, W. N., Copeland, N. G. and Jenkins, N. A. (1996). Absence epilepsy in tottering mutant mice is associated with calcium channel defects. *Cell* **87**, 607-617.

Fox, E. J., Stubbs, S. A., Kyaw Tun, J., Leek, J. P., Markham, A. F. and Wright, S. C. (2004). PRELI (protein of relevant evolutionary and lymphoid interest) is located within an evolutionarily conserved gene cluster on chromosome 5q34-q35 and encodes a novel mitochondrial protein. *Biochem J* **378**, 817-825.

Friggi-Grelin, F., Coulom, H., Meller, M., Gomez, D., Hirsh, J. and Birman, S. (2003). Targeted gene expression in Drosophila dopaminergic cells using regulatory sequences from tyrosine hydroxylase. *J Neurobiol* **54**, 618-627.

Fujioka, M., Jaynes, J. B., Bejsovec, A. and Weir, M. (2000). Production of transgenic Drosophila. *Methods Mol Biol* **136**, 353-63.

Fuller, M. (1993). Spermatogenesis. In *The development of drosophila melanogaster*, (ed. M. a. M. A. Bate, A), pp. 71-147. New York: Cold Spring Harbour Laboratory Press.

Gindhart, J. G., Chen, J., Faulkner, M., Gandhi, R., Doerner, K., Wisniewski, T. and Nandlestadt, A. (2003). The kinesin-associated protein UNC-76 is required for axonal transport in the Drosophila nervous system. *Mol Biol Cell* **14**, 3356-3365.

Gindhart, J. G., Jr., Desai, C. J., Beushausen, S., Zinn, K. and Goldstein, L. S. (1998). Kinesin light chains are essential for axonal transport in Drosophila. *J Cell Biol* **141**, 443-454.

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E. et al. (2003). A protein interaction map of Drosophila melanogaster. *Science* **302**, 1727-1736.

Gogos, J. A., Santha, M., Takacs, Z., Beck, K. D., Luine, V., Lucas, L. R., Nadler, J. V. and Karayiorgou, M. (1999). The gene encoding proline dehydrogenase modulates sensorimotor gating in mice. *Nat Genet* **21**, 434-439.

Goldschmidt-Clermont, P. J., Machesky, L. M., Baldassare, J. J. and Pollard, T. D. (1990). The actin-binding protein profilin binds to PIP2 and inhibits its hydrolysis by phospholipase C. *Science* **247**, 1575-1578.

Goldstein, L. S. (2001). Kinesin molecular motors: transport pathways, receptors, and human disease. *Proc Natl Acad Sci U S A* **98**, 6999-7003.

Goldstein, L. S. and Philp, A. V. (1999). The road less traveled: emerging principles of kinesin motor utilization. *Annu Rev Cell Dev Biol* **15**, 141-183.

Golemis, E. A. and Brent, R. (1992). Fused protein domains inhibit DNA binding by LexA. *Mol Cell Biol* **12**, 3006-3014.

Gonczy, P., Matunis, E. and DiNardo, S. (1997). bag-of-marbles and benign gonial cell neoplasm act in the germline to restrict proliferation during Drosophila spermatogenesis. *Development* **124**, 4361-4371.

Goodman, C. a. D., CQ. (1993). Embryonic development of the Drosophila central nervous system. In *The development of Drosophila melanogaster*, (ed. M. a. M. A. Bate, A), pp. 1131-1206. New York: Cold Spring Harbour Laboratory Press.

Gorczyca, M. G., Budnik, V., White, K. and Wu, C. F. (1991). Dual muscarinic and nicotinic action on a motor program in Drosophila. *J Neurobiol* **22**, 391-404.

Granato, M., van Eeden, F. J., Schach, U., Trowe, T., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C. P., Jiang, Y. J. et al. (1996). Genes controlling and mediating locomotion behavior of the zebrafish embryo and larva. *Development* **123**, 399-413.

Granderath, S., Stollewerk, A., Greig, S., Goodman, C. S., O'Kane, C. J. and Klambt, C. (1999). loco encodes an RGS protein required for Drosophila glial differentiation. *Development* **126**, 1781-1791.

Greene, J. C., Whitworth, A. J., Kuo, I., Andrews, L. A., Feany, M. B. and Pallanck, L. J. (2003). Mitochondrial pathology and apoptotic muscle degeneration in Drosophila parkin mutants. *Proc Natl Acad Sci U S A* **100**, 4078-4083.

Grenningloh, G., Rehm, E. J. and Goodman, C. S. (1991). Genetic analysis of growth cone guidance in Drosophila: fasciclin II functions as a neuronal recognition molecule. *Cell* **67**, 45-57.

Gunawardena, S. and Goldstein, L. S. (2001). Disruption of axonal transport and neuronal viability by amyloid precursor protein mutations in Drosophila. *Neuron* **32**, 389-401.

Guzman-Rojas, L., Sims, J. C., Rangel, R., Guret, C., Sun, Y., Alcocer, J. M. and Martinez-Valdez, H. (2000). PRELI, the human homologue of the avian px19, is expressed by germinal center B lymphocytes. *Int Immunol* **12**, 607-612.

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.

Hall, D. H., Gu, G., Garcia-Anoveros, J., Gong, L., Chalfie, M. and Driscoll, M. (1997). Neuropathology of degenerative cell death in Caenorhabditis elegans. *J Neurosci* **17**, 1033-1045.

Hall, J. C. (1994). The mating of a fly. *Science* **264**, 1702-1714.

Hall, J. C. (1998). Molecular neurogenetics of biological rhythms. *J Neurogenet* **12**, 115-181.

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

Hartenstein, V. (1993). Atlas of Drosophila Development. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Hayward, D. C., Delaney, S. J., Campbell, H. D., Ghysen, A., Benzer, S., Kasprzak, A. B., Cotsell, J. N., Young, I. G. and Miklos, G. L. (1993). The sluggish-A gene of Drosophila melanogaster is expressed in the nervous system and encodes proline oxidase, a mitochondrial enzyme involved in glutamate biosynthesis. *Proc Natl Acad Sci U S A* **90**, 2979-2983.

Heisenberg, M. (2003). Mushroom body memoir: from maps to models. *Nat Rev Neurosci* **4**, 266-275.

Hidalgo, A. and Brand, A. H. (1997). Targeted neuronal ablation: the role of pioneer neurons in guidance and fasciculation in the CNS of Drosophila. *Development* **124**, 3253-3262.

Hidalgo, A., Urban, J. and Brand, A. H. (1995). Targeted ablation of glia disrupts axon tract formation in the Drosophila CNS. *Development* **121**, 3703-3712.

Hodges, D. D., Lee, D., Preston, C. F., Boswell, K., Hall, L. M. and O'Dowd, D. K. (2002). tipE regulates Na+-dependent repetitive firing in Drosophila neurons. *Mol Cell Neurosci* **19**, 402-416.

Hoheisel, J. D., Lennon, G. G., Zehetner, G. and Lehrach, H. (1991). Use of high coverage reference libraries of Drosophila melanogaster for relational data analysis. A step towards mapping and sequencing of the genome. *J Mol Biol* **220**, 903-914.

Homyk, T., Jr., Szidonya, J. and Suzuki, D. T. (1980). Behavioral mutants of Drosophila melanogaster. III. Isolation and mapping of mutations by direct visual observations of behavioral phenotypes. *Mol Gen Genet* **177**, 553-65.

Hoogenraad, N. J., Ward, L. A. and Ryan, M. T. (2002). Import and assembly of proteins into mitochondria of mammalian cells. *Biochim Biophys Acta* **1592**, 97-105.

Horowitz, H. and Berg, C. A. (1995). Aberrant splicing and transcription termination caused by P element insertion into the intron of a Drosophila gene. *Genetics* **139**, 327-335.

Hoshino, M., Suzuki, E., Miyake, T., Sone, M., Komatsu, A., Nabeshima, Y. and Hama, C. (1999). Neural expression of hikaru genki protein during embryonic and larval development of Drosophila melanogaster. *Dev Genes Evol* **209**, 1-9.

Hoshino, M., Suzuki, E., Nabeshima, Y. and Hama, C. (1996). Hikaru genki protein is secreted into synaptic clefts from an early stage of synapse formation in Drosophila. *Development* **122**, 589-597.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. *Nature* **425**, 686-691.

Hummel, T., Krukkert, K., Roos, J., Davis, G. and Klambt, C. (2000). Drosophila Futsch/22C10 is a MAP1B-like protein required for dendritic and axonal development. *Neuron* **26**, 357-370.

Hurd, D. D. and Saxton, W. M. (1996). Kinesin mutations cause motor neuron disease phenotypes by disrupting fast axonal transport in Drosophila. *Genetics* **144**, 1075-1085.

Ito, K., Sass, H., Urban, J., Hofbauer, A. and Schneuwly, S. (1997). GAL4-responsive UAS-tau as a tool for studying the anatomy and development of the Drosophila central nervous system. *Cell Tissue Res* **290**, 1-10.

Iyengar, B., Roote, J. and Campos, A. R. (1999). The tamas gene, identified as a mutation that disrupts larval behavior in Drosophila melanogaster, codes for the mitochondrial DNA polymerase catalytic subunit (DNApol-gamma125). *Genetics* **153**, 1809-1824.

Jacobs, H. T. (1997). Mitochondrial deafness. *Ann Med* **29**, 483-91.

Jing, J. and Weiss, K. R. (2002). Interneuronal basis of the generation of related but distinct motor programs in Aplysia: implications for current neuronal models of vertebrate intralimb coordination. *J Neurosci* **22**, 6228-6238.

Johnston, R. M., Consoulas, C., Pfl ger, H. and Levine, R. B. (1999). Patterned activation of unpaired median neurons during fictive crawling in manduca sexta larvae. *J Exp Biol* **202 (Pt 2)**, 103-113.

Johnston, R. M. and Levine, R. B. (1996). Locomotory behavior in the hawkmoth Manduca sexta: kinematic and electromyographic analyses of the thoracic legs in larvae and adults. *J Exp Biol* **199 (Pt 4)**, 759-774.

Kaliss, N. (1939). The effect on development of a lethal deficiency in Drosophila melanogaster: with a description of the normal embryo at the time of hatching. *Genetics* **24**, 270.

Kanamori, T., Nishikawa, S., Shin, I., Schultz, P. G. and Endo, T. (1997). Probing the environment along the protein import pathways in yeast mitochondria by site-specific photocrosslinking. *Proc Natl Acad Sci U S A* **94**, 485-490.

Katoh, K., Shibata, H., Suzuki, H., Nara, A., Ishidoh, K., Kominami, E., Yoshimori, T. and Maki, M. (2003). The ALG-2-interacting protein Alix associates with CHMP4b, a human homologue of yeast Snf7 that is involved in multivesicular body sorting. *J Biol Chem* **278**, 39104-39113.

Kawasaki, F., Felling, R. and Ordway, R. W. (2000). A temperature-sensitive paralytic mutant defines a primary synaptic calcium channel in Drosophila. *J Neurosci* **20**, 4885-4889.

Kidwell, M. (1986). P-M mutagenesis. In *Drosophila: A practical Approach*, (ed. D. Roberts), pp. 59-82. Oxford: IRL Press.

Kim-Han, J. S., Reichert, S. A., Quick, K. L. and Dugan, L. L. (2001). BMCP1: a mitochondrial uncoupling protein in neurons which regulates mitochondrial function and oxidant production. *J Neurochem* **79**, 658-668.

King, F. J. and Lin, H. (1999). Somatic signaling mediated by fs(1)Yb is essential for germline stem cell maintenance during Drosophila oogenesis. *Development* **126**, 1833-1844.

Kitamoto, T. (2001). Conditional modification of behavior in Drosophila by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J Neurobiol* **47**, 81-92.

Koehler, C. M. (2004). The small Tim proteins and the twin Cx3C motif. *Trends Biochem Sci* **29**, 1-4.

Kozak, M. (2000). Do the 5'untranslated domains of human cDNAs challenge the rules for initiation of translation (or is it vice versa)? *Genomics* **70**, 396-406.

Kranz, A., Kinner, A. and Kolling, R. (2001). A family of small coiled-coil-forming proteins functioning at the late endosome in yeast. *Mol Biol Cell* **12**, 711-723.

Kretzschmar, D., Hasan, G., Sharma, S., Heisenberg, M. and Benzer, S. (1997). The swiss cheese mutant causes glial hyperwrapping and brain degeneration in Drosophila. *J Neurosci* **17**, 7425-7432.

Lai, C. H., Chou, C. Y., Ch'ang, L. Y., Liu, C. S. and Lin, W. (2000). Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics. *Genome Res* **10**, 703-713.

Lassing, I. and Lindberg, U. (1985). Specific interaction between phosphatidylinositol 4,5-bisphosphate and profilactin. *Nature* **314**, 472-474.

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Curr Opin Neurobiol* **11**, 475-480.

Laughlin, S. B., de Ruyter van Steveninck, R. R. and Anderson, J. C. (1998). The metabolic cost of neural information. *Nat Neurosci* **1**, 36-41.

Lee, T. and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451-461.

Lemmon, S. K. and Traub, L. M. (2000). Sorting in the endosomal system in yeast and animal cells. *Curr Opin Cell Biol* **12**, 457-466.

Li, X., Xie, Z. and Bankaitis, V. A. (2000).
Phosphatidylinositol/phosphatidylcholine transfer proteins in yeast. *Biochim Biophys Acta* **1486**, 55-71.

Lim, M. L., Lum, M. G., Hansen, T. M., Roucou, X. and Nagley, P. (2002). On the release of cytochrome c from mitochondria during cell death signaling. *J Biomed Sci* **9**, 488-506.

Lin, D. M., Auld, V. J. and Goodman, C. S. (1995). Targeted neuronal cell ablation in the Drosophila embryo: pathfinding by follower growth cones in the absence of pioneers. *Neuron* **14**, 707-715.

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

Lin, H. (2002). The stem-cell niche theory: lessons from flies. *Nat Rev Genet* **3**, 931-940.

Littleton, J. T., Bellen, H. J. and Perin, M. S. (1993). Expression of synaptotagmin in Drosophila reveals transport and localization of synaptic vesicles to the synapse. *Development* **118**, 1077-1088.

Littleton, J. T., Chapman, E. R., Kreber, R., Garment, M. B., Carlson, S. D. and Ganetzky, B. (1998). Temperature-sensitive paralytic mutations demonstrate that synaptic exocytosis requires SNARE complex assembly and disassembly. *Neuron* **21**, 401-413.

Liu, C. (2002). Functional analysis of arouser, an eps8 homologue, during Drosophila development. In *Biological Sciences*, Coventry: University of Warwick.

Liu, L., Wolf, R., Ernst, R. and Heisenberg, M. (1999). Context generalization in Drosophila visual learning requires the mushroom bodies. *Nature* **400**, 753-756.

Lopez, M. C., Nicaud, J. M., Skinner, H. B., Vergnolle, C., Kader, J. C., Bankaitis, V. A. and Gaillardin, C. (1994). A phosphatidylinositol/phosphatidylcholine transfer protein is required for differentiation of the dimorphic yeast Yarrowia lipolytica from the yeast to the mycelial form. *J Cell Biol* **125**, 113-127.

Loughney, K., Kreber, R. and Ganetzky, B. (1989). Molecular analysis of the para locus, a sodium channel gene in Drosophila. *Cell* **58**, 1143-1154.

Manseau, L., Baradaran, A., Brower, D., Budhu, A., Elefant, F., Phan, H., Philp, A. V., Yang, M., Glover, D., Kaiser, K. et al. (1997). GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of Drosophila. *Dev Dyn* **209**, 310-322.

Marder, E. and Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Curr Biol* **11**, R986-96.

Martin, J. R., Ernst, R. and Heisenberg, M. (1998). Mushroom bodies suppress locomotor activity in Drosophila melanogaster. *Learn Mem* **5**, 179-191.

Martin, J. R., Keller, A. and Sweeney, S. T. (2002). Targeted expression of tetanus toxin: a new tool to study the neurobiology of behavior. *Adv Genet* **47**, 1-47.

Martin, J. R., Raabe, T. and Heisenberg, M. (1999). Central complex substructures are required for the maintenance of locomotor activity in Drosophila melanogaster. *J Comp Physiol [A]* **185**, 277-288.

Martin, M. A., Hurd, D. D. and Saxton, W. M. (1999). Kinesins in the nervous system. *Cell Mol Life Sci* **56**, 200-216.

Martin-Blanco, E., Pastor-Pareja, J. C. and Garcia-Bellido, A. (2000). JNK and decapentaplegic signaling control adhesiveness and cytoskeleton dynamics during thorax closure in Drosophila. *Proc Natl Acad Sci U S A* **97**, 7888-7893.

McNabb, S. L., Baker, J. D., Agapite, J., Steller, H., Riddiford, L. M. and Truman, J. W. (1997). Disruption of a behavioral sequence by targeted death of peptidergic neurons in Drosophila. *Neuron* **19**, 813-823.

Meldolesi, J. (2001). Rapidly exchanging Ca2+ stores in neurons: molecular, structural and functional properties. *Prog Neurobiol* **65**, 309-338.

Muqit, M. M. and Feany, M. B. (2002). Modelling neurodegenerative diseases in Drosophila: a fruitful approach? *Nat Rev Neurosci* **3**, 237-243.

Murray, M. J., Merritt, D. J., Brand, A. H. and Whitington, P. M. (1998). In vivo dynamics of axon pathfinding in the Drosophilia CNS: a time-lapse study of an identified motorneuron. *J Neurobiol* **37**, 607-621.

Nakai, M., Takada, T. and Endo, T. (1993). Cloning of the YAP19 gene encoding a putative yeast homolog of AP19, the mammalian small chain of the clathrin-assembly proteins. *Biochim Biophys Acta* **1174**, 282-284.

Nakano, Y., Fujitani, K., Kurihara, J., Ragan, J., Usui-Aoki, K., Shimoda, L., Lukacsovich, T., Suzuki, K., Sezaki, M., Sano, Y. et al. (2001). Mutations in the novel membrane protein spinster interfere with programmed cell death and cause neural degeneration in Drosophila melanogaster. *Mol Cell Biol* **21**, 3775-3788.

Neupert, W. and Brunner, M. (2002). The protein import motor of mitochondria. *Nat Rev Mol Cell Biol* **3**, 555-565.

Niu, S., Antin, P. B. and Morkin, E. (1996). Cloning and sequencing of a developmentally regulated avian mRNA containing the LEA motif found in plant seed proteins. *Gene* **175**, 187-191.

Nusbaum, M. P. and Beenhakker, M. P. (2002). A small-systems approach to motor pattern generation. *Nature* **417**, 343-350.

Nusslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795-801.

O'Connell, P. O. and Rosbash, M. (1984). Sequence, structure, and codon preference of the Drosophila ribosomal protein 49 gene. *Nucleic Acids Res* **12**, 5495-5513.

Odden, J. P., Holbrook, S. and Doe, C. Q. (2002). Drosophila HB9 is expressed in a subset of motoneurons and interneurons, where it regulates gene expression and axon pathfinding. *J Neurosci* **22**, 9143-9149.

O'Hare, K. and Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. *Cell* **34**, 25-35.

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

Palladino, M. J., Hadley, T. J. and Ganetzky, B. (2002). Temperature-sensitive paralytic mutants are enriched for those causing neurodegeneration in Drosophila. *Genetics* **161**, 1197-1208.

Palladino, M. J., Keegan, L. P., O'Connell, M. A. and Reenan, R. A. (2000). A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. *Cell* **102**, 437-449.

Patil, N., Cox, D. R., Bhat, D., Faham, M., Myers, R. M. and Peterson, A. S. (1995). A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nat Genet* **11**, 126-129.

Peck, J. W., Bowden, E. T. and Burbelo, P. D. (2004). Structure and function of human Vps20 and Snf7 proteins. *Biochem J* **377**, 693-700.

Preston, C. R., Sved, J. A. and Engels, W. R. (1996). Flanking duplications and deletions associated with P-induced male recombination in Drosophila. *Genetics* **144**, 1623-1638.

Prokop, A. (1999). Integrating bits and pieces: synapse structure and formation in Drosophila embryos. *Cell Tissue Res* **297**, 169-186.

Puupponen-Pimia, R., Saloheimo, M., Vasara, T., Ra, R., Gaugecz, J., Kurten, U., Knowles, J. K., Keranen, S. and Kauppinen, V. (1993). Characterization of a birch (Betula pendula Roth.) embryogenic gene, BP8. *Plant Mol Biol* **23**, 423-428.

Reeve, S. (2002). Analysis of kisir, a gene required for fertility in Drosophila melanogaster. In *Biological Sciences*, Coventry: University of Warwick.

Renn, S. C., Park, J. H., Rosbash, M., Hall, J. C. and Taghert, P. H. (1999). A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in Drosophila. *Cell* **99**, 791-802.

Rizzuto, R., Brini, M., De Giorgi, F., Rossi, R., Heim, R., Tsien, R. Y. and Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants in vivo. *Curr Biol* **6**, 183-188.

Rizzuto, R., Brini, M., Pizzo, P., Murgia, M. and Pozzan, T. (1995). Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Curr Biol* **5**, 635-642.

Roberts, D. (1986). Drosophila: A practical approach. Oxford: IRL press.

Rohrbough, J., O'Dowd, D. K., Baines, R. A. and Broadie, K. (2003). Cellular bases of behavioral plasticity: establishing and modifying synaptic circuits in the Drosophila genetic system. *J Neurobiol* **54**, 254-271.

Roman, G. and Davis, R. L. (2001). Molecular biology and anatomy of Drosophila olfactory associative learning. *Bioessays* **23**, 571-581.

Roman, G. and Davis, R. L. (2002). Conditional expression of UAS-transgenes in the adult eye with a new gene-switch vector system. *Genesis* **34**, 127-131.

Roman, G., Endo, K., Zong, L. and Davis, R. L. (2001). P[Switch], a system for spatial and temporal control of gene expression in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **98**, 12602-12607.

Roseman, R. R., Johnson, E. A., Rodesch, C. K., Bjerke, M., Nagoshi, R. N. and Geyer, P. K. (1995). A P element containing suppressor of hairy-wing binding regions has novel properties for mutagenesis in Drosophila melanogaster. *Genetics* **141**, 1061-1074.

Ryckebusch, S. and Laurent, G. (1994). Interactions between segmental leg central pattern generators during fictive rhythms in the locust. *J Neurophysiol* **72**, 2771-2785.

Saint-Amant, L. and Drapeau, P. (2001). Synchronization of an embryonic network of identified spinal interneurons solely by electrical coupling. *Neuron* **31**, 1035-1046.

Sambrook, J., Fritsch, EL, Maniatis, T. (1989). Molecular Cloning - A laboratory Manual. Cold Spring Harbor Laboratory Press: Cold Spring Harbor Laboratory Press.

Schiavo, G., Matteoli, M. and Montecucco, C. (2000). Neurotoxins affecting neuroexocytosis. *Physiol Rev* **80**, 717-766.

Schmidt, H. A., Strimmer, K., Vingron, M. and von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502-504.

Schnapp, B. J. (1997). Retroactive motors. *Neuron* **18**, 523-6.

Schon, E. A. and Manfredi, G. (2003). Neuronal degeneration and mitochondrial dysfunction. *J Clin Invest* **111**, 303-312.

Sese, J., Nikaidou, H., Kawamoto, S., Minesaki, Y., Morishita, S. and Okubo, K. (2001). BodyMap incorporated PCR-based expression profiling data and a gene ranking system. *Nucleic Acids Res* **29**, 156-158.

Shaver, S. A., Riedl, C. A., Parkes, T. L., Sokolowski, M. B. and Hilliker, A. J. (2000). Isolation of larval behavioral mutants in Drosophila melanogaster. *J Neurogenet* **14**, 193-205.

Shibata, N., Arita, M., Misaki, Y., Dohmae, N., Takio, K., Ono, T., Inoue, K. and Arai, H. (2001). Supernatant protein factor, which stimulates the conversion of squalene to lanosterol, is a cytosolic squalene transfer protein and enhances cholesterol biosynthesis. *Proc Natl Acad Sci U S A* **98**, 2244-2249.

Shigenaga, A., Kimura, K., Kobayakawa, Y., Tsujimoto, Y. and Tanimura, T. (1997). Cell ablation by ectopic expression of cell death genes, ced-3 and Ice, in Drosophila. *Dev Growth Differ* **39**, 429-436.

Siekhaus, D. E. and Fuller, R. S. (1999). A role for amontillado, the Drosophila homolog of the neuropeptide precursor processing protease PC2, in triggering hatching behavior. *J Neurosci* **19**, 6942-6954.

Sink, H. and Whitington, P. M. (1991). Early ablation of target muscles modulates the arborisation pattern of an identified embryonic Drosophila motor axon. *Development* **113**, 701-707.

Skaer, H. (1993). The Alimenary Canal. In *The Development of Drosophila melanogaster*, vol. 2 (ed. M. a. M. A. Bate, A), pp. 941-1012: Cold Spring Harbor Laboratory Press.

Spradling, A., Drummond-Barbosa, D. and Kai, T. (2001). Stem cells find their niche. *Nature* **414**, 98-104.

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., Misra, S. and Rubin, G. M. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics* **153**, 135-177.

Spradling, A. C., Stern, D. M., Kiss, I., Roote, J., Laverty, T. and Rubin, G. M. (1995). Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. *Proc Natl Acad Sci U S A* **92**, 10824-10830.

Stauffer, D. R., Howard, T. L., Nyun, T. and Hollenberg, S. M. (2001). CHMP1 is a novel nuclear matrix protein affecting chromatin structure and cell-cycle progression. *J Cell Sci* **114**, 2383-2393.

Stowers, R. S., Megeath, L. J., Gorska-Andrzejak, J., Meinertzhagen, I. A. and Schwarz, T. L. (2002). Axonal transport of mitochondria to synapses depends on milton, a novel Drosophila protein. *Neuron* **36**, 1063-1077.

Strauss, R. (2002). The central complex and the genetic dissection of locomotor behaviour. *Curr Opin Neurobiol* **12**, 633-638.

Strauss, R. and Heisenberg, M. (1993). A higher control center of locomotor behavior in the Drosophila brain. *J Neurosci* **13**, 1852-1861.

Sun, Q. and Dale, N. (1998). Developmental changes in expression of ion currents accompany maturation of locomotor pattern in frog tadpoles. *J Physiol* **507 (Pt 1)**, 257-264.

Suster, M. L. and Bate, M. (2002). Embryonic assembly of a central pattern generator without sensory input. *Nature* **416**, 174-178.

Suster, M. L., Martin, J. R., Sung, C. and Robinow, S. (2003). Targeted expression of tetanus toxin reveals sets of neurons involved in larval locomotion in Drosophila. *J Neurobiol* **55**, 233-246.

Suzuki, D. T., Grigliatti, T. and Williamson, R. (1971). Temperature-sensitive mutations in Drosophila melanogaster. VII. A mutation (para-ts) causing reversible adult paralysis. *Proc Natl Acad Sci U S A* **68**, 890-893.

Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. and O'Kane, C. J. (1995). Targeted expression of tetanus toxin light chain in Drosophila specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14**, 341-351.

Tabuchi, K., Sawamoto, K., Suzuki, E., Ozaki, K., Sone, M., Hama, C., Tanifuji-Morimoto, T., Yuasa, Y., Yoshihara, Y., Nose, A. et al. (2000). GAL4/UAS-WGA system as a powerful tool for tracing Drosophila transsynaptic neural pathways. *J Neurosci Res* **59**, 94-99.

Takahashi, I., Nobukuni, T., Ohmori, H., Kobayashi, M., Tanaka, S., Ohshima, K., Okada, N., Masui, T., Hashimoto, K. and Iwashita, S. (1998). Existence of a bovine LINE repetitive insert that appears in the cDNA of bovine protein BCNT in ruminant, but not in human, genomes. *Gene* **211**, 387-394.

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

Theodosiou, N. A. and Xu, T. (1998). Use of FLP/FRT system to study Drosophila development. *Methods* **14**, 355-365.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876-4882.

Thummel, C. a. P., V. (1992). New pCaSpeR P element vectors. *Dros. Info. Service* **71**, 150-150.

Toby, G. G. and Golemis, E. A. (2001). Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. *Methods* **24**, 201-217.

Toivonen, J. M., O'Dell, K. M., Petit, N., Irvine, S. C., Knight, G. K., Lehtonen, M., Longmuir, M., Luoto, K., Touraille, S., Wang, Z. et al. (2001). Technical knockout, a Drosophila model of mitochondrial deafness. *Genetics* **159**, 241-254.

Toma, J. G., Akhavan, M., Fernandes, K. J., Barnabe-Heider, F., Sadikot, A., Kaplan, D. R. and Miller, F. D. (2001). Isolation of multipotent adult stem cells from the dermis of mammalian skin. *Nat Cell Biol* **3**, 778-784.

Torroja, L., Ortuno-Sahagun, D., Ferrus, A., Hammerle, B. and Barbas, JA. (1998) scully, an essential gene of Drosophila, is homologous to mammalian mitochondrial type II L-3-hydroxyacyl-CoA dehydrogenase/ amyloid-β peptide-binding protein. *J Cell Biol* **141**(4), 1009-1017.

Truscott, K. N., Brandner, K. and Pfanner, N. (2003). Mechanisms of protein import into mitochondria. *Curr Biol* **13**, R326-337.

Tsien, R. Y. (1998). The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544.

Tu, J., Vallier, L. G. and Carlson, M. (1993). Molecular and genetic analysis of the SNF7 gene in Saccharomyces cerevisiae. *Genetics* **135**, 17-23.

Vanden Berghe, P., Kenyon, J. L. and Smith, T. K. (2002). Mitochondrial Ca2+ uptake regulates the excitability of myenteric neurons. *J Neurosci* **22**, 6962-6971.

Verheyen, E. M. and Cooley, L. (1994). Profilin mutations disrupt multiple actin-dependent processes during Drosophila development. *Development* **120**, 717-728.

Vidal, M. and Legrain, P. (1999). Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Res* **27**, 919-929.

Waddell, S. and Quinn, W. G. (2001). What can we teach Drosophila? What can they teach us? *Trends Genet* **17**, 719-726.

Wallace, D. C. (1999). Mitochondrial diseases in man and mouse. *Science* **283**, 1482-1488.

Wang, J. W., Soll, D. R. and Wu, C. F. (2002). Morphometric description of the wandering behavior in Drosophila larvae: a phenotypic analysis of K+ channel mutants. *J Neurogenet* **16**, 45-63.

Wang, J. W., Sylwester, A. W., Reed, D., Wu, D. A., Soll, D. R. and Wu, C. F. (1997). Morphometric description of the wandering behavior in Drosophila larvae:

aberrant locomotion in Na+ and K+ channel mutants revealed by computer-assisted motion analysis. *J Neurogenet* **11**, 231-254.

Wang, S. and Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. *Nature* **369**, 400-403.

Wiedemann, N., Frazier, A. E. and Pfanner, N. (2004). The protein import machinery of mitochondria. *J Biol Chem* **279**, 14473-14476.

Wieschaus, E., Nusslein-Volhard, C. and Kluding, H. (1984). Kruppel, a gene whose activity is required early in the zygotic genome for normal embryonic segmentation. *Dev Biol* **104**, 172-186.

Wilson, C., Pearson, R. K., Bellen, H. J., O'Kane, C. J., Grossniklaus, U. and Gehring, W. J. (1989). P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. *Genes Dev* **3**, 1301-1313.

Xiong, W. C. and Montell, C. (1995). Defective glia induce neuronal apoptosis in the repo visual system of Drosophila. *Neuron* **14**, 581-590.

Xiong, W. C., Okano, H., Patel, N. H., Blendy, J. A. and Montell, C. (1994). repo encodes a glial-specific homeo domain protein required in the Drosophila nervous system. *Genes Dev* **8**, 981-994.

Xu, T. and Rubin, G. M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. *Development* **117**, 1223-1237.

Yamamoto, D., Jallon, J. M. and Komatsu, A. (1997). Genetic dissection of sexual behavior in Drosophila melanogaster. *Annu Rev Entomol* **42**, 551-585.

Yassin, L., Gillo, B., Kahan, T., Halevi, S., Eshel, M. and Treinin, M. (2001). Characterization of the deg-3/des-2 receptor: a nicotinic acetylcholine receptor that mutates to cause neuronal degeneration. *Mol Cell Neurosci* **17**, 589-599.

Yeh, E., Gustafson, K. and Boulianne, G. L. (1995). Green fluorescent protein as a vital marker and reporter of gene expression in Drosophila. *Proc Natl Acad Sci U S A* **92**, 7036-7040.

Younossi-Hartenstein, A., Salvaterra, P. M. and Hartenstein, V. (2003). Early development of the Drosophila brain: IV. Larval neuropile compartments defined by glial septa. *J Comp Neurol* **455**, 435-450.

Zhang, Y. Q., Roote, J., Brogna, S., Davis, A. W., Barbash, D. A., Nash, D. and Ashburner, M. (1999). stress sensitive B encodes an adenine nucleotide translocase in Drosophila melanogaster. *Genetics* **153**, 891-903.

Zhou, L., Schnitzler, A., Agapite, J., Schwartz, L. M., Steller, H. and Nambu, J. R. (1997). Cooperative functions of the reaper and head involution defective genes in the programmed cell death of Drosophila central nervous system midline cells. *Proc Natl Acad Sci U S A* **94**, 5131-5136.

Zucker, R. S. (1999). Calcium- and activity-dependent synaptic plasticity. *Curr Opin Neurobiol* **9**, 305-313.

Zuker, C. S. (1992). Phototransduction in Drosophila: a paradigm for the genetic dissection of sensory transduction cascades. *Curr Opin Neurobiol* **2**, 622-627.

Zuker, C. S. (1996). The biology of vision of Drosophila. *Proc Natl Acad Sci U S A* **93**, 571-576.

Zuo, J., De Jager, P. L., Takahashi, K. A., Jiang, W., Linden, D. J. and Heintz, N. (1997). Neurodegeneration in Lurcher mice caused by mutation in delta2 glutamate receptor gene. *Nature* **388**, 769-773.

Appendix 1.

LOCUS slmo
DEFINITION 6kb BamHI genomic fragment (6176bp)
ORGANISM Drosophila melanogaster
REFERENCE Carhan (1999)
AUTHORS Ahmet Carhan

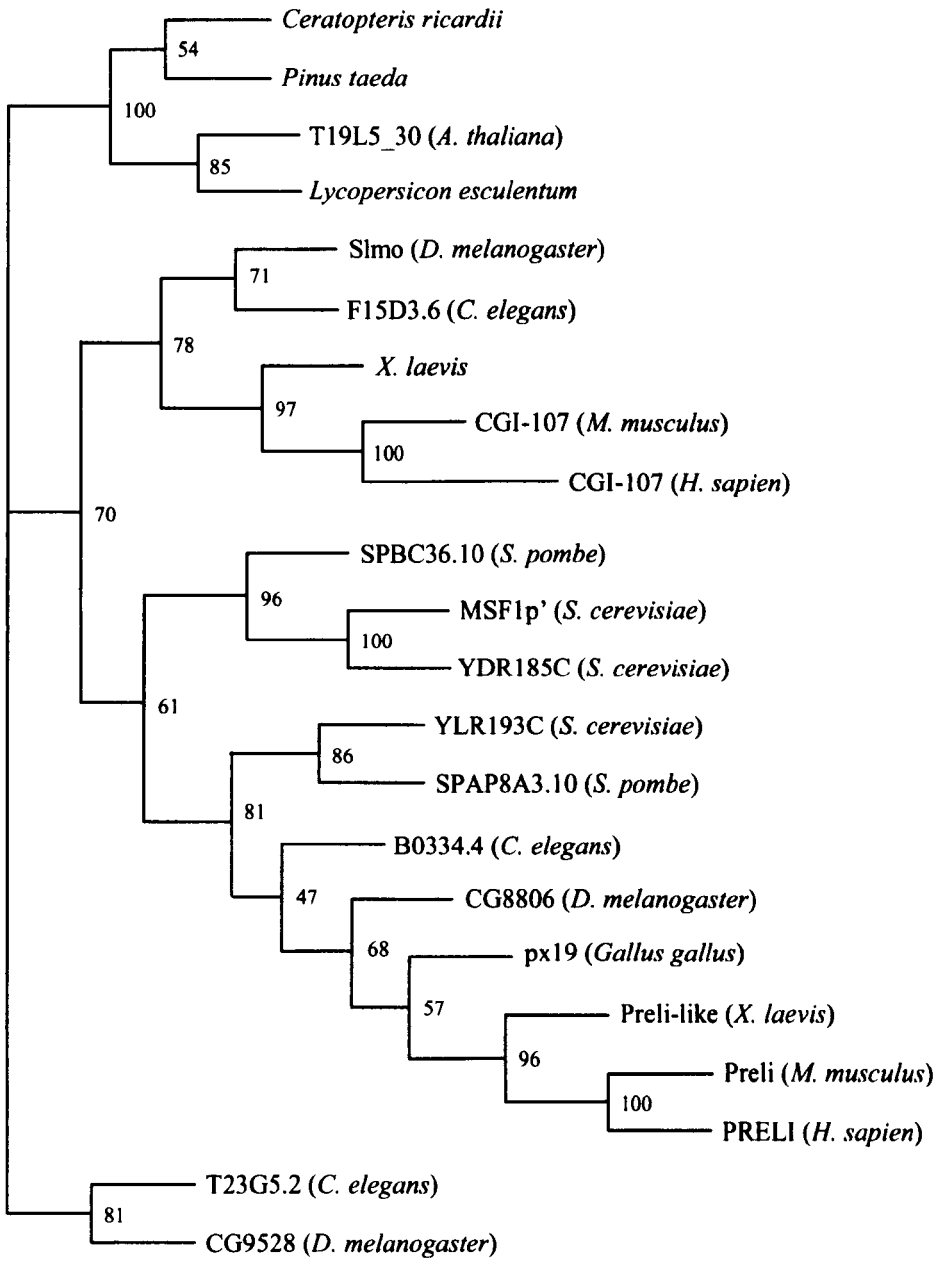
```
   1 CCGGACATCG GGATAGAAGA CGCTGTTCGC GATGGCAGGA GCGGGCGTGG GACGGCGCTG
  61 GGCCCTTGGG CGTTCCGCTG GGGCTTCCGT TTCCGTTTGC GTCTGCAAGT AGAAAGTGCC
 121 TGTAGCCTGG CTTTGGGAAT ATGCAACAGA TTTTGCAGAG AATGTCGCTT AGTGGCATTC
 181 CGCTGACTAA ATTAGATCGG CTCTGGGTAA ATGGCAACGA TAAGCTGGCA ATCGAAGTTC
 241 GGATTATACT GAGAGCACTC TGCTCTTATC TCTAGCTTCT ACTAACGGTA GACTCACATT
 301 GTATTGGGAC AAACAAACAG CTGGCTTTGG GGCCGAAAAA TGCCGTATGA TTGTATTTAG
 361 GCAAAGTTGT AGTTCTGCGG TTATCTATTG GCCAAGTTTC CCACTTAAAA TTAGCTTTTA
 421 AAACTCAAAG CCCCGAACAT GGTGATTCAT TTCGTCTTTC AATTTGGTCA CTTCCGTTCA
 481 GTTAATTTTT GCATTACTAT ACCCACAGCA TTGTACATCC ATTTTAAGAG TACGTGTTCA
 541 TACCTTGCTT GCTCATTAGC TTCTCCTGCT GCAGGCGCGC GTTGCTCGAT CCGAAGTATC
 601 CGTTGATAAA TGGCAGGGGC GACATGACTA CTCCGCCCCT CAGCGAACTG TCTGGTGGAG
 661 GAGGAGCGGG TCTGTTGAAT GGATTTGTCC GCCTGGGCGG CGGTCTAGGC TTCGCACTCT
 721 CCGATGCATT TTGTGGTCCG TCCAAGGCGC TGCCGAGCAT CTCTGTAGGT TAAAGAACGC
 781 AGCGGATTAG TGCACAGGGA TTGTAGGATT AGCCATTGTG CAGTCTCACT GTTTCTTGCG
 841 TCCTCCATCT CTTGCAACTG GCGTTTTAGC TCGGCCAGGA TGATGCGCTT ATACTCGTAG
 901 CCGGCGCATG TTAGCGCCTT CATCCACGTC TCCATGCTGT CTTGGTTTTC GGCGGCGAGG
 961 ATGTAGGTGC GATTGCCGTT GAAGGCGATC TCGAAGCAGT AGTTGTCCAC CTCGTTAGAC
1021 AGCTCAATGG TGCAGCCCTC GACAATGATG AGGCCCAAGG GCTCCTTGTC CAAGCGCGAC
1081 TCGAAGTAGA AGAGCAGGTT GCCTTTCAGC ACGAAGTAGC GTCGCTGGAA GGCCTTGTTG
1141 ACCTCGCCCC GTTTGTTGAG GAAGCCTGCA AGGCGTATCG GTGTGTGATA AATGGATTAT
1201 GATTATTAAT AGACGAGCGA AAGGGCTCGC TTGCATACCT TCCATGTCGA AGGGCGGCGT
1261 TCTCGCGAAC ACGTACAGGT TCTTCTCGTT GATCTTCATG GCTGCACTGT CTCGGGCTGC
1321 ATTCGAAGCA AATATGGGTG TCCTATTTCC GCCTTTCGTC TCTTATTATT GTGATGAACA
1381 GGTAAAAAAA ATGCAAAGCG AGAAACAATT GACGCTGCGC GAACAGCTGA TGGCCAGTGC
1441 TGCCACACCT CCACCGGTTA TCGATTGCTG GGTGCGGAAT TTCTAGTTTT CAATGTAGAT
1501 TGCAAACATA AACGAATTTT GAATCCTGCG AAATGTCGTA CTTACTCACG GAAATCGCGC
1561 TGGAGATGCT GGGCTACAAG TTCTACGACA CCAATGGCAC ATTCCACCTG ACCAATTTTC
1621 AAAACAACTC CACGCTCAGC GAAACGCATC CCGATGAGCC CTCCCTGCAG GAATTCATTG
1681 AGTACACAAA TGAGGAGCTG AAGCAAGACG CGCAACAGCA GGAGCCGCAC ACGCTGGCGG
1741 TGCAGGAGGA GTCGTACCTG GATGCCCCCA CCATACCCGA GTCACGCTAT GCGCGGCGCA
1801 CGCCGGCCAA GGCGCTGGCC AAGATAATGG ACTACGAGAG CGAGAAGATA GTCAACCAGA
1861 TCAGCACTCG TCTGGAGGGC TCGGTGTCAC GTCGGGCGGC GGCGGACAGT GACGAGCAGC
1921 TGGCCGAGAT CATCAACGAG GAGAGGAAAC TAATCGAGGC GGATGCCGAG GTAATACACA
1981 ACTTTCGGGA GCTCGTGGAG AGCCGACTAC AGCGATTGGA TAAACGGCCC TACGAGAAAT
2041 ACGTACGTAC CTTCGGGCCG CAATAGGAAG CTGCACGTGC AATGTCAAAG TGCCGCCATT
2101 ACTTTTAGAA TTAATTTAGT TTATAAATTT AATAAATAAG TTACCGCTAC AGGGTGCCTG
2161 CGGCACTTTC ATTTTCTCTT ACTGCGCCAG TGTGACCACC CTAGCCCATG CACTGGCCAG
2221 CGATAGCCCC ATCTATCGGA GCCACTGTCA TCTCTAGCAG CACAGCACGA GTAAAGGTCA
2281 GCTCTCGTAA CTTGTCTCGG AAATATTCAA TAATTCCGAG ATTGCCGTCA AAATTCGAGT
2341 CGTGTACTTC AGCTAACGCT CCCCCTGTTT AATTCCGTTC CGTCTGCCCC CGCAGGCTCA
2401 GCTCAGCCCG CATCTCGGCT GCCAATTAAG GATATAAAGC CCTCCCCGGA CGGTGACGAC
2461 AACGGCGCTT GCATAATGAA AATCTGGACA TCGGAGCACA TATTCAACCA CCCGTGGGAG
2521 ACGGTCACGC AGGCGGCGTG GCGCAAGTAC CCCAATCCGA TGACGCCCTC CATCATTGGC
2581 ACGGACGTGG TCGAGCGCCG TGTGGTGGAC GGCGTCCTGC ACACGCATCG CCTGGTGCAG
2641 TCCAAGTGGT ACTTCCCCAA GTGGACGCAC GCGCTGATCG GCACGGCCAA AACTTGTTTC
2701 GCCAGCGAAC GCTCCACGGT GGACCCGGAG CGCAAACAGA TGGTGCTGAA GACAAACAAC
2761 CTCACCTTCT GCCGCAACAT CAGCGTGGAC GAGGTGCTCT ACTATGAGCC GCATCCGTCG
2821 GATGCCAGCA AGACGCTGCT CAAGCAGGAG GCCACCGTGA CCGTCTTTGG TGTGCCGCTG
2881 TCCCACTACA TGGAGGACCT GCTCACCTCG ACCATTAGCA CAAATGCCGG CAAGGGACGC
2941 CAGGGCCTGG AGTGGGTGAT CGGGCTGATC AACACGGAGG TCAAGGGCAT CGCTCGCGGC
3001 ACCGACGAAC TTCTGCACAA CACACGGCGC TCCATCGACG AGGTGACGGA GAGCGCGCGG
3061 AAGAGCATGG ACGAGATCAG TGCGCAGGCC GCCAAGGCGG CGAAGGCGAT GCACATTACG
3121 TAGGTGTCGA CAAGGCGACA GTGCAAGGTC GCGCGGGGGC TGTTCACTAG TGAATTATTT
3181 AGATGGCGCG CGGTCGTCGT TTTAGTACAA CTAATTAATC GATGGCACGA TGGACATGGA
3241 CAGCAACATG TGGGACAGCA ACACCTTCCT CTATTCTCAA GCCATTCAAT GCATATATTT
3301 TATAATTATT ACTCCCCTTC GAACGAATCG ACTGCTAATG AGTCATGTCG ATAAAGAAGT
3361 GCCCCCGAAT GGCTGAACTG GAAGGAACTC GCGCTAATGC TGGCCCCCAT TCATGCAGAT
3421 CAATATATAT ATAAACCTAA AGTACATAGT CAATGTAAGC TTTGTAAGCC GCCGCAGATT
3481 TCGAAGAACG AGAGGAGCAA TTAAACGACA TATTTGCCTA CGACCAGCCA GTGTCAAGTT
3541 CAATTGTTTA ATTTTAATTC GCAGCACACT TGTACTGCGT CTTTTTTTTT TGCATTTGCA
3601 TGACGTTCGC ATGTGCAACT GCAGCAATTT TAAGCATTCT GTGTATAAAT AAAACCCAAG
3661 AAAGTCGATT TGTGGGTTTT TTCTCTACAT ACATACTTGA TTAGCGATTT TTGGCGTCTG
3721 TTTTATTCGC TTATCATCAA AAGCCCACGT GCCTAGTGGC TCAACAACTG TGTACCTAGT
```

3781 TAGCATCGCA GACATGGAAA CTGATATGTC CCTTATCTGG TTCAACAGTT GCGAAGAATG
3841 TGATTGAAGG GGAAAGTGGG TGCTTGGAAA CAAGTTGATG ATCAAACAGC AGGTGTTCTA
3901 GGGACAAGGT GAAATACTTT TACCTAACTT ATCAAAAGAA GTAACTTATT TATTTAGCAG
3961 TTTTCGTAAG CTTTTAATCA AAATAAACAA GTAATCCTTA CGAATTACAA TTAAACTATT
4021 TGGTTCATTG AACCCGTGTA GCATTAACCC TTGCAAGTTA CGCAATGGGT ACAATTTACG
4081 AAAGCTTTTA AAGCTCAATT CGCCTAAGGG TTAATTTGAA TTCCAGCGGC AGTCGCTGTG
4141 CAGCACTGGC CAACCGAACG CTGTGGTGGA TTGAGCCGCG AAATGAGGGG AAACCGAGCT
4201 GGAGCCTCAA AACGATCGTA ATTTGCAGCT AAGACGCACG TGTGCTCCGC TCAGTGTTAG
4261 TGTTATTGTG CTTCAATTTG GGGTATTTGG CGTTAATTAT TTACCTAAAT TCCGACCATT
4321 CAGCAACGCA GGTTTCCGTT CGTAATGGAA AATATGTTTA CGCCGGCTTA TCAGCAAACT
4381 CCCAGCCGGG GAAAAACAAA AACAAGTTGT GACGCCAGTG CAGCGATTGC TGCAGGCGAT
4441 TCCATCGATT TCGCTCGGCT GTAACCTGAT CCGCTGCTGT GGGTTGTGGG CTGTGTGGCT
4501 GTGGTATAGG TGCCGCCAGT TGATAACCCC CAGCGTGGGC ACGGACCCAC TGGTCGTCAA
4561 CCGCTTCGCA ACTCCCAAGA TCATTCTCGA TTCGACGAGC AAGGAGTGCA CACTGTCTGT
4621 ACTCAATCCG CATTTTTCGT TTCTTGCAGC CCCACGACAT GGTCATCCTT CGCTACTACG
4681 ATGTGCAGGC CCACTCCGCG GCCGAGGAGG AGAGTGTCCT GCGTCGGTTG CGCGAGGAAG
4741 ACGGCGCCGT GGTGTCCGTG CGCATGGAGC GCTGCTATCA TCTGGAGTAC AGCGCTCAGG
4801 CAGAGCACTC ACTGGCCCTG GACGAGCTGC TGGTGTGGCT GGTCAAGCAA CCGCTGAGCA
4861 AAGGCCAGAG CTTGTCCAGG CAACCTGCCC TGCAGTCAAC TGGGTCGAGT CAGTTGCTCC
4921 TGGAGATCGG GCCGCGTTTT AACTTCTCCA CGCCGTACTC CACGAACTGC GTGAACATAT
4981 TCCAGAATCT CGGGTACTCA GAGGTGCGTC GCATGGAAAC CTCCACCCGC TATCTGGTTA
5041 CTTTTGGCGA GGGATCAAAG GCGCCGGAGG CAGCCAGGTT TGTTCCTCTG CTCGGTGACC
5101 GCATGACCCA GTGCTTGTAC ACCGAGGAGA ATACCCCCAA GGCGAGCTTT GACGAGCAGC
5161 TACCTGAGCG CCAGGCCAAC TGGCATTTCG TGCCCGTTTT GGAGGAGGGT AGGGCGGCAC
5221 TGGAGCGGAT TAATCAGGAG CTGGGCTTAG CCTTCAACGA CTACGATTTG GACTACTACC
5281 ACGACTTGTT TGCCAAGGAG CTGGGCCGCA ATCCCACCAC TGTGGAGCTC TTCGATTGCG
5341 CCCAGAGCAA CAGTGAGCAC TCGCGCCACT GGTTTTTCCG CGGACGTATG GTGATCGACG
5401 GCGTGGAGCA GCCCAAGTCG CTGATTCGCA TGATCATGGA CACGCAGGCC CACACGAACC
5461 CGAACAACAC CATTAAGTTC AGCGACAACA GCAGTGCCAT GGTGGGATTC GATCACCAGA
5521 CCATAGTTCC GTCCTCCGTG GTCGCTCCCG GCGCAGTGCG TCTGCAGAGC GTGCAGTCTG
5581 ACCTGATTTT CACGGCGGAG ACCCACAACA TGCCCACTGC AGTGGCGCCT TTCAGCGGAG
5641 CCACCACCGG CACTGGCGGA CGACTGCGTG ATGTCCAGGG CGTGGGCAGA GGAGGCGTGC
5701 CGATCGCCGG CACCGCTGGC TACTGTGTTG GCGCTCTTCA CATTCCAGGT GAGCATTGCT
5761 TTTTTTTTGA GTATTTAGTT CCTATTAGTT TTTATTATAC ATCACGACTT TCTTGTCGTC
5821 GCAGGTTACA AACAGCCGTA CGAGCCTTTG GACTTTAAAT ACCCTGCGAC GTTTGCGCCC
5881 CCACTTCAGG TGCTCATTGA GGCGAGCAAT GGCGCCTCCG ACTACGGAAA TAAGTTCGGC
5941 GAGCCAGTGA TCTCTGGTTT TGCCCTCTCC TATGGACTGA ACAGTGCTGC CGATGCCAGC
6001 CAGCGGGATG AGTACGTGAA ACCGATCATG TTCAGCGGTG GCTTGGGCAC CATGCCCGCT
6061 ACCATGCGCG AGAAGCTGCC GCCGGCACGC GGTCAGTTGC TGGCCAAGAT CGGGGGCCCA
6121 GTGTACAGGA TTGGAGTGGG TGGCGGCGCC GCAAGCTCTG TGGAGATACA AGGATC

Appendix 2: List of ESTs with similarity to *slmo*.

Species	EST Name	GenBank Accession
Ancylostoma caninum	pa35e01.y1	BE352448
Anopheles gambiae	4A3B-AAE-D-07-R	AJ284971
A. thaliana	AV549314	AV549314
Bombyx mori	AU000429	AU000429
Bos taurus	1A10C08	BG467246
Bos taurus	AV616134	AV616134
C.elegans	C71074	C71074
Ciona intestinalis	AV672110	AV672110
Ceretopteris ricardii	Cri2_8_P09_SP6	BE643491
Danio rerio	fb04b02.y1	AI330985
Dictyostelium discoidum	AU076350	AU076350
H.sapien	601437019F1	BE893122
H.sapien	601561340F1	BE728046
H.sapien	EST390625	AW978516
Lycopersicon esculentum	EST303596	AW218413
Meloidogyne incognita	ra44b02.y1	AW829756
M.musculus	ue18h06.y1	AI036726
M.musculus	vb09b11.r1	AA272853
Oryzias latipes	AV668326	AV668326
Oryzias latipes	AV670130	AV670130
Oryza sativa	AU057724	AU057724
Phytophthora sojae	7-11F-MY	BE583294
Pinus taeda	ST28D02	AW043024
Solanium tuberosum	EST427241	BE923472
Strongyloides stercoralis	kp55g09.y1	BE224559
Xenopus laevis	db31a05.y1	BE026022
Xenopus laevis	b138f03.w1	

Appendix 3. Phylogeny of the Slmo family of proteins (2000).



Phylogeny was generated from a ClustalX protein sequence alignment by Neighbor-joining using the PHYLIP package (Felsenstein, 1989). Bootstrap values were calculated using 100 replicate data sets. Evolutionary distances were calculated using the Dayhoff PAM matrix.

Appendix 4.

LOCUS *prel* (CG8806)
DEFINITION *prel* gene region (2041bp)
FLYBASE ID FBgn0033413
ORGANISM *Drosophila melanogaster*

```
   1 GCACACGCAC CATCTGGCAA CTTTGCGACC AGATGTTTCC GGATGGCCTT TGAACAGGCG
  61 TAAATTGTAT CCAAATTGGC AGCAAACCCT TGGAAAGCAA GCAATGAAAC CGTGTACAGA
 121 CAGGAGAAGG GGGCACTGAA GGTCGAGGCA ACGCAGGAAT CGAGGGATAT CGCCAGCAAA
 181 TTGCCAGAAA TTGAAGGTAA CGAACCCGAA GCACCCTGTG CCGGAGTTTG TGACGCAGTT
 241 ACACAACTAA ATTCTTAGTA GGCGCCCAAA GTCGGGGCAA CAACAAGCTG CTAGCTGAAT
 301 CAACGTGTGC CAGCTAATTG GTGCCAATCA CACTGCACTC TGGAGCTCTA AAACCGTAAG
 361 AGACCCCCGT TCTTGGACCC CCTCGCAAAG GTACGTCCAT CGATTCGGCG ATCTATTTTT
 421 AGCCGGAAAC ATGCGATTGC TATTGATGAT TATCGGTTTA TTTTTCAGAT CAGGGACACG
 481 GCGGCTTTCA CTTTCGAACA CATCGGCGAG TTCATTGGTC CACCTTCAAC ATGGTCGTGG
 541 CAGCATCCAC CTGCCGCACG GAGACGGTCT TCGATTACAG CTGGATGAAC GTGGTGGTCG
 601 CCTACTGGAA CCGCTACCCG AATCCCTCAA GCACGCACGT CCTCACCGAG GACACTATTC
 661 AGCGGGAGGT GCGCGATGGC AAGCTCTTCT CGCGCCGGCT GCTCTCCAAG ACGAATCCGG
 721 TGCCGAAGTG GGGAGCTCGC TTCTACAACA ATGTGCCGGT CAAGATTGTT GAGGACTCCG
 781 TGCTGGACCC GGTTAAAAAG ACATTCACCA CTTTCACCAG GAACCTGGGA ATGACCAAGA
 841 TAATGGCAAG TGATCGCACA CCTACTACGA TTGCTTTAAA TAGGTCACAT ATCAATTTAA
 901 AGAAATGTCT AGAAGAATTC AGCCACTAGC TGTTGGGCAT GACCGAAATA CCGGCATAAT
 961 ACAAAAGCAT AAGTGTTACC TAACAGATCG ATGGTCAATG CTGTGGCATT GTTTGTTATA
1021 TTTAACCCAT GAAACGCATA CAGAACATCG GGATCCGGTT AAGATTATTG TTCGATTTAT
1081 TTTACTGCAC CCTTTTAGTG AATATGTGAA ATATGTTACT TAATCTTTTC AAGTTTAATT
1141 TAATCCTAAT TAATGATCAT TGTTTTTTCC TTTTAGAAAG TCGATGAAAT TGTCGTCTAC
1201 AGCGAGCAGA AGGATGGTAG TACTCTGGCG GTGCGAAGAG CATACATTAG TTCGCAGGTG
1261 TTCGGATTCT CGCGGGCCAT TCGCGCCTTC GGCATCGAAC GGTTTAAGGC CAATGGCAAC
1321 AAGGCCTCCA ATGGATTCAA CTACGTGCTG CGGCGCATGT TTCCGGACAG TTTGGTTGGT
1381 GGAGGACATC ACCAGCACGC GGTGACGACG ACTTCGCCAG CTGGAGAACT TCCTGCAACG
1441 ACCATCACCG TATCCACAAC AAACGGCAGT CTCAACAACC AAGGAGCTCT CAAGAGCGCC
1501 GCAAAAGTGG GCTACGAGTT CTTCAAGAGT CATGCCTCCA AGATTGCGCA ACTGTTTTCT
1561 GTCAAGAACT GAGGAACTTC TAGTCTGTAC TTTTTGGAAT AGATGTCGAA TGGGACCGAA
1621 GATTTTAATG CTTTCGGCTC CGGGAGTTAA GGAGTTTGTG GATTTGACCA GGACGGAATT
1681 AGAATGACGA GCGGAGGAGT CGTGTCGCCG GGTGTTTTTA AACATTTATT CCAAAGCAAA
1741 TCCAACTACA TACTTTGCTG CATTTCATTC CGTGTGTGCG CGAATCAAAC GGCTGATCTG
1801 AACGGAGACA CAATTTGCTG AACACATTTC CTGCATGCTA CCAAATTCTA TTATTTATCG
1861 TTACTTATTT GCACTACATT ATTATGTATC ATTCTGTTAG GTCAAGAGCA GGAAATGAAA
1921 AGCCGTGTCA AAATTGGTAT TGCTATGTAA TATGTTGTAT ATTTTTATCA AAATATAAGA
1981 AAACCAACAC TTGTAGTGGT TAAAAGTGTA GTTATTAAAA TAAACTTAGT CTGTTAGAAA
2041 T
```

Appendix 5.

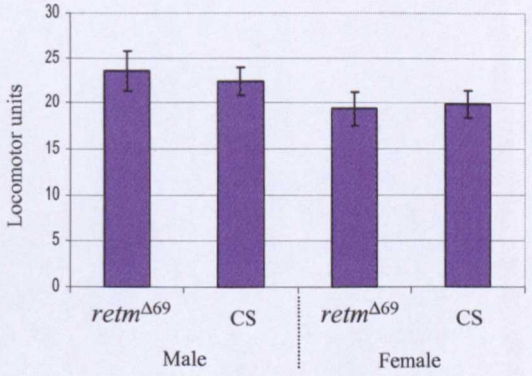
LOCUS *retm* (CG9528)
DEFINITION *retm* gene region (7134bp)
FLYBASE ID FBgn0031814
ORGANISM *Drosophila melanogaster*

```
   1 ACACCATTGG CGTCTGTTAA TCAAACTCGG AGAAAGGCTT TCTCTGTTTT CCACGGCCAG
  61 GCGGTTCTCT AGAACTCCTA CGCTAGGAGC CAACGGATAC CAGACACAAC TCAGTGAAAT
 121 AAATAGCAGT GGCGTGAACT GTCACGAGCA AAAAAACAAA CCCAGCTCGG TATTCCAGTG
 181 CCCAAGTTTG GTCACCTTCC CCGCCCCCGC GCGCTCATTA TACTGCAACG GAAAAGCGCA
 241 GAGTGAAAAT TGATTTTTTA AACAGTGCGA ACTCCACAAT CAGTGCAACG AAAATTTGAA
 301 ATAAAAAACA CGGAACGGAC GATTGAAAGG CAGCAAGATG GTGCAAAAAT TCCAGTCACC
 361 CGTTCGCGTC TACAAATATC CTTTTGAGCT GGTGATGAAG GTGTGTACAT TGCATTTTCC
 421 CCTCTCCATT TCGTCGTGTG GAGTGGGCGG CCTTATAGGC GTGTCCCACT GCCTGTTTAC
 481 ACATACCCAT CAATTTAAGT TTAAATTGTG CATTATGCGA TGGCATAGCC ACTAATTCAC
 541 TGCGACTCGC AAAGGGTCGT TGTGTATCGT TGTTTCCGCT CGGACACAGT TGCTTCATTT
 601 ATTCAAATCC AAGCTGCAGA TAAATAAAAT TAACCTTGGA CAGCAATTCC TCCCCCCGCC
 661 CAACATGCTC TCTCTTTCGC AGTTGCGCTC GAAGAGCGAG ATAGAGATAG ATTCTTTATC
 721 CAATTGAATA ATACACACTG CATGTACGTG TGCGTGTGTG TGCGTGTGTA AATAGTTTGT
 781 CGGTGGGAGG GGGCGGTTTG GTTGAAACGA CTTGTTTGTG AATTTTTCAT TTATTTTTAT
 841 TAATCTACTG ATTGTTTGCG TCCAGCTTTG CGTTTTAGTC TCATTTATGC CATAGAAATT
 901 TGTTCTTTTT TTAATGCTGT TAATTTACAT AATGCGCCTT TTTGAATGTA TTAACATTTA
 961 TCGTAAATTA TCGAAATTGG GGTTATAAAT TTATAAGATA AAAAATGGTT GTCTTTTTTA
1021 TTGAAGCTTT TCCTCCCTCC TACCTCTGTT TCCTTTGCAA ATCATTTTAT TTTTTGCCTT
1081 TTCACTGACT GATAAGAAAA CGAACTGTTT TTTACTTCAA AAGTTTATTT AAGCAACAGA
1141 AATCACTTCA AGTGCCGTTT TTTGTACTTA AAATAACTAT ACAGTCATAT ATTAGACTAA
1201 AAGTGTAGTC CAAACACAAA TCACCTAAAA TGAAAAGCCT ATTATGGTTA TTCGATACTT
1261 TGAGTGTTTA TGTATGTTTT CCCTTGTAGA CAGAAGTAAG AAAGCGCTTT TCGTGACAGA
1321 AACTCAAAAG TCTCAAATGT CAAAGGTCAA ACATTTATTC TACGTACGTC ACCGACAGGG
1381 GTAATTTAAA TTTAAGCAAC ATTTATGAGC CGCACGATAG GACACCTGTC CACTCGCATG
1441 TCCCCAAAGA TGCAGGCAAA GTTCCAGCAT TTTAGACCAT TACCGCCCGC CAGGACAATT
1501 AAGAAAGCCT ATCTGAATAA TTAAGGACAC CAGTGCCAAC TACACAGTTT GCTAGCTGGG
1561 TCTGGGAAAA ATATGGATAT TGTAAGACAT GTCCATGTGA GGTGAGGAGT CGCCACGTGT
1621 TGGATGGCTC CCACCGCACA CTGAATTATT CAACGTGTGA ATATCCACAT GAATATCATC
1681 ATGAATATCC ACACAGCGAT TTATGACCTT TGCACTAGTA ACAAGGACTC GCAACCACTA
1741 GATAGAATTA TGATTTGGAT TGCACAATTA TCGATATGGA AAATCTATTT TTGATTTTCA
1801 TTGTTTATTG CTTGCAAACA AGGTATGGAG TTCAAACGCA ACAAGAATAA GAATAGATCT
1861 TGAGCAATAA GAACGTATTT TTCCACTGTC TTAAGAATTT TACATTTTTA AACCCAACAA
1921 TTTCATGAAA ATAGCGAAAT ATTATATGCA AAAGAACACG TTTGAGTATA TTATTTCAAT
1981 TGTATTGATT TGTTCAATCA GCAACCAATT TTTGCAGCGA TTAGGTTGTA TTCTAAATGA
2041 AAAGCTTTTC AACTGCGCAA ATTTTTGATA CAGCGATAAT TTTCTACTTT TGACCGATCA
2101 TTGATTTGGG TCCTATATAA TATAGTCGTG TTGATCAGCG GTATTTAGCA TGCATGAATG
2161 GAGCTATTCT AAGTTTTGTT TATAAGTAAA TGTAATTTTT TTTGCTAAAT ATATTTAGCA
2221 ATTTCGGAAT TTGCATTTAT TATATATGGA AGAAAACACA AGTAAACAAC ATTTTCAGGC
2281 ACTGCTCTCA CCTGCTTAAA TATTGCGCTT TTGCGTATTT ACGCCATACC CAAATAAATA
2341 ATCAAAATAA ATGTGCGTCT GCGAATTGAG TTAAAAATCC ATCACTCATA CGCGCTGTTG
2401 ACTTTAATCA AAAAGTTTTG TACTTGCAAA TATGCGATGC TTCTTTTTTG CTATTAAATA
2461 ATTATCTATT TAGTATAGAA ATATGTATAT TATAGATTTA TATAAAATAT AGGAAACAAC
2521 ATTTCGCAAC CGCCCTAATA ATTTGACCTT TTTCATCGAA AACCCAGGCA CGATTGGCCA
2581 AAAATAGACC TTCGAAATTG TTAAAGGCTT CTGAGACGTT GGCTGCAATC TGTGACGGAG
2641 GATGAGCAAA ACGGGGCGGG ATGGGGTGCA TGTACATAGA TGGCTGCAGG GCCAAGTTAT
2701 AAATTGGCCA GCGTTCAAGG TCGCTTGGCA GGGAACCGGA ATATAGAAAA TAAATACATA
2761 CACTTACCAA AACTGTTGCT AGCAATTTGA ATTCACTAAA TACTTAGCTA TCAAAGAGAG
2821 TCCTTAAATA TACATATTCA ACATACATCT AACTTATTAT CCTCGTTGCT ACTGCAGGCC
2881 TATGAGCGGC GGTTTCCCAC ATGCCCGCAG ATGCCCATCG TCCTGGACTG TGAGGTCATT
2941 AAGGACGAGT CCCTGGAGGA CGGGGCCAAG CGAAACACGA GCCGACGCTG CAAGCTGGCG
3001 GTCGACGCGC CGTACATCTT CAAAAAGCTT ATTGGCGTGG ACCATGTGTA CTTCCTGCAG
3061 CACAACTTCC TGGACCTGGC CAACCGCACG CTGAGCATCG AGGCGGTCAA CGAGAGCTTC
3121 TCGTCGCGCA TCGAAATCTT CGAGCGGTGT CGCTACTACG CCCATCCGGA CAACTCCGAA
3181 TGGACGTGCT TTGACCAGTC GGCCACGCTG GACATTAAGA ACTTCTTCGG CTTCGAGCAC
3241 TCGATGGAAA AGATGGGCAT GAAGCAGTAC ACCCAGACGA CGCTGAAGGG CAAAGAAATC
3301 ATTGAGTTCT TTATCGGGCA GCTGCGCGAG GAAGGCATCA CGCATGTGGA GCGTTGGACA
3361 TCGCCCAGTG ATGCAACCAA ATCGCCAACC CTCGATCAGG CTTCGGATCA GCAGCATAGC
3421 ATCTTGCTGG ACGGCGACTT TATCGCACGC AGCCTGGGCC AGCTGTCGCC CATGCAGGAG
```

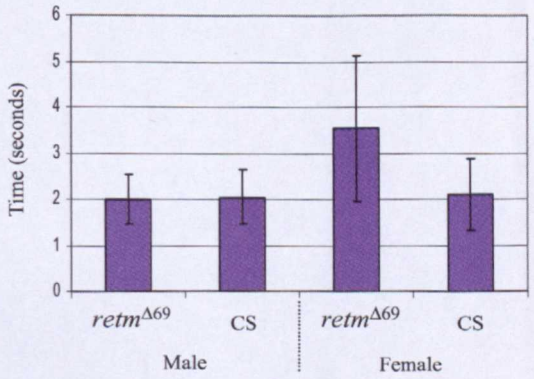
3481 TCCAAGCTGC TGGAGCTGCG AAAAATGCTG GATGGTGTCG ATGACTTGGA GCGTGTGCCC
3541 AGTTATCAGA CCATCCTGCG TTTCTTAGCA GCCAGGGACT GGCACGTTAG CCAGGCCTAC
3601 GCCATGCTCT GCGACTCGCT TCGCTGGCGT CGGGAACATC GCATCGACGC ACTGTTGGCG
3661 GAGTACTCCA AGCCAGCCGT GGTAGTGGAA CACTTTCCGG GCGGCTGGCA TCATTTGGAC
3721 AAGGACGGTC GACCCGTTTA CATTCTGCGC CTGGGACACA TGGACGTCAA GGGCCTGCTA
3781 AAGTCACTGG GAATGGATGG ACTACTGCGA TTGGTAAGCA AGCCGGCGGA GAACTTATCG
3841 CTGAGATTAC ATTTTAAACA ATATGCTGTC TGTCTATAGG CTTTGCACAT CTGCGAGGAA
3901 GGCATTCAGA AGATCAATGA ATCCGCCGAA CGCCTGGAGA AGCCAGTTCT CAACTGGTCA
3961 CTGCTTGTCG ATCTGGAGGG TCTTTCTATG CGCCACCTCT GGCGTCCCGG CATCAAAGCC
4021 CTGCTCAACA TCATCGAGAC TGTGGAGCGC AATTACCCAG AGACCATGGG CCGTGTCCTA
4081 GTGGTGCGAG CGCCCAGAGT GTTTCCCATT GCATGGACAA TCGTTAGTGC TTTTATTGGT
4141 GAGTCTGAAC AAGCCGTCAT ATGGAATTGC CTGTTAACAA CGCTTTGAAA TCCTGCAGAT
4201 GAGCACACCC GATCCAAGTT CCTGTTCTAC GGCCCCGACT GTGCGCACAT GAAGGATGGC
4261 CTGGCCCAGT ATCTCGACGA GGAGATTGTG CCTGACTTCC TGGGCGGACC GTGCAAGGTA
4321 AGCAAGCAAG GCGGCAGCCA CGACTTCCGG TAGCGCGTCT TCGGTTGGGA GCCCATTACT
4381 GCTGCATCTG TAGCAACGTA CTTACTATAT TATGTTATTA TTGTATATGT GCGTTAACTC
4441 TCTATCCCTC TGGTTTCGGC CTTGATTCCG GCTCGCTGGG GCTGCTCAGA CAGAGAACCA
4501 GCTAGTCTCG CATGTGGAGC AAGTCAGTCA GAAATACTTT AGCCATTGTT ATTTAACCAA
4561 ATGCGATTGT GTAAGCATCC TACTCTCTTG TTCTTTTGCT CCTGCTCTCG ATTTCGTTTG
4621 TCACCATCCA CAACGACAAA TTGTACTGTA CTCGTCCGCC TTTGTCTTAC TTCCTCGTTT
4681 TTGTTTCTCA CACTTTGTTC GCTGTGCGGT TCTGTTGCAA ACTGTGTCAA GCCCATGTGC
4741 GCAGCTTAGC CATGACATTC CTGCTGTAAT ATTTCCTCCA TTGTGCCCCA GCGTCCATGT
4801 CCACGTCACT TAATCTCTCT GTTCACCCGG CGGCTGTCAT ACATGTGCTC AGTCCAAGCC
4861 AACAGCCACG GAGCCTGCCC ACCAGAAACT GCTCTAACGT ATTGTCCTCT TTCCACAAAG
4921 ACCATGATTC ACGAGGGCGG ACTGGTGCCT AAGACGTTGT ACAAGATGAA TTCGCTGGAG
4981 GATCACGACG ACGAAGTGAC CGCCGAGCTG CCCACCACTG CCGCGGCACA GGCTTTGGTT
5041 CCAGGCAAGC GCCTGTCTGC CAACCAGCAA CACGATCATC GCAATCTGTA CAAGAGTGTG
5101 GATCTCAAGG CCGGATTCGC GCACGAGCTG CTCATCCGGA ACGAGGATCC CAAGAGTGTG
5161 CTCACCTGGG ACTTCGATGT GATGCGCAAC GATCTGCACT TTACGCTCTA CCGGGTCACC
5221 CAGGAGCTTC CGGAAAAGAA TGGTGAGTTA ATAATGCACT TACAAAAGGT TTTCATGTTG
5281 CTTTTTGTTT TTCTTTTGTT CTGAGAACTG GTTTTGGTTT CAAAGGTGCG AATATCACGA
5341 TGTTGTTGCC TCATTAGAAG TTGTTTTTGT TTACTTAGCA CTTTTATAAA AACACAAAGC
5401 CACGATAAAA AGCGAAAAAT GCGGAAAGTG GAAACTCTTA CACCCTTTCG AAAAGAAAAG
5461 AAAATTTGTT GTCAGGCAAC TTAAAAACAA ATTAATAACG TTTGGCTTAA ATTGAATAGT
5521 TTACTATTAC AATTAAGTAT TAAAACGTAT GTTAATTTGT GTATATATTT ATTTATATAA
5581 ATTTCCAAGG TATCTGTTGC TAGGGCTCTA TTAAATTGAT AACGCGCGCA TTTTGGTTCC
5641 GGTTTATTGA TAGGGATTGC TTTTCATATG TGTTTTGATA ATGTAGTAAA ATACGAAAAA
5701 GTTGAACCTT ATTTTATAAA GCTTTTTGTA CTTTGAACCT TTTAATTAAT TAACTTGATT
5761 AGCTTGACAG CTTACATATA ATTAAATGCC CTTCGCCCAC AGATTCCGCC GTCTCGTACT
5821 TTGATCTGCA GGACTTCGTC GAGGGCGTCA ATTATTTCCG TGAGGAGCCA ACGCTGATTT
5881 GCCGGCACAA GGAGAGCGTC CAGGGGTCGC ACGTGATGCA TCACAACGAC AGCTATCTTA
5941 TGCACTGGTT TAGTCCATCG GGTGCCCAGT TAAACGTCTT CTACGAGGTG CTCAGCTCGG
6001 CCAACTACAA GGGTTCGATG ACCAGCCTGC AGTCGGCATT CTCATCGAAC TCGTCGGCGG
6061 CCAGCTCGGT GCAGAGTCGA TGATATGAGA AAGCCGGGGA TGGCGGAACG AACGGCGCCG
6121 AATGGAGATC CCCGGCATTG GAGGAGCCCA TAGCGGCGGA CATTCTCACG GGCACCCTGA
6181 TGGAGAAGAG CCTGAATCTA TTTGAAACGC CGGACGTATT CTTTTAAGAC CGAAAGTGGT
6241 AATATCTTGT GTTTTAGTTG TCTTTAGTTG TTTAGCAGAA TGTTAGTTTG ATTTTAAATT
6301 GTATTTAGCT CAAATAGGCC ATTTATTTCC TTTTAAATTT ATTATATTCA TTCAATATAT
6361 ACAAGCATTC ATATTATTGA TAATCTTAAG AAAATTAGAA ATATCAGCGG TTTAAATGAG
6421 CAACACCAAA ATATCACGCA ACGTTTTATT TGAATCGTCA ATTTGAATTT TTCGTAAGCA
6481 GCTCTATAAA TGAAGTTTTA ATGCGTACAT ATATATGTAA AAGCGACGTA ATTTTTATAG
6541 CTTTTTCTAA TGGAACATTT TTTAGCAACA ATTAAGGAAC TTATTGCGAT TGCGCAGCCA
6601 CTGAAACGAA AATAGAAAAA ATTGAAATTG TAATCAATAA CTACGTGTTA CCTAATTAGT
6661 TTAATAAACA AATTGCCGAA ATGGCTTATT GCAAGGATTT TAGCCAGAAC TTATACAAAA
6721 GAGTGAAATG TAAGATTTGA ACTCGAAATA CACACGTAGG TTATAGCTAG ACGTCAGATT
6781 ATAATTATAA ATGTAAAATT GTATTTTAAA CATTCGAATG CAGAACGTTT TTGATTGCAC
6841 ACTAATTACC CCGTCATTTG TACGCCCTTA ATTGCTAATA TCAATCAATC TCCAAAGTGT
6901 TTTTAAAATC AAAAAAGGCA ACTTCAGTGT GTTTTTTTGG TTATTATTTA TTGTATTAAA
6961 AACGAATTAC TATGCGGATG TTTACAATTT AGTTTTATAA ATAATCTGAC GTGTATTGTA
7021 AGAATAAACT GACAAGCCAA TAACGAAGGA AAAATGAGTG CATGTTAACT AAGTGTGTAA
7081 GTTGTATTAA GAACGTAATC AAATAAATGC TATTAGATCA ATTACATGTA CATA

Appendix 6

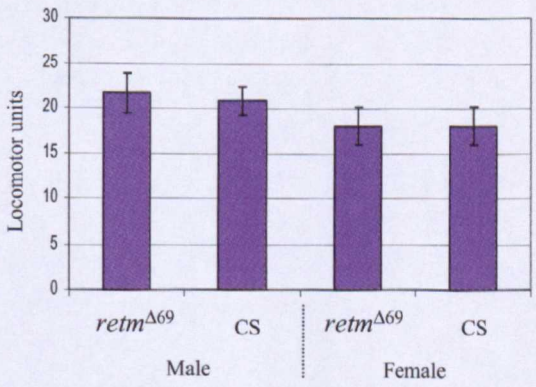
A.



B.



C.



Preliminary behavioural analysis of *retm*$^{\Delta 69}$ deletion mutants. For all behavioural assays, male and female *retm*$^{\Delta 69}$ homozygotes were tested separately and compared to CS wild-type controls. All assays were performed during the first four hours after subjective dawn. n=20. **A. Spontaneous locomotor activity**. Flies were mouth-pipetted and placed individually in a clear perspex 10cm x 10cm maze marked with a 1cm^2 grid. After 5secs for acclimatisation gridlines crossed in the following 30secs were counted. **B. Inversion assay** was performed immediately following the locomotor activity test. Singly, flies were knocked supine and the time taken to return to an upright position measured. **C. Post-shock spontaneous locomotor activity**. Immediately following the inversion assay, locomotor activity was tested again using the same method as in **A**. No significant differences were observed between mutant and wild type in any of these behaviours.