The Molecular, Cellular and Functional Organization of Metabotropic Glutamate Receptors in *Caenorhabditis elegans*

by

James Charles Dillon MBioSci

A thesis presented for the degree of Doctor of Philosophy

of the

University of Southampton

in the

Faculty of Medicine, Health and Life Sciences School of Biological Sciences

February 2007

UNIVERSITY OF SOUTHAMPTON
ABSTRACT
FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
SCHOOL OF BIOLOGICAL SCIENCES
Doctor of Philosophy
The Molecular, Cellular and Functional Organization of Metabotropic Glutamate Receptors in *Caenorhabditis elegans*
by James Charles Dillon

The role of glutamate as a neurotransmitter is evolutionarily conserved across phyla [Walker, R.J. et al 1996]. In the mammalian central nervous system it is involved in signalling pathways that subserve complex behaviours [Nakanishi, S. et al 1994]. Two classes of receptor conduct glutamatergic signals, these are the ionotropic receptors [iGluRs] and the metabotropic receptors [mGluRs] [Nakanishi, S. et al 1998]. The experimental tractability of the model organism *Caenorhabditis elegans [C. elegans]* has provided insight into the function of iGluRs [Brockie, P.J. and Maricq, A.V. 2006], however, there is very little known about the *C. elegans* mGluRs (*mgls*). We have utilised this subclass of receptor and *C. elegans* as models by which to further understand the cellular and molecular mechanisms of receptor organisation underlying behaviour.

The *C. elegans* genome predicts three metabotropic glutamate receptors and these are denoted *mgl-1, mgl-2* and Y4C6A.2a (referred to as *mgl-3*). A combination of *in silico* and cDNA analysis has identified the *mgl* proteins harbour conserved structural domains and protein interaction motifs that are important to mGluR function. The cellular expression of *mgl-1* and *mgl-2* suggests that each subtype is differentially expressed in the *C. elegans* nervous system. *mgl-1* has the more widespread expression of the two subtypes and is selectively associated with the pharyngeal nervous system, which directs and modulates *C. elegans* feeding behaviour. Pharmacological assays performed on the pharynx of *mgl* deletion mutants using mammalian mGluR agonists identified *mgl-1* and *mgl-3* regulate the activity of the pharynx and serve as neuromodulators of the pharyngeal neural network.

It is now apparent that the C-terminal domain of membrane bound receptors serves as the foundation upon which large multi-protein intracellular signalling complexes are constructed. Such complexes enlist scaffolding proteins that orchestrate the receptors compartmental targeting, subcellular anchoring and signalling. To identify *mgl-1* scaffolding proteins the intracellular C-terminal was utilised to perform a LexA yeast-2-hybrid screen. Three proteins were prioritised from the screen that harboured protein interaction domains with known functions in scaffolding receptor signalling complexes. These were encoded by the genes *ptp-1*, *mpz-1* and *tag-60*. The cellular expression of each gene has been defined and discussed. The co-expression of *mpz-1* and *mgl-1* in the pharyngeal nervous system suggests that *mpz-1* may organise *mgl-1* signalling in this neural circuit. To examine this, the function of *mgl-1* was tested in the two available *mpz-1* mutant strains; *mpz-1(tm1136)* and *mpz-1(gk273/+)* and the trans-heterozygote *mpz-1(tm1136/gk273)*, by using the pharyngeal preparation as an assay of *mgl-1* function.

cDNA analysis predicts both of the mutant transcripts do not encode MPZ-1 protein, however, the function of *mgl-1* was intact in the mutant strains. In the absence of any disruption to *mgl-1* signalling in the *mpz-1* mutants further work is required to determine the relevance of the *mpz-1* interaction to *mgl-1* function. This study has provided the first functional description of an *mgl* subtype in *C. elegans*. In doing so this has identified a bio-assay for investigating the organisation of *mgl-1* signalling by interacting proteins *in vivo*.

LIST OF CONTENTS

List of Figures

Acknowledgements

Publications and Presentations

Dillon, J., Zeng, T., Hopper, N.A., Holden-Dye, L., and O'Connor. V. (2005) Metabotropic glutamate receptors in *C. elegans*: Models for studying the organization of receptor function. Biochemical Society. Molecular Determinants of Synaptic Function: Molecules and Models. Southampton

Lidwell, K., Dillon, J., Sihota, A., O'Connor, V. and Pilkington, B. Determining calmodulin binding to metabotropic glutamate receptors with distinct protein-protein interaction methods. (2004) Biochemical Society Transactions. **32**. 868-870.

Dillon J Holden-Dye L and O'Connor. V. (2004) Metabotropic glutamate receptors and *C. elegans*: Models for synaptic function Synapses: Genesis and Function, Ringberg, Germany. 2004.

Dillon J, Pilkington B, Cook A, Willson J, Holden-Dye L and O'Connor. V. Metabotropic glutamate receptors in *C.elegans*. (2003) Society of Experimental Biology **134**. S70.

Abbreviations

Amp	Ampicilin
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
CIAP	Calf Intestinal Alkaline Phosphatase
CHO	Chinese Hamster Ovary
CNS	Central Nervous System
$CuSO_4$	Copper Sulphate
cfu	Colony forming units
DAG	Diacylglycerol
DEPC	Diethylpyrocarbonate
DMF	N,N-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribose nucleic acid
dNTP	deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
EtOH	Ethanol
GABA	γ-amino-butyric acid
GDP	Guanine nucleotide diphosphate
GIRK	G-protein activated inward rectifying K^+ Channel
GluCl	Glutamate gated chloride channel
GPCR	G-protein Coupled Receptor
GRKs	G-protein Receptor Kinases
GST	Glutathione-S-transferase
GTP	Guanine nucleotide triphosphate
GTPase	Guanine nucleotide triphosphatase
5-HT	5-hydroxytryptophan
iGluR	Ionotropic glutamate receptor
IP3	Inositol tris-phosphate
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kainate	2-Carboxy-3-carboxymethyl-4-isopropenylpyrrolidine
LB	Luria Broth
LTD	Long Term Depression
LTP	Long Term Potentiation
mGluR	Metabotropic Glutamate Receptor
mRNA	messenger Ribose Nucleic Acid
NCBI	National Centre for Biotechnology Information
NGM	Nematode Growth Media
NMDA	N-methyl-D-Aspartate
NSM	Neurosecretory motorneuron
PCR	Polymerase Chain Reaction
PKC	Protein Kinase C
PLC	Phospholipase C
RACE	Rapid Amplification of cDNA Ends.
RGS	Regulator of G-protein Signalling
RNA	Ribose nucleic Acid
rpm	revolutions per minute
TAE	Tris-Acetate EDTA
TE	Tris-EDTA

TM	Transmembrane
SDS	Sodium dodecyl sulphate
X-Gal	5-bromo-4-chloro-3-inolyl-β-D-galactopyranoside
YTH/Y2H	Yeast Two Hybrid

Amino Acids

A	Alanine	Ala
C	Cystine	Cys
D	Aspartate	Asp
E	Glutamate	Glu
F	Phenylalanine	Phe
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
K	Lysine	Lys
L	Leucine	Leu
M	Methionine	Met
N	Asparagine	Asn
P	Proline	Pro
Q	Glutamine	Gln
R	Arginine	Arg
S	Serine	Ser
T	Threonine	Thr
V	Valine	Val
W	Tryptophan	Trp
Y	Tyrosine	Tyr

Nucleotides

a	Adenosine
c	Cytosine
g	Guanine
t	Thymine
u	Uracil

CHAPTER 1

Introduction

1.1   The nervous system and behaviour

The activity of the central nervous system [CNS] underlies motor behaviours that include movement and breathing, together with more complex cognitive and affective behaviours, such as emotions, learning and thinking [Purves, D. et al 2001]. The CNS initiates commands on the basis of information it receives from the periphery. The processing of this information is achieved with nerve cells or neurons and the connections between them. Co-ordinated signalling between neurons gives rise to the complexity of behaviours mentioned above [Purves, D et al 2001].

1.2   The neuron

Studies performed by Santiago Ramon y Cajal [Histologie du Système Nerveux, Paris, 1909] revealed the neuronal cell is composed of 3 morphologically distinct regions, the cell body/soma, the axon and the dendrites *(see figure 1.1A)*. The cell body contains the nucleus and is the metabolic centre of the neuron. It gives rise to two types of process, the axon and dendrite [Bradke, F. and Dotti, C.G. 2000]. A single neuron can possess several branching dendrites, the function of the dendritic compartment is to receive signals from other neurons. Certain neurons, for example pyramidal cells, possess two different types of dendrites. Apical dendrites that extend from the apex of the cell body and shorter, basal dendrites that emerge from the base of the cell body [Larkman, A.U. 1991]. The axon arises from the axon hillock of the cell body and can extend for up to 1 metre. It is the main conducting unit of the neuron responsible for transmitting electrical signals. Towards the end of the axon it branches out into fine processes which give rise to specialised transmitting compartments required for conveying information from one neuron to another [Peters, A. and Palay, S.L. 1996].

1.3   The chemical synapse

Signal transmission between nerve cells occurs at specialised points of communication, one type found in the nervous system is the chemical synapse *(see figure 1.1B)*. The chemical synapse is characterised by a presynaptic neurotransmitter releasing nerve terminal [Katz, B. and Miledi, R. 1965] opposed to a receptor expressing postsynaptic membrane specialization [Palay, S.L. 1958 and Shepherd, G.M. 1972]. Neurotransmitters recognise and bind to specific

receptors leading to the activation of intracellular signalling events. Mechanisms contributing to synaptic signalling include the reuptake of released neurotransmitter, the organization of neurotransmitter releasing vesicles, the location of neurotransmitter release sites relative to synaptic receptors and the organisation of receptors and the signalling cascades to which they couple. [Garner, C. and Nash, J. 2001].

1.4   The neurotransmitter glutamate

Glutamate is the main excitatory neurotransmitter of the mammalian central nervous system [Greenamyre, J.T., Porter, R.H. 1994]. The significance of glutamate as a signalling molecule is highlighted by the large number of different physiological processes it is involved in [Meldrum, B.S. 2000]. These include sensory pathways [Nakanishi, S. et al 1998], motor co-ordination and cognition [Homayoun, H. et al 2004]. Furthermore glutamatergic signalling within neural networks is conserved across phyla [Walker, R.J. et al 1996]. It is involved in behaviours such as mechanosensory responses [Maricq, A.V. et al 1995], learning and memory [Morrison, G.E. and Van der Kooy, D. 2001] in the nematode *C. elegans* and signalling at the neuromuscular junction of *Drosophila Melanogaster* [Bogdanik, L. et al 2004]. The involvement of glutamate in mammalian conditions of drug addiction and neuropathalogies [Meldrum, B. et al 1990] such as ischemia, epilepsy and schizophrenia make glutamatergic pathways an attractive target for drug therapeutics and further highlights the importance of understanding how glutamate signalling is carried out.

There are two classes of receptor capable of binding glutamate and mediating downstream effects. These are ionotropic [iGluR], ligand gated ion channels responsible for fast transmission [msecs] [Popescu, G. and Auerbach, A. 2004] and metabotropic [mGluR], G-protein coupled receptors responsible for slower [100's msecs] modulatory signalling [Nakanishi, S. et al 1992].

A

Dendrites

Pre-synaptic
Bouton

Cell soma

Axons

Synapse

B

Mito

AZ

pre

SVs

post

PSD

glia
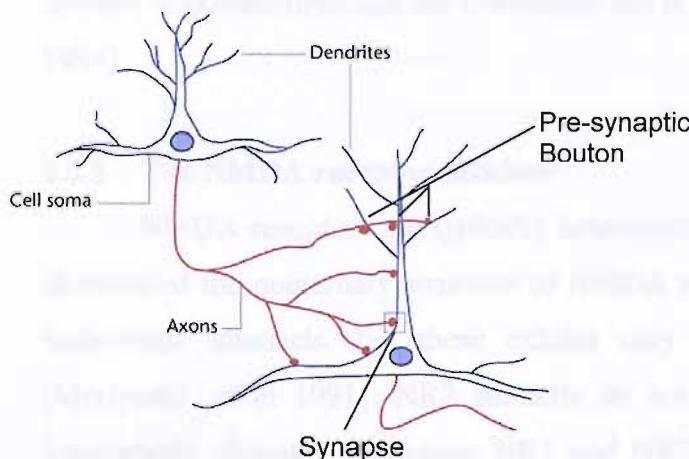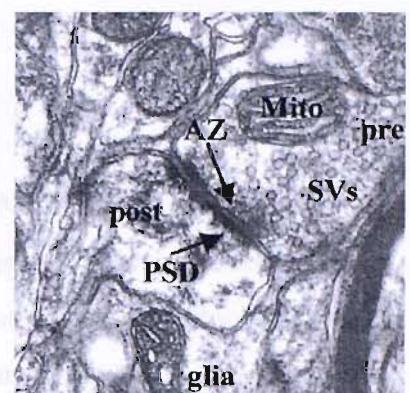
Figure 1.1 General description of a neuron and synapse. **A. A diagram of neurons making synaptic connections.** Neurons consist of a cell body, dendrites and axons. The axon extends from the cell body and branches out to make synaptic connections with the dendrites of a neighbouring cell. Two neurons are shown making synaptic connections. **B. An electron micrograph of a chemical synapse.** The pre-synaptic domain [pre] exhibits synaptic vesicles [SVs]. When the neuron is electrically stimulated the SVs fuse with the plasma membrane at the active zone [AZ] to release their neurotransmitter contents. The post-synaptic compartment [post] is characterized by an electron dense region [Postsynaptic Density: PSD], which is rich in neurotransmitter receptors and intracellular signalling molecules. [A and B taken from Garner, C.C. and Nash, J. 2001]

1.5   The ionotropic glutamate receptors

The mammalian ionotropic glutamate receptors are divided into two classes based on their pharmacological properties and named according to the agonist they are responsive to. The two classes are N-methyl-D-aspartate [NMDA] receptors and non-NMDA receptors, which are 2-amino-3[-3-hydroxy-5-methylisoxazol-4-yl] proprionate receptor [AMPAR] and Kainate [2-Carboxy-3-carboxymethyl-4-isopropenylpyrrolidine] receptors.

The iGluR channel is an oligomeric protein. Seven NMDA [NR1, NR2A-D, NR3A-B], 4 AMPA [GluR1-4] and 5 Kainate [GluR5-7, KA1-2] ion channel subunits have been cloned from rats [Riedel, G. et al 2003]. The iGluR subunits exhibit a characteristic membrane topology. It consists of three transmembrane helical domains [TMI, TMIII, TMIV], and a hydrophobic loop region, TMII that

makes a hairpin turn within the membrane *(see figure 1.2]*. The N-terminus domain is extracellular and the C-terminal tail is intracellular [Hollmann, M. et al 1994].

1.5.1 The NMDA receptor subclass

NMDA receptors are typically heteromers [McIlhinney, R.A. et al 1996]. In terms of the quaternary structure of NMDA receptors NR1 subunits can form homomeric channels, but these exhibit very small agonist-induced currents [Moriyoshi, et al 1991]. NR2 subunits do not form homomeric channels but heteromeric channels containing NR1 and NR2 subunits are highly active and have much larger agonist-induced currents [Ikeda, K. et al 1992], which suggests the majority of NMDA receptors in the brain are likely to be heteromeric. NMDA receptor specific, radio-ligand binding studies have identified NMDA subunits in different anatomical regions of the brain, but the NMDA receptors are predominantly in the forebrain with highest levels in the CA1 region of the hippocampus [Maragos, W.F. et al 1988 and Subramaniam, S. et al 1991]. In-situ hybridisation histochemistry has shown the NR1 subunit mRNA to be widespread in the brain [Kutsuwada, T. et al 1992] and compared to NR1, the distribution of NR2 transcripts in the rodent brain is developmentally regulated. [Watanabe, M. et al 1992]. Subsequently, during brain development the temporal and spatial changes in NMDA receptor subunit expression gives rise to functionally diverse, tetrameric channels.

Targeted gene disruption of NR1 and NR2B subunits has identified them as being essential for survival, mutating either NR1 or NR2B causes death from respiratory failure [Forrest, D. et al 1994 and Kutsuwada, D. et al 1996]. Disruption of either NR2 subunit A [Sakimura, K. et al 1995], D [Ikeda, K. et al 1995] or C [Ebralidze, A. et al 1996] is not lethal. However, in mice lacking both NR2A and NR2C subunits, for example, the NMDA receptor channel currents are disrupted, leading to impairment of fine co-ordination [Kadotani, H. et al 1996]. This illustrates the combined disruption of more than one NNR2 subunit can have a more detrimental effect upon brain function than disruption of either on its own.

NMDA receptors display characteristic ion channel properties. At resting membrane potentials the ion channel is blocked by Mg^{2+} [Mayer, M.L. et al 1984] depolarisation of the neuronal membrane and activation by agonist leads to

an influx of Ca^{2+} [Mayer, M.L. and Westbrook, G.L. 1987]. Efficient activation requires the co-agonist glycine. NMDA receptor activation and deactivation is slower (50-500ms) than AMPA and Kainate receptors (0.2-2ms) and NMDA receptors remain active after glutamate has been removed from the cleft [Popescu, G. and Auerbach, A. 2004]. These properties have earned NMDA receptors the reputation of being 'coincidence detectors' because physiologically, activation is dependent upon both presynaptic release of agonist and postsynaptic depolarisation.

1.5.2   The AMPA subclass of receptor

The GluR subunits are approximately 900 amino acids in length, with 68-73% identity at the level of primary structure [Boulter, J. et al 1990]. Their association into either homomeric or heteromeric channels generates receptors with different pharmacology and conductance properties [Jonas, P. and Burnashev, N. 1995 and Jonas, P. et al 1994]. Homomeric and heteromeric channels containing the GluR2 subunit conduct the divalent cation Ca^{2+} poorly, whereas homomers of GluR1, 3 and 4 are highly permeable to Ca^{2+}. This variation is caused by RNA editing of the GluR2 subunit [Sommer, B. et al 1991]. The GluR2 precursor mRNA is chemically modified to convert a Glutamine codon, encoded by the gene, into an Arginine codon. Both the location of this residue within TMII, the pore forming region, and the ionic charge it possesses results in it having a significant influence over the conductance properties of both hetero- and homomeric AMPA receptor channels [Schoepfer, R. et al 1994].

The expression of the AMPA subclass of receptor throughout the mammalian central nervous system implicates AMPA receptors are involved in a number of different processes. However, the expression of different GluR subunits in different anatomical regions of the brain [Keinanan, K. et al 1990] suggests AMPA receptors with different subunit compositions play specific roles within specific regions of the brain. Consistent with the abundant expression of AMPA receptors in the hippocampus and cerebral cortex they have been identified as having a role in the molecular mechanisms underlying memory and learning [Malenka, R.C. et al 2003]. In contrast to NMDA receptors AMPA receptors have rapid gating kinetics and are responsible for fast excitatory transmission at the postsynaptic membrane [Ozawa, S. et al 1998].

13

1.5.3 The Kainate subclass of receptor

The GluR5-7 subunits are ~900 amino acids in length and exhibit 75-80% amino acid sequence identity [Ozawa, S. et al 1998]. The KA1-2 subunits are slightly longer, ~970 amino acids and display a similar amino acid identity, 70%. Amino acid identity between GluR5-7 and KA1-2 is however lower, at ~40% [Seeburg, P.H. 1993]. Radioligand binding studies have identified Kainate binding sites through out the CNS, with intense labelling in CA3 of the hippocampus and the granule cell layer of the cerebellum [Foster, A.C. et al 1981 and Monaghan, D.T. 1984] . In situ hybridisation has confirmed this pattern is representative of Kainate receptor binding and has further highlighted the distinct, but overlapping distribution of the various subunits [Bahn, S. et al 1994]. The use of recombinant Kainate receptors to study channel properties has suggested they are likely to function as homomers (in the case of GluR5 and 6) (Egebjerg, J. and Heinemann, S.F. 1991) and heteromers composed of different subunit combinations and ratios, with distinct pharmacological properties [Herb, A. et al 1992]. The use of selective pharmacological tools has allowed further insight into the physiological function of Kainate receptors [Christensen, J et al 2004].

A presynaptic role in the regulation of neurotransmitter release has been demonstrated at the glutamatergic Schaffer collateral-CA1 synapses of hippocampal slices and at GABAergic interneurons of the hippocampus [Rodriguez-Moreno, A et al 1997]. In addition to this, postsynaptic Kainate receptors responsible for mediating excitatory neurotransmission have been described at synapses formed by mossy fibres onto CA3 neurons in the hippocampus [Castillo, PE 1997]. Transgenic mice with disrupted genes encoding certain Kainate subunits [either KA2 or GluR5 or GluR6] [Contractor, A et al 2001 and 2003] have been generated and are viable, with normal behaviour and comparable spatial learning abilities to wild type. However, they do exhibit deficits in synaptic plasticity. The use of these transgenic lines has enabled the identification of the subunit composition of Kainate receptors in specific cell types and signalling pathways [Mulle, C et al 2000]. In addition to ion channel activity, Kainate receptors can couple to second messenger signalling cascades. Such metabotropic activity of Kainate receptors has been shown to regulate the

release of GABA in the hippocampus [Rodriguez-Moreno, A. and Lerma, J. 1998].

1.6   The metabotropic glutamate receptors

The G-protein coupled receptor [GPCR] family is a large class of membrane receptors. In model organisms genome sequencing identifies the abundance of this receptor family. The *Drosophila* genome contains ~200 genes encoding GPCRs [Brody, T et al 2000] [~1.5% of the genome]. The mouse genome is estimated to contain 1000-2000 GPCRs, comprising >1% of the genome [Offermans, S 2003]. In *C. elegans* GPCRs are predicted to account for ~5% of the genome [Bargmann, C et al 1998]. The GPCR family members are involved in a multitude of different physiological processes including vision, olfaction, neurotransmission and hormonal signalling [Bockaert, J et al 1998]. Such diversity is reflected by the variety of ligand binding domains amongst GPCRs. This characteristic and sequence homology has been used to subdivide GPCRs into five families called A, B, C, D and E *(see table 1.1)*. Other than the conserved membrane topology there is very little structural or sequence homology between the individual families.

The metabotropic glutamate receptors [mGluRs] are members of Family C, which includes the metabotropic $GABA_B$ receptor, calcium sensing receptor [CaR] and several putative pheromone receptors. The majority of receptors in this family have a large, globular extracellular N-terminal ligand-binding domain [Hermans E., and Challiss, R.A. 2001). Eight mammalian mGluR subtypes have been identified and further subdivided into three groups based upon pharmacological properties and sequence homology. mGluR1 and mGluR5 belong to Group I, mGluR2 and mGluR3 belong to Group II and mGluR4, mGluR6, mGluR7 and mGluR8 belong to Group III.

FAMILY		Representative Receptors	Ligand Binding Mechanism
A	**I**	Rhodopsin, Adenosine, Adrenergic, Serotoninergic, Opioid	Group 1a GPCRs bind small ligands, such as catecholamines and rhodopsin. The ligand binding site is located within the 7-transmembrane domain.
	II	Peptide, Cytokine, Thrombin	Group 1b GPCRs bind peptides, the binding site is comprised of the N-terminal, extracellular loops and superior parts of the transmembrane domain.
	III	Glycoprotein hormone receptors [Leutinising hormone receptor]	The Group 1c GPCRs have a large extracellular domain, this makes up most of the ligand binding site, although contacts are also made with extracellular loops e1 and e2
B		High molecular weight hormone receptors, glucagon, Calcitonin, secretine, PTH, VIP receptors and the a-latrotoxin receptor	The Family 2 GPCRs are similar in morphology to Group 1C and exhibit a very similar ligand binding domain, yet there is no sequence homology between the two groups.
C		mGluRs, Ca2+ sensing receptors, GABAB, pheromone receptors	The ligand binding site is comprised of a large globular, bilobed extracellular domain
D		Pheromone receptors associated with Gi	
E		'Frizzled' and 'smoothened' receptors involved in embryonic development and cell polarity.	

Table 1.1 GPCRs can be classified into 3 main families based on sequence homology and molecular mechanisms of ligand binding. [Adapted from Bockaert, J et al 1999] The mGluRs are classified as Family C.

1.7   The functional relationship between the mGluR and iGluR subclass

The Group I mGluRs and the NMDA receptors are regularly found in close proximity on the postsynaptic membrane [Lujan, R. et al 1996] and can interact bi-directionally with each other. Both mGluR1 and 5 have been shown to potentiate NMDA evoked currents in different neuronal populations [Awad et al 2000 and Heidinger et al 2002] and low concentrations of NMDA can potentiate mGluR-mediated increases in phosphoinositide hydrolysis [Algarsamy et al 1999b]. Recent studies suggest NMDA modulates mGluR5 by the activation of a calcium-dependent protein phosphatase 2B, calcineurin. It is proposed calcineurin exists in a complex with mGluR5 and potentiates mGluR5 signalling by dephosphorylating sites in the C-terminal phosphorylated by Protein Kinase C (PKC) responsible for desensitization *(see section 1.13)* [Algarsamy S. et al 2005]. Interactions between mGluRs and NMDA receptors may underlie several pathophysiological and physiological processes, such as NMDA-receptor dependent forms of learning [Collingridge, G.L. and Bliss, T.V. 1995].

1.8   mGluR structural domains that underlie function

The mGluRs contain several important structural domains and each of these performs a specific role important for the receptors function and transmembrane signal transduction [Pin, J et al 2003]. There are five domains; the N-terminal ligand binding domain, a cysteine rich domain, the 7 transmembrane α-helices, a G-protein coupling domain and an intracellular carboxyl-terminal *(see figure 1.2)*. When glutamate binds to the extracellular N-terminal ligand-binding domain of the mGluR the receptor undergoes a conformational change. This is transmitted through the 7 α-helical domain and across the plasma membrane, enabling the activation of G-proteins. The G-proteins are positioned on the intracellular face of the plasma membrane. When activated by the receptor the G-proteins can activate intracellular effectors, these include ion channels and enzymes [Conn, P.J. and Pin, J.P 1997] .

1.8.1   The extracellular domain

The large mGluR extracellular domain consists of the agonist binding region and the cysteine rich domain *(see Section 1.8.3)*. The resolution of the mGluR1a extracellular domain by X-ray crystallography established that mGluR1a is a homodimeric molecule. The extracellular domain of each subunit is composed of 2 lobes, separated by a cleft where the agonist binds [Kunishima, N et al 2000]. The bi-lobate structure of each subunit in the dimer is capable of oscillating between two relative states. In the absence of agonist the dimer is in a 'resting state' and the bi-lobate structure of each subunit is open. In the presence of agonist the dimer is in an 'active state' and either one or both of the bi-lobate structures are closed. More recently it has been established that the binding of agonist to one subunit of an mGluR5 dimer is sufficient to activate the receptor but agonist binding to the second subunit is necessary for full activity. Full activity of the receptor is achieved when the bi-lobate strucutre of both subunits is in the closed conformation [Kniazeff J et al 2004].

1.8.2   GPCR Dimerization

GPCR dimerization has been shown to be functionally important [Jingami, H et al 2003]. For example, the heterodimerization of the GABA$_B$

receptor subunits is necessary for both signal transduction [Galvez et al 2001 and 2000] and intracellular trafficking from the endoplasmic reticulum to the plasma membrane [Margeta-Mitrovic et al 2000]. This works by the B2 subunit masking an endoplasmic retention signal in the C-terminal of the B1 subunit. This facilitates and selects the transport of a functional heterodimer to the cell surface, since it is the B1 subunit that binds GABA but it is the B2 subunit only that couples to the G-protein. Heterodimerization between different mGluR subtypes has not been reported. However, mGluR1 is capable of heterodimerization with the calcium sensing receptor [CaR] which belongs to the same GPCR family as the mGluRs and the adenosine A1 receptor [discussed in Jingami, H et al 2003]. Homodimerization of mGluR subtypes is dependent upon the N-terminal [Robbins, MJ et al 1999] [described in Section 1.6.2]. Though the inability of the mGluR1$_a$ and mGluR1$_b$ C-terminal splice variants to heterodimerize suggests interactions between other structural domains, such as the C-terminal, may be involved [Reviewed in Hermans, E et al 2001]. Indeed, the heterodimerization of GABA$_B$ subunits involves interactions between the C-terminals of the individual receptors [Margeta-Mirovic et al 2000].

1.8.3 Cysteine rich Domain

The cysteine rich domain connects the bi-lobed structure to the transmembrane domain. It contains 9 conserved cysteine residues but its purpose is currently unknown. However, it is required for mGluR function [Tsuji, Y et al 2000, Ray, K et al 2000] and may be involved in dimerization [Romano, C et al 2001] and/or transmitting conformational changes at the N-terminal upon ligand binding to the transmembrane and intracellular domains. This would make the cysteine rich domain an important intermediate in translating receptor activation into G-protein coupling.

1.8.4 The intracellular loops

The second and third intracellular loops of mGluRs couple the receptors to G-protein signalling molecules *(this is discussed in section 1.10)* [Francesconi, A and Duvoisin R.M 1998]. The second intracellular loop is long and poorly conserved amongst mGluR subtypes. It determines the selectivity of the G-protein activated by the receptor [Gomeza, J et al 1996]. The second intracellular loop

contains an evolutionarily conserved glutamine residue that can regulate receptor activity [Yamashita, T et al 2004]. The third intracellular loop is shorter and highly conserved, it plays a modulatory role in the coupling efficiency of the receptor to the G-protein [Francesconi, A and Duvoisin R.M 1998].

1.8.5 The intracellular C-terminal Domain

The intracellular C-terminal of mGluRs is important for directing the receptors function. It is the least conserved domain amongst mGluR subtypes. Alternative splicing of this domain produces further receptor heterogeneity. C-terminal splice variants of mGluR1[1_a,1_b,1_c,1_d], mGluR4[4_a, 4_b] mGluR5[5_a, 5_b], mGluR7[7_a, 7_b] and mGluR8[8_a, 8_b] have been identified and cloned. The C-terminal of mGluRs has been identified as mediating interactions with intracellular signalling molecules *(see table 1.5)* [Fagni, L et al 2004]. Long and short C-terminal splice variants of individual receptor subtypes can therefore have contrasting signalling properties [Tu JC et al 1998, Flor, PJ et al 1996 and Mary, S et al 1998].

The C-terminal can undergo post-translational modifications, these include phosphorylation and dephosphorylation of specific residues and this provides one mechanism for modulating signalling at the C-terminal [Cai Z et al 2001]. The contribution of the intracellular C-terminal to G-protein coupling [Pin, J-P. et al 1994 and El Far, O. et al 2001] *(discussed in section 1.15)*, targeting and trafficking of the mGluRs to neuronal compartments *(see section 1.21)*, demonstrates the importance of this domain to mGluR function.

1.8.6 Cross-talk between mGluR domains

The mechanisms involved in translating ligand binding to the receptor, across the transmembrane to the intracellular G-protein coupled domains are beginning to emerge. The efficiency of agonist binding to the *Drosophila* mGluR homologue DmGluRA has been shown to be regulated by the properties of the plasma membrane traversed by the heptahelical domain. High affinity agonist binding is dependent upon the membrane ergosterol content and the incorporation of the receptor into sterol induced lipid rafts. Furthermore cholesterol can directly interact with DmGluRA. This suggests interaction between the heptahelical

domain and the lipid bilayer may play an important role in modulating ligand binding to the extracellular N-terminal [Eroglu, C et al 2003].

The Rhodopsin receptor belongs to the Class A family of GPCRs. It has been used as a model GPCR in the elucidation of molecular mechanisms of translating ligand binding into G-protein activation. It has been shown that the receptor undergoes a conformational change, within the 7-transmembrane domain, when in the active state. This involves a change in the orientation of transmembrane helices III and VI, relative to each other, which unveils a G-protein binding site within intracellular loops 2 and 3 [Farrens, DL et al 1996]. By use of fluorescence resonance energy transfer [FRET] it has been shown that agonist binding causes conformational changes in specific intracellular domains of the mGluR1a dimer [Tateyama, M, Hideki, A et al 2004]. One domain that undergoes rearrangement is the 2nd intracellular loop, a domain that is essential for G-protein coupling and activation *(see section 1.8.4)*. The rearrangement of the 2nd intracellular loop by agonist binding may represent a mechanism of effector activation and signal transduction to the intracellular side of the cell membrane.
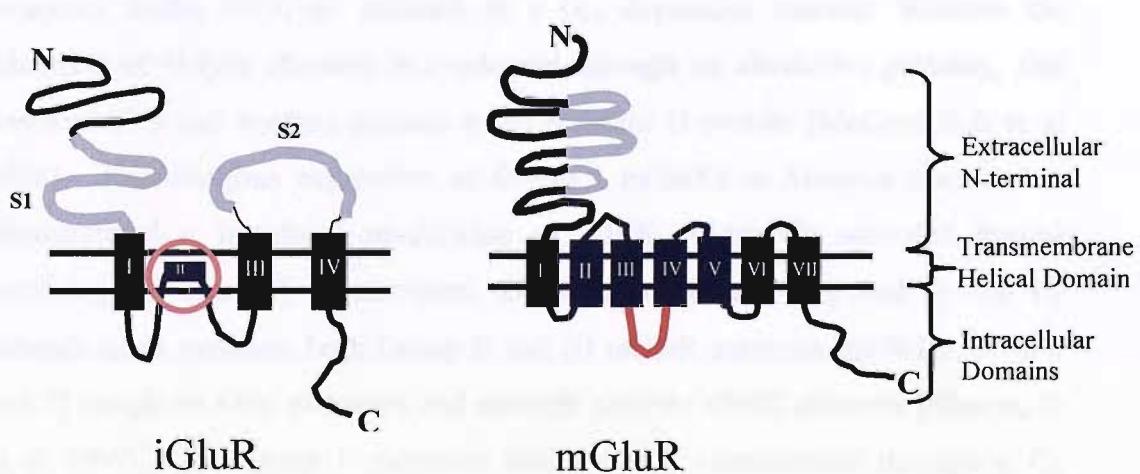
Figure 1.2 **Topological organisation of glutamatergic receptor subunits and annotation of structural domains underlying function.** The crystal structure the ligand binding domain of the NMDA subunit NR1 [Furukawa, H and Gouaux, E 2003] and AMPA GluR2 subunit (Armstrong N et al 2000) has identified the two extracellular loops, denoted S1 and S2 are important regions involved in ligand binding. Ligand binds to a cleft formed between the S1 and S2 lobes. The region of the iGluR circled in red is involved in formation of the ion pore. The grey represents the agonist binding site and is an indication of the regions involved in agonist recognition and binding.

The 2nd intracellular loop of the mGluR is highlighted in red. This region is involved in G-protein coupling.

1.9 The physiological function of mGluR subtypes in the mammalian CNS

The phenotypic characterisation of mGluR knockout transgenic lines has provided insight into the role mGluR signalling plays in the physiological processes modulated by glutamate *(see table 1.2)*. Knocking out mGluR subtypes does not cause lethality, unlike ionotropic receptor subunit NR1 [Forrest D et al 1994]. Instead, the phenotypes generated are consistent with the neuro-modulatory role mGluRs perform in synaptic transmission.

GROUP	SUBTYPE	PHENOTYPE
I	mGluR 1	Ataxia gait, intention tremor, motor disco-ordination, deficient LTD, ↓ hippocampal LTP[3,4,5,6]
I	mGluR 5	↓LTP in hippocampus, reduced spatial and contextual learning[7]
II	mGluR 2	Impaired hippocampal LTD[2]
II	mGluR 3	
III	mGluR 4	Impaired motor learning and short term synaptic plasticity[10]
III	mGluR 6	Loss of ON response in retinal bipolar cells[11]
III	mGluR 7	Deficient in aversion learning [fear response and conditioned taste aversion][8]
III	mGluR 8	↑ Anxiety related behaviour[9]

Table 1.2 Phenotypes associated with mGluR subtype knockout in either mice or rats. [1] Masu, M et al 1995. [2] Yokoi, M et al 1996. [3] Conquet, F et al 1994. [4] Ichise, T et al 2000. [5] Aiba, A et al 1994. [6] Aiba, A et al 1994. [7] Lu, Y-M et al 1997. [8] Masugi, M et al 1999 [9] Linden, A-M et al 2002 [10] Pekhletski, R et al 1996

1.10 mGluR pharmacology

mGluR agonists tend to fall into three categories, those that are receptor specific, those that are subgroup specific and those that are broad-spectrum compounds that act at more than one subgroup *(see table 1.3)*. Receptor specific

compounds that allow discrimination between different receptors of the same subgroup have not been developed for every mGluR subtype. Instead the majority of compounds are subgroup selective but exhibit different rank potencies at the receptor subtypes within the subgroup they are selective for. Broad-spectrum agents can provide a preliminary starting point when the pharmacology of the receptor being investigated is largely unknown, such as in the case of mammalian receptor homologues in different species. However, although broad-spectrum agents can interact with more than a single subgroup, the potency of such compounds can vary between subgroups by upto 100 fold [Conn, P.J and Pin, J-P 1997]. The mGluR agonists have been identified as having actions outside of the mammalian phyla and have been used to characterize the pharmacology of mammalian mGluR receptor homologues [Parmentier ML et al 2000, Funada M et al 2004].

GROUP	RECEPTOR	AGONIST
I	mGluR1a	quisqualate > *3, 5-dihydroxyphenylglycine [3, 5-DHPG] glutamate*>**1S,3R-1-amino-1, 3-cyclopentanedicarboxylate [ACPD]** = ibotenate > **[2S, 1'S,2'S]-2[carboxycyclopropyl] glycine [L-CCG-I]**>3-hydroxyphenylglycine [3-HPG] > *trans*-azetidine-2, 4-dicarboxylate [t-ADA]
II	mGluR2	[2S, 10R, 20R, 30R]-2-[2, 3-dicarboxycyclopropyl] glycine [DCG-IV]=**L-CCG-I**>*2R, 4R-4-aminopyrrolidine-2, 4-dicarboxylate [APDC]* > glutamate > 1S, 3S-ACPD = **1S, 3R-ACPD** > 4C3HPG> ibotenate
III	mGluR4a	*L-amino-4-phosphonobutyrate [L-AP4]* > L-serine-O-phosphate [L-SOP] > glutamate **L-CCG-I** > 1S, 3S-ACPD >> **1S, 3R-ACPD**

Table 1.3 The rank order of potency of agonists characterized at prototypic group members [Conn, P.J.and Pin, J-P. 1997]. Broad-spectrum agonists that interact with all three groups of mGluR are shown in bold and group specific agonists are italicized.

1.11 Mechanism of signalling by GPCRs

GPCRs are functionally coupled to their target proteins by trimeric GTP-binding proteins [Wess J 1997]. The targets of mGluR activated G-proteins can be intracellular enzymes or membrane bound ion channels. G-proteins are GTPases and can switch between two states; active when GTP is bound and inactive when GDP is bound [Kaziro Y et al 1991]. In the active state the G-protein can switch on signalling cascades. Ligand binding to the extracellular domain of the mGluR causes a conformational change in the receptor and the G-protein is activated by exchanging the bound GDP for GTP. The G-protein enters an inactive state when it hydrolyzes the bound GTP, turning it back into GDP. [Gudermann, T et al 1997]

Trimeric GTP-binding proteins contain three subunits, called α, β and γ. The α subunit binds GTP and hydrolyzes it to GDP, the β and γ complex anchors the G-protein to the plasma membrane through a prenyl group covalently linked to the γ subunit [Fukada Y et al 1990]. When activated the α subunit exchanges GDP for GTP, this causes it to disassociate from the $\beta\gamma$ complex. The α-GTP subunit and the $\beta\gamma$ complex have specific protein targets and trigger distinct signalling events [Gilman, AG 1987].

The trimeric G-proteins can be divided into three major families, this classification is based upon amino acid sequence homology between α subunits. Members of each family are evolutionarily conserved [Wilkie, T M 1999]. In *C. elegans* the $G_{i/o}$, G_s, G_q and G_{12} orthologs are goa-1, gsa-1, egl-30 and gpa-12 respectively. The sequence identity is 80%-60% between the *C. elegans* $G\alpha$ subunits and their mammalian counterparts [Wikie, TM et al 1999]. In mammals about 16 α subunits, 5 β subunits and 12 γ subunits have been identified. The genome of *C. elegans* encodes 20 α subunits, 2 β subunits and 2 γ subunits [Bargmann, CI et al 1998]. This variety of G-proteins allows GPCRs to activate a diverse range of signalling molecules.

1.12 Differential G-protein coupling of mGluR subgroups and subtypes

The heterologous cell expression of mGluR subtypes has identified the G-proteins they activate and the signalling cascades they couple to. The activation of

Group I receptors expressed in *Xenopus* oocytes activates G_q, stimulating phospholipase C and the subsequent production of the second messengers inositol-1,4,5-trisphosphate (IP$_3$) and diacylglycerol (DAG) from phosphatidylinositol 4,5-bisphosphate (PIP2) [Sugiyama, H et al 1987]. The production of IP$_3$ stimulates the release of Ca^{2+} from the endoplasmic reticulum (ER) by binding to IP$_3$-gated calcium channels in the plasma membrane of the ER and this causes an increase in the concentration of cytosolic calcium. DAG activates the enzyme PKC, a serine/ thronine protein kinase that can phosphorylate specific targets. The activation of Group II [Schoepp, DD et al 1995] and III [Saugstad, JA et al 1994, Flor, PJ et al 1995, Saugstad, JA et al 1997] receptors activates $G_{i/o}$ and inhibits the adenosine 3':5'-cyclic monophosphate (cAMP) biosynthetic enzyme Adenyl Cyclase.

mGluR subtypes can utilise a variety of signal transduction mechanisms. This is exemplified by the Group I mGluRs because heterologous expression of rat Group I mGluRs in HEK 293 cells has revealed multiple pathways couple the receptors to subtypes of neuronal voltage-gated Ca^{2+} channels that trigger fast neurotransmitter release at synapses. In response to agonist activation Group I receptors inhibit P/Q-type channels in a $G_{i/o}$ dependent manner. Whereas the inhibition of N-type channels is conducted through an alternative pathway, that involves Gi/o and another pertusis toxin resistant G-protein [McCool B.A et al 1998]. Heterologous expression of Group I mGluRs in *Xenopus oocytes* has demonstrated a functional modulation of GIRK [G-protein activated inward rectifying K$^+$ channel] conductance. GIRK channels are activated by the βγ subunits of G_i proteins, both Group II and III mGluR subtypes [mGluR2, 3, 4, 6 and 7] couple to Gi/o pathways and strongly activate GIRK channels [Sharon, D et al 1997]. The Group I receptors inhibit GIRK conductance through a G_q mediated pathway but at higher levels of receptor expression, GIRK channels have been shown to be activated by mGluR1$_a$ through a $G_{i/o}$ pathway. Furthermore distinct regions of the mGluR1a receptor have been shown to trigger the activation of G_s and couple the receptor to different intracellular targets [Francesconi, A et al 1998].

Depending on the recombinant cell system mGluR subtypes are expressed in, different G-proteins can be switched on and different targets activated. In BHK

cells, mGluR1a can activate both $G_{q/11}$ and $G_{i/o}$ subclasses of G-protein to modulate signal transduction through PLC [Hermans, E et al 2000]. This, however, was not reproducible when mGluR1$_a$ was recombinantly expressed in CHO cells [Selkirk, J.V. et al 2001]. It would appear mGluRs have the capacity to activate more than one G-protein in a cell type dependent manner.

1.13   The regulation of mGluR signalling

The mGluRs are regulated by heterologous and homologous mechanisms [Hermans, E et al 2001, De Basi et al 2001]. Group I receptors undergo agonist-induced desensitisation through a PKC specific mediated phosphorylation. PKC directly phosphorylates Serine and Threonine residues of the intracellular loops and C-terminal of mGluR5 [Gereau, RW et al 1998]. GPCR kinases [GRKs] have been identified as an additional mechanism of Group I regulation [Sallese, M et al 2000]. GRKs target the C-terminal of the agonist bound receptor for phosphorylation, creating an interaction site for β arrestin, which is involved in disrupting G-protein coupling and receptor sequestering [Krupnick, JG et al 1998].

1.14   The Regulators of G-protein signalling [RGS]

The RGS family contains approximately 20 different subtypes, which can control the signalling properties of GPCRs and their G-proteins. RGS proteins interact directly with members of the G-protein families to negatively regulate their activity. This is achieved by accelerating the intrinsic GTPase activity of the G_α subunit and by the modulation of G_α-GTP interactions with intracellular mediators[Ishii M et al 2003]. Group I and Group III mGluR subtypes have been identified as substrates for RGS. The RGS4 subtype is specifically expressed in brain and abolishes mGluR1$_a$ and mGluR5$_a$ mediated signalling via G_q in *Xenopus oocytes* and hippocampal CA1 neurons [Saugstad, J.A et al 1998]. The RGS2 subtype has been shown to modulate the activation of different G-protein pathways by mGluR1a when heterologously expressed in rat sympathetic neurons [Kammermeir, P.J. et al 1999]. In the retina mGluR6 couples to a G_{o1} transduction pathways to mediate the ON bipolar cell response [Nawy S.1999]. Using the constitutively active α subunit of G_{o1} as bait in a GAL4 yeast-2-hybrid screen a

retinal specific RGS was isolated and classified as Ret-RGS1 [Dhingra, A et al 2004]. Immunocytochemistry has identified Ret-RGS1 is localised to the dendritic compartments of bipolar cells along with mGluR6 and G_{o1}. The co- expression of Ret-RGS1 in *Xenopus oocytes* with mGluR6 and G_{o1} accelerates the inactivation of the signalling cascades and reveals a role for Ret-RGS1 as a potent GTPase activator for G_{o1} and regulator of mGluR6 signalling.

1.15   The mGluR C-terminal can direct receptor signalling pathways

The C-terminal can direct subtype specific coupling to neuronal voltage-gated calcium channels. The swapping of mGluR7 and mGluR2 C-terminals causes each to take on the other receptors functional properties. The activation of mGluR2 and 7 stimulates characteristic cellular responses; mGluR 2 inhibits N- and L-type Ca^{2+} channels and mGluR 7 blocks P/Q-type Ca^{2+} channels [Perroy, J et al 2001]. A chimera of mGluR7 and the mGluR2 C-terminal coupled mGluR 7 activation to the inhibition of N and L-type Ca^{2+} channels. A chimera of mGluR2 and the mGluR7 C-terminal coupled mGluR2 activation to the inhibition of P/Q-type Ca^{2+} channels.

The C-terminal domain is also capable of regulating ion channels independently of G-protein activation. The C-terminal of the $mGluR1_a$ receptor subtype can interact directly with the C-terminus and long intracellular loop of the $Ca_v2.1$ subunit of P/Q-type Ca^{2+} channels to modulate their conductance properties [Kitano, J et al 2003]. Through this mechanism $mGluR1_a$ has the potential to co-ordinate Ca^{2+} signalling events.

1.16 Receptor scaffolding complexes

It is now apparent receptor signalling can no longer be thought of as a linear process, consisting of a GPCR coupling to a G-protein and activating a intracellular messenger. Instead receptor signalling cascades are emerging as complex three-dimensional protein networks *(see figure 1.3)* and the interactions between components of these networks leads to the modulation and regulation of the signal being conducted. Scaffolding proteins have been identified that organize and cross-link signalling molecules within these networks, at specific membrane domains, for efficient intercellular communication. In-light of this scaffolding proteins can be thought of as performing three important roles for receptor function:

- **Signalling:** *The assembly of the receptors intracellular signalling cascade*

- **Targeting:** *The localization of receptors to specific compartments*

- **Anchoring:** *The retention and stabilization of receptors within ultrastructural domains*

Scaffolding proteins that target and anchor receptors within distinct sub-cellular compartments perform an important role in the cellular polarisation of signalling pathways *(see section 1.21).*
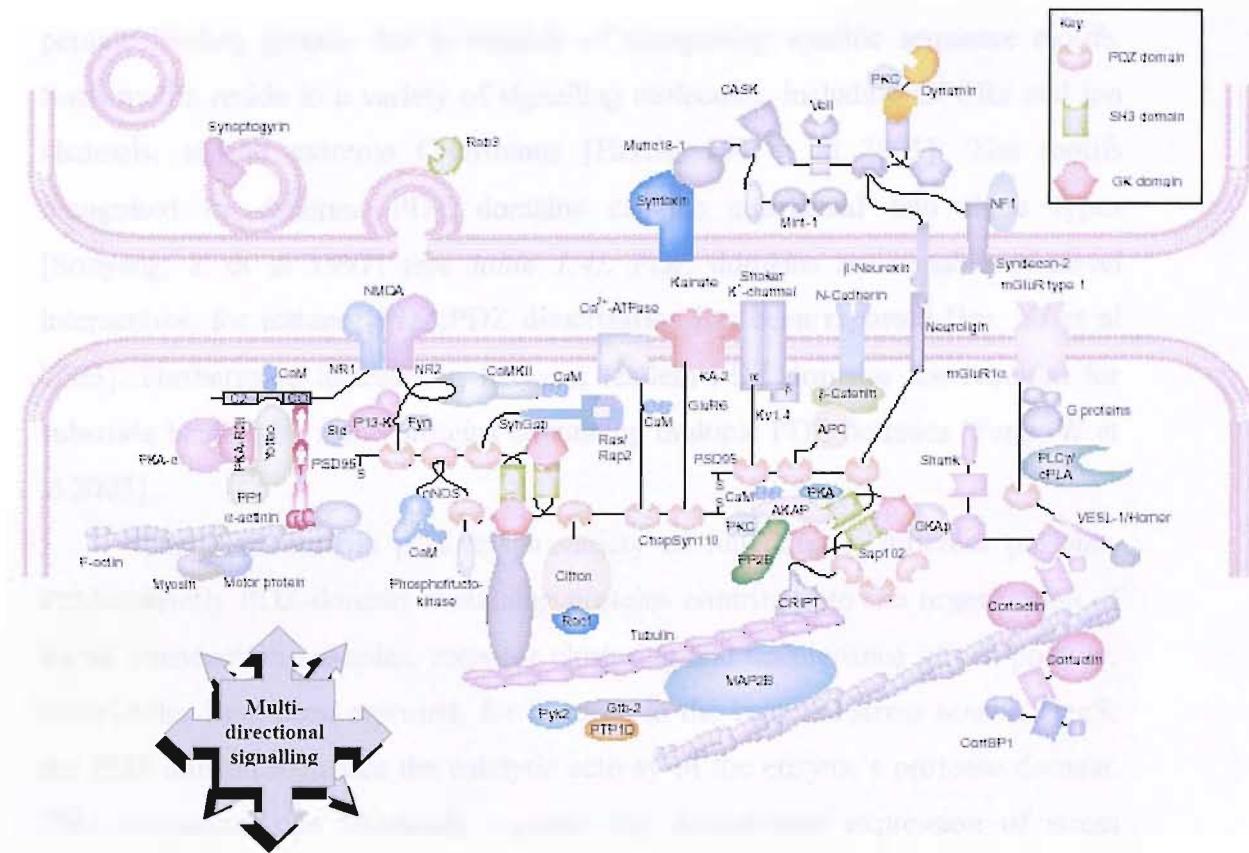
Figure 1.3 Receptor signalling cascades are 3-dimensional networks of protein-protein interactions. [Adapted from Grant, S. et al 2001]. The diagram represents the complexity and highly ordered nature of signalling networks. PSD-95 is a prominent example of a scaffolding protein of the postsynaptic density and organises NMDA receptor signalling. The arrows represent how communication within these networks is multi-directional and not solely in a linear field, allowing cross talk between different receptors and downstream components of different signalling pathways.

1.17 The PDZ Domain- a backbone for organizing scaffolding networks

The PDZ domain is so called because of its initial isolation in proteins that direct protein-protein interactions: **P**SD-95, a mammalian postsynaptic density protein; the *Drosophila* septate junction protein, **D**iscs-large; the mammalian epithelial tight-junction protein **Z**ona occludens. The human genome predicts 440 PDZ domains in 259 different proteins [Hung, A.Y et al 2002]. Crystallization has shown the structure of the PDZ domain is highly organized and globular [Doyle, D et al 1996]. A PDZ domain is typically ~90 amino acids and folds to generate a

29

peptide binding groove that is capable of recognising specific sequence motifs. Such motifs reside in a variety of signalling molecules, including GPCRs and ion channels, at the extreme C-terminus [Harris, BZ et al 2001]. The motifs recognised by different PDZ domains can be subclassed into three types [Sonyang, Z et al 1997] *(see table 1.4)*. PDZ domains are capable of novel interactions, for instance PDZ:PDZ dimerization has been reported [Im, YJ et al 2003]. Furthermore interactions between tandem PDZ domains are required for substrate binding by some proteins containing multiple PDZ domains [Feng, W et al 2003].

The PDZ domain performs a variety of functions in different proteins. Predominantly PDZ-domain containing proteins contribute to the organization of signal transduction cascades, receptor clustering and maintenance of cell polarity. Novel roles have been reported, for instance in the bacterial stress sensor, DegS, the PDZ domain regulates the catalytic activity of the enzyme's protease domain. This mechanism can ultimately regulate the downstream expression of stress response genes during adverse external conditions [Walsh, N.P. et al 2003].

PDZ DOMAIN	Consensus binding sequence
Class I	S/T X Φ - COOH
Class II	Φ X Φ - COOH
Class III	X X C - COOH

Table 1.4 Classification of PDZ Domain Recognition motifs. X denotes any amino acid, Φ denotes a hydrophobic amino acid, usually Valine, Isoleucine or Leucine.

1.18 mGluR subtype interactions with scaffolding proteins

The C-terminal domain of the mGluR subtypes is an important domain for scaffolding mGluR function in terms of its signalling, targetting and trafficking *(discussed in sections 1.19.1 and 1.21 respectively)*. The functional importance of the mGluR subtypes intracellular C-terminal is further highlighted by the interactions it specifies with a variety of scaffolding proteins *(see table 1.5)*, not all of which contain a PDZ domain. The most characterized scaffold protein interactions are known to perform important cellular functions. The scaffolding

proteins in *table 1.5* are involved in either targeting or anchoring of mGluRs or they are involved in mGluR signalling. In some cases a single scaffolding protein is involved in all three processes. The importance of mGluR targeting and anchoring for synaptic signalling is discussed in *section 1.21* and the molecular mechanisms involved are highlighted and using Homer as an example, the role performed by scaffolding proteins in mGluR signalling is discussed in the next section.

GROUP	RECEPTOR SUBTYPE	INTERACTORS
I	mGluR 1	Homer[1], Calmodulin[2], Siah-1A[3], Tubulin[4], Tamalin[6] CaV2.1[2]
	mGluR 5	Calmodulin[3], Siah-1A[3], Homer[1], Tamalin[6], Filamin A[5]
II	mGluR2	Microtubule associated serine/threonine kinase [MAST] [7] Syntrophin associated serine/thronine kinase [SAST][7] Rhophillin, Protein Interacting[7] Cousin of Thyrodoxin [PICOT][7], Syntenin[8], GRIP[8]
	mGluR 3	
III	mGluR 4	Syntenin[8], GRIP[8], G-protein βγ subunit complex,[13] Filamin A[5], Calmodulin[13]
	mGluR 6	Syntenin[8], GRIP[8]
	mGluR 7	G-protein βγ subunits[10], PICK1[9,12], Calmodulin[13], PKC [11], Syntenin[8], GRIP[8], Filamin A[5]
	mGluR 8	Filamin A[5], Calmodulin[13], G-protein βγ subunit complex[13]

Table 1.5 Identified mGluR scaffolding proteins. [1] Tu, JC et al 1998. [2] Kitano, J et al 2003 [3] Ishikawa, K et al 1999. [4] Ciruela, F et al 1999 [5] Enz, R et al 2002 [6] Kitano, J. et al 2002 [7] Pilkington, B [unpublished] 2004 [8] Hirbec, H et al 2002. [9] Enz, R et al 2003. [10] O'Connor, V et al 1999. [11] Nakijima, Y et al 1999. [12] El Far, O et al 2000 [13] El Far, O et al 2001.

1.19   Homer: a family of scaffolding proteins

In mammals three genes encode Homer proteins and give rise to the three major isoforms: Homer1 [H1], Homer2[H2] and Homer3[H3] [de Bartolomeis A, Iasevoli F. 2003]. A single Homer-related gene *(homer)* has been isolated and characterized in *Drosophila* [Diagana, T.T. et al 2002]. However, searches of the *C. elegans* genome with either mammalian or *Drosophila* protein sequences did not identify a *C. elegans* homologue of Homer. The three mammalian Homer proteins share a common domain architecture. An EVH1 [Drosophila Enabled/vasodilator-stimulated phosphoprotein Homology 1] domain is located at the N-terminal [Kato, A et al 1997] and the C-terminal contains a coiled-coil [CC] motif [Sun, J et al 1998]. Both domains are involved in protein-protein interactions. The coil-coiled domain mediates the assembly of Homer proteins into multimeric complexes [de Bartolomeis A, Iasevoli F. (2003)]. The EVH1 domain is required for interactions with the IP$_3$-gated calcium channel [Tu, JC et al 1998], potentially ryanodine receptors [Hwang, SY et al 2003], Group I mGluRs [Tu, JC et al 1998] and SHANK proteins [Tu, JC et al 1999].

Splice variants of each isoform have been identified and can be subdivided into short [H1a] and long variants [H1b, H1c, H2a, H2b and H3] [de Bartolomeis, A et al 2003]. With the exception of H1a all variants are constitutively expressed [Xiao, B et al 1998]. H1a was the first Homer protein to be identified and characterised as an immediate-early gene in the hippocampus and cortex [Brakeman, PR et al 1997]. Synaptic activity regulates transcription from the H1 gene. Increased synaptic activity switches transcription of the H1 gene from the constitutive isoforms H1c and H1b to the shorter transcript for H1a. Transcription terminates in an intron after the EVH1 domain and before the CC motif, consequently the H1a protein lacks the CC domain [Bottai, D et al 2002]. This has important functional implications. The long Homer isoforms can multimerize, via the CC domain [Kato, A et al 1998] and this is an important feature for cross linking signalling proteins. As H1a lacks the CC domain, it cannot self-associate but can compete with the longer isoforms for binding to signalling molecules. Because H1a cannot oligomerize it disrupts the assembly of scaffolding networks and functions as a dominant negative regulator [Tu JC et al 1998 and Xiao, B et al 1998] .

1.19.1 Homer interacts with the C-terminal of Group I receptors and scaffolds their signalling

The Homer proteins bind to a proline rich motif in the C-terminal of Group I mGluR subtypes mGluR1$_a$, mGluR5$_a$ and mGluR5$_b$. Homer recognises similar ligand motifs in the IP$_3$-gated Ca^{2+} channel, ryanodine receptors, TRP channels and SHANK proteins [Tu JC et al 1998]. Multimers of Homer assemble targets of mGluR signalling. The mGluR, IP$_3$-gated receptor and Homer have been shown to assemble into a complex using biochemical techniques [Tu, JC et al 1998]. In this complex Homer functions as a physical cross-link between the mGluR and the IP$_3$ receptor to optimise the mobilisation of calcium from the ER upon mGluR activation. The SHANK proteins are known to interact with the NMDA receptor scaffolding protein PSD-95 [Naisbitt, S et al 1999]. Complex formation between Homer-shank-PSD-95 may cross-link the mGluRs to the NMDA receptors [Tu, JC et al 1999].

The expression of mGluR1$_a$ and mGluR5$_a$ in cultured cerebellar granule cells has identified interactions with Homer that can directly regulate the activity of the receptor. The disruption of the interaction between the mGluRs and the endogenously expressed H3 protein, by either RNA anitisense knockdown of H3 or mutagenesis of the Homer ligand motif, leads to constitutive receptor activity. The induction of native H1a expression by exposure of the cultured neurons to NMDA and kainate also switches on constitutive signalling by mGluR1$_a$ and mGluR5$_a$ [Ango, F et al 2001]. This reveals a Homer dependent mechanism for the regulation of mGluR receptor activity in response to neuronal excitation.

1.20   Neurons are morphologically and functionally polarised structures

As highlighted previously *(see section 1.2)*, neuronal cells are polarized structures containing two morphological specializations, these are dendritic branches and a single axon. The two compartments are functionally distinct from each other and can extend from opposite poles of the cell body [Bradke, F and Dotti CG 200]. The dendrites receive extracellular signals from neighbouring cells and at the synapse the dendritic membrane is postsynaptic. The axon transmits signals to neighbouring cells and at the synapse the axon membrane is presynaptic. This functional polarization within the same cell is achieved by the

selective targeting of proteins to the two different compartments and their anchoring within these compartments.

1.21 The differential targeting and anchoring of the mGluR subtypes is important for synaptic signalling

The metabotropic glutamate receptor subtypes have been identified as being differentially targeted to specific neuronal compartments. The Group I mGluRs are predominantly targeted to the postsynaptic membrane and localised to perisynaptic sites [Nusser, Z et al 1994 and Lujan, R et al 1997], forming an annulus around a central core of NMDA receptors. The Group II mGluRs are targeted to both the pre and postsynaptic membranes, here they are localised to preterminal membranes [Lujan, R et al 1996 and Azkue, J.J et al 2000]. The Group III mGluRs are predominantly targeted to the presynaptic membrane, where they are localised close to the site of vesicle fusion and neurotransmitter release [Shigemoto, R et al 1996 and Blumcke, I et al 1996]. The mGluR subtypes expressed presynaptically can regulate the neurotransmitter release properties of the presynaptic membrane in response to glutamate [Dietrich, D et al 2002 and Valenti, O et al 2003]. The Group III receptor subtype specific agonist LAP-4 can inhibit glutamate and GABA release from presynaptic terminals [Matsui, T et al 2003, Lorez, M et al 2003].

Postsynaptic mGluRs are involved in modulating components of the postsynaptic signalling cascade. Reciprocal cross talk between Group I mGluRs and NMDA receptors has been shown to occur and contributes to long term changes in synaptic transmission *(discussed in section 1.7)* [Cho, K et al 2002]. The mGluR1$_a$ subtype interacts with the TRPC1 cation channel biochemically and colocalizes in perisynaptic domains of Purkinge neurons of the cerebellum. The activation of mGluR1$_a$ at purkinje neurons evokes a slow excitatory postsynaptic potential [EPSC] mediated by the TRPC1 cation channel [Kim, S.J et al 2003].

The spatial arrangement of mGluRs in sub-structural microdomains at the synapse has an important biological function. One important property of the presynaptic membrane is the amount of neurotransmitter it releases into the synaptic cleft. This is predominantly governed by the strength of neuronal stimulation. Under high frequency stimulation the amount of neurotransmitter

released into the synaptic cleft is higher compared to low frequencies [Katz, B and Miledi, R 1965]. Due to the physical location of the mGluR subtypes their activation is dependent upon the quantity of neurotransmitter released. This provides a mechanism for the neurotransmitter release properties of the presynaptic membrane to be interpreted and translated into distinct signalling events, depending on where the receptors are positioned relative to the source of transmitter release *(see below)* [Scanziani, M. et al 1997].

The Group I receptors are located towards the edge of synapses and have evolved a relatively high affinity for glutamate [Conn, PJ and Pin, JP 1997], they are activated by bursts of synaptic activity [Batchelor, A.M et al 1994]. The Group II receptors are located further away from the synaptic cleft, they are activated by the diffusion of glutamate from the synaptic cleft to extra synaptic sites during high frequency stimulation [Tempia, F et al 1998]. This is referred to as 'spill over' and can lead to the activation of mGluRs located at the same sites on neighbouring neurons. A role for mGluRs as heteroreceptors has been described in the cerebellum. The release of glutamate from mossy fibres onto the axon terminal of inhibitory Golgi cells depresses GABA release [Mitchel, S.J et al 2000]. The Group II receptors are exposed to low levels of glutamate and so have evolved a high affinity for the molecule [Conn, PJ and Pin, JP 1997]. The Group III receptor subtype mGluR 7 is activated by low frequency stimulation. Under these conditions the receptor experiences a high concentration of neurotransmitter because of its localisation in the active zone membrane [Shigemoto, R et al 1997]. Consequently mGluR 7 has evolved with a lower affinity for glutamate [Okamoto, N et al 1994].

1.22 The mGluR intracellular C-terminal plays a key role in specifying receptor targeting and polarisation

The signal for this highly selective targeting resides in the intracellular C-terminal [Stowel, J et al 1999]. mGluR1$_a$ and mGluR2 are targeted to the dendrites and excluded from the axons of cultured hippocampal neurons. The mGluR 7 subtype is targeted to both sites. The substitution of the mGluR 7 C-terminal for the mGluR 2 C-terminal causes mGluR 7 to be excluded from axons. Whereas the reciprocal of this causes axonal targeting of mGluR 2 [Stowell, J et al 1999]. Thus the expression of mGluRs is controlled by both exclusion and

targeting signals residing in the intracellular C-terminal domain. How these signal are interpreted by the cell to achieve this specific targeting is unclear and may involve directing the receptors into specific vesicular populations, directing specific degradation or stabilization of the receptors within microdomains at the membrane [Marchand, S. and Cartaud, J 2002]. Alternative splicing has been reported to unveil targeting motifs in the C-terminal domain [Francesconi, A et al 2002] of the mGluR1 subtype. When expressed in cultured retinal neurons mGluR1$_b$ and mGluR1$_a$ are addressed to axonal and dendritic compartments respectively. The signal for this differential distribution has been mapped to a tripeptide motif common to all mGluR1 subtypes. The longer C-terminal of mGluR1$_a$ can mask this motif and contains a dominant dendritic targeting signal.

1.23   Homer directs the targeting and anchoring of Group I mGluRs

The role of Homer proteins in targeting Group I mGluR subtypes has been addressed using various heterologous cell expression systems. In HEK-293 and COS-7 cells Homer 1c increases the amount of cell surface mGluR1$_a$ clusters and enhances the signalling capacity of the receptor in response to agonist [Ciruela et al 2000]. Whereas Homer1b has been demonstrated to cause retention of mGluR5 and mGlu1$_a$ in the ER of HeLa cells [Roche, K.W. et al 1999]. In cultured cerebellar granule cells the distribution of mGluR5$_a$ has been shown to be regulated by neuronal excitation and is dependent upon the Homer 1 subtype variant that is being expressed [Ango et al 2000]. In undifferentiated primary cultures the Homer 1b/c variants organise mGluR5$_a$ into clusters in the neurite processes through their capacity to multimerize. This is dependent upon a direct interaction between Homer 1b and the mGluR5$_a$ receptor, disruption of the Homer ligand motif causes mGluR5$_a$ to be restricted to the soma. When co-expressed with Homer 1a mGluR5$_a$ shows a diffuse expression in the cell body and neurites.

Differentiated cerebellar granule cells are fully polarised and have been used to demonstrate Homer 1b and 1a differentially localise mGluR5$_a$ [Ango, F et al 2000]. In the presence of Homer 1b, mGluR5$_a$ is targeted to the cell body and the dendrites but not the axon. In the presence of Homer 1a the pattern of distribution is different, with mGluR5$_a$ being targeted to the axons and dendrites. The distribution of mGluR5$_a$ can be modulated by the induction of Homer 1a in response to neuronal activity. At rest mGluR5$_a$ is detected throughout the cell

body and neurites of undifferentiated cultured neurons, strong depolarisation causes endogenous Homer1a expression and exclusive mobilization of mGluR5$_a$ to neurite processes.

The trafficking of mGluR5$_a$ to the cell surface of cerebellar granule neurons has been shown to also be regulated in a Homer 1a dependent manner. The induction of Homer 1a expression in response to NMDA and kainate exposure, reverses the intracellular retention of mGluR5$_a$ by Homer1b [Ango, F et al 2002]. Functional cell surface signalling by mGluR5$_a$ is demonstrated by cells co-expressing the receptor with Homer 1a but not Homer 1b. This was recorded as an increased amplitude of Ca^{2+} signalling in response to selective activation of mGluR5$_a$. An increased latency in calcium signalling was also recorded from cells co-expressing mGluR5$_a$ with Homer1a, which likely to be due to a lack of direct interaction between the mGluR5$_a$-Homer1a complex and the IP3 receptor.

1.24   Scaffolding proteins contribute to complex mammalian behaviours

Metabotropic glutamate receptors are distributed throughout the cortical and limbic compartments of the brain [Kenny, PJ et al 2004]. The transmission between structures of these brain regions comprises the brain reward circuitry. Drugs of abuse reduce the threshold for signalling through the brain reward circuitry and cause an increase in brain reward function [Bespalov, A et al 1994]. The use of selective pharmacological agents have demonstrated mGluR2/3 receptors are involved in the development of drug addiction and symptoms of withdrawal [Kenny, PJ et al 2003], mGluR5 drives drug consumption [Chiamulera, C et al 2001] and mGluR1 facilitates activation of brain reward function by addictive drugs [De Vry, J et al 2001].

More recently, proteins known to scaffold glutamate receptors have been demonstrated to regulate drug induced behaviours. The knockout of Homer 2 causes mice to exhibit behaviours characteristic of cocaine addiction [Szumlinski, K.K et al 2004]. These include self administration of cocaine, locomotor hyperactivity and conditioned rewarding.

A microarray screen of mouse models of cocaine addiction, generated by the targeted knockout of either individual plasma membrane monoamine (dopamine and norepinephrine) transporters or the vesicular monoamine

transporter (VMAT), identified a down-regulation in the levels of PSD-95 mRNA in the striatum [Yao, W et al 2004]. PSD-95 is a prominent scaffolding protein of the post-synaptic density and has a number of interaction partners, including NMDA receptor subunits [Kornau, HC et al 1995]. Furthermore, the ablation of PSD-95 recapitulates specific behaviours characteristic of cocaine addiction.

1.25 Outstanding issues and how *C. elegans* can assist

An array of mammalian mGluR scaffolding molecules have been identified *(see table 1.5)* and studies performed at the molecular, biochemical and cellular levels have enhanced our understanding of how these interactions occur and their relevance to receptor function. However, with few exceptions *(described in section 1.24)* the physiological significance of these receptor signalling complexes remains unclear, primarily because of the complexity of the mammalian nervous system.

The nematode *C. elegans* provides a model system in which to better understand the link between interacting genes, signalling complexes, neuronal circuitries and behaviour within a single animal. The advantage of *C. elegans* is that the physiological relevance of the receptor interaction identified *in-vitro* can be immediately assessed by determining if the receptor and candidate interacting protein are co-localized within the same neurons at the level of the intact animal.

The behavioural assays *C. elegans* offers then enable the function of the receptor scaffolding complex to be assessed. Characterizing the biological relevance of a mammalian protein-protein interaction is a much longer process. For example the role of the scaffolding protein PSD-95 in addiction behaviour has only just begun to be understood [Yao, W et al 2004], 10 years after it was initially identified as interacting with the NR1 subunit in 1995 [Kornau, H.C. et al 1995]. Therefore *C. elegans* provides a complimentary approach to further our understanding of how glutamatergic signalling complexes are organised and how signalling through protein-protein interaction networks mediate behaviour.

1.26 The use of model organisms in neuroscience

The human nervous system is comprised of billions of neurons that make billions of connections. These neurons are arranged into neural circuits that function to process information and mediate behaviour. Model organisms with simpler nervous systems provide a basis for studying how neural networks operate to control behaviour.

Invertebrate model organisms have made a significant contribution to our understanding of neuronal function. The nervous systems of invertebrates share many of the same biological principles as the mammalian nervous system. This is highlighted by the historical contribution invertebrates have made to our understanding of how neurons work, in terms of axon conductance [Hodgkin AL, Huxley AF 1952] and synaptic transmission [Katz, B. et al 1965]. The use of invertebrate model systems has now progressed to analysing the relationship between genes, molecules, neurones, circuitries and behaviour [de Bono, M and Maricq, A.V 2005].

The nematode *C. elegans* is one such model system that is being used to probe complex behaviours, such as social behaviour at the genomic, cellular and systems biology level [de Bono, M. 2003]. Such studies have been made possible by the availability of the sequenced genome together with: the powerful genetics of *C. elegans*; its amenability to transgenic techniques and the behaviours it displays *(see Section 1.27)*.

1.26.1 General description of *C. elegans*

C. elegans is a small non-parasitic nematode found free-living in soil, where it feeds on bacteria. It has a life cycle of 3 days, consisting of 4 larval stages before reaching adulthood. The mature adult grows to 1mm in size and is fertile for four days. The two sexes are male and hermaphrodite, the latter produce oocytes and sperm enabling self-fertilization and easy laboratory cultivation. The basic body plan is composed of two concentric tubes. The outer tube consists of cuticle, hypodermis, neurons and muscles, which surround the pseudocoelomic space that contains the intestine and gonad *(see figure 1.4)*.

A

ovary intestine

pharynx oviduct oocytes uterus eggs vulva rectum anus

B
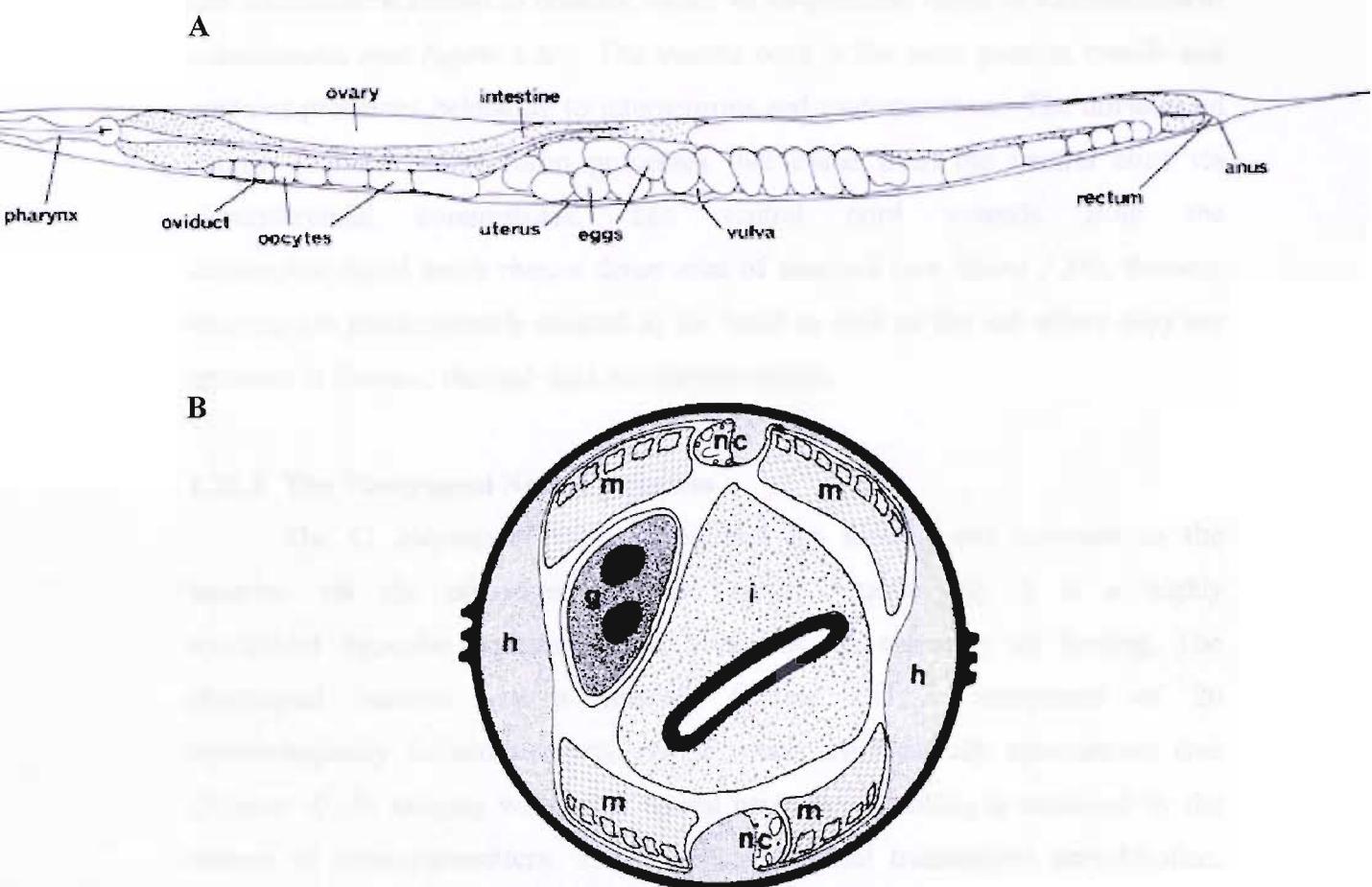
Figure 1.4 The *C. elegans* hermaphrodite. **A.** General anatomy of an adult hermaphrodite. [Taken from Wood, W.B 1988] **B.** Cross-section through the hermaphrodite worm revealing the basic body plan. [m] muscle, [h] hypodermis, [nc] nerve chord, [i] intestine, [g] gonad [Taken from Wood, W.B 1988]

1.26.2 Organization of the *C. elegans* nervous system

The adult hermaphrodite nervous system contains 302 neurons, with a total of 5000 chemical synapses [White, J.G. et al 1986]. The connectivity of the nervous system has been reconstructed and mapped using serial electron microscopy [Albertson, D et al 1976 and White, J et al 1986]. The *C. elegans* nervous system is segregated into the pharyngeal and central nervous system. The two systems function almost independently of each other but are connected by two interneurons [RIP]. In the central nervous system processes of motorneurons

and interneurons extend in bundles, either as longitudinal tracts or circumferential commissures *(see figure 1.5a)*. The ventral cord is the main process bundle and contains processes belonging to interneurons and motorneurons. The dorsal cord mainly contains motorneuron processes that come from the ventral cord via circumferential commissures. The ventral cord extends from the circumpharyngeal nerve ring, a dense area of neuropil *(see figure 1.5b)*. Sensory neurons are predominantly located in the head as well as the tail where they are involved in chemo-, thermal- and mechanosensation.

1.26.3 The Pharyngeal Nervous System

The *C. elegans* pharynx is required for feeding and connects to the intestine via the pharyngeal-intestinal valve. Anatomically it is a highly specialised muscular structure, this optimises the efficiency of feeding. The pharyngeal nervous system controls feeding and is comprised of 20 morphologically defined neurons, six of which are bilaterally symmetrical *(see Chapter 4)*. In keeping with other neural networks signalling is achieved by the release of neurotransmitters. These include classical transmitters acetylcholine, glutamate, octopamine and 5-HT as well as non-classical signalling molecules that include neuropeptides [Rogers, C. et al 2001] and pharyngeal function is modulated by the concerted actions of these transmitters.
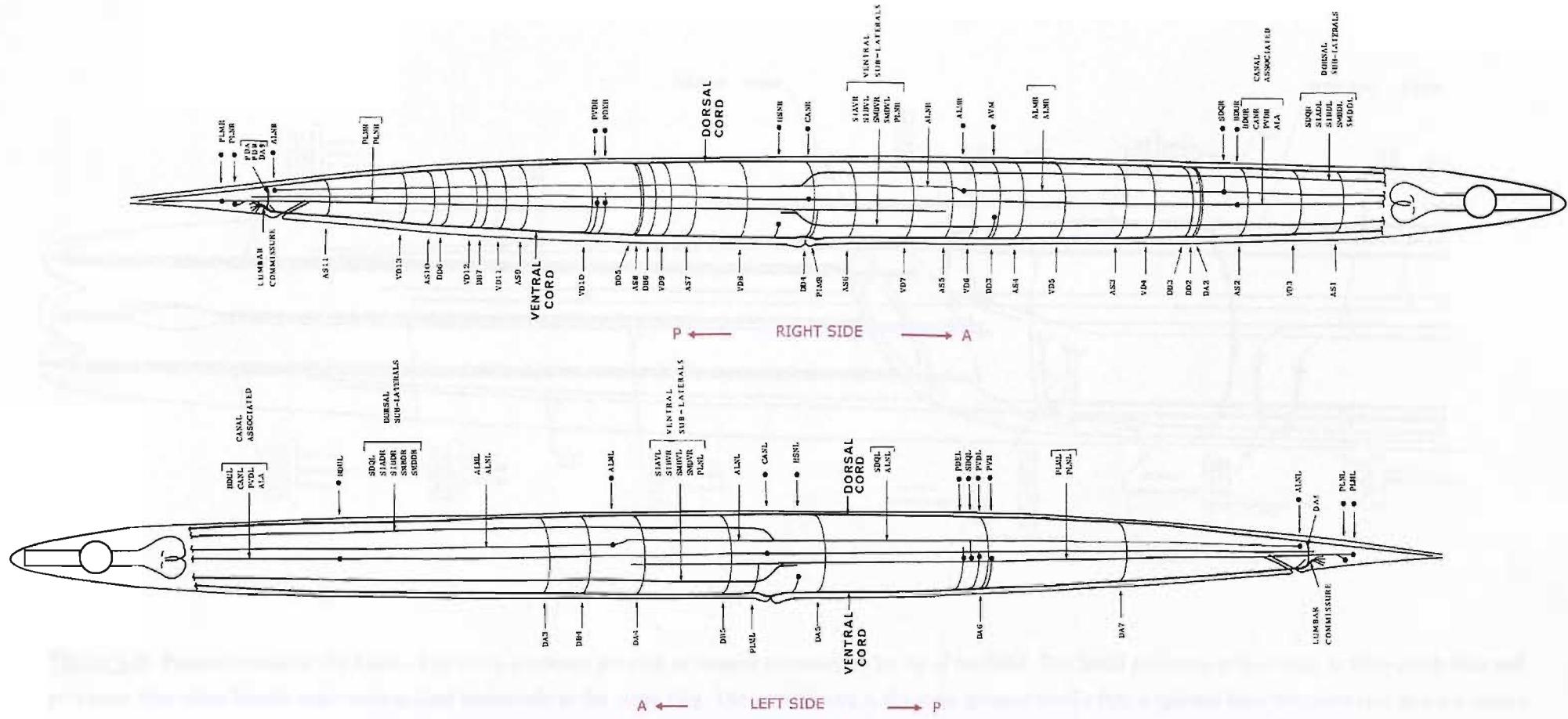
Figure 1.5a Organization of the *C. elegans* somatic nervous system. Process tracts in the left and right hand side of the body are shown [White JG et al 1986]. Neurons are given names consisting of three upper case letters. The last letter can alternatively be a number of up to two digits. Bilateral pairs of neurons are referred to together by a 3-letter name and separately by the same name followed by L or R to indicate the left or right member of the pair and D or V to indicate the dorsal or ventral of the pair. Ventral cord motorneurons have 2-letter names, followed by a number that corresponds to their place in order along the ventral cord.
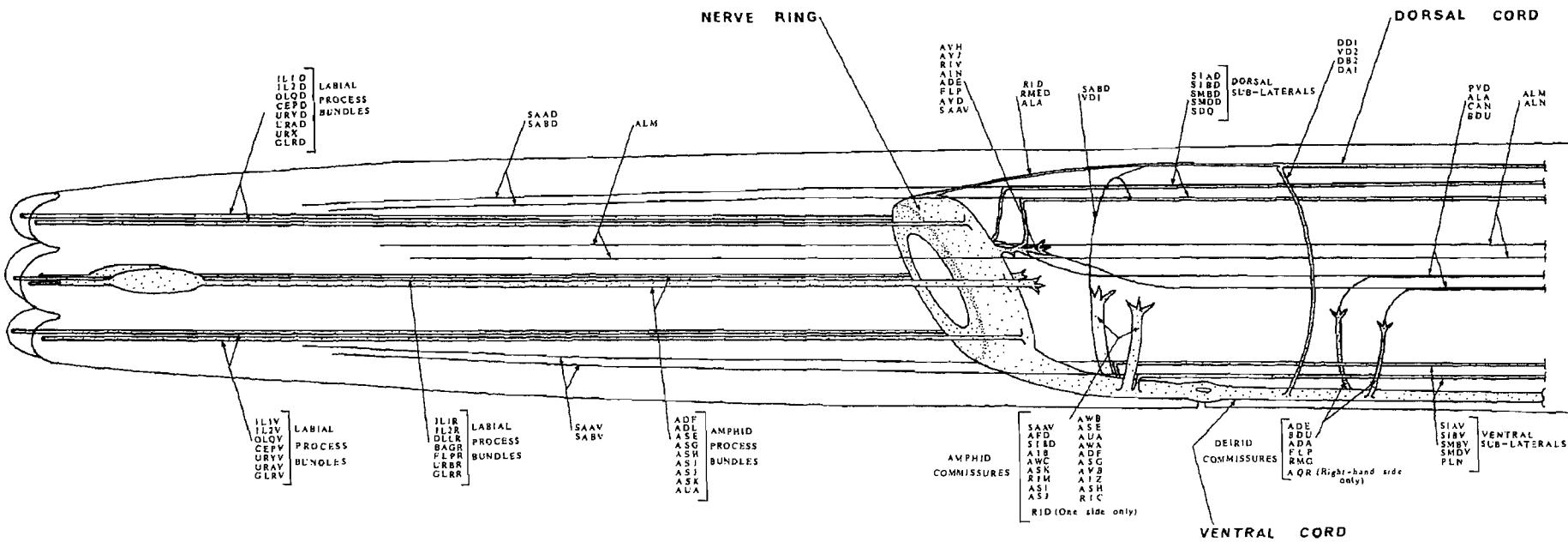
Figure 1.5b Process tracts in the head. The labial processes are rich in sensory receptors at the tip of the head. The labial processes project back to their cell bodies and processes from these bipolar cell bodies project posteriorly to the nerve ring. The ventral cord is the main process bundle that originates from the nerve ring and it contains processes of interneurons and motoneurons. Most of the processes in the dorsal cord originate in the ventral cord and enter the dorsal cord via commissures. There are four sub-lateral process bundles, made up of processes from motoneurons and interneurons that come from the nerve ring. [Taken from White, J.G. et al 1986].

1.27 *C. elegans* exhibits simple behaviours that allow receptor function to be assayed

In *C. elegans* the motor circuits that mediate locomotion [Chalfie, M et al 1985], egg-laying [Waggoner, LE et al 1998], defecation [Avery, L et al 1997], pharyngeal pumping [Avery, L et al 1997] and proprioception [Driscoll, M and Kaplan, J 1997] have been defined. The relative simplicity of these neural circuits and their behavioural outputs makes them tractable for study. Components of each behaviour can be quantified and so provide a basis for the characterization of gene function within neural networks. The components of *C. elegans* locomotion that can be measured include body bend frequency, thrashing, frequency of spontaneous backward movement, radial locomotion and run duration [Tsalik, EL and Hobert, O 2003].

The *C. elegans* sensory nervous system can process basic senses that include olfaction [smell], chemosensation [taste], mechanosensation [touch], osmolarity, ambient temperature, noxious heat and food. The integration of these sensory inputs within motor circuits controlling for example locomotion, enables the worm to modulate how it behaves in response to its external environment. How the worm modulates its behaviour can be measured using more sophisticated assays, such as the chemotaxis maze [Bargmann, CI et al 1993]. Here the worm must modulate several components of its locomotion in order to navigate its way through a maze constructed from an aversive chemical stimuli to reach a chemical attractant.

The behaviours described provide simple models for assaying receptor signalling involved in complex behaviours such as sensory integration, habituation, learning and memory [Hobert, O 2003]. This feature allows *C. elegans* to be used as a tool to further understand receptor signalling in circuitries controlling complex mammalian behaviours. Indeed, *C. elegans* has provided further insight into the function of several receptor classes *(see table 1.6)*.

Receptor Group	Receptor Class	Class Representative	Biological Function
IONOTROPIC	nAChR	*unc-38* [nAChR α-subunit]	Cation selective channel subunit required for body wall muscle contraction during locomotion [Fleming, JT et al 1997]
	GABAR	*unc-49* [GABA$_A$]	Inhibitory anionic channel required for body wall muscle relaxation during locomotion [Bamber, BA et al 1999]
		exp-1	Excitatory cationic channel required for enteric muscle contraction [Beg, AA et al 2003]
	5-HTR	*mod-1*	Inhibitory Chloride channel that modulates locomotion in response to food [Ranganathan, R et al 2000]
	iGluR	*nmr-1* [NMDA]	Cation selective channel subunit involved in the timing of backward and forward movement [Brockie, PJ et al 2001
		glr-1 [AMPA]	Cation selective channel subunit required for mechanosensory responses [Mellem, JE et al 2002]
		avr-15 [GluCl]	Inhibitory Chloride Channel subunit involved in pharyngeal muscle relaxation [Dent, JA et al 1997]
GPCR	mAChR	*gar-3*	GAR-3 couples to PLC when expressed in CHO cells and is involved in pharyngeal muscle relaxation [Park, Y-S et al 2003]
	DopR	*dop-1* [D1-Like]	Modulation of locomotion and mechanosensory behaviours [Sanyal, S et al 2004]
	5-HT	*ser-7b* [5-HT$_7$]	SER-7b couples to the production of cAMP when expressed in COS-7 cells and is predicted to be involved in pharyngeal function [Hobson, RJ et al 2003]
	Peptide Receptors	CO5H2.1 [FSHR]	Positive regulation of growth rate and embryonic development
		npr-1 [NPY-R]	Social aggregation [de Bono, M et al 1998]

Table 1.6 Receptor classes identified from *C. elegans* and their biological functions. The grey boxes highlight receptors belonging to classical sub-classes but with novel signalling characteristics.

1.28   Mediators of Glutamatergic Signalling in *C. elegans*

Glutamate is an important signalling molecule in the mammalian and *C. elegans* nervous system. A number of proteins involved in glutamatergic transmission are conserved between mammals and *C. elegans*, they include members of the iGluR subclass, bio-synthetic enzymes and synaptic vesicular transporter [Bargmann, CI 1998] *(see Figure 1.6)*. In addition to this the *C. elegans* genome encodes novel mediators of glutamatergic signalling, such as the glutamate gated chloride channels [GluCl]. The simple behaviours exhibited by *C. elegans* and its amenability to transgenic approaches has allowed novel insights into the biological function of conserved molecules and has expanded our overall understanding of the molecular, cellular systems and behaviours regulated by glutamate.
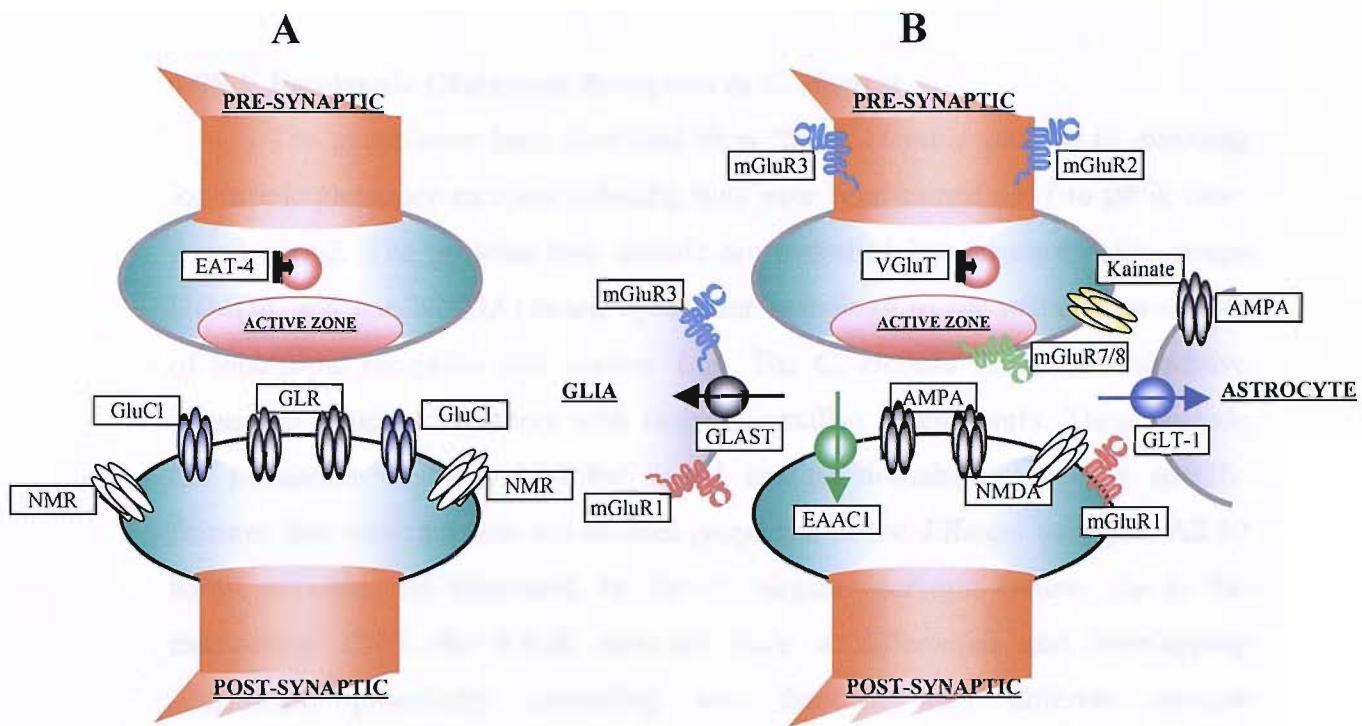
Figure 1.6 An idealized organization of glutamatergic receptors and transporters at the mammalian and *C. elegans* synapse. A. The majority of the key components of glutamatergic transmission in mammals, are also conserved in *C. elegans*. These include classical post-synaptic iGluRs (NMR and GLR) and vesicular glutamate transporter (EAT-4). The *C. elegans* glutamatergic synapse also displays the GluCl channels, a novel class glutamate receptor. **B.** A diverse profile of glutamate receptors and transporters are differentially distributed at the synapse, in both neuronal and non-neuronal cells, in mammals. Group I mGluRs are predominantly peri-synaptic at the post-synaptic membrane (the peri-synaptic compartment of the synapse is shaded in green) (Lujan et al 1996). Group II mGluRs (mGluR2 and 3) are located at extra-synaptic sites (highlighted in orange) (Lujan et al 1997). Group III mGluRs, mGluR7 and 8 are localised immediately adjacent to the active zone (Shigemoto et al 1996).

[Abbreviations: EAAC-1, neuronal glutamate and aspartate transporter in rat brain; GLAST, rat glial glutamate and aspartate transporter; GLT-1, rat glial glutamate transporter; V-GLUT, vesicular glutamate transporter; NMR-1, NMDA receptor homologue; GluCl, Glutamate gated chloride channel; EAT-4, vesicular glutamate transporter; GLT-1, excitatory amino acid transporter homologue.]

47

1.28.1 Ionotropic Glutamate Receptors in *C. elegans*

Ten genes have been identified from the *C. elegans* genome as encoding ionotropic glutamate receptor subunits, they have been named *glr-1* to *glr-8, nmr-1* and *nmr-2*. The proteins they encode are classified into characteristic groups [NMDA and non-NMDA] based upon their homology to the mammalian classes of ionotropic receptors *(see section 1.5)*. The *C. elegans* iGluR subunits have conserved structural features with their mammalian counterparts. These include the membrane topology and the ligand binding domain, as well as specific features that determine the ion channel properties of the different subtypes. All 10 iGluR subunits are expressed by the *C. elegans* nervous system. As in the mammalian CNS the iGluR subunits have a differential and overlapping distribution potentially generating ion channels with different subunit compositions and functional properties [Brockie, PJ et al 2001].

1.28.2 *C. elegans* locomotion and the iGluRs

The ability of an organism to modulate its movement in response to its environment is intrinsic to its survival. Locomotion is a complex behaviour, it encompasses a number of different aspects, such as speed, direction, duration and frequency of altering direction. All of these parameters must be tightly controlled and optimized if the animal is to survive within its environment. For example, when food is sparse an animal increases it duration of movement in one direction to optimise its search strategy. *C. elegans* has provided insight into how neuronal circuits achieve this and glutamatergic transmission is involved.

The *glr-1* gene was the first to be identified as an iGluR subunit in 1995. *glr-1* encodes a subunit with 40% amino acid identity to the rat GluR1 subunit. GLR-1 is the best studied of the *C. elegans glrs* and is expressed in the command interneurons of the locomotory circuit of *C. elegans (see figure 1.7)*. This circuit mediates backward and forward locomotory responses. The central role of GLR-1 in this circuitry was first established by phenotypic analysis of the *glr-1* knockout, using a simple 'nose-touch' assay. The *glr-1* mutant was unable to reverse its direction of movement in response to a mechanical stimuli applied to the nose. Expression of a dominantly active GLR-1 in the command interneurons showed the circuit functions to gate whether the worm moves forward or backwards

movement and that the activity of this circuit is modified by sensory inputs from external cues, such as touch.

The *C. elegans* genes *nmr-1* and *nmr-2* encode iGluR subunits that most resemble the mammalian NMDA subtype. The *nmr-1* subunit is most similar to the NR1 subunit of the mammalian iGluR family, both in terms of sequence homology and biophysical properties. These similar properties allow *C. elegans* NMDA receptors to perform roles in common with the mammalian receptors at the synapse.

Wild-type worms modulate their movement through their environment as part of a foraging strategy. Phenotypic analysis of the *nmr-1* deletion strain *nmr-1(ak4)* first established a role *nmr-1* in the decision to switch from a forward to backward direction of movement. Subsequently it has been shown NMR-1 signalling controls the probability of backward movement by regulating the balance of neural activity between components of the locomotory control circuit that initiate forward and backward movement [Brockie, PJ et al 2001].

In the mammalian CNS the properties of the NMDA receptors contribute to neural mechanisms of synaptic plasticity [Collingridge, GL and Bliss, TVP 1995]. In *C. elegans* the conserved channel properties of the NMR-1 receptor are hypothesised to provide a mechanism for the temporal integration of sensory stimuli. A slowly desensitizing GLR-1 variant, with similar kinetic properties to the NMR-1 dependent current, was expressed in the *nmr-1(ak4)* strain under the control of the *nmr-1* promoter. The behavioural phenotypes caused by the deletion of *nmr-1* were successfully rescued by the GLR-1 variant [Brockie, PJ et al 2001], supporting the view that it is the slow kinetics of the NMR-1 receptor conductance that underlies its function and allows the worm to optimize its movement in complex environments..

1.28.3 The conductance properties of the *C. elegans* AMPA-like receptors are similar to their mammalian counterparts

The subunits GLR-1 and GLR-2 are coexpressed in the command interneuron AVA, where they assemble into a functional heteromeric ion channel. The amino acid sequence of GLR-1 and GLR-2 displays the highest percentage similarity to the AMPA receptor subunit GluR1, in rats. Despite this, the

pharmacological profile of the heteromeric channel is more characteristic of Kainate receptors. The application of Kainate elicits large, rapidly activating and inactivating currents in AVA that is dependent upon the expression of the GLR-1 and GLR-2 subunits. Furthermore, the GLR-1/GLR-2 dependent current displays properties characteristic of Kainate receptors in the mammalian CNS [Mellem, JE et al 2002]. This highlights that conservation at the level of amino acid sequence does not always correlate with conserved functional properties.

1.28.4 Novel mechanisms regulating GLR-1 function

The GLR-1 receptor conductance is the subject of a novel mode of regulation, involving a transmembrane protein encoded by the gene *sol-1* [Zheng, Y et al 2004]. Initially identified as a suppressor of a constitutively active variant of GLR-1, the SOL-1 protein co-localizes with GLR-1 at postsynaptic sites but is not required for receptor targeting. Co-immunoprecipitation and electrophysiological analysis suggests SOL-1 directly interacts with GLR-1 at the plasma membrane and is required for GLR-1 channel conductance. The SOL-1 protein is almost entirely extracellular and contains protein-protein interaction domains that may recruit proteins important for GLR-1 function.
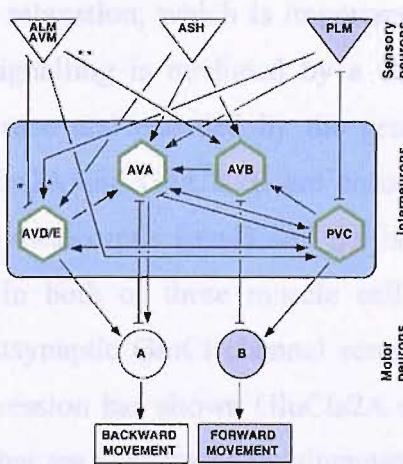
Figure 1.7 **Locomotory command circuit in *C. elegans*.** Reciprocal inhibitory connections between the forward (AVB, PVC) and backward (AVA, AVD/E) command interneurons controls the initiation of either forwards or backwards movement. Sensory information is integrated into the command circuitry to activate distinct classes of motorneuron [A class and B class]. Synaptic connections are indicated by arrows and gap junctions as [⊥]. Forward neurons are grey and backward neurons are white. Boxes outlined in green express iGluR subunits. [Figure taken from Zheng, Y et al 1999].

1.28.5 GluCl Channels

The glutamate gated chloride channel [GluCl] family is a novel class of ionotropic receptor, expressed only in invertebrate muscle and nerve cells. To date eight GluCl genes have been identified. The GluCls are closely related to the GABA$_A$ and glycine gated chloride channels in vertebrates. The features they have in common are a large extracellular N-terminal domain, four transmembrane domains and putative N-linked glycosylation sites in the extracellular domain. The first GluCl subunits to be cloned were GluClα1 and GluClβ [Cully, D.F et al 1994]. Coexpression of the two subunits in *Xenopus oocytes* identified they can assemble to create a heteromeric ion channel. The GluClα1/GluClβ ion channel was selectively permeable to chloride and reversibly activated by glutamate. It was also sensitive to the anthelmintic ivermectin, the application of IVMPO4 (Ivermectin-4"-O-phosphate, monosodium salt) activates rapid and irreversible chloride currents.

GluCl subunits are required for the normal electrical activity of the pharynx. Specific subtypes have been identified as having an important role in mediating glutamate signalling at the neuromuscular junction in the *C. elegans* pharynx. The M3 motorneuron controls the timing of pharyngeal muscle hyperpolarization and relaxation, which is important for efficient feeding [Heyi, Li et al 1997]. M3 signalling is mediated by a GluCl channel located on the muscle plasma membrane and encoded by the gene *avr-15*. Two alternatively spliced subunits GluClα2A and GluClα2B are encoded by *avr-15*. The synaptic targets of M3 are the metacorpus [pm4] and the isthmus [pm5], the GluClα2A subunit is expressed in both of these muscle cell types. It is likely to be a component of the postsynaptic GluCl channel responsible for M3 transmission. Recombinant cell expression has shown GluClα2A can assemble into functional homomeric channels that are activatable by glutamate and sensitive to ivermectin [Dent, JA et al 1997].

1.28.6 *eat-4* encodes a glutamate vesicular transporter

The polypeptide encoded by *eat-4* is 48% identical to a mammalian brain specific sodium-dependent inorganic phosphate cotransporter [BNPI] [Avery, L et al 1993]. The protein BNPI transports glutamate into synaptic vesicles. Based

upon this homology and phenotypic analysis of *eat-4* mutants, it is likely to perform the same function in *C. elegans* neurons. The expression of EAT-4 is widespread in the *C. elegans* nervous system. It is expressed in pharyngeal and extrapharyngeal neurons. An *eat-4*::reporter gene construct identified a total of 34 different cell types express EAT-4 [Lee, RY et al 1999]. In the pharynx the EAT-4 protein is expressed in the M3 motorneuron, neurosecretory motorneuron [NSM] and the interneuron I5. The remaining neurons include both glutamatergic neurons and neurons with unknown neurotransmitter content. The co-expression of *eat-4* in the 5-HT containing neuron NSM may represent a novel interplay between the two neurotransmitter molecules, glutamate and 5-HT, that underlies the control of NSM signalling *(see Results Chapter 4)*.

Phenotypic analysis of *eat-4* mutants show that it is required for glutamate release from presynaptic terminals and the glutamatergic control of pharyngeal pumping by M3 requires *eat-4*. [Raizen, DM and Avery, L 1994]. The mechanosensory neurons PLM, AVM and ALM express *eat-4*. A presynaptic role for *eat-4* in chemical transmission between these sensory neurons and their interneuron targets has been described [Lee, R.Y.N et al 1999].

1.28.7 The Metabotropic Glutamate Receptors

The *C. elegans* genome predicts three genes encoding metabotropic glutamate receptors [*mgls*]. Existing analysis [Ishihara et al 1996] has defined the genes as *mgl-1* [ZC506.4], *mgl-2* [F45H11.2] and Y4C6A.2a [in the absence of a *C. elegans* Genetics Centre [CGC] approved name this will be referred to as *mgl-3*]. Existing analysis of the *mgl-1* and *mgl-2* amino acid sequences has predicted they couple to classical intracellular signalling molecules, $G_{\alpha i}$ and $G_{\alpha q}$ respectively [Ishihara, T et al 1996]. Furthermore, HEK293 cells expressing *mgl-2* cDNA increase phosphoinositide turnover in response to glutamate [Ishihara, T et al 1997]. However little is known about the signalling role *mgls* perform in *C. elegans*.

1.29 Cellular assays of *C. elegans* receptor function

C. elegans receptors have been successfully expressed in mammalian cell lines and by doing so a number of them have been functionally characterized

52

[Mertens I et al 2006]. Cellular assays of receptor signalling have made it possible to profile *C. elegans* GPCRs pharmacologically and to determine the molecular basis of their function. This has been particularly beneficial in the ongoing process of de-orphanizing *C. elegans* receptors and it has been applied successfully to members of the *C. elegans* family of neuropeptide GPCRs [Mertens, I. et al 2005].

1.30   Pharmacology of function

The genetics of *C. elegans* has been used to identify and characterize the molecular targets of therapeutic molecules [Kaleta, T. and Hengartner, M.O 2006]. 'Chemical genetic screens' have been utilised to identify *C. elegans* genes that either confer resistance or hypersensitivity to a compound of interest and these genes can either be the primary drug target or a component of the same pathway as the primary drug target. In using this approach the *in vivo* mode of action of a drug can be explored and all targets potentially identified. For example, it is known that Fluoxetine (Prozac™) affects 5-HT levels, but other than this its therapeutic mode of action is ill-defined. To address this a genetic screen performed on a transgenic strain of *C. elegans* deficient in 5-HT identified a second drug target, involving a G-protein signalling pathway [Dempsey, C.M. et al 2005]. Such genetic screens have been used to establish the molecular networks underlying the mode of action of a number of other drugs and these are summarised in *table 1.7*

The reverse genetic approach is where a gene of interest has been identified from the genome and this guides the pharmacological agents that are used to perform the functional characterization of the gene. For example 42 genes have been identified from the *C. elegans* genome that could encode nicotinic acetylcholine receptors, however, the function of most of these remains to be fully defined. The broad-spectrum mammalian nAChR agonist, 1,1-dimethyl-4-phenylpiperazinium (DMPP) has been shown to cause developmental defects in *C. elegans*. Knockdown of these drug effects by RNAi of nAChR genes has identified those nAChR genes that signal in the timing of *C. elegans* development. Subsequently chemical genetic screens using DMPP has identified genes required for nAChR signalling in *C. elegans* development.

COMPOUND	FUNCTION	REFERENCE
Levamisole	Anthelmintic	Brenner, S. 1974
Aldicarb	Anthelmintic	Nguyen, M et al 1995
Benzimidazole	Anthelmintic	Driscoll, M et al 1989
Ivermectin	Anthelmintic	Vassilatis, D.K. et al 1997, Dent, J.A. et al 1997 & 2000
Nicotine	Drug of addiction	Waggoner, L.E. et al 2000
Alcohol	Drug of addiction	Davies, A.G. et al 2004 and 2003
Volatile general anaesthetics	Human drug	Hawasli, A.H. et al 2004
Farnesyl transferase inhibitors	Anticancer drug	Lackner, M.R. et al 2005
Fluoxetine (Prozac)	Anti-depressant	Ranganathan, R. et al 2001 & Dempsey, C.M. et al 2005
Imipramine	Anti-depressant	Dempsey, C.M. et al 2005

Table 1.7 Drugs at the centre of 'chemical genetic screens.' *C. elegans* has been used to characterise the molecular targets of drugs with varied uses; ranging from anthelmentics to human anti-depressants.

1.31   Tool kit for *C. elegans* research

A major asset of *C. elegans* as a model organism is the availability of a sequenced genome, which can be screened for homologues of human genes of interest. Once a candidate gene is identified the genomic sequence can be used to design molecular reagents for the analysis of cDNA and characterization of the encoded transcripts. Analysis of the translated cDNA sequence, by comparisons to mammalian homologues, allows conserved structures, motifs and residues underlying functionality to be identified [Harteneck, C .et al 2000].

Once the protein sequence and the gene structure is known transgenic reporter constructs can be generated to characterize the spatial and temporal expression of both the gene and the protein at the level of the whole animal [Dupuy, E. et al 2005]. Identification of the neurons expressing the gene and phenotypic analysis of transgenic mutants allows the physiological role of the gene within the animal to be assessed. Assays used to perform the phenotypic analysis can be behavioural , pharmacological or electrophysiological and the type

of assay used can be guided by the processes involved with the neural circuits containing the gene [Rankin, C.H. et al 2002 and Suo, S. et al 2004].

1.32 Scaffolding mechanisms in *C. elegans*

As outlined in *section 1.16* receptor scaffolding is functionally relevant. In a similar way the results discussed for *sol-1* reflect similar processes in the organization of *C. elegans* receptor function. Indeed, scaffolding mechanisms underlie a number of different biological processes in *C. elegans*, including embryonic development [Caruana G 2002], vulval differentiation [Ferguson, EL et al 1985], defecation rhythm [Jee, C et al 2004], cell junction assembly and synaptogenesis [Yeh, E 2005] . This highlights scaffolding is involved in both essential and non-essential processes in *C. elegans*. As previously described, the PDZ domain plays a major role in organizing cell signalling complexes and the sequenced genome of *C. elegans* predicts 138 PDZ domains in 96 proteins [Hung AY et al 2002]. The PDZ domain has been identified as an important module in embryonic development [Hung, T.J. et al 1999], vulval differentiation [Kaech, S.M. et al 1998] and possibly the specialization of axon identity [Hallam, S.J. et al 2002].

1.32.1 Polarization in the *C. elegans* nervous system

The principles of a pre- and postsynaptic domains as discussed in the mammalian nervous system *(see sections 1.2 and 1.3)* are conserved in *C. elegans* neurons, but there are some subtle differences. The invertebrate nervous system is predominantly composed of unipolar cells. This class of neuron is characterised by a single process or neurite extending from the soma. Both presynaptic and postsynaptic membrane compartments can exist along the length of a single unbranching process. In this case the process is not strictly a dendrite or an axon *(see figure 1.8)*.

The D-type ventral cord motorneurons of *C. elegans*, DD and VD, control the inhibition of dorsal and ventral body wall muscle contraction during locomotion. They are pseudo-unipolar cells. This type of neuron has a single process emerging from the soma that splits into two processes. In this example one process functions as the axonic compartment exclusively sending synaptic signals and the other functions as the dendritic compartment and receives synaptic

signals [White, JG et al 1986]. This is similar to the sensory neurons of the dorsal root ganlion in mammals, these cells are also classified as pseudo-unipolar and have a single axon that extends from the cell body and splits into two branches that are functionally distinct. The *C. elegans* sensory amphid neurons that detect external stimuli are similar to the mammalian olfactory and retinal bipolar cells, where a single dendrite and axonal process extend from the cell body *(see figure 1.8)*. The *C. elegans* amphid neurons have a ciliated dendrite that extends to the nose and is exposed to the external stimuli [White, J.G. et al 1986]. The cilia represent a specialised compartment and proteins are targeted to the cilia that are important for detecting external stimuli [Scholey, J.M. et al 2003].

A

Pre-synaptic

Post-synaptic

NEURITE

CELL BODY

B

Dorsal Muscle
CONTRACTED

DORSAL
CORD

VENTRAL
CORD

VD

Dorsal ACh
Motorneuron

Ventral Muscle
RELAXED

C

Cilia

Dendrite

Cell-body

Axon
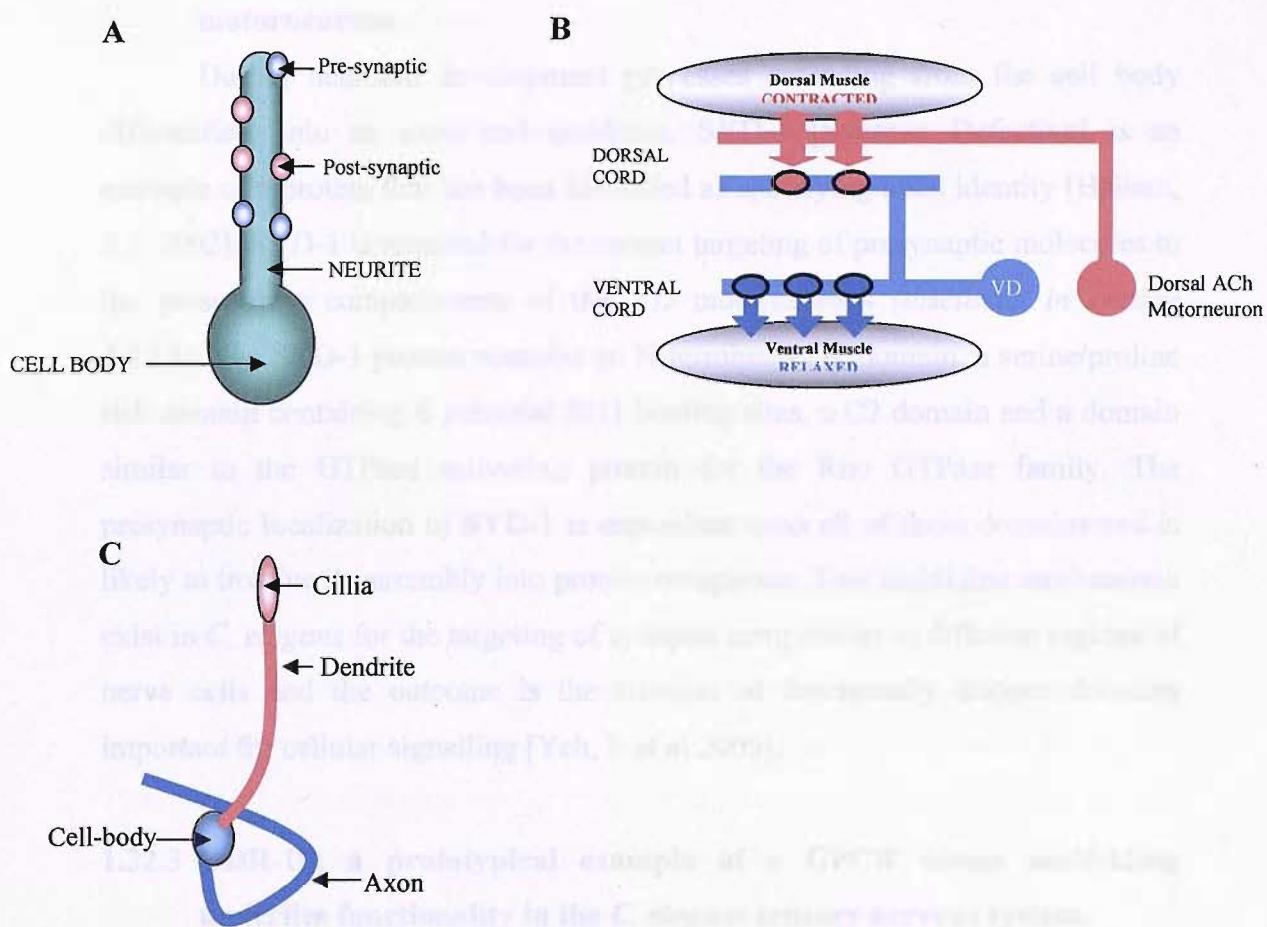
Figure 1.8 *C. elegans* **neurons are polarized cells. A. Unipolar cell.** Pre- and postsynaptic compartments exist within a single process (for example nerve ring interneurons, such as AIA are unipolar). **B. Pseudo-unipolar cell.** The GABAergic VD motorneuron (blue) of the adult worm is shown as an example of a pseudo-unipolar cell. The VD motorneuron has an anatomically defined presynaptic (blue ovals) and postsynaptic (red ovals) compartments. The VD motorneuron is involved in the co-ordination of movement. A cholinergic motorneuron is shown sending inputs to the dorsal body muscle and also the GABAergic motorneuron. Release of ACh (red) causes dorsal muscle to contract and stimulates VD to release GABA onto the muscles opposite (ventral), which stimulates relaxation and together this causes the body to bend. **C)** A *C. elegans* amphid sensory neuron (for example ASH).

1.32.2 Targeting of presynaptic components to the axon of ventral cord motorneurons

During neuronal development processes extending from the cell body differentiate into an axon and dendrites. SYD-1 [**S**ynapse **D**efective] is an example of a protein that has been identified as specifying axon identity [Hallam, S.J. 2002]. SYD-1 is required for the correct targeting of presynaptic molecules to the presynaptic compartments of the VD motorneurons *(discussed in section 1.32.1)*. The SYD-1 protein contains an N-terminal PDZ domain, a serine/proline rich domain containing 6 potential SH3 binding sites, a C2 domain and a domain similar to the GTPase activating protein for the Rho GTPase family. The presynaptic localization of SYD-1 is dependent upon all of these domains and is likely to involve its assembly into protein complexes. This highlights mechanisms exist in *C. elegans* for the targeting of synaptic components to different regions of nerve cells and the outcome is the creation of functionally distinct domains important for cellular signalling [Yeh, E et al 2005].

1.32.3 ODR-10: a prototypical example of a GPCR whose scaffolding underlies functionality in the *C. elegans* sensory nervous system.

The GPCR ODR-10 is an olfactory receptor targeted to the cilia of the olfactory neuron AWA. Cilia are specialised microdomains of the *C. elegans* amphid sensory neurons and are abundant in both intracellular and membrane bound signalling molecules. ODR-10 is specifically targeted to the cilia of AWA [Dwyer, N.D et al 1998] by *odr-4* and *odr-8*. The targeting of other signalling molecules to the cilia, such as ODR-3 [Gα protein subunit], TAX-2 [cyclic nucleotide-gated channel subunit] and OSM-9 [cation channel] have been shown to be independent of *odr-4* and *odr-8* [Dwyer, N.D et al 1998].

ODR-4 is predicted to be almost entirely cytoplasmic with a single C-terminal membrane anchoring region. Models proposed for how ODR-4 targets ODR-10 include; a role as a chaperone involved in correctly folding ODR-10 into the membrane of the ER; a cargo receptor that sorts ODR-10 into secretory vesicles addressed to the cilia; an accessory protein for the trafficking of vesicles to the cilia through interactions with neuronal motor proteins [McClintock, T.S. et al 2003]. This illustrates the sophistication of the *C. elegans* nervous system and

demonstrates scaffolding is employed by complex physiological processes in the worm, such as chemosensation.

1.32.4 Scaffolding molecules can be shared between different cell types and receptor types in *C. elegans*

The LET-23/Egfr targeting complex [LIN-10, LIN-7 and LIN-2] that has been implicated in basolateral sorting in vulval cells has also been shown to recognise GLR-1 and target it to the basolateral membrane when it is heterologously expressed in vulval precursor cells. As like LET-23 the GLR-1 receptor contains a C-terminal Type 1 PDZ domain allowing it to utilise the same trafficking pathways as LET-23 [Rongo, C et al 1998]. This highlights that the same molecular signals are recognised by the targeting machinery employed by epithelial and neuronal cells. Infact proteins that are targeted to the apical membrane of epithelial cells are often targeted to the axon when expressed in neurons. Whereas basolateral proteins go to somato-dendritic domains [Bredt, D.S 1998]. The LIN-10 protein of the LET-23 targeting complex is common to both neurons and epithelial cells. In interneurons it co-localises with GLR-1 at the synapse and it is required for the targeting of GLR-1 to the synapse [Rongo, C et al 1998]. This demonstrates that scaffolding proteins can be employed by more than one signalling pathway in more than one cell type *(see figure 1.9)*.

1.32.5 The regulation of receptor abundance at the synapse is a conserved feature of the mammalian and *C. elegans* nervous system

The stabilization and retention of receptors within sub-domains is another means of modulating receptor localization and synaptic transmission. In the mammalian CNS the make-up of excitatory synapses is continuously being modified by the rapid recycling of glutamate receptors between the membrane and intracellular compartments [Palmer, C.L. et al 2005; Ashby, M.C. et al 2004; Martin, S. et al 2004]. Such mechanisms are conserved in the *C. elegans* nervous system [Grunwald, ME et al 2004]. The abundance of synaptically localized GLR-1 is reported as being regulated by ubiquitin mediated endocytosis into clathrin coated vesicles. Assessment of *C. elegans* locomotory behaviours supports this process as a mechanism for changing the strength of glutamatergic transmission. Mutations that prevent the internalisation of GLR-1 recapitulate the

same behavioural phenotypes exhibited by transgenic lines expressing a constitutively open GLR-1 receptor [Burbea, M et al 2002].
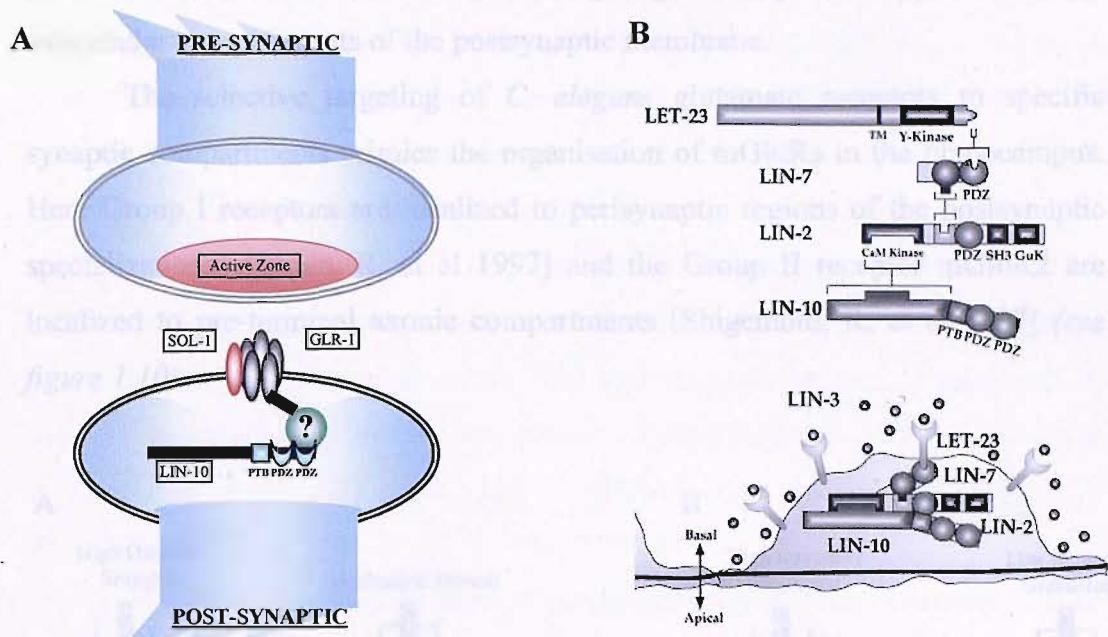
Figure 1.9 A model for LET-23 targeting to the basolateral membrane of vulval epithelial cells and the GLR-1 targeting to the postsynaptic domain of interneurons. **A** Although LIN-10 is required for GLR-1 targeting in interneurons, a direct interaction between GLR-1 and LIN-10 has not be established, other neuronal specific elements (represented by the grey circle and question mark) may be involved in this process. **B,** The targeting apparatus of LET-23 assembles in a co-ordinated manner through specific protein-protein interaction domains [Kaech, SM et al 1998].

1.32.6 Polymodal signalling - A paradigm for glutamate receptor scaffolding

Scaffolding mechanisms are important for physiological processes in *C. elegans,* which is well demonstrated by the targeting complex of LET-23 in vulval development [Kaech, S.M. et al 1998]. In *C. elegans* mechanisms of polymodal signalling by the ASH sensory neuron has now established a role for the subcellular targeting and anchoring of iGluRs in decoding distinct sensory stimuli. This shows clear parallels with mammalian glutamate receptors *(as described in section 1.21 and figure 1.6)* Mechanical and osmotic stimuli detected by the

sensory neuron ASH lead to the differential activation of postsynaptic iGluR subunits [Mellem, J.E. et al 2002]. The activation of distinct receptor combinations is achieved by the sensory stimuli-specific modulation of glutamate release from the ASH neuron and the targeting of receptor subtypes to distinct subcellular compartments of the postsynaptic membrane.

The selective targeting of *C. elegans* glutamate receptors to specific synaptic compartments mimics the organisation of mGluRs in the hippocampus. Here Group I receptors are localized to perisynaptic regions of the postsynaptic specializations [Lujan, R. et al 1997] and the Group II receptor mGluR2 are localized to pre-terminal axonic compartments [Shigemoto, R, et al 1997] *(see figure 1.10)*.

A

High Osmotic Strength   Mechanical Stimuli   High frequency simulation   Low frequency stimulation

B

ASH   ASH   PRE

POST

⋔ NMDA   ⋔ GluCl   ⋔ NON-NMDA   🌀 mGluR2   🌀 mGluR1
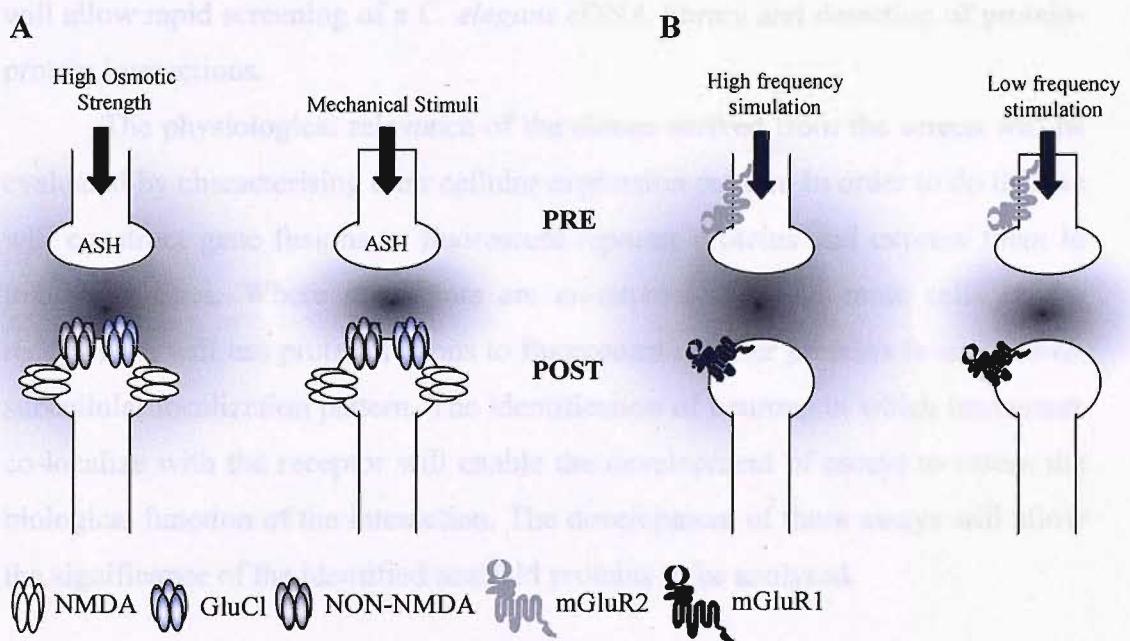
Figure 1.10 Subcellular organization of glutamatergic receptors at the synapse, in *C. elegans* sensory pathways and the mammalian hippocampus. **A.** *C. elegans* ionotropic receptors are differentially localized to distinguish between sensory stimuli. High osmotic pressures stimulate enhanced glutamate release from the sensory neuron ASH, compared to mechanical stimuli. Increased amounts of glutamate can diffuse further causing a different profile of receptor activation, which encodes the type sensory stimuli encountered. **B. mGluR subtypes are targeted to distinct post- and pre-synaptic compartments of hippocampal neurons.** As in *C. elegans,* the amount of neurotransmitter release from the presynaptic terminal can co-ordinate the signalling profile, through activation of different mGluR subtypes. For instance, high frequency stimulation [100Hz] causes the release of sufficient glutamate to activate mGluR2 receptor subtypes, which are localized away from the synaptic cleft.

61

1. 33   Aims and objectives

At the outset very little was known about the molecular, cellular, systems or behavioural contribution made by the *C. elegans mgls* to glutamatergic transmission. Accordingly we adopted a functional genomic approach to the study of mGluR scaffolding in the model organism *C. elegans*. The plan of research is outlined in *figure 1.11* and commences with the identification of three mGluR genes [*mgls*] from the sequenced *C. elegans* genome. *In silico* techniques will be used for the close examination of these predicted protein to validate them as mGluRs and identify functional domains and motifs. Bearing in mind the capacity of the intracellular C-terminal to scaffold mGluR function, this domain will be used for the isolation of *mgl* binding proteins. Using the yeast-2-hybrid technique will allow rapid screening of a *C. elegans* cDNA library and detection of protein-protein interactions.

The physiological relevance of the clones derived from the screen will be evaluated by characterising their cellular expression pattern. In order to do this we will construct gene fusions to fluorescent reporter proteins and express them in transgenic lines. Where interactors are co-expressed in the same cells as the receptor we will use protein fusions to fluorescent reporter proteins to assess their subcellular localization pattern. The identification of neurons in which interactors co-localize with the receptor will enable the development of assays to assess the biological function of the interaction. The development of these assays will allow the significance of the identified scaffold proteins to be analysed.
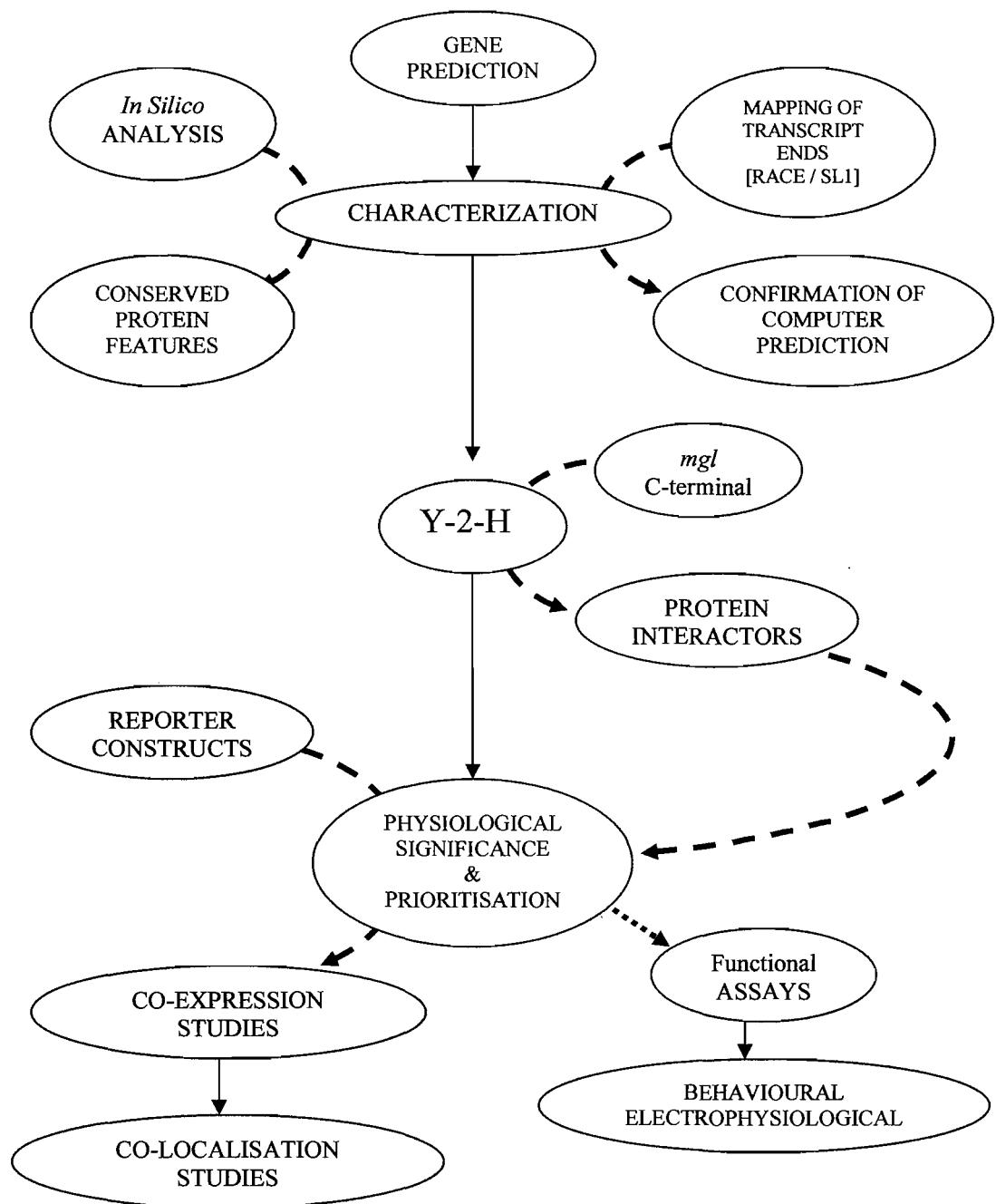
Figure 1.11 A flowchart summarising the approach we have adopted to study *mgl* scaffolding in *C. elegans.* RACE: Rapid Amplification of cDNA Ends. SL: Splice leader. Y-2-H: Yeast-2-Hybrid.

CHAPTER 2

Methods and Materials

2.1 Molecular Techniques for Bacteria

2.1.1 Bacterial Strains

XL-1 Blue E.coli
Electrically
competent cell

Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173
endA1 supE44 thi-1 recA1 gyrA96 relA1
lac [F' proAB lacIqZΔM15Tn10 (Tetr)].

Stratagene   Routine cloning

TOP10 E.coli
Chemically
competent cells

F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lac
ZΔM15 ΔlacX74 recA1 araD139 Δ(araleu)
7697 galU galK rpsL (StrR) endA1 nupG

Invitrogen   Topo & TA Cloning

OP50 E.coli   Uracil auxotroph   Nematode cultivation

2.1.2 Bacterial Cultures

Bacterial cultures were grown under sterile conditions using LB (Luria Bertani) media (Tryptone 10g/L, Yeast Extract 5g/L, NaCl 10g/L, pH 7) at 37°C with shaking. Bacterial colonies were selected for on LB agar plates (Tryptone 10g/L, Yeast Extract 5g/L, NaCl 10g/L, Agar 20g/L, pH 7). Media was sterilized by autoclaving (LTE 300SH 121°C 15min, 1 BAR), allowed to cool to 50°C and supplemented with the appropriate selective antibiotic (see Table 2.1).

Antibiotic	Stock Solution – Concentration	Final Concentration
Ampicilin	100mg/ml in water	50μg/ml
Tetracycline	5mg/ml in ethanol	50μg/ml
Kanamycin	50mg/ml in water	50μg/ml

Table 2.1. Antibiotic stock solutions and working concentrations. Stock solutions of antibiotic were stored in frozen aliquots at −20°C

2.1.3 Generation of Electrocompetent Cells

Single colonies of XL-1 Blue were generated by streaking from a single glycerol stock onto tetracycline selective LB agar. From this a single colony was picked and grown at 37°C overnight in 50ml of YENB (Yeast Extract, Bacto Nutrient Broth) media (7.5g/L Yeast Extract, 8g/L Bacto Nutrient Broth, pH 7). The following day the starter culture was diluted in 1L of YENB to an optical density (OD_{600}) of 0.1 and grown to an OD_{600} of 0.6. All vessels at this stage had

been incubated at 4°C and all steps from this point onwards were performed at 4°C. The 1L culture was divided in half and spun at 4 000g (Sorvall Legend RT) for 10 minutes and the supernatant removed. The remaining pellet was resuspended in 20ml of sterile 10% (v/v) glycerol and spun at 4 000g for 10 minutes (Falcon Tubes), the supernatant was removed and the pellet resuspended in a final volume of 3mls of 10% Glycerol. This was divided into 80μl aliquots and snap frozen in liquid nitrogen. Cells were stored at −80°C until use.

2.1.4 Electroporation

The XL-1 Blue strain of electrocompetent cells was used for standard transformations with 1ng of double stranded plasmid DNA. A single aliquot of XL-1 blue was thawed on ice and the DNA added. Gentle mixing was then followed by transfer to a prechilled 2mm gap electroporation cuvette (Qbiogene). The cuvette was placed in the electroporation chamber and 2.4KV passed through. The cells were recovered with 1ml of LB and incubated at 37°C with shaking for 1 hr. Selective LB plates were prewarmed to room temperature and dried under sterile conditions. The culture was then spread onto selective plates and incubated for 16 hrs at 37°C. Colonies were then identified for further analysis.

2.1.5 Extraction of plasmid DNA from bacteria

Bacterial colonies were initially screened for plasmid DNA by using a crude mini-prep protocol based on an alkaline lysis method (Bimboim, HC and Doly, J. 1979). Bacterial colonies were picked and grown overnight in 3mls of selective LB media. The next day 1.5mls of the culture was pelleted and resuspended in 100μl of Buffer P1 (50mM TrisCl, pH 8.0, 10mM EDTA 100μg/ml RNase A). To this 200μl of Buffer P2 (200mM NaOH, 1%SDS w/v) was added. The sample was mixed by gentle inversion and the lysate neutralized by the addition of 150μl Buffer P3 (3.0M potassium acetate, pH5.5). The mix was again gently inverted to avoid shearing genomic DNA. The lysate was cleared by spinning at 14,000rpm for 5 minutes. The cleared lysate (~450μl) was removed into 900μl 100% EtOH, desalting and precipitating the plasmid DNA. The solution was then vortexed and spun, the supernatant removed and the pellet overlaid with 250μl of 70% Ethanol to remove residual salt and spun again. The

supernatant was removed and the pellet allowed to air dry for 30 minutes before being resuspended in 20µl of H_2O. High quality plasmid DNA was generated using the QIAGEN Plasmid Purification Kits according to the manufacturers (Qiagen) microcentrifuge protocol. Plasmid DNA prepared in this way was used for sequencing, cloning and eventually micro-injection.

2.1.6 Details of the *C. elegans* LexA Library

The *C. elegans* Lex-A cDNA library was obtained from OriGene Technologies Inc. The library has a complexity of 3.8 x10^6 independent clones, with inserts ranging in size from 0.2-3.6kb. The library was constructed from cDNA made with a oligo d(T) primed reverse transcription of mixed stage *C. elegans* mRNA.

2.1.7 Titration

The 1ml of library was thawed on ice and aliquoted into 10X 100µl volumes, 9 tubes were frozen at −80°C and 1 tube was left on ice and stored at 4°C. From this serial dilutions were made (1:10^5, 1:10^6, 2:10^6, 1:10^7). Serial dilutions of the library were plated onto LB plates containing ampicilin (LB$_{amp}$) and incubated at 37°C overnight. The original tube containing 100µl of the library was stored at 4°C overnight. The next day the number of colony forming units (cfu) were counted on each plate and used to calculate the number of cfu/ml of library. This was calculated as 9.55x10^9 cfu/ml. The volume of library required for an amplification 3x the complexity of library (1.1x10^7 cfu) was calculated as 1.2µl.

2.1.8 Plating

1.2µl of the library was diluted into 50µl of cold LB$_{amp}$ and 10µl of this was diluted into 13ml of cold LB$_{amp}$. The 13ml volume was plated out onto LB$_{amp}$ plates. 100x, 15 diameter plates (Greiner Bio-1) were plated at a density of 85,000 cfu/plate and 6x, 20x20cm plates were plated at a density of 450,000 cfu/plate. The remaining 40µl volume of the diluted library was diluted and plated in the same way. When performing the amplification the library was titred by making serial dilutions in LB$_{amp}$ and calculated as 5x10^9cfu/ml, giving an amplification of

5×10^6 cfu. A second amplification was performed to combine with the first amplification and obtain a total amplification 3x the complexity of the library. 2μl of the library (1×10^7 cfu) was diluted into 50μl cold LB$_{amp}$, 10μl of this was then diluted into 13ml and plated as before. The library was titred during the amplification as 6×10^9 cfu/ml and from this the amplification was calculated as 12×10^6 cfu. The final amplification of the library was calculated as the sum of the first amplification and second amplification, which was 17×10^6 cfu.

2.1.9 Extraction of the LexA library plasmid DNA

Amplification plates were incubated at 4°C for 3hrs to solidify and bacteria scrapped from the plates into a total volume of 4x 1L of LB$_{amp}$. Cultures were incubated for 3hrs at 37°C, then spun at 6 000g for 15 minutes and the supernatant removed. The plasmid DNA was extracted using the QIAGEN plasmid purification reagents. Twelve 50ml centrifuge tubes were sterilized by washing with detergent, rinsing 4x with distilled water and then drying. Once dry, tubes were exposed to UV by using a UV Stratalinker 1800 (Stratagene). The bacteria yielded from the two amplifications were pooled by resuspending the pellets in 160ml of buffer P1 (Qiagen) and 13.4ml was aliquoted out into 12x 50ml centrifuge tubes. To each centrifuge tube 13.4ml of buffer P2 (Qiagen) was added and mixed by inversion, then left at room temperature for 5 minutes. Then 13.4ml of chilled buffer P3 (Qiagen) was added to each tube, mixed immediately and left at 4°C for 30 minutes. The centrifuge tubes were spun at 4°C and 20 000g (Sorvall RC-5B Refrigerated Superspeed Centrifuge, Rotor: SLA-600TC), the supernatant from each tube was removed and pooled. The supernatant was aliquoted into 24mls in 50ml falcon tubes and 17mls of isopropanol added. The tubes were inverted to mix and spun at 4600rpm for 1hr 30 minutes at 4°C. The supernatant was removed from each tube and the pellet washed with 70% cold ethanol and spun again at 4600rpm for 30 minutes at 4°C. The pellets from each of the 20 falcon tubes were resuspended in a total volume of 25ml sterilised ddH$_2$O. The library was quantified by spectrophotometry, the concentration of DNA was 2.5mg/ml. The DNA was aliquoted into 1ml volumes and stored at -80°C.

2.1.10 Glycerol Storage of Bacterial Cultures

Bacteria were grown in 3ml of selective liquid media, 0.85 ml of the bacterial culture was added to 0.15 ml of 100% glycerol (sterilized by autoclaving). The cryogenic tube (sterilized by autoclaving) was then vortexed to mix the glycerol and frozen by placing immediately into liquid nitrogen. Tubes were transferred to the −80°C freezer for long term storage. The bacteria were recovered by taking a scrape from the frozen surface of the culture, using a sterilized loop and streaking onto selective LB plates. Plates were incubated overnight at 37°C and the glycerol stock returned to −80°C.

2.1.11 Glycerol storage of Cosmid strains

Cosmids were freely available from the 'Sanger Genetics Centre' and requests made directly by email. Cosmids were received as bacterial agar stabs. Immediately upon receipt an 'instant' glycerol stock was made by taking an inoculum from the stab into 0.1ml of LB media + 30µl glycerol (i.e 30% v/v). This was vortexed, snap frozen in liquid nitrogen and stored at −80°C. In parallel a sterile loop was used to streak the bacteria onto selective LB plates. Single colonies were selected, large blobby and pin prick sized colonies were avoided. Selected colonies were grown overnight in 4ml LB selective media and used to make glycerol stocks.

2.1.12 Extraction of Cosmid DNA

Cosmid DNA was extracted using the QIAGEN Plasmid Purification Miniprep Kit (Qiagen). The standard microcentrifuge protocol for the isolation of plasmid DNA was modified in the following ways, according to the suppliers' (Qiagen) recommendation. A single colony was picked from a freshly streaked LB agar plate and used to inoculate a 10ml volume of LB media, containing the appropriate antibiotic. When preparing the lysate all mixing was performed with very gentle inversion and pipetting was performed using tips with enlarged openings. In both cases this minimises shearing of the delicate cosmid DNA. An additional wash step with buffer PB, containing isopropanol was included. DNA

was eluted from the column using either buffer EB or H_2O preheated to 70°C. Cosmids were analysed by restriction digests for confirmation.

2.2 DNA Manipulation

2.2.1 Restriction Digests

Restriction digests were used to characterize plasmid DNA and generate complementary DNA ends for subcloning. Restriction endonucleases were supplied by Roche or New England Biolabs and used in accordance with their recommended incubation conditions. Typically restriction reactions were based upon an activity of 1Unit of enzyme digesting 1μg of lambda DNA in 1hr. Restriction digests were optimised for linear and circular DNA based upon supplemental data provided by the supplier.

2.2.2 Electrophoresis

DNA was analysed on ethidium bromide agarose gels underneath ultra violet light. Agarose gels were prepared at various percentages to achieve optimal separation of DNA fragments. A 1% mini-gel was made by adding 0.5g of agarose to 49mls of water and 1ml of 50x Tris-acetate (50x Concentrated Stock Solution: 100ml/L 0.5M EDTA pH 8.0, 2M TrisBase, 57.1 ml/L Glacial Acetic Acid, make up to 1L with ddH_2O). The agarose was dissolved by heating the mixture in a microwave for 2 minutes. After 2 minutes the molten gel mix was allowed to cool slightly and 1μl of Ethidium Bromide (EtBr) [10mg/ml] was added. The gel was then poured, allowed to set and run in an electrophoresis chamber containing 1X Tris-acetate (TAE) buffer. Large gels were run at 120mV and mini-gels were run at 75mV. When DNA was extracted from the gel visualization was performed using low intensity UV and the DNA purified using the QIAGEN Gel Extraction Kit (Qiagen) and the microcentrifuge protocol. RNA was run on ethidium bromide agarose mini-gels, at high voltage (~100mV) for a short time span (20-30 min) for analysis.

2.2.3 PCR Reactions

Enzymes routinely used - *platinum pfx*   Invitrogen

Hi-Fidelity PCR Master	Roche
Expand Long Template PCR System	Roche
Taq polymerase	Qiagen

The polymerase chain reaction (PCR) was routinely used to generate DNA fragments for analytical and cloning purposes. *Platinum pfx* is a proof reading enzyme with 3'-5' exonuclease activity. It was used to generate fragments ranging in size from 100 – 3kb from cDNA. Hi-Fidelity PCR Master is a mixture of the proof reading *tgo* DNA the *taq* DNA Polymerase. PCR products are a combination of blunt ended and 3' single Adenosine overhangs. The kit was used for amplifications of up to ~5Kb from both genomic and cDNA templates. Expand long Template PCR system is a proof reading system and was used for amplifications ranging from 3kb-22kb, primarily using genomic or cosmid DNA as a template. In all cases the PCR reactions were setup on ice. Initial reactions were setup using the concentrations recommended by the manufacturer of the enzyme for each reagent. The cycling parameters were optimised based upon the manufacturers guidelines. Oligonucleotides used in PCR reactions were ordered from Invitrogen and melting temperatures predicted according to A/T = 2°C and C/G = 4°C. This was done as routine and where specific cycling conditions have been used it is stated in the figure legend. PCR products intended for use in cloning reactions were purified either directly, using the QIAGEN PCR Purification kit (Qiagen), or purified by gel extraction (Qiagen).

2.2.4   Sequencing

Double stranded DNA and PCR products were sequenced using the ABI PRISM DNA Sequencing kit. Sequencing reactions were setup on ice. A typical Dye-labelled terminator reaction consisted of :

TERMINATOR Ready Reaction Mix   4μl

(The ready reaction mix contains BigDye terminators (labelled with dRhodamine acceptor dyes), dNTPs, AmpliTaq DNA Polymerase, FS, r*T*th pyrophosphatase, magnesium chloride and buffer premixed into a single tube)

Template	Double stranded DNA	250-300ng
	PCR product	50-100ng
Primer		1.6pmol
ddH₂O		up to a final volume of 10µl

The sequencing cycle used was :   96°C   10s

                                                 50°C   5s

                                                 60°C   4min   for 30 cycles.

2.2.5   Precipitation of Sequenced DNA

To each sequencing reaction 10µl of ddH₂O, 2.0µl of 3M sodium acetate – pH 4.6 and 50µl of 95% ethanol was added in a 1.5ml eppendorf tube. The tubes were vortexed to mix and left at room temperature for 1 hour to precipitate the dye-incorporated extension products. Samples were then spun in a microcentrifuge for 20 minutes at 14,000 rpm. The supernatant was removed and the pellet overlaid with 250µl of ice cold 70% ethanol, vortexed briefly and spun for a further 10 minutes at 14,000 rpm. Again the supernatant was removed and the pellet dried in a heat block at 90°C for 1 minute. Automated sequencing was performed by the School of Biological Sciences, University of Southampton. Chromatogram files were analysed manually using chromas software (Tecnelysium Pty Ltd). High quality sequencing data produced a clean signal with even peak heights. However, where bases could not be called by the computer software, due to poor sequence data, it was performed manually. In circumstances where the base could not be called either by the software or manually the DNA was re-sequenced.

2.2.6 TOPO Cloning Reactions

The TOPO cloning kits (supplied by Invitrogen) were used to directly clone PCR products into a plasmid vector for sequencing. Two types of kit were used in parallel, the TOPO TA Cloning kit and the Zero Blunt Topo cloning kit. In brief, the TA cloning kit takes advantage of the non-template dependent addition of a single deoxyadenosine at the 3' end of PCR products amplified by *Taq* DNA polymerase. The Zero Blunt cloning kit allows PCR products with blunt ends, typically generated by proof reading DNA polymerases to be effectively cloned. TOPO cloning reactions were performed according to the suppliers' protocol (Invitrogen) and suggested modifications made where applicable. A typical TOPO cloning reaction consisted of:

Vector	1μl (10ng)
PCR Product	4μl
Salt Solution (1.2M NaCl, 0.06M MgCl$_2$)	1μl
Total Volume	6μl

In situations where DNA had been purified by gel extraction for TOPO cloning, 3' A-overhangs were added by incubating with 1x Taq polymerase buffer (Qiagen), 1μl of 10mM dATP (Promega) and 2 Units of *Taq* polymerase (Qiagen) for 15 minutes at 72°C. 4μl of this was used to transform chemically competent TOP-10 cells (Invitrogen).

2.2.7 Ligation of cohesive ends for plasmid cloning of DNA fragments

DNA fragments for cloning were digested with the appropriate restriction endonucleases to generate cohesive ends and then purified using the QIAGEN PCR Purification kit (Qiagen). Plasmid DNA was prepared in the same way and where necessary gel purified using the QIAGEN Gel Extraction Kit (Qiagen). The plasmid vector was dephosphorylated using Calf Intestinal Alkaline Phosphatase (CIAP) (Roche) and purified using the QIAGEN PCR Purification kit (Qiagen). A typical dephosphorylation reaction consists of:

Vector DNA	25μl (~ 4-5ug)

10x Buffer	4μl
CIAP (1U/μλ)	2μl
ddH2O	9μl
Total Volume	40μl

Vector DNA and DNA fragments prepared in this way were used in ligation reactions. A typical ligation reaction consists of :

Vector	50-100ng
3:1 molar ratio of insert:vector	
10x Ligase Reaction Buffer (Roche)	2μl
T4 DNA Ligase (Roche)	2 Units
ddH20	20μl final volume

The ligation reaction was incubated at 16°C overnight and 2μl used to transform XLI-Blue (Stratagene) electrocompotent cells.

Name	Vector Author	Method Used
PGILDA	Origene	
pGILDA-MGL1CT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGEX-MGL-1CT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGILDA-MGL2iiCT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGEX-MGL2iiiCT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGILDA-MGL3iCT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGILDA-MGL3iiCT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGILDA-MGL3iiiCT	JD	PCR amplification and insert into *Eco*RI/*Sal*I
pGEX-MGL3iiCT		
pPD95-97	NK	
pPD95-97-*tag-60(C01F6.6a)*	JD	PCR amplification and insert *Bam*HI/*Age*I
pPD95-97-*mpz-1(i)*	JD	PCR amplification and insert *Hind*III/*Age*I
pPD95-97-*mpz-1(ii)*	JD	PCR amplification and insert *Sph*I/*Bam*HI into pPD95-97-*mpz-1(i)*
pHAB200-GFP	HB	
pHAB200-GFP-*mgl-1*	JD	Asc1/Not1 lift out of pCR2.1-*mgl-1*
pHAB200-GFP-MGL1	JD	PCR amplification and insert *Ngo*MIV/Apa1 into pHAB200-GFP-*mgl-1*
pHAB200-CFP	HB	
pCR2.1	Invitrogen	
pCR2.1-*mgl-1*	SP	PCR amplification and insert TA
pCR2.1-*ptp-1*	JD	PCR amplification and insert TA
pHAB220-CFP	HB	
pHAB200-CFP-*ptp-1*	JD	*Sal*I/*Not*I lift from pCR2.1-*ptp-1*
PHAB200-CFP-*mpz-1 (iii)*	JD	*Sph*I/*Bam*H1 lift out of pPD95-97-*mpz-1(ii)* and PCR amplification and insert *Bam*HI/*Not*1
PHAB200-mRFP-1-*mpz-1(iii)*	JD	CFP in PHAB200-CFP-*mpz-1 (iii)* substituted for mRFP-1
PHAB200-CFP-*ptp-1a*	JD	pCR2.1-*ptp-1* cut and inserted into pHAB200-CFP
PHAB200-mRFP-1-*ptp-1a*	JD	CFP in PHAB200-CFP-*ptp-1a* substituted for mRFP-1
pBX-*pha-1*	Neline Kriek	

Table 2.2 Plasmid Constructs. Abbreviations: NK – Neline Kriek, HB – Howard Baylis, SP-Stuart Pullen, JD – James Dillon.

PRIMER NAME	APPLICATION	SEQUENCE 5'→3'
RACE		
MGL-1 GSP-1	mgl-1 5'RACE	GATGGCGAGCTCTCGTCATCTGTCATCCTC
MGL-1 GSP-2	mgl-1 3'RACE	CTCCGCCTGGGTCAACGGCATCAAGGTGC
CeMGL1 Gene Antisense Internal	mgl-1 5' RACE	CCGCTTTTTATCCATGGATTTTGG
CeMGL1 Gene Sense Internal	mgl-1 3' RACE	CCAAAATCCATGGATAAAAAGCGG
MGL-1 Full Sequencing (A/S)	nested mgl-1 5' RACE	
mgl-2 GSP-1 (A/S) 5'RACE	mgl-2 5' RACE	CGATGCTGTTCGTGTGACTCCTCCACCCATAC
mgl-2 GSP-1 (S) 3'RACE	mgl-2 3'RACE	GTCGGAGTTGGTTTGATGCGGGATTGGCCGGATG
mgl- 2 GSP (A/S)	nested mgl-2 5' RACE	GTCGGATAAGTCTGGAGTGGTGGCGGAG
mgl-3 GSP-1 (A/S) 5' RACE	mgl-3 5'RACE	GAAACCGGAGGAATTGGCACGCAAAAATGAAGAG
mgl-3 GSP-2 (S) 3'RACE	mgl-3 3'RACE	GAATGGTGACGGAATCGGACGATATGATGTCTT
mgl-3 GSP (A/S)	nested mgl-3 5' RACE	CGGCTCCTGTTGAACTGTAGCTGACTTGAGG
C01F6.6a 5' RACE	c01f6.6a 5' RACE	GCAGTCGATGATTCCTTGTATTGATAAGCCGAG
C01F6.6a 3' RACE	c01f6.6a 3' RACE	GGACCTTCCAAATATCGAACGAGTTTCTCCAATG
MPZ-1 5' RACE	mpz-1a 5' RACE	CAACTTCTGCTCCTGTTGCTGCTGCTGTTTCTGCTC
MPZ-1 3' RACE	mpz-1a 3' RACE	GAATCCCCACCGCTTCCACCTCCACCGCCGGAAG
PTP-1 5' RACE	ptp-1a 5' RACE	GTGTTATAGGCCTGCTGATTGGTTGGTGGAAGTC
PTP-1 3' RACE	ptp-1a 3' RACE	GCACCGTCTCTCCGTCAGCAACGTCTCAGCAAAGA
SL RACE		
SL-1		GGTTTAATTACCCAAGTTTGAG
SL-1(A)		GGTTTAATTACCCAAGTTTGAG(A)
SL-1(T)		GGTTTAATTACCCAAGTTTGAG(T)
SL-1(C)		GGTTTAATTACCCAAGTTTGAG(C)
SL-1(G)		GGTTTAATTACCCAAGTTTGAG(G)
SL-2		GGTTTTAACCCAGTTACTCAAG
YEAST-2-HYBRID		
CemGluR1 5' Prey	mgl-1 CT (*Eco*RI)	CAGAATTCGAAAAACACAAAAACGTCCGAAAG
CemGluR1 3' Prey	mgl-1 CT (*Sal*I)	CAAGTCGACTCATAAGAAAGTATCGTGAGC
CemGluR2 5' Prey	mgl-2 CT (*Eco*RI)	CAGAATTCCATCCTGAGAAGAATATCAGA
CemGluR2 3' Prey	mgl-2 CT (*Sal*I)	CAAGTCGACTCAAAAGATTTGCTTGAAATC
CemGluR3.2a 5' Prey	mgl-3a CT (*Eco*RI)	CAGAATTCCAACCATCAAAAATGTGAGG
CemGluR 3.2a 3' Prey	mgl-3a CT (*Sal*I)	CAAGTCGACTCAAAGAAAAGTGGAATTAGTGTC
glr-1 1F	glr-1 CT Sense primer (*Eco*RI)	CAGAATTCTTCTTGTATCGAAGTAGGATTGAAGC
glr-1 1R	glr-1 CT Antisense (*Sal*I)	CAGTCGACTCAGACAGCTGTGTTGTAGAGATGT
LET-23CT (A/S) STOP *Xho*I	let-23 CT (*Xho*I)	CACTCGAGTTAAAGACAAGTTTCCTTTTGTGATAC
LET-23CT (S) *Nco*I	let-23 CT (*Nco*I)	CACCATGGGACAAAAAGCTGGGAGCCGGAGC
LET-23 CT (A/S) *Xho*I 3'UTR	let-23 CT + 3'UTR (*Xho*I)	CACTCGAGCAACGGGAATGACAAAAAATGTTCTG
NMR-1 (S)	nmr-1 CT (*Eco*RI)	CAGAATTCGAGGTTTCTTATGGAAGACGGCTAG
NMR-1 (A/S)	nmr-1 CT (*Sal*I)	CAGTCGACTCACACATAAAATCTAGTTGATCTTG
NMR-1 3'UTR	nmr-1 CT + 3'UTR (*Sal*I)	TGGTCGACGAAATGCCAAATTCACACATAAAATC
3'TFP	pJG4-5 sequencing primer (Antisense)	GCCGACACCCTTGATTG
5'TFP	pJG4-5 sequencing primer (Sense)	CTGAGTGGAGATGCCTCC
pGILDA 3' Seq - AKA BFP 3'	pGILDA Sequencing	CGTCAGCAGAGCTTCACC
REPORTERS		
5KB USTR *Asc*I Sense	MGL-1::GFP gene/protein fusion	GGCGCGCCGTCATATTTACGCCTCCTCCG
3KB USTR *Asc*I Sense	MGL-1::GFP gene/protein fusion	GGCGCGCCGAATCGTCTAACGGACGATTC
1KB DSTR *Apa*I (A/S) Ver2	MGL-1::GFP protein fusion	CAAGGGCCCGTCTTTTGGAGACTTATTCTC

cDNA Stop *NgoMI*V Antisense Ver2	MGL-1::GFP protein fusion	CAAGCCGGCTCATAAGAAAGTATCGTGAGC
*Not*I Sig.Seq-GFP Antisense	MGL-1::GFP gene/protein fusion	GCGGCCGCTGTCTGAAAGATCTCTTTTTTTG T
GFP-MGL-1 Link *NgoMI*V Ver2	MGL-1::GFP protein fusion	CAAGCCGGCAAACGGAAGCCGTCGACCTCG CTCC
GFP-MGL-1 Link *EcoR*I Sense	MGL-1::GFP protein fusion	AAGAATTCACGGAAGCCGTCGACCTCGCTC C
C01F6 Rev	Construction of *tag-60a*::dsRED2 promoter construct	CAACCGGTGGACTATCCGGATCGACGGTAC CAACG
C01F6 Forwd	Construction of *tag-60a*::dsRED2 promoter construct	CAGGATCCGAGCTCAGCTTCAACTTCTCAG ACAGG
T19E10 Rev	Construction of *mpz-1(i)*::dsRED2 promoter construct	CAACCGGTGCTCCTTTCGTATCGTCGGAGG AGAC
T19E10 Forwd	Construction of mpz-1(i)::dsRED2 promoter construct	GCAGAAGCTTCTGTTCATCTACGTGGATGC
MPZ-1/2 (AS) *Not*I	mpz-1(iii)::CFP promoter construct	CAAGCGGCCGCAGGACTGCCTGGTAGAATA G
MPZ-1/2 (S)	mpz-1(iii)::CFP promoter construct	GACGGGCACCGAGACTTCGTGAGCATG
mpz-1 5UTR Sense *Sph*I	mpz-1(ii)::dsRED2 and mpz-1(iii)::CFP promoter constructs	CATGCATGCATCTCCTCTATACTTTTCCTC
mpz-1 *BamH*I Antisense	mpz-1(ii)::dsRED2 and mpz-1(iii)::CFP promoter constructs	TTTGGATCCAAAAGGCCATCAAGGAC
PTP-1 5'UTR-EXON3 (A/S)	PTP-1(i)::CFP promoter construction	GCGGCCGCCAACAACGCAAGTGATCCTTC
PTP-1 5'UTR-EXON3 (S)	PTP-1(i)::CFP promoter construction	CGAGCTTGGCAAACGGAAGTTTCG
PTP-1 Intron1 Sense	PTP-1(ii)::CFP promoter construct	GGCAGGTAGGTAGGCAGGCCGCAGG
PTP-1 Common Exon	PTP-1(ii)::CFP promoter construct	GTCAGCAACGTCTCAGCAAAGAGGCG
mRFP-1 PCR out Sense *Not*I	Substitution of CFP with mRFP-1	CAGCGGCCGCTATGGCCTCCTCCGAGGACG TCATC
mRFP-1 PCR out Antisense *Nhe*I	Substitution of CFP with mRFP-1	CTGCTAGCCAGGGCGCCGGTGGAGTGGCGG CC
pHAB GFP N-terminal	Sequencing (Originating from Vector Backbone)	CAACAAGAATTGGGACAACTCCAGTG
pHAB GFP C-terminal	Sequencing (Originating from Vector Backbone)	GCTGCTGGGATTACACATGGCATG
pHAB GFP (A/S) Sequencing MCS2	Sequencing (Originating from Vector Backbone)	AATTTGTGCCCATTAACATCACCATC
pHAB 5'	Sequencing (Originating from Vector Backbone)	CACTCACAACGATGGATACGCTAAC
pHAB 3'	Sequencing (Originating from Vector Backbone)	CAAATAGGGGTTCCGCGCAC
MUTANTS		
mpz-1(gk273)		
gk273 EXTLEFT (A/S)	Genomic characterization	TCTGCCGCTTTCATCTTTTT
gk273 EXTRIGHT (S)	Genomic characterization	AAACGAGAGCACAAGCCAAT
gk273 INTLEFT (A/S)	Genomic characterization	AGGTGCTGTTTGGCAGAGAT
gk273 INTRIGHT (S)	Genomic characterization	GACGCATTTCGGTTTTTAGC
gk273 cDNA (A/S)	cDNA characterization	CTTTTGTCTTTTTACAACCACCACCTC
gk273 cDNA (S)	cDNA characterization	CTCCGGAAACCTTCAAATCAGATAATC
mpz-1(tm1136)		
tm1136 EXT (A/S)	Genomic characterization	CGTAGTGTTTTCAGTCGCTA
tm1136 EXT (S)	Genomic characterization	TGGCCGGAATCTGCTAAACT
tm1136 INT (A/S)	Genomic characterization	GTGTGTTGCGGTCAGACTAT
tm1136 INT (S)	Genomic characterization	CAGTCCTGCAGATAAAGTGA
tm1136 cDNA (A/S)/ *mpz-1-GSP-RT*	cDNA characterization/first strand cDNA synthesis	CCACAATCAACGTTACAGGAGCTCCAG
tm1136 cDNA (S)	cDNA characterization	CAAATAGGCAATATCAACTCACACGGGATG AG
EST Sequencing		
Seq YK1004e08 (1) S		GTCTCCTCCGACGATACGAAAGG
Seq YK1004e08 (2) S		GCAGACAAACCAATAGATACTCCAG
Seq YK1004e08 (3) S		GTTAGTCGTAGCAGTAAACGGAATAG
Seq YK1004e08 (4) S		GGAGACGCAAGTGATTGATGTGAAG
Seq YK1004e08 (5) S		CCATTTCTGGCGGCAAACCAGG
Seq YK1004e08 (6) S		CAGGAAACTGTGTGGATTGGTACAG
Seq YK1004e08 (7) S		CACATACATCACATCAGCGGATGC
Seq YK1004e08 (8) S		GTTCAACCAACTCCTCGGGATCTTG
Seq YK1004e08 (9) S		GTCAAGAGTCACGTCAAGGACACC
Seq YK1004e08 (10) S		GCTACCACCGACTGCGCTCGTAAC
Seq YK1004e08 (11) S		GATGAGCCACAGGCAATGTCACC
Seq YK1004e08 (12) S		GTACGAAGTGACACGATCACCAGC

Table 2.3 Oligonucleotide List. Oligonucleotides were stored at –20°C as stock solutions at 50pmol/µl. Oligonucleotides were supplied by Invitrogen.

2.2.8 *In silico* analysis.

The *C. elegans* gene repository, Wormbase, was used to access the gene details for specific proteins of interest. This resource allowed the cDNA sequence, genomic DNA sequence and protein sequence to be downloaded and used for further *in silico* analysis.

The transmembrane topology of the receptor was predicted using two separate programmes in parrallel, they were HMMpredict (http://www.enzim.hu/hmmtop/) and TMHMM Server v 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/) The protein sequence of each receptor was separately copied into the programmes. Hydropathy plots were generated, compared to identify predictions in common and used to map the transmembrane domains.

Protein sequence allignments, for the identification of conserved regions, were performed using the ClustalW version 1.82, accessed via the EMBL-EBI site (http://www.ebi.ac.uk/clustalw/#). Phylogenetic trees were constructed using Clustal W 1.82, default settings. NCBI (National Centre for Biotechnology Information) database screens and homology searches were performed with the protein sequence of interest, using the Basic Local Alignment Sequence Tool (BLAST).

Signal peptide cleavage sites were predicted using the SignalP V2.0 server (http://www.cbs.dtu.dk/services/SignalP-2.0/)(Nielsen, H et al 1999). The protein sequence of interest was downloaded from Wormbase and inserted into the programme. The parameters were set to analyse eukaryotic data.

2.3 Molecular techniques for Yeast

2.3.1 Media

Yeast Media Stock Solutions

100x Tryptophan (T)	4mg/ml
100x Uracil (U)	2mg/ml
100x Leucine (L)	6mg/ml
100x Histidine (H)	2mg/ml

All amino acid solutions were filter sterilized. All other media was sterilised by autoclaving.

10x SD Droupout mix (-Histidine, -Uracil, -Tryptophan, -Leucine)	
20x Glucose	40% w/v glucose
20x Galactose	40% w/v galactose
40x Raffinose	40% w/v raffinose
5x Yeast Nitrogen Base (YNB)	33.5g/L
2x Yeast Nitrogen Base Agar	6.7g YNB, 20g Agar in 500ml H_2O

YNB agar was autoclaved and supplemented with SD, glucose or galactose/raffinose and the appropriate amino acid combinations.

2.3.2 Glycerol Stock Cultures of Yeast Strains

Yeast strains were grown in 5 ml of the appropriate selective SDO medium overnight and 0.75 ml of the yeast culture was added to 0.25 ml of 100% glycerol (slerilized by autoclaving). The cryogenic tube was vortexed to mix and placed at $-80^{\circ}C$.

2.3.3 Medium scale yeast transformation

A 2-3 day old yeast colony was inoculated into 5ml of selective media and grown overnight with shaking. The next day the overnight culture was diluted to an OD_{600} of 0.1 in a total volume of 60ml of selective media and grown until the OD_{600} was 0.6-0.8. The cells were pelleted by centrifuging at 1 500g for 5

minutes, the supernatant removed, the pellet resuspended in 20ml of ddH2O and spun again at 1 500g for 5 minutes. The supernatant was removed and the pellet resuspended in 1ml of ddH2O in a microcentrifuge tube. The sample was centrifuged for 30 seconds at 14 000g and all water removed from the tube. The pellet was resuspended in 300μλ TE/LiOAc (made immediately beforehand: 1 part 10x TE, 1 part LiOAc, 8 parts ddH2O). In a separate eppendorf tube 100ng of each plasmid being transformed was added to 50ug of salmon carrier DNA, maintaining the smallest volume possible. Then 100μl of the resuspended cell suspension was added to the DNA together with 300μl of TE/LiOAc/PEG (made immediately beforehand: 1 part 10xTE, 1 part LiOAc, 8 parts PEG). The DNA was mixed with the cells by gentle inversion. The samples were placed at 30°C, after 30 minutes 70μl of DMSO was added and the cells heat shocked at 40-42°C. After 15 minutes the cells were pelleted at 14 000g for 5 minutes. The supernatant was removed and the pellet resuspended in 500μl of ddH2O. The cells were centrifuged again and the pellet resuspended in 500μl of ddH2O. The cells were plated onto selective agar media and incubated at 30°C for 3-4 days.

2.3.4   β-galactosiadase Filter Lift Assay

Whatman filters (sterilized by wrapping in aluminium foil and autoclaving) approximately the same diameter as the plate of yeast transformants being assayed were soaked in 5ml of Z-Buffer/X-gal (100ml Z Buffer, 0.27ml β-mercaptoethanol, 0.334μl X-gal [100mg/ml]) solution. A clean filter was placed over the surface of the plate of colonies being assayed. The filter was gently rubbed with the edge of the forceps. This improves adhesion of colonies to the filter but does not completely remove the colony from the agar surface. Holes were poked through the filter at three asymmetric positions allowing the filter to be orientated with the agar. The filter was then lifted from the plate using forceps and transferred to liquid nitrogen for 30 seconds and allowed to thaw. This was repeated twice, the filter was then placed onto the pre-soaked filter with the colonies facing upwards taking care to avoid creating air bubbles between the two filters. The lid, containing the filter lift, was placed back onto the agar plate, inverted and incubated at 30°C. Placing the filter lift in the lid of the assayed plate prevented the soaked filter from drying before blue colour was observed. The

filter lifts were checked for blue colour every hour. The β-galactosidase producing colonies were identified by aligning the filter lift with the agar plate using the asymmetric marks. The corresponding colonies were picked from the plate and streaked onto fresh selective media and incubated at 30°C. Where the colony has been completely lifted from the agar surface or cannot be identified accurately the colony was picked from the filter and used to make the lift and restreaked onto fresh media.

2.3.5   Crude Plasmid Isolation form Yeast

A 2-3 day old colony, 2-4mm in size was inoculated into 0.5ml of selective SDO liquid media in a 1.5ml microcentrifuge tube. The culture was incubated overnight at 30°C with shaking (~230 rpm). The next day samples were spun at 14,000 rpm for 5 minutes and the supernatant removed. The cell pellet was resuspended in the residual supernatant, 10μl of lyticase solution (5units/μλ in TE buffer) was added. The sample was vortexed to mix and incubated at 30°C. After 60 minutes the tube was removed, 10μl of 20% SDS was added, the sample was vortexed for 1 minute and placed at –20°C. The tubes were removed after 30 minutes, allowed to thaw and vortexed for another 1 minute. At this stage the yeast cells should be completely lysed and the plasmid DNA can be purified. The sample volume was made up to 200μl with TE buffer (pH 7.0), 200μl of phenol:choloroform:isoamyl:alcohol (25:24:1) was added and the tubes vortexed. After 5 minutes of vigorous vortexing the tubes were spun at 14,000 rpm for 10 minutes, the upper aqueous phase removed to a fresh microcentrifuge tube and 8μl of 10M ammonium acetate was added together with 500μl of 95% ethanol. The sample was placed at –80°C or 1hr. After 1hr the tubes were centrifuged at 14,000 rpm for 10 minutes, the supernatant removed and the pellet allowed to air dry for 20 minutes. The dried pellet was resuspended in 20μl of ddH2O

2.3.6   Large Scale Library Transformation

EGY48 co-transformed with pGILDA*mgl-1ct* and pSH18-34 was grown overnight in –HU YNB(glu) selective media. The culture was diluted to an OD_{600} of 0.2-0.3 in 1L of –HU YNB(glu). The culture was grown for 4hrs at 30°C to an OD_{600} of 1 and pelleted by centrifuging at 1 000g for 5min. The SNT

(supernatant) was removed, the cells were resuspended in 500ml TE and pelleted again. The cells were washed in 20ml TE/LiOAC, pelleted and resuspended in a final volume of 20ml TE/LiOAC. The cells were mixed gently with 10mg of the *C. elegans* cDNA library plasmid (pJG4-5) and 20mg of carrier DNA and 150ml of TE/LiOAC/PEG added. The suspension was gently mixed and after incubation of 30min at 30°C the mix was removed and 17ml of DMSO added. The cells were the heat shocked at 42°C for 15min, with gentle mixing every 1min. Following this cells were immediately placed on ice for 5min, washed twice with 20ml TE and a volume of 185μl plated onto 110 –HUT YNB (glu) selective agar plates. Serial dilutions were made and plated onto selective plates, allowing number of transformants obtained to be calculated.

Plates were incubated at 30°C for 3 days and transformants harvested by scrapping cells into water. The cells were pelleted at 1 000g for 5min, the supernatant removed and resuspended in a volume of 50% glycerol equal to the volume of the cell pellet. The resuspended cells were alqouted as 1ml volumes and frozen at –70°C.

2.3.7 Titration of the number of viable cells

The transformation efficiency of the *C. elegans* cDNA library into the EGY48 cells containing pGILDA*mgl-1ct* and pSH18-34 was calculated as 1.1×10^5/ug of plasmid. A 1ml aliquot of frozen transformants was thawed on ice and diluted into 10ml of –HUT YNB(galactose) and grown for 4hrs to induce the Gal1 promotor and fusion-protein expression. 100μl of the cell culture was used to make serial dilutions and 100μl of each dilution was plated onto –HUT YNB(gal) selective plates. Plates were incubated at 30°°C for 3 days. The number of viable transformants was titrated to 120×10^7 /1ml.

2.3.8 The Screen

1 aliquot of transformants was thawed on ice, diluted into 9ml of –HUT YNB(gal) and grown at 30°C for 4hrs. After incubation this was added to 110ml of –HUTL YNB(gal). 200μl of this was spread onto –HUTL YNB(gal) selective plates, 200 plates were spread. A total of 400×10^6 transformants were screened, at a density of 2×10^6 transformants per plate. Plates were incubated for a total of 15

days. Clones were selected on the basis of their growth under selection for a positive interaction at days 3, 5 and 7 of incubation. By day 10 a large background growth was evident and β-galactosidase filter lift assays were performed to isolate clones expressing the *lac-Z* reporter gene. Colonies identified as expressing reporter genes were picked and spotted onto –HUT YNB(glu) plates as back ups and onto –HUTL YNB(gal) + X-gal. Colonies were taken through three rounds of selection for reporter gene expression by streaking onto –HUTL YNB(gal) + X-gal. Colonies still exhibiting reporter gene expression after 3 rounds of selection were prioritised. Colonies were picked and spotted onto –HUTL YNB(gal) + X-gal, the amount of each colony was standardized by using a 10μl pipette tip to perform the picks. The incubation of the prioritisation plates at 30°C were synchronised. Colonies were classified into 24, 48, 72 and 96 hour groups based on their strength of growth under leucine selection and the intensity of blue colouration they exhibited.

2.3.9 Characterization and Cataloguing of clones

Crude plasmid DNA preparations were made from the yeast colonies yielded by the screen. Existing oligonucleotides designed against flanking regions of the pJG4-5 library vector multiple cloning site (5'-TFP and 3'-TFP) were used to amplify cDNA inserts. 2μl of the crude yeast extract was used in the PCR reaction, which was performed using *Taq* enzyme (Qiagen). 10μl of the PCR products were analysed by agarose/EtBr and 5μl of the PCR reaction was used for restriction digest with the restriction enzyme HaeIII (Roche). Comparisons allowed unique clones to be isolated on the basis of their size and HaeIII restriction profile. The identity of unique clones was determined by the sequential purification of the PCR amplified product and sequencing. PCR products were purified using μ-CLEAN (Microzone Ltd) reagents. Sequencing was performed using the oligonucleotides: 5'-TFP and 3'-TFP. Clones were characterised by screening the sequencing data against genome databases (Wormbase) for alignments.

2.3.10 Plasmid Rescue from Yeast

The crude yeast extract was a mixture of the prey, bait and reporter plasmids, it was therefore necessary to isolate the pJG4-5 library vector that contains the cDNA of interest. This was done by transforming XLI-Blue electrocompotent cells with 5μl of the crude extracts prepared from the appropriate yeast colonies. Transformed bacteria were spread onto LB agar plates containing ampicilin (LBamp) together with 40μl of IPTG (100mM) and 10μl X-Gal (100mg/ml). The XLI-Blue strain of bacteria are deficient for the *lac-Z* gene, therefore in the presence of IPTG the pSH18-34 reporter plasmid will confer *lac-Z* expression and the ability to metabolise X-Gal. Colonies containing pSH18-34 could therefore be discriminated by their blue colouration. White colonies were picked and crude DNA extractions made. Bacteria containing the prey vector (library clone) were characterised by Xho-1/EcoR1 restriction digest, to release the cDNA clone.

2.4 Molecular and Genetic Techniques for *C. elegans*

2.4.1 *C. elegans* nomenclature

Gene names conform to a structure consisting of three of four italicized letters, a hyphen and an italicized number. When the gene name is followed by an italicized Roman numeral this indicates the linkage group to which the gene maps. The protein product of a gene is written as the gene name in non-italic capitals. Transgenes exisitng as extrachromosomal arrays are given italicized names consisting of the laboratory allele prefix, the two letters *Ex* and a number. Integrated transgenes are designated in the same way but *Is* is used instead of *Ex*. Mutations are given names consisting of one or two italicized letters followed by an italicized number. When the gene and mutation name is used together, the mutation name is included in parantheses after the gene name.

2.4.2 *C. elegans* Strains

Wild-type	N2 Bristol Strain
Transgenic strains	*pha-1(e2123)III ; mgl-1::GFP(utIs35)*
	mgl-1::GFP(utIs35)
	mgl-2::GFP(utIs44)

mgl-2::GFP(utIs35) and *mgl-1::GFP(utIs35)* are integrated lines, which have been outcrossed 3 times. The lines were obtained from Ishihara, T. National Institute of Genetics, Mishima, Japan. The *pha-1(e2123)*III (Schnabel, H et al 1990) strain was obtained from the *C. elegans* Genetics Centre. *(See table 2.3 for transgenic mutant strains).*

Available transgenic mutant worm strains were readily obtainable from two sources, the '*C. elegans* gene knockout consortium' [http://celeganskoconsortium.omrf.org] and the Japanese consortium (Mitani Laboratory at the Tokyo Women's Medical University School of Medicine. http://shigen.lab.nig.ac.jp/c.elegans/index.jsp). Strains were despatched as starved cultures on NGM media. Acquired transgenic mutant worm strains are shown in Table 2.4.

Gene Name	Allele	Genotype	Outcrossed
mgl-1	*tm1811*	Homozygous	✓(x4)
mgl-2	*tm0355*	Homozygous	✓(x4)
mgl-3	*tm1766*	Homozygous	✓(x4)
mpz-1	*tm1136*	Homozygous	X
mpz-1	*gk273*	Heterozygous *gk273/mIn1[mIs14 dpy-10(e128)*	✓ (x1)
pha-1	*e2123*	Homozygous	✓

Table 2.4 A table of acquired mutant strains. Outcrossing of *mgl* mutants was performed by Ting Zeng. In the case of the other outcrossed strains this was performed by the originating laboratory.

2.4.3 *C. elegans* cultivation media

Nematode Growth Media (NGM) :

NaCl   3g/L

Agar   10g/L

Peptone   2.5g/L

Cholesterol   1ml/L of 5mg/ml dissolved in ethanol

H_2O 985ml

Sterilization by autoclaving then add:

1M $CaCl_2$ 1ml/L (Sterilized by autoclaving)

1M $MgSO_4$ 1ml/L (Sterilzed by autoclaving)

1M KH_2PO_4 25ml/L (Sterilized by autoclaving)

2.4.4 Growing OP50

OP50 was generated by picking a single colony from a freshly streaked plate (LB) and transferring to 10ml LB media. Cultures were grown overnight at 37°C without shaking.

2.4.5 Maintenance of *C. elegans* strains

Worms were maintained on NGM plates (9cm diameter) seeded with 150μl *E.coli* OP50 at 20°C (According to Brenner, 1974), unless otherwise stated.

Strains were passaged every 4-6 weeks by cutting $1cm^2$ chunks from existing plates and transferring to freshly seeded plates. Cutting was performed using a scapel blade sterilized with ethanol and flamed over a bunsen.

2.4.6 Trizol extraction of Total RNA from *C. elegans*

When preparing RNA several precautions were taken to eliminate contamination from RNases. Disposable gloves and lab coat were worn at all times to eliminate RNase contamination from the skin. All plasticware used was exclusive for RNA preparation and sterilised either by autoclaving or certified as RNase and DNase free. The bench space was washed with 0.5M NaOH and partitioned off to prevent cross contamination. Distilled water was treated with Diethylpyrocarbonate (DEPC), a strong inhibitor of RNase activity. DEPC was added at 0.1% v/v, left overnight with stirring and before autoclaving the DEPC treated solution (LTE 300SH, 121°C, 30min 1BAR).

Well fed worms 5-7 days old were washed from 40X 9cm diameter NGM plates into a total volume of 80ml (2ml/plate) of M9 buffer (KH_2PO_4 22mM, Na_2HPO_4 42mM, NaCl 86mM, autoclave and then add sterilised 1M $MgSO_4$ to a final concentration of 1mM in $DEPC.H_2O$). Plates were washed taking care to minimise disturbance of the bacterial lawn. Worms were spun down at 4 500g for 10 minutes (RT Sorvall Legend). The supernatant removed and the worm pellet resuspended to a final volume of 2.5ml of M9 buffer. The resuspended worms were incubated for 1hr at 4°C. Worms settle towards the base of the tube while bacteria remain suspended. The supernatant was removed and the mass of the worms measured. 1ml of Trizol (Invitrogen) reagent was added per 50mg of worms and homogenized with a power homogenizer. This stage makes the disruption of the tough surrounding cuticle more efficient, improves cell lysis and maximises RNA yield. Insoluble material was removed from the homogenate by centrifugation at 12 000g for 10 minutes at 4°C. The cleared lysate was left to stand at room temperature for a further 30 minutes and then split into 1ml aliquots in 1.5ml eppendorf tubes. To each 1ml of lysate 0.2ml of chloroform was added, the samples were vortexed and incubated at room temperature for 3 minutes. The samples were then centrifuged at 12 000g for 15 minutes at 4°C. The mix separates into a lower, red, phenol-choloroform organic phase containing DNA and protein. RNA remains exclusively in the colourless upper aqueous phase. This

was removed to a fresh tube and 0.5 ml of isopropyl alcohol added to each tube to precipitate the RNA. The samples were then left at room temperature for 10 minutes and spun at 12 000g for 10 minutes at 4°C. The supernatant was removed and the pellet washed with 1ml of 75% DEPC treated EtOH by vortexing. At this stage RNA pellets were safely stored at −80°C until ready for redissolving and quality analysis.

2.4.7 Routine analysis of RNA preparations to assess quality and integrity

RNA pellets stored at −80°C were spun at 7 500g for 5 minutes at 4°C and the supernatant removed. The pellet was air dried for 10 minutes and resuspended in 500μl of DEPC treated H_2O with heating at 60°C. The quality and quantity of redissolved RNA was routinely analysed by spectrophotometry prior to mRNA purification. The concentration of RNA in solution was calculated from the absorbance at 260nm (A_{260}). In parallel the integrity and size distribution of RNA was analysed on EtBr agarose mini-gels. Electrophoresis tanks were cleaned with 0.5% SDS, rinsed 3 times with DEPC treated H2O, then ethanol and allowed to dry. RNA was denatured by heating at 70°C for 5 minutes prior to loading onto the EtBr agarose mini-gel. The gel was run at high voltage (~100mV) for 20-30 minutes for analysis.

2.4.8 First Strand cDNA Synthesis.

First strand cDNA synthesis was performed using Superscript II Reverse Transcriptase (Invitrogen). The reaction was setup in a 200μl thin walled PCR tube (Starlab). Reverse transcription was performed on 5ug of Total RNA, using either 800ng of Oligo(dT) primer (New England Biolabs - NEB), 200ng of Random Primers (9mers) (NEB) or 2pmol of gene specific primer (GSP). A typical reaction consisted of:

1.	Random primer (50ng/μλ) or	4μl
	Oligo(dT) primer (400ng/μλ) or	2μl
	GSP (2pmol/μλ)	1μl
2.	5ug totalRNA	1μl
3.	dNTP Mix (10mM each)	1μl

4. H$_2$O to a final volume of 12μl

Heated incubations were performed using a Perkin Elmer Thermal Cycler. The mixture was heated to 65°C for 5min, then placed on ice for 2 minutes and the following added:

5x First Strand Buffer	4μl
0.1M DTT	2μl
H$_2$O	1μl

The tubes were mixed by gently flicking and the contents brought to the bottom of the tube by a brief centrifugation. When using GSP or Oligo(dT) the tubes were placed at 42°C for 2min, when using random primers tubes were placed at 25°C for 2min. After 2min 1μl of 200units/μl of Superscript II was added to the tube (when using random primers tubes were incubated at 25°C for 10min). All samples were then incubated at 42°C for 50min and 70°C for 15min. 1μl (2units) of RNase H (GIBCO) was added to the tube and incubated at 37°C for 20min. 2μl of the reverse transcription reaction was used for PCR.

2.4.9 5'and 3' SMART RACE technique

1μg of total RNA was used as the starting material. For the preparation of 5'RACE cDNA 1-3μl of the RNA sample was combined with 1μl of 5'-CDS primer and 1μl of the SMART II A oligo and this was made up to final volume of 5μl with RNase free water in a 200μl thin walled PCR tube. For the preparation of 3'RACE cDNA 1-3μl of the RNA sample was combined with 1μl of 3'-CDS primer and made upto a final volume of 5μl with RNase free water in a separate tube. From this point onwards the 3'RACE and 5'RACE cDNA was prepared in the same way. Contents were mixed by a brief vortex and collected at the bottom of the tube by spinning briefly in a microcentrifuge. The tubes were then incubated at 70°C for 2 minutes and then cooled on ice for 2 minutes. Tubes were spun briefly and the following was added to each tube: 2μl 5x First-strand buffer, 1μl DTT, 1μl dNTP mix (10mM), 1μl PowerScript Reverse Transcriptase. The contents of the tube were mixed by gentle pipetting and spun briefly. Tubes were

then incubated for 1.5 hours at 42°C in an air incubator. 100μl of Tricine-EDTA buffer was then added and tubes were heated at 72°C for 7 minutes. 2μl of the RACE cDNA was used for each RACE PCR. PCR conditions used to perform RACE reactions and nested RACE reactions are as follows:

3' and 5' RACE Cycle: 5x(5sec @ 94°C, 3min @ 72 °C), 5x(5sec @ 94 °C, 10sec @ 70 °C, 3min @ 72 °C), 25x(5sec @ 94 °C, 10sec @ 68 °C, 3min @ 72 °C).

Nested RACE Cycle: 2min @ 94 °C, 30x(30sec @ 94 °C, 30sec @ 68 °C, 3min @ 72 °C), 7min@72 °C.

In all cases the Advantage 2 Polymerase Mix, supplied with the kit, was used to perform the RACE reactions (unless otherwise stated). PCR was performed using a PE GeneAMP 9600 hot-lid thermal cycler.

2.4.10 Extraction of Genomic DNA

Well fed worms 5-7 days old were washed from a 9cm diameter NGM plate with 2ml of M9 buffer. Worms were placed at 4°C for 1hr and the supernatant decanted, removing suspended bacteria. The remaining worms were resuspended in 100μl of lysis buffer (10mM Tris-HCl pH8.3, 50mM KCl, 2.5mM MgCl$_2$, 1ml of 45% v/v NP40, 1ml of 45% v/v Tween 20, 0.01% w/v gelatin) and 1μl of Proteinase K (10mg/ml) freshly added. The sample was frozen at -80°C for 20 minutes and then incubated for 1hr at 60°C and 15 minutes at 95°C in a PCR machine. The concentration of the genomic DNA was assessed by spectrophotometry before use in PCR amplifications.

2.4.11 Long term storage of worm strains

Stocks of wild-type and transgenic *C. elegans* strains were maintained indefinitely by freezing at –70°C. A plate of starved worms were washed off with ~1.5ml of M9 buffer. 0.75ml of this worm suspension was added to 0.75ml of freezing solution (NaCl 0.1M, KH$_2$PO$_4$ 0.05M, glycerol 30% w/v, NaOH, 5.6mM; autoclave and then add 0.1M MgSO$_4$ (sterilised by autoclaving) to a final concentration of 0.3mM). The mix was vortexed and placed at –80°C in a

Styrofoam holder, this enabled slow cooling necessary for survival. The viability of the frozen worm strain was assessed by taking a frozen aliquot, rapidly thawing at room temperature and distributing onto a large seeded NGM plate. After 24hrs at 20°C worms exhibiting normal movement were picked onto a fresh seeded NGM plate and counted.

2.4.12 Decontamination of *C. elegans* Cultures

10μl of bleach was added to 10μl of 4M NaOH and spotted onto the surface of a seeded NGM plate. Gravid adult worms were placed into the solution and incubated at room temperature. The adult worms dissolve and the resistant eggs hatch, larvae crawl to the bacterial lawn and commence feeding. The bleached spot was cut out to remove any contamination that has survived bleaching.

2.4.13 Microinjection of DNA

Preparing needles for injection

Injection needles were prepared from aluminosilicate glass capillaries (1.0mm OD x 0.5m I.D.). A single capillary was used to generate two injection needles with the model P-2000 (Salter Instruments Co).

Preparing Injection solutions

Plasmid DNA prepared from 3 individual colonies using the QIAGEN mini-prep kit was pooled, precipitated with 1M KAc at -20°C and resuspended in a final volume of 20μλ ddH$_2$O. A typical injection solution consisted of:

Rescue construct pBX (pha-1)	30-50ng/μl
Test construct	30ng/μl
5x Injection Buffer	1μl
ddH20	To a final volume of 50μl

The mix was purified by centrifugation at 14000rpm for 30 minutes. 40μl of the supernatant was removed, leaving 10μl behind, to a fresh tube and centrifuged

under the same conditions. This cycle of centrifugation and removal of the supernatant for centrifugation, leaving 10µl behind each time, was repeated 3 times in total.

2.4.14 Mounting worms on agarose pads and injection

Injection needles were inverted and placed into the injection solution, loading occurred by capillary action. The filled needle was mounted on the microinjector arm. A microscope slide mounted with a broken fragment of coverslip provided a suitable jagged edge to break the tip of the needle under high power magnification. With this done the flow of the needle was checked to ensure it had been sufficiently broken without blocking. A drop of oxygen permeable oil was placed on one side of the agarose pad and then placed onto an inverted petri dish beneath a light microscope. Agarose pads were made by placing a drop of molten 2% agarose onto a microscope slide (76 x 26mm ground edges 90°C frosted end, manufacturer: Menzel GmBH), then placing another slide on top. This was left for 5 minutes and then the upper slide removed. The slide with the agarose pad was backed for 1hr at 80°C. Three worms were placed into the oil with a fine pick. One worm was selected, transferred to the agarose pad so that its body is as straight as possible and overlaid with oil. The pad was placed onto the microinjection stage with the body aligned for insertion of the injection needle into the germ line distal core region. Following injection of the gonad arm the worm was rehydrated by applying a wash of the injection mix and removed to a seeded NGM plate.

2.4.15 Generation of the *mgl-1*::GFP ; *pha-1* selectable marker strain

5x *mgl-1::GFP(utIs35)* homozygous males were crossed with 3x *pha-1(e2123)*III homozygous hermaphrodites. After 3 days 3 L4 hermaphrodite worms from the F1 progeny were picked to individual plates and self crossed at 15°C. After 7 days 48 individual hermaphrodites from the F2 progeny of the self crossed worms, were picked to individual plates and placed at 15°C for 5 days. Half of the total number of worms from each of the 48 lines were passaged to seeded plates and placed at the non-permissive temperature of 25°C to select for the *pha-1(e2123)* homozygous genotype. Lines identified as being homozygous for the

pha-1*(e2123)* deletion were not viable at 25°C. Visual inspection of *pha-1(e2123)* homozygous worms allowed identification of the homozygous *mgl-1::GFP(utIs35)* genotype. 15 individuals from each line were analysed, only if all 15 exhibited *mgl-1::GFP(utIs35)* fluorescence was the line considered to be homozygous.

2.4.16 DiI Filling of Amphid Sensory neurons

200µl of stock DiI (Molecular Probes) (2mg/ml in N,N-dimethylformamide and stored at –20°C in a foil wrapped tube) was diluted 1:200 in M9 Buffer. Worms were incubated in 150µl of the diluted DiI solution for 2-3hrs at room temperature in a 1.5ml eppendorf. After incubation the contents of the eppendorf were emptied onto an unseeded NGM plate. Viable worms were picked to fresh, seeded plates and allowed to crawl for 1 hour to destain.

2.4.17 Preparation of Worms for Microscopy

Worms were mounted on slides (76x26mm) with 2% Agarose pads and immobilized in 20mM sodium azide. Agarose pads were made by dissolving 1g of agarose in 50ml of distilled H_2O. Three micorscope slides were placed side by side, the two slides on either side of the central slide were wrapped in autoclave tape. A drop of the molten agaorse was placed onto the central microscope slide using a pasture pipette and a micorscope slide placed ontop at a 90° angle. The agarose was allowed to set on the slide for 5 minutes and the top slide was then removed. A 20µl of 20mM sodium azide was droped onto the agarose pad. Worms were picked into the sodium azide and a coverslip (No. 1.5 18mmx18mm, Chance propper LTD) placed on top. Coverslips were fixed in position using nail polish.

2.4.18 Confocal and Epifluorescence Microscopy

Images were captured sequentially using the Zeiss confocal laser microscope (settings for Dsred2/GFP coexpression lines: GFP: 488nm excitation, 505-530nm band pass filter. DsRed2: 543nm exciation, 560nm long pass filter, 40x Oil objective). DiI stained worms were imaged using the Leica QFluro

epifluorescence microscope, settings: 40x objective, excitation 543nm, emission 560nm.

2.4.19 De-sheathing of worms for imaging

Worms were picked into DENTS saline solution, in a 3cm petri dish. The worms were immobolized by placing the petri dish at -20°C. After 3 minutes the worms were removed and placed underneath a light mocroscope. The tip of the worms nose was cut using a razor blade, contraction of the surrounding cuticle exposed the pharyngeal muscle for imaging.
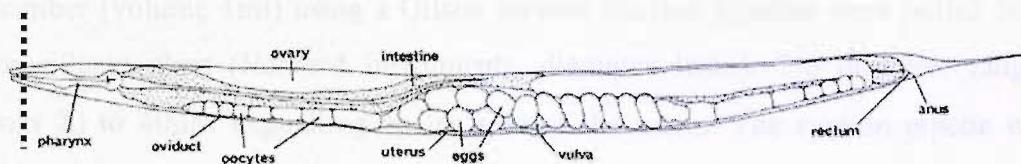
Figure 2.1 The desheathed pharynx preparation. The hashed line indicates the site of the transverse cut performed at the tip of the nose to desheath the pharyngeal muscle.

2.5 Electrophysiological Techniques for *C. elegans*

2.5.1 Electropharyngeogram (EPG) Recordings

5 adult (3-4 day old animals) hermaphrodites were picked from a non-starved plate and placed into a 3cm petri dish containing Dent's saline (10mM D-Glucose, 140mM NaCl, 1mM MgCl$_2$, 3mM CaCl$_2$, 6mM KCl and 10mM HEPES, pH7.4 with NaOH). A sharp razor blade in a blade holder was then used to cut transversely through the worm directly behind the terminal bulb at the pharyngeo-intestinal valve. Dissected heads were then transferred to a custom built recording chamber (volume 1ml) using a Gilson pipette. Suction pipettes were pulled from borosilicate glass (Harvard Instruments, diameter 1mm). Tip diameter ranged from 20 to 40µM depending on the size of the worm. The suction pipette was filled with Dent's saline and mounted in a holder with a tubing port through which suction was applied to capture the nose of the worm. The suction pipette was connected to a silver electrode, which was connected to a HS2 head stage (Axon Instruments), which in turn was connected to an Axoclamp 2B-recording amplifier. The reference electrode was a silver chloride-coated silver pellet in Dent's connected to the recording chamber by an agar bridge in Dent's saline. Data was acquired using axoscope 8.0 (Axon Instruments). All drugs were prepared in Dent's saline at the final concentration and applied by exchanging the entire solution in the recording chamber using a 1ml Gilson pipette.
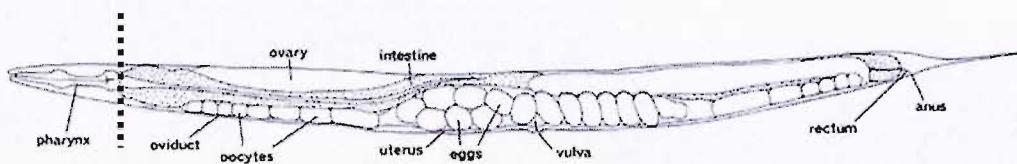
Figure 2.2 Dissection of the pharynx for EPG recordings. The hashed line indicates the transverse cut site at the pharyngeo-intestinal valve.

2.5.2 Intracellular Recordings

The recording chamber was mounted on an inverted microscope and perfused via gravity feed with Dent's saline at a rate of 5 ml/min. The preparation was secured by means of a suction electrode applied to the terminal bulb region of the pharynx and impaled with an aluminosilicate glass microelectrode (1.0-mm OD glass, with filament, pulled on a Sutter P-2000 microelectrode puller; 60–80 M_when filled with 4 M KAcetate, 10 mM KCl) connected to an Axoclamp 2B recording amplifier. The reference electrode was a silver chloride-coated silver pellet in 3 M KCl connected to the recording chamber by an agar bridge. All drugs were applied by addition to the perfusate. Data were acquired and analyzed using pclamp 8 (Axon Instruments). A hard copy of the data, membrane potential, and spike frequency was obtained on a Gould chart recorder (taken from Franks, C.J et al 2002). All intracellular recordings presented were performed by Christopher J Franks.

CHAPTER 3

The molecular characterization of the *mgl* transcripts and the cellular expression of *mgl-1* and *mgl-2*

3. 1   Introduction

There are two classes of glutamate receptor, ionotropic and metabotropic, that both function to co-ordinate glutamatergic transmission in the central nervous system *(see section 1.5 and 1.6)*. The use of model organisms has provided insight into the function of mGluRs and the importance of mGluR signalling. Phenotypic analysis of mGluR knockout mice and rats has shown the neuro-modulatory properties of mGluRs are a very important component of glutamatergic circuitries that underlie complex physiological processes, such as motor-coordination, vision and memory *(see table 1.2)*. The identification and cloning of an mGluR homologue, DmGluRA (Parmentier, M-L et al 1996), from the sequenced genome of the invertebrate, *Drosophila melanogaster,* encouraged the use of this model organism to gain further insights into mGluR function. DmGluRA has similar signalling capabilities and pharmacological properties to mammalian counterparts (Parmentier, M-L et al 2000). As well as this its neuro-modulatory role in glutamatergic transmission is well conserved, within the context of the invertebrate neuromuscular junction (Bogdanik, L et al 2004). In comparison there is limited information about mGluRs in *C. elegans.* As previously highlighted *(see section 1.28)* the use of *C. elegans* has provided insight into the molecular and cellular organisation of glutamatergic signalling and this is particularly well demonstrated by the insights gained from *C. elegans* into the organisation of iGluR function.

Subsequently the sequenced genome of *C. elegans* has been utilised to identify genes encoding mGluRs. The proteins encoded by the identified genes have been annalysed using *in silico* techniques to assess if they have the archetypal mGluR domain architecture *(see figure 1.2)* and to identify protein motifs that are important directing for mGluR function. In parallel molecular approaches have been used to define the 5' and 3' ends of these transcripts.

3.1.2   Sequence comparison of *mgl* subtypes to their mammalian counterparts

The protein sequence of the representative mammalian mGluR subtype, mGluR1$_\alpha$ was obtained from the NCBI database and used to perform a screen of the *C. elegans* genome database to identify candidate *C. elegans* mGluRs.

Amongst the ~19,000 *C. elegans* genes, 3 genes were identified as encoding proteins with homology to mGluR1$_\alpha$. The use of the Basal Local Alignment Search Tool (BLAST) isolated the 3 genes F45H11.4 (*mgl-2*), ZC506.4 (*mgl-1*) and Y4C6A.2 (*mgl-3*). The protein sequences of the candidate genes identified were used to perform individual BLAST searches of the NCBI database. The profile of each screen revealed the *C. elegans mgl* proteins display the highest homology to different mGluR subtypes. The three *C. elegans mgls* display a higher percentage similarity (~40%) to members of mammalian mGluR subgroups than to each other. A percentage similarity of 40% is comparable to the percentage similarity between iGluR homologs in *C. elegans* and mammals, which have been demonstrated as having conserved signalling properties. *(see section 1.28.3).*

This prompted us to construct a cladogram to ascertain the evolutionary conservation between the identified *mgls* and members of the mGluR GPCR family. The protein sequences of mGluR subtypes were obtained from the rat genome database, together with other representative Group III GPCRs. Both the cladogram *(see figure 3.1)* and BLAST screens suggest that the *mgls* are distinct from other Group III GPCRs, such as GABA$_B$ subunits and CaR and more closely related to the mGluR family. Indeed, *mgl-2* is most like mammalian Group I mGluRs, whereas *mgl-1* and *mgl-3* are Group II/III like. Perhaps surprisingly a cladogram constructed from the mammalian mGluR subtypes, DmGluRA and the predicted protein sequence of the *C. elegans mgl* subtypes has revealed little evolutionary conservation between the *Drosophila* metabotropic glutamate receptor (DmGluRA) and the *mgls*.

```
                                                          GABAB1
                                                          GABAB2
                                                          CaR
                                                          mGluR1  ⎤
                                                          mGluR5  ⎬ Group I
                                                          mgl-2   ⎦
                                                          mGluR2  ⎤
                                                          mGluR3  ⎬ Group II
                                                          DmGluRA ⎦
                                                          mGluR4  ⎤
                                                          mGluR8  ⎬ Group III
                                                          mGluR7  ⎦
                                                          mGluR6  ⎤
                                                          mgl-1   ⎬ Group II/III like
                                                          mgl-3   ⎦
```
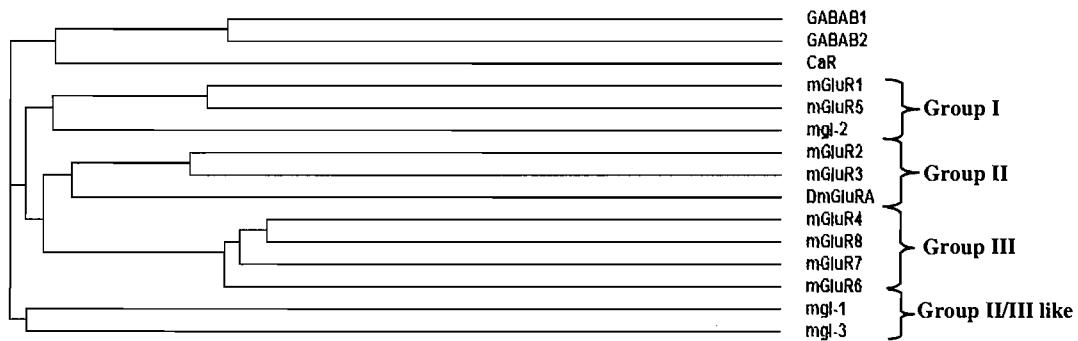
Figure 3.1. A cladogram of the sequence homology between mammalian mGluRs and the *C.* *elegans* mgls. The amino acid sequence of the three *C. elegans* *mgls* were alligned with the mammalian (rat) mGluRs and the *Drosophila* mGluR, DmGluRA. The amino acid sequence of the other Group III GPCRs were also included. These are mammalian (rat) GABA receptor subtypes GABAB$_1$ and GABAB$_2$ and the calcium sensing receptor (also from rat) . The *C. elegans* *mgl* subtypes can be classified according to their mammlian counterparts. MGL-1 and MGL-3 are Group II/III-like receptors and MGL-2 is a Group I-like receptor.

3.1.3 *In silico* analysis of the *mgl-1*, *mgl-2* and *mgl-3* predicted protein sequence

To gain insight into the molecular architecture of the *C. elegans* *mgl* receptors we have utilised various bio-informatic tools and this has shown that the *mgls* have conserved structural features and motifs characteristic of metabotropic glutamate receptors and that are important for function. However, this approach has also identified poorly conserved features and highlights that the predicted transcripts are likely to be incorrect in these poorly conserved regions of the translated sequence.

Hydropathy plots were performed to define the topology of the *mgls* and both *mgl-1* and *mgl-3* have the archetypal mGluR structure, consisting of a large extracellular N-terminal domain, a seven transmembrane domain and an intracellular C-terminal. However, hydropathy plots predict the *mgl-2* protein contains an additional 8th transmembrane domain in the C-terminal *(Figure 3.2)*, introducing a 5th intracellular loop and giving rise to a receptor with an extracellular C-terminal.

Alignment between the predicted amino acid sequence of the *mgls* and the mammalian mGluR subtypes were used to identify conserved residues, shown to be important to mGluR function. The N-terminal domain of each *mgl* contains

100

conserved residues for glutamate binding (Tyr236, Thr188, Ser165, Asp208 and Arg78. (Sato, T et al 2003)). This suggests each *mgl* has the capacity to bind glutamate as a ligand. Intermolecular interactions are also likely to be conserved. Cysteine residues in the extracellular region, with a defined role in mGluR homodimerization are conserved. The predicted amino acid sequence of the *mgl-1* and *mgl-3* receptors predicts both C-terminals contains a Type I PDZ ligand motif, -TFL, indicating that this domain is capable of instigating protein-protein interactions.

The predicted amino acid sequence of the *mgls* was also analysed for elements important for the receptors targeting. The signal peptide is a stretch of amino acids located at the N-terminal end of a polypeptide chain. It functions as a sorting signal to direct proteins to specific compartments of the cell. As the *mgls* are likely to be cell surface receptors, they would be initially directed to the ER and sorted from there to the cell surface. The ER signal peptide is typically rich in hydrophobic residues and 20-30 amino acids in length. *In silico* analysis identifies *mgl-2* and *mgl-3* possess conventional signal peptides. Whereas the predicted signal peptide of *mgl-1* is less conventional, it is unusually long, ~60 residues and the density of hydrophobic residues it contains is lower than would be expected *(figure. 3.2)*.

The current gene models for *mgl-1, mgl-2* and *mgl-3* are based upon a combination of EST data and *in silico* predictions. The transcripts of the *mgl* subtypes have yet to be fully characterised and as highlighted analysis of the predicted protein sequence has identified descrepancies that may represent predictive errors. Defining the gene model and the transcript of the gene is therefore important since the production of downstream molecular reagents (for example expression constructs, *see section 3.9.2*) will be dependent upon an accurate protein sequence.

3.2 Alignment of *C. elegans mgl* predicted amino acid sequence with the rat mGluRs identifies conserved residues important for function

```
mGluR2.   ------------------------------------MESLLGFLALLLLWGA------------VAEGPAKKVLTLEGDLVLGGLFPVHQ---KGGP-AEECGPVNEHRGIQ  61
MGL-1     *                                                  EAVDLAPPPKVRQIRIPGDILIGGVFPVHS---KSLNGDEPCGEIAETRGVH  117
mGluR4a.  ----------------------MSGKGGWAWWWARLPLCLLLSLYAPWVPSSLG------KPKGHPHMNSIRIDGDITLGGLFPVHQ---RGS-EGKACGELKKEKGIH  78
mGluR8a.  ------------------------------MVCBG---KRLASCPCFFLLTAKFYWILTMMQ--------RTHSQEYAHSIRVDGDIILGGLFPVHA---KGE-RGVPCGELKKEKGIH  75
mGluR7a.  ------------------------------MVQLGKLLRVLTLMKFPCCVLEVLLCVLAAAA--------RGQEMYAPHSIRIEGDVTLGGLFPVHA---KGP-SGVPCGDIKRENGIH  78
mGluR6.   ------------------------------MGRLPVLLLWLAWWLS--------QAGIACGAGSVRLAGGLTLGGLFPVHA---RGA-AGRACGALKKEQGVH  62
MGL-3     ------------------------MTCVSFHTPPLLRRYSSILQNLFFFLLHIP---------ISST-VKQLTVPGQIVLGGLFPIHE---AGRNASHQCGKIKALQGVQ  75
mGluR1a.  ------------------------MVRLLLIFFPMIFLEMSILPRMPDRK------VLLAGASSQRSVARMDGDVIIGALFSVHHQPPAEKVPERKCGEIREQYGIQ  78
MGL-2     ------------------MLYSHVTLLMLRVVDVVAKT---------VQQSARMLVAEIHGEIQIGALFPIHR----QIAGSESCGEIWEQYGIQ  66
                                                            : ^ :.:*.:*.:* ^^ :. *.::*
```

```
mGluR2.   LEAMLFALDRINRDPHLLFGVRLGAHILDSCSKDTHALEQALDFVRASLSRG--ADGSRHICPDGS--YATHSDAPTAVTGVIGGSYSDVS1QVANLLRLFQIPQISYASTSAKLSDKSR  177
MGL-1     VEAMLYALDQINSQNDFLRGYKLGALILDSCSNPAYALNQSLDFVRDMIGS---SEASDYVCLDGSDPNLKKQSQKKNVAAVVGGSYSSVSVQLANLLRLFRIAQVSPASTNADLSDKNR  234
mGluR4a.  LEAMLFALDRINNDPDLLPNITLGARILDTCSRDTHALEQSLTFVRALIEK----DGTEVRCGSGGPPIITKP---ERVVGVIGASGSVSIMVANILRLFKIPQISYASTAPDLSDNSR  191
mGluR8a.  LEAMLYAIDQINKDPDLLSNITLGVRILDTCSRDTYALEQSLTFVQALIEK----DASDVKCANGDPPIFTKP---DKISGVIGAALSVSIMVANILRLFKIPQISYASTAPELSDNTR  188
mGluR7a.  LEAMLYALDQINSDPNLLPNVTLGARILDTCSRDTYALEQSLTFVQALIQK----DTSDVRCTNGEPPVFVKP---EKVVGVIGASGSVSIMVANVLRLFAIPQISYASTAPELSDDRR  191
mGluR6.   LEAMLYALDRVNADPELLPGVRLGARLLDTCSRDTYALEQALSFVQALIRGRGDGDEASVRCFGGVPPLRSAPP--ERVVAVVGASASVSIMVANVLRLFAIPQISYASTAPELSDSTR  180
MGL-3     MVAMLFALEKVNRDRQLLPQASLGAQILDTCSVDSYALEQTLEFIKSVMSN-----GDGVTCADGSTGSYTRQ----PVVAVVGAAGQVSVHVASMLQLFKIPQVSYSSTGAELSEKPR  186
mGluR1a.  VEAMFHTLDKINADPVLLPNITLGSEIRDSCWHSSVALEQSIEFIRDSLISIRDEKDGLNRCLPDGQTLPPGRT-KKPIAGVIGPGSSSVAIQVQNLLQLFDIPQIAYSASIDLSDKTL  197
MGL-2     SEIAMLTVKQLNEELPFK----LGLSIRDSCWTERIAMEQTIAFLREGVAQ---------CSCCQTPGCNKKS--VPVVAVIGPGKSSTIAVQNLLQVFRIPQVGYSATPDLSDKEQ  170
          ::::.::^*:     ** ^:** ^:: *^:*  *::: *:^ :   ^ *^         :.*^ .*. :.::*:^ *.^:. :^*   ^^*:.
```

```
mGluR2.   YDYFARTVPPDFFQAKAMAEILR-----------------FFNWTYVSTVASEGDYGETGIEAFELEARARN--ICVATSEKVG--RAMSRAAFEGVVRALLQKP--SARVAVLFTRSEDA  275
MGL-1     FEYFARTVPSEDYQAMAMVEIAV-----------------KFKWSYVSLVYSADERGELGADAFKKEARKKG--ICIALEERIQNKKESFTESINNLVQKLQPEKNVGATVVVLFVGTEYI  336
mGluR4a.  YDFFSRVVPPEDTYQAMAMVDIVR---------------ALKWNYVSTLASEGSYGESGVEAFIQKSRENGG-VCIAQSVKIP--REPKTGEFDKIIKRLLETS--NARGIIIFANEDDI  290
mGluR8a.  YDFFSRVVPPESYQAQAMVDIVT---------------ALGWNYVSTLASEGNYGESGVEAFTQISREIGG-VCIAQSQKIP--REPRPGEFEKIIKRLLETP--NARAVIMFANEDDI  287
mGluR7a.  YDFFSRVVPPESYQAQAMVDIVR---------------ALGWNYVSTLASEGSYGEKGVESFTQISKEAGG-LCIAQSVRIPQERKDRTIDFDRIIKQLLDTP--NSRAVVIFANEDDI  292
mGluR6.   YDFFSRVVPPESYQAQAMVDIVR---------------ALGWNYVSTLASEGNYGESGVEAFVQISREAGG-VCIAQSIKIP--REPKPGEFHKVIRRLMETP--NARGIIIFANEDDI  279
MGL-3     FAFFSRVVPPENLQAQVMARVVCFSSLFKSSWHEIVQIGALEWTYVHAIADTGSYGERGMDSFRAAAENG--ICIDGDVQKIS-RRWTEKNFRDLLIRMHRTR--KARGVVMFVDEDNL  301
mGluR1a.  YKYFLRVVPSETLQARAMLDIVK---------------RYNWTYVSAVHTEGNYGESGMDAFKELAAQEG--LCIAHSDKIYS--NAGEKSFDRLLRKLRERLP-KARVVVCFCEGMTV  296
MGL-2     FGYFLRVVPSEVFQAQAINRLLH---------------HYNMTYVAVLYSAGNYGEKGFESLEKLIAHRSSSVCIAYSEKIKT--LASEQEYRQVLTRLDSQNS-RPQVVVCFCEGASM  271
          : :: * *.^*^.* **  ::   :            ^.*^  :  :****  *:::  ::^    :.:^      *:^      :: :  .  ^.:
```

```
mGluR2.   RELLAATQR--------LNASFTWVASDGWG-ALESVVAGSER-AAEGAITIELASYPISDFASYFQSLDPWNNS--RNPWFREFWEERFHCSPRQ----------------RDCAAHS  366
MGL-1     PDILRYTAERMKLT-SGAKKRIIWLASESWDRNNDKYTAGDNRLAAQGAIVLMLASQKVPSFEEYFMSLHPGTEAFERNKWLRELWQVKYKCEFDTPPG-----------STASRCEDIK  444
mGluR4a.  RRVLEAARRANQTG------HFFWMGSDSWGSKSAPVLRLEEV--AEGAVTILPKRMSVRGFDRYFSSRTLDNNR--RNVWFAEFWEDNFHCKLSRHAL--------KKGSHIKKCTNRE  392
mGluR8a.  RRILEAAKKLNQSG------HFLWIGSDSWGSKINPVYQQEEI--AEGAVTILPKRASIDGFDRYFRSRTLANNR--RNVWFAEFWEENFGCKLGSHG--------KRNSHIKKCTGLE  388
mGluR7a.  KQILAAAKKRADQVG------HFLWVGSDSWGSKINPLHQHEDI--AEGAITIQPKRATVEGFDAYFTSRTLENNR--RNVWFAEYWEENFNCKLTISGS--------KKEDTDRKCTGQE  394
mGluR6.   RRVLEATRQANLTG------HFLWVGSDSWGSKISPILNLEEE--AVGAITILPKRASIDGFDQYFMTRSLENNR--RNVWFAEFWEENFNCKLTSSGG-------QSDDSTRKCTGEE  381
MGL-3     KRLLKTLDLLVAEGHTELDRHFWFVASDSWGIKQSVVRGLEHR--TYGAITIAPMVREETGYLEYFRSLSPKGFV--FLEEFFEYLGCSATVDVKTFGDCFDMVNITLKQVYIHRSRTII  417
mGluR1a.  RGLLSAMRRLGVVG------EFSLIGSDGWADRDEVIEGYEVE--ANGGITIKLQSPEVRSFDDYFLKLRLDTNT--RNPWFPEFWQHRFQCRLPGHLLE---------NPNFKKVCTGNE  398
MGL-2     RMFFKAQKHLADGK--MQMKRFQWIGSDGWADRNDVVEDLEEE--AEGSFSIRIHAPKIPGFRQYYTALHPENNT--MNPWFREFWQQKFNCQFAVSKED--------KNNENIRICSGDE  378
          .:             ::   *^:*^ ^    :^ ^*.: ::  ^..  *^            :  *^     *    ^
```

```
mGluR2.   LRAVP--FEQESAIMFVVNAVYAMAHALHNMHRALCPN------------TTHLCDAMRPVNGRRLYKDFVLNVKFD-------APFRPADTDDEV----RFDRFGDGIGRYNIFTYLR  460
MGL-1     QSTEG--FNADLDVQFVIDAVYAIAHGLQSMKQAICPDDAIENHWISRYSKQPEICHAMQNIDGSDFYQNYLLKVNFTGKTISIFSSFRLSPFSDIVGKRFRFSPQGDGPASYTILTYKP  562
mGluR4a.  RIGQDSAYEQEGKVQFVIDAVYAMAHGLQNMHHDRLCP------------GRVGLCPRMDPVDGTQLLK-YIRNVNFS--------------GIAGNPVTFNENGDAPGRYDIFQYQI  482
mGluR8a.  RIARDSSYEQEGKVQFVIDAVSMAYALHNMHKERCP------------GYIGLCPRMVTIDGKELLG-YIRAVNFN--------------GSAGTPVTFNENGDAPGRYDIFQYQT  478
mGluR7a.  RIGKDSNYEQEGKVQFVIDAVYSMAYALHNMNMKDLCA------------DYRGVCPEMEQAGGKKLLK-YIRHVNFN--------------GSAGTPVMFNKNGDAPGRYDIFQYQT  484
mGluR6.   RIGQDSAYEQEGKVQFVIDAVYAIAHALHSMHQALCP------------GHTGLCPAMEPTDGRTLLH-YIRAVRFN--------------GSAGTPVMFNENGDAPGRYDIFQYQA  471
MGL-3     FINVLKEFLQESYVPFVVDTVKIIAKAISMYIEDDCGK-----------IPFHKCTLAQSGFRGERLQR-YYRNMSLI--------------KNE--PALIDANGDGIGRYDVFQLDI  507
mGluR1a.  SLEEN----YVQDSZMGFVINAIYAMAHGLQNMHHALCPG------------HVGLCDAMKPIDGRKLLD-FLIKSSFV--------------GVSGEEVWFDEKGDGPRDYDIMNLQY  486
MGL-2     --NLDEQYKEDFELSQVINSIRVVALGLKAMYQDRCRD------------NSTLCTEMLSRNG-TLLYEYLLNVTYS---------------DQFKQPVYPDRNGDPPAWYDILNYIG  466
          ^       *:::.     :^       :^ ^  :.     :           ^.  *^            :  ^      ^
```

```
mGluR2.   AGSGRYR-----------YQKVGYWA--EG-LTLDTSFIPWASPSAG-PLPASRCSEPCLQNEVKSVQPGEV-CCWLCIPCQPYEYRLDEF-TCADCGLGYWPNASLTGCFELPQ----EY  560
MGL-1     KSMDKKRRMTDDESSPSDYVEIGHWS--ENNLTIYEKNLWWDPDHT----PVSVCSLPCK IGFRKQLIKDEQ-CCWACSKCEDYEYLINET-HCVGCEQGWWPTKDRKGCFDLSLSQLKY  674
mGluR4a.  RNGS------------AEYKVIGSWT----DHLHLRIERMQWPGSGQQ--LPRSICSLPCQPGERKKTVKGMA-CCWHCEPCTGYQYQVDRY-TCKTCPYDMRPTENRTSCQPIPIVK---  580
mGluR8a.  NNKS-----------TEYKIIGHWT----NQLHLKVEDMQWANREHT--HPASVCSLPCKPGERRKTVKGVP-CCWHCERCEGYNYQVDEL-SCELCPLDQRPNINRTGCQRIPIIK---  576
mGluR7a.  TNTTN----------PGYRLIGQWT---DELQLNIEDMQWGKGVRE--IPSSVCTLPCKPGQRKKTQKGTP-CCWTCEPCDGYQYQFDEM-TCQHCPYDQRPNENRTGCQNIPIIK---  583
mGluR6.   TNGSASS---------GGYQAVGQWA---EALRLDMEVLRWSGDPHE--VPPSQCSLPCGPGERKKMVKGVP-CCWHCEACDGYRFQVDEF-TCEACPGDMRPTPNHTGCRPTPVVR---  572
MGL-3     N-----------GVYQKVGKWRSTDDFLSVEVEKIRHAFKTAHGERPMSVCSTDCPRG-HYRAYQDQT-CCWACIPCDTSTSIHNET-SCEECAVGMVPDRTLHFCVPIPPVS---  606
mGluR1a.  TEANR---------YDYVHVGTWH--EGVLNIDDYKIQMNKSGMV----RSVCSEFCLKGQRKVIRKGEVSCCWICTACKENEFVQDEF-TCRACDLGWWPNAELTGCEPIPVRY---  585
MGL-2     TKDLDNPYN-----EVGSFKSINDYG--VEELDMTAKSMFFDKTELL---PESVCSRPCGIGQRQRE---TMACCWICESCLDIQYVNKTTNQCMNCTLGSWPNANRTGCEYIIPEV---  570
          ^ ^         ^.  *^ :    .     ^:^ .: .  ::  ^     * .* ^  *   :  ^     *.**:^..***: *:: .  * :::. :
```

```
mGluR2.   IRWGDAWAVGPVTIACLGALATLFVLGVFVRHNATPVVKASGRELCYILLGGVFLCYCMTFVFIAKPSTAVCTLRRIGLGTAFSVCYSALLTKTNRIARIFGGARE--GAQRPRFISPAS  678
MGL-1     MRWRSMYSLVPTILAVFGIIATLFVIVVYVINETPVVKASGRELSYILLISMIMCYCMTFVLLSKPSAIVCAIKRTGIGFAFSCLYSAMFVKTNRIFRIFS---TR--SAQRPRFISPIS  790
mGluR4a.  LEWHSPWAVLPLFVAVGIAATLFVVVTFVRYNDTPVVRASGRELSYVLLAGIFLCYATTFLMIAEPDLGTCSLRRIFLGLGMSISYAALLTKTNRIYRIFEQGKK--SVSAPRFISPAS  698
mGluR8a.  LEWHSPWAVVPVFIAILGIIATTFVIVTFVRYNDTPIVRASGRELSYVLLTGIFLCYSITFLMIAAPDTIICSFRRIFLGLGMCFSYAALLTKTNRIHRIFEQGKK--SVTAPKFISPAS  694
mGluR7a.  LEWHSPWAVIPVFLAMLGIIATIFVMATFIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRVFLGLGMCISYAALLTKTNRIYRIFEQGKK--SVTAPRLISPTS  701
mGluR6.   LTWSSPWAALPLLLAVLGIMATTIMATFMRHNDTPIVRASGRELSYVLLTGIFLYAITFLMVAEPCAAICAARRLLLGLGTTLSYSALLTKTNRIYRIFEQGKR--SVTPPPFISPTS  690
MGL-3     MQWDTTWSLIPAAFSTLGIASTIFVVSVFLKFSNTPVIMASGRELCYCMMSGIGMCYTLTFFLVSQPTVITCSMTRILMGLSMSAIYIAAITITKTNRLARVFKPD-----SAQRPRFITPKA  722
mGluR1a.  LEWSDIESIIAIAFSCLGILVTLFVTLIFVLYRDTPVVKSSSRELCYIILAGIFLGYVCPFTLIAKPTTTSCYLQRLLVGLSSAMCYSALVTKTNRIILAGSKKKICTRKPRFMSAWA  705
MGL-2     VSWTSFGHILALVLAVTGIITSMATLAVFLRHNSTPVVKSTTRELSYIILSGLVACYAVSFALLATPSTTSCFITRVIPPIAFAVVYSALLTKTNRIARILAGSKKRILTKKPRFLTTFS  690
          : *   .    . :   :: . **:: ^: ***.^ ^  ^ .:* .^^ ^:  ^.. *.:*^.****:.*:: .  ^ :::. :
```

```
mGluR2.   QVAICLALISGQLLIVAAWLVVEAPGTGKETAP-----ERREVVTLRCNHRDASMLGSLAYNVLLIALCTLYAFKTRKCPENFNEAKFIGFTM------------------YTTCIIWL  774
MGL-1     QVVMTAMLAGVQLIGSLIWLSVVPPGWRHHYP-------TRDQVVLTCNHRPHFLYSLAYDGFLIVLCCTTRKVVPENFNETKPIGFSM--------------TRDQVVLTCNHRPHFLYSLAYDGFLIVLCCTTRKVVPENFNETKPIGFSM------------------YTTCVVWL  884
mGluR4a.  QLAITFILISLQLLGICVWFVVDPSHSVVDFQDQRTLDPRFARGVLKCDISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTM------------------YTTCIVWL  799
mGluR8a.  QLVITFSLISVQLLGVFVWFVVDPPHTIIDYGEQRTLDPENARGVLKCDISDLSLICSLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTM------------------YTTCIIWL  795
mGluR7a.  QLAITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNPEQARGVLKCDITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTM------------------YTTCIVWL  802
mGluR6.   QLVITPGLTSLQVVGVTAWLGAQPPHSVIDYEEQRTVTDPEQARGVLKCDMSDLSLIGCLGYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTM------------------YTTCIIWL  791
MGL-3     QVGICMGIVVSLQLVIGTFVW1LFDPPGTMIVF-------PTRTEAVLTCKATTSHLLISLLYNILLIVACTVYAFKTRKIPENFNETRHIGFTM--------------------YSTCILWL  816
mGluR1a.  QVIIASILISVQLTLVVTLIIMEPPMPILSYP-------SIKEVYLICNTSNLGVVAPVGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTM------------------YTTCIIWL  799
MGL-2     QVVITWILVAVQCVIVGVGLMRDWPDATYAKY------ALPRKLILECDTETKSFLIPFFWDFFLITLCTLYAFKTRNLPENFNEAKFIGFTISQPAKSLPFFQKEHVIFQYCTVVVWI  804
          ^: ^ ^     . :: ^     ^:    ^    :   .^^ :  ^^  * *   *.^  ^ :: *.***.^ ^  .^^ ^  ^  ^ ^.::^
```

```
mGluR2.   AFLPIFYVTSS---DYRVQTTTMCVSVSLSGSVVLGCLFAPKLHIILFQPQKNVV------------------------------SHRAP---------------------  831
MGL-1     SWIFFFFGTGS--DFQIQTSSLCISISHSANVALACIFSPKLWIILFEKHKNVR-----------------------------KQEGES-------------------------  942
mGluR4a.  AFIPIFFGTSQSADKLYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPEQNVP----------------------------KRK-------------------------  857
mGluR8a.  AFIPIFFGTAQSAEKMYIQTTTLTVSMSLSASVSLGMLYMPKVYIIFHPEQNVQ-----------------------------KRK-------------------------  853
mGluR7a.  AFIPIFFGTAQSAEKLYIQTTTLTIISMNLSASVALGMLYMPKVYIIIFHPELNVQ----------------------------KRK-------------------------  860
mGluR6.   AFVPIYFATQS---DFRIQISLCMCISLSGTVALICFFAPKVYIIVLFQPYKNVR---------------------------KRK-------------------------  849
MGL-3     AFGPIYFATGS---DFRIQITSLCMCISLSGTVALICPFPMKMYIIIIAKPERNVRSAFTTSDVVRMHVGDGKLPCRSNTFLNIFRRKKPGAGNANSNGKSVSWSEPGGRQAPKGQHVWQR  904
mGluR1a.  AFVPIYFGSNY------KIITTCFAVSLSVTVALGCMFTPKMYIIIAKPERNVRSAFTTSDVVRMHVGDGKLPCRSNTFLNIFRRKKPGAGNANSNGKSVSWSEPGGRQAPKGQHVWQR  912
MGL-2     AFLVLHMGTTH------KALVMSFSYSLSASVALALLPFPKLYIILMHPEKNIRASYTITKLIRCHFGNSQAAYDSTSKQQHLGSKTTARTSVQSGSASKSSSMGGGVTRTASVHVPVS  917
          ^^:^ :.^ ^                  .  ..:* .^ * ^^  ::^^ ^:       ^.^
```

```
mGluR2.     ----TSR-FGSAAPRASAN-LGQGS-----------------------------------GSQFVPTVCN----GREVVDSTT[RR]------------------------------ 872
MGL-1       MLNKSSRSLGNCSSRLCANSIDEPN-----------------------------------QYTALLTDSTRRRSSRKTSQPTST[SSAHD][RY]---------------------------- 999
mGluR4a.    RSLKAVVTAATMSNKFTQKGNFRPN-----------------------------------GEAKSELCENLETPALATKQTYVTYTNHAI----------------------------- 912
mGluR8a.    RSFKAVVTAATMQSKLIQKGNDRPN-----------------------------------GEVKSELCESLETNTSSTKTTYISYSNHSI----------------------------- 908
mGluR7a.    RSFKAVVTAATMSSRLSHKPSDRPN-----------------------------------GEAKTELCENVDPNSPAAKKKYVSYNN[IVI]----------------------------- 915
mGluR6.     RSLKKTSTMAAPP---------------------------------------------QNENAEDAK------------------------------------------------- 871
MGL-3       QSYKPSTEESSHTSNAQVQQPTRAIPPHVIEQLAASLPISDKNDLVKKLSLQDKFFNNNNNNENSEEIRRRRPSSVHTVGAVTNGSSVAHIPPRSYTDEPKSSMIRQDTAKSRTSLAESH 1024
mGluR1a.    LSVHVKTNETACNQTAVIKPLTKSYQGSGKSLTFSDASTKTLYNVEEEDNTPSAHFSPPSSPSMVVHRRGPPVATTPPLPPHLTAEETPLFLADSVIPKGLPPPLPQQQPQQPPPQQPPQ 1032
MGL-2       RGSTHSTDVSTQTEAASKFSRSFSIVGRKKQGLDDDVQQLVDACRRYQDEKINSSAANLLLEESEDEVGALLADSIENSMRTVLSTVAGKAVVPLVPMVPMIPVVTLPTAPSQEDNFEQA 1037
                                            ▲[GTLRSIFPLIMFKIVFLSNFDILG]

mGluR2.     -----------------------------------------------------------------
MGL-1       -----------------------------------------------------------------
mGluR4a.    -----------------------------------------------------------------
mGluR8a.    -----------------------------------------------------------------
mGluR7a.    -----------------------------------------------------------------
mGluR6.     -----------------------------------------------------------------
MGL-3       QVDLILEEIAAD[TNS][TFT]--------------------------------------------------
mGluR1a.    QPKSLMDQLQGVVTNFGSGIPDFHAVLAGPGTPGNSLRSLYPPPPPPQHLQMLPLHLSTFQEESISPPGEDIDDDSERFKLLQEFVYEREGNTEEDELEEEDLPTASKLTPEDSPAL[RI] 1152
MGL-2       GFKEKCQLSKDLSRKSPKQFLHVAPKSRCSTSRAVSSYTIMSGNQQLMSPMGSVSSSAGSGSPSPMSSFDHLSDEELAQISVRQLNQKLMGQDRNVVMQWKQKRRTLKNRGYALNCRARR 1157

mGluR2.     ----------------------------------------------------
MGL-1       ----------------------------------------------------
mGluR4a.    ----------------------------------------------------
mGluR8a.    ----------------------------------------------------
mGluR7a.    ----------------------------------------------------
mGluR6.     ----------------------------------------------------
MGL-3       ----------------------------------------------------
mGluR1a.    [PGPPTL]SVASGSSVPSSPVSESVLCTPPNVTYASVILRDYKQSS[STL]------------------------------
MGL-2       VNNQVQLEADNMMLRNQIKTLREALSEAQMRLHYYEPVFYQAYPSIVPSTTSPVTVSLPIIARPIPVPPDSTATVTKFATDFKQIF
```

Figure 3.2. Alignment of *C. elegans mgl* receptors with rat mGluRs (ClustalW). MGL protein sequence is taken from wormbase version WS160 (31 July 2006). Sequence alignment reveals conserved residues identified from crystal structure analysis of mGluR1a (Kunishima, N et al 2000), that are involved in binding glutamate. Transmembranous domains (predicted using HMMTOP : http://www.enzim.hu/hmmtop/) are highlighted in grey. The 8th transmembrane domain of MGL-2 is highlighted in pink, this has recently been amended in the database, which is why the sequence was not included in the alignment. Signal peptide sequences (predicted using SignalP 3.0 Server) are highlighted in yellow, note the length of the MGL-1 signal peptide is considerably longer than the other *mgls* and mGluRs and it does not have the characteristic motif features of a signal peptide sequence. Type I PDZ ligand motifs *(see table 1.4)* are highlighted in blue; note *mgl-2* is the only subtype whose C-terminal lacks a PDZ ligand motif. Where the MGL protein sequence is underlined this indicates the location of the primers used to amplify the intracellular C-terminal. The Homer binding motif of the mGluR1a C-terminal is highlighted in dark red, note how this motif is not conserved in the C-terminal of the *C. elegans* MGL-2 receptor.

RESULTS

3.3 Predictive PCR has identified that alternative splicing generates variation in the C-terminal of the *mgls*

The identification of protein-protein interaction motifs within the C-terminal domain of *mgl* subtypes predicts that it is likely to perform an important role in the receptors function. In view of this we set about using molecular techniques to clone the C-terminal of the *mgl* subtypes for use in functional studies and to clarify discrepancies highlighted from the predicted amino acid sequence. Primers were generated using the database (www.wormbase.org) predicted sequence *(figure 3.2)* to PCR amplify the *mgls* C-terminal domain from *C. elegans* cDNA. In all cases the antisense primer was designed against the predicted stop codon, which was taken as the end of the C-terminal domain and sense primers were designed immediately after the seventh transmembrane domain *(figure 3.2)*. PCR products were subcloned into the yeast-two-hybrid vector pGILDA and the pGEX vector for generating GST fusions, providing the tools for the genetic and biochemical detection of protein-protein interactors.

Analysis of PCR products by ethidium bromide electrophoresis and subsequent sequencing suggests the C-terminal of *mgl-2* and *mgl-3* subtypes are subject to splicing events *(see figures 3.3)*. PCRs were performed using two independent sources of template, which were the LexA cDNA Library (Origene) and random primed cDNA (O'Connor, V). A single band, of the correct size, corresponding to the *mgl-1* C-terminal was amplified from both templates. This was cloned and sequenced and it confirmed the sequence of the *in silico* gene model in the region corresponding to the intracellular C-terminal *(see alignment in section 3.2 for the MGL-1 C-terminal protein sequence)*

In the case of *mgl-2* and *mgl-3* multiple bands were amplified from the two independent templates *(figure 3.3)*. The C-terminal cDNA fragments were cloned and sequenced and comparison of the sequence data to the available *C. elegans* genome sequence established the multiple products amplified were generated by alternative splicing of the C-terminal domain *(see table 3.1)*. Sequence analysis was performed using 5' and 3' sequencing primers complementary to the pGilda (5' Bait Fusion Primer and 3'Bait Fusion Primer) and pGEX vectors.
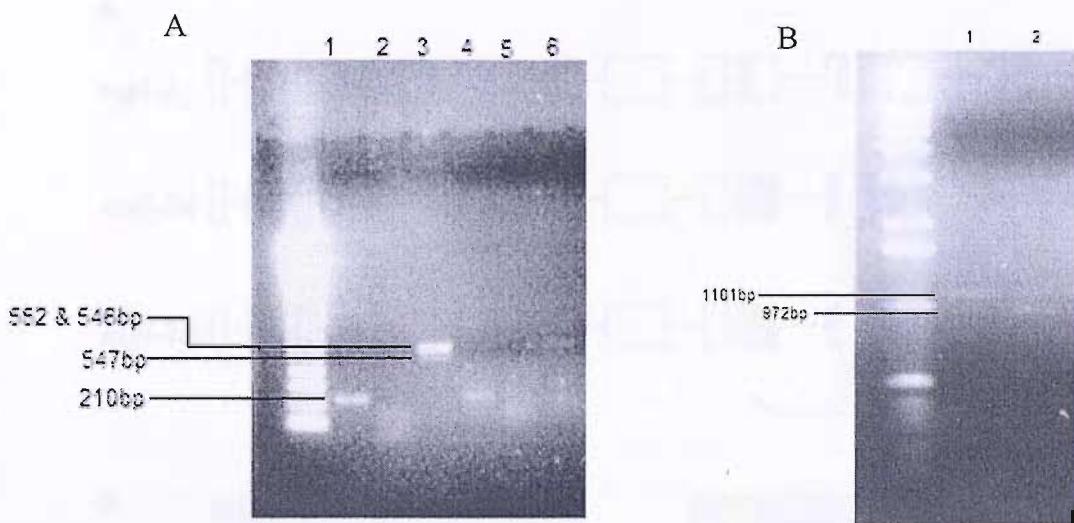
Figure 3.3. PCR amplification of *mgl-1*, *mgl-2* and *mgl-3* C-terminal domains. PCR was performed with the Platinum *Pfx* DNA Polymerase (Invitrogen) using the LexA cDNA Library (Lanes 1-3 in Figure A and Lane 1 in Figure B) and Random primed cDNA (Lanes 4-6 in Figure A and Lane 2 in Figure B) as templates. **A).** Amplification of *mgl-1* C-terminal (Lanes 1 and 4) and *mgl-3* C-terminal (Lanes 3 and 6). *Cycle*: Hot start, 94°C 2min, (94°C 15sec, 55°C 30sec, 68°C 1 min) x 30 cycles, 68°C 10 min. Under these cycling condition a product for *mgl-2* C-terminal could not be amplified (Lanes 2 and 5) **B).** Amplification of *mgl-2* C-terminal using a lower annealing temperature. *Cycle:* Hot Start, 94°C 2min, (94°C 15sec, 50°C 30sec, 68°C 2min) x 30 cycles, 68°C 2 min.

3.3.1 *mgl-2* C-terminal

cDNA analysis identified discrepancies in the C-terminal domain of the predicted *mgl-2* gene model *(figure 3.4)*. Two C-terminal variants (named *mgl-2ii* and *mgl-2iii*) were identified and both possess a common Type I PDZ ligand motif, -TPL *(figure 3.5)*. In addition the region corresponding to the predicted 8th transmembrane domain does not exist in either variant and confirms *mgl-2* conforms to the archetypal mGluR structure. The PCR analysis did not amplify a cDNA fragment corresponding to the *in silico* predicted C-terminal.
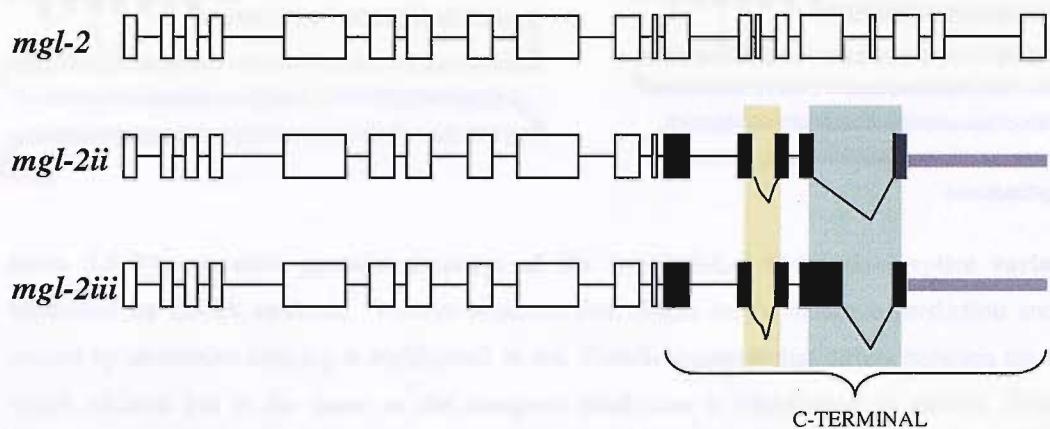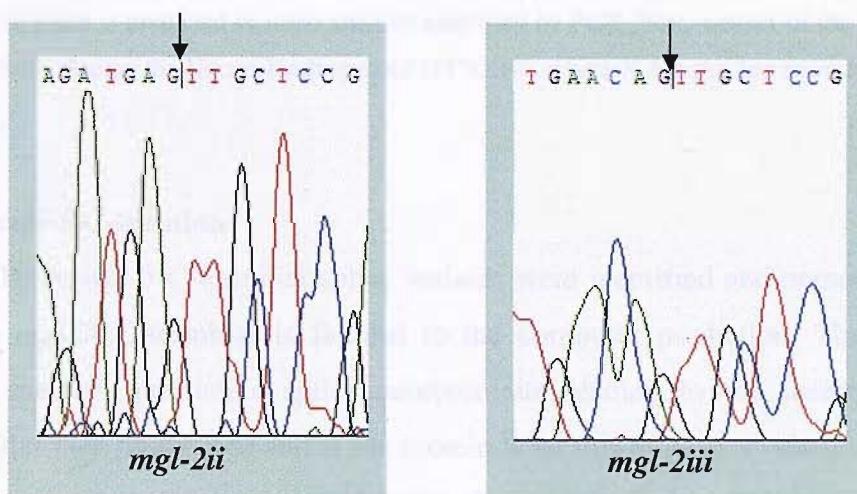
Figure 3.4 The analysis of C-terminal cDNA fragments PCR amplified identified two alternative *mgl-2* splice variants. A) The *in silico* predicted gene model is shown and denoted *mgl-2*. The gene structure established by the PCR amplification of the *mgl-2* C-terminal from cDNA is shown below. Two variants were identified, denoted *mgl-2ii* and *mgl-2iii*. The green and yellow shading corresponds to alternative gene structure to the predicted gene model. In addition the green shading corresponds to alternative splicing of the *mgl-2ii* and *mgl-2iii* variants. The grey line indicates 3' untranslated sequence **B)** Sequence to support the alternative splicing in the green shaded region is shown and the black arrow indicates the newly defined exon boundary of the *mgl-2ii* and *mgl-2iii* variants in this region.
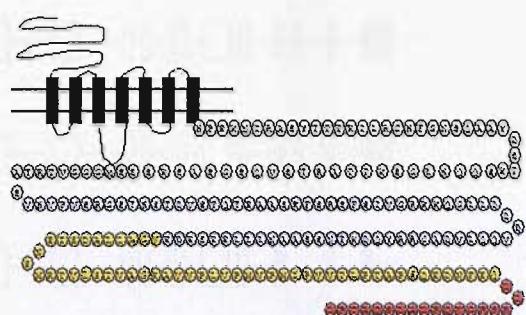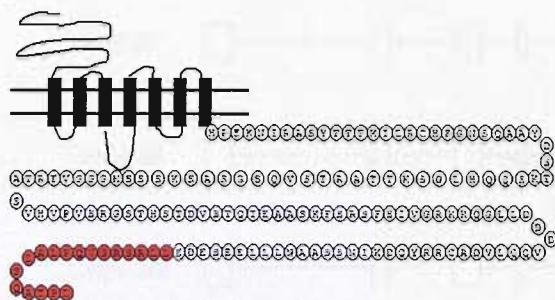
igure 3.5 The putative protein sequence of the two MGL-2 C-terminal splice variants identified by cDNA analysis. Protein sequence that differs to the database prediction and is caused by alternative splicing is highlighted in red. Protein sequence that differs between the two splice variants but is the same as the computer prediction is highlighted in yellow. Protein sequence in white is predicted *in silico* and was amplified by PCR. Note: neither of the C-terminal splice variants display the Homer binding motif (PPXXFR, where X denotes any amino acid).

3.3.2 *mgl-3* C-terminal

Three *mgl-3* C-terminal splice variants were identified and named *mgl-3i-iii*. The *mgl-3*i C-terminal is faithful to the computer prediction. The *mgl-3*ii variant uses an alternative splice acceptor site leading to the insertion of 6 nucleotides *(see figure 3.6)* and at the protein level this induces a substitution of a Threonine residue for the 3 amino acids, Isoleucine, Serine and Alanine. Both *mgl-3*i and *mgl-3*ii contain a Type I PDZ ligand motif, -TFL *(see figure 3.7)*. In the *mgl-3*iii variant, splicing generates a much shorter C-terminal. Splicing causes a downstream frame-shift which gives rise to an alternative reading frame and a premature stop codon in the C-terminal. Translation of the altered reading frame shows this C-terminal splice variant is unlike the others because it does not possess a PDZ ligand motif *(see figure 3.7)*.

107

Figure 3.6 The analysis of C-terminal cDNA fragments PCR amplified identified two alternative _mgl-3_ splice variants. A) The _in silico_ predicted C-terminal was amplified from cDNA and this is denoted _mgl-3i_. The gene structures defined by PCR amplification of the _mgl-3_ C-terminal from cDNA are shown. Two alternative variants were identified, denoted _mgl-3ii_ and _mgl-3iii_. The red line in the gene model for _mgl-3ii_ indicates a 6 nucleotide insertion. The yellow shading corresponds to the region of alternative splicing in the three variants. The grey line indicates 3' untranslated sequence **B)** Sequence to support the alternative splicing in the yellow shaded region is shown and the black arrow indicates the newly defined exon boundary of _mgl-3ii_ and _mgl-3iii_. In the case of _mgl-3i_ the black arrow indicates the _in silico_ and experimentally verified exon boundary.
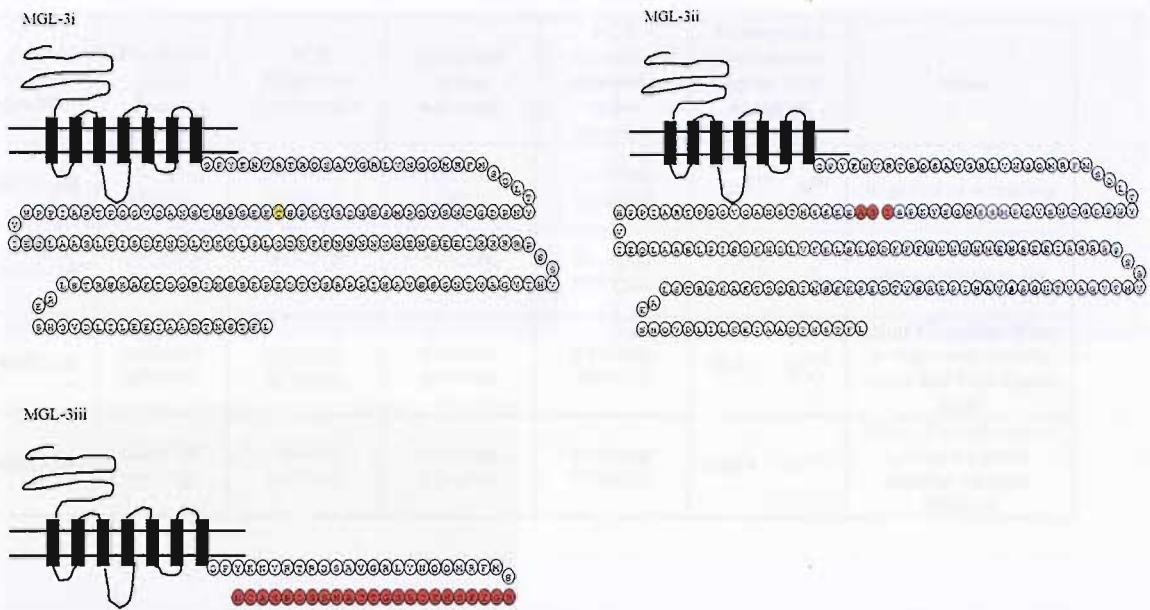
Figure 3.7 The putative protein sequence of the three MGL-3 C-terminal splice variants identified by cDNA analysis. Protein sequence that differs to the database prediction and is caused by alternative splicing is highlighted in red. Protein sequence that differs between the two splice variants but is the same as the computer prediction is highlighted in yellow. Protein sequence in white is predicted *in silico* and was amplified by PCR. Note the C-terminal of MGL-3iii does not express a PDZ ligand motif, unlike MGL-3i and MGL-3ii. Note the three amino acid substitution in the C-terminal of MGL-3ii, Iso, Ser, Ala. Serine residues in particular are targets for phosphorylation, which can modulate receptor function.

Splice Variant Identified	Predicted splice donor	PCR identified splice donor	Predicted splice acceptor	PCR identified alternative splice acceptor	Interrupted Codon and amino acid at end of exon	Notes
MGL-3iii	TATGAG gtaggt	TATGAG gtaggt	ttctag CCAATT	tttcag CAGAGG	AG*C S^{886}	Exon 15 is spliced out to give a new reading frame
MGL-3ii	CTAGTA gtgagt	CTAGTA gtgagt	tttcag CAGAGG	ataaag TTTCAG	A*TT I	In frame 3 amino acid ISA substitution for T^{911}
MGL-2ii	GAACAG gtttgt	<u>GATGAG</u> <u>gtagga</u>	ttgtag GCAGGA	tttcag TTGCTC	GAG* E^{1008}	Exon 17 is spliced out giving a new reading frame and PDZ ligand motif
MGL-2iii	GAACAG gtttgt	GAACAG gtttgt	ttgtag GCAGGA	tttcag TTGCTC	CAG* E^{1008}	Exon 17 is spliced out giving the same reading frame as *MGL*-2ii

Table 3.1 A summary of the splice sites identified by sequencing of PCR generated C-terminals. This table summarises each newly identified splice donor and acceptor site for each individual C-terminal splice variant. Only the splice donor and acceptor sites that do not correlate with the original computer predictions are shown. The splicing of *mgl-3*i corresponds exactly with the computer prediction and is therefore not included in the table. All splice variants identified use alternative splice acceptor sites to those predicted by the sequenced *C. elegans* genome and only *mgl-2*ii uses an alternative splice donor site to that predicted. Alternative splicing of *mgl-2*ii and *mgl-2*iii generates a C-terminal with predicted PDZ ligand motifs.

3.4 Characterization of the 5' and 3' ends of the *mgl* transcripts

3.4.1 Introduction

The predictive PCR analysis of the *mgls* C-terminal highlighted that the description of *mgl* transcripts by the database is likely to be incomplete and therefore further sequence data for the *mgl* cDNAs was required. Furthermore, as highlighted previously *mgl-1* has an uncharacteristic signal sequence motif and it is known that the start of genes is often poorly defined by *in silico* gene prediction programmes and so requires experimental confirmation (Hwang, B.J et al 2004). In view of this our approach to molecular characterization of the *mgl* subtypes was extended to incorporate a RACE characterization of the 5' and 3' ends. This has important relevance for the production of gene reporter constructs *(see Chapter 6)*. In parallel to the RACE characterization of the *mgl* transcripts an SL (spliced leader) based approach was adopted. In doing so this took advantage of a naturally occurring RNA processing event in *C. elegans (see section 3.4.3)* and facilitated the isolation of *mgl* 5' cDNA fragments for the characterization of important functional motifs.

3.4.2 Overview of RACE primer design

The *C. elegans* 5' and 3' RACE gene specific primers (GSP) were designed against regions of the *mgl* subtypes, *mgl-1, mgl-2* and *mgl-3* that were strongly conserved in protein sequence alignments between the *C. elegans* subtypes and mammalian subtypes. The 5' and 3' RACE GSP were designed in each case to amplify overlapping 5' and 3' RACE products *(see figure 3.8)*. In designing primers that give a 400-600bp overlap in the RACE products this provided a positive control for the PCR (this is referred to as the Internal Control in subsequent sections and figure legends); a control for the quality of the reverse transcription and would enable the joining together of RACE fragments by using restriction sites located in the overlap. The 5' RACE and 3' RACE GSP were both used separately with the adaptor primer (subsequently referred to as AP) in the first round RACE PCR.

A nested 5'RACE GSP (nGSP) was also designed for each *mgl* subtype and was used in conjunction with the nested adaptor primer (subsequently referred

to as the nAP) to perform nested PCR on the primary 5' RACE amplification performed using the 5'RACE GSP and AP. Nested 5'RACE PCR was used when the level of background was too high with the single GSP. The predicted product size range is indicated for each RACE reaction, it was established using the most current gene model available from wormbase at the time and was used as the criteria to select RACE products for sequencing.
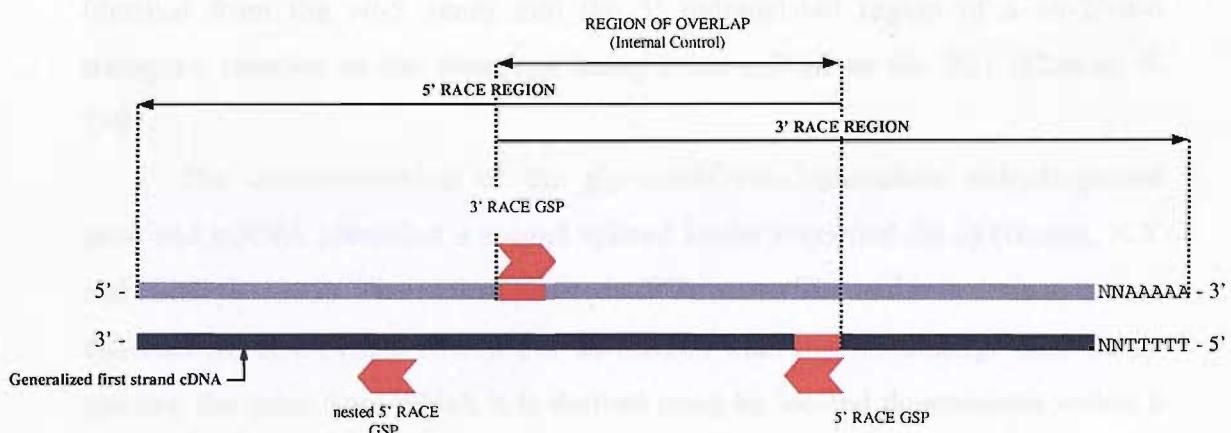
Figure 3.8 A conceptualization of RACE gene-specific primers and their relation to the first strand cDNA template. A generalized first strand cDNA synthesis is shown. First strand cDNA synthesis is performed using a modified oligo(dT) primer. The overlap between the 5' and 3' RACE GSP is highlighted in the diagram and the use of the 3'RACE GSP and 5'RACE GSP together served as a PCR control (subsequently referred to as 'Internal Control'). The location of the nested 5'RACE GSP is also indicated.

3.4.3 Overview of *C. elegans* splice leader (SL)

Spliced leader (SL) addition is a specific type of trans-splicing that occurs in *C. elegans*. It was initially identified by studies undertaken to investigate the structure and expression of actin-encoding genes (Krause, M. Hirsh, D 1987). Analysis of the transcripts from each of the four *C. elegans* actin genes revealed a 22nt 5' terminal leader sequence common to three of the four transcripts. The spliced leader (SL1) was subsequently shown to be the first 22 nucleotides of a non-polyadenylated RNA, ~100nt in length. The signals necessary for *trans*-splicing to occur are very similar (Table 3.2) to those signalling *cis*-splicing

(intron removal). The splice donor site in the spliced leader RNA and the sequence for the 5' intron splice site are highly conserved. The *trans*-splice site and 3' intron splice site sequence are the same. *C. elegans* transcripts can undergo both *trans*-splicing and *cis*-splicing. The signal for *trans*-splicing is the presence of an outron. An outron is an A and U rich, intron like sequence ending in a 3' splice site located at the 5' end of the pre-mRNA without a functional 5' splice site upstream. Indeed, the insertion of an intron without the 5' splice site (derived from the *vit*-5 gene) into the 5' untranslated region of a *vit-2/vit-6* transgene resulted in the transcript being *trans*-spliced to the SL1 (Conrad, R 1991).

The characterization of the glyceraldehyde-3-phosphate dehydrogenase gene and mRNA identified a second spliced leader sequence (SL2) (Huang, X.Y and Hirsh, D 1989). The SL2 sequence is different to SL1 and it is derived from a different SL RNA (SL2 RNA). For an mRNA transcript to undergo SL2 *trans*-splicing the gene from which it is derived must be located downstream within a cluster of closely spaced genes transcribed in the same orientation. The promoter at the 5' end of the gene cluster is responsible for controlling the transcription of the entire cluster, generating a polycistronic mRNA transcript. This is converted to monocistronic mRNA by cleavage and 3' polyadenylation of each gene. At this stage SL2 is *trans*-spliced onto the 5' end of genes located downstream. *Trans*-splice sites located close to the promoter accept only SL1. *Trans*-splice sites located in some downstream genes can accept either SL1 or SL2, whereas others will only accept SL2. The reason for this is not known.

It is estimated that mRNA transcripts of ~70% of *C. elegans* genes undergo *trans*-splicing events, with SL1 accounting for 57% and SL2 16% and SL *trans*-splicing in *C. elegans* is predicted to perform an important biological role in the maturation of mRNA transcripts. However, the precise biological function of this process is not known. It is not restricted to mRNA encoding a specific class of protein and is unlikely to be required for translation or transport of the mRNA because some mRNA transcripts are not *trans*-spliced. The trimethylguanosine cap at the 5' end of the SL sequence has been identified as regulating translation efficiency. *Trans*-spliced mRNA transcripts undergo translation with greater efficiency compared to the same transcripts possessing a monomethylguanosine cap in an *Ascaris lumbricoides in-vitro* system (Maroney et al 1995).

	-7	-6	-5	-4	-3	-2	-1	+1
	U	U	U	U	C	A	G	A/G
Cis	53	89	98	70	83	100	100	74
SL1	57	92	97	64	82	100	100	79
SL2	66	66	84	66	78	100	100	75

Table 3.2 **The percentage of** *cis* **and** *trans*-**splice sites that match the consensus 3' splice acceptor site.** (Riddle, DL *C. elegans* Book II 1997).

3.5   Assessment of total RNA quality and integrity

Trizol extraction from a 500mg mass of mixed stage worms yielded ~7mg of total RNA. The examination of the total RNA by EtBr agarose gel electrophoresis produced a smear ranging in size from 12kb-0.2kb, with prominent ribosomal RNA bands *(figure 3.9)*. The UV absorbance of the sample was measured by spectrophotometry and the ratio of readings at 260nm and 280nm was taken as an estimate of the purity of the RNA sample. 1:100 dilutions were made in DEPC treated H_2O. The total RNA gave a 260/280 ratio of 1.7, with a ratio of 1.5 being considered representative of a high purity (Okamoto, T and Okabe, S 2000). Total RNA was used as the starting material in first strand cDNA synthesis reactions for RACE and SL-1 experiments.

A

20 μg *C. elegans* Total RNA
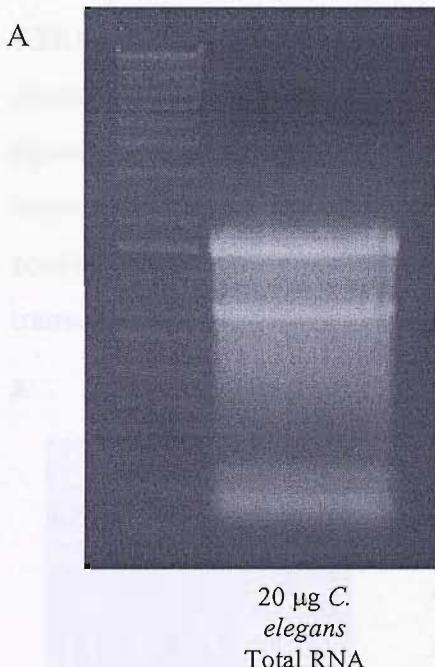
Figure 3.9 Analysis of the *C. elegans* total RNA used in SMART RACE. A. 20ug of *C. elegans* total RNA produced a smear on an agarose gel, intense bands represent ribosomal RNA.

3.6   SMART RACE system

Two RACE approaches were used involving poly A$^+$ mRNA or total RNA as the starting material. The Marathon RACE approach used poly A$^+$ mRNA as the starting material and did not work and so instead the SMART RACE approach that uses total RNA as the starting material was focused on. 3' and 5' RACE cDNA was generated from the human placental total RNA (which served as the positive(+) control RACE template) and in parallel the *C. elegans* total RNA (generated from a mixed stage population of animals) *(described in section 3.5)*. The (+)control RACE reactions were performed on the 3' RACE and 5' RACE cDNA generated from the human placental total RNA and amplify the ends of the transferrin receptor (TFR) cDNA. Both the internal control and the 3' and 5' RACE positive control reactions amplified cDNA fragments of the correct size, 0.3Kb, 2.9Kb and 2.6Kb respectively *(see figure 3.10)*. When performing the control RACE reaction the touchdown PCR cycling parameters *(see legend to figure 3.10)* in combination with the enzyme BD Advantage II were identified as improving specificity and reducing non-specific background. These reagents and conditions were subsequently adopted for the 3' and 5' RACE analysis of *mgl* transcripts.

A
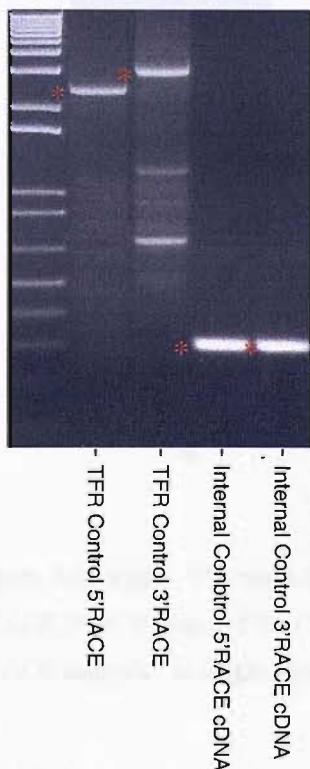
Figure 3.10 Control RACE reactions A) Amplification of the transferrin receptor 3' and 5' cDNA from the human placental control RACE cDNA. *Cycle: 1min 94°C, 5x(94°C 30s, 72°C 3min), 5x(94°C 30s, 70°C 30s, 72°C 3min), 25x(94°C 30s, 68°C 30s, 72°C 3min).* Enzyme: BD Advantage II. TFR 5'RACE product: 2.6kb. TFR 3'RACE product: 2.9kb. Internal controls were performed on both 3' and 5' RACE cDNA. In both cases the correct sized band (0.3kb) was amplified.

TFR Control 5'RACE

TFR Control 3'RACE

Internal Conbtrol 5'RACE cDNA

Internal Control 3'RACE cDNA

116

3.6.1 *mgl-1* RACE Characterization

The *mgl-1* internal control PCR amplified a prominent band in both the 3' and 5'RACE reactions that was approximately 50bp shorter than the predicted product size, suggesting the *mgl-1* gene model is ill defined in this region. The *mgl-1* 5'RACE reaction amplified multiple bands. The two largest bands amplified were within range of the predicted product size *(see figure 3.11A)*. However, other smaller bands were present and to select for *mgl-1* specific 5'RACE cDNA a nested PCR was performed on the first round RACE reaction. Several prominent bands were amplified by the nested 5'RACE PCR. The most abundant fragments amplified *(indicated by the hashed box in Figure 3.11B)* were the most similar to the expected product size and subsequently these were selected by gel extraction and were cloned and sequenced. The 1st round *mgl-1* 3'RACE PCR amplified a single band that was ~100bp larger than the predicted product size of ~1677bp *(see figure 3.11C)*.
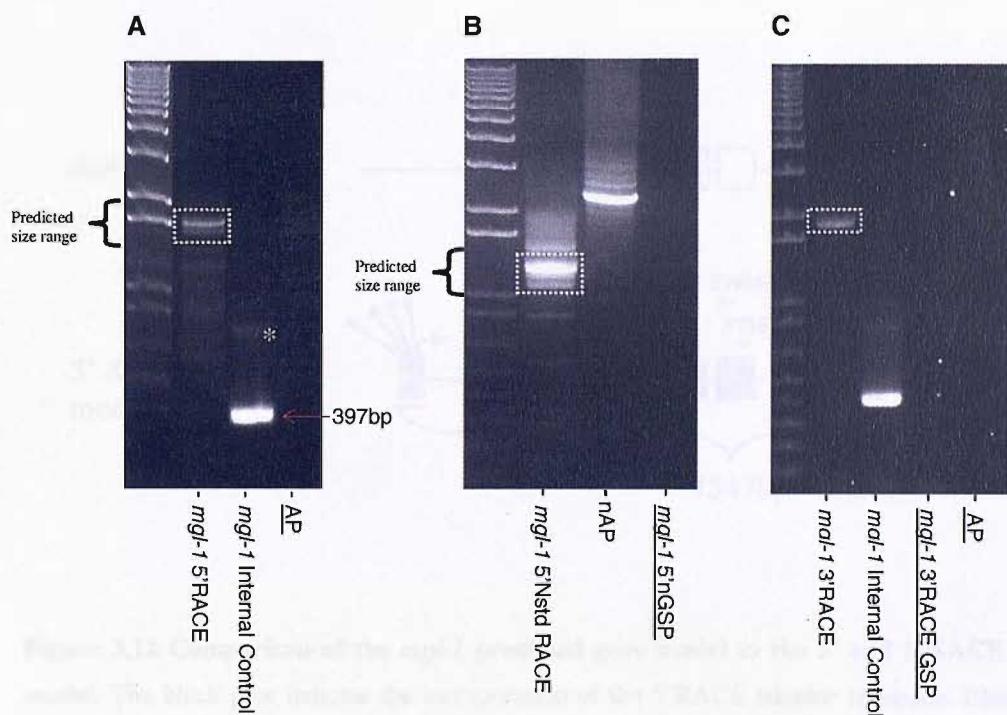
Figure 3.11 *mgl-1* **1st round 5' RACE, nested 5' RACE and 3' RACE analysis. A.** *mgl-1* 5'RACE PCR. **B.** *mgl-1* 5'RACE nested PCR performed on the 1st round amplification. **C.** *mgl-1* 3'RACE analysis. In single primer control PCR reactions the name of the primer is underlined.

Sequencing of the *mgl-1* 5'nested RACE products revealed the multiple products amplified were caused by the alternative incorporation of the SMART adaptor sequence at the 5' end of the *mgl-1* cDNA. The longest *mgl-1* nested 5'RACE cDNA sequenced was not SL-1 *trans*-spliced *(see section 3.7 for the SL-1 analysis of mgl-1)*. It identified an alternative 5' exon structure to that defined by the database predicted gene model *(shown in figure 3.12)* and identifies an alternative start codon to the database prediction.

The size difference between the predicted 3' RACE product and the PCR amplified product corresponds to the presence of a 129bp 3'UTR (27bp of which is the polyA tail) and the splicing out of a 48bp region located within the cysteine rich domain of *mgl-1* and in the internal control region of overlap. This alternative splicing occurs at exon 12 and it causes the removal of 16 amino acids (GKTISIFSSFRLSPFS) that were poorly conserved in alignments of *mgl-1* to the other *C. elegans* subtypes *mgl-2* and *mgl-3* and the mammalian mGluR subtypes *(see section 3.2)*.
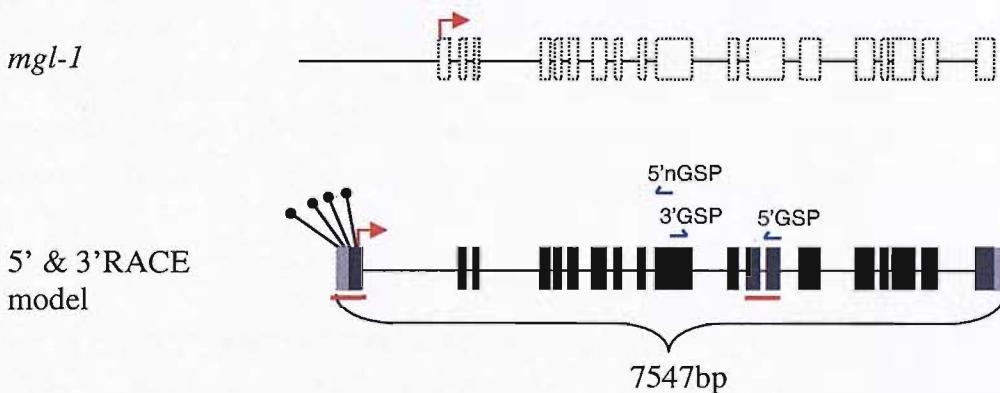
Figure 3.12 Comparison of the *mgl-1* predicted gene model to the 5' and 3'RACE defined model. The black pins indicate the incorporation of the 5'RACE adaptor sequence. Black boxes indicate confirmed exons and grey boxes indicate untranslated sequence. Red arrows represent the position of translational start codons. Dashed boxes represent exons defined by the predicted *mgl-1* gene model. Alternative exon structure to that predicted is underlined in red.

3.6.2 *mgl-2* RACE analysis

The *mgl-2* internal control amplified multiple bands from the 3' and 5'RACE cDNA, but the most prominent band amplified was very close in size to the size of the predicted internal control product. Other bands amplified by the internal control *(indicated by the asterisks in figure 3.13A)* were not seen in subsequent internal control PCRs *(data not shown)* performed on first strand cDNA synthesised with either random primers, a gene specific primer or an independent oligo(dT) primer *(see section 3.7.1)*, which suggests these alternative bands are likely to be non-specific amplifications.

The *mgl-2* 1st round 5'RACE PCR amplified several bands that were located in the same size range as the predicted product (~2.7Kb) *(see figure 3.13A)*. To select for *mgl-2* 5'RACE fragments a nested RACE reaction was performed on the first round PCR. A series of prominent bands were amplified close to the predicted size *(see figure 3.13B)*. The two longest *mgl-2* 3'RACE cDNA fragments amplified *(see figure 3.13C)* are very close in size to that predicted for the two alternative splice variants identified by PCR amplification of the *mgl-2* C-terminal *(see section 3.3.1)*.

A B C

mgl-2 5'RACE | mgl-2 Internal Control

mgl-2 5'Nested RACE | nAP | mgl-2 5' RACE nGSP

mgl-2 3'RACE | mgl-2 Internal Control | mgl-2 3'RACE GSP | AP
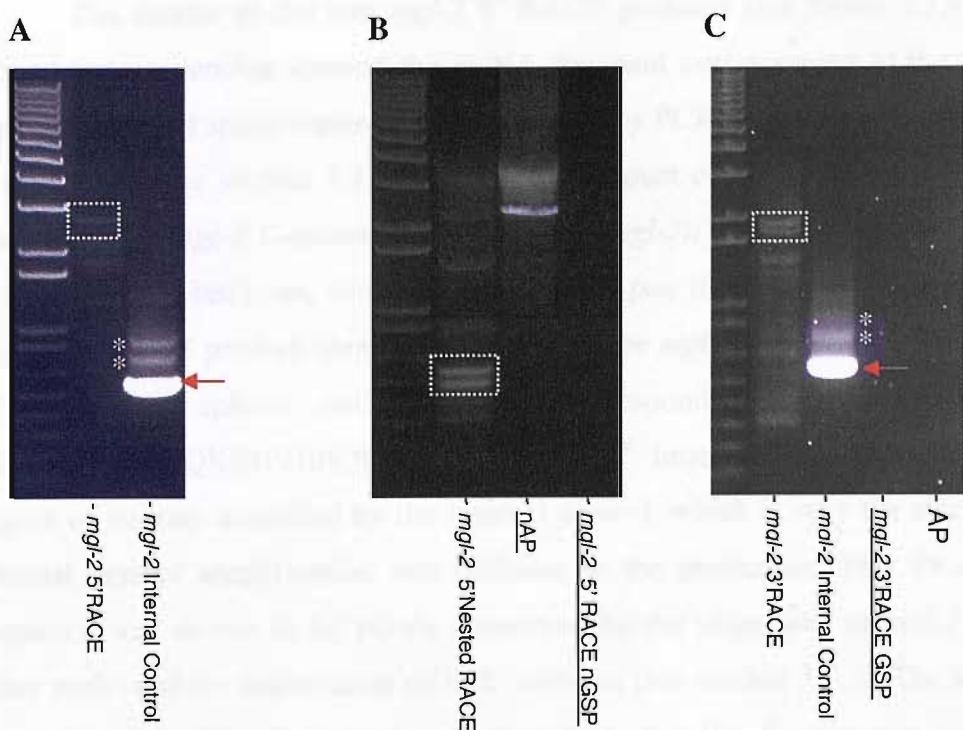
Figure 3.13 *mgl-2* 1st round 5'RACE, nested 5' RACE and 3' RACE analysis. A. *mgl-2* 5'RACE PCR. The red box highlights the PCR product in the predicted size range. Yellow asterisks indicate non-specific amplification products. **B. *mgl-2* 5'nested PCR reaction.** Red asterisks indicate bands gel extracted, cloned and sequenced **C. *mgl-2* 3'RACE PCR.** The red box highlights the PCR product in the predicted size range that was cloned and sequenced.

The longest *mgl-2* 5'nested RACE product that was cloned and sequenced was not *trans*-spliced to either SL-1 or SL-2. The RACE adaptor sequence was located 50 nucleotides upstream of the assigned start codon. The translation of the 50 nucleotide 5' extension in the same frame as the MGL-2 protein sequence did not identify an alternative start codon to that already assigned by the *in silico* gene model. The RACE product supports the splicing boundaries of the first 4 exons of the *mgl-2* gene model and it supports the currently assigned translational start codon of *mgl-2* by the predicted gene model. Alignment of the shorter *mgl-2* 5'RACE fragments amplified showed they were produced by the alternative incorporation of the 5'RACE adaptor sequence at the 5' end and are likely to be incomplete reverse transcriptions.

The shorter of the two *mgl-2* 3' RACE products *(see figure 3.13C)* was cloned and sequencing showed the cDNA fragment corresponded to the shorter *mgl-2* C-terminal splice variant, *mgl-2ii*, defined by PCR amplification of the *mgl-2* C-terminal *(see section 3.3.1)*. The RACE product confirms the exon splicing pattern of the *mgl-2* C-terminal splice variant *mgl-2ii* and it identifies a 397bp extension of the last exon, which is in the 3'UTR *(see figure 3.14)*. In addition the *mgl-2ii* 3'RACE product identified exon 11 of the *mgl-2* gene model defined by Wormbase is spliced out. Exon 11 corresponds to 19 amino acids (SQPAKSLPFFQKEHVIIPQ) located in the 3rd intracellular loop and in the region of overlap amplified by the internal control, which is why the size of the internal control amplification was different to the prediction. This 19 residue sequence was shown to be poorly conserved by the alignment of *mgl-2* to the other *mgls'* and the mammalian mGluR subtypes *(see section 3.7.5)*. The splicing causes an I to M substitution and other than this the downstream sequence remains the same. The upper band amplified corresponds in size to the longer splice variant, *mgl-2iii* and therefore provides further evidence for the presence of this splice variant.
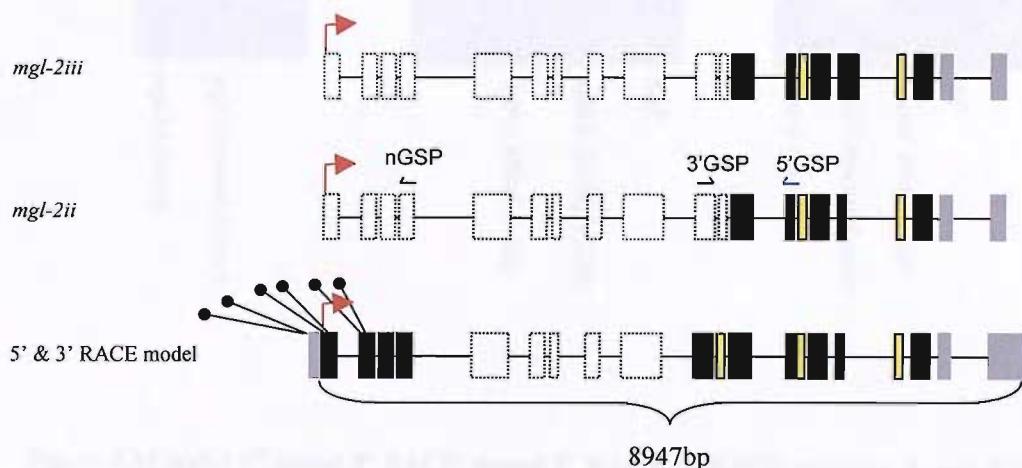
Figure 3.14 *mgl-2* 5' and 3'RACE defined gene model. The black boxes represent exons confirmed by the analysis of cDNA fragments. The dashed boxes represent predicted exons defined by the available *mgl-2* gene model in Wormbase. Yellow boxes represent exons predicted by Wormbase but were not identified by the analysis of *mgl-2* cDNA. The black pins represent the point of incorporation of the 5'RACE adaptor sequence. The red arrow indicates the position of the translational start codon. Grey boxes represent untranslated sequence.

121

3.6.3   *mgl-3* RACE characterization

The *mgl-3* internal RACE control amplified a product close to the predicted size. The *mgl-3* 5' RACE reaction produced a single band that was very close in size to the predicted product *(see figure 3.15A)*. A nested 5' RACE reaction was performed on the first round RACE PCR and multiple bands in the size range of the predicted product were obtained *(see figure 3.15B)*. The *mgl-3* 3'RACE amplification produced what appeared as a single, prominent band close to the predicted product size *(see figure 3.15C)*.
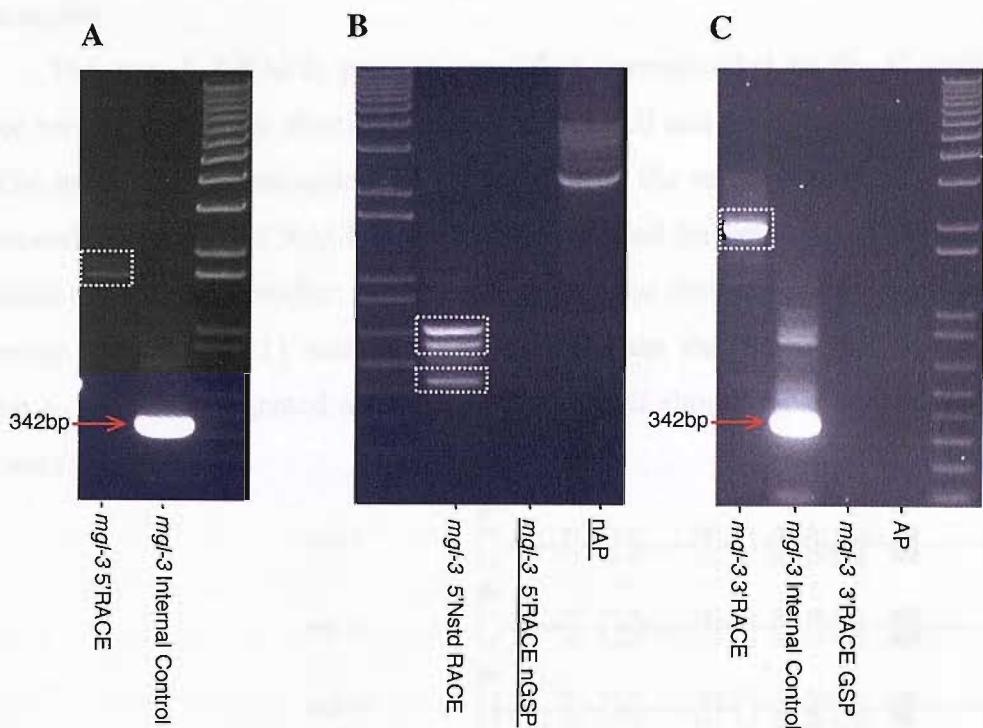
Figure 3.15 *mgl-3* **1ˢᵗ round 5' RACE, nested 5' RACE, 3' RACE reaction. A.** *mgl-3* 5'RACE PCR. The white hashed box highlights the PCR product in the predicted size range. **B.** *mgl-3* nested (Nstd) 5' RACE PCR. The white hashed box indicates PCR product gel extracted, cloned and sequenced **C.** *mgl-3* 3'RACE PCR. The white hashed box indicates PCR product in the predicted size range that was cloned and sequenced. The red arrow (figure A and B) indicates the internal control PCR product of the predicted size.

Sequencing showed the multiple *mgl-3* nested 5' RACE products were because of the alternative incorporation of the 5'RACE adaptor sequence. The sequence of the longest *mgl-3* 5' nested RACE product identified a 57 nucleotide exon located ~15kb upstream of the first assigned exon of the predicted gene model and a 68 nucleotide 5' extension of the first assigned exon *(Figure 3.16)*. Together this identified a combined total of 126 nucleotides upstream of the start of the *mgl-3* transcript that the *in silico* model had failed to identify. Translation of the upstream 126 nucleotide sequence produced a nonsense amino acid sequence when in the same frame as the MGL-3 protein sequence and did not identify an alternative translational start to that currently assigned by the predicted gene model.

The *mgl-3* 3'RACE product amplified corresponded to the C-terminal splice variant previously identified *(see section 3.6.3)* and assigned the name *mgl-3i*. The *mgl-3i* C-terminal splice variant agrees with the *mgl-3* gene model defined by wormbase and the 3'RACE product has provided further support for this. In addition the 3'RACE product confirmed the database defined splicing boundaries of exons 7, 8, 9, 10, 11 and 12. 3'RACE products that correspond to the C-terminal variants designated as *mgl-3ii* and *mgl-3iii* should have been amplified but were not identified.
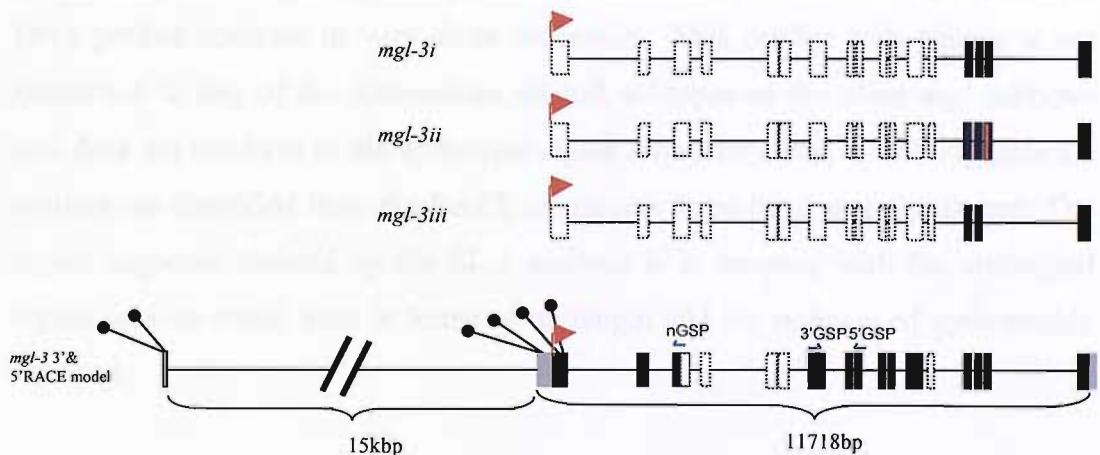
Figure 3.16 *mgl-3* 5' and 3'RACE defined gene model and comparison to the predicted gene model. The black boxes represent exons confirmed by the analysis of cDNA fragments. The dashed boxes represent predicted exons defined by the available *mgl-3* gene model in Wormbase. The black pins represent the point of incorporation of the 5'RACE adaptor sequence. The arrow indicates the position of the translational start codon. Grey boxes represent untranslated sequence. Analysis of genomic sequence ahows that the 15kb inbtron indetified does not contain any genes.

3.7 SL1 characterisation identifies discrepancies in the database prediction of the *mgl-1* 5' end

An SL screen of cDNA reverse transcribed from *C. elegans* total RNA using an *mgl-1* gene specific primer was used as the template to perform the SL1 characterization of *mgl-1*. A PCR reaction using the primers SL1 and 5' RACE GSP amplified a cDNA fragment ~1.7kb in size and a shorter fragment ~1.6kb in size *(data not shown)*. The two fragments where cloned and sequenced. The SL2 amplification performed with the 5' RACE GSP did not produce any amplification products.

SL1 *trans*-splicing identified two *mgl-1* transcripts with different 5' ends and bona fide *trans*-splice acceptor sites. The two transcripts differ from each other by a ~100 nucleotide non-coding exon and differ from the gene annotation predicted by the wormbase database *(figure 3.17A)*. Sequence analysis of the SL1 *trans*-spliced fragments identified an alternative 5' end of the *mgl-1* gene to that defined by RACE analysis *(see figure 3.17A)*. This difference defines an alternative transaltional start codon to that defined by the RACE analysis *(see figure 3.17B)*.

One pronounced feature of signal peptide motif defined from the RACE cDNA analysis is that the first predicted methionine is followed by a series of three proline residues in very close succession. This proline rich feature is not conserved in any of the mammalian mGluR subtypes or the other *mgl* subtypes and does not conform to the archetypal signal sequence motif, which suggests the methionine identified from the RACE is unlikely to be the translational start. The signal sequence defined by the SL-1 analysis is in keeping with the archetypal signal peptide motif, both in terms of its length and the richness of hydrophobic residues.
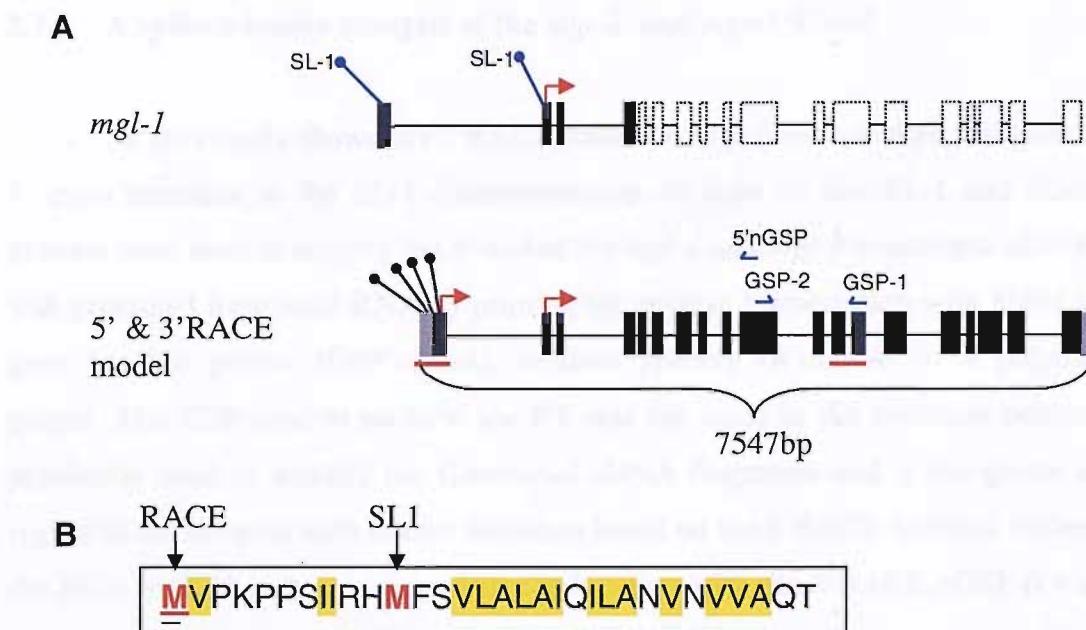
Figure 3.17. SL-1 *trans*-splicing generates two *mgl-1* transcripts. A. A comparison of the SL-1 and RACE characterization of the *mgl-1* gene 5' exon structure. The blue pins correspond to the alternative *trans*-splicing of the SL-1 sequence to the 5' end of the *mgl-1* gene. **B.** The first 32 amino acids of the translated MGL-1 sequence, upto the predicted signal peptide cleavage site, is shown. The methionine residues defined by the RACE analysis and the two SL-1 *trans*-splice variants are indicated. Hydrophobic residues are highlighted in yellow.

3.7.1 A spliced-leader analysis of the *mgl-2* and *mgl-3* 5' end

As previously shown the 5'RACE result for *mgl-1* identified an alternative 5' exon structure to the SL-1 characterization. In light of this SL-1 and SL-2 primers were used to amplify the 5' end of the *mgl-2* and *mgl-3* transcripts. cDNA was generated from total RNA by priming the reverse transcription with either a gene specific primer (GSP'cDNA), random primers (R'cDNA) or a polydT primer. The GSP used to perform the RT was the same as the antisense primer previously used to amplify the C-terminal cDNA fragments and it recognises a region of the receptor with known sequence based on the 3'RACE analysis. Either the SL-1 or SL-2 primer was used in combination with the 5'RACE nGSP (i.e a different primer to the GSP used to perform the reverse tranascription). The amplification of *mgl-2* and *mgl-3* 5' ends from R'cDNA, polydT'cDNA and GSP'cDNA using SL primer did not amplify any product within the predicted size range, however, both of the *mgl* transcripts were detected by the internal controls *(as described for the RACE experiments, see section 3.4.2; data not shown)*.

Subsequently the 1st round 5' RACE amplified cDNA for both *mgl-2* and *mgl-3* was used for the SL PCR because it served as a cDNA pool enriched for *mgl* 5'end. The SL characterization was done for *mgl-1* as well, to confirm the results of the SL-1 characterization described in *section 3.7*. The SL-1 amplification of the *mgl-1* 5'end amplified a single prominent band, that was not present in the negative primer controls *(data not shown)*. This was sequenced and it corresponded to the longest SL-1 *trans*-splice variant previously identified using unamplified cDNA *(see section 3.7)* . However, the shorter SL-1 *trans*-splice variant that was previously identified *(see section 3.7)* using unamplified cDNA was not amplified in this instance from the first round *mgl-1* 5'RACE cDNA *(data not shown)*.

Both the SL-1 and SL-2 amplification of the *mgl-2* cDNA 5'end yielded multiple bands but none of the abundant cDNA fragments were located in the size range that would be predicted for the *mgl-2* 5' end amplification using either SL primer with the 5'RACE nGSP *(see figure 3.18A)*. Determining which of the PCR products amplified and outside of the predicted size range might be specific and correspond to the *mgl-2* cDNA 5' end was hampered because of the non-specific

amplifications in single primer controls by both SL-1 and 5'RACE nGSP *(see figure 3.18A)*.

The SL-1 amplification of the *mgl-3* 5'end yielded a single, prominent band of the predicted size that was not amplified by either SL-1 primer or the 5'RACE nGSP in single primer controls *(see figure 3.18B)*, whereas the SL-2 amplification did not amplify anything close to the size predicted. Sequencing of the *mgl-3* SL-1 amplified product showed that SL-1 is *trans*-spliced to the 5'end of *mgl-3* 68 neucleotides upstream of the start codon identified by the 5'RACE analysis *(see figure 3.18C and D)*. This 68 nucleotides correspond to a 5'extension of the first coding exon of the *mgl-3* gene model. The exon 15kb upstream of the first coding exon and identified by RACE *(see figure 3.16)* was not identified by the SL-1 characterization of *mgl-3*. The SL-1 characterization confirms the 5'RACE defined splicing boundaries of the first 2 coding exons at the 5'end of the *mgl-3* gene model, this corresponds to the minimum region of coding sequence necessary for constructing a gene fusion reporter construct.

The findings from the molecular analysis of the three *mgl* subtypes have been used to ammend the protein sequence of each *mgl* subtypes. The ammended sequences were used to perform an alignment, which is shown in *figure 3.19* and a cladogram *(figure 3.20)*. The improvements to the alignment and the cladogram are outlined in the respective figure legends (see below).
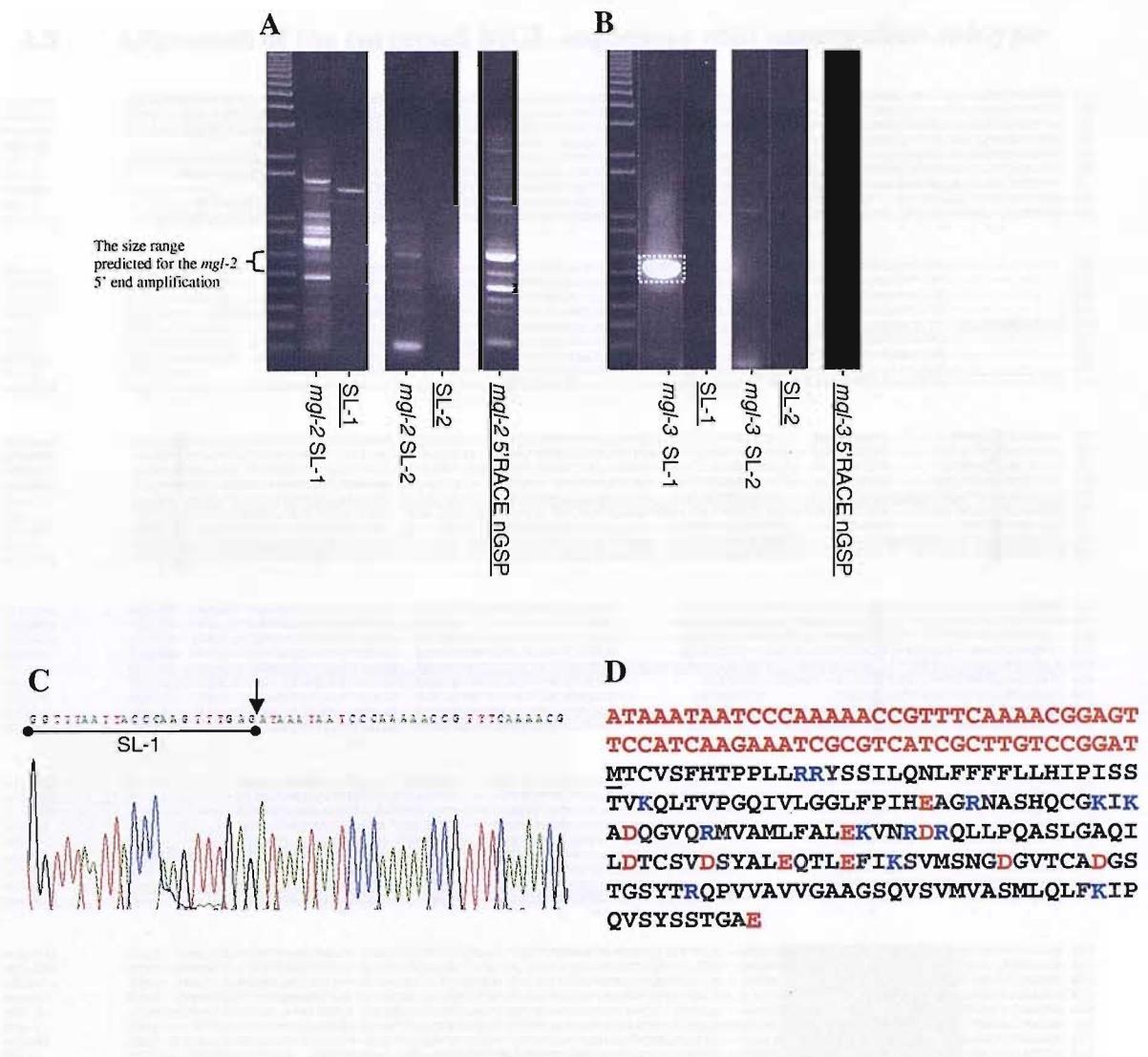
Figure 3.18 SL-1 and SL-2 amplifications of the *mgl-2* and *mgl-3* 5' cDNA end performed on the 1st round RACE amplification. Either SL-1 or SL-2 primer was used in combination with the appropriate *mgl* subtype 5'RACE nGSP (previously used for the SMART RACE characterization) to perform a PCR on both the *mgl-2* and *mgl-3* 1st round RACE amplified cDNA. **A)** The SL-1 and SL-2 amplification of *mgl-2* cDNA 5'end. **B)** The SL-1 and SL-2 amplification of *mgl-3* cDNA 5'end. The white hashed box indicates PCR product in the predicted size range that was cloned and sequenced. **C)** A selected region of the chromatogram from sequencing the *mgl-3* SL-1 amplification product. The point at which SL-1 is *trans*-spliced to *mgl-3* is indicated by the arrow. The SL-1 nucleotide sequence is underlined. **D)** The protein sequence translated from the amplified cDNA. The 68 nucleotide non-coding sequence located between the putative start codon and point of SL-1 *trans*-splicing is highlighted in red. The putative start methionine is underlined, this residue is also proposed by the *in silico* prediction.

3.8   Alignment of the corrected MGL sequences with mammalian subtypes

```
mGluR4a  MSGKGGWAWWWARLPLCLLLSLYAPWVPSSLGKPKGHPHMNSIRIDGDITLGGLFPVHGRGS-EGKA---CGELKKEKGIHLEAMLFALDRINNDPDLLPNITLGARILDTCSRDTHAL 116
mGluR8a  MVCEG---KRLASCPCFFLLTAKFYWILTMMQRTHSQEYAHSIRVDGDIILGGLFPVHAKGE-RGVP---CGELKKEKGIHLEAMLYAIDQINKDPDLLSNITLGVRILDTCSRDTYAL 113
mGluR7a  MVQLGKLLRVLTLMKFPCCVLEVLLCVLAAAARGQEMYAPHSIRIEGDVTLGGLFPVHAKGP-SGVP---CGDIKRENGIHLEAMLYALDQINSDPNLLPNVTLGARILDTCSRDTYAL 116
mGluR6   ----------------MGRLPVLLLWLAAWWLSQAGIACGAGSVRLAGGLTLGGLFPVHARGA-AGRA---CGALKKEQGVHLEAMLYALDRVNADPELLPGVRLGARLLDTCSRDTYAL 100
MGL-3    MTCVSFHTPPLLRRYSSILQNLFFFLLHIP----ISSTVKQLTVPGQIVLGGLFPIHEAGRNASHQ---CGKIKADQGVCMVAMLFALEKVNNRDRQLLPQASLGAQILDTCSVDSVAL 113
MGL-1    ------MFSVLALAIQILANVVVAQTTEAVDLAPPPKVRQIRIFGDILIGGVFPVHSKSLNGDEP---CGEIAETRGVHLEAMLYALDQINSQNDFLRGYRLGALILDSCSNPAYAL 110
mGluR2   ----------------MESLLGFLALLLLWGAVAEGPAKKVLTLEGDLVLGGLFPVHQKGG-PAEE---CGPVNEHRGIQRLEAMLFALDRINRDPHLLPGVRLGAHILDSCSKDTHAL 99
MGL-2    ---------MLYSHVTLLMLRVVDVVAKTTVQQS---ARMLVAEIHGEIQIGALFPIHRQIAGSES----CGEIWEQYGIQSEIAMLTVKQLNEE---LP-FKLGLSIRDSCWTERIAM 100
mGluR1a  ----MVRLLLIFFPMIFLEMSILPRMPDRKVLLAGASSQRSVARMDGDVIIGALFSVHHQPPAEKVPERKCGEIREQYGIQVEAMFHTLDKINADPVLLFNITLGSEIRDSCWHSSVAL 116
                                              :  *    : *:.:* *:*                   **  *    *::*  : :::.::*   *    **  : *:*    *:

mGluR4a  EQSLTFVRALIEK----DGTEVRCGSGGPPIITKP---ERVVGVIGASGSVSIMVANILRLFKIPQISYASAPDLSDNSRYDFFSRVVPSTYQAQAMVDIVR--------------214
mGluR8a  EQSLTFVQALIEK----DASDVKCANGDPPIFTKP---DKISGVIGAAASVSIMVANILRLFQIPQISYASAPELSDNTRYDFFSRVVPPSYQAQAMVDIVT--------------211
mGluR7a  EQSLTFVQALIQK----DTSDVRCTNGEPPVFVKP---EKVVGVIGASGSVSIMVANILRLFQIPQISYASAPELSDDRRYDFFSRVVPFSFQAQAMVDIVR--------------214
mGluR6   EQALSFVQALIRGRGDGDEASVRCPDGVPPLRSAPP--ERVVAVVGAAGSVSIMVANVLRLFAIPQISYASAPELSDSTRYDFFSRVVPSDYQAQAMVEIAV--------------203
MGL-3    EQTLEFIKSVMSN-----GDGVTCADGSTGSYTRQ----PVVAVVGAAGQVSVMVASMLQLFKIPQVSYSSTGAELSEKPRFAFFSRVVPPDNLQAQVMARVVCFSSLFKSSWHEIVQI 224
MGL-1    NQSLDFVRDMIGS---SEASDYVCLDGSDPNLKKQSQKKNVAAVVGGSYSVSVQLANLLRLFRIAQVSPASENADLSDKNRFEYFARTVPSDYQAMAMVEIAV--------------212
mGluR2   EQALDFVRASLSRG--ADGSRHICPDGSYATHSDAP--TAVTGVIGGSYSDVSIQVANLLRLFQIPQISYASSAKLSDKSRYDYFARTVPFFFQAKAMAEILR--------------200
MGL-2    EQTIAFLR---------EGVAQCS-CCQTPGCNKK-SVPVVAVIGPGKESSTIAVQNLLQVFRIPQVGYSAITPDLSDKEQFGYFLRVVPSDVFQAQAINRLLH--------------193
mGluR1a  EQSIEFIRDSLISIRDEKDGLNRCLPDGQTLPPGRT-KKPIAGVIGPGSSSVAIQVQNLLQLFDIPQIAYSAKSIDLSDKTLYKYFLRVVPSETLQARAMLDIVK--------------220
           :*::. *::            *                : :*.:*  .* : .*:::*:.:**  *.::.   *:.    *:. **:*.*.* **  :

mGluR4a  -ALKWNYVSTLASEGSYGESGVEAFIQKSRENGG-VCIAQSVKIP--REPKTGEFDKIIKRLLETSNA---RGIIIFANEDDIRRVL--EAARRANQTG------HFFWMGSSWGSKSAP 320
mGluR8a  -ALGWNYVSTLASEGNYGESGVEAFTQISREIGG-VCIAQSVKIP--REPRPGEFEKIIKRLLETPNA---RAVIMFANEDDIRRIL--EAAKKLNQSG------HFLWIGSSWGSKIAP 317
mGluR7a  -ALGWNYVSTLASEGSYGEKGVESFTQISKEAGG-LCIAQSVKIP--FQERKDRTIDFDRIIKQLLDTPNS--RAVVIFANEDDIKQIL--AAAKRADQVG-----HFLWVGSSWGSKINP 322
mGluR6   -ALGWNYVSTLASEGNYGESGVEAFVQISREAGG-VCIAQSIKIF--REPKPGEFHKVIRRLMETPNA---RGIIIFANEDDIRRVL--EATRQANLTG------HFLWVGSSWGSKISP 309
MGL-3    GALEWTYVHAIADTGSYGERGMDSFRAAAAENG--ICIDGDVQKIS-RRWTEKNFRDLLIRMHRTRKA--RGVVMFVDEDNLKRLL--KTLDLLVAEGHTELDRHFWFVASDSWGIKQSV 337
MGL-1    -KFKWSYVSLVYSADEYGELGADAFKKEARKKG--ICIALEERIQNKKESFTESINNLVQKLQPEKNVGATVVVLFVGTEYIPDILRYTAERMKLTSG---AKKRIIWLASESWDRNNDK 326
mGluR2   -FFNWTYVSTVASEGDYGETGIEAFELEARARN--ICVATSEKVG--RAMSRAAFEGVVRALLQKPSA--RVAVLFTRSEDARELL--RVAAVLTG------FTWVASDGWGALESV 303
MGL-2    -HYNWTYAVLYSAGNYGEKGFESLEKLIAHRSSSVCIAYSEKIKT--LASEQEYRQVLTRLDSQNSR-PQVVVCFCEGASMRMFFK-----AQKHLADGKMQMKRFQWIGSGWADRNDV 305
mGluR1a  -RYNWTYVSAVHTEGNYGESGMDAFKELAAQEG--ICIAHSDKIYS--NAGEKSFDRLLRKLRERLFK-ARVVVCFCEGMTVRGLLS----AMRRLG----VVGEFSLIGSSWADRDEV 326
           .**.  .** .****.* ::::     .:   :.  .:       *   *    :      :* *        *        ..:  .:*.*
```

```
mGluR4a     ----------------------------------------------------------------------------------------------------------
mGluR8a     ----------------------------------------------------------------------------------------------------------
mGluR7a     ----------------------------------------------------------------------------------------------------------
mGluR6      ----------------------------------------------------------------------------------------------------------
MGL-3       ----------------------------------------------------------------------------------------------------------
MGL-1       ----------------------------------------------------------------------------------------------------------
mGluR2      ----------------------------------------------------------------------------------------------------------
MGL-2       -----------------------------YQDEKIN---------SSAANLLLE-------------ESEDELLRSRGVQP---LALSQATPI----------------------- 983
mGluR1a     GPGTPGNSLRSLYPPPPPPQHLQMLPLHLSTFQEESISPPGEDIDDDSERFKLLQEFVYEREGNTEEDELEEEEDLPTASKLTPEDSPALTPPSPFRSVASGSSVPSSPVSESVLCTPP 1180

mGluR4a     --------------------
mGluR8a     --------------------
mGluR7a     --------------------
mGluR6      --------------------
MGL-3       --------------------
MGL-1       --------------------
mGluR2      --------------------
MGL-2       --------------------
mGluR1a     NVTYASVILRDYKQSSTPI 1199
```

Figure 3.19 An alignment of amended MGL receptor protein sequence to representative members of the mammalian mGluR family. The sequence underlined was not contained in the 3' and 5' cDNA fragments analysed and remains to be confirmed experimentally. Note the improved homology of transmembrane domain 6; the signal sequence motif of MGL-1 is now in-keeping with other mGluR subtypes and the MGL-2 C-terminal displays a Type I PDZ ligand motif. (NB. MGL-2 and MGL-3 in the alignment refers to the MGL-2ii and MGL-3i C-terminal splice variants).
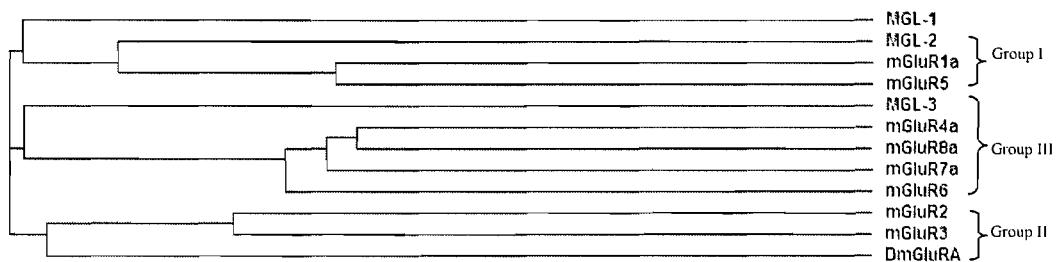
Figure 3.20 A revised cladogram produced using the revised protein sequences of the MGL subtypes. MGL-1 is now distantly related to the Group I mGluR subtypes rather than Group II. However, the BLASTp algorithm scores the Group II receptors as those most closely related to MGL-1. The position of MGL-3 and MGL-2 remains unchanged. Taken together this suggests that MGL-2, MGL-1 and MGL-3 are evolutionarily related to Group I, Group II and Group III mGluR subtypes respectively. However, further experimental verification will be required to test this. The cladogram was generated with sequence for MGL-2ii and MGL-3i C-terminal splice variants.

3.9   *mgls* display a differential and selective expression in *C. elegans*.

3.9.1   Analysis of existing transgenic lines

Existing transgenic worms expressing either an MGL-1 or MGL-2 GFP fusion protein *(see figure 3.21a and 3.21b respectively)* revealed *mgl* receptors are differentially expressed within the *C. elegans* nervous system. MGL-1 expression is more widespread, in the nerve ring, pharyngeal motorneurons and interneurons, whereas MGL-2 expression is restricted to the pharyngeal nerve ring and suggests the two receptor subtypes are involved in different physiological processes.

MGL-1::GFP expression in cell bodies of the tail and their processes is both punctate and diffuse. The punctate clusters indicate the involvement of targeting machinery in MGL-1 localisation, whereas the diffuse expression may be a result of mislocalisation caused by interference of the C-terminal GFP fusion with the receptors targetting. Imaging of the unsheathed pharynx identified MGL-1::GFP is expressed within neurons of the pharyngeal neural circuit, the fusion protein can be seen in cell bodies and their branching projections, where the distribution is diffuse. The gene *cat-1* encodes a monoamine transporter involved in 5-HT vesicular packaging, the *cat-1*::GFP functional fusion has been used to visualize 5-HT expressing neurons (Sze, JY et al 2002). The gene *tph-1* encodes tryptophan hydroxylase, the serotonin biosynthetic enzyme. The fusion of *tph-1* to GFP has also been used to visualise the NSM neurons (Sze, JY et al 2002). Our comparisons with these published images of NSM *(shown in figure 3.21a)* have enabled the identification of the neurosecretory motorneurons (NSM) as expressing MGL-1::GFP.

Expression of MGL-1::GFP was also identified in processes that extend towards the nose region *(indicated by the yellow arrow in figure 3.21a)*. These processes could correspond to either amphid sensory neurons, since the mammalian receptor mGluR 4a is involved in taste sensation, or to interneurons of the pharyngeal nervous system (such as I1 or I2) whose processes extend towards the nose. As already described MGL-1 is expressed in both the nerve ring, where the amphid cell bodies are located and the pharyngeal nervous system where the interneurons I1 and I2 are located. DiI staining of the amphid sensory neurons however did not identify any colocalization with MGL-1::GFP *(see section 3.10)* and analysis of confocal images *(see figure 3.22)* argues these

processes belong to either the interneurons I2 or I1 of the pharyngeal nervous system.
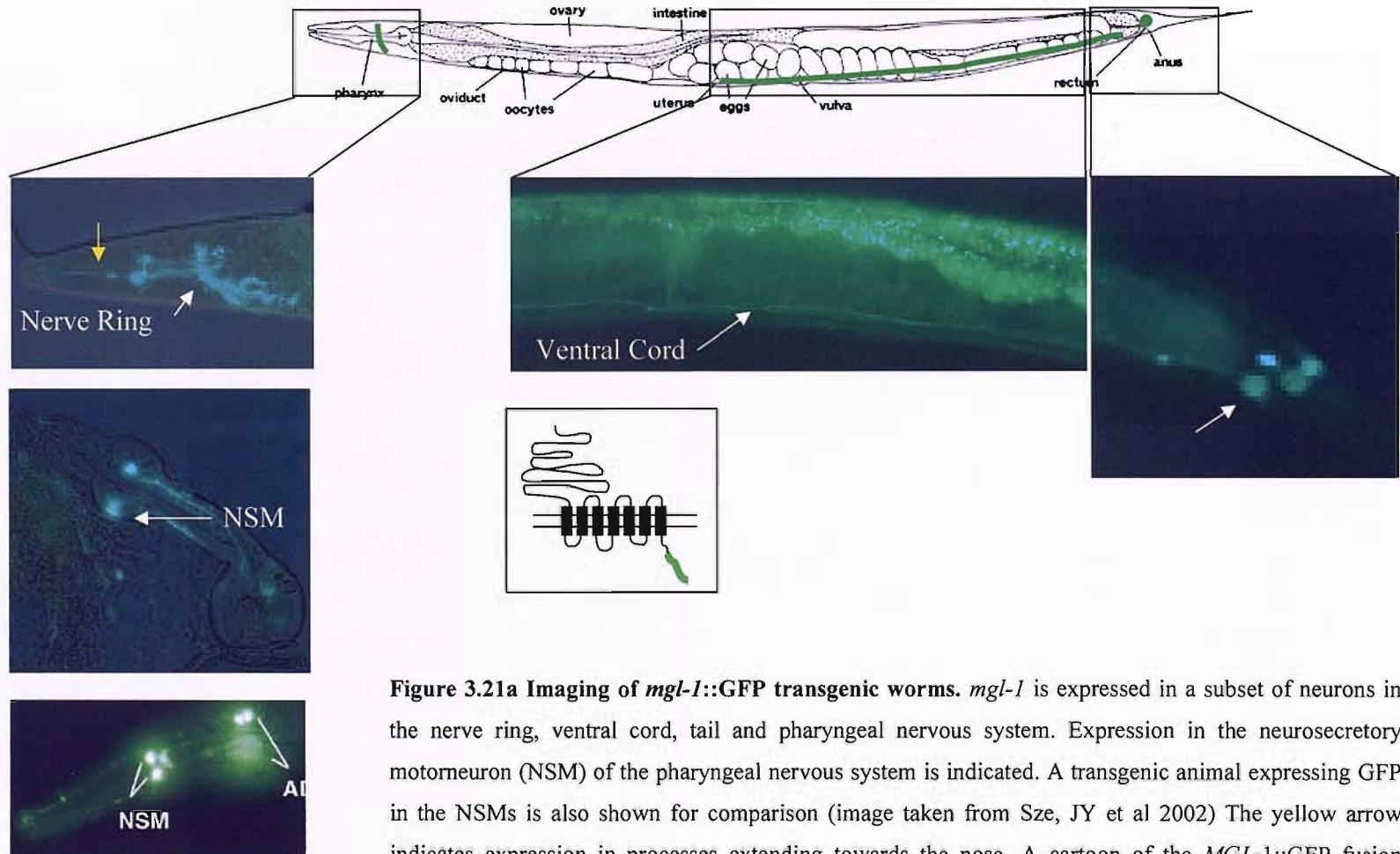
Figure 3.21a Imaging of *mgl-1*::GFP transgenic worms. *mgl-1* is expressed in a subset of neurons in the nerve ring, ventral cord, tail and pharyngeal nervous system. Expression in the neurosecretory motorneuron (NSM) of the pharyngeal nervous system is indicated. A transgenic animal expressing GFP in the NSMs is also shown for comparison (image taken from Sze, JY et al 2002) The yellow arrow indicates expression in processes extending towards the nose. A cartoon of the *MGL*-1::GFP fusion construct expressed by transgenic worms is shown. GFP is fused to exon 18 of the *mgl-1* gene at position 2081 of the ORF, generating a 59 amino acid C-terminal truncation (the transgenic line was supplied by Isao Katsura, National Institute of Genetics, Mishima, Japan).
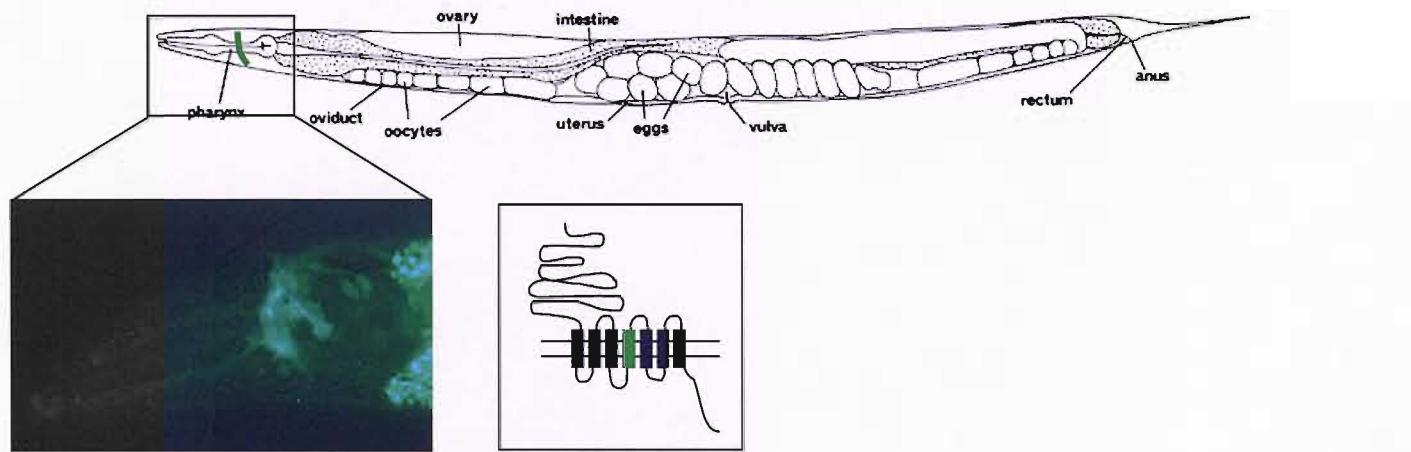
Figure 3.21b Imaging of *mgl-2*::GFP transgenic worms. *mgl-2* is expressed in a subset of neurons in the nerve ring. A cartoon of the MGL-2::GFP fusion construct is shown. GFP is fused to exon 10 of the *mgl-2* gene at position 2090 of the ORF using a BamH1 restriction site generating a 571 amino acid C-terminal truncation. Green represents the GFP fusion site and the location of this tag in relation to the receptors structure (the transgenic line was supplied by Isao Katsura, National Institute of Genetics, Mishima, Japan).

135
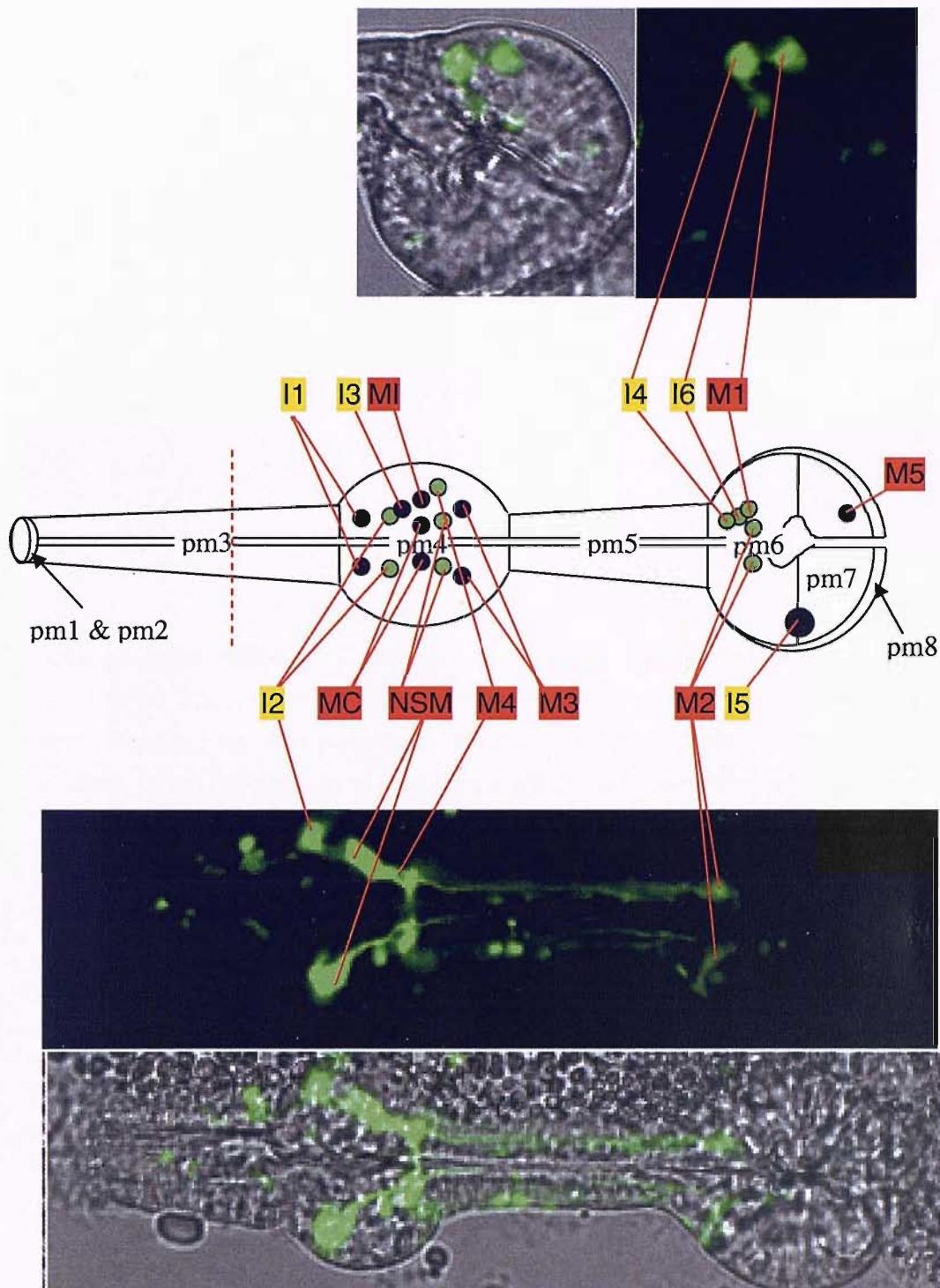
Figure 3.22 Expression of *mgl-1* in pharyngeal nervous system. Confocal images taken from the same animal that best represent the expression pattern of MGL-1::GFP in the pharyngeal nervous system are shown. The positions of the neuronal cell bodies are cartooned in 2D. The neurons suggested to express MGL-1::GFP are indicated by the red arrows. Neurons highlighted in red and yellow are motorneurons and interneurons respectively. The hashed red line indicates the position at which the nose of the worm was cut, to desheath the pharnx.
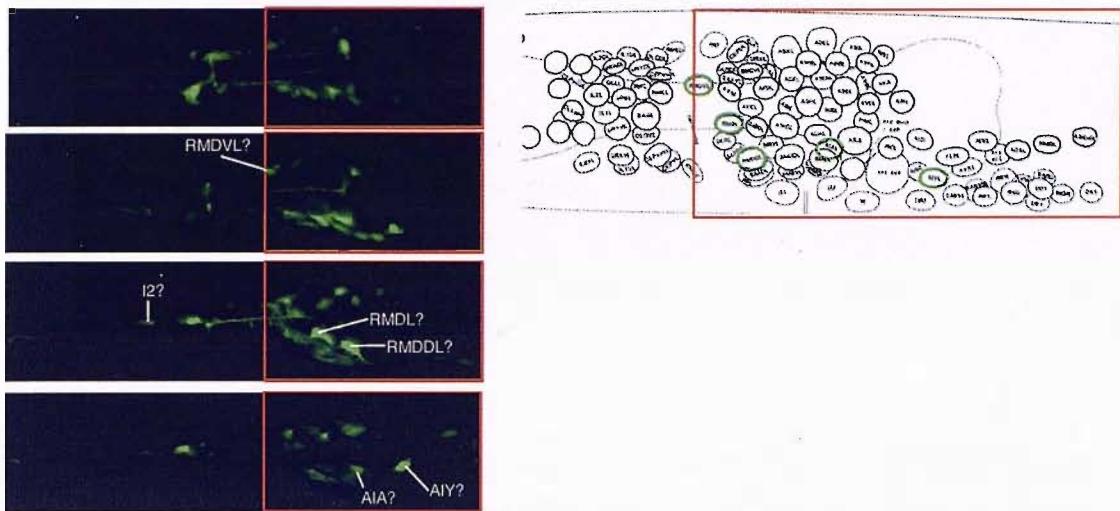
Figure 3.23 Depiction of MGL-1::GFP expression in the adult hermaphrodite nerve ring. A. Confocal images MGL-1::GFP expression in the nerve ring. Images shown were selected at 5 µm intervals from a z-stack performed on a single animal. **B. A diagram of the arrangement of neuronal nuclei in the circumpharyngeal nerve ring of the hermaphrodite.** Neurons described as expressing *mgl-1* are highlighted in green and the putative corresponding cell body is labelled in Figure A. The red box highlights corresponding regions of the head. The I2 interneuron of the pharyngeal nervous system is indicated, it is this neuron that is suggested to be responsible for the processes expressing MGL-1::GFP that extend towards the nose.

3.9.2 Independent assessment of *mgl-1* expression in *C. elegans* by using an N-terminal GFP protein fusion construct

Placing the GFP at the C-terminal of the NMR-1 receptor has been shown to prevent correct subcellular targetting. Correct subcellular targetting was achieved by placing the GFP at the N-terminal. Therefore, to better define the precise subcellular localization of MGL-1 an N-terminally tagged *mgl-1* fusion protein reporter was generated. Based upon the SL-1 and RACE analysis the GFP was inserted into the second coding exon of the *mgl-1* gene *(see figure 3.24A)*. The precise position corresponds to 2 amino acids after the predicted signal peptide cleavage site. 2.5kb of genomic sequence 5' of the first exon and 1kb of 3'UTR was included in the construct. The expression of an N-terminally tagged *mgl-1* protein fusion reporter independently confirmed the gross expression of *mgl-1* in neurons of the nerve ring, the tail and the pharyngeal nervous system *(figure 3.24B)*. The flourescence of the N-terminally fused GFP reporter appeared as a halo, encapsulating the nucleus of the cell body. Very little expression was identified in neuronal processes, compared to the MGL-1::GFP expression.
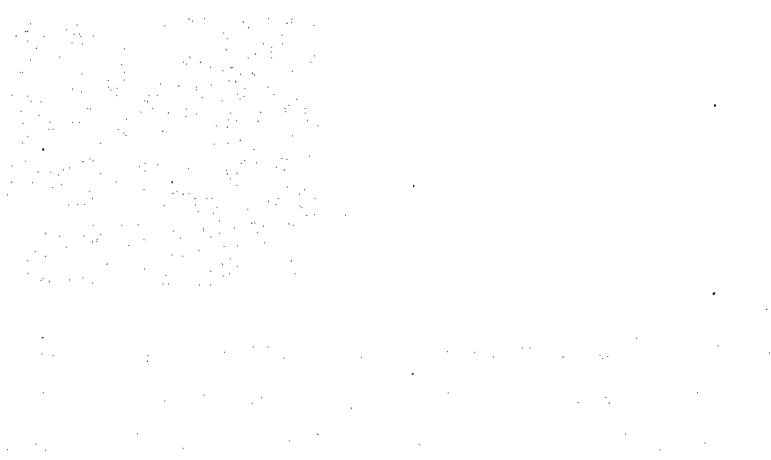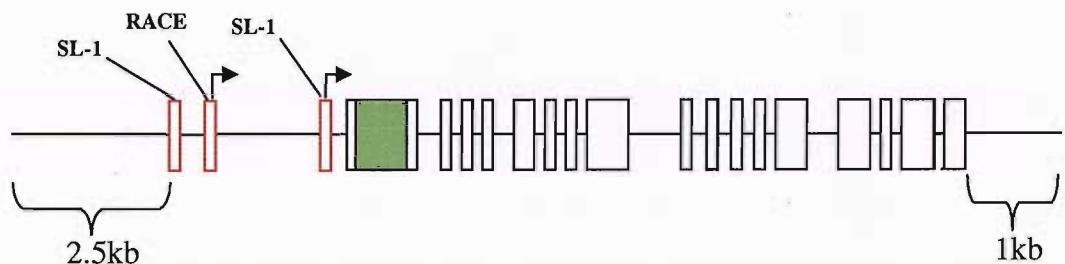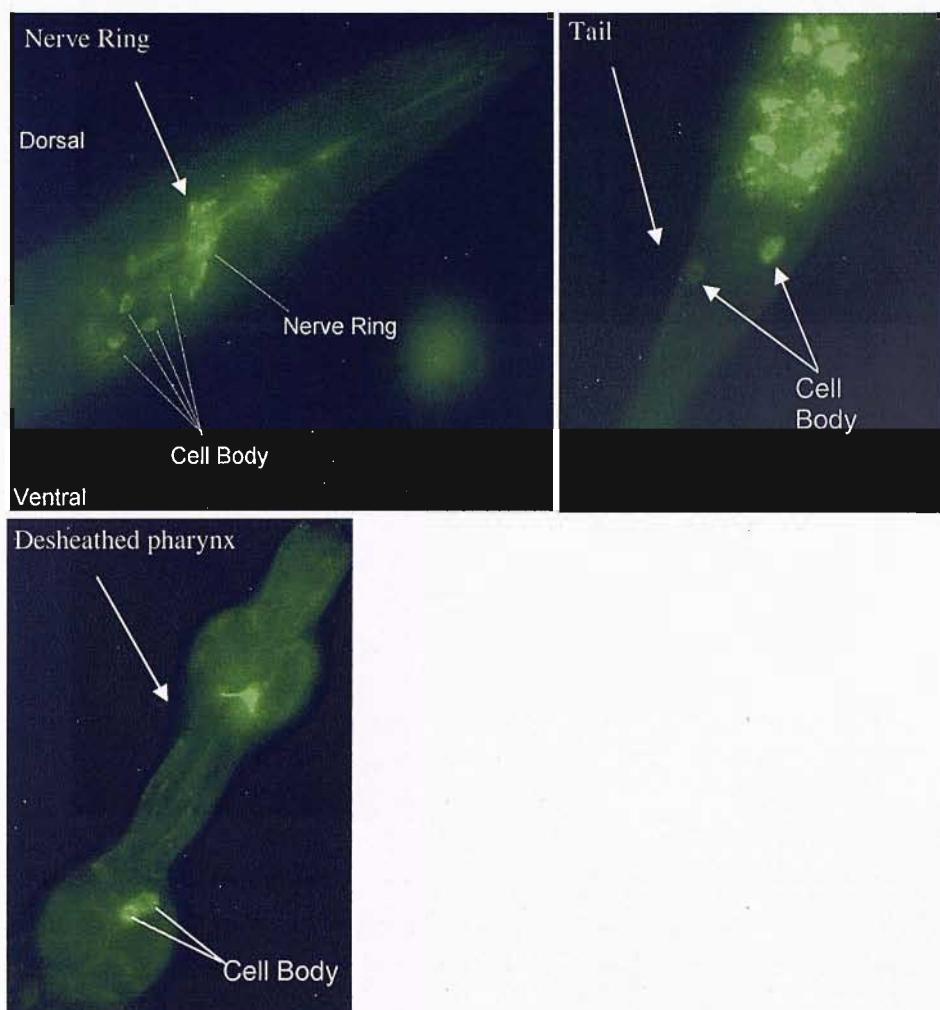
Figure 3.24 The expression pattern of the *mgl-1* N-terminal GFP fusion protein reporter. A. The site of GFP fusion to the *mgl-1* gene. Putative translational start codons are defined by the black arrows. The position of the SL-1 *trans*-splice sites and the 5'RACE adaptor incorporation are indicated. **B. The expression of the *mgl-1* N-terminally tagged GFP protein fusion reporter.** Expression was identified in the nerve ring, tail and the pharyngeal nervous system. Exposure time 10sec. Objective: x60 Oil.

### 3.10   DiI Staining of *C. elegans* expressing *mgl::GFP*

Based on the strong expression of MGL-1::GFP in processes extending towards the tip of the nose the lipophilic flourescent dye DiI (1,1-dioctadecyl-3,3,3,3-tetramethyl indocarbocyanine) was used to stain the amphid chemosensory neurons ASK, ADL, ASI, AWB, ASH and ASJ in the head and phasmid chemosensory neurons PHA and PHB in the tail of existing transgenic worms expressing either MGL-1::GFP or MGL-2::GFP. The ciliated dendrites of these chemosensory neurons are exposed to the exterior environment and are essential for detecting chemical cues that specify food, population density and aversive conditions. DiI staining did not colocalise in either MGL-1::GFP or MGL-2::GFP expressing worms. Consequently neither MGL-1 nor 2 are expressed in the amphid or phasmid chemosensory neurons that take up DiI *(see figure 3.25A and B respectively)*. The overlap between DiI staining of the nerve ring with MGL-1::GFP and MGL-2::GFP worms confirms these two receptors are associated with this dense neuropil.
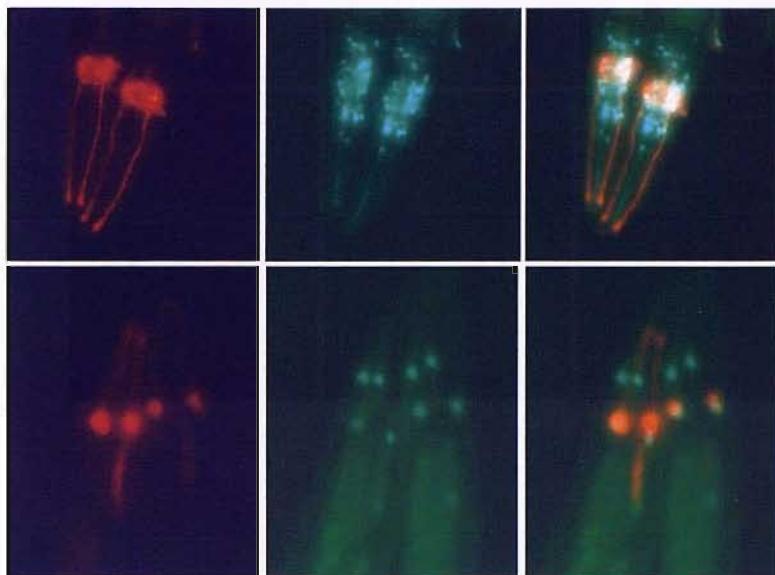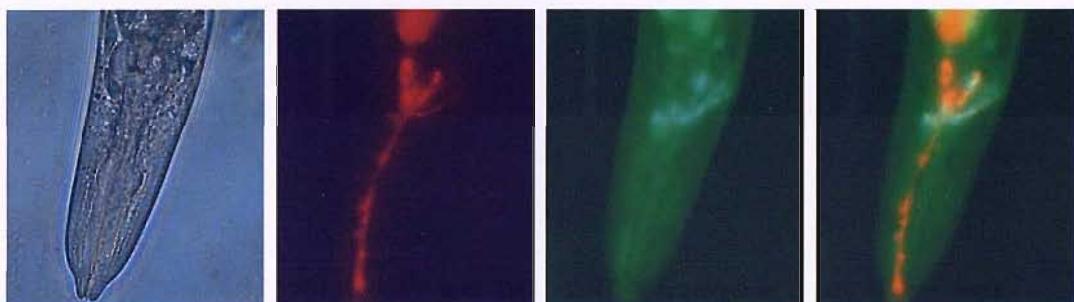
Figure 3.25 DiI Staining of MGL-1::GFP and MGL-2::GFP transgenic Lines. Worms were incubated in DiI (1mg/ml in DMF diluted 1:200 in M9 Buffer) for 3 hrs, followed by a 1hr destaining period to remove DiI from the worms intestine. **A.** MGL-1::GFP DiI staining of the **i.** pharyngeal chemosensory amphid neurons and the **ii.** phasmid tail neurons **B.** MGL-2::GFP DiI staining of the pharyngeal chemosensory amphid neurons.

3.11 Discussion

In silico analysis corroborates that the 3 genes designated as *mgl-1, mgl-2* and *mgl-3* do encode metabotropic glutamate receptors. The amino acid sequence predicts characteristic features that are shared with their mammalian counterparts. The demonstration that *mgls* can be segregated into distinct subgroups that are evolutionarily conserved further substantiates their identity as metabotropic glutamate receptors. Predictive PCR and RACE revealed discrepancies in the gene structure annotated by the Wormbase database for *mgl-2* and an incomplete prediction of the transcripts for *mgl-2* and *mgl-3*. Expression analysis of existing transgenic worms revealed *mgl-1* and *mgl-2* are expressed within different neuronal populations suggesting they perform distinct physiological roles.

With this in mind we sought to authenticate the *in silico* prediction for the *mgl-1* gene using two approaches, RACE and an SL1 characterisation. In doing so this would assist us in the construction of a GFP::MGL-1 N-terminal fusion protein to better define the subcellular localization of the receptor and confirm the expression pattern reported by the exisitng transgenic line. As previously outlined the proper localization of a receptor can be achieved more accurately by placing the GFP tag at the N-terminal, just after the signal peptide motif, as opposed to the C-terminal (Brockie, P.J et al 2001). It is therefore pivotal that the 5' end of the *mgl-1* transcript is known, so as GFP is fused correctly.

A combination of RACE and SL-1 analysis has identified, in the case of the *mgl-1* transcript three different 5'ends, all of which are different to the *in silico* prediction. This further highlights the weakness of *in silico* based tools to predict the true 5' ends of *C. elgans* genes [Lamesch, P et al 2004]. Biologically this may reflect alternative promoters are involved in the transcriptional control of the *mgl-1* gene [Choi, J. and Newman, A.P. 2006]. The use of alternative promoters is a prominent aspect of gene control and confers features such as the sex-specific, developmental stage-specific and cell-specific expression of genes [McGhee, J.D. and Krause, M.W. 1997]. This is embodied by the organisation of the promoter region belonging to the inositol-1,4,5-triphosphate gene, *itr-1*. The *itr-1* gene has three promoters, each imparts cell specific and individualised expression patterns when fused to the reporter gene GFP (Gower, N.J. et al 2001). Our findings intimate *mgl-1* may be the subject of control by distinct regulatory

elements. One could envisage the binding of cell specific trans-acting proteins to different regulatory elements of the *mgls* upstream region as a mechanism for directing and regulating a differential neuronal or cell type expression.

The SL1 and RACE based cDNA characterisation of the 5' end of the *mgl-1* transcript combined with *in silico* analysis provided sufficient information for the suitable fusion of GFP to *mgl-1* just after the signal peptide. The construct for the GFP::MGL-1 N-terminal fusion protein confirms the gross expression pattern of *mgl-1,* as reported by the MGL-1::GFP C-terminal fusion reporter. However, its expression appears to be restricted to a perinuclear domain of the cell body and the expression appears to be less pronounced. This may represent improper processing and targetting of the receptor, caused by the presence of GFP at the N-terminal. In addition the N-terminal GFP fusion to *mgl-1* was not integrated into the *C. elegans* genome, unlike the C-terminal GFP fusion to *mgl-1*, which could be a reason for why expression was weaker and less widespread with this construct.

mgl-3 5'RACE identified a small exon in the promotor region defined by the *in silico* gene model that introduces a 15kb intron. The 15kb intron does not contain any other genes and may encompass important promotor elements. However, this region is poorly conserved in *C. briggsae.* The SL-1 PCR identified a single SL-1 *trans*-splice variant and this did not contain the 15kb intron indentified by the RACE. Both SL-1 and RACE analysis do, however, confirm the *in silico* defined start codon of *mgl-3.* This highlights a combination of SL-1 and RACE allows for a more expansive characterization of *C. elegans* transcripts. This chracterization of *mgl-3* lends itself to the more efficient design of promotor::reporter constructs to define the cellular expression of *mgl-3.*

3.11   Insights into MGL scaffolding and function

The expression of *mgl-1* is widespread by comparison to other glutamate receptors, the NMR-1 subunit is only expressed in 6 neurons throughout the *C. elegans* nervous system (Brockie, P.J et al 2001). The selective inclusion of *mgl-1* and *mgl-2* in a number of different neurons suggests they perform an important role in neuronal signalling. The expression of *mgl-1* in neurons projecting towards the nose region is particularly interesting and implicates it has a role to play in sensory signalling. However, DiI staining of the animals expressing MGL-1::GFP

did not identify MGL-1 in the amphid neurons. Furthermore analysis of confocal microscopy suggests the processes extending towards the nose, that express MGL-1::GFP are likely to belong to the I2 interneuron of the pharyngeal nervous system.

A role for the C-terminal domain of the *mgls* in scaffolding their function can be extrapolated from sequencing data obtained from predictive PCR and 3'RACE. Analysis revealed the *mgl-2* and *mgl-3* C-terminals are subjected to splicing events. The alternative splicing of mammalian mGluR subtypes can regulate cis-acting signals in the C-terminal, causing them to be excluded or incorporated into regional compartments of neurons [Francesconi, A. and Duvoisin R.M. 2002]. The C-terminals obtained display protein binding motifs. Our PCR analysis clarified PDZ interaction domains in the C-terminal of all three *mgls (see section 3.5)*. The presence of a PDZ ligand motif at the extreme end of the C-terminals makes them likely substrates for interacting proteins. Subsequently it would seem that there is sufficient criteria to suggest the potential for scaffolding of *mgl* function. Further analysis, such as Northern blot, is required to confirm the presence of the alternative splice-variants of the *mgl* receptor subtypes. Subsequently the biological relevance of such alternative splicing can be assessed.

CHAPTER 4

Functional analysis of *mgl* subtypes

4.1.1 Introduction

Glutamate neurotransmission has been shown to be involved in controlling various aspects of *C. elegans* behvaiour, such as locomotion, pharyngeal pumping and responses to sensory stimuli. The NMR-1 receptor regulates locomotion during foraging behaviour (Brockie PJ et al 2001) and GLR-1 is required for the integration of mechanical stimuli with the locomotory control circuits (Maricq A et al 1995; as earlier described in *section 1.28.2*). The mutations *eat-4* and *avr-15* have highlighted the importance of glutamate in the control of pharyngeal pumping. Both mutations cause a longer pharyngeal muscle contraction (Lee, RY et al 1999 and Dent et al 1997), either by reducing or blocking glutamatergic transmission respectively. Consequently glutamate is an important transmitter in the control of feeding and feeding is impaired in worms where glutamatergic transmission is disrupted.

Both locomotion and feeding can be measured in *C.elegans*. In terms of locomotion this can be measured by, for example, timing the duration of forward and backward movement counting body bends on plates and counting thrashing, a thrash is defined as the change in the direction of bend at the mid-body of the worm when swimming in liquid [Miller, K.G. et al 1996]. Pharyngeal function can be measured by counting the rate of contraction and relaxation by the terminal bulb, which provides a measurement of the pumping rate in the intact animal. Alternatively electropharyngeograms can be used, which can provide a more detailed insight into the electrical events that occur during pumping *(see section 4.1.4)*. The expression of *mgl-1* in the pharyngeal nervous system suggests that it could be involved in signalling within this network and the control of pharyngeal function.

4.1.2 Physiology of the *C. elegans* pharynx

The *C.elegans* pharynx is a neuromuscular organ located in the head and is used for feeding. *C.elegans* is a filter feeder and it feeds by ingesting bacteria suspended in fluid and it then traps bacteria in the pharynx and expels the fluid. Normal feeding behaviour is achieved by the co-ordination of two processes, which are pharyngeal pumping and peristalsis [Albertson, D.G. and Thomson, J.N. 1976]. The pharynx is comprised of three muscle groups: the corpus, the isthmus and the terminal bulb [Altun, Z.F. and Hall, D.H. et al 2005]. Pumping of

146

the pharynx involves the synchronized contraction and relaxation of the corpus, anterior isthmus and the terminal bulb [Avery, L. and Horvitz, H.R. 1989]. Isthmus peristalsis occurs after the relaxation phase of the pump has completed and usually only occurs every fourth pump. During persitalsis a wave of contractions in the posterior isthmus transport bacteria from the anterior isthmus to the terminal bulb, where it is ground up [Avery, L and Thomas J.H 1997]. This illustrates pharyngeal function is complex [Franks, C.J. et al 2006 ; Avery, L. 1993] and dependent upon a sequence of motions that occur sequentially and in a controlled and tightly regulated manner [Avery, L. and Shtonda, B.B. 2003].
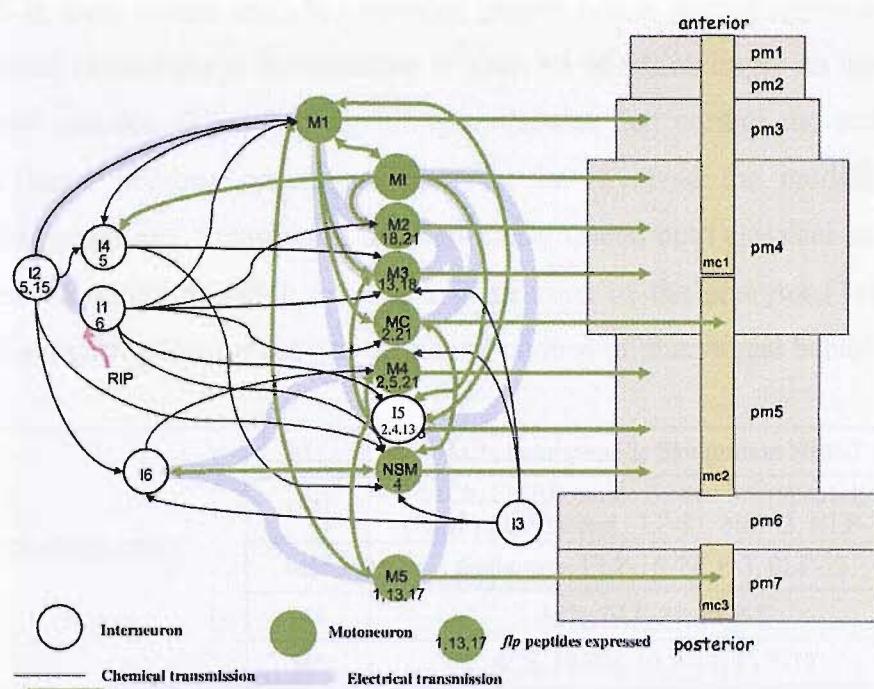
Figure 4.1 The *C. elegans* pharynx and the major neuronal inputs of the pharyngeal nervous system. The pharyngeal muscle cells (pm) and marginal cells (mc) are indicated. [Figure taken from Papaioannou, S et al 2005] NSM : Neurosecretory motorneuron; MI : Motor-interneuron; MC : Marginal Cell neurons.

4.1.3 The *C. elegans* Pharyngeal Nervous System

The *C. elegans* pharyngeal nervous system consists of 20 neurons *(figure 4.1)*, which comprises five classes of motorneurons, six classes of interneurons and a further three classes that do not fit into either of these categories *(See Table 4.1)*. Pharyngeal pumping can continue, with the corpus and terminal bulb contracting in synchrony, even when the entire pharyngeal nervous system has

been laser ablated [Avery, L and Horvitz, H.R. 1989]. However, synaptic inputs do play an important role in 'fine-tuning' pharyngeal pumping to make it a more efficient process and it is important for modulating pumping in response to environmental cues, such as the presence of food (bacteria) [Avery, L 1993]. In worms where all neurons except M4, which is required for viability, are laser ablated the frequency of pharyngeal pumping is slow (16 pumps/min) and irregular compared to wild-type (43 pumps/min) and unlike wild-type animals the pumping rate does not increase in response to bacteria [Avery, L et al 1989]. Other subtle phenotypes that infer the 'fine-tuning' of pharyngeal pumping is disrupted in these worms include a retarded growth rate, a starved appearance and the transport of bacteria to the intestine is poor, all of which imply an inefficient pharyngeal function. Therefore signalling molecules that control the activity of the pharyngeal nervous system are likely to be involved the modulation of pharyngeal pumping. A number of different transmitters, both classical and novel have been identified as being expressed in neurons of the pharynx (Table 4.1) further highlighting the intricacy of the neural control of pharyngeal behaviour.

MOTORNEURONES	M1	ACh, neuropeptide like protein NLP-3
	M2[§]	ACh, FMRF-amide related neuropeptides (FaRP) FLP-18 and FLP-21. NLP-3, NLP-13.
	M3[§]	Glutamate, FLP-18, NLP-3, FLP-13
	M4	ACh, FLP-21, FLP-5
	M5	ACh, FLP-1, FLP-13, FLP-17
INTERNEURONES	I1[§]	NLP-3, FLP-6
	I2[§]	NLP-3, NLP-8, FLP-15
	I3	NLP-3
	I4	NLP-3, NLP-13, FLP-5
	I5	5-HT, glutamate, FLP-2, FLP-4, FLP-13
	I6	ACh, NLP-3, FLP-4
OTHERS	NSM[§]	5-HT, glutamate, NLP-13, NLP-18, NLP-19, NLP-4, NLP-3 (right hand cell only)
	MI	
	MC[§]	Acetylcholine, FLP-21, FLP-2

Table 4.1. Neurons of the pharyngeal nervous system and the transmitters they express. The *C. elegans* pharyngeal nervous system consists 14 neuronal cell types and the putative transmitters they contain are indicated. § binuclear. Produced using wormatlas individual neuron pages (www.wormatlas.org)

148

4.1.4 Electrophysiological analysis of the *C. elegans* pharynx

4.1.4a The electropharyngeogram

An electropharyngeogram (EPG) is an extracellular technique that measures electrical currents that occur during pumping (Avery, L and Raizen, D 1995). The EPG represents electrical events that occur in the pharyngeal muscle membranes and the EPG can be modulated by neurochemical inputs *(see below)*. A single EPG corresponds to a single pump and consists of three distinct phases: the E phase is an upwards deflection that corresponds to a depolarization of the pharyngeal corpus and terminal bulb muscle; the P phase, where the muscle is both depolarized and contracted and the R phase, which is a downwards deflection and corresponds to repolarization and relaxation of the corpus and terminal bulb muscle [Raizen, D.M. and Avery L 1994] *(see figure 4.2)*. The timing of pharyngeal pumps is controlled by the motor neurons MC and M3 (Niacaris, T and Avery, L 2003). The E phase corresponds to currents elicited by the cholinergic motorneuron MC and the P phase typically displays several negative spikes that represent inhibitory postsynaptic potentials, mediated by glutamate release from the motorneuron M3 onto the pharyngeal muscle during action potential firing. EPG recordings can be recorded from intact worms, but typically the pharynx and the pharyngeal nervous system is dissected away from the rest of the worm by cutting transversely through the body just behind the terminal bulb *(see Methods section 2.5.1)*. This is beneficial during pharmacological experiments because it allows the drug direct access to the pharynx instead of having to cross the cuticle first.

4.1.4b Intracellular recordings

Intracellular sharp electrode recordings can be made from the muscle of the terminal bulb to examine the resting potential and the shape of the action potential in the pharyngeal muscle [Franks, C.J. et al 2002]. The technique has been used to characterize the role of various ion channel subtypes in pharyngeal muscle action potentials [for example see Davis, M.W. et al 1999].
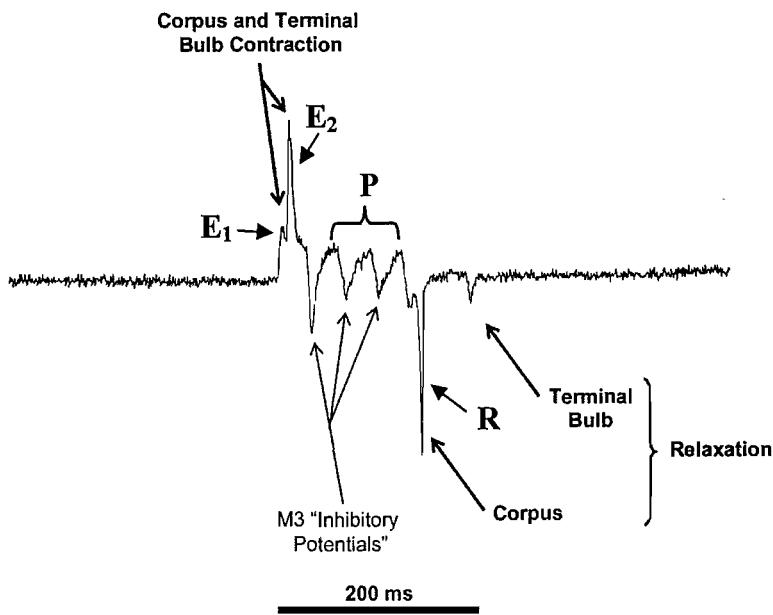
Figure 4.2 The idealized electropharyngeogram recording of a single pump. The different phases of the EPG and the pharyngeal motions that are associated with the different electrical transients are indicated.

4.1.5 The neurochemical modulation of pharyngeal pumping

Features of EPG recordings that can be measured include duration of single pumps (which is the time between contraction and relaxation phases), the rate of pumping and the shape of the pump and this provides information about the neural control of pumping [Cook, A et al 2006] . Distinct neurochemicals have been shown to regulate certain features of the EPG and this has lead to the identification or further characterization of their role in the pharyngeal nervous system and the modulation of pumping. For example, rapid pharyngeal pumping is stimulated by the exogenous application of 5-HT [Avery, L et al 1990], which is expressed in the NSMs and I5 neuron in the pharyngeal nervous system *(see table 4.1)*. This rapid rate of pumping mimics the response of *C. elegans* to food and so it has been proposed 5-HT is one neurochemical signal of the presence of food but that there is likely to be other neurochemicals involved aswell [Franks, C.J. et al 2006].

Glutamate is expressed in the M3s, NSMs and I5 and regulates the duration of single pumps. Reducing glutamate transmission, as in the *eat-4*

knockout mutant *(previously discussed in section 1.28.6)*, delays the relaxation phase of the pump and increases the duration of the EPG [Lee, R.Y. et al 1999], and the exogenous application of glutamate reduces the duration of the pump [Dent, J.A. et al 1997]. The shape of the EPG can provide useful information about the neuron activity [Cook, A et al 2006]. As previously outlined the inhibitory postsynaptic potentials (IPSPs) of the EPG plateau phase are mediated by glutamate release from M3 and so the loss of IPSPs reflects a loss of M3 activity and disruption of glutamatergic transmission in the pharynx [Raizen, D.M. and Avery, L. 1994].

Acetylcholine is expressed in M1, M2s, M4, M5, I6 and MCs and it controls the rate of pharyngeal pumps because genetic or laser disruption to cholinergic transmission causes a reduced rate of pumping [Raizen, D.M. et al 1995]. As highlighted in table 4.1 a number of neurons have been identified as expressing neuropeptides and the exogenous application of neuropeptides to the pharynx can either increase or decrease the rate of pharyngeal pumping suggesting this class of neuro-chemical is also important to the modulation of the pumping [Papaioannou, S. et al 2005].

This highlights how the activity of the pharyngeal nervous system is regulated by distinct neurotransmitters, in keeping with other neural circuits and the EPG provides a bio-assay for comparing the activity of this simple neural circuit under different experimental conditions [for example see Papaioannou, S. et al 2005]. Such comparisons could, for example, be made between different transgenic animals, different bio-molecules or different pharmacological agents (or a combination of each) to assay the biological activity of a gene, neuron, neurotransmitter or receptor signalling pathway within this neural circuit.

4.1.6   The pharmacological profile of mGluRs is conserved between phyla

The pharmacological profiling of the *Drosophilla* mGluR DmGlu$_A$R has shown that it is very similar to the mammalian receptor mGluR2, which it is closest to phylogentically. DmGlu$_A$R has a conserved capacity for binding and responding to agonists of the mammalian mGluR2 receptor. Furthermore the rank potency of mGluR2 agonists is also conserved at DmGlu$_A$R, which highlights the strong selective pressure during evolution to conserve the ligand recognition site of mGluRs across phyla (Parmentier M-L et al 2000). This supports concept of

using mammalian agonists as a basis for defining the pharmacology of orthologous receptors. Also, the *Drosophila* mGluR homologue DmXR highlights that divergence at the level of a few key residues within the ligand-binding domain is sufficient to confer different ligand recognition selectivity. Since DmXR has high sequence homology to the mGluRs but is completely insensitive to glutamate. DmXR orthologues have not been identified from either the *C.elegans* or human genomes and the identification of mXRs from *Anopheles gambiae* and *Apis melifera* could suggest this divergence has only occurred in insects (Mitri, C et al 2004).

Residues that are known from the crystallization of the mGluR1 ligand binding domain to directly contact glutamate are for the most part well conserved in the mXRs. The greatest divergence occurs in the part of the binding pocket that interacts with the γ-carboxylic part of glutamate [Mitri, C et al 2004] and residues Arg-78 and Lys-409 are not conserved in the DmXRs. Analysis of the alignment of the *C. elegans* subtypes to mammalian mGluRs identifies Arg-78 is conserved in all three *mgls* but Lys-409 is conserved only in *mgl-1* and *mgl-2*. In *mgl-3* a Tyrosine residue is at this position. However, nearly all other key residues of mGluRs involved in glutamate binding [Kunishima, N et al 2000] are conserved in all *mgls* (see section 3.2). In addition *mgl-2* is responsive to glutamate in heterologously expressed in HEK cells [Ishihara, T et al 1997]. Therefore there is good molecular and cellular evidence to suggest the *mgls* are capable of binding glutamate and therefore potentially mGluR agonists.

As highlighted key mGluR residues involved in ligand binding are mostly conserved in the *mgls* and the pharmacological profiling of DmGluRA from *Drosophila* provides evidence in support of the strong evolutionary conservation of the mGluR agonist binding domain in invertebrates. The expression of *mgl-1* in the pharyngeal nervous system suggests it could be important in this circuit. Accordingly broad-spectrum mammalian mGluR agonists were used in combination with EPG recordings and available *mgl* mutant strains to investigate if *mgls* regulate this network.

4.2.1 Measurement of basal pharyngeal pumping by wild-type animals in Dent's saline.

EPG recordings were made from wild-type worms (Bristol Strain N2) in the presence of Dent's saline. The basal rate of pharyngeal pumping was recorded for a continuous 15 minutes. This consisted of three 5 minute recording periods and each 5 minute period was followed by a Dent's saline wash *(figure 4.3A)*. A wash consisted of a complete exchange of the Dent's solution in the recording chamber (capacity: 1ml) for new Dent's. The basal rate of pharyngeal pumping was not significantly different between each 5 minute recording period and comparison of pumping in the first and last 5 minutes showed it did not depreciate over the total recording period of 15 minutes.

A single pump recorded from the wild-type strain consisted of the characteristic excitation phase (E), the plateau phase (P) punctuated by inhibitory postsynaptic potentials and the relaxatory phase (R). The E1 phase is also shown but this feature was less well defined and harder to consistently distinguish, it represents contraction of the corpus and terminal bulb *(figure 4.3B)*.

Furthermore the mechanical disturbance caused to the dissected pharynx by removing and exchanging the solution in the recording chamber during the wash did not affect pharyngeal pumping. The median basal rate of pharyngeal pumping in the first third was 73 pumps/5min, in the second third it was 79 pumps/5min and in the final third it was 74 pumps/5min *(figure 4.3C)*.
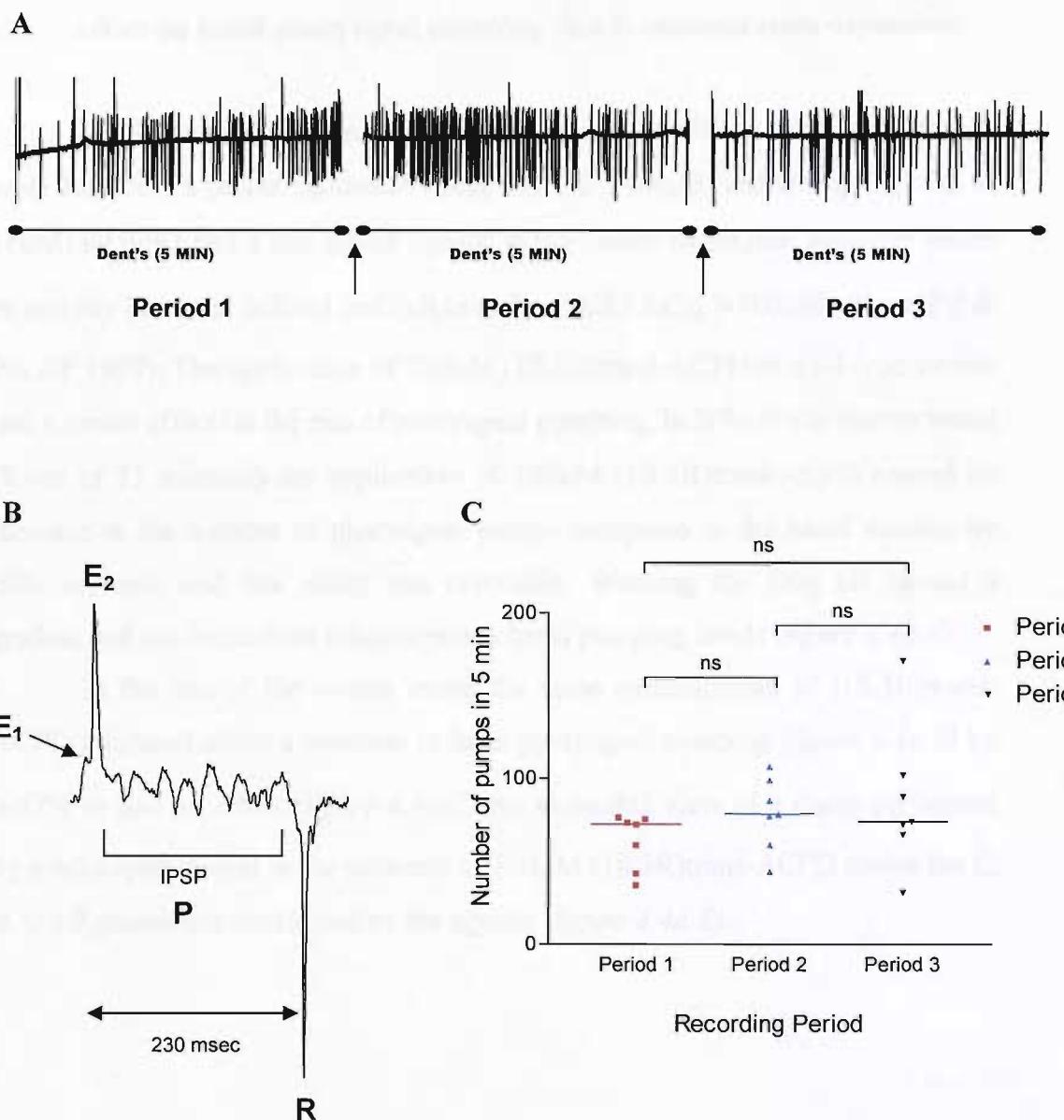
Figure 4.3 Measurement of the wild-type basal pumping rate in Dent's. **(A)** An EPG recording made from a single wild-type worm. Arrows indicate wash periods at the end of each 5 minute recording period. **(B)** An example of a single pump recorded from a wild-type worm in Dent's. **(C)** A comparison of the number of pumps in each of the three 5 minute recording periods. A two-tailed paired t-test between each 5minute period identified the pharyngeal pumping does not significantly change between 5 minute periods or over the 15 minute recording period and is not affected by mechanical disturbance. Bars indicate median values (n=7).

4.2.2 The mammalian mGluR agonist (1S,3R)trans-ACPD has a differential effect on basal pharyngeal pumping that is concentration dependent

(1S,3R)trans-ACPD is a broad spectrum mammalian mGluR agonist *(see table 1.X)*. It is a potent agonist at Group I (EC$_{50}$ 5-80μM) and Group II (EC$_{50}$ 8-18μM) mGluRs and a less potent agonist at the Group III mGluR subtypes where its activity has been defined (mGluR4a and mGluR6 EC$_{50}$ ≥300μM) (Conn P.J & Pin J-P 1997). The application of 100μM (1S,3R)trans-ACPD to wild-type worms had a varied effect on the rate of pharyngeal pumping. In 50% of the worms tested (8 out of 15 animals) the application of 100μM (1S,3R)trans-ACPD caused an increase in the number of pharyngeal pumps compared to the basal number by 20% or more and this effect was reversible. Washing the drug off caused a gradual and not immediate return towards basal pumping levels *(figure 4.4a.A)*.

In the rest of the worms tested the same concentration of (1S,3R)*trans*-ACPD produced either a decrease in basal pharyngeal pumping *(figure 4.4a.B)* by ≥37% or had no effect *(figure 4.4a.C)* An expanded view of a pump performed by a wild-type animal in the presence of 100μM (1S,3R)trans-ACPD shows the E, R and P phases are unaffected by the agonist *(figure 4.4a.E)*.

A

Dent's | 100μM (1S,3R)trans-ACPD | Dent's

B
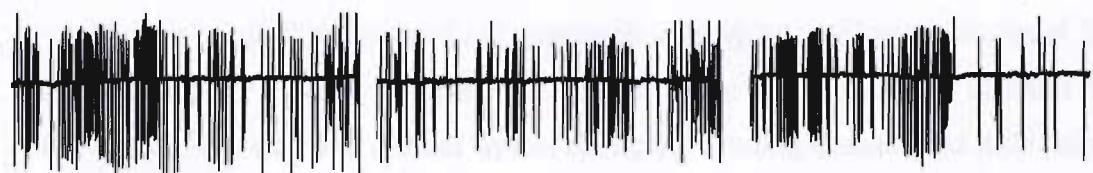
Dent's | 100μM (1S,3R)trans-ACPD | Dent's

C
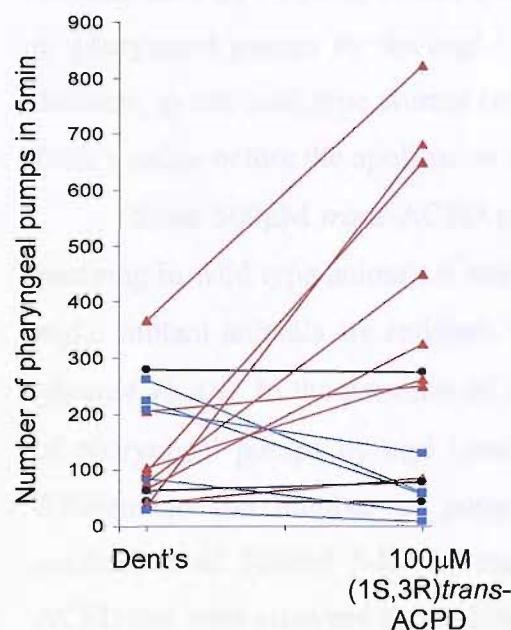
D
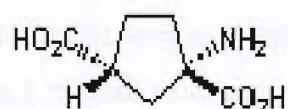
E
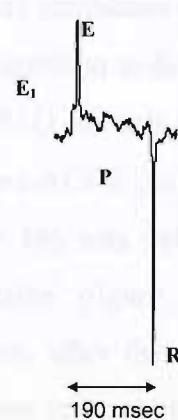
Figure 4.4a (1S,3R)trans-ACPD has a variable affect on the *C. elgans* pharynx. Two EPG recordings are shown that are representative of the (**A**) increase and (**B**) decrease in basal pharyngeal pumping observed in the presence of 100μM (1S,3R)*trans*-ACPD. (**C**) A graph of the change in pumping caused by 100μM (1S,3R)trans-ACPD. Each point corresponds to pumps made by a single animal and lines connect the pumps made by the same animal in Dent's and 100μM (1S,3R)trans-ACPD. Pumps connected by a red line signify increased network activity and blue lines signify a decreased network activity of the pharynx in the presence of drug. Black lines signify no change in network activity. (**D**) The molecular structure of (1S,3R)*trans*-ACPD. (**E**) An EPG of a single pump recorded from a wild-type worm in the presence of 100μM (1S,3R)trans-ACPD.

156

The basal rate of pharyngeal pumping of wild type animals was, however, consistently and reversibly inhibited by 500μM (1S,3R)*trans*-ACPD. The application of 500nM 5-HT after the 5 minute recovery period from 500μM *trans*-ACPD caused a rapid increase in pharyngeal pumping by all animals tested. This is the characteristic 5-HT response of a functional pharyngeal nervous system *(figure 4.4b.Ai)*

The mutant *mgl-1(tm1811)* expresses a receptor lacking a functional N-terminal ligand binding domain. The mutant allele *mgl-1(tm1811)* contains an 859bp genomic deletion located within the ligand binding domain and it therefore lacks key amino acid residues necessary for agonist binding to occur and is subsequently assumed to be non-functional. EPG recordings were made from homozygous *mgl-1(tm1811)* adult hermaphrodites. The mutant EPG displayed the characteristic E, P and R phases and the glutamate mediated IPSPs. The number of pharyngeal pumps by the *mgl-1* mutant worms (n=16) was not significantly different to the wild type worms (n=22) during the 5 minute recording period in Dent's saline before the application of (1S,3R)*trans*-ACPD *(figure 4.4b.B)* .

Since 500μM *trans*-ACPD produced a consistent inhibition of pharyngeal pumping in wild type animals it was used at this concentration to determine if the *mgl-1* mutant animals are resistant. Indeed *mgl-1(tm1811)* animals were resistant *(figure4.4b.Aii)*. In the presence of 500μM (1S,3R)*trans*-ACPD the basal number of pharyngeal pumps by *mgl-1(tm1811)* worms (n = 16) was not significantly different to the number of pumps recorded in saline *(figure 4.4b.B)*. The application of 500nM 5-HT to *mgl-1(tm1811)* worms, after the (1S,3R)*trans*-ACPD had been removed by washing, produced the same response as in wild type worms. This was observed as a robust increase in the frequency of pharyngeal pumps *(figure 4.4b.Aii)* and further confirms the pharyngeal nervous system of *mgl-1(tm1811)* mutants is functional at the level of exogenous stimulated activity. Individual pumps by *mgl-1(tm1811)* in the presence of 500μM (1S,3R)trans-ACPD were normal. A comparison of a pump produced by the same mutant worm in Dent's and 500μM (1S,3R)trans-ACPD shows that there is no difference between the two at any phase. The glutamate mediated IPSPs and the pump duration are unaffected by the application of 500μM (1S,3R)trans-ACPD *(figure 4.4b.C)*.
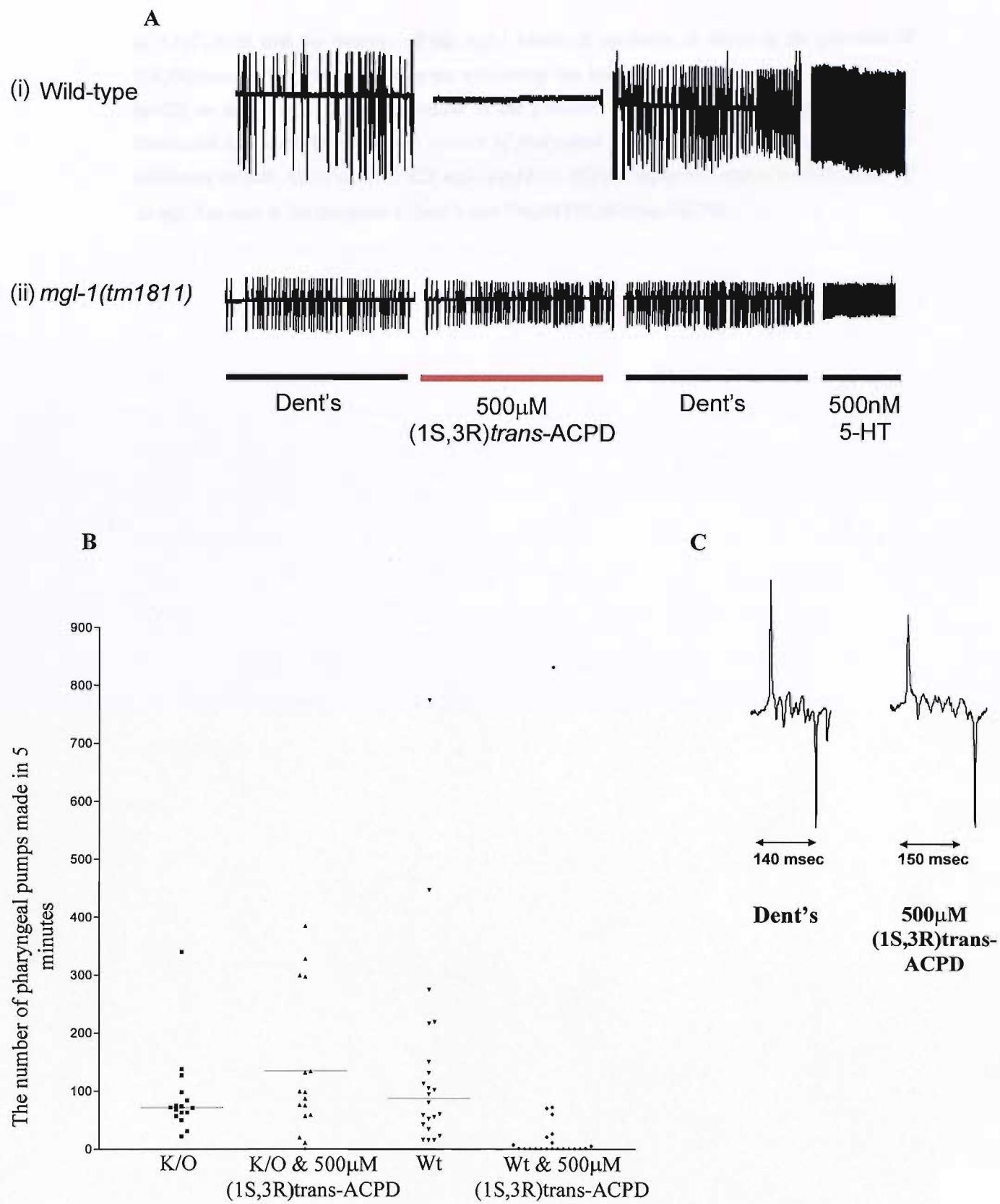
Figure 4.4b. A comparison of the response of wild-type and *mgl-1*(tm1811) to (1S,3R)trans-ACPD. (A) EPG recording of the basal pharyngeal pumping by a wild-type and *mgl-1(tm1811)* adult hermaphrodite in the presence and absence of 500μM (1S,3R)*trans*-ACPD and the response

to 5-HT. Note that the pharynx of the *mgl-1* knockout continues to pump in the presence of (1S,3R)trans-ACPD. **(B)** A scattergram comparing the basal pharyngeal pumping of wild-type (n=22) to *mgl-1(tm1811)* (n=16) worms in the presence and absence of (1S,3R)*trans*-ACPD. Horizontal bars indicate the median number of pharyngeal pumps. Each point is the result for an individual animal. Abbreviations, KO: *mgl-1(tm1811)*. **(C)** An expanded view of a single pump by an *mgl-1* mutant in the presence of Dent's and 500μM (1S,3R)trans-ACPD.
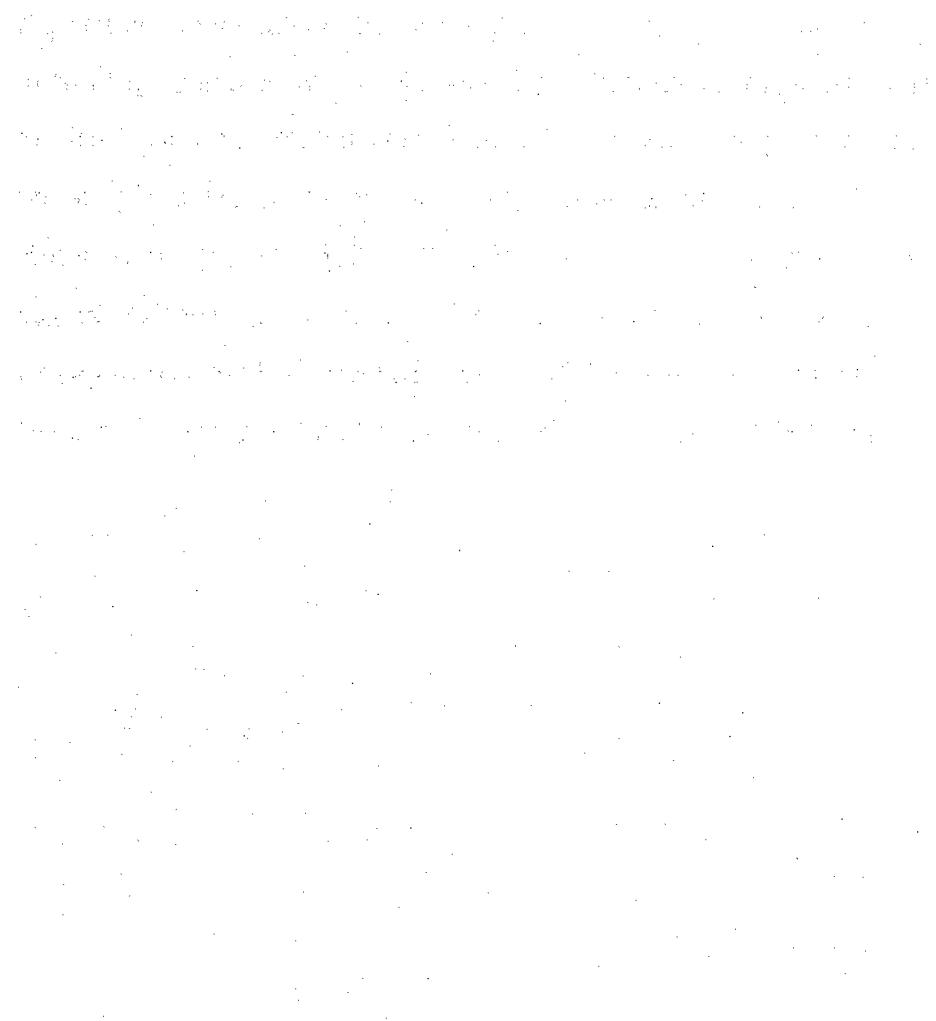
4.2.3 Intracellular recordings in the presence of (1S,3R)*trans*-ACPD

The variable response of basal pharyngeal pumping by wild-type worms in the presence of 100μM (1S,3R)*trans*-ACPD and the cessation of pumping at 500μM suggested a depolarization block could be occurring. Glutamate has been shown to inhibit the pharynx by causing a depolarization block (Pemberton DJ et al 2001). To address this sharp electrode intracellular recordings were made from the muscle cells in the terminal bulb of the pharynx of wild-type adult hermaphrodites. The same general protocol used for making the EPG recordings was applied when the intracellular recordings were made, i.e an initial recording in Dent's, followed by drug application and recovery after washing. The solution in the recording chamber was exchanged by perfusion as opposed to manual pipetting. Each action potential corresponds to a single pump of the pharynx. Spontaneous spikes were not affected by the application of 100μM (1S,3R)*trans*-ACPD but were inhibited by 500μM (1S,3R)*trans*-ACPD, At both concentrations there was no change in the resting membrane potential (-75mV recorded in Dent's) *(see figure 4.5)*. The effect of 500μM (1S,3R)*trans*-ACPD was reversible, spontaneous spiking was restored when the drug was exchanged for Dent's saline.
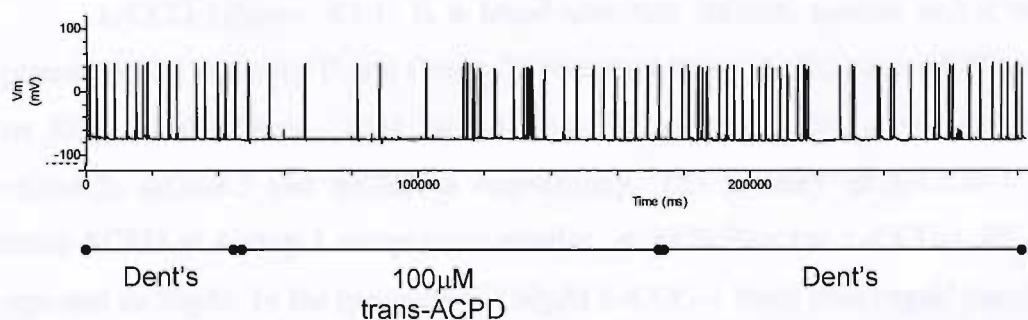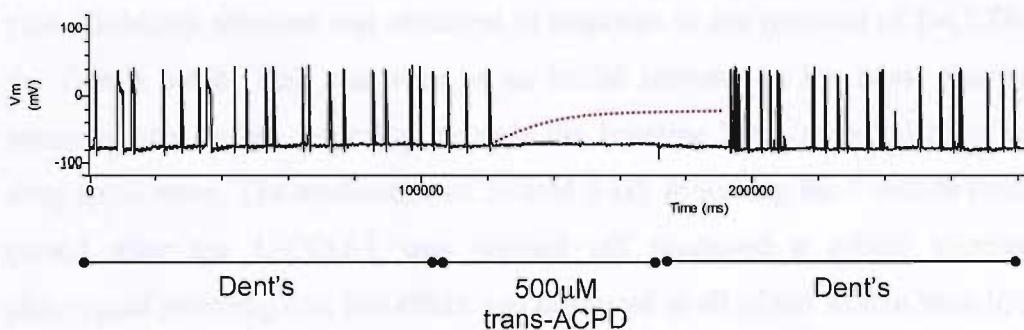
Figure 4.5. Examples of intracellular recordings made from the terminal bulb of wild-type adult hermaphrodites in the presence of (1S,3R)*trans*-ACPD. (A) 100μM (1S,3R)*trans*-ACPD **(B)** 500μM (1S,3R)*trans*-ACPD. Recordings demonstrate a depolarization block of the pharyngeal muscle does not occur when the agonist is applied. If the muscle were depolarized, the period of agonist application would be associated with the resting membrane potential becoming more positive (indicated by the red dashed line in B), but this is not seen and instead the resting membrane potential remains constant.

4.2.3 The inhibition of basal pharyngeal pumping by L-CCG-I requires *mgl-1*

L-CCG-1 *(figure 4.6A)* is a broad-spectrum mGluR agonist and a more potent agonist at Group II and Group III receptors than (1S,3R)*trans*-ACPD, with an EC_{50} of 0.3-04μM, 1μM and 9-50μM at representative group members mGluR2, mGluR3 and mGluR4a respectively. The potency of L-CCG-1 and *trans*-ACPD at Group I receptors is similar, at mGluR1a the L-CCG-1 EC_{50} is reported as 50μM. In the presence of 100μM L-CCG-1 basal pharyngeal pumping was inhibited in wild type animals, this effect was reversible. In some cases a post-inhibitory rebound was observed in response to the removal of L-CCG-1 by the Dent's wash. This was seen as an initial increase in the basal pharyngeal pumping rate before returning towards the baseline level recorded prior to the drug application. The application of 500nM 5-HT following the 5 minute recovery period after the L-CCG-1 was washed off produced a robust increase in pharyngeal pumping and this effect was observed in all of the worms tested*(figure 4.6B)*.

EPG recordings from *mgl-1(tm1811)* identified the inhibition of basal pharyngeal pumping by 100μM L-CCG-1 was reduced in mutant animals. The mean number of pharyngeal pumps recorded over 5 minutes in the presence of L-CCG-1 was ~60% of the mean number of pumps recorded in the 5 minutes prior to the application of L-CCG-1. In common with wild-type worms, some of the *mgl-1(tm1811)* animals tested exhibited an increase in the basal rate of pharyngeal pumping after the drug was washed off. The application of 500nM 5-HT to *mgl-1(tm1811)* animals, after the L-CCG-1 had been washed off and after a 5 minute recovery in Dent's, produced the same robust increase in pharyngeal pumping as that observed in wild type worms. The 5-HT stimulation of rapid pharyngeal pumping was observed in all *mgl-1(tm1811)* animals tested *(figure 4.6B)*.

Inhibitory screening data recorded from the EPG assay were confirmed by
intracellular recordings of the pharynx. L-CCG-1 was applied
with the EPG recordings the L-ccg-1 study was confirmed as a post-inhibitory
rebound effect, marked by the reduced frequency of action potentials. The
further intracellular recordings of the pharynx were influenced by the
L-CCG-1, indicating the inhibitory effect...
caused by a glutamate-gate anion.

A   B

C   D

Figure 4.6. L-CCG-1 screening of wild-type and *mgl-1(tm1811)* worms. A) The molecular
structure of L-CCG-I. **B)** EPG recordings made from wild-type and *mgl-1(tm1811)* animals. Note
the pharynx of *mgl-1(tm1811)* animals continue to pump in the presence of 100μM L-CCG-I. The
red bars indicate the post-inhibitory rebound observed after the drug was washed off. **C)** An EPG
of a single pharyngeal pump by an *mgl-1(tm1811)* animal in the presence of 100μM L-CCG-I **D)**
A scattergram comparing the effect of 100μM L-CCG-I on basal pharyngeal pumping by wild-
type (n=15) and *mgl-1(tm1811)* (n=18) animals. Each point is the result of an individual animal.

4.2.5 Intracellular recording of the L-CCG-I inhibition of basal pharyngeal pumping by wild-type animals

Intracellular recordings were made from the terminal bulb of wild-type adult pharynxes and the application of 10µM L-CCG-I caused a cessation of action potential firing *(figure 4.7)*. This inhibition was reversible and as observed with the EPG recordings the Dent's wash was followed by a post-inhibitory rebound effect, marked by the increased frequency of action potential firing. The resting membrane potential (-80mV) was unaffected by the application of 10µM L-CCG-I, indicating the inhibition of basal pharyngeal pumping is not being caused by a depolarization block .
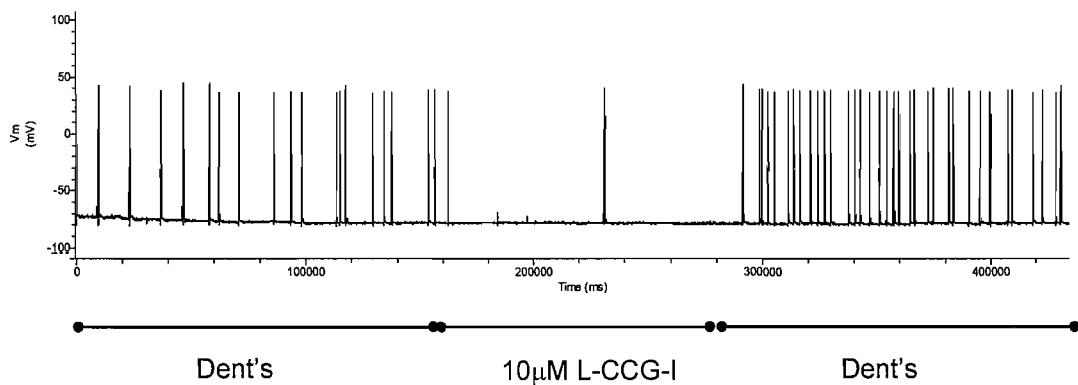
Figure 4.7 An intracellular recording made from a wild-type worm in the presence of 10µM L-CCG-I. The application of 10µM L-CCG-I to wild-type worms inhibited post-synaptic potentials without causing a change in the membrane potential.

4.2.6 Basal pharyngeal pumping is inhibited in a dose-dependent manner by L-CCG-I in wild-type animals and this response involves *mgl-1*

Since L-CCG-1 was identified as a more potent inhibitor of basal pharyngeal pumping than (1S,3R)*trans*-ACPD and its pharmacological profile in the pharyngeal pumping assay was subsequently characterised. The dose response analysis was performed by taking single worms and sequentially applying increasing concentrations of L-CCG-I, over the range $1 \mu M$ $-100 \mu M$. The drug was applied for 2.5 minutes, with a 5 minute recovery period in Dent's between each drug application *(figure 4.9)*. The change in pharyngeal pumping, in the presence of the drug, was calculated as a percentage of the number of basal pharyngeal pumps in Dent's. By doing so the drug mediated effect in each animals was normalised to the basal pumping level for each individual animal. The percentage inhibition of pharyngeal pumping caused by the drug was

$$\left(\frac{\text{Number of pumps in 2.5 min in DRUG} - \text{Number of pumps in 2.5min in Dent's}}{\text{Number of pumps in 2.5min in Dent's}} \right) \times 100$$

calculated as follows:

The number of pumps in 2.5 minutes in Dent's corresponds to the second half of the five minute recovery period, before the drug was applied.

In the presence of increasing concentrations of L-CCG-1 a direct increase in the percentage inhibition of the basal pharyngeal pumping of wild-type worms was recorded. The EC_{50} of L-CCG-1 in wild-type animals was $3 \mu M$. The basal pharyngeal pumping of *mgl-1(tm1811)* animals was not inhibited by concentrations of L-CCG-1 in the range $1 \mu M$-$12.5 \mu M$. At $25 \mu M$ there was an increase in the percentage inhibition of the basal pharyngeal pumping of *mgl-1(tm1811)* animals. At $25 \mu M$ the percentage inhibition of basal pharyngeal pumping plateaud at ~44%, increasing the concentration of L-CCG-1 to $100 \mu M$ did not produce any further increase of the inhibition of basal pharyngeal pumping *(figure 4.8)*.
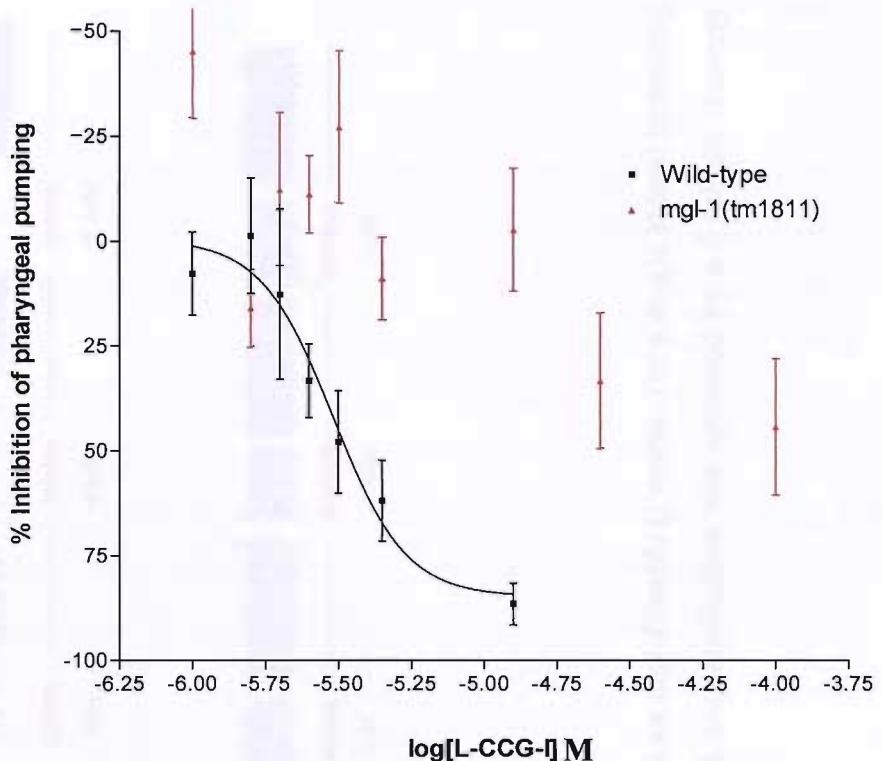
Figure 4.8. Dose-response curve for the % inhibition of basal pharyngeal pumping by wild-type and *mgl-1(tm1811)* worms in the presence of **L-CCG-1**. Wt (IC$_{50}$ 3μM $\pm$ 12.2) (n=9) and *mgl-1(tm1811)* (n=12). Each point represents the mean % inhibition of the basal pumping, $\pm$ S.E.M).

A. Wild-type

1µM 1.5µM 2µM 2.5µM 3µM 4.5µM 12.5µM

B. *mgl-1(tm1811)*

1µM 1.5µM 2µM 2.5µM 3µM 4.5µM 12µM

25µM 100µM
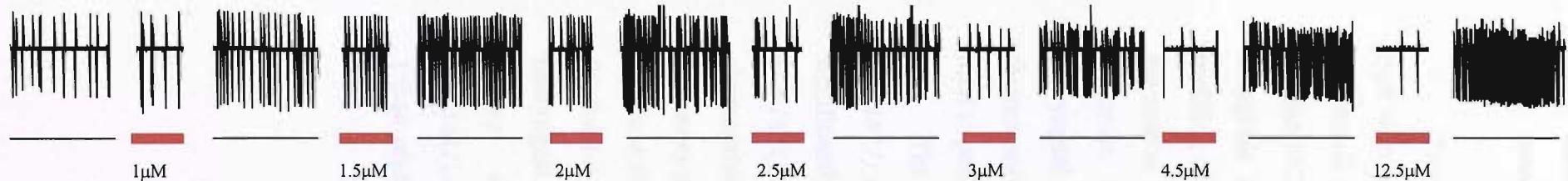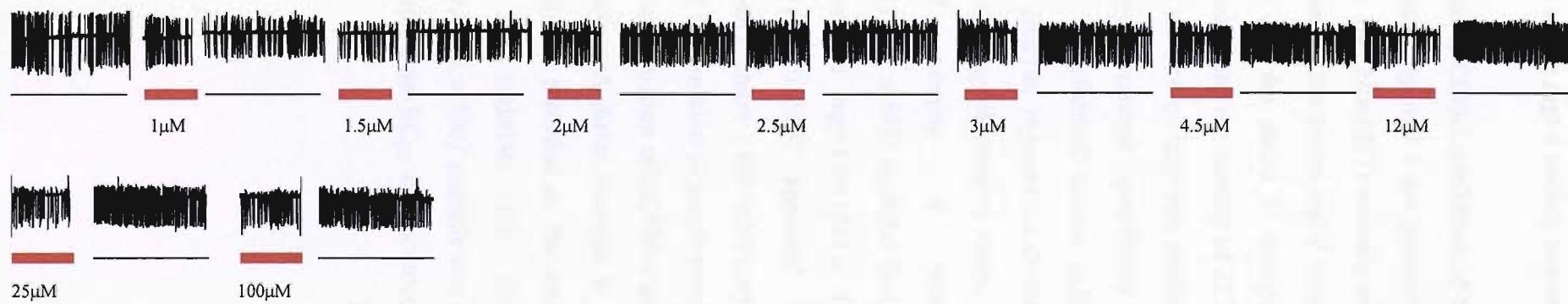
Figure 4.9. An L-CCG-1 dose response trace recorded from a wild-type worm and an *mgl-1(tm1811)* worm. For a single animal increasing concentrations of L-CCG-I were sequentially applied to the recording chamber. Each drug application was separated by a 5 minute recovery period in Dent's saline (black line). Drugs were applied for 2.5 minutes (red line).

167

4.2.7 Rescue of the L-CCG-1 response in *mgl-1(tm1811)* animals by re-introduction of the *mgl-1* coding sequence

To confirm the L-CCG-1 inhibition of basal pharyngeal pumping of wild type worms is mediated by *mgl-1* the genomic coding sequence of *mgl-1* was introduced back into *mgl-1(tm1811)* animals as an extrachromosomal array. The cosmid ZC506.4 contains the entire *mgl-1* coding region, including 3.6Kb of 5' intergenic sequence and the entire 3' intergenic sequence. Restriction digest profiling and PCR confirmed the identity of ZC506 before it was co-injected with the marker pPD118.33. The co-injection marker consists of GFP and the *myo-2* promoter. The *myo-2* promoter specifically drives GFP expression in the pharyngeal muscle. Transformed worms exhibiting bright green, non-mosaic fluorescence in the pharynx represented co-transformants for ZC506 and were used to perform the pharyngeal pumping assay.

The basal number of pharyngeal pumps by *mgl-1(tm1811);Ex[pPD118.33;zc506]* over the first 5 minute period in Dents was not significantly different to *mgl-1(tm1811)*. The EPGs recorded from *mgl-1(tm1811);Ex[pPD118.33;zc506]* appeared normal and consisted of the characteristic excitatory phase , the inhibitory post-synaptic potentials and the relaxatory phase and the duration of single pumps was similar to wild type worms *(figure 4.10A)*. The introduction of ZC506.4 into *mgl-1(tm1811)* animals rescued the wild-type phenotype. A direct increase in the percentage inhibition of basal pharyngeal pumping was recorded as the concentration of L-CCG-1 increased *(figure 4.10B)*. In addition the pharyngeal pumping of *mgl-1(tm1811);Ex[pPD118.33;zc506]* animals was slightly more sensitive to L-CCG-1 than wild-type animals, the EC_{50} was recorded as 2.4μM *(figure 4.10C)*.

A

B

| Dents | 1µM
L-CCG-1 | Dents | 3µM
L-CCG-1 | Dents | 10µM
L-CCG-1 | Dents |

C

Figure 4.10 Genomic rescue of the L-CCG-I response in *mgl-1(tm1881)* **animals A)** A recording of a single pharyngeal pump from an *mgl-1(tm1811);Ex[pPD118.33;zc506]* hermaphrodite. **B)** A trace recorded from a single *mgl-1(tm1811);Ex[pPD118.33;zc506]* hermaphrodite in the presence of increasing concentrations of L-CCG-1: 1µM, 3µM and 10µM. **C)** A dose response curve for *mgl-1(tm1811);Ex[pPD118.33;zc506]* and L-CCG-1. The data plotted is the mean % inhibition of basal pharyngeal pumping $\pm$ SEM, recorded from two independent lines (n = 5/line). The wild-type and mutant dose response data in the same drug range is included on the graph for comparison.

169

### 4.2.8 *mgl*-2 and *mgl*-3 mutant responses to L-CCG-1

Analysis of transgenic lines expressing an *mgl-2* gene fusion reporter construct did not identify this receptor as being expressed in the pharyngeal nervous system and the expression of *mgl-3* remains to be fully characterized. To further investigate if either *mgl-2* or *mgl-3* receptor subtypes contribute to the L-CCG-1 modulation of basal pharyngeal pumping, EPG recordings were made from *mgl-2* and *mgl-3* mutants. The strain *mgl-2(tm355)* has a deletion mutation in *mgl-2* that generates a disrupted gene that enocodes a protein lacking the N-terminal ligand binding domain. Worms that are homozygous for the mutation are viable and were used in the L-CCG-1 pumping assay. The allele *mgl-3(tm1766)* is a null mutation in the *mgl-3* receptor. Worms that are homozygous for the mutation are viable and were used in the L-CCG-1 pharyngeal pumping assay.

An expanded view of a single pump by each mutant in Dent's shows the characteristic E, P and R phases. The glutamate mediated IPSPs appear as normal and the pump duration is normal compared to wild-type animals *(figure 4.11B and 4.12B)*. Subsequently recordings were made from the strains *mgl-2(tm355)* and *mgl-3(tm1766)* in the presence of increasing concentrations of L-CCG-1 *(figure 4.11A and 4.12A)*, using the protocol previously described *(section 4.2.6)*.

The response of basal pharyngeal pumping by both mutant strains showed the same general trend, which was inhibition by L-CCG-I, however this was also assosciated with subtle differences. A direct increase in the percentage inhibition of *mgl-2(tm355)* animals was recorded in response to increasing concentrations of L-CCG-I above and including 3μM and the inhibition was reversible. The percentage inhibition achieved at 3μM and 10μM concentrations tested was similar to wild-type animals, whereas at 1μM L-CCG-I basal pumping was more similar to the *mgl-1* mutant because it increased by 25% *(figure 4.11C)*.

The dose-response of *mgl-3(tm1766)* worms to L-CCG-1 had a similar trend, at concentrations of 3μM and above basal pharyngeal pumping was inhibited and this effect was reversible with washing. However, unlike both *mgl-1(tm1811)* and *mgl-2(tm355)* at the lowest concentration of L-CCG-I tested, which was 1μM, the basal pharyngeal pumping of the *mgl-3* mutant was inhibited. In addition there was very little increase (3%) in the percentage inhibition of basal

pharyngeal pumping when the concentration of L-CCG-I was increased from 1 μM to 3 μM *(figure 4.12C)*.

A

B

180 msec
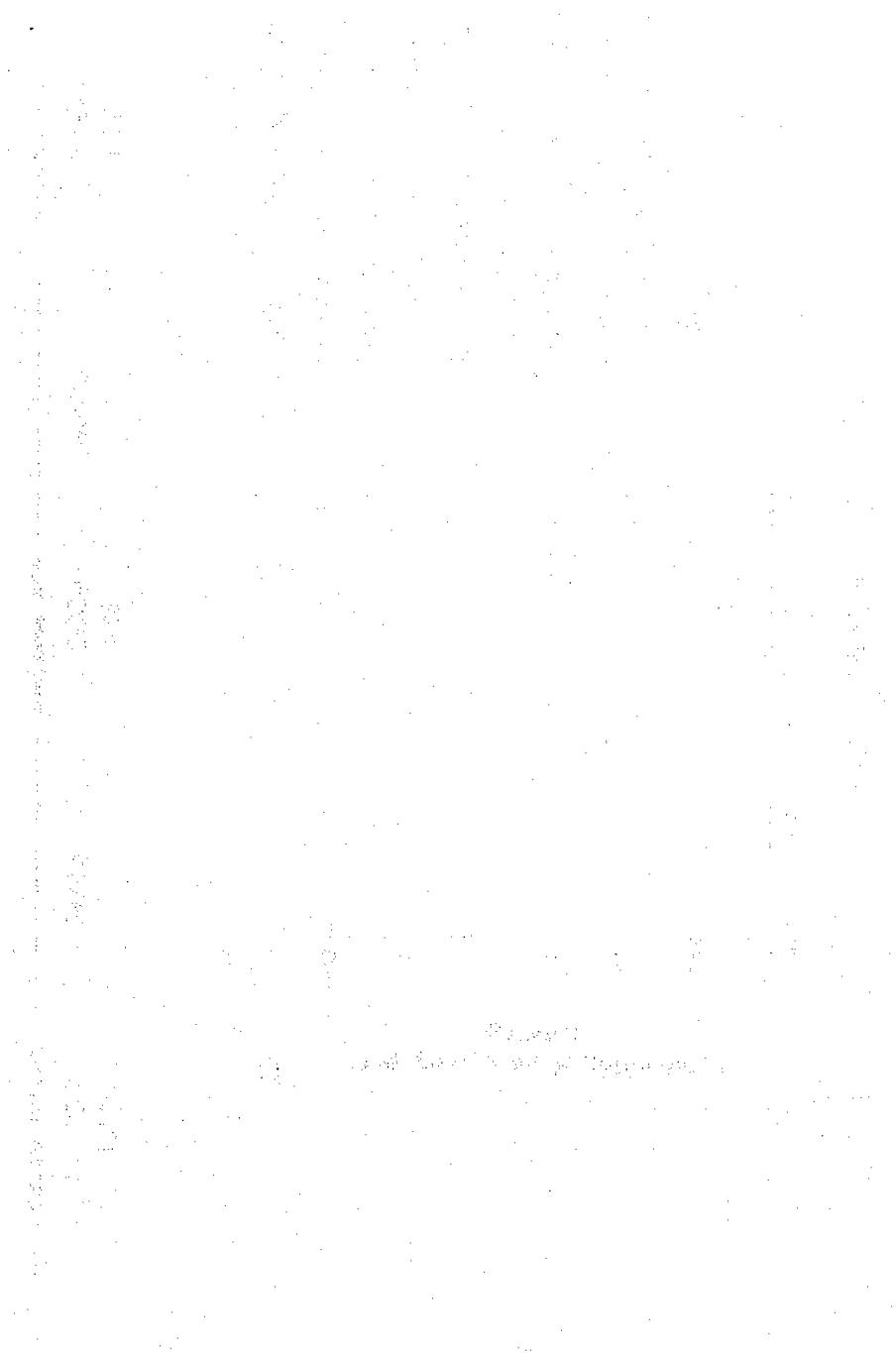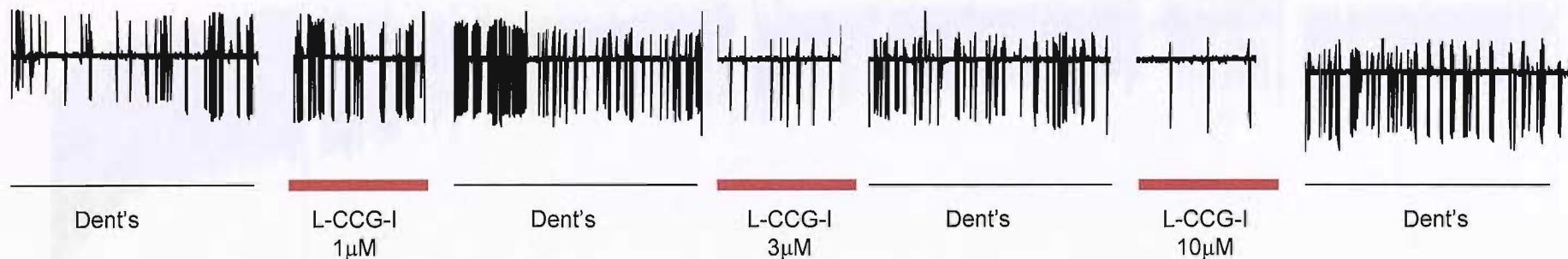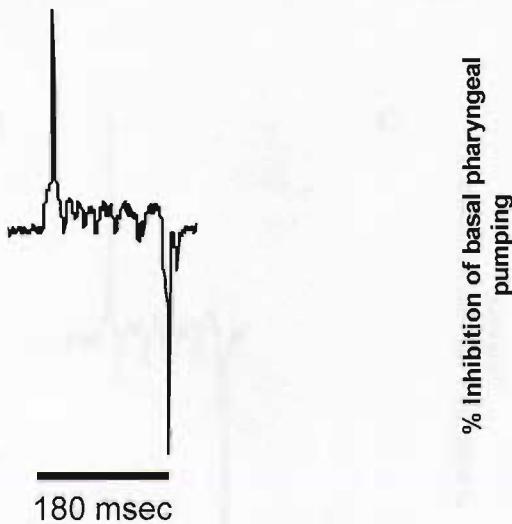
C
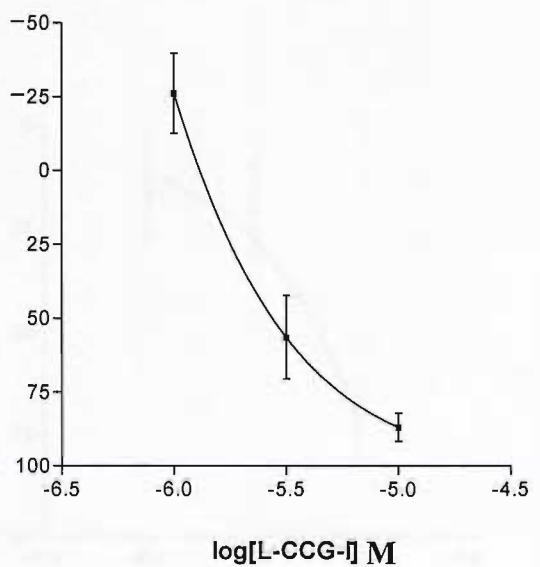
Figure 4.11 The *mgl-2(tm355)* response to L-CCG-I A) A recording made from an *mgl-2(tm355)* worm in the presence of increasing concentrations of L-CCG-I. Drugs were applied for 2.5 minutes (red bar) with a 5 minute recovery period in Dent's (black line) between each drug application. B) An EPG of a single pump recorded in Dent's from the *mgl-2* mutant. C) A graph of the percentage inhibition of basal pharyngeal pumping by *mgl-2(tm355)* animals in the presence of increasing concentrations of L-CCG-I (n=6). Each point represents the mean % inhibition +/- SEM.

172
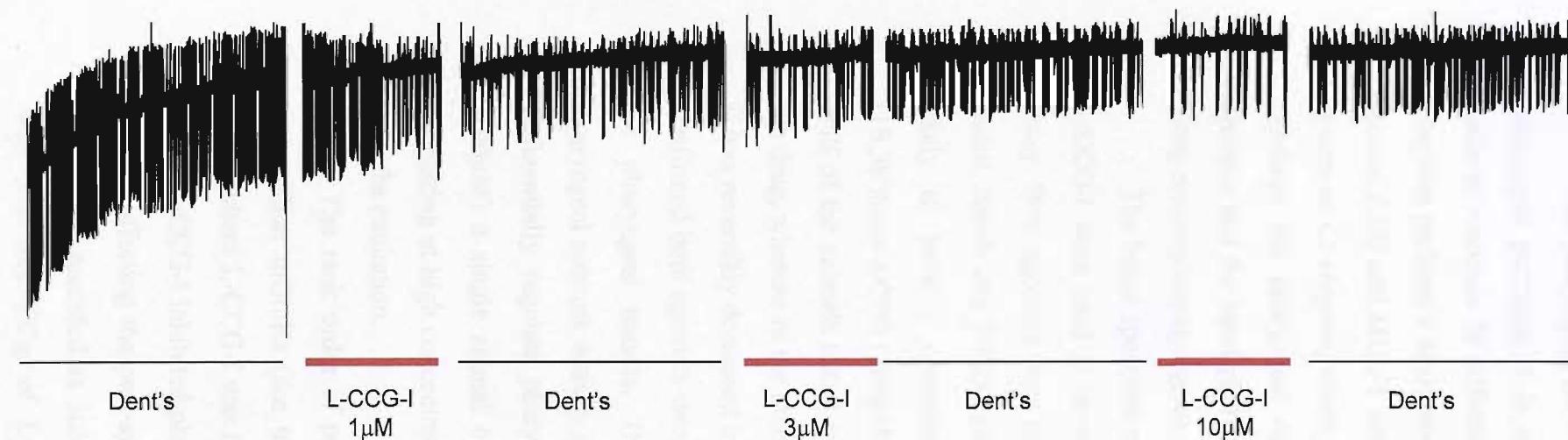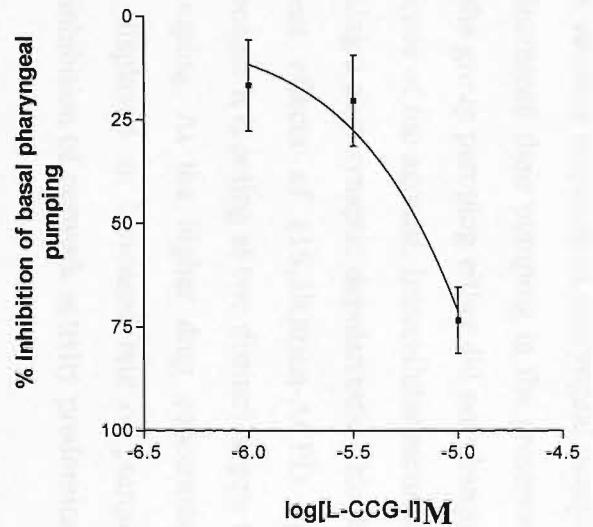
Figure 4.12 The *mgl-3(tm1766)* response to L-CCG-I A) A recording made from an *mgl-3(tm1766)* worm in the presence of increasing concentrations of L-CCG-I. Drugs were applied for 2.5 minutes (red bar) with a 5 minute recovery period in Dent's (black line) between each drug application. **B)** An EPG of a single pump recorded in Dent's from the *mgl-3* mutant. **C)** A graph of the percentage inhibition of basal pharyngeal pumping by *mgl-3(tm1766)* animals in the presence of increasing concentrations of L-CCG-I (n=6). Each point represents the mean % inhibition +/- SEM.

173

4.3   Discussion

The pharyngeal nervous system is known to be important for regulating pharyngeal pumping, it is especially important for modulating how the worm feeds in response to different environmental cues. The metabotropic glutamate receptors perform a neuromodulatory role in the mammalian nervous system *(see section 1.19)* and MGL-1 may perfo.rm a similar role in the pharyngeal nervous system of *C. elegans*, where it has a widespread expression. To assay if MGL-1 regulates the pharyngeal circuitry mammalian agonists were applied to the pharynx and the basal pharyngeal pumping was measured electrophysiologically, using electropharyngeogram recordings.

The broad spectrum mammalian mGluR agonists (1S,3R)trans-ACPD and L-CCG-I were used to investigate MGL signalling in the pharyngeal pumping assay. Both agonists were predominantly inhibitory at the highest concentration tested, which was 500μM and 100μM respectively, suggesting that each agent is likely to have a common mode of action. The lower concentration of (1S,3R)trans-ACPD (100μM) caused a variable response in pharyngeal pumping. Half of the animals tested reversibly increased their pumping in the presence of the drug, whereas in the other half of the group pumping either did not change or it was reversibly decreased in the presence of the agonist. Intracellular recordings confirmed both agonists were not causing a post-synaptic depolarization block of the pharyngeal muscle. The different effects of (1S,3R)trans-ACPD upon pharyngeal network activity may be because it is acting at two distinct targets that differentially regulate pharyngeal pumping. At the higher drug concentration (500μM) a single animal out of 22 displayed an increased rate of pumping, suggesting at high concentrations the inhibition of network activity predominates over the excitation.

The rank order of potency of L-CGG-I and (1S,3R)trans-ACPD at the mammalian mGluRs (see table 2) was conserved in the pharyngeal pumping assay, where L-CCG-I was the more potent inhibitor of pharyngeal pumping. At 10μM L-CCG-I inhibited pharyngeal pumping by 80-100% in wild-type animals, without affecting the post-synaptic resting membrane potential. The EC_{50} of L-CCG-I was recorded as 3μM in the EPG dose response analysis, which agrees well with the EC_{50} of L-CGG-I established at other invertebrate mGluR

homologues, DmGluRA EC_{50} = 8.95µM and the mammalian Group II receptor (Parmentier, M-L et al 2000), mGluR3 EC_{50} = 1µM (Conn PJ and Pinn J-P 1997), and mGluR3 belongs to the subgroup MGL-1 has the closest amino acid identity to.

Analysis of the *mgl-1(tm1811)* mutant, which is a putative null, suggests the inhibitory effect of L-CCG-I and (1S,3R)trans-ACPD are predominantly mediated by the MGL-1 receptor subtype. In the presence of increasing concentrations of L-CCG-I, within the range 1-12.5µM, pharyngeal pumping by *mgl-1* mutants was resistant. The inhibitory effect of L-CCG-I in the range 1-12.5µM could be rescued by re-introducing the *mgl-1* gene into the *mgl-1* mutant background. This argues the inhibitory effect of L-CCG-I is mediated by *mgl*-1. At 25µM and 100µM L-CCG-I began to inhibit the pharyngeal pumping by *mgl-1* mutants, and this plateaud at 44% inhibition. This may be due to signalling through a second inhibitory site, with a low affinity for L-CCG-I.

Basal pharyngeal pumping by *mgl-1* mutant animals was otherwise normal, as was the response of the *mgl-1* mutant to 500nM 5-HT. This suggests MGL-1 is not essential for basal pharyngeal pumping, however, when activated MGL-1 can modulate pharyngeal pumping. Subsequently MGL-1 may function to modulate the network activity of the pharyngeal nervous system and could regulate pharyngeal pumping under non-basal conditions, such as when the worm must modify its feeding strategy during changes to its environment.

In wild-type animals the application of high concentrations of (1S,3R)trans-ACPD and L-CCG-I inhibits pumping, but with a variable effectiveness. In a minority of wild-type animals the pharynx continued to pump but at a lower rate, when in the presence of (1S,3R)trans-ACPD. Comparison of pumps performed in the presence and absence of (1S,3R)trans-ACPD did not identify any differences, in the EPG waveform or the pump duration. This may instead reflect the low potency of (1S,3R)trans-ACPD, since the more potent agonist in the mammalian systyem, L-CCG-I, was a more efficient inhibitor.

The basal pharyngeal pumping of *mgl-2* and *mgl-3* mutants was normal. The reporter construct *mgl-2::GFP* was not identified in the pharynx, which is consistent with the normal pumping behaviour displayed by the *mgl-2* mutant animals. In addition pumping by the *mgl-2* mutant was inhibited by L-CCG-I at

3µM and 10µM, with a greater than 80% inhibition of basal pumping recorded at 10µM, suggesting that this receptor is not responsible for the inhibitory response to L-CCG-I. The expression of *mgl-3* has not been fully described and since the basal pharyngeal pumping of the *mgl-3* mutant was inhibited by L-CCG-I at 1, 3 and 10µM it is unlikely to be involved in the inhibitory response. However, at the lowest concentration of L-CCG-I used (1µM) the pharyngeal pumping of both *mgl-1* and *mgl-2* mutants was increased above the basal pharyngeal pumping rate. This response was not seen in the *mgl-3* mutant, since at the same concentration pharyngeal pumping was inhibited to about the same extent as when 3µM L-CCG-I was applied. This would suggest that at the lower concentrations of L-CCG-I the MGL-3 receptor subtype could be responsible for the upregulation of pumping. This would be consistent with the dual response recorded in the presence of 100µM (1S,3R)trans-ACPD and the increased pharyngeal pumping may be due to (1S,3R)trans-ACPD activating the MGL-3 receptor to produce an upregulation of pharyngeal pumping. This highlights how the responses to both L-CCG-I and (1S,3R)trans-ACPD are complex and since both are broad spectrum agonists there may be activity at other *mgl* receptor subtypes, other than *mgl-1*. By crossing *mgl* subtype mutants together a number of different strains could be produced, that are defficient for any combination of the three *mgl* subtypes. This would allow the function of individual receptor subtypes to be studied in isolation, without interferrence from other subtypes.

Both the *mgl-1* mutant and wild-type animals exhibited a rebound excitation after agonist was washed off. This was observed as rapid action potential firing (in intracellular recordings) and rapid pumping (in EPG recordings) immediately after removal of the drug. Since the rebound excitation was common to both the wild-type and *mgl-1* mutant animals it is likely to involve the drug acting upon a pathway that is independent of *mgl-1* signalling.

The EPG can provide information about the activity of the glutamatergic motorneuron M3 *(see section 4.1.5)*. In animals where M3 transmission is disrupted the IPSP transients that occur in the plateau (P) phase are absent and the duration of the EPG is doubled from ~200ms to ~400ms [Raizen, D.M. and Avery, L 1989]. The two mammalian agonists did not disrupt M3 transmission to the pharyngeal muscle in wild-type animals because the P-phase transients and the

duration of the EPG were not affected by applying the agonists. Furthermore, M3 transmission to the pharyngeal muscle is not affected by the absence of individual *mgl* subtypes and it is therefore likely that M3 synaptic inputs to the pharyngeal muscle do not require *mgl* receptor subtypes for normal function.

Glutamate and ivermectin have been shown to inhibit the pharynx, eliciting a chloride-dependent, depolarizing block (Pemberton DJ et al 2001). Intracellular recordings from the pharynx of wild-type animals showed the resting membrane potential is not affected by either 10μM L-CCG-I or 500μM (1S,3R)trans-ACPD, but pharyngeal action potentials were inhibited as was pumping. This supports the view that L-CCG-I and (1S,3R)trans-ACPD do not inhibit the pharynx by stimulating a glutamatergic depolarization block. Furthermore because the pharynx of wild-type animals is almost completely inhibited by L-CCG-I it is unlikely it is doing this by inhibiting glutamatergic transmission, because as described earlier mutants where glutamate signalling in the pharynx is inhibited (i.e *eat-4* (Lee RY et al 1999), *avr-15* (Dent et al 1997), M3 laser ablation (Heyi Li et al 1997)) have a distinctive EPG profile and can still pump but with a longer pharyngeal muscle contraction.

The motorneuron MC is the main excitatory neuron in the pharynx, its firing triggers pharyngeal muscle action potentials by releasing acetylcholine that acts on the muscle nicotinic receptor, *eat-2* (Raizen D.M et al 1995). The function of *eat-2* requires *eat-18* (McKay JP et al 2004). In mutants where cholinergic transmission is disrupted, such as *eat-18* mutants (Niacaris T and Avery L 2003) and MC ablated animals pumping still occurs, but at a lower rate (approximately one fith of wild-type) (Avery, L and Horvit, R et al 1989). This supports the argument that L-CCG-I is acting on multiple transmitter pathways, as opposed to a single pathway, to inhibit pumping.

Normal feeding involves the co-ordination of two processes, pharyngeal pumping and peristalsis *(see section 4.1.2)*. The cholinergic neuron M4 is required for isthmus peristalsis to occur (Avery, L and Horvitz, H.R. 1987) and worms with a laser ablated M4 do not grow (Avery, L., Horvitz, H.R. 1989) but do continue to pump (Avery, L and Horvitz, H.R. 1987). M4 synapses posteriorly onto the pm5 muscle cells to control peristalsis of the isthmus but M4 postsynaptic potentials are not detected in the EPG (Avery, L and Thomas, J.H.

1997). In Chapter 5, it is suggested MGL-1::GFP is expressed in the M4 motorneuron and therefore *mgl-1* may regulate the neural control of persitalsis, however, this cannot be determined from the EPG analysis.

As previously described (see Chapter 5) MGL-1::GFP was identified as being expressed in the motorneuron NSM and there is evidence that NSM is capable of releasing transmitters, including glutamate, 5-HT and neuropeptides (Table 4.1). It has been sugested 5-HT underlies neural mechanisms that are involved in the adaptation of pharyngeal behaviour in the presence and absence of food. 5-HT stimulates rapid pumping by both reducing the duration of action potentials and increasing their frequency in the pharyngeal muscle cells. These effects are dependent upon 5-HT acting on MC and M3 motorneurons (Niacaris, T and Avery, L 2002) . NSM and I5 are the only source of 5-HT in the pharyngeal nervous system, however, laser ablation of the NSM cells does not cause a complete inhibition of basal pharyngeal pumping, instead it is reduced by only 11% (in *unc-31(e928)* animals that diaply constitutive pumping) (Lee, RY et al 1999). MGL-1 may therefore act presynaptically to inhibit 5-HT release from NSM, to inhibit the activity of the pharyngeal network. Of course, NSM expresses other transmitters and so it is possible MGL-1 activation may regulate the release of any of these also to inhibit network activity.

As shown by intracellular recordings the resting membrane potential of pharyngeal muscle is not affected by the presence of either agonist, which suggests the mechanism of inhibition is not a direct effect on the pharyngeal muscle cells but instead it is an effect on the pharyngeal nervous system. Furthermore *mgl-1* is not expressed in the pharyngeal muscle, therefore the activation of MGL-1 affects network signalling in the pharyngeal nervous system.

In summary the pharynx responds to metabotropic glutamate agonists and this response is lost in the *mgl-1* knockout, supporting the view that MGL-1 is a metabotropic glutamate receptor. Furthermore this shows that the pharyngeal preparation provides a bioassay for studying *mgl-1* receptor function and therefore be exploited to study the role of scaffolding proteins.

CHAPTER 5

A yeast-2-hybrid screen to identify MGL scaffolding proteins

5.1   Introduction

The yeast-two-hybrid system is a genetic technique that allows the detection of protein-protein interactions *in vitro*. It is an established technique for the detection of protein-protein interaction and suitable for large scale high-throughput screening. It has been used to generate large protein-protein interaction networks (the interactome) within *C.elegans* (Vidal, M et al 2004), *S.cerevisiae* (Uetz, P et al 2000) and the bacterium *Helicobacter pylori* (Rain, J.C et al 2001).

Variations of the yeast-2-hybrid technique exist (Chien, CT et al 1991 and Mendelsohn, AR & Brent, R 1994), but they are all based upon the same principle. At its core the yeast-2-hybrid system is based on the reconstitution of a functional transcription factor by the interaction of two proteins. Eukaryotic transcription factors are comprised of a DNA Binding Domain (DBD) and Activation Domain (AD). The DBD recognises and binds to specific promoter sequences while the AD directs RNA polymerase II transcription of downstream genes. This allows two different hybrid proteins to be made, a protein (X) fusion to the DBD and a protein (Y) fusion to the AD. The interaction between proteins X and Y brings the AD and DBD together in the nucleus; this is detected by the expression of specific reporter genes with promoters that are inducible by that transcription factor. The X-DBD hybrid is commonly referred to as the 'BAIT' whereas the Y-AD hybrid is the 'PREY'.

5.2   Variations upon the 2-hybrid theme.

Variations of the 2-hybrid system have been used to gain alternative insights into protein-protein interactions. The reverse hybrid system (Leanna C.A. & Hannink M 1996) allows events that disrupt protein-protein interactions to be selected for. These disruptive events can either be mutations, drugs or competing proteins. The reverse hybrid system utilises a yeast strain in which the expression of two interacting proteins induces the transcription of a toxic marker. In yeast where this interaction is disrupted the toxic marker is not expressed and instead these colonies are able to grow, thereby facilitating detection. The 3-hybrid system has been used to study macromolecular interactions consisting of three or more proteins (Zhang, J., Lautar, S. 1996). As the 2-hybrid is a heterologous system some types of post-translational modification, necessary for an interaction

to occur in the homologous system, do not take place in yeast. The kinase 3-hybrid system addresses this and allows the detection of protein-protein interactions dependent upon tyrosine phosphorylation (Osborne, M.A. & Dalton, D et al 1995). This is a type of post-translational modification intrinsic to a number of signalling cascades. In the kinase 3-hybrid a mono-specific cytosolic tyrosine kinase is introduced into the yeast, together with the 2-hybrid components. The protein 3-hybrid system is an extension of the 2-hybrid system, whereby a third protein is required to direct the interaction between two proteins. The third protein achieves this by either directly mediating the interaction or causing a conformational change in one of the other two proteins that enables the interaction to occur (Zhang, J & Lautar, S 1996).

Alternative yeast-based systems to the 2-hybrid have been developed to detect protein-protein interactions. These include, for example, the reverse Ras recruitment system (Stagljar, I & Fields, S 2002), G-protein fusion technology (Ehrhard, K.N et al 2000) and the split-ubiquitin assay (Stagljar, I. et al 1998). These systems have been developed to identify interactions involving proteins that are membrane bound, such as receptors. Full-length receptors make poor candidates for study using the 2-hybrid system. Primarily because the 2-hybrid system relies upon the interaction occurring in the yeast nucleus, a poor environment for the correct post-translational modification of membrane bound proteins. This is often why a smaller portion of the receptor, corresponding to a functionally important domain is used instead. For example, the receptor intracellular C-terminal is an important domain that underlies receptor function *(see Introduction section 1.15, 1.19 and 1.22)* and it is ideally positioned on the intra\cellular side of the membrane to mediate protein-protein interactions, which is why it is often used to screen with using the yeast-2-hybrid system (for example Flajolet, M. et al 2003).

5.3   The proteomic approach to characterizing signal complexes.

Protein interaction networks are being dissected using proteomic methods as well. These are broadly based on mass spectrometry and biochemical techniques such as protein affinity chromatography, immunoprecipitation, two-dimensional gels and affinity immunoblotting. The NMDA receptor complex was characterised in this way (Husi, H & Ward, M.A. 2000). It was identified as being

comprised of 77 proteins that are classified into adaptor, signalling, receptor, cytoskeletal and novel proteins. Similar proteomic studies have been conducted on the C-terminal domain of the 5-HT$_{2C}$ GPCR (Becamel, C & Galeotti, N et al 2002) and the mGluR5 GPCR (Farr, C.D. & Gafken, P.R. et al 2004). In the case of the mGluR5 receptor the approach used revealed novel protein interactions, previously unidentified as components of the signalling complex. This highlights receptor signalling complexes can be extensive and a variety of techniques are required to characterise them fully.

The most appealing quality of using the yeast-2-hybrid is that it allows direct access to the cDNA encoding the interacting protein. Additionally there is minimal requirement for the initiation of a 2-hybrid screen. Only the full-length or partial cDNA of the gene of interest is required and the signal amplification produced by the reporter gene system allows weak or transient interactions to be detected. Biochemical approaches can be dependent on good quality antibodies, sometimes-high amounts of purified protein and often protocols require specific optimisation.

5.4   Overview of the LexA Yeast-2-Hybrid Approach

The LexA system is a variant of the original yeast-2-hybrid approach and the main difference is that the DBD and AD are derived from prokaryotic proteins as opposed to eukaryotic proteins (van Criekinge, W and Beyaert, R. 1999). In the LexA system the AD is comprised of an 88-residue acidic blob domain. The DBD is comprised of the bacterial LexA protein and specifically binds to LexA operators upstream of the two reporter genes, *leucine* and *lac-Z*. The *leucine* reporter gene is integrated into the host strain genome and the *lacZ* reporter resides on a plasmid (pSH18-34). Reporter gene expression can be assayed with the appropriate selection media. *Leucine* expression enables growth on *leucine* deficient media and *lac-z* expression gives rise to β-galactosidase, which turns the colourless substrate X-Gal blue.

The LexA system provides several advantages over other yeast-2-hybrid systems. The inducible expression of the DNABD-fusion and AD-fusion proteins in the presence of galactose, conferred by the GAL1 promoter, reduces the threat of potentially toxic foreign fusion proteins to the yeast host cell. This

subsequently prevents their elimination from the protein pool being screened. Both BAIT and PREY plasmid copy number can be amplified *in vivo* on glucose media prior to screening on galactose containing media, increasing the opportunity for detecting protein interactors present at low numbers. Additionally because the *LacZ* reporter gene is located on an autonomously replicating plasmid, its high copy number amplifies signals from weak interactions. This enables β-galactosidase expression to be assayed directly on the culture plate by including the substrate X-Gal in the media (van Criekinge, W and Beyaert, R. 1999).

The yeast-2-hybrid system can generate false positives, these are typically proteins with an intrinsic ability to transactivate the reporter gene expression. Likely candidates include transcription factors and ribosomal subunits. False positives can be eliminated from the pool by performing strict controls, for instance if a true two-hybrid interaction is occurring reporter gene expression will not occur in the absence of the 'BAIT'.

5.5   Organization of the Screen

The selective expression of MGL receptors and the splicing of the C-terminal, together with the presence of protein-protein interaction motifs, indicate this domain is involved in directing MGL receptor scaffolding, targeting and anchoring. Subsequently, the C-terminal domains of the MGL receptor subtypes were utilised in the LexA yeast-2-hybrid system to identify interactors involved in scaffolding their function. The cDNA encoding C-terminal was isolated from the rest of the receptor and used for the screen because in the yeast-2-hybrid system full-length transmembrane receptors are unsuitable candidates.

The organisation of the screen is summarised in *Figure 5.1*. Initial controls were performed using splice variants of MGL-1, MGL-2 and MGL-3 amplified by PCR, to identify suitable bait constructs for screening with. The proceeding stages are designed to manage the screen output and incorporates the cataloguing of clones by PCR amplification, restriction digest analysis and sequencing. Unique clones yielded from the screen were prioritised upon the basis of their known or inferred biological function and their screen profile.

Amplification of the LexA *C.elegans* cDNA library

Construction of the MGL receptor C-terminal DBD-fusion protein 'BAIT'

Controls

Controls determine the suitability of MGL C-terminals for use in a two-hybrid screen. Controls include assessment of the nuclear targeting and expression of the MGL-CT DNABD fusion protein (the BAIT), transactivation of the reporter genes by the BAIT, toxicity to the host yeast cell and plating density.

Screen

The screen will be initially optimised. The MGL CT will be used to screen a *C.elegans* LexA cDNA Library.
Colonies identified under selection for reporter gene expression will be selected at predetermined time points.
Colonies will be subjected to a further four rounds of selection to help eliminate false positives.

Cataloguing of Clones

Cataloguing allows efficient management of large clone numbers and prioritisation. The clones identified from the screen will be catalogued based on :

- strength of interaction with the MGL-CT as determined by reporter gene expression
- size as assessed by PCR amplification from crude yeast extracts
- HAEIII restriction profile

At this stage unique clones will be sequenced, the resulting data can be used for database

Verification

The two-hybrid interaction will be confirmed by assessing any intrinsic ability of the rescued interacting protein to transactivate reporter gene expression.

Figure 5.1. The progression of the yeast-two-hybrid technique for the identification of MGL scaffolding proteins.

5.6   Assessment of the amplified *C.elegans* LexA cDNA library

Before the *C.elegans* LexA cDNA library could be screened it was necessary to amplify it. The amplification of the *C.elegans* LexA cDNA library yielded 17 million clones *(see Section 2.1.8)*. The quality of the amplification was then assessed. The *mgl-3* receptor is representative of a low abundance transcript. Robust bands corresponding to the *mgl-3* receptor C-terminal splice variants were generated from the amplified library (50ng) by PCR *(see below)*.

A.
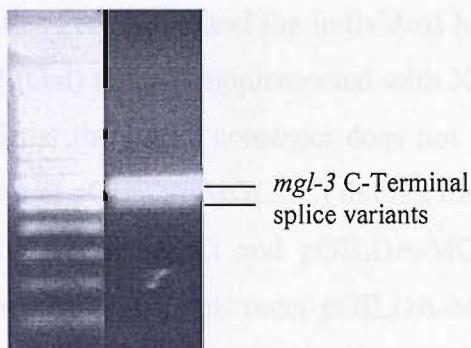
Figure 5.2. PCR amplification of the *mgl-3* C-terminal splice variants from the amplified *C.elegans* cDNA library. *Cycle:* (94°C 2min, 30 cycles of 94°C 15s, 55°C 30s, 68°C 1min and final extension 68°C 10min)

5.7   LexA Yeast-2-Hybrid Control Assays

The intracellular C-terminals MGL-1, MGL-2ii and MGL-3i were each individually fused in-frame to the LexA DNA Binding domain protein (BAIT constructs pGILDA-MGL-1Ct, pGILDA-MGL-2iiCt and pGILDA-MGL-3iCt). Prior to performing a screen of the *C.elegans* LexA cDNA library each of the MGL C-terminal fusion constructs were tested for their suitability.

To assess the nuclear targeting and expression of the MGL 'BAIT' constructs each was transformed into the host yeast strain (EGY48 : Erica Golemis Yeast) with the vector pJK101. The vector pJK101 provides a visual assay for the nuclear targeting and expression of the BAIT fusion protein in EGY48. It contains the *lacZ* reporter gene under the control of the inducible GAL-1 promoter. Two LexA operators are located between the *lacZ* gene and the GAL-1 promoter. The binding of LexA fusion proteins to the LexA operators represses *lacZ* gene expression. This is seen as a repression of the strength of blue colour,

conferred by pJK101 to yeast cells grown on media containing galactose and X-Gal. Each of the MGL 'BAIT' fusion proteins repressed the expression of *lacZ*, confirming each is being nuclear targeted and expressed. The vector pEG202-Max was included as a positive control for the repression assay. It constitutively expresses the LexA-Max fusion protein. The intrinsic ability of the MGL 'BAIT' constructs to transactivate the reporter genes *lacZ* and *leucine* was assessed. Transactivation occurs when the BAIT possesses the intrinsic ability to activate the expression of a reporter gene, in the absence of a 2-hybrid interaction.

Yeast co-transformed with the vector pSH18-34, which contains the reporter gene *lacZ*, and the individual MGL 'BAIT' constructs were spotted onto –HU (Gal) media, supplemented with X-Gal. Yeast growing without turning blue confirms the BAIT construct does not transactivate the reporter gene *lacZ*. The construct pGILDA-MGL-1Ct did not transactivate *lacZ* (table 5.1). The constructs pGILDA-MGL-2iiCt and pGILDA-MGL-3iCt strongly transactivated *lacZ*. To address this the constructs pGILDA-MGL-2iiCt and pGILDA-MGL-3iCt were each co-transformed with the less sensitive reporter constructs pJK103 (medium sensitivity) and pRB18-40 (low sensitivity). The MGL-2iiCt and MGL-3iCt fusion proteins transactivate the pJK103 lacZ reporter construct but not the weaker lacZ reporter construct pRB18-40 (table 5.2).

To test if the MGL BAITs transactivate *leucine*, EGY48 were co-transformed with an individual MGL BAIT construct, pSH18-34 and the empty library vector pJG4-5. The co-transformed yeast cells were plated onto the media –HUTL (Gal), which selects for *leucine* expression and in parallel –HUT (Gal). The inhibition of growth by a *leucine* deficiency confirms the BAIT cannot transactivate *leucine* expression. The construct pGILDA-MGL-1Ct did not transactivate *leucine* expression (table 5.1). The constructs pGILDA-MGL-2iiCt and pGILDA-MGL-3iCt both transactivated *leucine* expression (table 5.2).

The toxicity of each of the MGL C-terminal BAITs to the host yeast strain EGY48 was tested. The MGL-1, MGL-2ii and MGL-3i C-terminal BAIT constructs were not toxic. The galactose induction of each of the C-terminal fusion proteins did not cause a significant reduction in the number of yeast colonies, compared to a non-inducing plate (containing glucose) spread with the same density of cells (data not shown).

Control Assay	Transformation	Predicted	RESULT
Nuclear Targeting	pJK101	Blue colour	
		Repression of blue colour / LacZ	
	pJK101 pGILDA.MGL-1C-Tail	Repression of LacZ confirms targeting	
β-galactosidase reporter gene autoactivation	pSH18-34 pGILDA.MGL-1 C-Tail	White colour infers autoactivation is not occurring	
Leucine reporter gene autoactivation	pSH18-34 pGILDA.MGL-1 C-Tail pJG4 - 5	Lack of growth on Leucine deficient media confirms autoactivation is not occurring	-HUTL (GAL) -HUT (GAL)

Table 5.1 Control assays performed with the MGL-1 C-terminal BAIT construct. The control assays confirmed the MGL-1 C-terminal is expressed by the host yeast strain and does not transactivate *leucine* or *lacZ* expression. The *leucine* and *lacZ* transactivation assays were incubated for 20 days at 30°C, which exceeds the intended duration of the screen. Nuclear targeting assays were incubated for 24 hours at 30°C.

Control Assay	Plasmids Co-Transformed with MGL C-Tail Construct	pGILDA MGL-2ii CT	pGILDA MGL-3i CT	
Nuclear targeting	pJK101			✓
LacZ Transactivation	pSH18-34 (Strong Sensitivity)			✗
	pJK103 (Medium Sensitivity)			✗
	pRB18-40 (Weak Sensitivity)			✓
Leucine Transactivation	pRB18-40	-HUTL(gal)	-HUTL(gal)	✗

Table 5.2 Control assays performed with the C-terminal constructs pGILDA-MGL-2iiCt and pGILDA-MGL-3iCt. Both MGL-2ii and MGL-3i C-terminals are capable of transactivating *leucine* and to a lesser extent *lacZ* when the low sensitivity reporter pRB18-40 is co-transformed. Images are recorded after a 15 day incubation at 30°C. A tick indicates the control assay has been passed successfully, a cross indicates a failure.

The control assays identified the MGL-1 C-terminal was the most suitable BAIT for the 2-hybrid screen, it is nuclear targeted and does not transactivate either of the 2-hybrid reporter genes. The PDZ binding motif contained in the MGL-1 C-teminal suggests it possesses the ability to interact with intracellular molecules that scaffold its function. Reporter constructs have shown it is selectively expressed and widespread in the *C.elegans* nervous system. To identify molecules that participate in the functional scaffolding of MGL-1 a 2-hybrid screen was initiated with the MGL-1 C-terminal (Construct pGILDA-MGL-1Ct).

5.8   Screening Parameters

The transformation efficiency of the host yeast strain (EGY48) by the *C.elegans* library was optimised by comparing it to an established mammalian brain LexA cDNA library, previously screened with the mGluR2 and mGluR3 C-terminals (Pilkington, B). Yeast (EGY48) was sequentially transformed with pSH18-34, MGL-1Ct-pGILDA and then these cells were transformed with the *C.elegans* cDNA library (pJG4-5). The number of transformants screened was titred to 400 million, which corresponds to approximately three times the titre of viable transformants. The screen was performed on media selecting for the *leucine* reporter gene. The transformants were plated at a density of $2x10^6$ per plate. Colonies able to grow in the absence of leucine were selected on days 4, 5 and 7 of the screen. On day 10 the screen plates exhibited a large background of growth *(figure 5.3C)*. For the further selection of interacting clones β-galactosidase filter lift assays were performed on the screen plates *(figure 5.3D)*. During the course of the screen 640 colonies were selected on the basis of their growth (days 4, 5 and 7) and 1650 colonies were identified by the β-galactosidase filter lift assay on days 10 and 11.

The colonies identified were subjected to sequential rounds of selection, by streaking them onto media selecting for both leucine and β-galactosidase expression *(figure 5.3A)*. Only colonies positive for both reporter genes were passaged through each round, negative colonies were eliminated. Screen controls were included in each round. The positive control was yeast co-expressing SAST (Syntrophin Associated Serine/Threonine Kinase) and the mGluR5 C-tail. This was previously characterized as a strong interaction in yeast (Pilkington, B). The negative control was yeast co-expressing MGL-1Ct-pGILDA and the empty library vector pJG4-5. After the three rounds of selection 1000 colonies were selected for prioritisation.

5.9   The Prioritisation of Selected Colonies

Colonies were prioritised for further characterization based upon the strength of reporter gene expression they exhibited . This was taken as an indication of the strength of the protein-protein interaction occurring in the yeast

and allows high-affinity interactions to be discriminated from medium to low-affinity interactions. The strength of interaction predicted by the two-hybrid
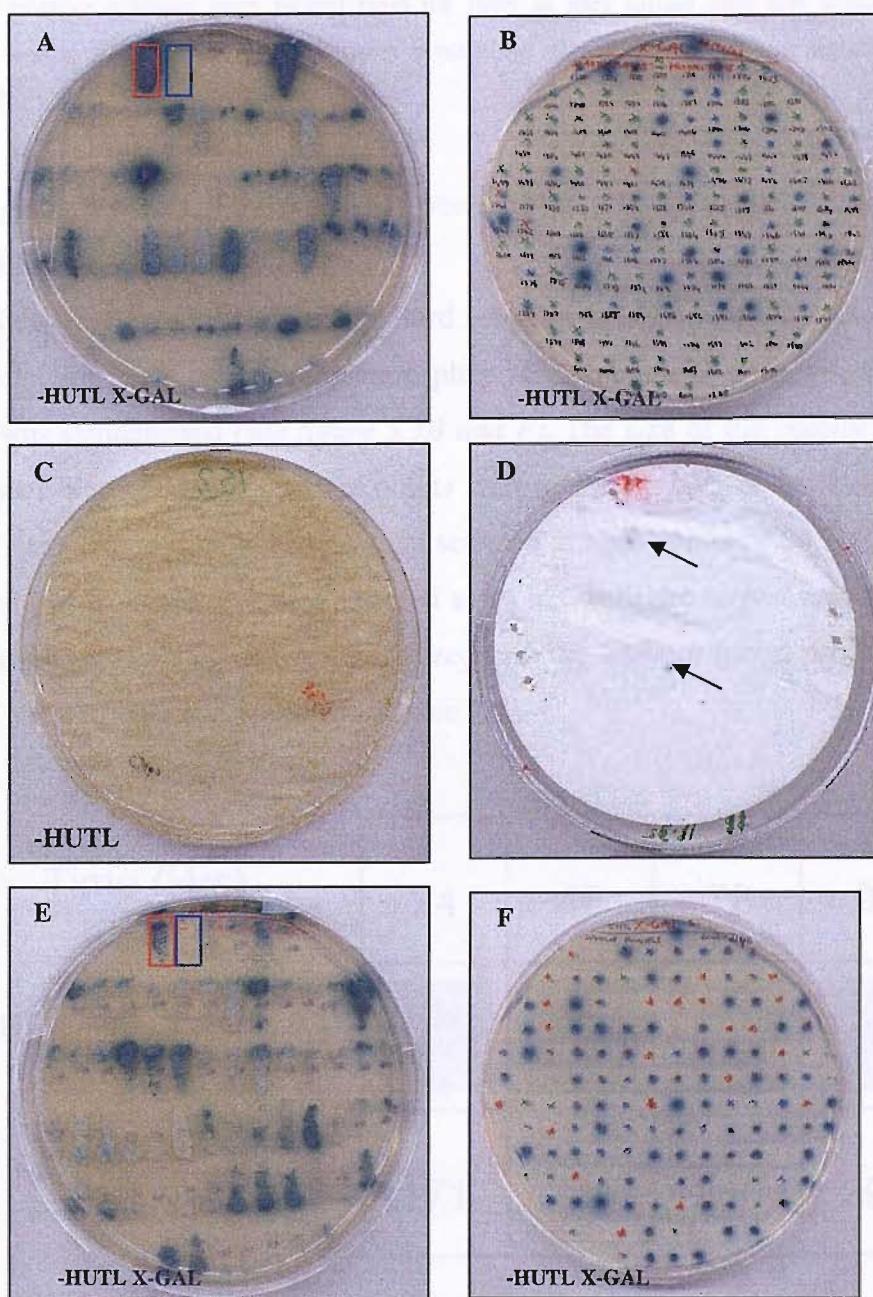
Figure 5.3 Colony selection and prioritisation. A. Colonies identified on days 4, 5 and 7 were picked on day 7 and streaked onto plates selecting for *lacZ* and *leucine* expression. Only colonies exhibiting growth and blue colour after 5 days at 30°C were picked and passaged to the next round of selection. The red box indicates the position of the positive control *(see text)* and the negative control *(see text)* is indicated by the blue box. **B.** Colonies remaining positive for both reporter genes after three rounds of selection were picked and spotted onto prioritisation plates *(image shown)*. Colonies were prioritised on the strength of the reporter gene expression. Image recorded at 96 hours **C.** On day 10 the screen plates had developed a background growth of yeast colonies.

Image taken following 10 days at 30°C. **D.** β-galactosidase filter lift assays were performed on the screening plates to isolate colonies that are expressing the *lacZ* reporter genes. Filters were allowed to develop for upto 16 hours. Arrows indicate the position of positive yeast colonies. **E.** The *lacZ* positive colonies were picked from the filter as they turned blue and streaked onto selective media. **F.** Filter selected colonies remaining after three rounds of selection were prioritised.

system generally correlates well that occuring *in vitro* (Estojak J., Brent R., Golemis E.A. 1995).

Colonies were taken from the third round of selection and transferred to a fourth selection plate, the prioritisation plate. The amount of each yeast colony spotted was standardised *(see figure 5.3B and F)*. The size of the colony and the intensity of blue colour at set time points were representative of the strength of reporter gene expression. Colonies were selected at time points of 24, 48, 72 and 96 hours. For example, colonies selected at 24 hrs were the largest and bluest at this time point. The 171 clones categorized into the 24-hour group proceeded to the next stage of characterization *(see table 5.3)*.

Time (Hrs)	24	48	72	96
Example Clone				
N° Identified	171	284	255	290

Table 5.3 Representative colonies selected at 24, 48, 72 and 96 hour time periods. Colonies were spotted onto –HUTL X-Gal media. The number of colonies classified into each time period is included.

5.10 The Cataloguing of Library Clones

The cataloguing of the prioritised colonies enabled clones to be selected for sequencing. Performing PCR on crude yeast extracts with pJG4-5 specific

191

primers flanking the insert site assessed the size of the clone and *Hae*III restriction digest of the PCR product distinguished between clones of the same size *(see figure 5.4)*. Comparing the size and restriction profile to previously catalogued clones enabled the identification of unique clones. The PCR amplified products of unique clones were sequenced using the pJG4-5 specific primer (5'TFP). This primer sequences over the site of fusion between the library clone and the acidic blob activation domain. The nucleotide sequence of the library clone must give rise to the protein sequence it encodes when it is translated in the same reading frame as the activation domain it is fused to. The identity of the unique clones was established by screening the sequence obtained against the NCBI database using the BLAST facility and the frame of the clone was subsequently confirmed. In addition to the 24-hour clones, 100 clones belonging to the 48 hour subgroup were catalogued. This did not yield any additional or unique clones.
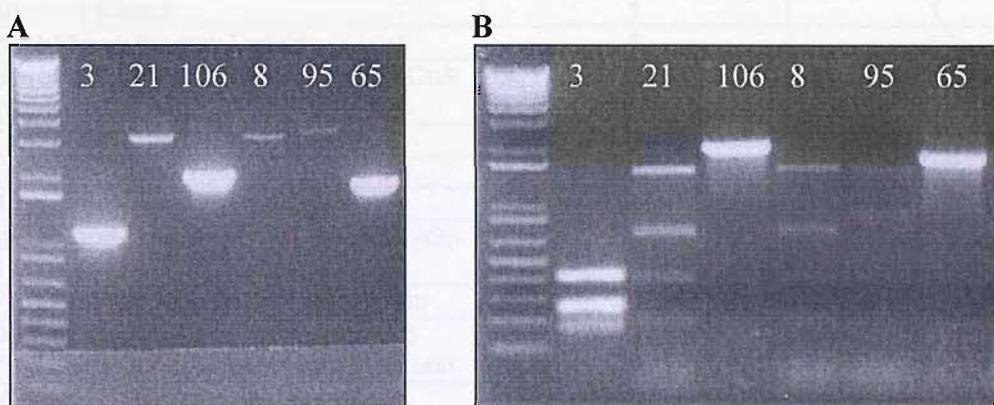
Figure 5.4 The cataloguing of 24hr selected clones.

(A) PCR amplification of library clones, using oligonucleotides (5'TFP and 3'TFP) specific for the library vector (pJG4-5), was performed on crude DNA preparation from selected yeast colonies. A representative agarose gel is shown and the clone number is indicated.

(B) HAEIII restirction digest of the PCR products allowed individual clones to be further characterized and enabled unique clones to be identified for sequencing. For example clone 21 and 8 yielded the same size PCR fragment and *Hae*III restriction pattern and were therefore taken as the same clone.

192

5.11   The Screen Profile.

The identity of clones yielded from the screen are summarised in table 5.4. Three proteins were identified more than once and by two independent clones, they were C01F6.6a, MPZ-1 (C52A11.4a) and PTP-1 (C48D5.2). In all cases the clones identified do not represent the full-length protein sequence. The full-length sequence of the proteins represented by each clone was screened for conserved domain architecture. The region encompassed by the 2-hybrid clone was then mapped to establish which domains it contained.

Clone Identified	Number of Times Identified	Number of Independent Clones
C01F6.6a 2 PDZ Domain Protein	9	2
MPZ-1 9 PDZ Domain Protein	4	2
PTP-1 FERM, PDZ, Protein Tyrosine Phosphatase Domain Protein	4	2
Arginine Kinase	1	1
Heat Shock Protein Daf-21	1	1
Sterol Carrier Protein-x-3-ketoacyl CoA thiolase	1	1
Troponin C (tnc-2)	1	1
Troponin C (tnc-1/pat-10)	1	1
Superoxide Dismutase SOD-1	1	1
gpd-2, gpd-3 genes for glyceraldehyde-3 phosphate dehydrogenase	1	1
Putative protein with coil-coil domain	1	1
Dehydrogenase E1 component and Transketoalse-pyridine binding domain	1	1
Malate Dehydrogenase	1	1
Putative protein family member (XI402)	1	1
Putative protein with coiled-coil 4 domain (1B46)	1	1
Putative protein with 2 coiled-coil domains (5c349)	1	1
2 5-diketogluconic acid reductase family member	1	1
Phosphoserine aminotransferase	1	1

Table 5.4 A Summary table of the MGL-1 C-terminal 2-hybrid Screen. Clones identified as being 'out of frame' with the Acid Blob activation Domain, therefore not giving rise to the protein predicted by the nucleotide sequence have been excluded. As have nuclear proteins, transcription factors and ribosomal subunits, these represent low priority clones based on their likely intrinsic ability to transactivate the reporter genes.

5.12   C01F6.6a

The gene *c01f6.6* encodes two splice variants, C01F6.6a and C01F6.6b, both contain 2 PDZ domains but have alternative N-terminals. The clone identified from the screen corresponds to C01F6.6a; it is almost full length and contains both predicted PDZ domains.
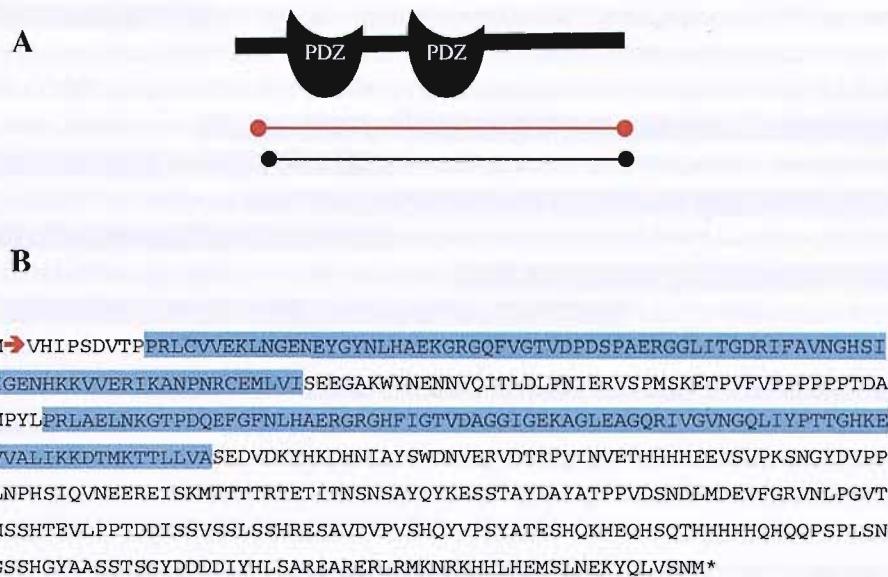
Figure 5.5 The C01F6.6a protein and its predicted sequence. A. The domain architecture of C01F6.6a. The arrows map the region of the protein contained within the clones isolated from the screen. The red arrow represents the clone used to perform the 2-hybrid confirmations. The black arrow represents a distinct clone identified from the screen. **B.** The predicted C01F6.6a protein sequence. The red arrow signifies the commencement of the 2-hybrid clone, represented by the red bar in figure A and identified by sequence generated with the 5'TFP. The regions highlighted in blue correspond to sequence comprising PDZ domains predicted by CDART.

5.13   MPZ-1a

The gene *mpz-1* encodes two splice variants MPZ-1a and MPZ-1b. The protein sequence of MPZ-1a contains 10 PDZ domains, MPZ-1b is a shorter splice variant corresponding to the 5' end of MPZ-1a and contains 5 PDZ domains. The clone identified corresponds to the C-terminal end of MPZ-1a and contains 4 PDZ domains.

A
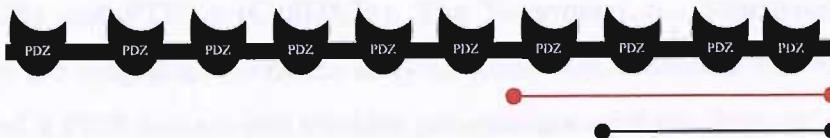
B

MVLSGDWTQVEVIHLNTETGGLGFGIVGGTSTGVVVKTILPGSPADKDGRLQPGDHILQIGNINSHGMSSQ
QVATILRHQHPTVDMIVGRPIAYADKPIDTPECFTLATKAALCATTLEEALQRQMATSVPSTTANQTPNTL
SPSISCVPIEEEKEEKEPEPEPTTSKEQPIKITTYLPDTIDIRKPCDCAPNYNRRDKRKSSCDDGEHGDGD
GGGGCKKTKEDHQEEENEIDSKDELGEMVRSPTVETGQSDENSVVTSVSRSSKRNSNSITSRISLKSLQEM
ALTVFLRENWESTKFELIDVALHRDPALGLGITVAGYVHKKEEIGGIFVKSLVPRSAASSSGVIKVHDLIL
EVNGTTLEHMSHADSVRTLVKSGDQVKLKLVRFPLSSPQAQCLKMLQQQETETQVIDVKLSNPDLVKEWKS
RIGDDIEIIAAVVKPDRQSVDGGLGISLEGTVDVLNGAQLCPHHYIESIRQDGPVAKTKVLQAGDELLQVN
HSPLYGESHVTVRQALTRAVHSGAPVTLIVARRSQHLHVFEPTTYEKQLPLSFPFLAANQETVVKAKSDID
LTSTRAETTHLLQRVSRRLRSKSLENFHGLAVWNCVPLVIHLCKDSRGLGFSIVDYKDPTHRDESVIVVQS
LVPGGVAQADGRVVPGDRLLFVNNHDLSNSSLERAVAVLKAARMGPVRLGLAKPIPVDQKHINFKKSHPVP
LQVRKLCGLVQLNNIESFILKAIIRQEVRQQHEVNGVSRSEFSMTSWSPCSTRSVSPCGSPISLRGSWAYD
VVFLPTHLERTVKLQKGALPLGAVLDGDKDKGVNGCVVKSICGKKAVALDGRIQVGDFITKINTESLRNVT
NSQARAILKRTNLVGTFCNVTYITSADAKTWKERFQRPSESSSPIINRLSPKVFPKFYRSPFMQRQESQSK
TEMTDDETEAPSIMTDSMSEHIKTFDLAEGSSRSHHDEQEVRNRMSRLIDGVEVDDFVNLIIKEAITDATI
ELRVLHKTKDWSNCNHKREERLESPPLPPPPPEVLSSPPKSPVASVQPTPRDLEEVLTRSTASSLEYHSGQ
RTSQLHILSTEEEVLQATPPSSPENKSSEVPPSISPSGIKLAGEVTAPEEIEVVKEQAENVDRAETAATGA
EVEATSTPTAEAIGNSKDDESTTTSISQQSVGLQTQALNSTEVNSNMSRVTSRTPSTGSESLQNQARQLVR
SKYWGEARTVTLVREPNKSFGISIVGGRVEVSQKGGLPGTGNTVCGIFIKSVLPNSPAGRSGQMNMGDRVI
SVNDVDLRDATHEQAVNAIKNASNPVRFVLQSLHTNQQNMINSASNSTVGSVRFENAKPEEELPPTALVTP
LKPMISSSGSSTSKPAMNFPPPSISTTTTTSMESESKEEEA➔TSSSPDIQRENTVKRRSMDVEREQPEAN
PVEETSIVKEKTPPRKISAKSNSSQKESNDRKEIKKQKSVEREMSVESKKSVRSIKKHRQESTIIKKSPSN
ETAPLIVSDVSSSETHEDEPQAMSPSTSFDTRDADAMRALGIDDDSAAFQIKNDGEEPSKFFYTAGKIERK
YDSDGGELVLVACERPDGGLGISLAGNKDRDKQNVFVVNVRPSCPLAIRPGDELLEINGRLLNKISHVAAS
AVVRECCDQHQNIEIVLRRRNGALNECAVRSDTITSSQSLPSPPTSPSTLAPSESKSTAPQMPTTTEGIEN
AAPNQEELSRKKSFSQERTQAIENGRETMIEIDKDKGKGLGLSIVGGADTVLGTVVIHEVYSDGAAAHDGRL
KPGDQVLEVNGTSLRGVTHDQSIAYLRRTPPKVRLLIYRDVNLQLSLLDPTQIYNIFEIDLVKKTGRGLGI
SIVGRKNEPGVYVSEIVKGGLAESDGRLMTGDQILEVNGKDVRGCMQEDVAAMLKTITGKVHLKLGRWKIT
ETANRVHAATQALAKSATTPRTTENNNDPVRPVAVAVSSPAATPSTSASVTVTSAPATSVPEPNVVVEPIP
GPTSAREERNDVPPPAPPMRPIITHTSPEGCETQQEPAGLSPVTEEPSSGCDFMSVQEEEQPRTPTTTTSS
SNNNNSLAIDIIHDLKEEGSDTLLVELKKVVDQQLGMGIGKRSRGILVTSLQPGSAAAEKLKVGDRILAVN
ALPVSDQLSAVTFVKASGQRLYLQIARPHSIPQQ*

Figure 5.6 The MPZ-1a protein and its predicted sequence. A. The domain architecture of MPZ-1a. The arrows map the region of the protein contained within the clones isolated from the screen. The red arrow represents the clone used to perform the 2-hybrid confirmations. The black arrow represents a distinct clone identified from the screen. **B.** The predicted MPZ-1a protein sequence. The red arrow signifies the commencement of the 2-hybrid clone, represented by the red bar in figure A and identified by sequence generated with the 5'TFP. The regions highlighted in blue correspond to sequence comprising PDZ domains predicted by CDART.

5.14 PTP-1

The gene *ptp*-1 encodes three splice variants PTP-1a (C48D5.2A), PTP-1b (C48D5.2b) and PTP-1c (C48D5.2c). The N-terminal is alternatively spliced, PTP-1a is the longest and contains an N-terminal FERM domain and towards the C-terminal a PDZ domain and tyrosine phosphatase catalytic domain. The splice variants PTP-1b and PTP-1c are very close in size and do not contain the FERM domain. The longest 2-hybrid clone corresponds to a homologous region of the three splice variants, consisting of the PDZ and tyrosine phosphatase domain. It does not distinguish between them.
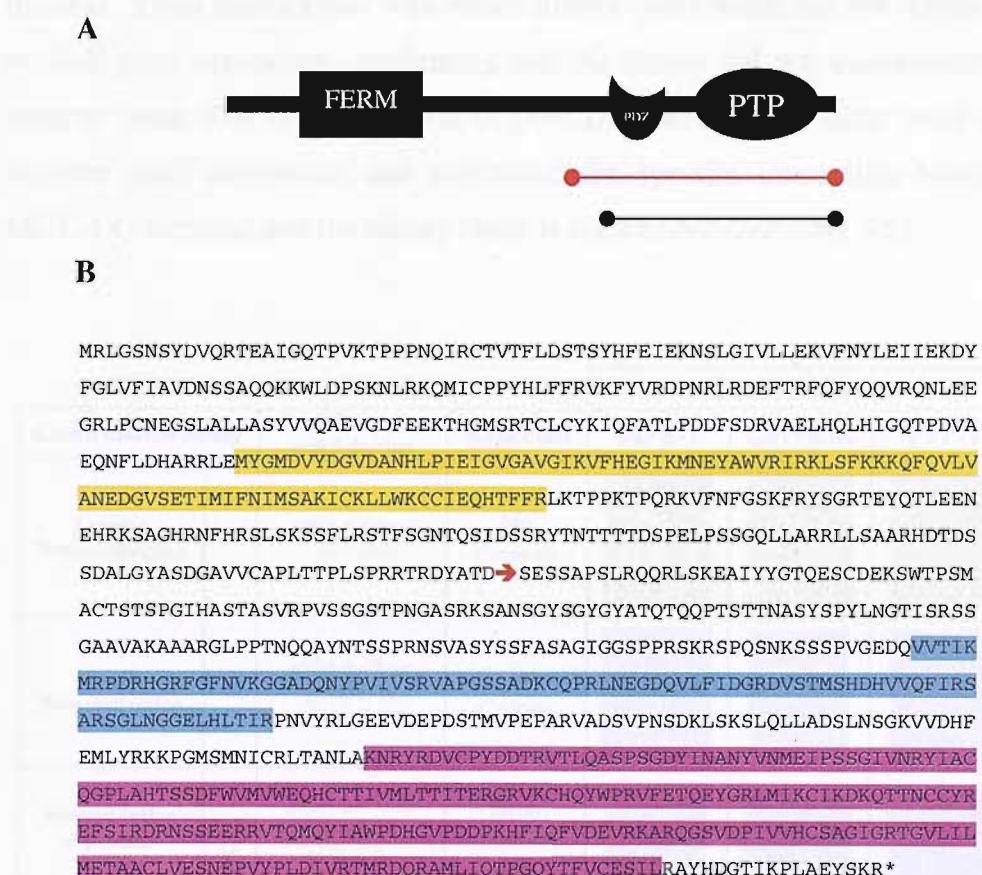
Figure 5.7 The PTP-1a protein and its predicted sequence. A. The domain architecture of PTP-1a. The arrows map the region of the protein contained within the clones isolated from the screen. The red arrow represents the clone used to peform the 2-hybrid confirmations. The black arrow represents a distinct clone identified from the screen. **B.** The predicted PTP-1a protein sequence. The red arrow signifies the commencement of the 2-hybrid clone, represented by the red bar in figure A and identified by sequence generated with the 5'TFP. The highlighted sequence corresponds to the FERM (Band 4.1, Ezrin, Radixin, Moesin) domain (yellow), the PDZ domain (blue) and the tyrosine-phosphatase (pink), predicted by CDART.

5.15 Confirmation of protein-receptor C-terminal interaction

The longest clones of the three prioritised proteins, MPZ-1, C01F6.6a and PTP-1 were rescued in the electrocompetent bacterial strain XLI Blue and crude DNA preparations were made. Restriction digest analysis (EcoR1/Xho1) of the DNA extracted selected for bacterial colonies carrying the pJG4-5 vector containing the interacting clone. High quality DNA preparations were subsequently made and sequenced to confirm the fusion protein is in-frame. The clones were transformed back into the yeast EGY48 host strain alone and in combination with the pGILDA-MGL-1Ct construct and the pSH18-34 reporter plasmid. Yeast transformed with either library clone alone did not report *leucine* or *lacZ* gene expression, confirming that the clones did not transactivate either reporter gene. The re-introduction of pGILDA-MGL-1Ct to these yeast enabled reporter gene expression and confirmed the specific interaction between the MGL-1 C-terminal and the library clone is responsible *(see table 5.5)*.

Confirmation Assay		Transformation and Selection	Expected	Confirmation		
				MPZ-1	C01F6.6a	PTP-1
Leucine Transactivation	✘	pJG4-5-Clone -TL GAL	No Growth			
LacZ Transactivation	✘	pJG4-5-Clone pSH18-34 -T X-GAL	No Blue Colour			
Reconstitution of protein-protein interaction	✓	pJG4-5-Clone pSH18-34 pGILDA-MGL1Ct -HUTL X-GAL	Blue colour and growth			

Table 5.5 Confirmation of MGL-1 C-terminal interactions with rescued clones identified from the yeast-2-hybrid screen. Images are of representative colony streaks after 10 days incubation at 30°C, in each case an approximately equal colony pick was streaked.

5.16 Characterizing the specificity of the MGL-1 C-terminal interactors.

The specificity of the interaction between the MGL-1 C-terminal and the proteins prioritised from the screen was assessed using the yeast-two-hybrid system. C-terminal domains that have known roles in directing protein-protein interactions important for a receptors scaffolding were screened against the MGL-1 interactors. The C-terminal domains used belonged to the ionotropic glutamate receptors GLR-1 (an AMPA-like receptor) and NMR-1 (an NMDA-like receptor) and the tyrosine kinase receptor LET-23 (the EGFR homolog) *(figure 5.8)*. The C-terminal of GLR-1 contains a type I PDZ binding motif and directs the receptors targeting in neurones. The C-terminal of NMR-1 contains a type II PDZ ligand and is required for the targeting of NMR-1 to the synapse. The C-terminal of LET-23 contains a type I PDZ binding motif. It is responsible for interacting with the protein LIN-7, a component of the heterotrimeric complex that targets LET-23 to the basolateral membrane of the vulva precursor cells.

NMR-1 C-terminal

```
EVSYGRRLADKGRRRRIVTRYFQKWHDLTLGKKRRPYRLKYNLDRMIVRRGFSGLERC
SFQELRERRQIRGLPTSKVDPYCFWPDLDFKDKPLVLFCSRCRNIVESDVHRETGLFA
FLSCFLFAILFLWPCSPLPCFLSSFSDFVHICPLCSHIMGRFRRARSTRFYV
```

GLR-1 C-terminal

```
FLYRSRIEARKSNSNSMVANFAKNLKSALSSQLRLSVEGGAVAQPGSQSHNAIRRQQV
AAFLPANEKEAFNNVDRPANTLYNTAV
```

LET-23 C-terminal

```
DKKLGAGAFGTVFAGIYYPKRAKNVKIPVAIKVFQTDQSQTDEMLEEATNMFRLRHDN
LLKIIGFCMHDDGLKIVTIYRPLGNLQNFLKLHKENLGAREQVLYCYQIASGMQYLEK
QRVVHRDLATRNVLVKKFNHVEITDFGLSKILKHDADSITIKSGKVAIKWLAIEIFSK
HCYTHASDVWAFGVTCWEIITFGQSPYQGMSTDSIHNFLQELLRCWMADPKSRPGFEI
LYERFKEFCKVPQLFLENSNKISESDLSAEERFQTERIREMFDGNIDPQMYFDQGSLP
SMPSSPTSMATFTIPHGDLMNRMQSVNSSRYKTEPFDYGSTAQEDNSYLIPKTKEVQQ
SAVLYTAVTNEDGQTELSPSNGDYYNQPNTPSSSSGYYNEPHLKTKKPETSEEAEAVQ
YENEEVS
```

Figure 5.8 The sequence of the receptor C-terminals used to screen against the MGL-1 C-terminal interactors. The PDZ interaction motif is highlighted in red and underlined.

The suitability of each receptor C-terminal domain for use as the 'BAIT' in the LexA two-hybrid system was verified *(table 5.6)*. Yeast co-transformed with the C-terminal domains and the individual interactors were spotted onto a single assay plate at the same time point to standardise the growth conditions and the incubation period. The clones C01F6.6a and MPZ-1 display a specific and strong interaction with the MGL-1 C-terminal. No 2-hybrid interaction was detected between the other C-terminal domains screened and either MPZ-1 or C01F6.6a. The clone PTP-1 interacted specifically with the GLR-1 C-terminal and to a similar extent in the 2-hybrid system as the MGL-1 C-terminal. No 2-hybrid interaction was detected between PTP-1 and either the NMR-1 or LET-23 C-terminal *(table 5.7)*.

ASSAY			Vector Transformation	Expected Result	Confirmation
Lac-Z Transactivation	BAIT TEST		pGILDA-GLR-1Ct & pSH18.34	No transactivation = No blue colouration	
			pGILDA-NMR-1Ct & pSH18.34		
			pGILDA-LET-23Ct & pSH18.34		
	(-)		pRFHM1 & pSH18.34	WHITE	
	(+)		pSH17.4 & pSH18.34	BLUE	
Leucine Transactivation			pGILDA-GLR-1Ct & pSH18.34	No transactivation = No Growth	
			pGILDA-NMR-1Ct & pSH18.34		
			pGILDA-LET-23Ct & pSH18.34		
Nuclear Targeting	BAIT TEST		pGILDA-GLR-1Ct & pJK101	Nuclear targeting = Repression of blue colouration	
			pGILDA-NMR-1Ct & pJK101		
			pGILDA-LET-23Ct & pJK101		
	(+)		pEG202MAX & pJK101	WHITE/FEINT BLUE	
	(-)		pJK101	BLUE	

Table 5.6 The yeast-2-hybrid controls performed using the C-terminals of GLR-1, NMR-1 and LET-23 (Constructs pGILDA-GLR-1Ct, pGILDA-NMR-1Ct and pGILDA-LET-23Ct). Each construct was nuclear targeted and did not transactivate either *leucine* or *lacZ* expression.

PREY \ BAIT	MGL-1 C-tail Type I PDZ Motif (-TFL)	GLR-1 C-tail Type I PDZ Motif (-TAV)	NMR-1 C-tail Type II PDZ Motif (-FYV)	LET-23 C-tail Type I PDZ Motif (-EVS)
C01F6.6a		✗	✗	✗
MPZ-1		✗	✗	✗
PTP-1			✗	✗

Table 5.7 A directed 2-hybrid screen of the MGL-1 C-terminal interactors. The 2-hybrid interaction between the MGL-1 C-terminal and each of the interactors is shown for reference. Where a 2-hybrid interaction was recorded it is shown in the table for comparison. The absence of a two-hybrid interaction is represented by a cross Images were recorded after 5 days @ 30°C. The screen was allowed to proceed for 10 days. The time course of the GLR-1 interaction with PTP-1 was similar that of MGL-1 and was detected on day 3.

5.17 Discussion

The LexA yeast-2-hybrid system was selected to identify scaffolding molecules of the metabotropic glutamate receptors, MGL-1, MGL-2 and MGL-3. 2-hybrid controls were performed with C-terminals belonging to each of the receptor subtypes. In the case of MGL-2 and MGL-3 the longest splice variants identified were used. The MGL-1 C-terminal was identified as the only suitable BAIT for the 2-hybrid screen.

The large-scale screen of a *C.elegans* cDNA library with the *mgl-1* receptor C-terminal harvested three proteins of specific interest. Based upon their domain architecture and similarities to known scaffolding proteins, the three clones selected for further analysis were, C01F6.6a, MPZ-1 and PTP-1. The characteristics of the screen suggest that the interactions are authentic. Each protein was identified more than once with independent clones, at early stages of the initial screening process and each belongs to the group of clones prioritised as the strongest interactions (24 hour colonies). Sequencing over the point of fusion with the acidic blob AD confirmed each clone gives rise to bona fide protein. 2-hybrid controls have confirmed each is a genuine two-hybrid interaction with the MGL-1 C-terminal.

The three proteins were screened against the C-terminal of the receptors NMR-1, GLR-1 and LET-23. These alternative receptor C-terminals were selected to test the specificity of the identified two-hybrid interaction with MGL-1 and to gain insight into the biological function of the three proteins. In *C.elegans* proteins involved in directing a receptors scaffolding can be conserved between different receptor types. An example of this is LIN-10, a protein containing two PDZ domains. It is a shared component of the scaffolding machinery belonging to both receptors GLR-1 and LET-23 and it performs this role in different cell types (Rongo, C. et al 1998).

The C-terminal of each receptor has been implicated in its scaffolding *(see section 1.32.4)*. Both LET-23 and GLR-1 contain a TYPE-I PDZ domain binding motif, as like MGL-1. In the case of LET-23, interactions with intracellular molecules are dependent only on the C-terminal PDZ motif. It directly mediates the interaction between the receptor and LIN-7, another component of its scaffolding complex (Kaech, S.M. et al 1998). In the 2-hybrid system MPZ-1 and C01F6.6a did not interact with either the GLR-1 or LET-23 C-terminal. This

suggests the specificity of the interaction between the MGL-1 C-terminal and both of the proteins MPZ-1 and C01F6.6a is dependent upon more than the terminal PDZ binding motif.

The MPZ-1 and C01F6.6a clones each contains more than a single PDZ domain. It has been established that in some cases tandem PDZ domains are necessary for PDZ mediated target binding (Feng, W et al 2003). Multiple PDZ repeats also provide a basis for dual specificity contained within a single molecule. An example of this is the mammalian homolog of MPZ-1, MUPP-1. It contains 13 PDZ domains and interacts with several different targets through alternative PDZ domains (see section 5.18.3). Also ionotropic and metabotropic glutamate signalling cascades within mammals are known to be interconnected through scaffolding networks, providing a circuitry for cross-talk to occur. The mGluR5 and mGluR1 subtypes are indirectly linked to NMDA receptor subtypes via the protein Shank (Ehlers, M.D. 1999). This is perhaps a mechanism for the strategic positioning of different receptor types relative to each other, enabling mGluR1 regulation of NMDA receptor signalling. The NMR-1 receptor C-terminal, which contains a TYPE-II PDZ binding motif, did not directly interact with any of the identified proteins. GLR-1 did interact with the protein tyrosine-phosphatase, PTP-1 in the 2-hybrid system. The PTP-1 could function as a physical bridge between GLR-1 and MGL-1 or it may function as a generic scaffolding protein of different receptor types, as in the case of LIN-10.

As LET-23 is a receptor tyrosine kinase, the phosphorylation of its C-terminal is intrinsic to its function. The activation state of the receptor regulates the protein interactions that occur at the C-terminal. Upon ligand induced receptor dimerization specific tyrosine residues are autophosphorylated in the C-terminus and are recognised by SH2 domains of target proteins belonging to the downstream signalling pathway. Therefore the regulation of the C-terminal phosphorylation state, by a tyrosine phosphatase, represents a potential mechanism for modulating the receptors signalling. For this reason PTP-1 was screened against the LET-23 C-terminal but did not interact. However, it is important to realise the C-terminal, when used in isolation in a heterologous cell system, is unlikely to be in the same phosphorylated state that exists under normal circumstances. Subsequently if the interaction between PTP-1 and the LET-23 C-

terminal is dependent upon the activated state of the receptor it could be undetectable in yeast.

5.18 Insights into MGL scaffolding by identified interactors

5.18.1 Protein Tyrosine Phosphatase (PTP-1)

Three splice variants of protein tyrosine phosphatase 1 are predicted from the C.elegans genome, the longest of these (C48D5.2a) exhibits an N-terminal FERM domain (4.1 protein, Ezrin, Radixin and Moesin) followed by a PDZ and C-terminal tyrosine phosphatase domain. The FERM domain is absent from the shorter N-terminal splice variants (C48D5.2b, C48D5.2c).

The FERM domain has been identified as a molecular linker connecting cell-surface proteins to the actin cytoskeleton and is located at the N-terminus of the majority of FERM containing proteins (Ollis, D et al 1985). The co-assembly of FERM and PDZ domains within the same protein is well established. Examples of such proteins can be found in mammals (PTP-MEG and PTP-BAS) and *C.elegans* (PTP-2).

A physiological role for protein tyrosine phosphatase regulation of mGluR1 signalling has been described at parallel fibre to Purkinje neuron synapses of the cerebellum, a region of the brain involved in motor co-ordination. The activation of mGluR1 can trigger a nonselective cation channel conductance and generate a slow EPSP. The coupling pathway involved is independent of the classical mGluR1 second messenger, PLC and can be turned on and off by tyrosine dephosphorylation and phosphorylation respectively (Canepari, M et al 2003). Similarly phosphatase activity has been demonstrated to modulate the depression by Group I subtypes of the slow duration after-hyperpolarization (sAHP) in hippocampal CA1 pyramidal neurons (Ireland, DR et al 2004). The depression of this current by the Group I selective agonist DHPG is gated by a delicate balance between tyrosine phosphatase and tyrosine kinase activity. Disruption of this steady state by inhibition of phosphatase activity promotes a heightened state of tyrosine kinase target phosphorylation and impedes the depression of the sAHP by GroupI mGluR activation in CA1 neurons of hippocampal tissue slices. The outcome is an enhancement of neuronal cell

excitability. Again, the mode of this regulation is independent of classical Group I signalling mediators (PKC, IP3 and PLC). A GAL-4 based Y-2-H screen performed with the C-terminal of mGluR 3 has identified a direct interaction with protein phosphatase 2C (PP2C) (Flajolet, M et al 2003). However, other than regulatory factors that control this interaction there is little functional data to define its physiological purpose.

5.18.2 C01F6.6a

C01F6.6a contains two PDZ domains and is predicted to be cytoplasmic. Its closest mammalian homologue is the Na/H-exchange regulatory factor 2 / EBP-50 but lacks the C-terminal ERM interaction domain. NHERF-2 has been shown to direct parathyroid hormone receptor (PTH1R) signalling, switching it from a cAMP to a PLCβ1 pathway by simultaneously binding PTH1R and PLCβ1 via PDZ interactions and coupling PTH1R to the inhibitory G-protein G_i (Mahon, M et al 2002). EBP-50 has been identified in the sorting of agonist induced internalised β2-adrenergic receptors into recycling endosomes, which is dependent upon intact interactions between the PDZ domain of EBP-50 and the receptor C-tail and the ERM domain of EBP-50 and the actin cytoskeleton (Cao, T et al 1999). C01F6.6a may be capable of mediating similar events in C.elegans with respect to MGL-1 signalling.

5.18.3 MPZ-1

In silico analysis predicts MPZ-1 contains 10 PDZ domains, this domain architecture suggests a scaffolding function. Its closest human homologue is the multiunit PDZ protein 1 (MUPP-1), which has been identified as containing 13 PDZ domains. Protein alignments show the PDZ domains of MPZ-1 are very similar to those in MUPP1 and the structural organization of the PDZ domains is evolutionarily conserved. (Ullmer, C et al 1998). Northern Blot analyses have shown tissue distribution of the MUPP1 transcript is widespread, in the brain, heart, liver, placenta, skeletal muscle, kidney and pancreas.

MUPP-1 was cloned from a GAL4 Y-2-H screen of a human fetal brain cDNA library with the C-terminal domain of the $5HT_{2C}$ receptor (Ullmer, C et al 1998). The $5HT_{2C}$ receptor possesses a Type I PDZ ligand motif (-SSV) and

interacts specifically with the 10th PDZ domain of MUPP-1 (Becamel, C et al 2001). This interaction is regulated by agonist dependent phosphorylation of Serine residues in the C-terminal PDZ ligand motif (Parker, L et al 2003). Serine phosphorylation causes a loss of MUPP-1 association, identifying this event as a modulator of PDZ:protein interactions, although in the case of MUPP-1 the purpose of this in 5-HT$_{2C}$ receptor function and signalling remains to be established. In situ hybridization and immunohistochemistry have verified the MUPP1 transcript is distributed to the same regions of the brain as the 5HT$_{2C}$ receptor. These include the choroid plexus, cerebral cortex, piriform cortex and hippocampus.

The mouse equivalent of MUPP-1 and MPZ-1 is Mpdz, this has been cloned (Simpson, E.H. et al 1999). QTL (Quantitative Trait Loci) analysis of genetic determinants of acute alcohol withdrawal liability has identified the Mpdz gene from genetic animal models of alcohol withdrawal syndrome (Fehr, C et al 2002). Quantitative RT –PCR has demonstrated the Mpdz mRNA transcript is more abundant in animal models of acute mild ethanol withdrawal compared to animal models of severe ethanol withdrawal (Shirley, R. et al 2004). Additionally sequencing has identified 18 polymorphosims in Mpdz alleles between acute mild and severe models of ethanol withdrawal. 10 of these have been predicted as changing the protein amino acid sequence. This implicates changes in Mpdz protein structure and Mpdz gene expression may be genetic determinates of alcohol withdrawal liability and ultimately the onset of alcoholism.
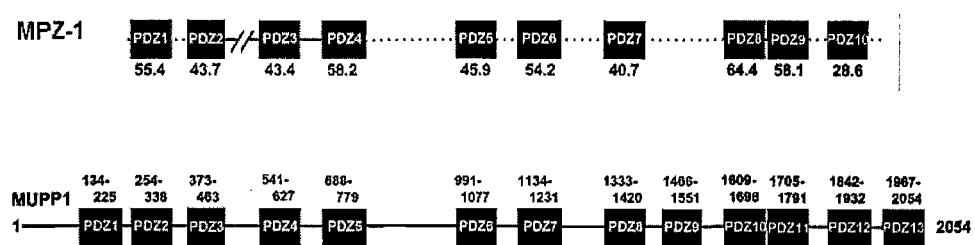
Figure 5.9 Alignment of MPZ-1 and MUPP-1 PDZ domains. The percentage identity between MUPP-1 and MPZ-1 PDZ domains is indicated by numbers beneath MPZ-1. Numbers above MUPP-1 indicate amino acid positions (Taken from Ullmer, C et al 1998).

CHAPTER 6

Molecular characterization and expression analysis of the yeast-2-hybrid prioritised clones

6.1   Introduction

The 2-hybrid screen provides a suitable starting point for the isolation of protein-protein interactions. Nevertheless, it is an *in-vitro* approach and further analysis is required to assess the physiological function of the interactions identified. An initial starting point for addressing this issue is the demonstration of a cellular co-expression pattern for the genes encoding the two interacting proteins. If the genes encoding the interacting proteins yielded by the 2-hybrid screen are not expressed in the same cells as the *mgl*-1 receptor *in-vivo,* the occurrence of a physical interaction between the identified protein and the receptor is very unlikely. At this time the cellular expression pattern of the three proteins prioritised from the 2-hybrid screen were not available from the published literature. Defining the cellular distribution of the *mgl*-1 interactors at the level of the whole organism was therefore necessary and an important requirement to facilitate their further prioritisation.

6.2   Methods for the detection of cellular expression patterns in *C.elegans*

There are various techniques that have been successfully used to define cellular   expression   patterns within *C.elegans*, although none are without their   own   caveats.   Antibody staining allows the localization of a protein, derived from the gene of interest, to be assessed. A substantial insight into a proteins function can be gained by determining the precise subcellular compartment it is localized to. However this is dependent upon the production or availability of a suitable antibody. Currently there are no antibodies available that are capable of recognising any of the three proteins prioritised from the 2-hybrid screen. In-situ hybridisation allows the localization of the mRNA transcript encoding the protein of interest. In-situ hybridisation has been used in *C.elegans* to study the expression of various genes (Tabara H et al 1996). However the non-specific recognition of homologous transcripts can yield high backgrounds and misleading results.

A third technique practised in *C.elegans* involves the production of transgenic animals expressing a promoter fused *in vitro* to a reporter gene such as *LacZ* or fluorescent proteins (eg. GFP). Such a chimera (*promoter::LacZ/GFP*) is commonly referred to as a gene fusion reporter construct and it allows the expression of a fluorescent reporter protein to be placed under the control of a

foreign promoter (Hope I., et al 1998 and 1991) This is typically the promoter belonging to the gene of interest. Hence, in transgenic animals transformed with the gene fusion reporter construct the fluorescent protein will only be expressed in cells where the gene of interest is being transcribed. The process takes advantage of the transparent nature of *C.elegans*, which enables the expression of a fluorescent or coloured reporter to be easily assessed in the intact, living organism throughout development. However, for reasons that will become apparent in subsequent sections, this technique is absolutely dependent upon a defined gene model and a defined ORF translation start site. We have initiated our approach to define the cellular expression pattern of the three proteins interacting with the *mgl*-1 C-terminal by generating transgenic lines expressing gene fusion reporter constructs for the three proteins, PTP-1, MPZ-1a and C01F6.6a.

6.3   Overview of the design approach used to generate gene fusion reporters, with reference to published criteria.

In eukaryotes the promoter encompasses the DNA sequences responsible for the correct regulation and initiation of transcription of a gene (Harbison, CT et al 2004). The promoter contains a spectrum of transcription control elements. In general the foundation of a promoter is the core promoter; this is the site at which the RNA polymerase and transcription factors assemble together to generate an initiation complex (Lee, TI and Young, RA 2000). Proteins that bind to other sequences (promoter elements) within the promoter (or elsewhere) can regulate the assembly of this complex, to mediate transcription control. Enhancer proteins, for example can bind to promoter elements upstream of the core promoter and promote gene expression, whereas repressor proteins serve as negative regulators [Brown, T.A. 2002]

The promoters of several *C.elegans* genes have been studied in depth to gain insight into the mechanisms that underlie the process of sex, stage and tissue specific regulation. For example, *unc*-54 is one of the four myosin heavy chain genes. The expression of *unc*-54 is the subject of multiple control elements that reside in the 5' flanking region of DNA, within the body of the gene and in the 3' untranslated region (Okkema, P.G. et al 1993). A 200bp region of 5' flanking sequence and the third intron can both independently confer tissue specific expression to *unc*-54, whereas the 3' untranslated region is required for expression

and not tissue specificity. This highlights that upstream and downstream sequences can serve as sites for activating enhancers that regulate the properties of a genes expression (Brown, TA 2002).

In our approach the promoter region for a gene of interest was defined as the expanse of genomic DNA between the ATG start codon of that genes open reading frame and the 3' extremities of the nearest upstream gene. This is commonly referred to as the intergenic region. In setting this as our definition we have ignored any potential downstream regulatory elements residing within introns or the 3' untranslated region. Despite this it has been demonstrated that the promoter fusion approach provides a suitable first-pass assessment of cellular expression patterns and it avoids the need to amplify the entire gene. This is a distinct advantage since the genes encoding the hybird identified proteins range from 12kb –35kb (this does not include the intergenic region or the 3'untranslated region) and long distance PCR can present a significant obstacle.

The amount of promoter sequence that is used in the promoter::reporter approach is highly variable. In an independent report, the analysis of 200 randomly selected publications describing an expression pattern derived from a promoter::reporter-approach identified the amount of promoter sequence utilised ranges between three hundred and several thousand base pairs (Dupuy. D et al 2004). In the majority of publications specific expression was obtained without using the entire intergenic region, but instead a small proportion of it and the size of the region used was typically governed by the presence of a convenient restriction site. As mentioned previously as little as 200bp of 5' flanking sequence was sufficient for the specific expression of *unc*-54.

This analysis formed a component of the strategy used to generate the *C.elegans* promoterome (Version 1.1) (Dupuy. D et al 2004). The *C.elegans* promoterome is a genome-wide resource that was generated by the PCR amplification and cloning of ~6600 promoters. The database represents approximately a third of the total number of unique *C.elegans* promoters. Similar gene expression databases are being compiled for other species that include humans, mice and *Xenopus*.

In generating the *C.elegans* Promoterome the upper size limit of the PCR fragment was set to 2kb. This was done primarily to facilitate cloning efficiency. It was justified on the basis that sequence homology between *C.elegans* and

C.briggsae dramatically decreases 1500bp upstream of the ATG start codon for the majority of genes and the consensus in the literature is that the entire intergenic region is not required to obtain expression.

In our approach, where a substantial amount of intergenic sequence was present, the amount of promoter sequence deemed to be acceptable was set initially, as a first pass, to a minimum threshold of 2kb (although the minimum used was actually 2.4kb). Alignments between the *C.elegans* and *C.brigssae* intergenic region of the genes of interest, was used to identify conserved promoter elements. This served as a guide for determining the minimum amount of intergenic sequence required for all potential upstream *cis*-regulatory elements to be included. The fluorescent marker protein was fused in-frame, to either the second or third exon of the gene by subcloning of the PCR amplified genomic fragment into either the vector pPD95.75.DsRED2 or pHAB.mRFP-1. The first intron was included to take account of any additional *cis*-acting elements or enhancers that are located within it and it is known that processing by the endogenous splicing machinery improves the expression efficiency and levels of promoter fusion reporter constructs (Fire, A., Seydoux, GC., Xu, SQ. 1995, unpublished observation). The vectors pPD95.75.DsRED2 and pHAB.mRFP-1 both contain generically used 3' untranslated regions from the genes *unc*-54 and *let*-858 respectively. The alternative to subcloning for the production of gene reporter constructs is the PCR-fusion based approach (Hobert, O 2002). Essentially two primary PCR products are merged together using nested primers to generate a *promoter::reporter* fusion. The PCR reaction is then injected into the gonad of *C.elegans* without further purification. This technique is dependent upon the primary PCR working, which can be hindered when large fragments, several thousand base pairs in size are required.

6.4 Practical steps taken in the production of gene fusion reporters and transgenic animals.

Primers designed to amplify the necessary region of genomic DNA for subcloning were initially based upon the Wormbase gene submission. It became apparent the gene models provided were subject to inaccuracies, primarily these inaccuracies resided in defining exons, intron boundaries and ATG start codons (*See Chapter 3*). It was determined that as a minimum requirement, it was

important to experimentally validate the 5' region of the gene using a RACE approach. In parallel a 3'RACE analysis was performed. The information gained from this was intended for use in the production of protein fusion reporter constructs (*promoter-ORF::reporter*), should the *promoter::reporter* construct identify a cellular co-expression with *mgl*-1. The definition of the 5' and 3' end of a transcript enables primers to be designed for the amplification of the full-length transcript, including 5' and 3' untranslated sequence. Furthermore characterization of the transcript is necessary if the fluorescent marker is to be fused correctly when generating protein fusion reporters.

PCR amplifications were performed on genomic DNA, prepared from wild-type worms and in parallel cosmid DNA. Cosmids were first characterized by restriction digest analysis to confirm their identity. The initial PCR amplification was analysed by electrophoresis, to ensure a product of the correct size was obtained. Both restriction digest analysis and partial sequencing was performed to confirm the correct region of genomic DNA was fused to the reporter, prior to microinjection.

Upon successful transformation of the germ line cells the transgene is expressed as an extra-chromosomal array (Stinchcomb, D.T et al 1985). To screen for transgenic lines expressing the gene reporter constructs, rescue of the temperature sensitive developmental mutation *pha*-1*(e1213)*ts was utilised as a co-injection marker *(see Methods section 2.4.15)*. The protein encoded by *pha*-1 is required for the correct development of the pharyngeal muscle during embryogenesis. At 25°C the pharynx does not develop in *pha*-1*(e1213)* mutants, inhibiting feeding in larvae and causing death. At 15°C pharyngeal muscle development is normal. Micro-injection was performed using the *mgl*-1::GFP;*pha*-1*(e1213)*ts selectable marker strain (generated by crossing the stable, integrated transgenic line *mgl*-1::GFP (Supplied by Katsura, I) with a strain carrying the allele *pha*-1(e1213)ts, *Methods and Materials Section 2.4.15*).

6.5 The RACE characterization of *mpz-1* and its cellular expression

6.5.1 Constructing a gene fusion reporter to elucidate the cellular expression of *mpz*-1.

The MPZ-1a protein is encoded by the gene *mpz*-1. At the time of identifying MPZ-1a as a 2-hybrid interaction partner of the MGL-1 receptor C-terminal, the *mpz*-1 gene model was derived from an *in-silico* prediction (http://ws100.wormbase.org). There were no available ESTs to support the areas of sequence defined as exons and introns and the predicted model represented the only available information regarding the structure of *mpz*-1. The design of the *mpz*-1 gene reporter construct, *mpz-1(i)*::DsRED2 was therefore based upon the *in-silico* defined gene model *(figure 6.1A)*.

The construct *mpz-1(i)*::DsRED2 contained 2.8kb of 5' un-translated sequence. This incorporates the entire intra-genomic sequence between *mpz*-1 and the nearest upstream gene T19E10.2 *(figure 6.1A)*. The fluorescent protein DsRED2 was fused in-frame to the second coding exon of *mpz*-1. The *mpz*-1 gene is separated over the two cosmids T19E10 and C52A11. The region of *mpz*-1 required to make *mpz-1(i)*::DsRED2 resided in the cosmid T19E10. This cosmid was obtained but failed to produce the correct restriction pattern when digested and failed to yield a product when used as the template for PCR. The desired region of *mpz*-1 was successfully PCR amplified from wild-type genomic DNA *(figure 6.1B)*.
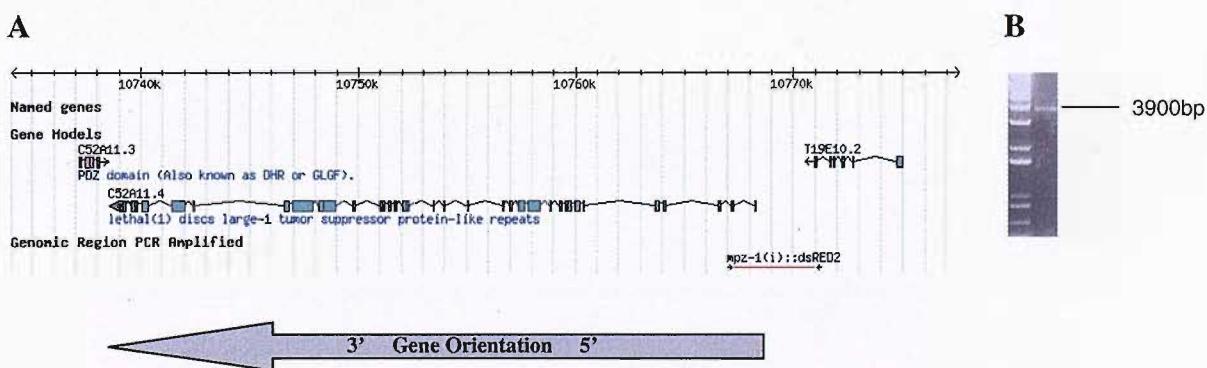
Figure 6.1 The molecular basis of the promoter construct *mpz-1(i)*::DsRED2. A. The *in-silico* predicted exon/intron structure of *mpz*-1 (denoted as C52A11.4 in http://ws100.wormbase.org; 10

May 2003). *Key: Boxes* = exons, *Chevrons* = introns. The red bar indicates the genomic region used to produce the reporter construct *mpz-1(i)*::DsRED2. The nearest upstream gene, T19E10.2, is shown. The grey arrow indicates the gene orientation of *mpz-1* (right to left) as it is documented in Wormbase. **B.** PCR amplification of the *mpz-1* genomic DNA. *Cycle parameters: 1x(2min@94°C), 10x(10s@94°C, 30s@57°C, 3min30s@68°C), 20x(15s@94°C, 30s@57°C, 3min35s@68°C), 7min@72°C.*

6.5.2 Establishment of an altered *mpz*-1 gene model and its implications for the fidelity of the reporter *mpz*-1(i)::DsRED2.

The construct *mpz*-1*(i)*::DsRED2 was partially sequenced at the 5' end. To confirm the sequence as being correct a BLAST search was performed of the *C.elegans* genome repository, Wormbase. Upon alignment of the *mpz*-1*(i)*::DsRED2 sequence to *mpz*-1 it was realised the original *mpz*-1 gene model had in the mean time been significantly altered and also a splice variant of *mpz*-1 (C52A11.4b) had been identified. The modified *mpz*-1 gene model is now denoted as *mpz*-1(C52A11.4a). The majority of the changes to the *mpz*-1 predicted gene model had occurred in the 5' region of *mpz*-1. Consequently the portion of *mpz*-1 that was used for the construct *mpz*-1*(i)*::DsRED2 did not possess the correct molecular components required to report on the cellular expression pattern of *mpz*-1 (*Figure 6.2*).
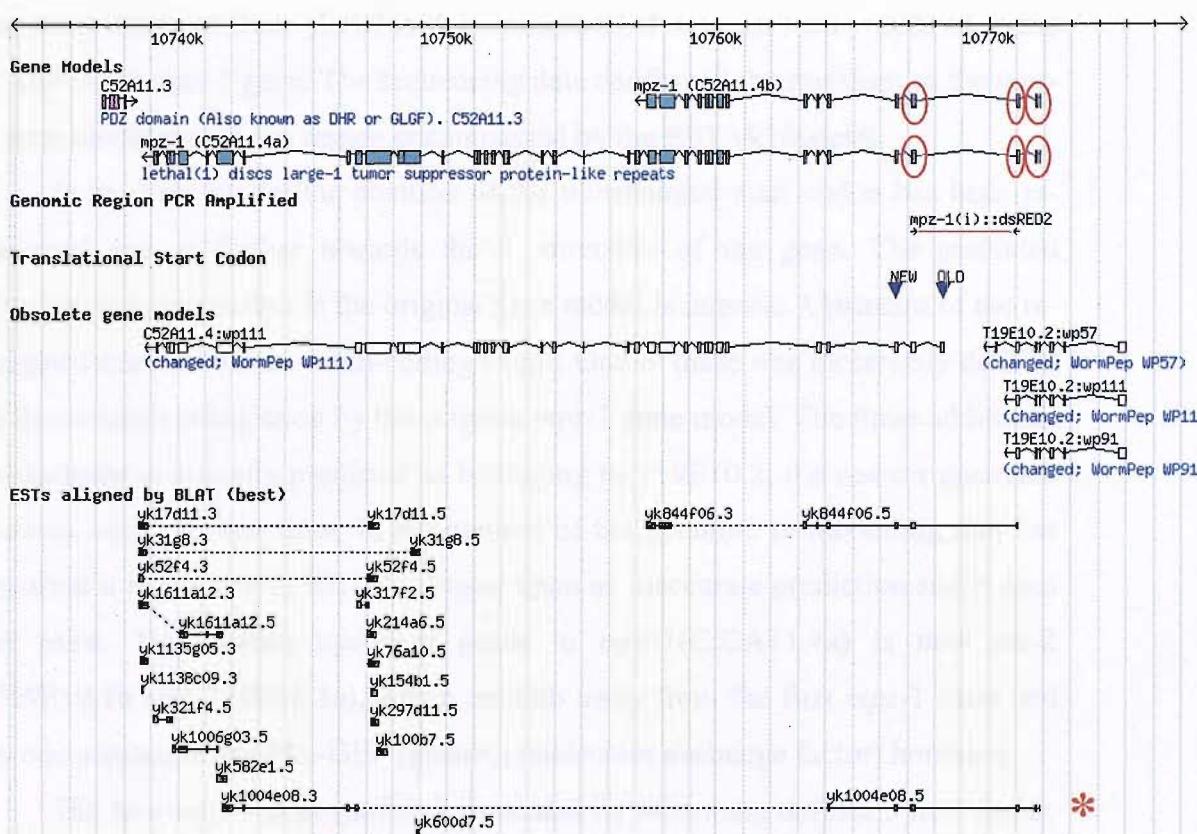
Figure 6.2. Comparison of the old and new *mpz*-1 gene models. The new *mpz*-1 gene structure is denoted as *mpz*-1(C52A11.4a) and it is depicted beneath the heading "Gene Models." The old gene model for *mpz*-1 (C52A11.4:wp111) and T19E10.2 (wp57, wp91 and wp111 represent consecutive modifications to the gene model) are depicted beneath the heading "Obsolete gene

models." The predicted gene T19E10.2 does not exist in the context of *mpz*-1(C52A11.4a). The newly identified splice variant *mpz*-1(C52A11.4b) is shown. The red bar indicates the "Genomic region PCR amplified" and used to make the construct *mpz*-1*(i)*::DsRED2. The position of the "new" and "old" translation start codon is indicated by the blue triangles. The ESTs aligning to *mpz*-1 are shown as "ESTs aligned by BLAT (best)." The 5' exons of the *mpz*-1 gene model defined as non-coding by the EST yk1004e08 (indicated by the red asterisk) are circled in red.

The 5' region of *mpz*-1 was changed because of the EST submission yk1004e08 (originating from the National Institute of Genetics, Mishima, Japan. *figure 6.2*) that defines an SL-1 *trans*-splice acceptor site; re-defines the 5' end of the *mpz*-1 transcript and re-defines the nearest upstream gene. Compared to the other available ESTs yk1004e08 provides the largest coverage of the *mpz*-1 transcript, ~8kb. The yk1004e08 cDNA had been only partially sequenced at the 3' and 5' end by the EST consortium and was fully sequenced by us. The sequence obtained from yk1004e08 contained SL-1 sequence *trans*-spliced to the 5' end of the *mpz*-1 gene. The sequencing data confirmed the revisions to the *mpz*-1 gene model within the region encompassed by the EST yk1004e08.

In the new model the position of the translational start codon has been re-assigned and is further towards the 3' direction of the gene. The predicted translational start codon in the original gene model is intronic. Upstream of the re-assigned start codon are 4 non-coding exons. One of these was incorrectly defined as the second coding exon by the original *mpz*-1 gene model. The three additional exons were incorrectly predicted as belonging to T19E10.2, the nearest upstream gene to *mpz*-1 at that time. In the context of the genomic re-modelling that has occurred it is evident T19E10.2 is based upon an inaccurate prediction and it does not exist. The nearest upstream genes to *mpz*-1(C52A11.4a) is now *ect*-2 (T19E10.1b and T19E10.1a), which are 6kb away from the first *mpz*-1 exon and encode a mammalian Rho-GEF (guanine nucleotide exchange factor) homolog.

The new *mpz*-1 gene model documented by wormbase defines 3 new exons, which are exons numbered 8, 33 and 36. When the new exons are introduced they do not alter the reading frame of the *mpz*-1a transcript or the domain architecture, 10 PDZ domains are still present.

6.5.3 Alignment of *mpz*-1 to the *C.briggsae* genome for the identification of conserved promoter elements

Comparisons between genomes can assist in the annotation of functional genomic elements, in particular the resolution of protein-encoding genes. *Caenorhabditis briggsae* is a soil dwelling nematode and one of *C.elegans'* closest known relatives. The two species diverged from each other ~100 million years ago; they are virtually indistinguishable in their morphology and possess the same developmental programming. Areas of genomic sequence that encode protein are mostly conserved between the two organisms (Stein, L.D. et al 2003). Alignment of *mpz*-1(C52A11.4a) to the *C.briggsae* genome identified strong sequence homology (indicated by the dark blue squares) over the coding regions. The nucleotide sequence of the four *mpz*-1(C52A11.4a) non-coding exons are also highly conserved in the intergenic region of the *C.briggsae* orthologue, CBP00807. This would suggest they are conserved elements that are important for the correct expression of *mpz*-1. There are additional elements also in the intergenic region between *mpz*-1 and T19E10.1a/b that are conserved in *C.briggsae* and are also likely to be important for the gene's expression (*Figure 6.3*).
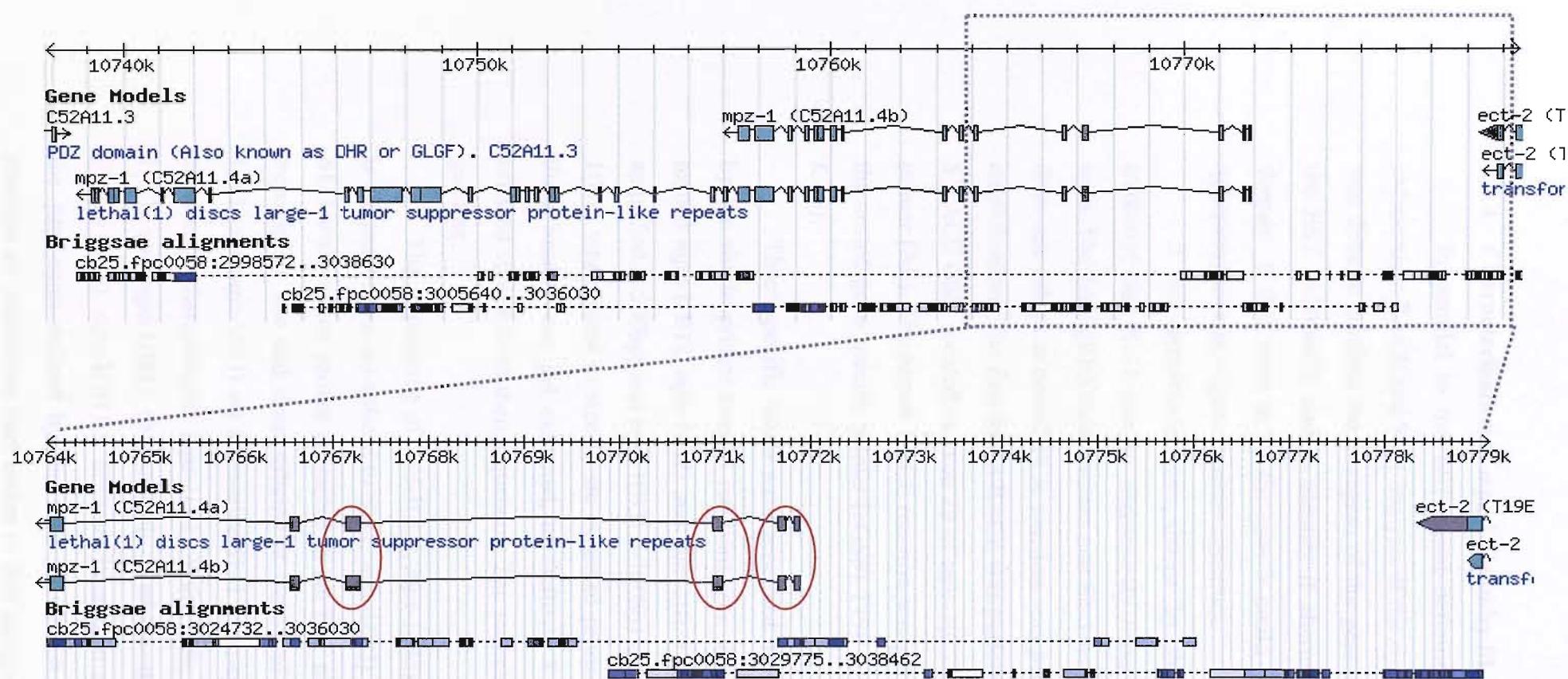
Figure 6.3 Alignment of the *C.elegans* gene *mpz*-1(C52A11.4a) to the *C.briggsae* genome identifies conserved features of the two genomes in the *mpz-1* 5'UTR. The gene *mpz*-1(C52A11.4a) is shown and its alignment to part of the *C.briggsae* supercontig cb25.fpc0058. The four non-coding exons defined by yk1004e08 are circled in red. The dark blue, light blue and grey boxes correspond to coding, strong and weak alignments as determined by the algorithm Woble Aware Bulk Aligner (WABA) (Stein, L et al 2003). Dotted lines represent regions that do not align. Strong blocks of identity cover the *mpz*-1(C52A11.4a) coding DNA sequence. The intergenic region (~6,000bp) of *mpz*-1(C52A11.4a) contains coding, strong and weak blocks of identity that may represent conserved functional elements of the upstream promoter region.

217

6.5.4 Characterization of *mpz*-1 5'end by SL-1 PCR and RACE

In parallel to the EST based definition of the *mpz*-1 gene model an independent RACE and SL-1 analysis of the *mpz*-1 5' region was performed. This was done to confirm the 5' region of the gene model established by sequencing the EST yk1004e08 and to identify if alternative splice variants exist in this domain. If this were to be the case it could have important consequences for generating an *mpz*-1 gene reporter construct.

It has previously been shown by the characterization of the *mgl*-1 transcript that SL-1 can be alternatively *trans*-spliced to generate different 5' ends. The EST yk1004e08 defines one SL-1 *trans*-spliced variant. To establish if there are others a nested PCR reaction was performed on a first round RACE amplification. The first round RACE was performed with the primer pairs, *mpz-1* 5'RACE GSP, located in exon 29 of *mpz*-1(C52A11.4a) and the adaptor specific primer (AP). The nested RACE reaction was performed with an SL-1 primer and the nested gene specific primer (nGSP) T19R located in exon 4 of *mpz*-1 *(figure 6.4B)*.

Three specific bands were generated by the PCR that were not amplified by the single primer control reactions *(see figure 6.4A)*. The three bands were named *mpz*-1(B1), *mpz*-1(B2) and *mpz*-1(B3). *mpz*-1(B1) is the longest product amplified (~500bp) and *mpz*-1(B3) (~100bp) is the shortest *(see figure 6.4)*. *mpz*-1(B2) was closest in size to the expected product, which is 381bp. Each of the three bands were gel extracted, to facilitate TA-cloning of each product and selected clones were then sequenced. The sequence was aligned to the *C.elegans* genome.

The sequencing of *mpz*-1(B1), *mpz*-1(B2) and *mpz*-1(B3) confirmed each is alternatively *trans*-spliced to SL-1. *mpz*-1(B2) corresponds to the 5' end of the SL-1 *trans*-splice variant contained in the EST clone yk1004e08. It confirms the *trans*-splice site and exon splicing at the 5'end identified by yk1004e08. *mpz*-1(B1) and *mpz*-1(B3) are alternative SL-1 *trans*-splice variants that have not been previously characterized. *mpz*-1(B1) is the more robust band compared to *mpz*-1(B2) and *mpz*-1(B3). Qualitatively it represents the most abundant transcript *(figure 6.4A)*. *mpz*-1(B1) is different because it contains two exons upstream of the first exon defined by the EST yk1004e08. The first exon of *mpz*-1(B1) contains an alternative start codon to that assigned by yk1004e08 *(figure 6.4B)*.

The protein sequence of *mpz*-1(B1) does not alter the existing PDZ domain architecture. *mpz*-1(B3) has the same start codon in common with *mpz*-1(B2) but contains only a single non-coding exon upstream of it *(figure 6.4C)*.

In addition the 5' region of the *mpz*-1 transcript was characterized by the SMART RACE technique. A nested RACE reaction was performed using the nested RACE adaptor primer (nAP: nested Adaptor Primer) and the nested gene specific primer (nGSP), T19R. The primer T19R was previously used to perform the SL-1 characterization of *mpz*-1. The nested RACE reaction confirmed the SL-1 *trans*-splice site and the exon splicing defined by the variant *mpz*-1(B1). It also identified shorter variants that were not *trans*-spliced to SL-1 (data not shown). The shorter variants contained the RACE adaptor sequence and alignment did not identify any alternative exon splicing to that already established by the SL-1 analysis. The size difference was assigned to the alternative incorporation of the SMART RACE adaptor sequence.
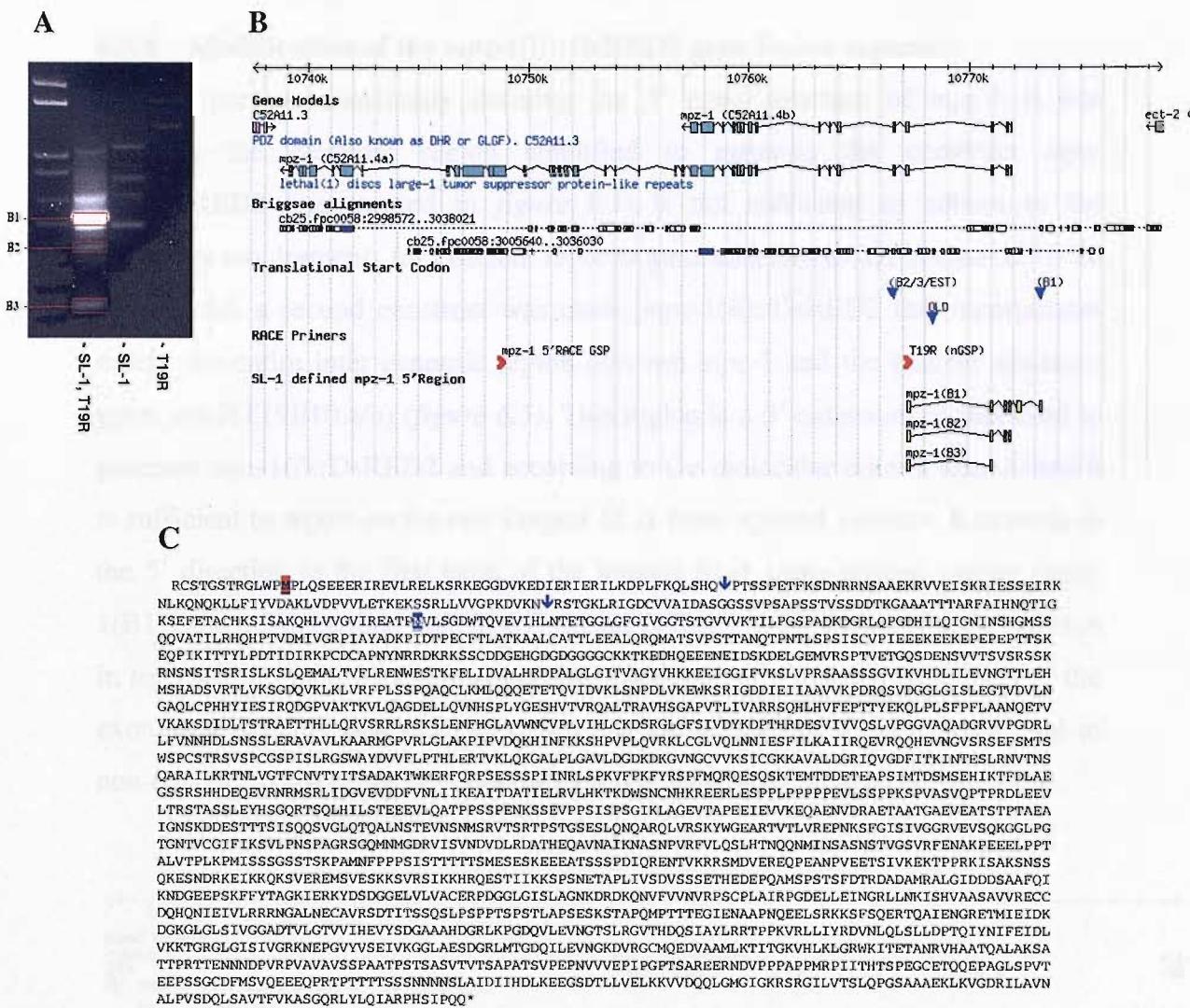
Figure 6.4. The SL-1 characterization of the *mpz*-1 5' region. A) The SL-1 characterization of the *mpz*-1(C52A11.4a) 5' region identified three alternative SL-1 *trans*-spliced variants (*mpz-1(B1)*, *mpz-1(B2)* and *mpz-1(B3)*). **B)** The exon configuration defined by each variant is indicated beneath "SL-1 defined *mpz-1* 5'Region." The position of the first round RACE primer, *mpz-1* 5'RACE GSP and the nested RACE primer, T19R(nGSP) are indicated. The position of the *mpz-1(B1)* defined translational start codon is represented by the blue triangle labelled as (B1). The position of the translational start codon defined by *mpz-1(B2), mpz-1(B3)* and the EST yk1004e08 is indicated by the blue triangle labelled as (B2/3/EST). **C)** The protein sequence of MPZ-1a is shown and has been modified to include the protein sequence of the *mpz-1(B1)* cDNA fragment. A blue arrow indicates the position of alternative SL-1 trans-splicing defined by *mpz-1(B2)* and *mpz-1(B3)*. The ATG start codon defined by the longest SL-1 *trans*-splice variant is in red and the start defined by the shorter SL-1 variants is highlighted in blue.

6.5.5 Modification of the *mpz*-1(i)::DsRED2 gene fusion reporter.

Upon experimentally defining the 5' exon structure of *mpz*-1 it was apparent the genomic region amplified to generate the construct *mpz*-1(*i*)::DsRED2 (represented in *figure 6.5*) is not sufficient to adhere to the necessary requirements for efficient reporter gene expression *(see section 6.3)*. To address this a second construct was made, *mpz*-1(*ii*)::DsRED2 that incorporates nearly the entire inter-genomic region between *mpz*-1 and the nearest upstream gene, *ect*-2(T19E10.a/b) *(figure 6.5)*. This region is a 5' extension of that used to generate *mpz*-1(*i*)::DsRED2 and according to the molecular criteria established it is sufficient to report on the two longest SL-1 *trans*-spliced variants. It extends in the 5' direction to the first exon of the longest SL-1 *trans*-spliced variant (*mpz*-1(B1)) and 4.5kb beyond that. The DsRED2 is fused at the same genomic position in *mpz*-1(*ii*)::DsRED2 as its predecessor. However, due to the remodelling of the exon structure this now corresponds to exon 6 of the *mpz*-1(B1) variant and to non-coding exon 4 of the yk1004e08/*mpz*-1(B2) variant.
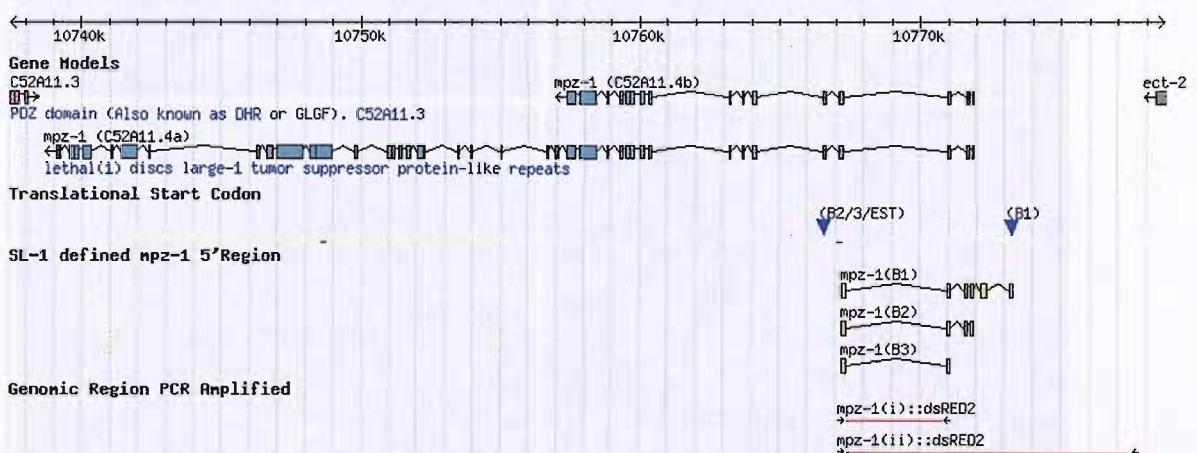
Figure 6.5. The molecular basis of *mpz*-1(*ii*)::DsRED2. The *mpz*-1(C52A11.4a) gene model is shown, together with the alternative 5' exon structures defined by SL-1 *trans*-splicing. The genomic region used to manufacture the construct *mpz*-1(*ii*)::DsRED2 is indicated by the solid red line, labelled as DsRED2. The region used to make the construct *mpz*-1(*i*)::DsRED2 is represented by the red line labelled as *mpz*-1(i)::DsRED2.

6.5.6 Expression analysis of *mpz*-1(ii)::DsRED2

The microinjection of the *mpz-1(ii)*::DsRED2 reporter construct into *mgl-1*::GFP expressing worms did not produce an identifiable co-expression pattern. The expressed *mpz-1(ii)*::DsRED2 reporter construct appeared as discrete punctate aggregations, that were irregular in both size and their intensity of fluorescence. The puncta were distributed in the head, along the length of the worms' body, on the ventral and dorsal surfaces and displayed a more concentrated aggregation at the vulva. The focal plane of the puncta was towards the periphery of the worm, in close proximity to the hypodermis and the cuticle *(figure 6.6)*. This gross distribution of *mpz-1(ii)*::DsRED2 was consistent between three stable transgenic lines.
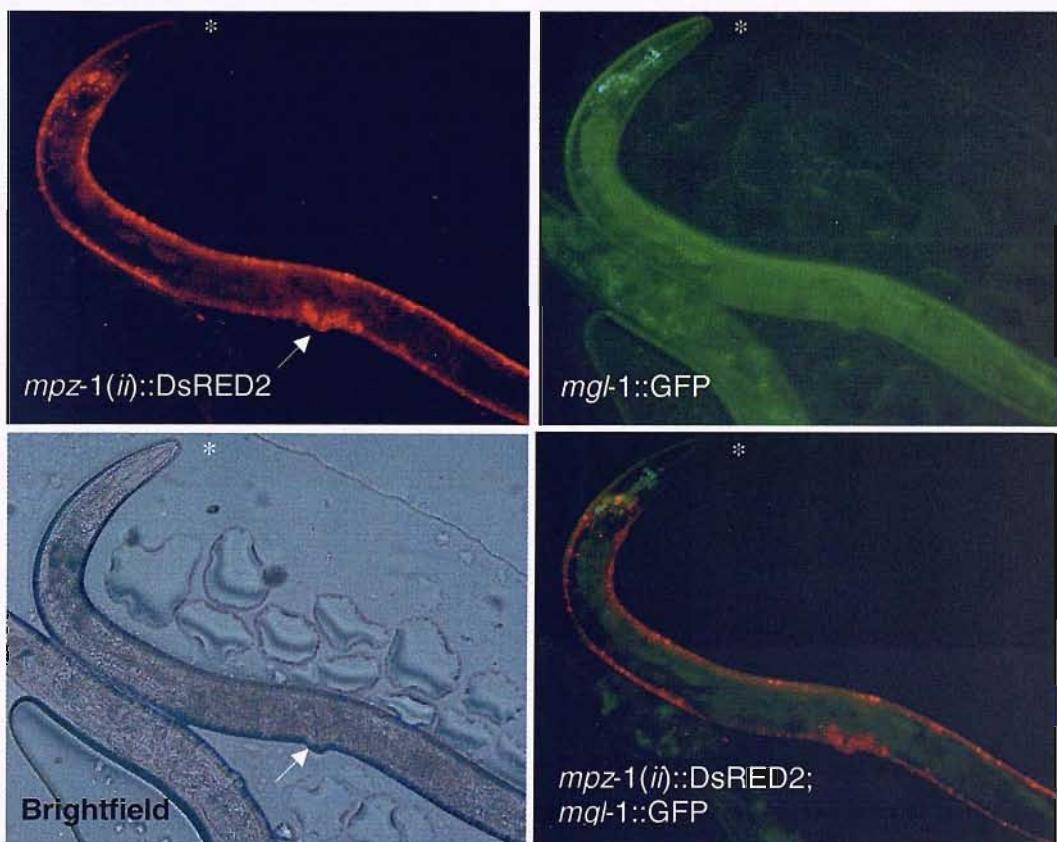
Figure 6.6 The expression pattern of *mpz*-1, defined by the construct *mpz-1(ii)*::DsRED2 in the adult hermaphrodite co-expressing *mgl*-1::GFP. Two worms are present in the image, the asterisk indicates the nose tip of a worm strongly expressing the transgene. The white arrow indicates the position of the vulva and ventral surface. Objective : x40 DIC.

6.5.7 Modification of the *mpz-1(ii)*::DsRED2 gene fusion reporter

The cellular distribution of *mpz-1* reported by the construct *mpz-1(ii)*::DsRED2 correlates poorly with existing descriptions (Xiao, H et al 2004). Abstract publications have described *mpz-1* as having a widespread distribution in *C.elegans*, where it is expressed in both the nervous system and distinct types of muscle. A comparison between the neurons named as expressing *mpz-1* and *mgl-1* by independent abstract publications indicates the two proteins are co-expressed in the nervous system. The nerve ring motorneuron, RMD is common to the expression patterns described for *mpz-1*(Xiao, H et al 2004) and *mgl-1*(Ishihara and Katsura 1996) in separate reports. In light of this, the construct *mpz-1(ii)*::DsRED2 ill defines the complete cellular expression pattern of the gene *mpz-1*.

The construct *mpz-1 (ii)*::DsRED2 was subsequently modified to generate *mpz-1(iii)*::CFP. The construct *mpz-1(iii)*::CFP incorporates two key molecular alterations from its predecessor. The first is the fluorophore DsRED2 was substituted for CFP (Cyan Fluorescent Protein), a modified version of GFP (Heim, R and Tsien, RY. 1996). An alternative fluorescent reporter was used because DsRED has been reported to oligomerize into aggregates when expressed in cells (Sacchetti, A et al 2002). The punctate aggregates yielded by the micro-injection of *mpz-1(iii)*::DsRED2 are a potential artefact of the DsRED2 undergoing self-oligomerization (Sahai, E. Wyckoff, J. et al 2005).

It is well established several promoters can drive the expression of a single gene. Abstract publications have reported, in brief, that *mpz-1* is no exception to this. Alternative promoters are capable of driving the expression of *mpz-1* in different cell types (Xiao, H et al 2004). In these circumstances the amount of genomic sequence used to generate the reporter construct *mpz-1 (ii)*::DsRED2 could have been insufficient and selected for an alternative promoter that drives a previously un-characterized expression pattern of *mpz-1*. To address this, the second modification to construct *mpz-1(ii)*::DsRED2 was made by extending the region of genomic DNA in the 3' direction. In doing so this incorporated the exon containing the start codon defined by the EST yk1004.e08 and the next exon downstream of that, which is fused to the fluorescent reporter protein *(figure 6.7 illustrates the region of promoter used to make the modified version of mpz-1(iii)::DsRED2, denoted mpz-1(iii):CFP/mRFP-1).*

The construct *mpz*-1*(iii)*::CFP was micro-injected into the line *pha-1*(e1213);MGL-1::GFP. Worms successfully injected were selected for and analysed by confocal microscopy. Specific CFP fluorescence could not be detected using filter settings that adhere to the excitation/emission profile of CFP (Clontech, Living Colors User Manual®, PT2040-1(PRY1Y691), 2001). The generalised CFP fluorescence observed was the same in wild-type worms, not transformed with the construct *mpz*-1*(iii)*::CFP and the transgenic lines. Subsequently the CFP was replaced in the construct by the fluorophore mRFP-1(monomeric red fluorescent protein-1) (Campbell, RE et al 2002 via Hannah, M). The construct *mpz*-1*(iii)*::mRFP-1 was injected into the line *pha-1*(e1213);MGL-1::GFP and the successful transgenics were analysed by normaski and confocal microscopy.
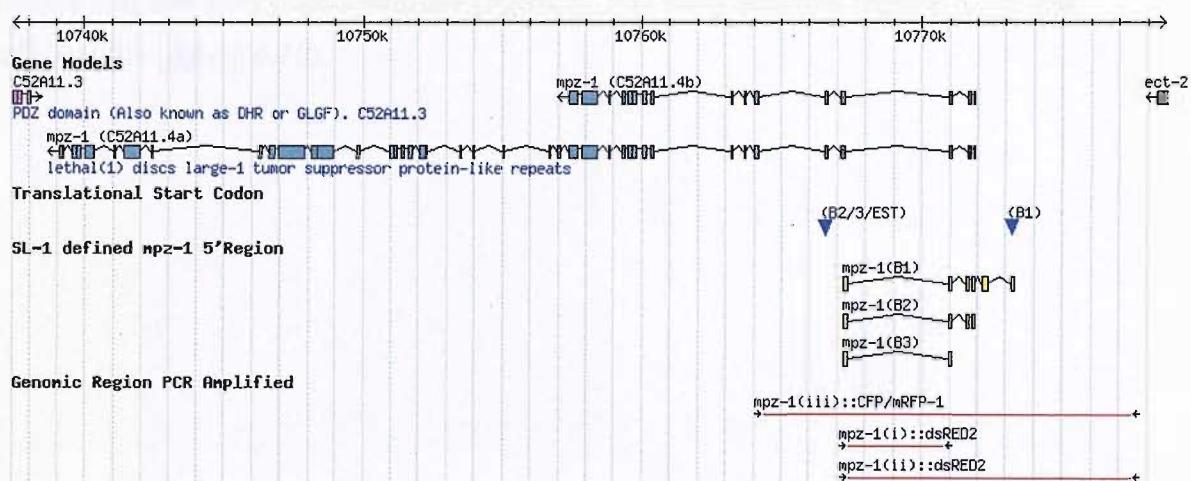
Figure 6.7 Molecular basis of the construct *mpz*-1(*iii*)::CFP. The red line labelled mpz-1(iii)::CFP/mRFP-1 indicates the genomic region used to make the constructs by the same name.

6.5.8 Analysis of the expression pattern reported by *mpz*-1(iii)::mRFP-1.

The cellular expression of *mpz*-1 reported by the construct *mpz*-1*(iii)*::mRFP-1 was significantly different to *mpz*-1*(ii)*::DsRED2. The intense punctate aggregates reported in the first instance with *mpz*-1*(ii)*::DsRED2, were predominantly absent. Instead, the fluorescence emitted by mRFP-1 was diffuse in the majority of cell types. In some circumstances the fluorescence appeared as puncta superimposed upon a weak, diffuse background and this was typically associated with both muscle and neuronal cells (*see figures 8C and 12*).

The cell types identified as expressing *mpz*-1 by the construct *mpz*-1*(iii)*::mRFP-1 generally correlated with the existing reports (Xiao, H et al 2004). Weak fluorescence was detectable in the body wall muscle cells, extending from the head to the tail, the vulva muscle cells (*see figure 6.8*) and the nervous system. The expression within the nervous system encompassed nerve-ring neurons (*figure 9B*) and pharyngeal neurons (*figure 6. 9A*) body neurons (*figure 6.10*) and tail neurons (*figure 6.10*),
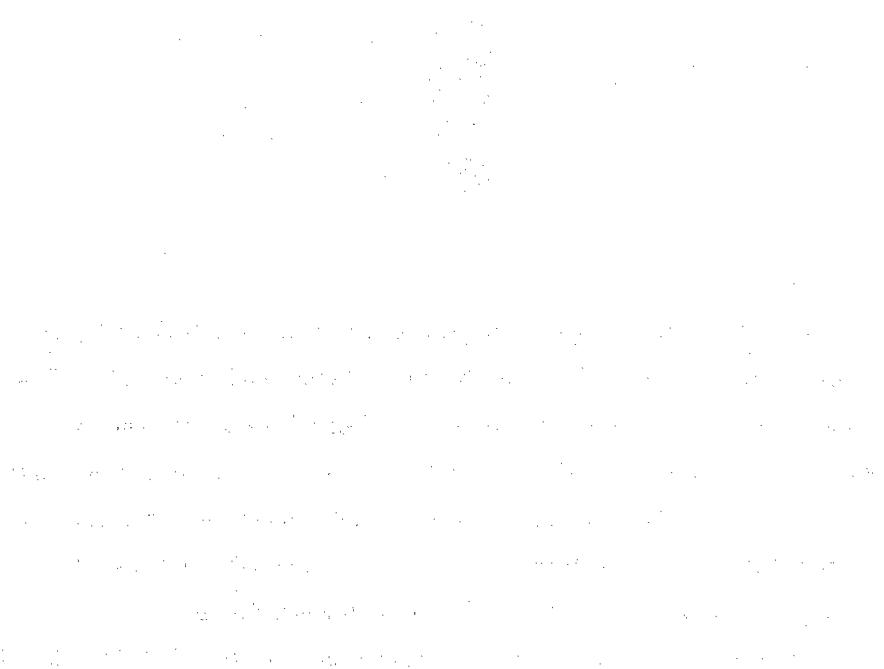
Ai

Aii

Vulval Opening

Vulval Muscle Cells

Vulval Muscle Cells

Vulval Muscle Cells

Vulval Muscle Cells

Vulval Opening

Lateral Hypodermis

Vulval Muscle Cells

B

Body wall muscle cells

Body wall muscle cells

Nerve ring Co-expression
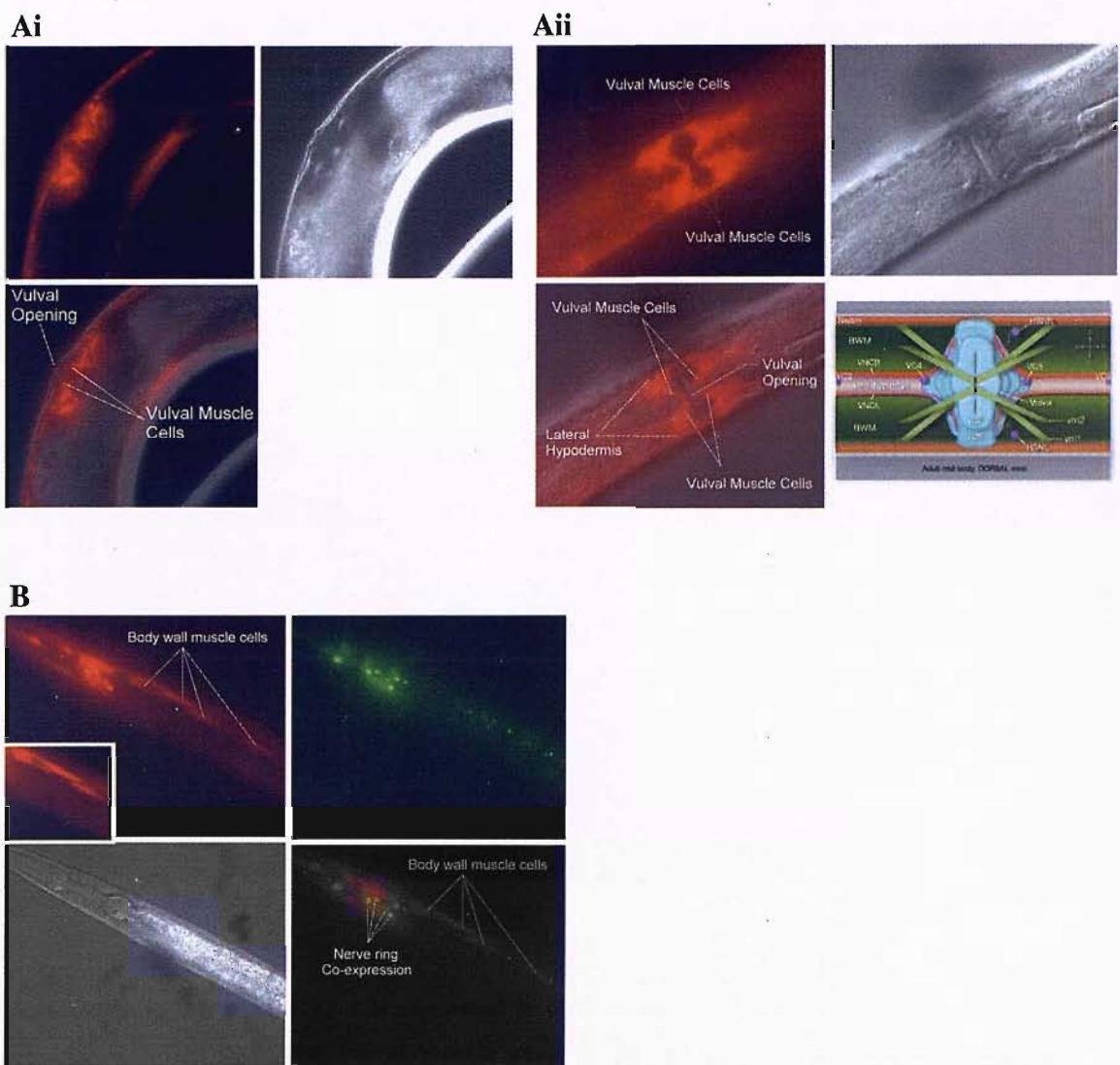
Figure 6.8 *mpz*-1::mRFP-1 **is expressed in distinct muscle cell types. A) The expression of** *mpz*-1::mRFP-1 **in the vulval muscle cells of the adult** *C.elegans* **hermaphrodite.** The expression in the vulva muscles is imaged in two planes. In **Ai)** the worm is positioned on its side and in **Aii)** the worm is positioned with the ventral surface facing upwards. In Aii a cartoon of the adult vulva (in the same orientation as the microscope images) depicts the positioning of the vulva muscle cells (vm) and the organization of the neuronal innervations. Objective 60x oil DIC. **B) The expression of** *mpz*-1(*iii*)::mRFP-1 **in the body wall muscles of a young L3 hermaphrodite.** MGL-1::GFP is co-expressed with *mpz*-1(*iii*)::mRFP-1 in neurons of the nerve ring and not in the body wall muscle. Objective 60x Oil DIC.

226

The cells expressing *mpz-1(iii)*::mRFP-1 in the pharyngeal nervous system were amongst the most consistent, between individual transgenic worms. In the de-sheathed pharynx *mpz-1* expression could be observed in 12 neurons of the pharyngeal nervous system. Of these 12 neurons MGL-1::GFP fluorescence was present in them all. Independent analysis has identified *mpz-1* is expressed in the motorneurons M4 and NSM of the pharyngeal nervous system, which both have distinct anatomical morphologies (*figure 9A*). Our own analysis has confirmed *mpz-1* is expressed in these cells and enabled the cellular expression of MGL-1::GFP to be further defined, as GFP fluorescence is observed in NSM and M4 also.

By comparing reports that describe separately the expression of *mgl-1* and *mpz-1*, a common neuron belonging to the nerve-ring was identified as expressing both genes this was RMD. Our analysis has identified *mpz-1* and MGL-1 are co-expressed by neurons within the vicinity of the nerve ring corresponding to the anatomical location of this cell. The expression of *mpz-1(iii)*::mRFP-1 can be seen in the nerve ring as early as the developmental stage L2 in the same neurons as MGL-1::GFP (*figure 6.8B*).

A.
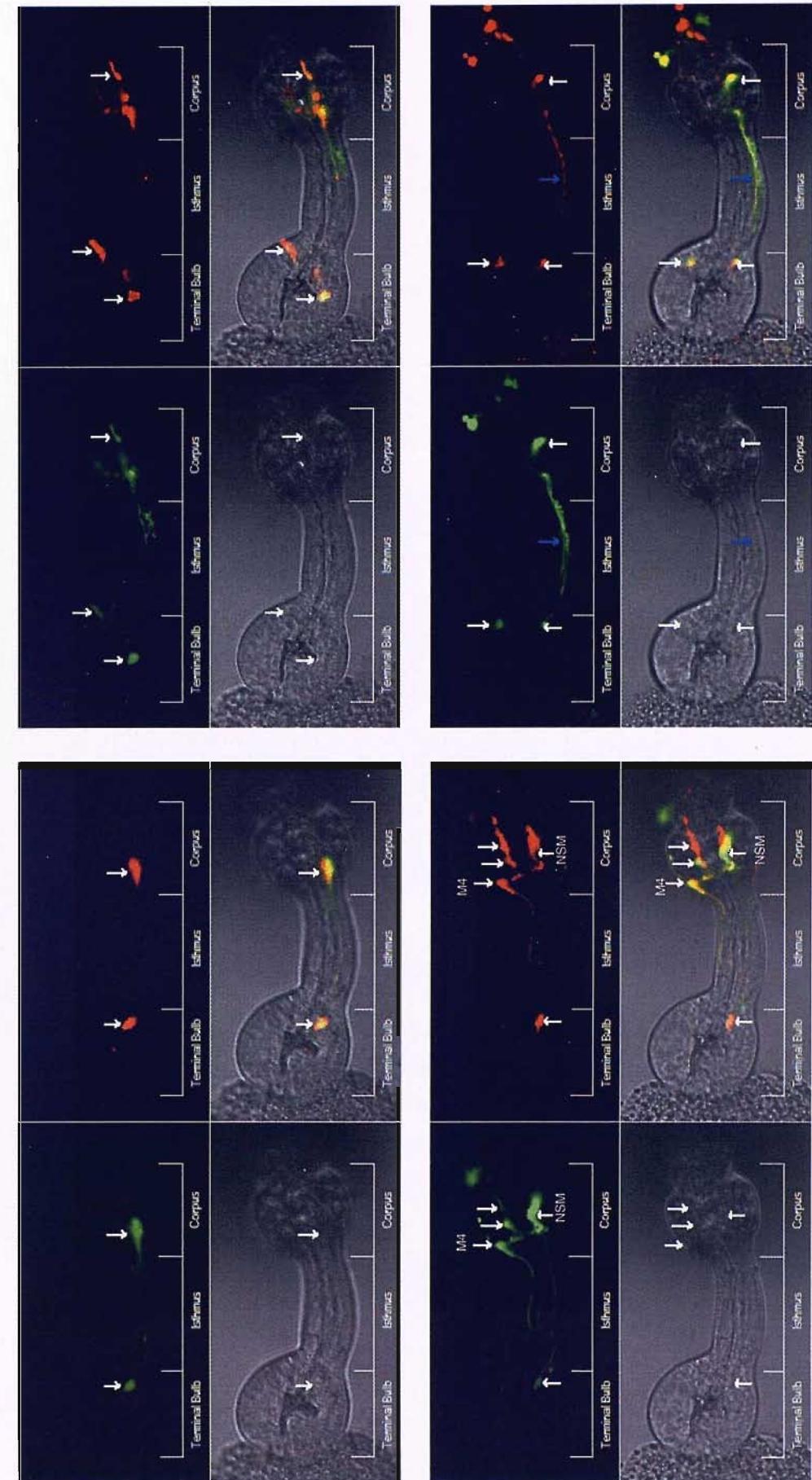
B.
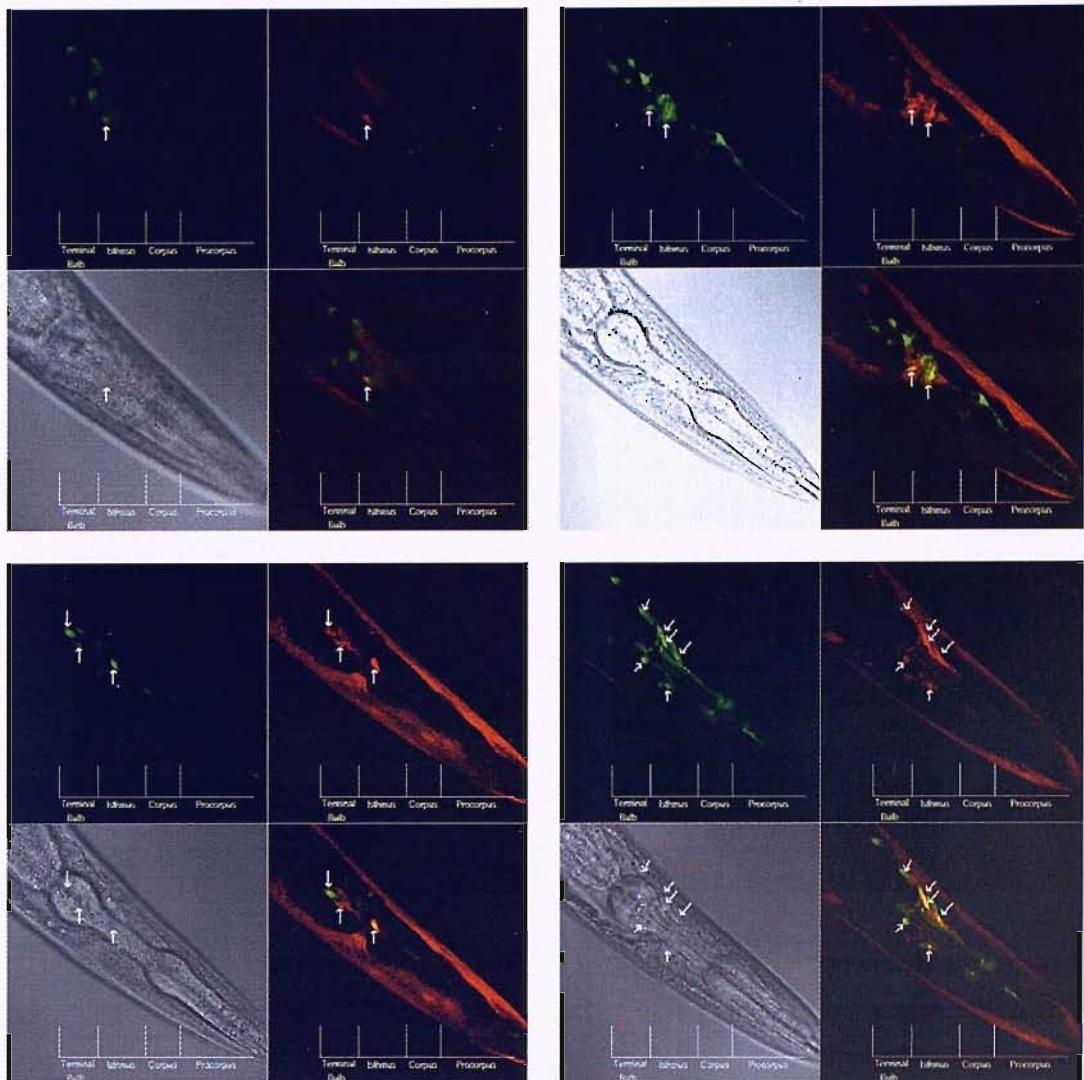
Figure 6.9 The expression of *mpz*-1::mRFP1 in the nerve ring and de-sheathed pharynx of an adult hermaphrodite *C.elegans* expressing MGL-1::GFP. The images displayed are representative sections taken from a Z scan performed on the de-sheathed pharynx (A) and head (B) of a single worm. White arrows point toward cells co-expressing *mpz*-1::mRFP-1 and *mgl*-1::GFP and blue arrows point towards co-expression in neuronal processes. Scan mode: Z Stack, interval 0.98μm; Objective: Plan-Neoflaur40x/1.3 Oil DIC. GFP excitation 488nm excitation, 505-530 band pass filter; mRFP1 excitation 453nm, 560-615 band pass filter.

Body neurons previously identified as expressing *mpz*-1 are AVM, SDQR, PVM, SDQL, HSNL/R. Using this description body neurons expressing *mpz*-1*(iii)*::mRFP-1 have been proposed. A neuronal expression corresponding to HSNL/R was not identified. Tail neurons independently identified as expressing *mpz*-1 are ALNL/R, PVCL/R, PVQL/R and PVNL/R. The expression of *mpz*-1*(iii)*::mRFP-1 in the tail neurons was inconsistent between transgenic worms. In

the majority of worms the expression of *mpz*-1*(iii)*::mRFP-1 in tail neurons was not overt. The mRFP-1 fluorescence was not observed in the same neurons of the tail as MGL-1::GFP.
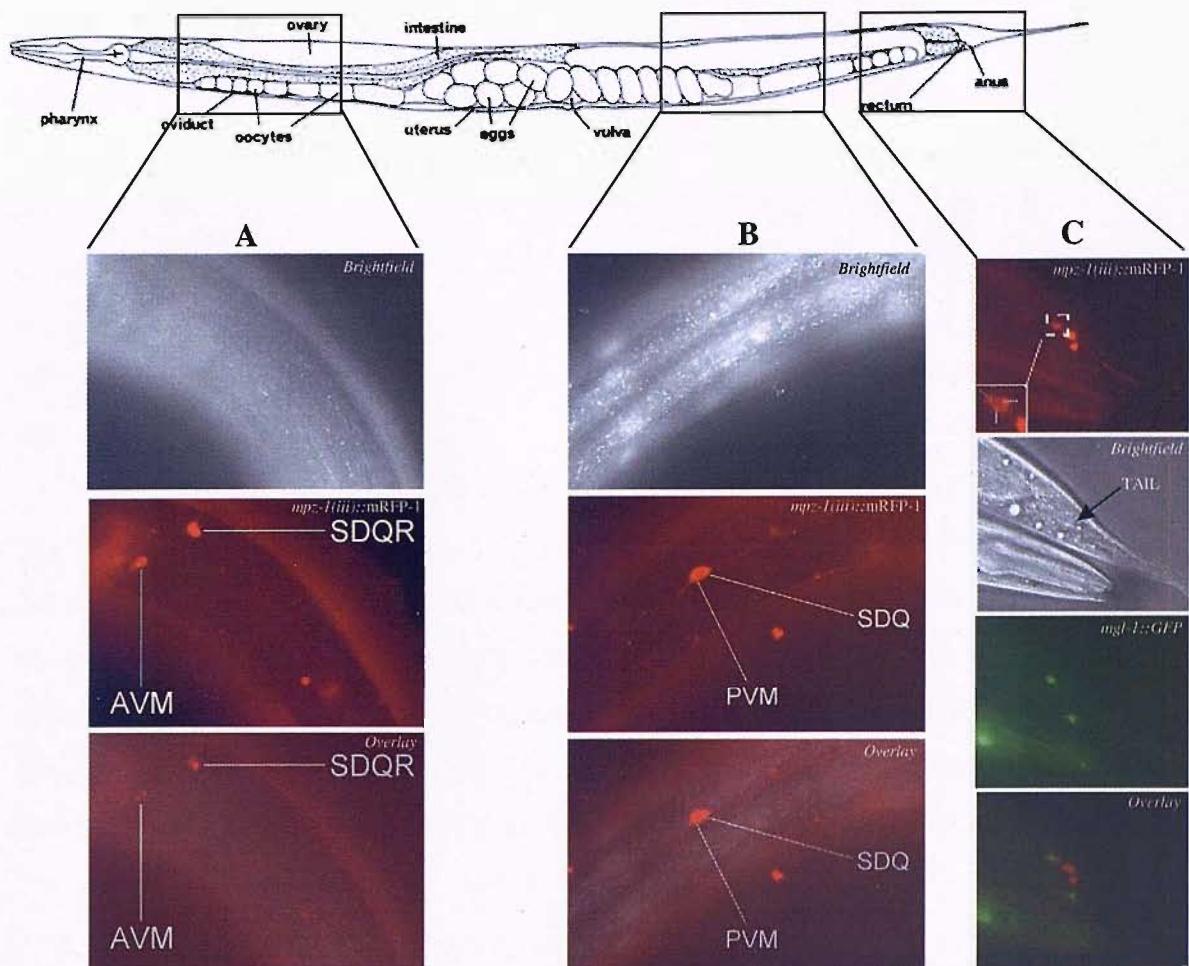
Figure 6.10 The expression of *mpz*-1*(iii)*::mRFP-1 in mid-body and tail neurons of an adult hermaphrodite. The identities of neurons expressing mRFP-1 are proposed on the basis of their similar anatomical localization to those neurons that have been previously reported as expressing MPZ-1 (Xiao, H et al 2004). The region of the worm the image was taken of is indicated **A)** Image recorded anterior to the vulva **B)** Image recorded posterior to the vulva. *Objective: 60x, Oil DIC. Exposure time 8.5 seconds.*

C) Image recorded from the tail. Three neurons of the tail were identified as expressing mRFP-1. The inset highlights an example of where a combination of diffuse and punctate expression was observed in the same cell. The overlay shows *mgl-1*::GFP was not co-expressed with *mpz-1(iii)*::mRFP-1 in the tail cells. *Objective: 60x Oil DIC. Exposure time: 5 seconds.*

230

6.6 The RACE characterization of *tag-60(C01F6.6a)* and its cellular expression

6.6.1 Introduction

The gene *tag*-60 (temporarily assigned gene name) encodes two splice variants, C01F6.6a and C01F6.6b *(figure 6.11)*. The two variants have alternative N-terminals and this produces the different domain architecture in the two proteins. The protein C01F6.6a contains an N-terminal PDZ domain whereas C01F6.6b does not, but both possess a common C-terminal PDZ domain. Since C01F6.6b was not identified from the 2-hybrid screen it is likely the C-terminal is not responsible for the interaction with the MGL-1 C-terminal. Currently there is no information available for the biological function of *tag-60* or the proteins it encodes, in *C. elegans*.

The clone identifying C01F6.6a as a 2-hybrid interaction partner of the MGL-1 C-terminal was incomplete and lacked the predicted start codon. Subsequently a first round 5' RACE characterization was performed with a GSP that recognised coding sequence common to both C01F6.6a and C01F6.6b and was located over exons 6 and 7. An internal control was performed with the 5' RACE GSP and the 3' RACE GSP. The 3' RACE GSP was designed sufficiently towards the N-terminal of *tag*-60(C01F6.6a), in exon 3, to allow specificity for this transcript.
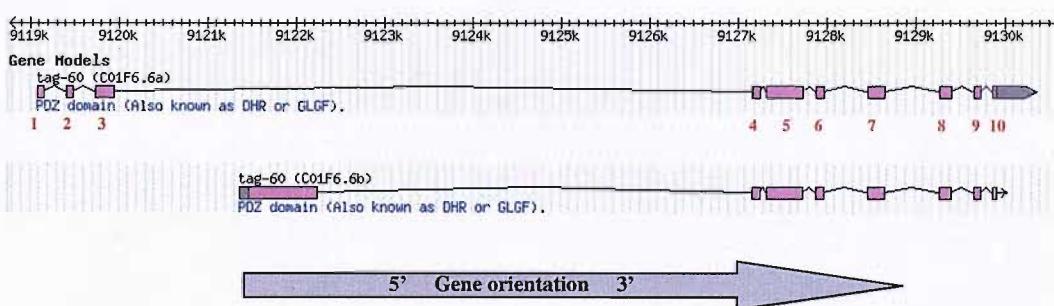
Figure 6.11 The *tag*-60(C01F6.6a) and *tag*-60(C01F6.6b) gene models (WormBase website, *http://www.wormbase.org*, WS152, 1/2006). The *tag*-60(C01F6.6a) gene model contains 10 exons and these are numbered in red. Exons 4 to 10 of *tag*-60(C01F6.6a) are common to *tag*-60(C01F6.6b).

6.6.2 5'RACE Characterization of *tag*-60(C01F6.6a)

Both the 5' RACE reaction and the internal control produced multiple bands that were not predicted from the *tag*-60(C01F6.6a) gene model *(figure 6.12B)*. To enrich for *tag*-60(C01F6.6a) RACE products a nested PCR was performed with the 5' RACE reaction as the template. The nGSP was located in the second exon of C01F6.6a and was used to select against the transcript C01F6.6b. The nested RACE reaction produced two prominent bands, superimposed upon a feint smear *(figure 6.12C)*. An aliquot of the nested PCR was used in a TA cloning reaction and four unique clones were selected by size.

The longest RACE product cloned (205bp) corresponds to the larger of the two prominent bands visualised. It identified C01F6.6a is *trans*-spliced to SL-1 and analysis of the genomic sequence confirmed that the *trans*-splice site is very similar to the consensus SL-1 *trans*-splice sequence. Alignment of the RACE products' sequence to the C01F6.6a gene confirmed the exon-intron boundaries of exon 1-intron 1- exon 2 and it identified the start codon is correctly defined by the available gene model *(figure 6.12C)*. The shorter RACE products were not *trans*-spliced to SL-1 and alignment of the sequences to the 5'end of C01F6.6a revealed they were not representative of alternative splice variants. Instead each product was generated by the alternative incorporation of the RACE adaptor at the 5' end of the transcript.
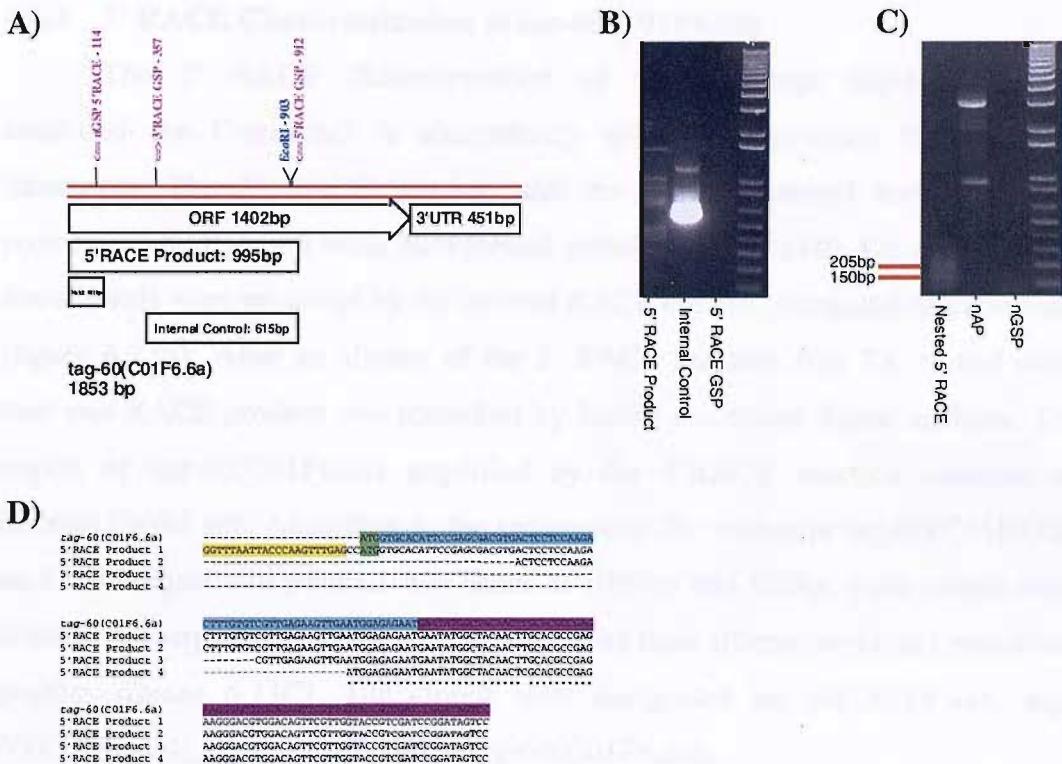
Figure 6.12 The 5' RACE characterization of the gene _tag_-60(C01F6.6a). A) An outline of the 5'RACE approach documenting the primer annealing sites and the product sizes predicted by the available _tag_-60(C01F6.6a) sequence. **B)** _tag_-60(C01F6.6a) 5'RACE reaction. Primer pairs used were: _lane 1_. AP & GSP; _lane 2_. 5' RACE GSP & 3' RACE GSP; _lane 3_. GSP (-tive control). _Cycle_: 5x (94°C 5sec, 72°C 3min); 5x (94°C 5sec, 70°C 10sec, 72°C 3min); 25x (94°C 5sec, 68°C 10sec, 72°C 3min). **C)** _tag_-60(C01F6.6a) nested 5'RACE analysis. Primer pairs used were: _lane 1_. nGSP & nAP; _lane 2_. nAP (-tive control); _lane 3_. nGSP (-tive control). _Cycle_: 25x (94°C 5sec, 68°C 10sec, 72°C 3min). 1µl of a 1:50 dilution of the 1st round reaction was used as the template. **D)** Alignment of the four RACE products' sequence to the _tag_-60(C01F6.6a) 5' end (Performed using ClustalW 1.82). SL-1 sequence is highlighted in yellow and the start codon is highlighted in green. Exon 1 and 2 are highlighted in blue and purple respectively.

233

6.6.3 3' RACE Characterization of *tag*-60(C01F6.6a)

The 3' RACE characterization of the transcript *tag*-60(C01F6.6a) identified the C-terminal is alternatively spliced to generate three distinct transcripts. The 3' RACE reaction and the internal control both yielded a prominent band upon a weak background signal *(figure 6.13B)*. On this occasion fewer bands were produced by the internal RACE control, compared to previously *(figure 6.12B)*. After an aliquot of the 3' RACE reaction was TA cloned more than one RACE product was identified by EcoRI restriction digest analysis. The region of *tag*-60(C01F6.6a) amplified by the 3'RACE reaction contains an internal EcoRI site. According to the sequence of the transcript *tag*-60(C01F6.6a) an EcoRI digest will produce two bands at 1050bp and 580bp. Four clones were selected for sequencing and this was based upon their alternative EcoRI restriction profiles *(figure 6.13C)*. The clones were designated *tag*-60(C01F6.6a)$_1$, *tag*-60(C01F6.6a)$_2$, *tag*-60(C01F6.6a)$_3$, *tag*-60(C01F6.6a)$_4$.

All of the variants contain a characteristic poly (A$^+$) tail attached at the 3' end. The four transcripts can be divided into two pairs and each pair shares a common attachment site of the poly(A$^+$) tail. The difference between the points of attachment is two nucleotides. The clone *tag*-60(C01F6.6a)$_1$ is consistent with the EcoRI restriction profile predicted by the gene model documented in the Wormbase repository *(figure 6.13D)*. The clone *tag*-60(C01F6.6a)$_2$ is the largest transcript characterised (1680bp) and it contains sequence defined as being intronic by the *tag*-60(C01F6.6a) gene model. The additional sequence corresponds to a new exon located between exons currently numbered as 3 and 4 and it causes one amino acid substitution (S^{120}→R) in addition to 20-odd additional amino acids *(figure 6.13E)*. Other than this the downstream reading frame, the domain architecture and the stop codon is the same as that for *tag*-60(C01F6.6a). The clone *tag*-60(C01F6.6a)$_3$ has this new exon in common with the transcript *tag*-60(C01F6.6a)$_2$ but in addition exon 7 is spliced out *(figure 6.13D)*. When these changes are made to the nucleotide sequence of C01F6.6a, the protein reading frame, the domain architecture and the stop codon remain the same *(figure 6.13E)*. The *tag*-60(C01F6.6a)$_4$ variant is the same as the transcript *tag*-60(C01F6.6a)$_1$ except for a point mutation that converts a T to a C and introduces a second EcoRI site. It is not known if this is a site of mRNA editing or an error introduced by the PCR. The amino acid encoded by the tri-nucleotide

containing the substitution (F^{126}) is unaltered. Sequence obtained from the Y-2-H clone, identified as interacting with the MGL-1 C-terminal is sufficient to determine it is consistent with the splice variant *tag*-60(C01F6.6a)$_1$. The full protein sequence of *tag-60(C01F6.6a)* is shown in *figure 6.13E,* the sequence confirmed by 5' and 3' RACE is highlighted for each variant.
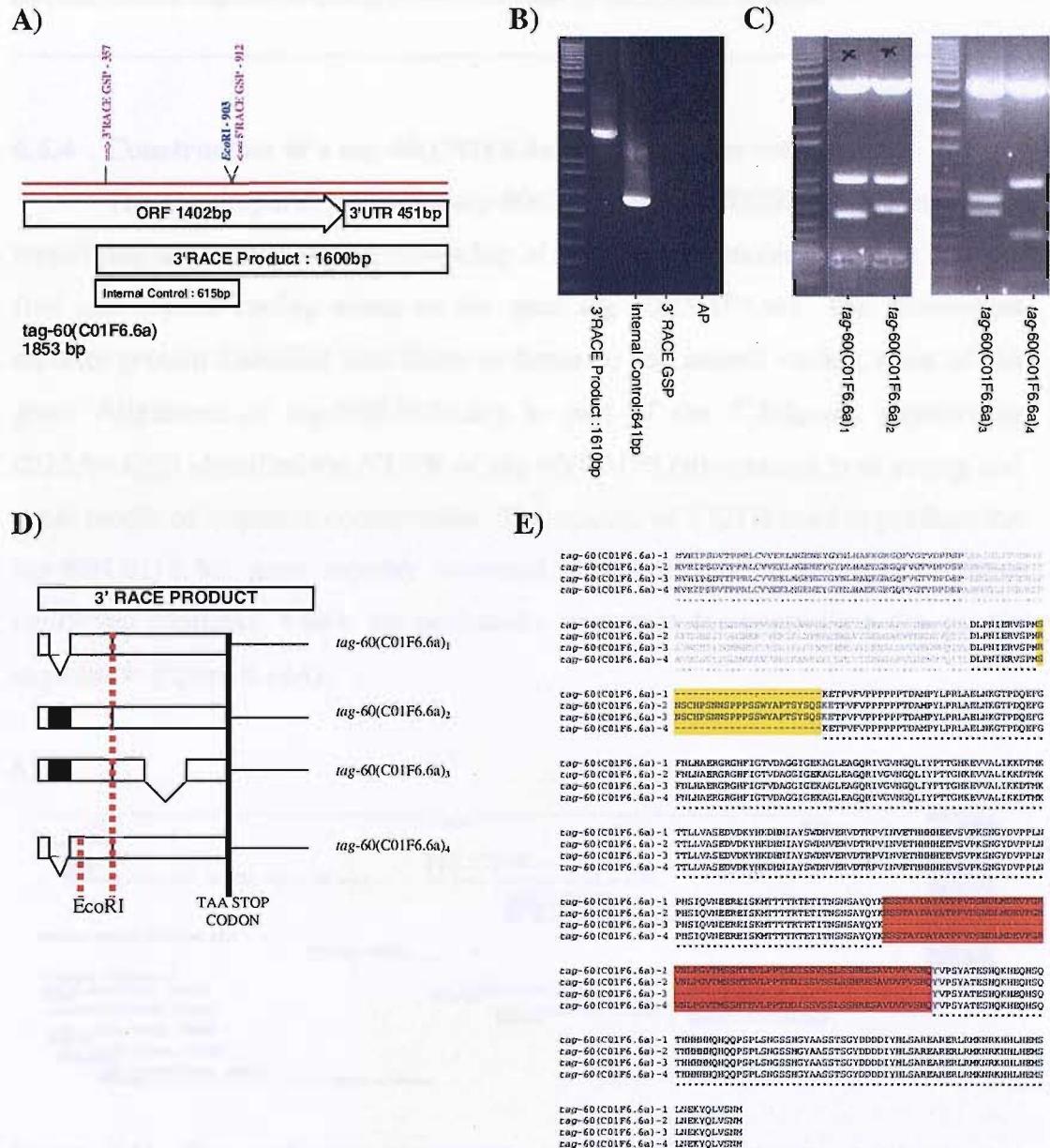
Figure 6.13 *tag*-**60(C01F6.6a) is alternatively spliced at the 3' end. A**) An outline of the 3' RACE approach documenting the predicted product sizes and the EcoRI restriction site in base pairs. **B**) The 3'RACE reaction and internal control PCR both produced a prominent band of the expected size. **C**) EcoRI restriction analysis of TA cloned 3'RACE products. The restriction profiles produced were: (C01F6.6a)$_1$ 580bp and 1030bp; (C01F6.6a)$_2$ 652bp and 1030bp;

(C01F6.6a)₃ 652bp and 826bp; (C01F6.6a)₄ 1030bp, 442bp and 138bp (the third fragment at 3.9kb corresponds to the pCR II vector backbone). **C)** Diagrammatic representation of the alternatively spliced *tag*-60(C01F6.6a) 3' RACE products. Chevrons indicate sequence spliced out corresponding to an exon; white boxes are coding sequence; the black box is a newly identified coding exon; lines are 3'UTR sequence; the dashed red line indicates EcoRI restriction sites. **D)** Alignment of the protein sequences encoded by the *tag*-60(C01F6.6a) mRNA splice variants. Protein sequence in grey is not contained within the region amplified by the 3'RACE or 5' RACE reaction. Protein sequence in dark grey was confirmed by the 5'RACE analysis.

6.6.4 Construction of a *tag*-60(C01F6.6a) gene reporter construct.

The gene reporter construct *tag*-60(C01F6.6a)::DsRED2 was generated by amplifying a genomic region consisting of 3.3Kb of promoter sequence and the first and second coding exons of the gene *tag*-60(C01F6.6a). The fluorescent reporter protein DsRED2 was fused in-frame to the second coding exon of the gene. Alignment of *tag*-60(C01F6.6a) to part of the *C.brigssae* supercontig cb25.fpc4260 identified the 5'UTR of *tag*-60(C01F6.6a) contains both strong and weak blocks of sequence conservation. The amount of 5'UTR used to produce the *tag*-60(C01F6.6a) gene reporter construct was sufficient to incorporate these conserved elements, which are potentially important for controlling the gene's expression *(figure 6.14A)*.

A) B)

—3700bp
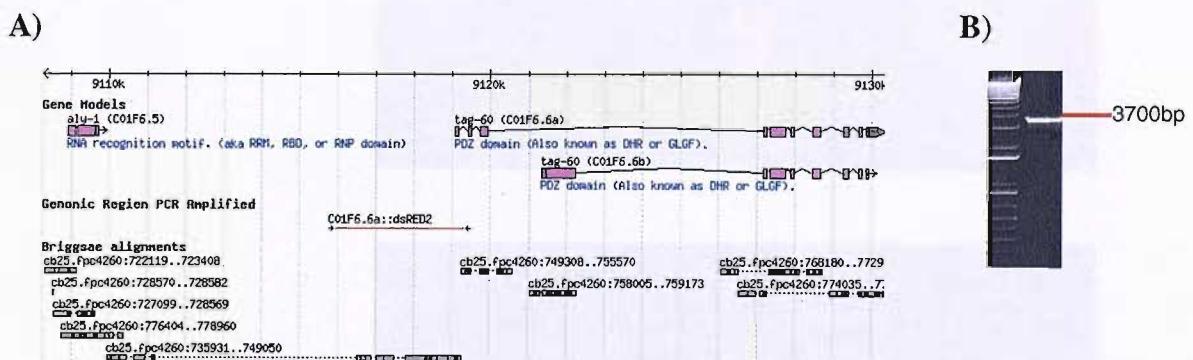
Figure 6.14 The molecular components of the gene reporter construct *tag*-60(C01F6.6a)::DsRED2. **A)** The gene *tag*-60(C01F6.6a) is shown in relation to the nearest upstream gene, *aly*-1 (located ~9kb away). The 'genomic region PCR amplified' and fused to DsRED2 is indicated by the red line labelled as C01F6.6a::DsRED2. The alignment of *tag*-60(C01F6.6a) to the *C.brigssae* supercontig cb25.fpc4260 is shown. **B)** PCR amplification of the *tag*-60(C01F6.6a) genomic fragment was performed on the cosmid C01f6.6a and produced a

single band of the correct size, 3.7Kb. *Cycle: 1x(2min@94°C), 10x(10s@94°C, 30s@57°C, 3min30s@68°C), 20x(15s@94°C, 30s@57°C, 3min35s@68°C), 7min@72°C.*

6.6.5 Expression analysis of the construct *tag*-60(C01F6.6a)::DsRED2.

The microinjection of *mgl*-1::GFP expressing lines with the gene reporter construct *tag*-60(C01F6.6a)::DsRED2 produced two stable lines with a consistent expression pattern. The gene is expressed anatomically within close proximity to *mgl*-1 but is not co-expressed within the same cells *(see below)*. The *tag*-60(C01F6.6a)::DsRED2 is expressed in both muscle cells and neurons. Intense fluorescence is distributed throughout the intestine but in the pharynx it is more specific and selectively confined to distinct pharyngeal muscle cells, pm3 and pm4. A neuronal expression of *tag*-60(C01F6.6a)::DsRED2 was detectable in both the pharynx and the tail. In the pharynx *tag*-60(C01F6.6a)::DsRED2 expression was identified in three neurons, whose cell bodies were situated in the terminal bulb.
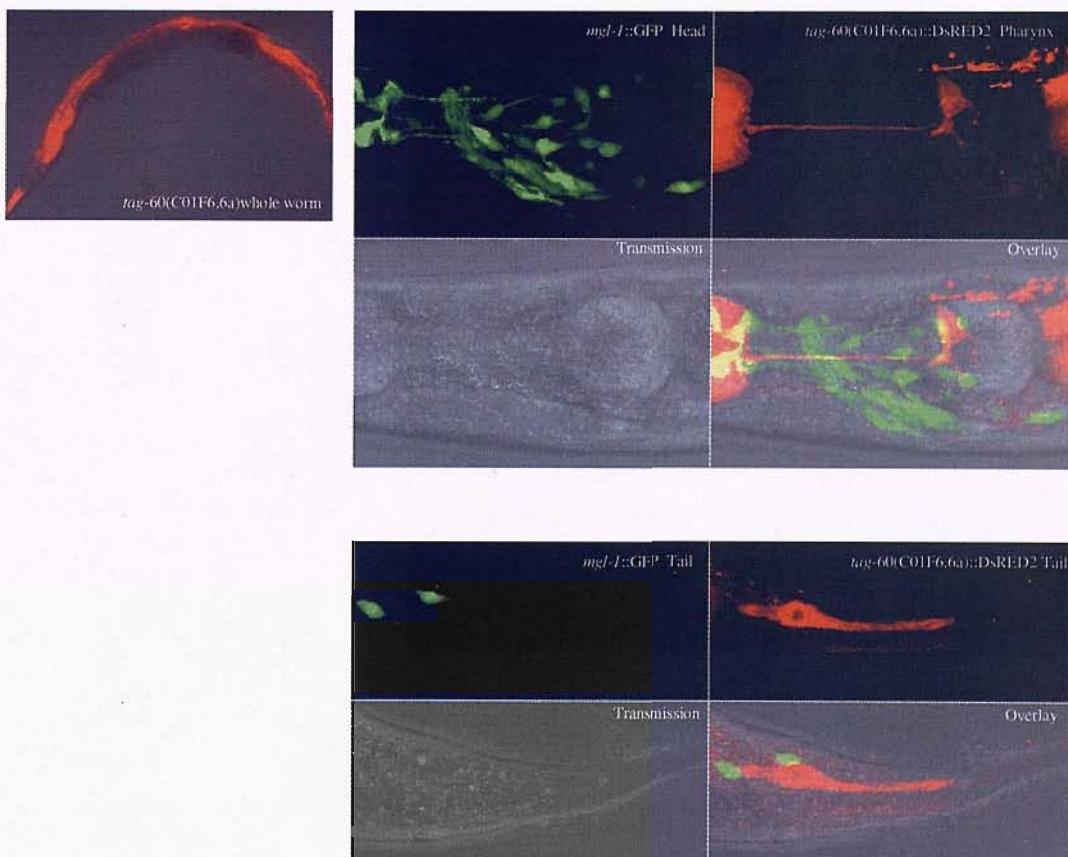
Figure 6.15 Transgenic lines expressing *mgl*-1::GFP and *tag*-60(C01F6.6a)::DsRED2. Worms were immobilised in 0.5% NaAzide and mounted on microscope slides with 2% agarose. **A)** *tag*-

60(C01F6.6a)::DsRED2 is expressed within a specific subset of pharyngeal muscle cells and the intestine. **B)** A more detailed analysis using confocal laser-scanning microscopy reveals a neuronal distribution in the region of the **(i)** head and **(ii)** tail. Images shown are representatives selected from a z-stack performed on an adult hermaphrodite. *Confocal settings: Objective: 63x Oil DIC. GFP excitation 488nm, band pass filter 505-530nm. DsRED2 excitation 543nm, long pass filter 560nm. Z-slice width: 0.9μM.*

6.7 The RACE characterization of ptp-1 and its cellular expression

6.7.1 Introduction

As described in *Section 5.14* the gene *ptp*-1 encodes three splice variants, PTP-1a, PTP-1b and PTP-1c, which are alternatively spliced at the N-terminal. PTP-1a is the longest splice variant and is the only one to contain an N-terminal FERM domain. The Y-2-H clone identified as interacting with the MGL-1 C-terminal corresponds to a common C-terminal region of each PTP-1 splice variant and does not distinguish between them. Without knowing which splice variants can interact with the C-terminal of MGL-1, it was decided a gene reporter construct would be initially produced for the longer splice variant PTP-1a.

6.7.2 5' RACE Characterization of *ptp-1a*

Both the 5'RACE GSP and the 3'RACE GSP were designed against exons common to each of the three *ptp*-1 splice variants. The 5'GSP was positioned in exon 11 of *ptp*-1a (designated exon 4 in *ptp*-1b and *ptp*-1c gene models). This was initially done to obtain race products corresponding to the three transcripts, should this information be required for the design of splice variant specific gene reporter constructs. The 3'GSP was positioned in exon 9 of *ptp*-1a (designated exon 2 in *ptp*-1b and *ptp*-1c gene models). The expected RACE products are shown in *figure 6.15A*.

The first round 5'RACE reaction produced multiple bands, none of which precisely correspond to the amplification size predicted for each of the three splice variants (ptp-1a: 1742bp; ptp-1b: 536bp and ptp-1c: 417bp. Sizes given include the adaptor sequence, which is 27nt) *(figure 6.15B)*. The internal control produced a prominent band of the predicted size (358bp) together with a weaker background signal. Subsequently a nested 5'RACE PCR was performed on the first round reaction. The nGSP was designed against the third exon of *ptp*-1a and was used to specifically enrich for 5'RACE products of this transcript.

The nested RACE reaction amplified three prominent bands, with a size range of 525bp, 433bp and 336bp *(figure 6.15C)*. The largest nested RACE product defined the ATG start codon assigned by the gene model; it identified the transcription start site as –42bp upstream of the ATG; it confirmed the intron/exon

boundaries of exons 1, 2 and 3. The alignment of the three RACE products identified the two shorter variants are produced by the alternative incorporation of the RACE adaptor sequence and do not identify any alternative splicing within the 5' end of the *ptp*-1a transcript *(figure 6.15D)*. The *ptp*-1a transcript was not identified as being *trans*-spliced to neither SL-1 nor SL-2.
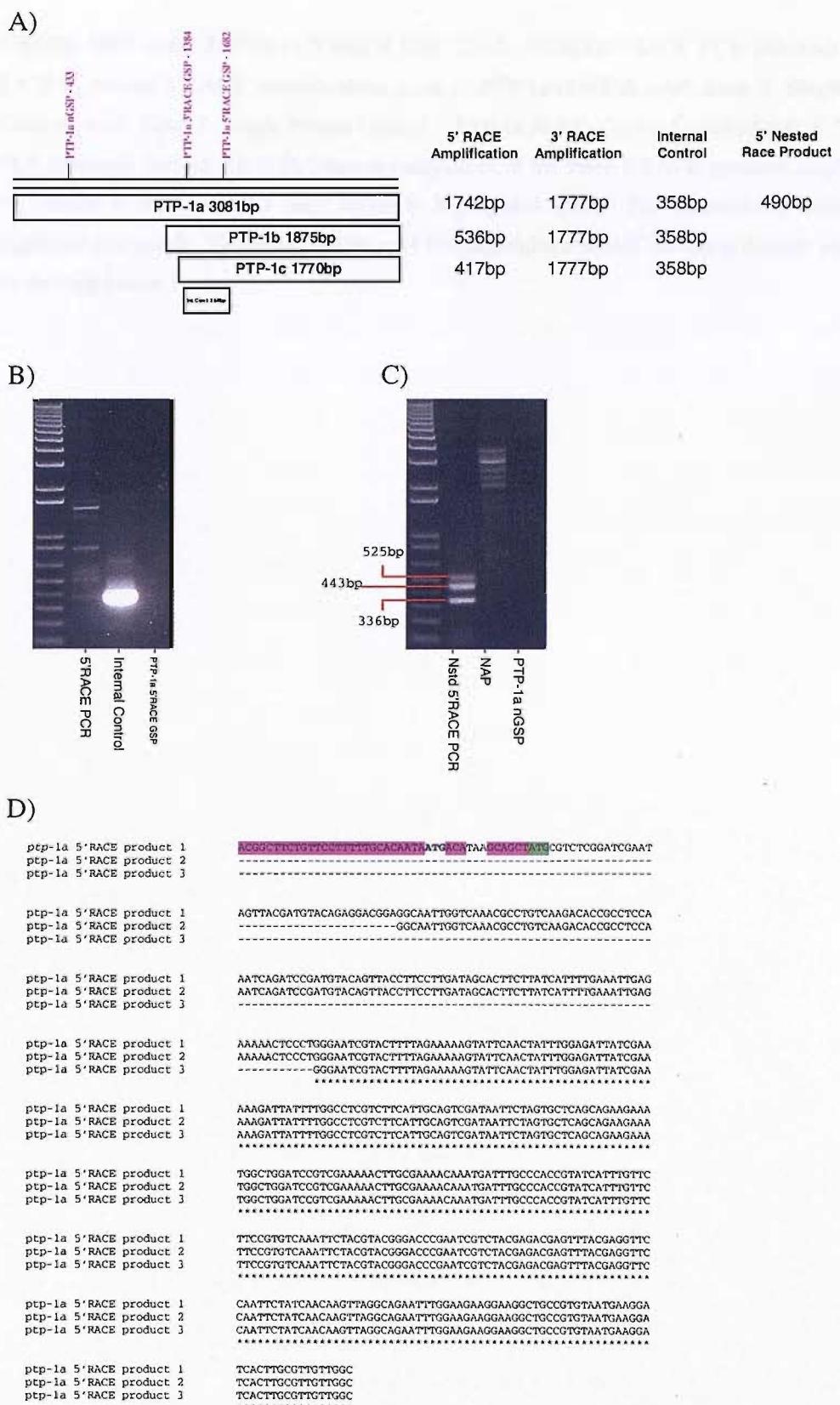
Figure 6.16 5'RACE characterization of *ptp*-1a. A. The annealing sites of the RACE primers PTP-1a 3'RACE GSP, PTP-1a 5'RACE GSP and PTP-1a nGSP are shown in relation to the three *ptp*-1 transcripts. Together with the predicted amplicon size for each RACE reaction. **B.** 1St Round

3'RACE GSP. *Lane 3:* PTP-1a 5'RACE GSP. *Cycle: 5'SMART RACE PCR (Methods Section 2.4.9)* **C.** Nested 5'RACE Amplification. *Lane 1:* PTP-1a nGSP & nAP. *Lane 2:* Single Primer Control –nAP. *Lane 3:* Single Primer Control – PTP-1a nGSP. *Cycle: 5'SMART RACE NESTED PCR (Methods Section 2.4.9)* **D.** Sequence alignment of the three 5'RACE products amplified by the nested PCR. The ATG start codon is highlighted green. The untranslated sequence is highlighted in purple. This does contain an ATG (highlighted white) but this is directly proceeded by the stop codon TAA.

6.7.3 3'RACE Characterization of *ptp-1* transcripts

The 3'RACE characterization did not discriminate between the three PTP-1 transcripts. The largest and most abundant RACE product from the first round reaction was ~2.2Kb and ~400bp larger than the predicted amplification size *(figure 6.16A)*. This band and a secondary band, very close in size, were gel extracted together and TA cloned. Upon the restriction mapping of clones with EcoRI three different sized RACE products were identified and then sequenced. The size difference between the *ptp*-1 3'RACE products and the predicted amplification is attributable to size of the 3'UTR of *ptp*-1. The size variation between the three RACE products identified upon cloning is accounted for by the alternative *trans*-splicing of the poly(A) tail to the 3'end. Sequencing of the PTP-1 2-hybrid clone confirmed the sequence of the 3'UTR established by the RACE products and identified a fourth alternative poly-adenylation site, localized close to the sites defined by the 3' RACE products 2 and 3 *(figure 6.13B)*.

A.

3'RACE PCR: 2100bp
Internal Control: 368bp
5'RACE GSP
3'RACE GSP

B.

```
3'RACE Product1  GATCAAAGGGCAATGCTTATTCAGACACCGGGCCAATACACATTTGTGTGTGAATCAATT
3'RACE Product2  GATCAAAGGGCAATGCTTATTCAGACACCGGGCCAATACACATTTGTGTGTGAATCAATT
3'RACE Product3  GATCAAAGGGCAATGCTTATTCAGACACCGGGCCAATACACATTTGTGTGTGAATCAATT
                 ************************************************************

3'RACE Product1  CTACGTGCCTACCACGACGGAACCATCAAACCGCTTGCTGAATATTCGAAACGTTGAGCT
3'RACE Product2  CTACGTGCCTACCACGACGGAACCATCAAACCGCTTGCTGAATATTCGAAACGTTGAGCT
3'RACE Product3  CTACGTGCCTACCACGACGGAACCATCAAACCGCTTGCTGAATATTCGAAACGTTGAGCT
                 ************************************************************

3'RACE Product1  TTTATCCAAAAAATACATATATCGTTTTTGTTTCTTCAAATTCTTCTTCCCATCGAACTC
3'RACE Product2  TTTATCCAAAAAATACATATATCGTTTTTGTTTCTTCAAATTCTTCTTCCCATCGAACTC
3'RACE Product3  TTTATCCAAAAAATACATATATCGTTTTTGTTTCTTCAAATTCTTCTTCCCATCGAACTC
                 ************************************************************

3'RACE Product1  TCATGAATCACGGATCCCGCGAGGTGCTAGCTATTTTTGCCTTTTTTCTTTCTTCTTTTT
3'RACE Product2  TCATGAATCACGGATCCCGCGAGGTGCTAGCTATTTTTGCCTTTTTTCTTTCTTCTTTTT
3'RACE Product3  TCATGAATCACGGATCCCGCGAGGTGCTAGCTATTTTTGCCTTTTTTCTTTCTTCTTTTT
                 ************************************************************

3'RACE Product1  TTTATTGCATAGTTAATTAGCTATTGTTTTCCTACACAAAACTAGTCAATGTTTTAAGTA
3'RACE Product2  TTTATTGCATAGTTAATTAGCTATTGTTTTCCTACACAAAACTAGTCAATGTTTTAAGTA
3'RACE Product3  TTTATTGCATAGTTAATTAGCTATTGTTTTCCTACACAAAACTAGTCAATGTTTTAAGTA
                 ************************************************************

3'RACE Product1  ATTAAATATCATCATTTAATATTTCAACAAAAAATCTATCTCAATGGGTCACCCGATGTG
3'RACE Product2  ATTAAATATCATCATTTAATATTTCAACAAAAAATCTATCTCAATGGGTCACCCGATGTG
3'RACE Product3  ATTAAATATCAT------------------------------------------------
                 ***********

3'RACE Product1  ATTTTTCGTACCAATTGTTCCCCATCACTACGTCATAATTGTACCACCCCCCCCCATCTTT
3'RACE Product2  ATTTTTCGTACCAATTGTTCCCCATCACTACGTCATAATTGTACCACCCCCCCCCATCTTT
3'RACE Product3  ------------------------------------------------------------

3'RACE Product1  CATGTACGAAAAATCGCCCAAACTTGTATGTAAAAAAAACAAAAAAGTCCTCTCAAATCA
3'RACE Product2  CATGTACGAAAAATCGCCCAAACTTGTATGTAAAAAAAACAAAAAAGTCCTCTCAAATCA
3'RACE Product3  ------------------------------------------------------------

3'RACE Product1  TCACAAACTTTCCTT↓CTTTTTCATATAAAAATGTTACAGTCTGTG
3'RACE Product2  T---------------------------------------------
3'RACE Product3  ---------------------------------------------
```

Figure 6.17 The 3' RACE characterization of *ptp*-1. A. The 3'RACE amplification of *ptp*-1. *Lane 1:* PTP-1a 3'RACE & AP. *Lane 2:* PTP-1a 3'RACE GSP & PTP-1a 5'RACE GSP. *Lane 3:* Single Primer Control - PTP-1a 5'RACE GSP. *Lane 3:* Single Primer Control – PTP-1a 3'RACE GSP. *Cycle: 3'SMART RACE PCR (See Methods Section 2.4.9)* **B.** Alignment of sequences of the three largest 3'RACE products amplified. The arrow indicates the poly adenylation *trans*-splice site of the PTP-1 2-hybrid clone. Sequence obtained from the individual RACE products is in black and the ORF stop codon assigned by the ptp-1a, b and c gene model is in red.

244

6.7.4 Making the *ptp-1a* gene reporter construct

The gene C48D5.3 is 8.5Kb away from *ptp*-1a and the nearest upstream gene. Alignment to the *C.briggsae* super-contig cb25.fpc2234 identified the intergenic region between *ptp*-1a and C48D5.3 contains blocks of sequence homology that are coding, strong and weak *(figure 6.18A)*. To generate a construct containing the entire intergenic region it was decided to perform the amplification in two halves and utilise an internal Sal-1 restriction site to clone the fragments sequentially into the expression vector pHAB.mRFP-1.

The construct *ptp*-1a::mRFP-1 represents the first sequential step in the cloning process. It was generated by the PCR amplification of a genomic region, encompassing exons 1-3 and 2.4kb of 5' un-translated sequence *(figure 6.18B)*. The fluorescent reporter protein mRFP-1 was fused in-frame to the third exon of *ptp*-1a. The construct *ptp*-1a::mRFP-1 was utilised as a gene reporter to preliminarily assess the expression pattern of the *ptp*-1 gene.
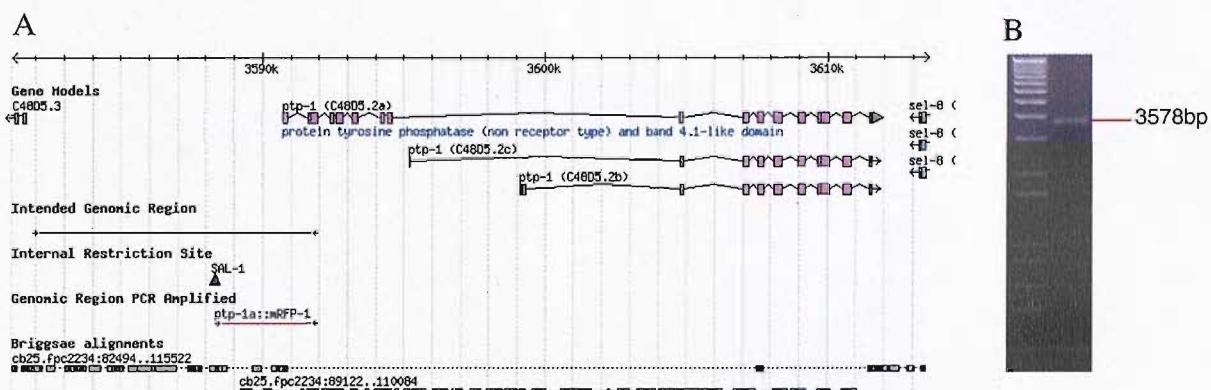
Figure 6.18 The *ptp*-1(a,b and c) gene models and alignment to the *C.brigssae* super-contig cb25.fpc2234. A) The intergenic region between *ptp*-1 and the nearest upstream gene, C48D5.3, is shown. The blue line indicates the entire genomic region intended for production of the ptp-1 gene reporter construct. The blue triangle indicates the position of the internal Sal-1 site intended for the sequential cloning of the two ptp-1 genomic fragments into the vector pHAB.mRFP-1. The red line indicates the first *ptp*-1 genomic fragment cloned and used to generate the construct *ptp*-1a::mRFP-1. The colour coding of the *C.briggsae* alignment is the same as for *Figure 6.3*. **B)** The genomic PCR of the *ptp-1a* promoter region.

6.7.5 Analysis of the *ptp*-1a reporter construct expression

The construct *ptp*-1a::mRFP-1 was injected into the transgenic line *mgl*-1::gfp; *pha*-1(e1213). The expression of the reporter construct *ptp*-1a::mRFP-1 was identified in the adult hermaphrodite's head, midbody and tail. In the head mRFP-1 fluorescence was observed in neuronal cells of the nerve ring and the pharyngeal nervous system. Neurons of the nerve ring and the pharyngeal nervous system were identified as co-expressing *ptp*-1a and *mgl*-1 *(figure 6.19 and 6.20 respectively)*. In the de-sheathed pharynx *mgl*-1::GFP and *ptp*-1a::mRFP-1 were identified as being co-expressed in one neuron of the pharyngeal nervous system. The cell body of the neuron was located in the terminal bulb of the pharynx and displays a single process extending towards the corpus *(figure 6.19 and 6.20)*. In the extrapharyngeal nerve ring at least four neurons co-express the two genes together. In addition there are neurons within the nerve ring that express the gene *ptp*-1a and not *mgl*-1.
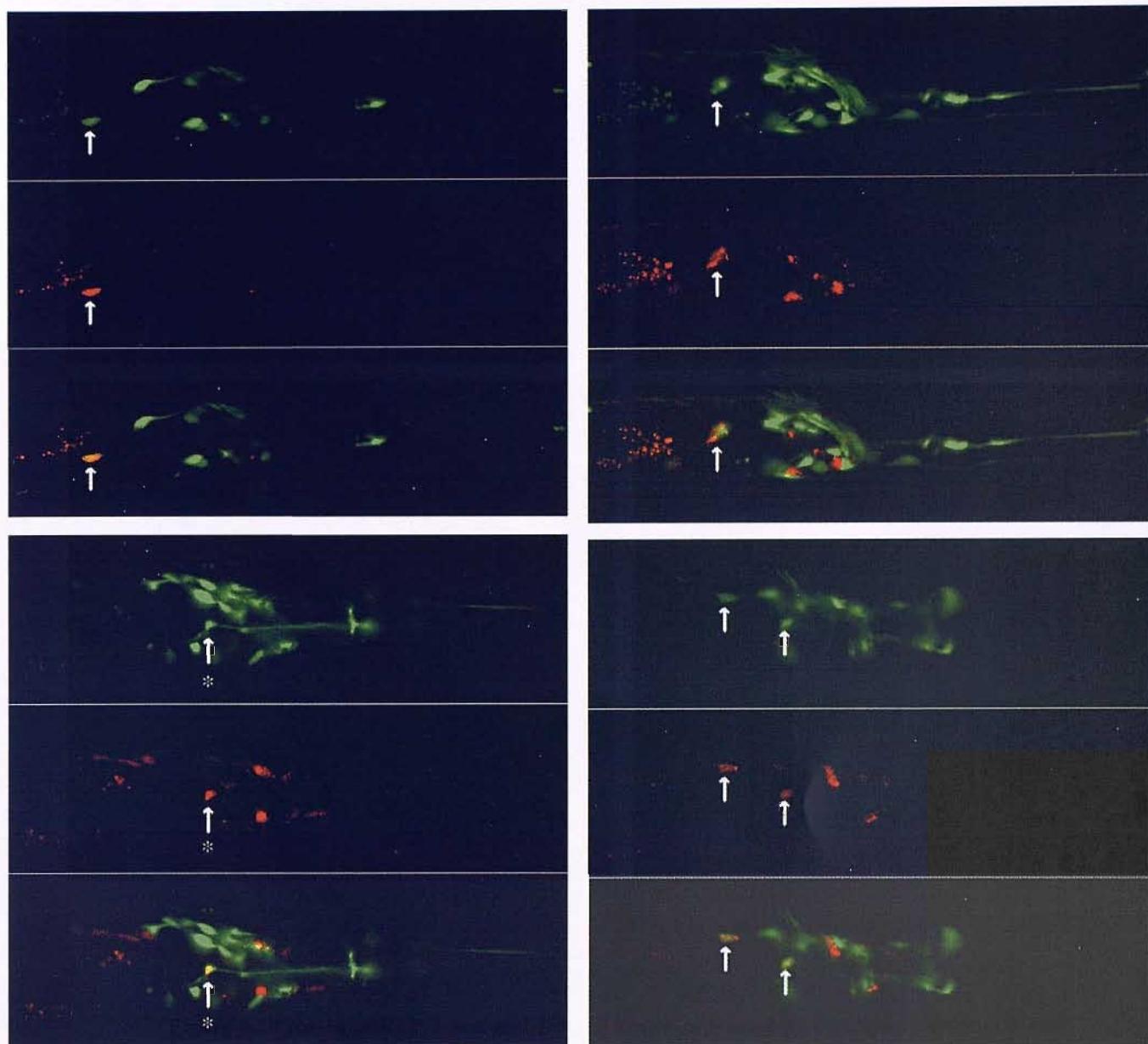
Figure 6.19 *ptp*-1a::mRFP-1 and *mgl*-1::GFP are co-expressed by neurons belonging to the **nerve ring and pharyngeal nervous system.** The expression of *ptp*-1a was also observed in neurons of the nerve ring independently to *mgl*-1. Representative sections of a Z-stack performed on the head of an adult hermaphrodite are shown. White arrows indicate cellular co-expression. White asterisk indicates a pharyngeal neuron *(see figure 6.21 for further detail) Confocal Settings: Objective: 40x Oil DIC. GFP excitation 488nm, band pass filter 505-530nm. mRFP-1 excitation 543nm, band pass filter 560-615nm. Z-Slice: 0.98μM.*
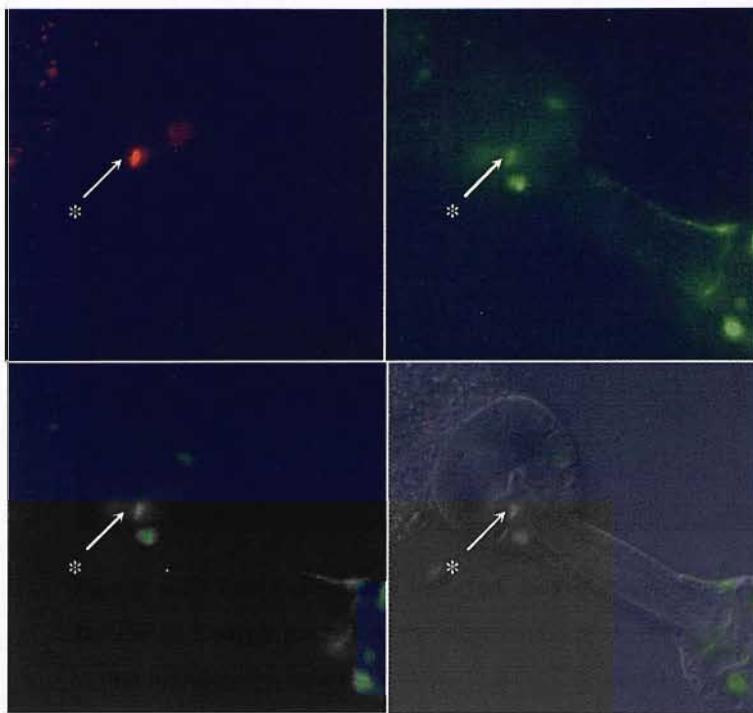
Figure 6.20 The desheathed pharynx identifies co-expression of *ptp*-1a::mRFP-1 and *mgl*-1::GFP in a single neuron of the pharyngeal nervous system. The neuron indicated corresponds to that indicated by an asterix in *figure 6.20*. The anatomical localization of the cell body suggests this neuron is either: M1, M2, I4 or I6. M1 and M2 are motorneurons. I4 and I6 are interneurons. Objective: x60, oil.

In the tail *ptp*-1a::mRFP-1 expression was not identified in the same four tail neurons expressing *mgl*-1::GFP *(figure 6.21)*. In the mid-body *ptp*-1a::mRFP-1 was identified in the canal neuron, CAN. Anatomically the cell body of CAN is located at the mid-body, lateral sides, within close proximity to the vulva *(figure 6.2B)*. The morphology of CAN is distinct with apical and distal processes extending from the cell body, along the excretory canal and towards the head and tail. The expression of *ptp*-1a::mRFP-1 in CAN was identifiable as early as the developmental stage L1 *(figure 6.22A)*.
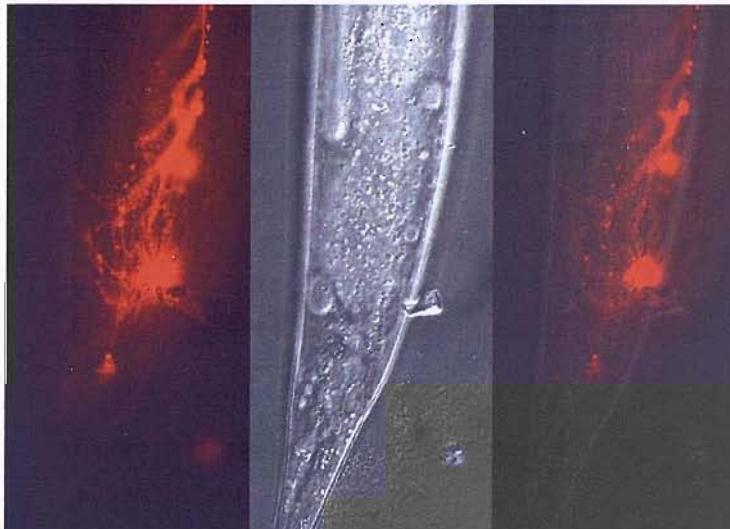
Figure 6.21 *ptp*-1a::mRFP-1 is expressed in the tail of the adult hermaphrodite. *ptp*-1a::mRFP-1 was not co-expressed in the same tail cells as *mgl*-1::GFP. Objective: x60, oil. Exposure time: 6 seconds.
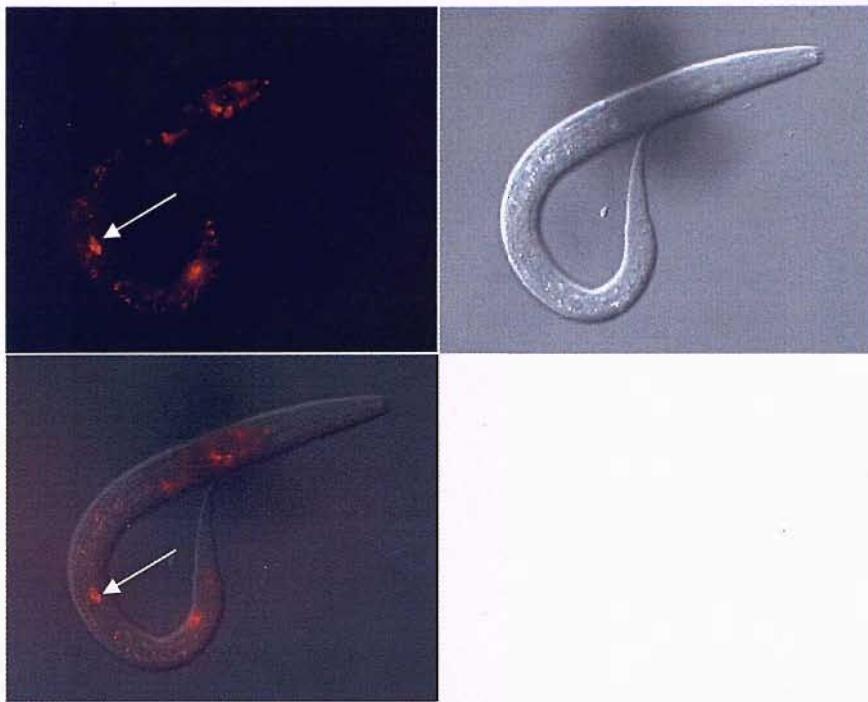
Figure 6.22 The expression of *ptp*-1a::mRFP-1 in the CAN neuron. A) L1 hermaphrodite. **B)** Adult hermaphrodite. Objective: x60 oil. Exposure time: 8 seconds. In *figure A* the white arrow indicates the position of CAN.

6.7.6 *ptp-1a::mRFP-1* expression in relation to *glr-1*

The *ptp-1* 2-hybrid clone, initially identified as interacting with the MGL-1 receptor C-terminal, was subsequently shown to interact with the C-terminal of the GLR-1 receptor in yeast. The published expression pattern of *glr-1* describes it in both interneurons and motorneurons of the nerve ring and interneurons of the tail [Brockie, P.J. et al 2001]. The expression of the gene reporter *ptp-1a::mRFP-1* was associated with neurons of the nerve ring also, but a more detailed analysis is required to establish if the cells expressing *ptp-1a* are the same as those expressing *glr-1*.
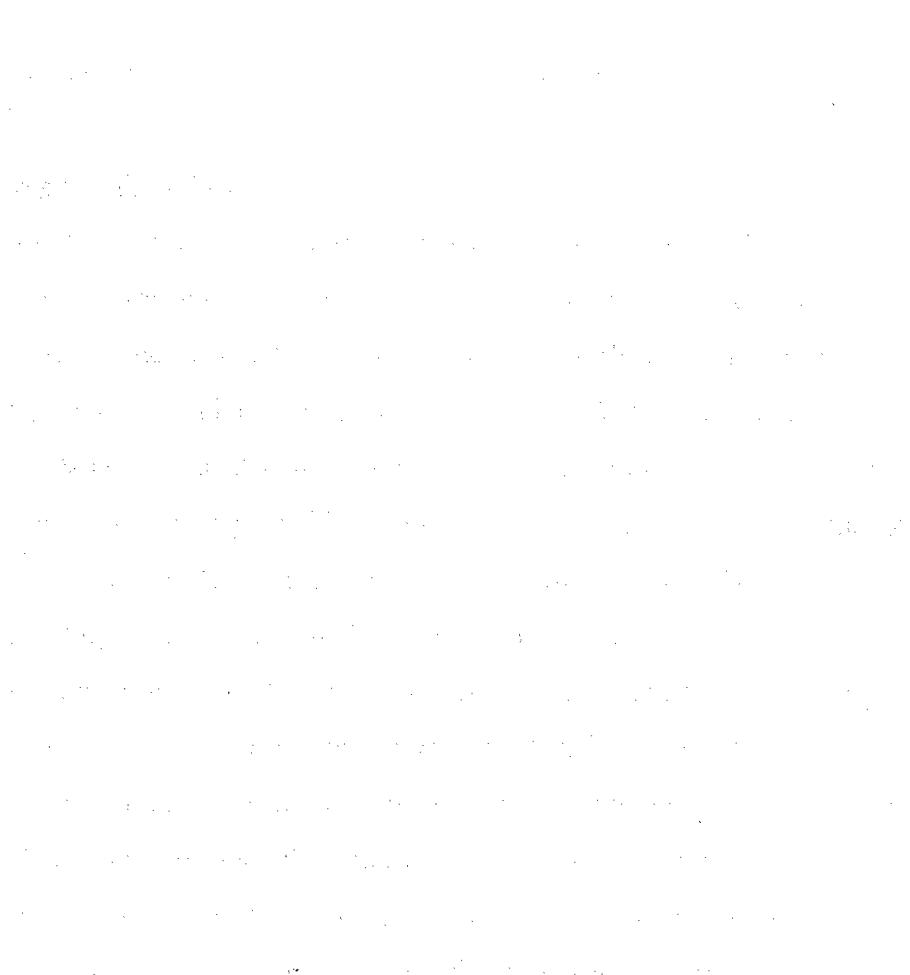
6.8   Discussion

The three proteins, MPZ-1a, PTP-1 and C01F6.6a were identified as interacting strongly with the MGL-1 receptor C-terminal in yeast, a heterologous system. As a first step to determining if the interaction is genuine it was necessary to confirm the 2-hybrid identified proteins are transcribed in the same cells as MGL-1 *in-vivo* at the level of the whole organism. This was done by taking transgenic *C.elegans* expressing *mgl*-1::GFP and transforming them with promoter::reporter fusion constructs that report on the cellular expression of the genes encoding each of the three proteins of interest. The technical issues encountered with the approach will be discussed together with the biological implications of them. Expression patterns were produced for the genes encoding each of the three proteins of interest. Analysis of the cellular expression pattern has provided insight into the biological function of the proteins and specifically the physiological significance of the 2-hybird identified interaction with the MGL-1 C-terminal.

6.8.1   *mpz-1(C52A11.4a)*

In total three gene reporter constructs were generated to characterize the cellular expression pattern of *mpz*-1(C52A11.4a), the gene encoding the protein MPZ-1a. Each successive gene reporter was a modification made to improve the reliability of the cellular expression pattern. The first gene reporter, designated *mpz*-1(i)::DsRED2 was designed and manufactured on the basis of an imprecisely defined *in-silico* gene model. Errors in the predicted 5' boundary of the gene and the exon/intron boundaries at the 5' end meant that the genomic region amplified was not appropriate for the reliable detection of gene expression.

The inaccuracies of the *in-silico* gene model were brought to light by the EST yk1004e08, which identified *mpz*-1 mRNA is *trans*-spliced to SL-1. This was confirmed independently and in the process two additional alternative SL-1 *trans*-spliced variants were characterized. As discussed previously (*see Chapter 3*) the precise function of SL-1 *trans*-splicing is poorly understood. It has been suggested that it is a mechanism for the post-transcriptional control of gene expression (Davis, R.E 1996) and infers *mpz*-1(C52A11.4a) cellular expression is regulated by different control elements. The alternative *trans*-splicing to SL-1

may reflect the processing of the *mpz-1* gene by the different cellular machinery of the different cell types it is expressed in.

The gene reporters *mpz-1*(ii)::DsRED2 and *mpz-1*(iii)::mRFP-1 both contain the entire 5' intergenic region of *mpz-1*(C52A11.4a), now more clearly defined by the SL-1 characterization of the *mpz-1* 5' boundary, but produced very different cellular expressions. The expression reported by *mpz-1*(ii)::DsRED2 was punctate and appeared as discrete aggregates towards the level of the hypodermis and cuticle, along the length of the worm. The puncta were not associated with cells expressing *mgl-1*::GFP. This was in marked contrast to the expression reported by *mpz-1*(iii)::mRFP-1. In transgenic animals carrying this reporter the fluorescence appeared more diffuse, as it was observed in both the processes and cell body of neurons and throughout individual muscle cells of the body wall and vulva. In addition the expression pattern was more extensive because red fluorescence was now identifiable in cells expressing *mgl-1*::GFP.

In generating this cellular expression pattern an increased amount of the *mpz-1*(C52A11.4a) gene was used and DsRED2 was substituted for the monomeric red fluorescent protein mRFP-1. The different cellular expression could be attributed to either of these changes or a combination of them both. Fusion of the marker further downstream in the *mpz-1* gene encapsulated two additional introns and the translation start codon defined by the EST yk1004e08 and two of the identified SL-1 *trans*-splice variants. This inclusion may be important for the gene's correct expression in an increased number of cells, since 5' intronic sequences are known to serve as regulatory sites for the cell-specific control of gene expression and gene expression levels. The cell type-specific expression of *mpz-1* and the intricate processing of the 5' end by *trans*-splicing to SL-1, generating alternative transcripts, would imply that the expression of *mpz-1* is regulated by complex promoter elements. The punctate appearance of the fluorescence in transgenic worms expressing DsRED2 could be an artefact of its aggregation. DsRED2 is known to form an obligate dimer during maturation and using a monomeric, non-dimerizing mutant (Campbell. RE et al 2002) instead produced a more diffuse fluorescence.

An abstract (Xiao, H et al 2004) has described *mpz-1* as being expressed in 40 neurons of the nerve ring, 6 neurons in the body, 8 neurons in the tail, 2 pharyngeal motorneurons, body wall and vulval muscles. The expression reported

by *mpz*-1(iii)::mRFP-1 in the nervous system was not as abundant. The abstract does not report how this expression was produced but the region of *mpz*-1 used could be a major source of variation and the reason for the observed differences. In generating the *mpz*-1 gene fusion reporters, *mpz*-1(ii) and *mpz*-1(iii) it was apparent that by increasing the amount of *mpz*-1 genomic coding sequence used, a wider expression pattern could be obtained. Hence, utilising the entire *mpz*-1 gene, including all 5' and 3' intergenic sequence, would incorporate all potential regulatory elements and based upon the co-expression of *mgl*-1 with *mpz*-1, reported by *mpz*-1(iii)::mRFP-1, this would be a purposeful process. Furthermore this would yield a protein reporter that would allow the subcellular localization of MPZ-1 to be determined, in relation to MGL-1.

6.8.2 Insights to *mpz-1* function

The *mpz*-1 gene was identified in two different cell types, neurons and muscle cells. Furthermore the expression was cell type-specific, *mpz*-1 was not expressed in the pharyngeal muscle but was expressed in the body wall and vulval muscle. In the nervous system *mpz*-1 is more widespread and expressed in the pharyngeal nervous system, the nerve ring, midbody neurons and tail neurons. The mouse homologue of *mpz*-1, m*pdz*, shares this characteristic of having a widespread expression pattern. It is expressed in different cell types of the brain (neurons and epithelia) and it is expressed in different tissues throughout the periphery and central nervous system, including heart, brain, placenta, kidney, liver, pancreas and skeletal muscle. The complexity of m*pdz*'s expression is reflected by the array of different signalling molecules it interacts with via the thirteen PDZ domains it contains. As MPZ-1 has a similar type of expression and protein architecture to m*pdz* it is conceivable it too binds different proteins and organises signalling within different cell types. The function of both muscle cells and neurons relies upon the correct localization of receptors and the organization of the signalling cascades they couple to. In addition *mpz*-1 was identified in cells not expressing *mgl*-1::GFP suggesting it is multi-functional. It has been reported to interact biochemically with both the 5-HT receptor SER-1 (Xiao, H and Komuniecki, R.W. et al 2004) and the PTEN homologue DAF-18 (Lin, L and Komunieki, R.W. et al 2004).

6.8.3 Insights into the physiological significance of the co-expression between *mpz*-1 and *mgl*-1

The genes *mpz*-1 and *mgl*-1 were co-expressed in neurons of the pharyngeal nervous system and in neurons of the nerve ring. However, co-expression was not identified in any neurons of the tail. The co-expression of *mpz*-1 and *mgl*-1 suggests MPZ-1 has a role in MGL-1 function and this is mediated by an interaction with the MGL-1 intracellular C-terminal. In the pharyngeal nervous system *mgl*-1 and *mpz*-1 were co-expressed in 12 neurons and this included the motorneurons NSM and M4 (*see Section 4.3* for a discussion of these neurons and their function). Taken together this suggests the scaffolding of MGL-1 by MPZ-1 has a role in pharyngeal function and feeding behaviour. Also co-expression in the nerve ring motorneuron RMD is proposed, this neuron is involved in controlling head movements of the worm and foraging behaviour,

As described previously (*see Section 5.18.3*) the 10th PDZ domain of MPZ-1 interacts with the 5-HT$_{2C}$ receptor homologue, SER-1, in yeast and biochemically. SER-1, as like *mgl*-1, is co-expressed with *mpz*-1 in NSM and M4. MPZ-1 may be responsible for organising both receptors function either independently or it provides a point of cross-talk between serotonergic and glutamatergic intracellular signalling pathways. The multi-PDZ domain structure of MPZ-1 makes it an ideal candidate for interacting with multiple binding partners simultaneously, enabling it to assemble larger protein complexes. The multi-PDZ *Drosophila* protein InaD exemplifies such a role, since it is central to the co-localization of signalling components of the light transduction cascade (Tsunoda, S et al 1997).

6.8.4 *tag-60(C01F6.6a)*

The 5' RACE characterization of the *tag*-60(C01F6.6a) transcript confirmed the intron/exon boundaries of the coding region fused to DsRED2 and this corresponded to exon 1 and 2 at the 5' end. A single RACE product was identified as being *trans*-spliced to the spliced leader sequence SL-1 and this corresponded to the longest *tag*-60(C01F6.6a) RACE product identified. The ATG start codon of the *tag*-60(C01F6.6a) open reading frame assigned by the gene model was confirmed and the ATG of DsRED2 was fused in frame with this.

The 3' RACE characterization identified that the C-terminal of C01F6.6a is alternatively spliced to generate three different splice variants, two of which have previously not been identified. A fourth variant, containing a silent nucleotide substitution is likely to have arisen by PCR error and is unlikely to be biologically relevant. The alternative splicing did not alter the PDZ domain architecture of the C01F6.6a protein but the alternative splicing of PDZ-domain flanking sequence has been identified as a source of functional diversity (Sierralta, J and Mendoza C, 2004). Insertions can confer: alternative cellular localization; intermolecular interactions that allow the alternative regulation of protein binding specificities; alternative sites of post-translational modification that regulate the proteins function. The C01F6.6a 3'RACE products differ in the use of alternative exons in two different sites. The first of these sites corresponds to a region localized between the two PDZ domains and the second is localized in the C-terminal distal of the second PDZ domain. They define the presence of 24 and 68 amino acid inserts respectively and may confer alternative functions to C01F6.6a. The 2-hybrid clone identified as interacting with the MGL-1 C-terminal corresponds to the identified splice variant containing the amino acid insertion at the C-terminal site but not at the site localized between the two PDZ domains. The cellular expression of C01F6.6a implies that it may interact with different proteins as part of its function.

The gene *tag*-60 was specifically expressed in the pharyngeal muscle subtypes pm3 and pm4, the intestine and neurons of both the tail and the pharyngeal nervous system. The use of approximately 3.3kb of the 9kb 5' intergenic region incorporated conserved promoter elements identified by alignments to the *C.briggsae* genome and this was sufficient to drive the expression of the fluorescent reporter DsRED2 in several different cell types. Furthermore it correlates with a subsequent description of the expression of *tag*-60 (McKay, SJ et al 2004). A gene fusion reporter containing a similar amount of the intergenic region identified *tag*-60 as being expressed in the pharynx, intestine and cells in the tail.

The interaction of *tag*-60 with the MGL-1 C-terminal was not recapitulated at the level of the whole organism, since *tag*-60 was not identified as being co-expressed in the same cells as *mgl*-1. A possible explanation for this is the inherent promiscuity of PDZ domains and PDZ recognition motifs. This is

particularly well demonstrated by the C-terminal of the neuroligin (Nlg) subtypes, Nlg1 and Nlg2. Neuroligin is a postsynaptic adhesion molecule that is postulated to have a role in aligning the postsynaptic membrane with the presynaptic active zone. The intracellular C-terminal of neuroligin contains a Type-I PDZ binding motif and is devoid of all other interaction motifs. A LexA yeast-2-hybrid library screened with the C-terminal identified 13 PDZ domain-containing proteins that interacted with the Nlgs and each of the 13 identified interaction partners also interacted with a large number of alternative protein targets as well (Meyer, G et al 2004). At least 32 other target proteins were identified as having a PDZ-domain binding specificity that overlapped with the Nlgs. This highlights that a single PDZ domain does not bind to a single target in-fact PDZ domains can be dual specific. That is a single PDZ domain can bind both a Class I and a Class II PDZ-binding motif and this further adds to the promiscuity and complexity of PDZ domain interactions. This was found to be the case for the 2nd PDZ domain of syntenin, a synaptic adaptor protein with two PDZ domains and a role in coupling transmembrane proteoglycans to cytoskeletal proteins (Grootjans, J.J et al 1997).

Specificity can instead be conferred by the differential cellular expression of PDZ domains and their alternative target ligands and post-translational modifications, such as phosphorylation. The first of these is demonstrated by the interaction of the MGL-1 C-terminal with C01F6.6a. The interaction between MGL-1 and C01F6.6a observed in yeast does not occur *in vivo* because of the spatial segregation of the two proteins. However, it should be noted that target proteins with overlapping specificities to Nlgs are expressed at the same types of synapse and are simultaneously present.

In the case of C01F6.6a the advantages of utilising the transgenics of *C.elegans* have been demonstrated. The biological relevance of an *in-vitro* identified protein-protein interaction can be immediately assessed at the systems level and categorized as a low priority for further investigation. It also highlights the caveats of the 2-hybrid system and the nature of using it as a tool to identify protein-protein interactions. Although it is suitable for isolating protein interacting partners it is in discriminate and has no preference for whether these interactions can occur within the context of a biological system. This disadvantage is further enhanced by protein-protein interactions that are promiscuous in nature, such as those mediated by PDZ domains and PDZ binding motifs.

6.8.5 *ptp*-1a

The gene *ptp*-1 encodes three splice variants, PTP-1a, b and c. The 5' RACE characterization of the longest transcript *PTP*-1a confirmed the intron-exon boundaries of the coding region fused to the reporter mRFP-1. This corresponded to the first three exons of *ptp*-1a. The longest 5' RACE product identified confirmed the ATG start codon of the PTP-1a open reading frame assigned by the gene model. The ATG of mRFP-1 was fused to PTP-1a in frame with the confirmed start codon, to generate a gene reporter fusion for the longest transcript. The 3'RACE characterization identified three different sized products, which were accounted for by the use of alternative poly adenylation sites.

The gene reporter fusion identified PTP-1a is expressed in the pharyngeal nervous system, the nerve ring, the tail and the midbody neuron CAN. Using 2.4kb of the 9kb 5' intergenic region generated this expression pattern. The proportion of the intergenic region used was governed by the presence of a convenient *sal*I restriction site and the construct *ptp*-1::mRFP-1 represents the first stage of the sequential cloning of the entire 5' intergenic region of *ptp*-1::mRFP-1. This then reports the first step in the process of generating a *ptp-1* reporter construct.

The canal neuron CAN is essential to the worms' survival but it has not been identified as making synapses and its function is unknown. Gene expression suggests CAN is capable of signalling, since it expresses markers of monoamine transmission (Duerr J.S et al 1999) and neuropeptide like proteins (Nathoo A et al 2001) as well as receptors that include a splice variant of the tyramine receptor SER-2 (Tsalik, E.L. et al 2003). The expression of *ptp*-1 was identified in the tail but it was not associated with neurons expressing the *mgl*-1::GFP reporter. A sub-population of the neurons in the nerve ring expressing *ptp*-1 were identified as co-expressing *mgl*-1::GFP. In the pharyngeal nervous system *ptp*-1 was expressed by a single neuron that also expressed *mgl*-1::GFP. This suggests PTP-1a is involved in the organization of MGL-1 function within both the pharyngeal and extra-pharyngeal nervous systems.

The *ptp*-1 gene is complex and the use of alternative ATG start codons is predicted to produce alternative splice variants that encode proteins with different domain architectures and therefore different functions. For this reason the *ptp*-1

promoter is likely to be intricate, with alternative promoters regulating the expression of the different splice variants. It cannot be said for certain that the gene reporter fusion *ptp*-1a::mRFP-1 is reporting on all three of the different *ptp*-1 splice variants, particularly since the structural organization of the *ptp*-1b and *ptp*-1c coding sequence suggests downstream promoter elements are likely to be involved in their expression (*See figure 6.24*). Furthermore, use of the entire intergenic region may yield a more comprehensive expression pattern. Since alignment to the *C.briggsae* genome identified conserved elements within the intergenic sequence that were not encompassed by the region used to generate *ptp*-1::mRFP-1. However, *ptp*-1::mRFP-1 has provided a preliminary assessment of where *ptp*-1 is expressed and its identification in the same cells as *mgl*-1 suggests larger gene reporter fusions, that report on all isoforms, may identify a more widespread co-expression pattern.
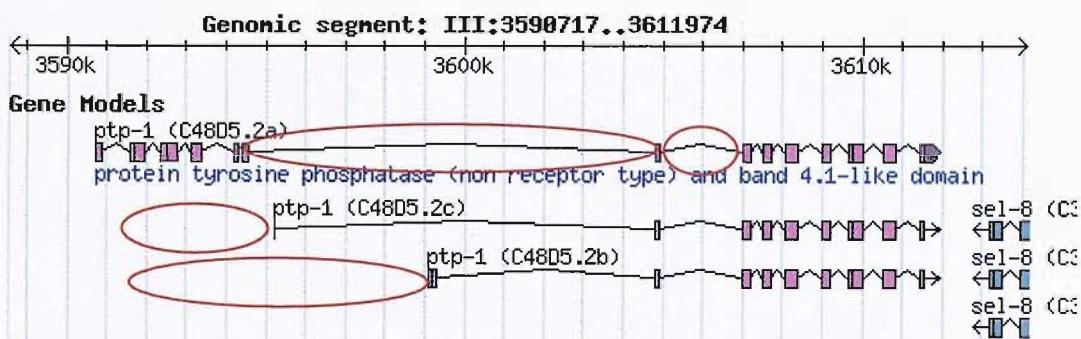
Figure 6.24 Potential promoter elements of the *ptp-1* gene. Large introns of the *ptp*-1 gene, such as those circled in red are candidates for harbouring promoter elements required for the cellular expression of *ptp*-1 isoforms. Inclusion of such regions in gene reporter fusions may be necessary to obtain a comprehensive expression pattern that encompasses all of the *ptp*-1 isoforms.

A BLAST search performed with the protein sequence of PTP-1a identified the protein PTP-MEG as the closest mammalian homologue. PTP-MEG is a protein tyrosine phosphatase that was initially cloned using cDNA from a human megakaryoblastic cell line [Gu, M. et al 1991]. PTP-MEG belongs to the PTPase subfamily and it has the same domain organisation as PTP-1a, from the N to C-terminal it contains a Band 4.1 domain, a PDZ domain and a PTPase domain. The N-terminus of PTP-1a is 42% identical to PTP-MEG (amino acids 28→362) and the C-terminus is 57% identical (amino acids 610→1026). PTP-MEG has been

shown to interact with the C-terminal of the ionotropic receptor subunits GluRδ2 and NR2A in a PDZ dependant manner and regulates the receptor activity through tyrosine phosphorylation [Hironaka, K et al 2000]. This suggests the interaction identified between PTP-1A and the C-terminal of the *C. elegans* ionotropic GLR-1 subunit in yeast may be an evolutionarily conserved feature of iGluR scaffolding and that PTP-1A may be a component of the GLR-1 signalling complex in *C. elegans*. As highlighted *(see section 6.7.6)* the expression of *glr-1* has been published. The expression of the gene reporter construct *ptp-1a::mRFP-1* in a *glr-1::GFP* integrated transgenic line would enable the interaction identified in yeast to be assessed *in vivo* and further insights to the role performed by the PTPase subfamily in ionotropic receptor scaffolding.

6.9   Summary

The expression pattern of three proteins, identified as interacting with the MGL-1 C-terminal were defined by gene reporter fusions. Two of the genes, *ptp-1*a and *mpz*-1a, were identified as being co-expressed by the same neurons expressing *mgl*-1. However, a population of neurons expressing *mgl*-1 still remains for which co-expression was not identified. This may reflect the expression pattern produced for MPZ-1a and PTP-1a is not complete and in both cases the production of a full-length gene reporter fusion is required. This is especially when there is evidence for a gene undergoing alternative processing and more than one isoform of the protein is being expressed. Despite this, the identification of *ptp*-1a and *mpz*-1(C52A11.4a) in the same neurons as *mgl*-1 provides a basis for producing protein reporter fusions to define the subcellular localization of each protein in relation to the MGL-1 receptor. A subcellular co-localization between the MGL-1 receptor and the 2-hybrid identified protein would provide conclusive confirmation that the interaction is biologically significant.

CHAPTER 7

The functional significance of *mpz-1* co-expression with *mgl-1* in the pharyngeal nervous system

7.1 Introduction

The co-ordination of synaptic signalling depends on the correct organization of the cellular machinery that conducts and terminates signalling. For instance neurons must direct neurotransmitter containing vesicles to their site of release and co-ordinate their fusion with the presynaptic membrane. Neurotransmitter transporters must be localized correctly at the membrane, in relation to the site of neurotransmitter release, to facilitate the re-uptake and recycling of the released neurotransmitter [Gonzalez, M.I. and Robinson, M.B. 2004] and membrane bound receptors and the intracellular signalling molecules they couple to must be targeted and anchored at the correct subcellular compartment of the neuron in relation to the neurotransmitter release site [Garner, C.C. et al 2000]. As previously described scaffolding proteins perform an important role in organizing this cellular machinery and signalling complexes *(see Introduction section 1.16)*. Therefore interfering with the normal function of a scaffolding protein has the potential to disrupt neurotransmission and impair the behaviours it underlies. As discussed in Chapter 6 *mpz-1* is co-expressed with *mgl-1* in neurons of the pharyngeal nervous system and therefore, MPZ-1 could be involved in scaffolding MGL-1 signalling in this circuit to modulate pumping behaviour.

As discussed previously glutamatergic receptor mutants combined with simple behavioural assays have provided insight into both the role of these proteins and glutamate in controlling *C. elegans* behaviour *(see section 1.28)*. In Chapter 4 the pharyngeal preparation was established as a bioassay for studying *mgl-1* receptor function and potentially the role of scaffolding proteins. The expression pattern of *mgl-1* and *mpz-1* has shown that the two genes are co-expressed by neurons of the pharyngeal nervous system. The pharyngeal preparation was used to assay *mgl-1* receptor function in the available *mpz-1* mutants and to investigate if *mpz-1* is required for scaffolding *mgl-1* signalling in neurons of the pharyngeal nervous system.

7.1.2 Introduction to the available *mpz-1* mutants

Two mutant *mpz-1* strains are currently available, these are *mpz-1(tm1136)*, which was obtained from the National Bioresource Project for the Experimental Animal Nematode *C.elegans* (Japan-NBP) and VC599, which was obtained from the *C.elegans* Genetic Consortium. The strain *mpz-1(tm1136)* is homozygous for the mutant *mpz-1* allele *tm1136* and has a viable phenotype. In the existing description of *tm1136* the mutation is defined as a genomic deletion of 1124bp and a 3bp insertion. The genomic region deleted contains two exons at the 5' end of *mpz-1*.

The strain VC599 carries the allele *mpz-1(gk273)*, that is a homozygous lethal deletion and chromosome balanced by the GFP- and *dpy-10* marked inversion *mIn1*. The genetic balancer *mIn1* is an inversion of a large central portion of Chromosome II. It carries the recessive morphological marker *dpy-10*, which encodes a cuticle protein. The *dpy-10* mutation causes animals to have a short, dumpy phenotype. In addition *mIn1* carries an integrated transgene comprised of a transcriptional fusion between GFP and the *myo-2* promoter. This confers GFP fluorescence to the pharyngeal muscle. Heterozygotes, *mpz-1(gk273/+)* have wild-type morphology and weak GFP expression in the pharyngeal muscle. The *mpz-1(gk273)* mutation is defined as an 1191bp genomic deletion and 8 nucleotide insertion (ATTTTTTT). The genomic region deleted encompasses a single exon located within the 5' end of *mpz-1*.

7.2   Confirmation of the genomic break-points of *mpz-1(tm1136)*

The primers defined by the Japanese-NBP (ExtFwd and ExtRev) were used to perform a PCR amplification on genomic DNA prepared from *mpz-1(tm1136)* animals. The primers flanked the genomic region defined as containing the 1224bp deletion and 3bp insertion. The PCR conditions were first optimised using a control template, which was genomic DNA, prepared from a plate of wild-type worms. The PCR amplification yielded a single band of the expected size (the expected/predicted size is based upon the sequence analysis of the est clone yk1004e08), which was 1783bp. The negative controls, which were the anti-sense and sense primers each on their own, did not produce any bands *(figure 7.2A)*.

The established PCR protocol was used to perform a single worm PCR on a wild-type worm and on *mpz-1(tm1136)* animals in duplicate. A single band of the expected size, 1783bp was amplified from the wild-type animal and both of the mutant duplicates produced a size shift in the genomic product of 1124bp, corresponding to the size difference caused by the deletion, to yield a band 662bp in size *(figure 7.2B)*. The product amplified by the *mpz-1(tm1136)* single worm PCR was cloned and sequenced. The sequence obtained from the PCR product was aligned to the *C.elegans* genome. The deleted region documented was confirmed *(See Figure 7.4 for the location of the deletion in the mpz-1 gene)* and the 3bp insertion at the site of the deletion was also confirmed as TTG *(Figure 7.2C)*.

A

1783bp

IntFwd/IntRev

IntFwd

IntRev

B

1783bp

662bp

N2 Genomic PCR

mpz-1(tm1136) Genomic PCR

C

CAGCTTAATCAAGTTGTGCTCAACCCTCT
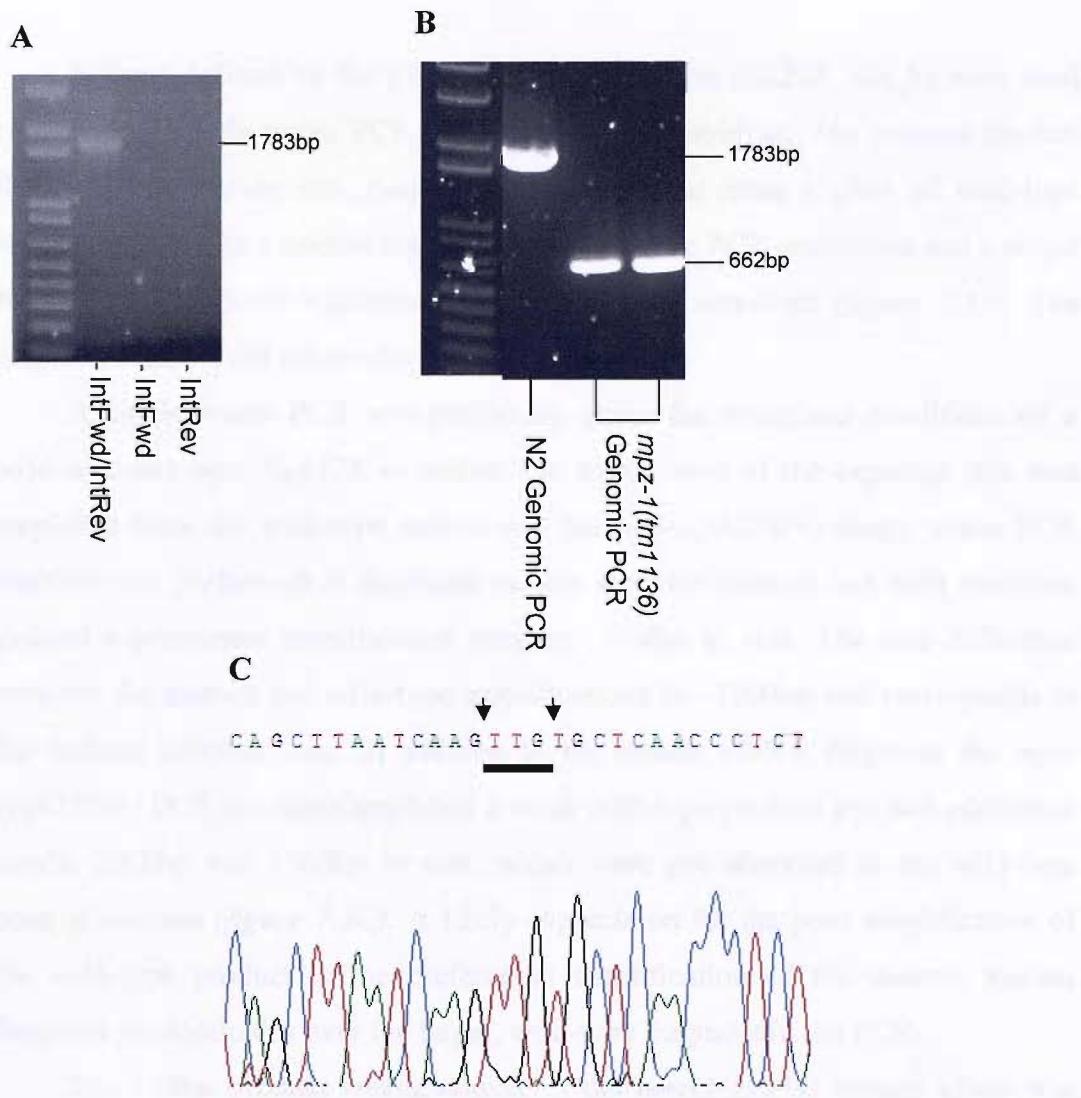
Figure 7.2 Defining the genomic break-points of *mpz-1(tm1136)* A) *mpz-1(tm1136)* primers were optimised on the wild-type type genomic preparation. A single PCR product of the correct size, 1783bp was amplified. **B)** the *mpz-1(tm1136)* single animal genomic PCR (performed in duplicate) amplified a single band, approximately 1100bp smaller than the wild-type product. **C)** A portion of the chromotogram from the sequenced *mpz-1(tm1136)* genomic fragment amplified, indicating the breakpoints (arrows) and the three neucleotide insertion, TTG (underlined).

7.2.1 Confirmation of the genomic break-points of *mpz-1(gk273)*

Primers defined by the CGC (GK273_Ext_f and GK273_Ext_b) were used to perform a single worm PCR on *mpz-1(gk273/+)* animals. The primers flanked the defined mutation site. Genomic DNA extracted from a plate of wild-type worms was used as a control template to optimise the PCR conditions and a single band of the predicted wild-type size, 1957bp, was amplified *(figure 7.3A)*. The negative controls did not produce any product.

A single worm PCR was performed under the optimised conditions on a wild-type and *mpz-1(gk273/+)* animal . A single band of the expected size was amplified from the wild-type animal and the *mpz-1(gk273/+)* single worm PCR reaction was performed in duplicate on two separate animals and both reactions yielded a prominent amplification product, 774bp in size. The size difference between the mutant and wild-type amplifications is ~1200bp and corresponds to the defined deletion size. In addition to the mutant cDNA fragment the *mpz-1(gk273/+)* PCR reactions amplified a weak wild-type product and two additional bands, 2000bp and 1500bp in size, which were not observed in the wild-type control reaction *(figure 7.3C)*. A likely explanation for the poor amplification of the wild-type product is the preferential amplification of the shorter, mutant fragment predominates over the larger, wild-type fragment in the PCR.

The 774bp product corresponding to the *mpz-1(gk273)* mutant allele was cloned and sequenced. The sequence obtained was aligned to the *C.elegans* genome and the mutation genomic break points were confirmed and so was the 8bp insertion, ATTTTTTT *(figure 7.3D)*. The region of the genomic deletion is shown in *figure 7.4*.
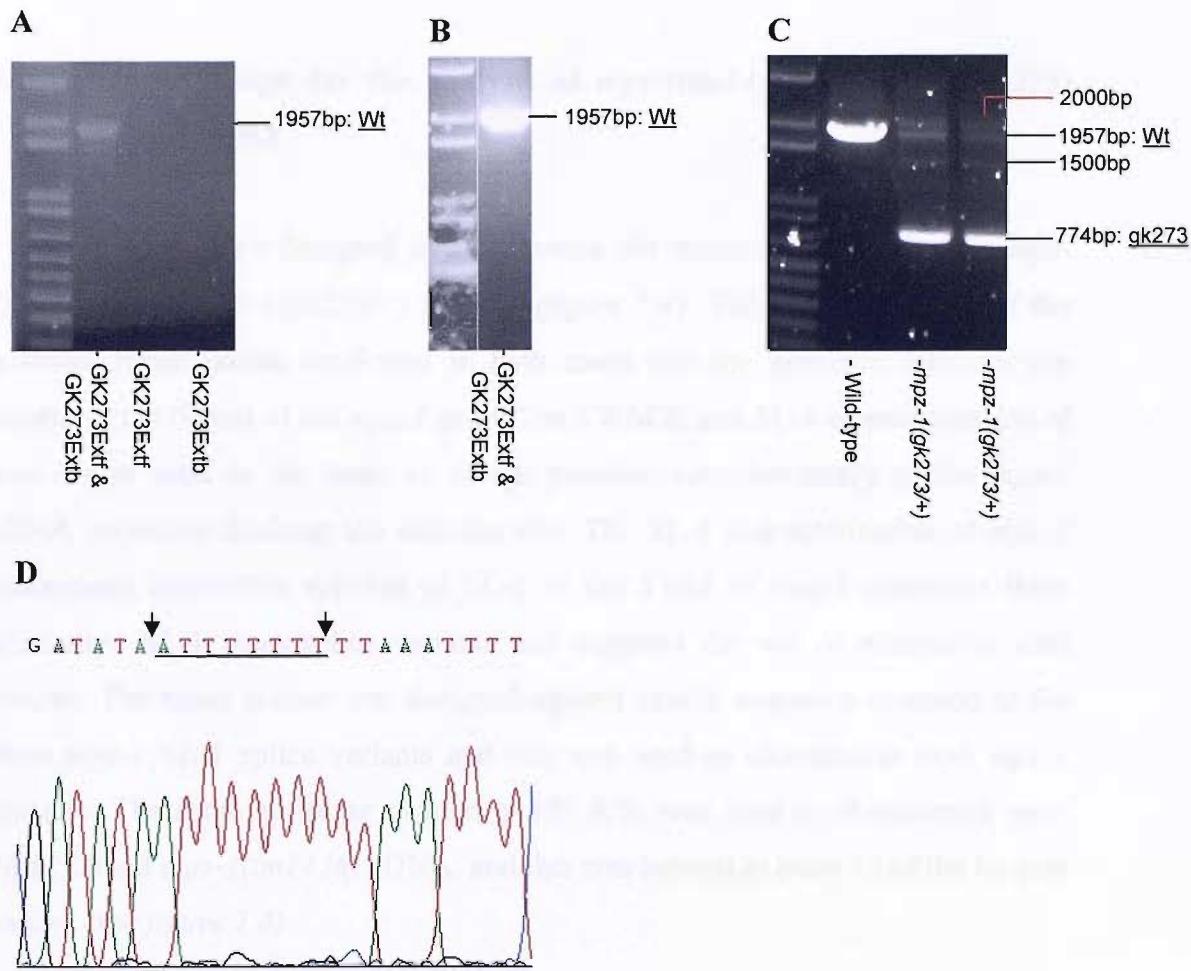
Figure 7.3 Defining the genomic break-points of *mpz-1(gk273)* (A) The primers GK273_Ext_f and GK273_Ext_b were optimised on genomic DNA prepared from a large population of wild-type animals. A single, specific band of the expected size was amplified. **B)** The optimised conditions were used to amplify from genomic DNA of a single wild-type worm. A single band was amplified of the predicted size. **C)** Amplification from genomic DNA of a single *mpz-1(gk273/+)* worm was performed in duplicate. Mutant product: 774bp; wild-type product: 1957bp. Two additional bands, 2kb and 1.5kb were amplified. **D)** A chromotogram of the seqeunced *mpz-1(gk273)* deletion product. The break points are indicated by the arrows and the 8 neucleotide insertion, ATTTTTTT, is underlined.

268

7.3 The significance of the mutations to the SL-1 defined gene model

7.3.1 Primer design for the analysis of *mpz-1(tm1136)* and *mpz-1(gk273)* mutant cDNA

Primers were designed to characterise the transcripts expressed in *mpz-1(tm1136)* and *mpz-1(gk273/+)* animals *(figure 7.4)*. The characterisation of the genomic break points confirmed in both cases that the genomic deletions are located at the 5' end of the *mpz-1* gene. The 5'RACE and SL-1 characterisation of *mpz-1* was used as the basis to design primers, complementary to the *mpz-1* cDNA sequence flanking the deletion site. The SL-1 characterisation of *mpz-1* determined alternative splicing of SL-1 to the 5'end of *mpz-1* generates three alternative SL-1 *trans*-splice variants and suggests the use of alternative start codons. The sense primer was designed against exonic sequence common to the three *mpz-1* SL-1 splice variants and this was used to characterise both *mpz-1* mutants. The same antisense primer, *gk273* A/S, was used to characterize *mpz-1(gk273)*and *mpz-1(tm1136)* cDNA, and this was located in exon 13 of the longest variant *(see figure 7.4)*.
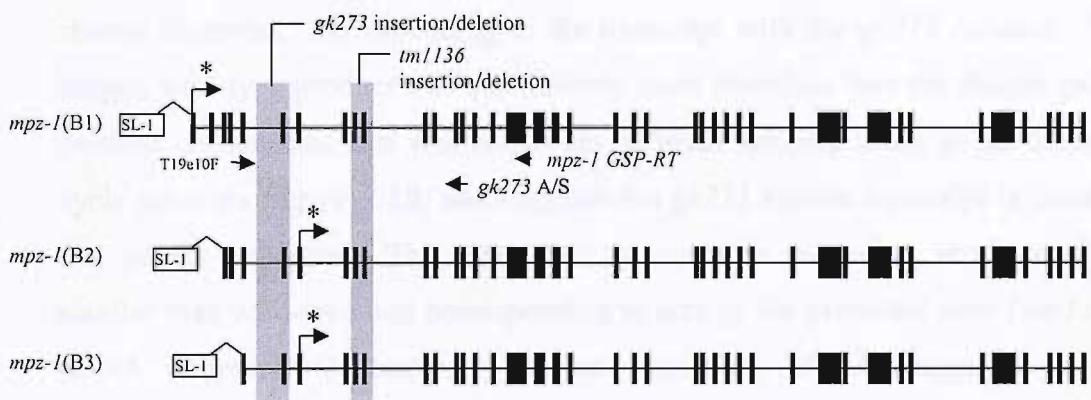
Figure 7.4. The genomic position of the mutations *tm1136* and *gk273* in the SL-1 defined mpz-1 gene models. The position of the genomic deletion/insertion of each mutant allele and defined by the analysis of *mpz-1(tm1136)* and *mpz-1(gk273)* mutant genomic DNA is indicated on each gene model by the greyed region. The location of mutant cDNA chracterization primers T19e10F and gk273 A/S are indicated. The asterisk defines start codons defined by alternative *trans*-splicing of *mpz-1* to SL-1. The primer *mpz-1 GSP-RT* was used as a gene specific primer to perform reverse transcription (see below)

7.3.3 Qualitative analysis of the *mpz-1(gk273/+)* and *mpz-1(tm1136)* mutant cDNA

Total RNA was extracted from wild-type and mutant worm strains in parallel and spectrophotometry was used to quantify the yield. The quantification was confirmed by the electrophoresis of equivalent amounts (5μg) of each preparation *(see figure 7.5A)*. Based upon the gel analysis an equivalent amount of each total RNA sample (5μg) was used to perform a reverse transcription using the same primer in each case (*mpz-1 GSP-RT; see figure 7.4)* and an equal volume of the first strand cDNA synthesised was used to perform the qualitative PCR analysis of the mutant transcripts.

Samples were taken from the wild-type and each mutant PCR at 18, 25, 32 and 40 cycles to compare the abundance of each cDNA species. Samples (10μl) were analysed by agarose gel electrophoresis. The wild-type control sample yielded a single PCR product of the expected size. The *mpz-1(gk273/+)* sample yielded two products, one corresponding in size to the wild-type transcript and a shorter fragment, corresponding to the transcript with the *gk273* deletion . The longer, wild-type product was qualitatively more abundant than the shorter *gk273* deletion product and this was consistent between samples taken at the different cycle intervals *(figure 7.5B)* and suggests the *gk273* mutant transcript is unstable and poorly expressed. The *mpz-1(tm1136)* sample yielded a single product, smaller than wild-type and corresponding in size to the predicted *mpz-1(tm1136)* cDNA fragment. Qualitatively the *mpz-1(tm1136)* cDNA fragment was as abundant as the *mpz-1* wild-type control *(figure 7.5B)*, suggesting the *tm1136* deletion does not affect the stability of this mutant transcript. Mutant cDNA fragments were cloned, sequenced and analysed.
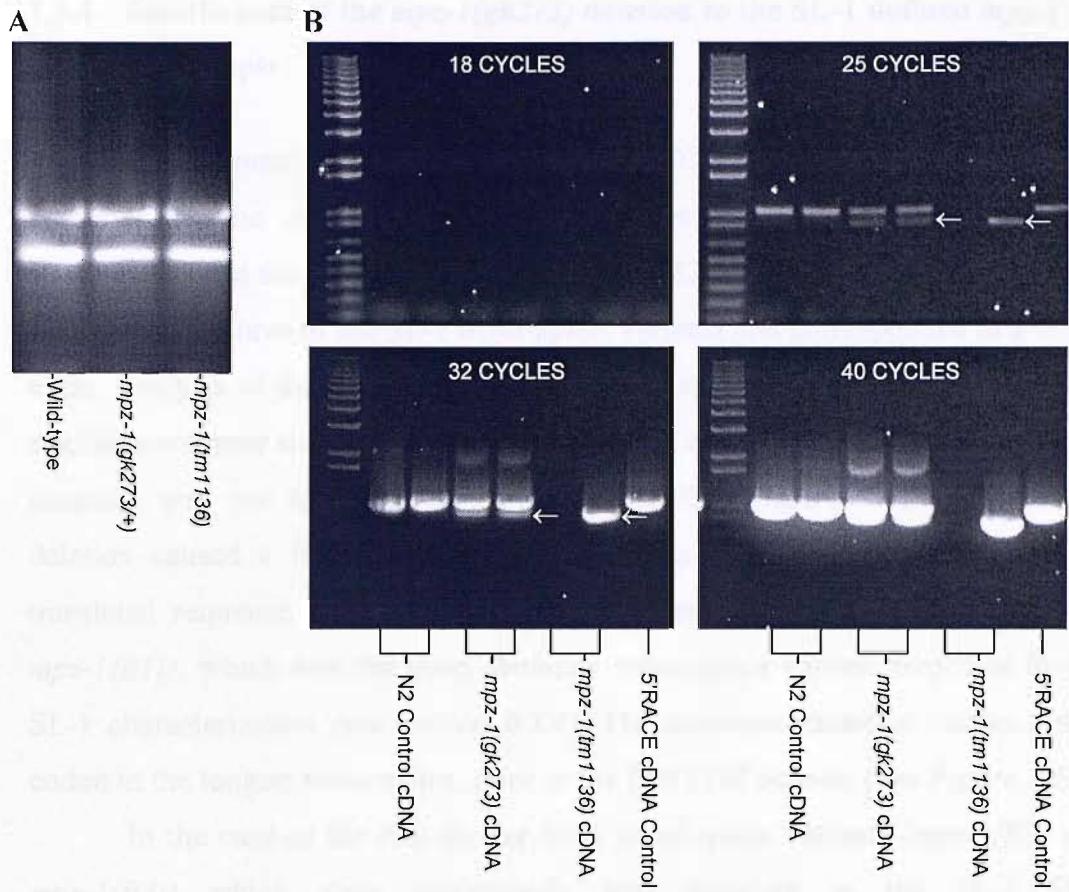
Figure 7.5 Qualitative analysis of the *mpz-1(gk273/+)* and *mpz-1(tm1136)* mutant cDNA A) 5μg of total RNA extracted from wild-type worms and worms expressing the *mpz-1* alleles *mpz-1(gk273/+)* and *mpz-1(tm1136)*. This amount was used to perform a first strand cDNA synthesis with the primer *mpz-1 GSP-RT* **B)** A qualitative cDNA characterization of the two *mpz-1* deletion alleles and wild-type allele. Sample volume run on gel: 10μl. The white arrows point to the mutant cDNA fragments. Note: in the case of gk273/+ the larger, wild-type fragment is more abundant than the mutant fragment in both duplicates. *Cycle: 94°C, 1min 30sec; 30x(94°C, 30sec; 68°C, 30sec; 72°C, 10min); 72°C, 10min.* Primers: T19F and *g273* A/S.

7.3.4 Significance of the *mpz-1(gk273)* deletion to the SL-1 defined *mpz-1* transcripts

The sequence of the *mpz-1(gk273)* cDNA fragment was aligned to sequence of the *mpz-1(C52A11.4A)* (confirmed by sequencing of the est yk1004.e08) and the mutation was defined as a 152 nucleotide deletion in a region common to all three of the SL-1 *trans*-splice variants and corresponded to a single exon. Analysis of the cDNA established the splicing machinery utilizes the next available acceptor site downstream of the deleted exon. The 8 nucleotide genomic insertion was not identified at the level of cDNA *(figure 7.6B)*. The cDNA deletion caused a frame shift that introduces a premature stop codon in the translated sequence of the longest SL-1 *trans*-splice variant of *mpz-1(denoted mpz-1(B1))*, which was the most abundant *trans*-splice variant amplified by the SL-1 characterization *(see section 6.5.4)*. The missense mutation causes a stop codon in the longest *trans*-splice, prior to the first PDZ domain *(See Figure 7.6A)*.

In the case of the two shorter SL-1 *trans*-splice variants (*mpz-1(B2)* and *mpz-1(B3)*) which were qualitatively less abundant in the SL-1 PCR characterization *(see figure 6.4)* the deletion is located upstream of the start codon, defined by the position of the SL-1 *trans*-splice site. The deletion is therefore predicted to be located within the un-translated region of these two *mpz-1* variants. It was predicted that a functional protein could still be made from the *mpz-1(gk273)* allele *(See Figure 7.6A)* if these shorter SL-1 *trans*-splice variants are biologically significant. However, as mentioned already the SL-1 analysis revealed these two transcripts were rare and are unlikley to be as biologically important as the longer SL-1 variant (*mpz-1(B1)*) because it was more abundant. Furthermore, as discussed in *section 7.3.2* the PCR analysis of mutant cDNA suggests the *gk273* transcript is unstable and very poorly expressed and so it is unlikely that the *gk273* mutant transcript would give rise to any protein.

A

gk273 insertion/deletion

STOP

*

SL-1

T19e10F

← gk273 A/S

mpz-1(gk273) PREDICTED PROTEIN PRODUCTS

*

SL-1

*

SL-1

B

GAATCGAAACGATA
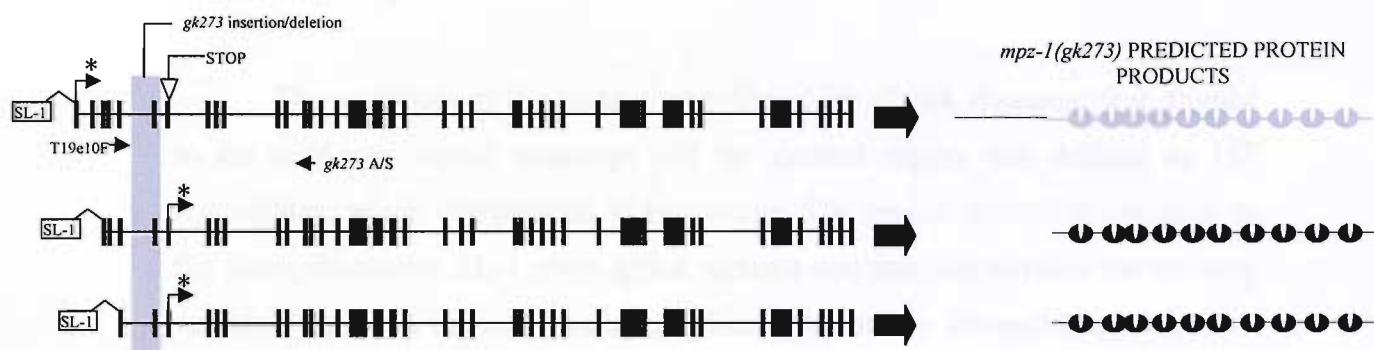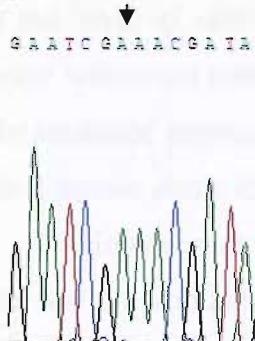
Figure 7.6 Significance of the *mpz-1(gk273)* deletion to the *mpz-1* gene model and transcript A) The three SL-1 defined gene models are depicted together with the deletion site defined by analysis of *mpz-1(gk273)* genomic DNA. The protein predicted to be translated from each SL-1 *trans*-splice variant is shown. Protein regions in black are predicted to be translated whereas those in grey are not, based upon the analysis of *mpz-1(gk273)* mutant cDNA. Note: The *mpz-1* SL-1 splice variants predict different proteins. The *gk273* deletion causes a frame shift in the longest SL-1 *trans*-splice variant leading to the introduction of a premature stop codon (indicated in the figure) prior to the first PDZ domain. However, the *gk273* does not interfere with the translated coding sequence of the other two *mpz-1* SL-1 splice variants. **B)** A chromatogram of the *mpz-1(gk273)* mutant cDNA. The arrow indicates the position of the deletion breakpoint.

7.3.4 Significance of the *mpz-1(tm1136)* deletion to the SL-1 defined *mpz-1* transcripts

The sequence of the mutant *mpz-1(tm1136)* cDNA fragment was aligned to the wild-type *mpz-1* transcript and the deleted region was defined as 167 nucleotides, which corresponds to two exons. The region deleted is common to the three alternative SL-1 *trans*-splice variants and analysis verifies the splicing machinery utilises the next available splice acceptor site downstream of the two deleted exons. This causes a frame-shift in the reading frame and the introduction of a premature stop-codon. The 3 nucleotide genomic insertion was not identified at the level of cDNA *(figure 7.7B)*. The premature stop codon is predicted to occur within the first PDZ domain of MPZ-1a *(See Figure 7.7A)* and subsequently the predicted protein translated from the *tm1136* would not contain the region of the 2-hybrid clone identified as interacting with the MGL-1 C-terminal.

However, as highlighted by the qualitative cDNA analysis *(section 7.3.2)* the *tm1136* mutant transcript is as stable as the wild-type *mpz-1* transcript. Subsequently it may be that protein translation could still be initiated from the *tm1136* transcript but at alternative Methionine residues downstream of the Methionine residues assigned as starts by SL-1 *trans*-splicing. This could give rise to alternative versions of MPZ-1 sufficient to perform the same functions as the wild-type protein and explain why *mpz-1(tm1136)* animals are viable *(see section 7.7 for further discussion)*.
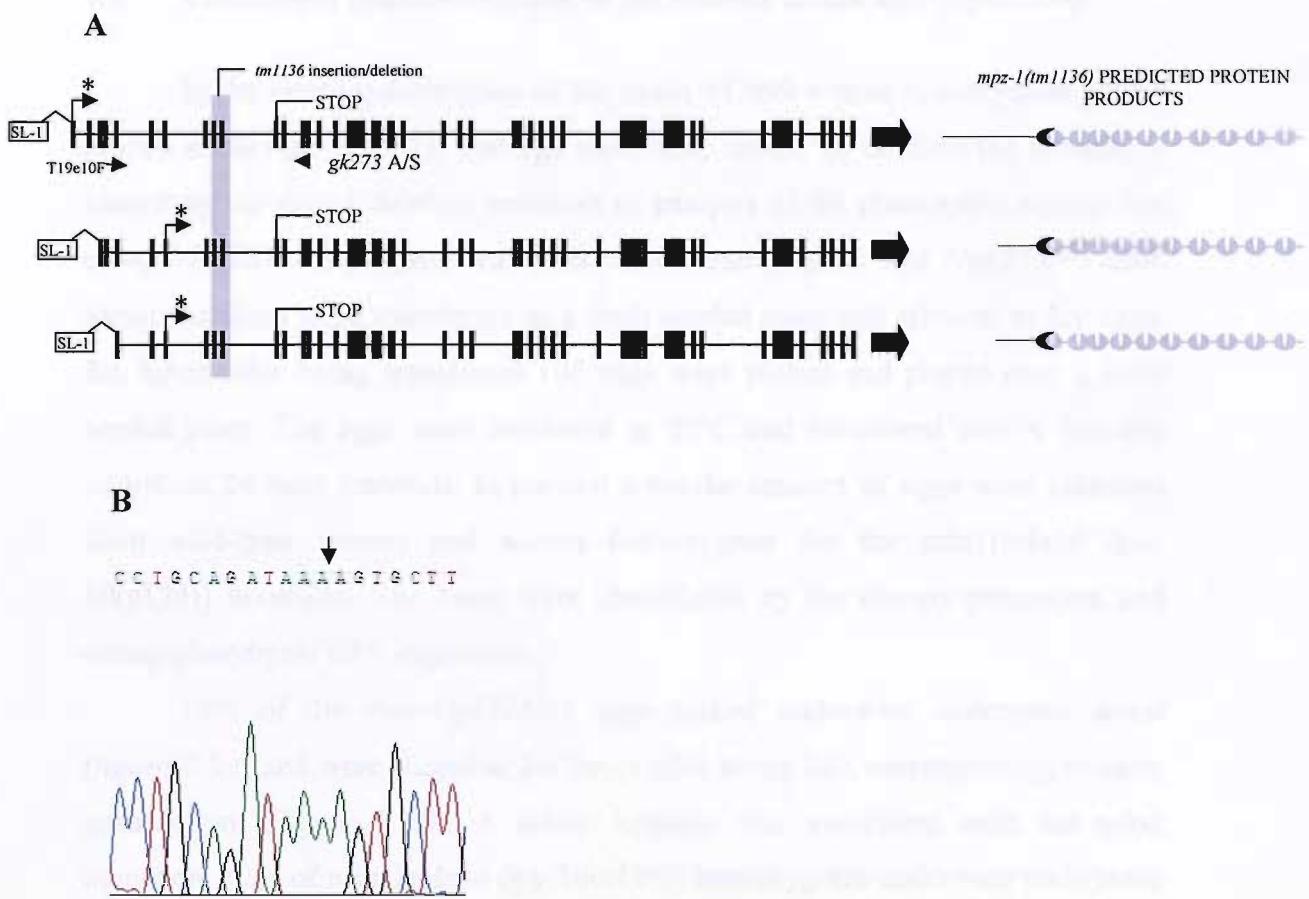
Figure 7.7 **Significance of the** *mpz-1(tm1136)* **deletion to the** *mpz-1* **gene model and transcript**

A) The three *mpz-1* SL-1 *trans*-splice variants are shown and the position of the region of the *tm1136* deletion defined by analysis of mutant genomic DNA is indicated in grey. The predicted protein product of each SL-1 splice variant is shown. Note: The same protein product is predicted from each of the SL-1 splice variants. The *tm1136* deletion introduces a premature stop codon into the transcript of each SL-1 variant. The stop codon causes the protein sequence to terminate in the first PDZ domain.

B) A chromotogram of the sequenced *mpz-1(tm1136)* cDNA PCR product. The black arrow indicates the breakpoint introduced by the deletion.

7.5   Phenotypic characterization of the mutant strain *mpz-1(gk273/+)*

In the existing description of the strain VC599 worms homozygous for the mutant allele *mpz-1(gk273)* undergo embryonic arrest. To confirm the lethality is caused by the *mpz-1* deletion mutation an analysis of the phenotypic segregation of *mpz-1(gk273/+)* progeny was performed. Heterozygous *mpz-1(gk273/+)* adult hermaphrodites were transferred to a fresh seeded plate and allowed to lay eggs. Six hours after being transferred 104 eggs were picked and placed onto a fresh seeded plate. The eggs were incubated at 20°C and monitored over a five-day period, at 24 hour intervals. In parallel a similar amount of eggs were collected from wild-type worms and worms homozygous for the mIn1[mIs14 dpy-10(e128)] inversion. The latter were identifiable by the dumpy phenotype and strong pharyngeal GFP expression.

28% of the *mpz-1(gk273/+)* eggs picked underwent embryonic arrest *(figure 7.8A)* and were staged at 3-6 hours after being laid, corresponding to early gastrulation *(Figure 7.8B)*. A subtle lethality was associated with the mIn1 inversion, 12% of mIn1[mIs14 dpy-10(e128)] homozygotes underwent embryonic arrest *(figure 7.8C)* and 2.5% underwent L1 larval arrest. This correlates with an independent report describing 10% of the progeny from homozygous mIn1[dpy-10] hermaphrodites fail to hatch from the egg or arrest as L1 larvae (Edgley, M.L. and Riddle, D.L. 2001). However no explanation for these observed phenotypes is provided. 50% of the *mpz-1(gk272/+)* progeny were heterozygous, displaying a weak pharyngeal GFP signal and wild-type body morphology and 20% were homozygous for the mIn1 marked inversion. All of the viable offspring expressed a pharyngeal GFP signal, indicating that animals homozygous for the allele *mpz-1(gk273)* and lacking a GFP signal account for the population of the offspring undergoing embryonic arrest.
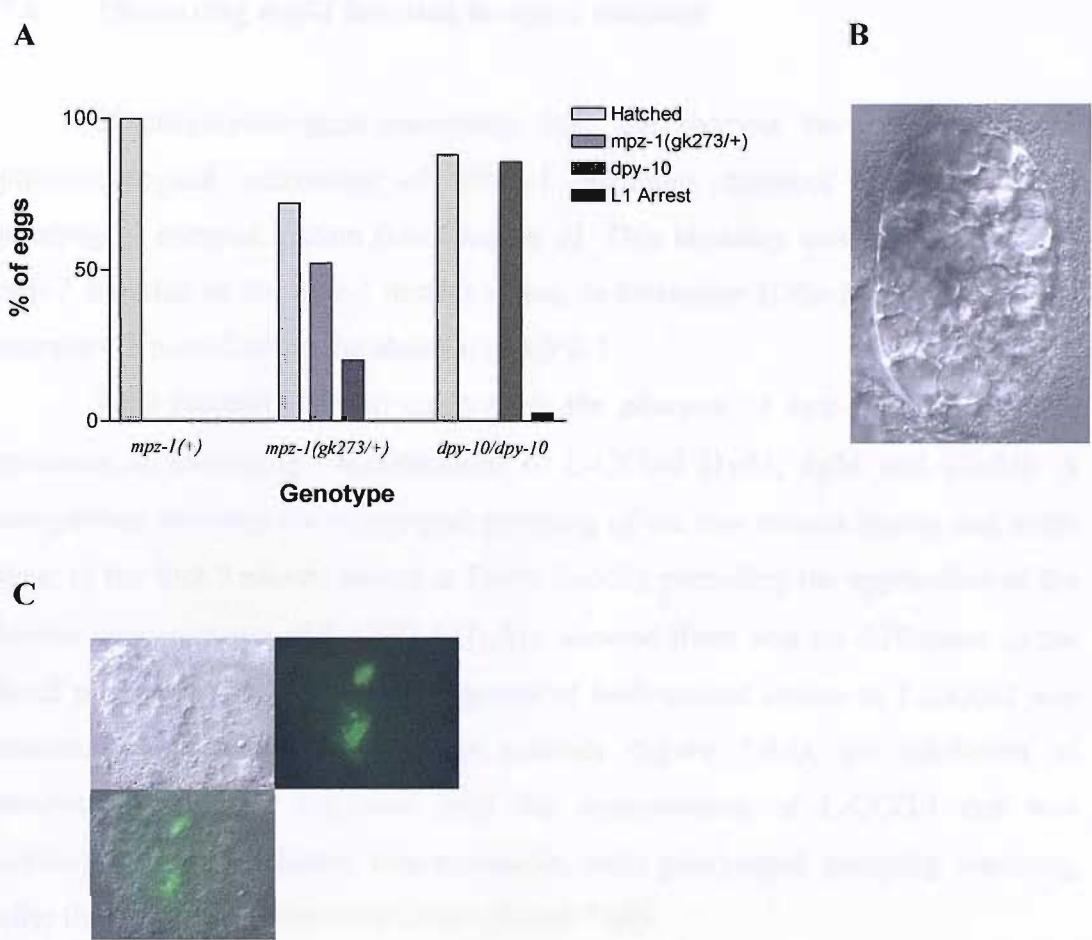
__Figure 7.8__ **Phenotypic characterization of the mutant strain** *mpz-1(gk273/+)*

A) The number of eggs that either: hatched, were heterozyous *mpz-1(gk273/+)* or homozygous mIn1[mIs14 dpy-10(e128)] are shown as a percentage of the total number of eggs that were picked. The genotypes tested were wild-type (*mpz-1*(N2)), heterozygous *mpz-1(gk273/+)* and homozygous mIn1[mIs14 dpy-10(e128)] hermaphrodites. **B)** DIC image of a homozygous *mpz-1(gk273)* egg, staged at 3-6 hours after fertilization. The image was taken 3 days after picking and did not exhibit GFP fluorescence. **C)** A homozygous mIn1[mIs14 dpy-10(e128)] egg, expressing GFP, that failed to hatch. Image recorded at 50hrs post picking.

7.6   Measuring *mgl-1* function in *mpz-1* mutants

Electrophysiological recordings from the pharynx have established the pharmacological activation of MGL-1 regulates network activity of the pharyngeal nervous system *(see Chapter 4)*. This bioassay was used to measure *mgl-1* function in the *mpz-1* mutant strains to determine if the MGL-1 signalling complex is perturbed by the absence of MPZ-1.

EPG recordings were made from the pharynx of *mpz-1* mutants in the presence of increasing concentrations of L-CCG-I (1µM, 3µM and 10µM). A comparison between the pharyngeal pumping of the two mutant strains and wild-type, in the first 5 minute period in Dents directly preceding the application of the lowest concentration of L-CCG-I (1µM), showed there was no difference in the basal pharyngeal pumping. The response of both mutant strains to L-CCG-I was essentially the same as wild-type animals *(figure 7.9A)*, the inhibition of pharyngeal pumping increased with the concentration of L-CCG-I that was applied and the inhibition was reversible, with pharyngeal pumping resuming after the drug was replaced by Dents *(figure 7.9B)*.

To determine if a detectable loss of MGL-1 signalling could be produced by combining the two mutant alleles the trans-heterozygote *mpz-1(gk273/tm1136)* was generated by crossing the two *mpz-1* mutant strains together. *mpz-1(tm1136)* males were mated with *mpz-1(gk273/+)* hermaphrodites and *mpz-1(gk273/tm1136)* hermaphrodites were selectable in the F1 generation by an absence of the pharyngeal GFP signal (marking the *mIn1* inversion). Subsequently EPG recordings were made in the presence of increasing concentrations of L-CCG-I and the animals were not resistant *(figure 7.9A)*.The dose-response profile of *mpz-1(gk273/tm1136)* was essentially the same as the wild type animals and both mutant animals *(figure 7.9B)*.
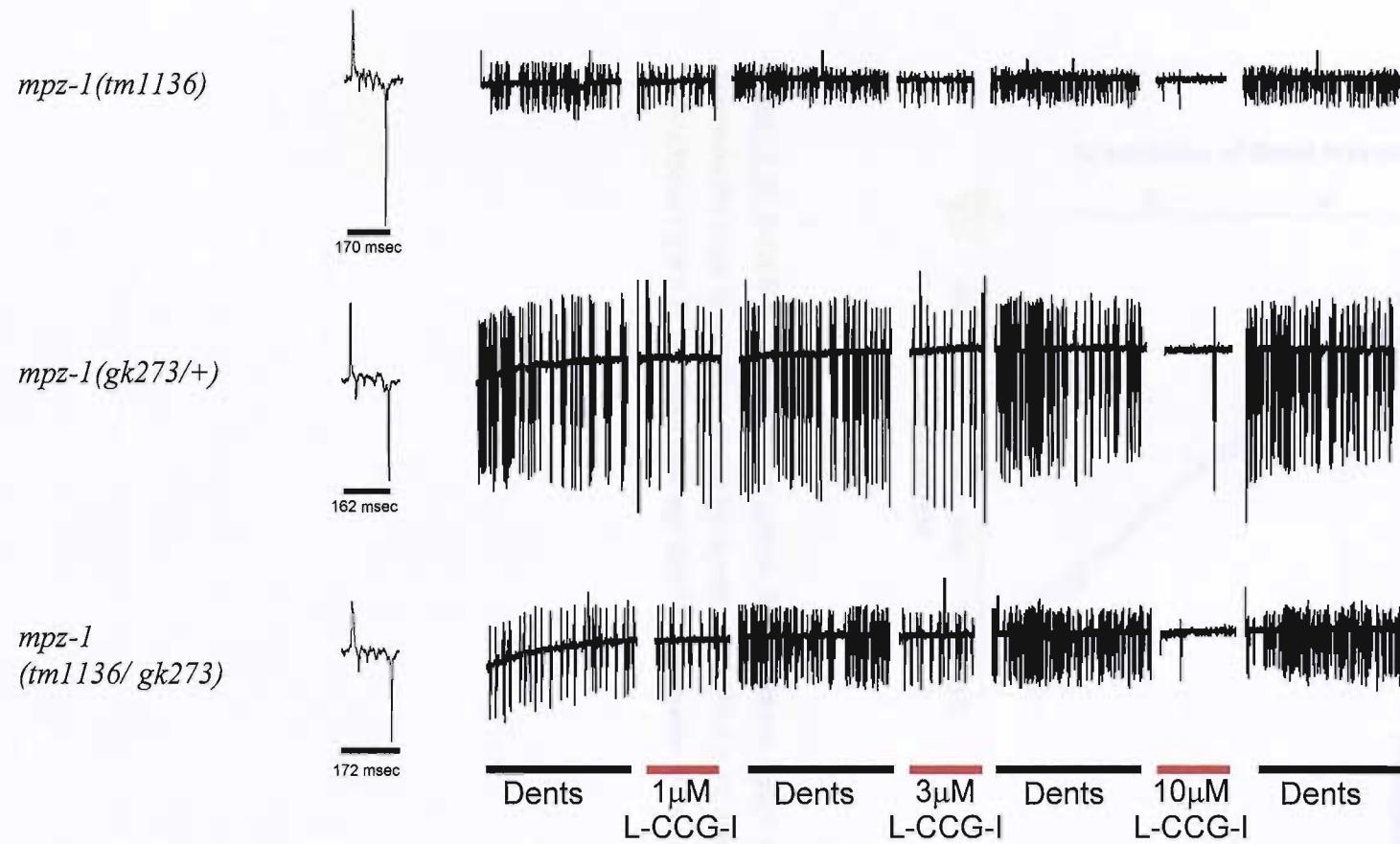
Figure 7.9 **EPG recordings made from *mpz-1* mutants *mpz-1(tm1136)*, *mpz-1(gk273/+)*, *mpz-1(tm1136/gk273)*** EPG recordings were made in the presence of increasing concentrations of the agonist L-CCG-I, 1μM, 3μM and 10μM. An expanded view of a single pump produced by each mutant strain in Dents saline is shown.
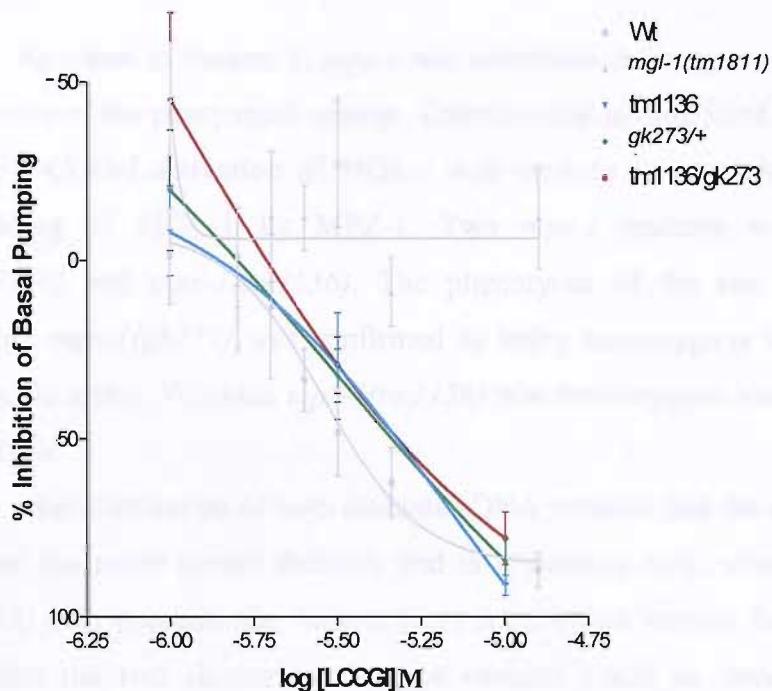
Figure 7.**10** **L-CCG-I dose response graph for the three *mpz-1* mutant strains.** Each point represents the mean % inhibition +/- the SEM. *mpz-1(tm1136)* n = 10, *mpz-1(gk273/+)* n = 12 and *mpz-1(tm1136/gk273)* n = 12. Wild-type and *mgl-1(tm1811)* responses are shown in grey for comparison.

7.7 Discussion

As stated in chapter 5, *mpz-1* was identified as being co-expressed with *mgl-1* in neurons of the pharyngeal system. Therefore the inhibition of pharyngeal pumping by the L-CGG-I activation of MGL-1 was used as an assay to measure the intact scaffolding of MGL-1 by MPZ-1. Two *mpz-1* mutants were available, *mpz-1(gk273/+)* and *mpz-1(tm1136)*. The phenotypes of the two mutants were very different, *mpz-1(gk273)* was confirmed as being homozygous lethal and underwent embryonic arrest. Whereas *mpz-1(tm1136)* was homozygous viable with no obvious phenotype.

Characterization of both mutants cDNA predicts that the allele *mpz-1(tm1136)* contains the more severe deletion and is a putative null, whereas the allele *mpz-1(gk273)* only prevents the longest SL-1 *trans*-splice variant from being expressed. Therefore the two shorter *trans*-splice variants could be translated to generate a functional protein despite the *gk273* deletion. The pharyngeal pumping of both mutants was inhbited by L-CCG-I suggesting the MGL-1 signalling complex is still intact. Furthermore combining the two alelles to generate the transheterozygote *mpz-1(tm1136/gk273)* did not affect the response to L-CCG-I and pumping was still reversibly inhbited.

A possible explanation for the discrepancy between the severity of the *mpz-1(tm1136)* mutation and phenotype is that the *mpz-1(tm1136)* transcript is alternatively spliced to generate a functional protein. As the primers used to characterize the mutant cDNA closely flank the deletion breakpoints the PCR amplification would not have detected alternative splicing to exons located distally of the antisense primer annealing site. The identification of alternative SL-1 *trans*-splice variants suggests that the splicing of *mpz-1* is complex and tightly regulated and the *mpz-1* gene could be alternatively spliced to generate shorter proteins that are still functional. Also the qualitative analysis of the *tm1136* cDNA suggests this mutant transcript is stable and suggests alternative Methionine residues could be utilised, downstream of the deleted region to initiate transaltion of from the *tm1136* transcript and generate MPZ-1 variants that retain the functional properties of the wild-type protein. This could also explain why the MGL-1 response to L-CCG-I is essentially normal in these animals, because splice variants of *mpz-1* could still give rise to proteins that are capable of scaffolding MGL-1 function.

Prediction based on cDNA analysis suggests the *gk273* deletion only affects the protein sequence of the longest SL-1 *trans*-splice variant, which was the most abundant variant amplified during the SL-1 characterization of the *mpz-1* gene 5' region. However, the qualitative analysis of the mutant cDNA suggests the *gk273* transcript is unstable and poorly expressed. If this reflects the situation *in vivo*, protein encoded by alternative SL-1 *trans*-splicing of the *gk273* transcript is also likely be very poorly expressed and may not even be translated due to the instability of the transcript. mRNA stability is thought to perform an important role in development [Newbury, S.F. 2006] and this could be a possible explaination for why the *gk273* deletion causes an embryonic lethality. Also it is known that proteins containing PDZ domains are involved in defining cell polarity during the initial stages of mitosis and the PAR (abnormal PARtitioning of embryonic cytoplasm) proteins are the prototypic example of such a class of proteins (Nance J 2005).

However, the MGL-1 regulation of pharyngeal pumping was not disrupted in the *mpz-1(gk273/+)* mutants. This may be because these animals are heterozygous and the wild-type *mpz-1* transcript is being stably made in these animals, which is sufficient for normal cellular function involving MPZ-1 to proceed. Furthermore as highlighted above alternative variants of MPZ-1 could be expressed from the *tm1136* transcript that retain functionality, which is why MGL-1 function may not have been disrupted in the *trans*-heterozygote either.

Therefore further analysis of the *mpz-1* mutant strains and the proteins expressed from the mutant alleles is required to ascertain how useful the mutants are for defining the role of MPZ-1 in MGL-1 scaffolding. Alternatively these findings raise the issue of how significant PDZ domains are to robust function. The deletion of the C-terminal PDZ ligand of the AMPA receptor GluR1 was recently shown to have little effect on GluR1 function in the hippocampus. Knock-in mice expressing GluR1 lacking the PDZ ligand displayed normal synaptic transmission and GluR1 synaptic localization (Kin C-H et al 2005). This is surprising, since GluR1 has been described as interacting with a number of different PDZ-domain containing proteins (Song I and Huganir RL 2002) but illustrates GluR1 can function normally in the absence of these interactions. However, in a more recent study synaptic transmission and synaptic trafficking of GluR2 was disrupted in the cerebellum of knock-in mice expressing GluR2 with a deleted C-terminal PDZ ligand (Steinberg JP et al 2006). This highlights that the significance of PDZ mediated interactions to robust function

is receptor specific and there is presedence supporting the findings reported here for the the non-essential role of PDZ domains in receptor function.

Chapter 8

General Discussion and Future Work

The evolutionarily conserved role of glutamate as a neurotransmitter and the significant role glutamate plays in a number of complex biological processes of the mammalian nervous system demonstrates it is an important bio-molecule. Understanding how glutamatergic signalling is organised could provide the key to unlocking higher brain functions, such as complex human behaviour. The aim of this study was to expand upon what is currently known about the organisation of glutamatergic signalling. In particular the project aimed to address the current deficit in our understanding of the relationship between receptor scaffolding mechanisms, neuronal cicuitries and the regulation of complex behaviour.

The important role scaffolding mechanisms perform in the organisation of the metabotropic glutamate receptors was introduced in Chapter 1. This study has exploited the mGluR receptor subclass as a model to study receptor scaffolding and combined it with the well-defined, simple neural circuits of *C. elegans* and the quantifiable behaviours they regulate (also introduced in Chapter 1). At the outset of this project there was very little known about the molecular, cellular or functional organisation of the mGluR subclass of receptor in *C. elegans* (designated *mgls*), compared to the iGluR subclass. The findings of this study have added to our current understanding of glutamate signalling in *C. elegans* and specifically the role performed by the *mgls*.

In Chapter 3 the *in silico* chacterization of three *C. elegans* mGluR homologues, *mgl-1*, *mgl-2* and *mgl-3* (Y4C6A.2a) identified conserved structural features of mGluR function. Analysis of cDNA ends experimentally defined the presence of motifs at the N and C-terminals important for both post-translational processing and scaffolding and the alternative splicing of the C-terminal further highlighted the conserved scaffolding role of this domain in *C. elegans*. Sequence alignments identified the *mgls* can be classified into characteristic subgroups based upon homology to their mammalian counterparts (*mgl-1* and *mgl-3* are Group II –like whereas *mgl-2* is Group I-like) and suggests that these subgroups are evolutionarily conserved. The characterization of the cDNA ends of the *mgl* transcripts will facilitate the future amplification of the full length *mgl* transcripts, which will serve further *in silico* and molecular/cellular studies.

The cellular characterization of *mgl-1* and *mgl-2* in Chapter 3 revealed a differential and widespread expression pattern in the *C. elegans* nervous system and a

selective expression of *mgl-1* in the pharyngeal nervous system. In keeping with other neural circuits the interaction between distinct neurotransmitters controls pharyngeal function and ultimately feeding behaviour in *C. elegans.* Accordingly in Chapter 4 mammalian mGluR agonists were used to identify if *mgls* regulate this network. Electropharyngeogram recordings were used to measure pharyngeal pumping and identified network activity was alternatively modulated by *mgl* receptor subtypes. The agonist L-CCG-I decreased network activity in an *mgl-1* dependent manner, whereas at low concentrations L-CCG-I increased network activity in an *mgl-3* dependent manner. Both findings support the contention that *mgls* perform a neuromodulatory role in the control of *C. elegans* feeding.

In order for an animal to survive it must be able to modulate its feeding behaviour in response to an ever changing environment. For example when food is becoming less available the animal may decide to increase its rate of feeding as a strategy for survival. In adult *C. elegans mgls* may be involved in this decision making process and may serve in the pharyngeal nervous system to modulate the worms feeding behaviour when it is exposed to different environmental cues, such as the abundance of food. This raises a number of interesting issues, such as how the sensory cue of food is detected, how it is integrated into the pharyngeal nervous system and what is the importance of *mgl* scaffolding to this process.

The role of the *C. elegans* iGluRs in the locomotory command interneurons (introduced in Chapter 1) sets a precedent for the importance of glutamatergic signalling in the process of sensory integration in the *C. elegans* nervous system. Interestingly the NSMs (putative neurons expressing *mgl-1*) have been suggested to have neurohumoral functions and signal the presence of food. This is unproven, but anatomically the NSMs have been identified as having sensory-like endings located at the point of bacterial accumulation in the pharynx and the NSMs contain a cocktail of transmitters that have established roles in regulating network activity of the pharynx; these include glutamate, 5-HT and various neuropeptides. Combined this would suggest the NSMs are ideal candidates for integrating sensory cues with the regulation of pharyngeal pumping and this may be dependent upon *mgl-1* signalling in this neuron. Initial observations have shown that *mgl-1* animals do not modulate their feeding behvaiour in the same way as wild-type animals in response to food being taken away (Sara Luedtke, unpublished).

The pharmacological approach taken in this study has highlighted *mgl* signalling is important in the pharynx but the major caveat is that it relies upon the use of mammalian agonists. Although the *Drosophila* mGluR sets precedent for the strong evolutionary conservation of the mGluR agonist binding domain, in the current study the agonists have been applied to a neural circuit which may contain alternative sites of action in addition to the *mgls*. In view of this an important aspect of future study would be defining the role of the *mgls* under native conditions and one way of tackling this issue would be to generate a transgenic animal expressing a constituvely active receptor and to then determine if the pharmacological phenotype is recapitulated. Recent studies performed on mammalian mGluR subtypes sets precedent for such an approach, where a constitutively active receptor has been made by mutating a conserved Glutamine residue to a Valine, in the C-terminal region of the second intracellular loop of mGluR8 that is involved in G-protein coupling [Yamashita, T et al 2004]. This Glutamine residue is conserved in all three of the *mgl* receptor subtypes.

The molecular characterization of protein scaffolding motifs in the C-terminal of the *mgls* (Chapter 3) supported the view that these receptors are subject to scaffolding mechanisms mediated by this domain. In Chapter 5 controls performed with the *mgl* receptor C-terminals identified the *mgl-1* C-terminal as the most suitable candidate for use in the LexA yeast-2-hybrid screen. Three interacting genes were prioritised from the screen with homology to known scaffolding proteins with established roles in receptor scaffolding. These were *mpz-1*/MUPP-1, *ptp-1*/PTP-MEG and *tag-60*/NHERF-2 (the mammalian homologues are underlined).

The mouse homologue of *mpz-1*, *Mpdz* has been identified as a genetic determinant of alcohol addiction. This is a complex behaviour that is likely to involve a number of different molecular targets and *Mpdz-1* has been identified as interacting with a diverse array of molecules. Ethanol has been shown to have a physiological effect on *C. elegans* by depresssing locomotory behaviour and this has encouraged its use as a model to identify the molecular targets of ethanol (Davies AG et al 2004). As like its mammalian counterpart, *mpz-1* has the molecular capabilities to interact with a number of different protein partners and it has been described elsewhere as binding to 5-HT receptors. Antagonists of Group I mammalian receptors have been shown to have anti-alcohol addiction effects in mouse strains that have a high preference for alcohol [Lominac, KD et al 2006]. This suggests that mGluR subtypes are involved in

ethanol effects upon the brain and behaviour. The scaffolding of *mgl-1* by *mpz-1* and in *C. elegans* may provide future insights into the molecular organisation of mGluR signalling underlying alcohol addiction.

In Chapter 6 the biological significance of the three proritised interactions was assessed *in vivo* within the context of the whole animal at the level of gene co-expression. Two genes, *mpz-1* and *ptp-1* were verified as being co-expressed with *mgl-1* in neurons and were prioritised as being biologically relevant. In the case of *tag-60(C01F6.6a)* gene co-expression was not identified and demonstrates *C. elegans* provides a strong model to study protein-protein interactions because it facilitates the rapid prioritisation of an *in vitro* identified interaction *in vivo* and bypasses the need for *in vitro* biochemical verifications dependent upon antibodies.

Furthermore in Chapter 6 the reliability of promoter::reporter constructs was adressed and the limitations of the genome annotatations were highlighted. This lead to a redefined understanding of the molecular criteria and experimental verifications required for generating promotor::reporter constructs to report upon the accurate expression pattern of genes, such as *mpz-1*, with complex promotors. The expression pattern established for *ptp-1* and *tag-60* has provided the first description of the cellular distribution of these genes and suggests scaffolding functions outside of the *mgl-1* receptor. It is important to realise, however, that subcellular co-localization of *mgl-1* with the 2-hybrid interactors at the level of protein would be a more definitive verification that the interaction occurs *in vivo*. The characterization of the cDNA ends of the interacting proteins in this study would enable full length cDNAs to be amplified for the production of such protein::reporter constructs.

In Chapter 6 the co-expression of *mpz-1* with *mgl-1* in the pharyngeal circuit is described and the putative neuronal correlates of this co-expression are suggested. The molecular and functional chracteristics of two neurons, NSM and M4, makes them good candidates for the *mgl-1* modulation of feeding behaviour. In Chapter 7 the pharyngeal assay was used to determine if *mpz-1* is required to scaffold *mgl-1* function in the pharyngeal neural network. The L-CCG-I pharmacological assay was performed on the two available *mpz-1* mutants; *mpz-1(gk273/+)*, *mpz-1(tm1136)* and the trans-heterozygote *mpz-1(gk273/tm1136)*. The analysis of mutant cDNA predicts both alleles are likely to be nulls, however the two mutant alleles give rise to very different phenotypes, *mpz-1(gk273)* is homozygous lethal and *mpz-1(tm1136)* is homozygous viable. Qualitative analysis of mutant cDNA suggests the *mpz-1(gk273)*

transcript is unstable and it is unlikely to encode protein. The *mpz-1(tm1136)* transcript looked as stable as the wild-type transcript and may therefore still encode protein if alternative downstream starts are utilised. This does highlight further work is required to expand upon the molecular characterization of *mpz-1* and the two mutants used in this study because of the complexities of the *mpz-1* promoter and *trans*-splicing events described in Chapter 6. This would potentially involve generating an antibody to analyse the MPZ-1 protein being expressed in the two mutant strains. *mgl-1* function was found to be intact in the *mpz-1* mutant pharynxes, which responded as wild-type to L-CGG-I and suggests the *mgl-1* signalling complex is intact in the *mpz-1* mutant animals.

This highlights the need to further characterise the neuronal basis of the L-CCG-I response and establish the cells through which *mgl-1* signalling is regulating neuronal activity in the pharynx. It may be that the response to L-CGG-I is being mediated by *mgl-1* through a cell type that does not express *mpz-1* or that in this cell type *mpz-1* is not required to scaffold *mgl-1* function, which may be why the mutant pharynxes still respond to the agonist. This issue could be addressed by neuronal ablation studies and by a process of elimination the relevant cell could be identified. For instance, if laser ablation of the NSMs were to cause a loss of the L-CGG-I response it would provide good evidence for *mgl-1* signalling in this neuron and identify the NSMs as the neuronal component of the pharyngeal circuit by which *mgl-1* modulates feeding.

Pharmacological assays have established *mgl-1* and *mgl-3* subtypes serve as neuromodulators in the pharyngeal neural network that controls feeding behaviour in *C. elegans*. However, the expression of *mgl-1* and *mgl-2* in neurons outside of the pharyngeal nervous system suggests these receptors may perform neuromodulatory roles in other circuits as well. The co-expression of *mgl-1* with *mpz-1* and *ptp-1* in neurons outside of the pharyngeal nervous system highlights the need to develop additional assays to study *mgl-1* scaffolding within these cells Phenotypic characterization of *mgl-1, mgl-2* and *mgl-3* mutant animals has highlighted a neuromodulatory role for the *mgls* in locomotory behaviours (Ting Zeng, unpublished findings) which may be exploited to assay *mgl* scaffolding outside of the pharynx.

Until now there has been little information about the molecular, cellular and functional organisation of the *mgl* class of glutamate receptor in *C. elegans*. Instead

research has focused primarily upon the ionotropic class, which is surprising since glutamate performs a number of important roles in *C. elegans* (as discussed in Chapter 1). The identification of *mgl-1* and *mgl-3* as neuromodulators of the pharyngeal nervous system provides the first functional description of any *mgl* receptor in *C. elegans*. In addition it has established that *mgl-1* is a metabotropic glutamate receptor and it is involved in the pharyngeal response to mammalian mGluR agonists. The expression of *mgl-1* and *mgl-3* in heterologous cell lines will enable the pharmacological profile of these receptors to be further probed and would enable them to be further classified into the subgroups of metabotropic glutamate receptors already established from the mammalian system and *Drosophila*.

The alternative splicing of the MGL-2 and MGL-3 C-terminals described in Chapter 3 could underlie functional diversity. Alternative splicing of mGluR1 has been shown to activate different patterns of calcium release in *Xenopus oocytes* (Pin JP et al 1992) and confer different sensitivities to agonist (Flor, P.J. et al 1996). As discussed in Chapter 1 the intracellular C-terminal performs an important role in mediating interactions with scaffolding proteins that organise the signalling, targeting and anchoring of the receptor. The alternative splicing of mGluR1 C-terminal generates different molecular signals that are responsible for targeting different splice variants to different cellular compartments (Francesconi, A. and Duvoisin R.M. 2002). In the case of the mGluR7 splice variant mGluR7b the C-terminal consists of several motifs that each are responsible for binding distinct proteins (Enz, R. Croci, C. 2003). It could therefore be envisaged that alternative splicing of the C-terminal can alter the profile of binding motifs and therefore give rise to C-terminal with different binding specificities. Future work may involve performing 2-hybrid screens with different C-terminal splice variants to identify alternative scaffolding proteins.

The identification of proteins interacting with the C-terminal of MGL-1 has in some cases supplemented what is already known about them, as is the case for *mpz-1*. Whereas in the case of *tag-60(C01F6.6a)* and *ptp-1* there is nothing published about their function in *C. elegans* and their interaction with MGL-1 has shed the first light upon the purpose these proteins serve. As highlighted in Chapter 6, gene promoter constructs are useful tool to define the cellular expression pattern of a protein, however, they are not infallable. Hence, in the case of *tag-60(C01F6.6a)*, where no co-expression was seen with *mgl-1*, future work may entail generating several gene reporter constructs that incorporate more extensive regions of the promoter to fully

define the expression of this gene in *C. elegans*. Both what is already known about *mpz-1* in *C. elegans*, in terms of its interaction with other GPCRs and signalling molecules and what is known about the role of mammalian homologues make this a very interesting protein for future study. Already it appears the multifunctional role of this protein is conserved betwen mammals and *C. elegans*.

The pharyngeal bioassay failed to establish if *mpz-1* is required to scaffold *mgl-1* function in this neural network, however as highlighted further characterization of the *mpz-1* mutant strains is needed. Furthermore, currently there are no availbale knockouts for *ptp-1* which was also identified as being co-expressed in the pharyngeal nervous system. Therefore the bio-assay may still be useful in shedding light upon the physiological significance of the interaction between *ptp-1* and *mgl-1*. As mentioned previously the development of additional assays for *mgl* function will assist in furthering our understanding of the role of these receptors in *C. elegans* and provide a basis by which to study the role of *mgl* scaffolding proteins outside of the pharyngeal nervous system.

REFERENCES

Aiba A, Kano M, Chen C, Stanton ME, Fox GD, Herrup K, Zwingman TA, Tonegawa S. (1994) Deficient Cerebellar Long-term Depression and Impaired Motor Learning in mGluR1 Mutant Mice. *Cell.* **79**, 377-388.

Aiba A, Chen C, Herrup K, Rosenmund C, Stevens CF, Tonegawa S. (1994) Reduced Hippocampal Long-Term Potentiation and Context-Specific Deficit in Associative Learning in mGluR1 Mutant Mice. *Cell.* **79**, 365-375.

Albertson DG, Thomson JN (1976) The pharynx of Caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci.* **275**(938):299-325.

Alagarsamy S, Marino MJ, Rouse ST, Gereau RW 4th, Heinemann SF, Conn PJ. (1999) Activation of NMDA receptors reverses desensitization of mGluR5 in native and recombinant systems. *Nature Neuroscience.* **3**, 234-240.

Alagarsamy S, Saugstad J, Warren L, Mansuy IM, Gereau RW 4th, Conn PJ. (2005) NMDA-induced potentiation of mGluR5 is mediated by activation of protein phosphatase 2B/calcineurin. *Neuropharmacology.* **49** Suppl 1:135-45.

Allison DW, Gelfand VI, Spector I, Craig AM. (1998) Role of actin in anchoring postsynaptic receptors in cultured hippocampal neurons: differential attachment of NMDA versus AMPA Receptors. *Journal of Neuroscience.* **18**, 2423-2436.

Allison DW, Chervin AS, Gelfand VI, Craig AM. (2000) Postsynaptic scaffolds of excitatory inhibitory synapses in hippocampal neurons: Maintenance of core components independent of actin filaments and microtubules. *The Journal of Neuroscience.* **20**, 4545-4554

Altun, Z. F. and Hall, D. H. (2005) Alimentary System. The Pharynx. In *WormAtlas.* http://www.wormatlas.org/handbook/excretory.htm

Ango F, Pin JP, Tu JC, Xiao B, Worley PF, Bockaert J, Fagni L. (2000) Dendritic and axonal targeting of of type 5 metabotropic glutamate receptor is regulated by Homer1 proteins and neuronal excitation. The *Journal of Neuroscience.* **20**, 8710-6.

Ango F, Prezeau L, Muller T, Tu JC, Xiao B, Worley PF, Pin JP, Bockaert J, Fagni L. (2001) Agonist independent activation of metabotropic glutamate receptors by the intracellular protein Homer. *Nature.* **411**, 962-965.

Ango F, Robbe D, Tu JC, Xiao B, Worley PF, Pin JP, Bockaert J, Fagni L. (2002) Homer-dependent cell surface expression of metabotropic glutamate receptor type 5 in neurons. *Molecular Cell Neuroscience.* **20**, 323-9.

Armstrong N, Gouaux E. (2000) Mechanism for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron.* **28**:165–181.

Aroian RV, Koga M, Mendel JE, Ohshima Y, Sternberg PW. (1990) The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes a tyrosine kinase of the EGF receptor subfamily. *Nature.* **348**, 693-699.

Ashby MC, De La Rue SA, Ralph GS, Uney J, Collingridge GL, Henley JM (2004) Removal of AMPA receptors (AMPARs) from synapses is preceded by transient endocytosis of extrasynaptic AMPARs. *J Neurosci.* **24**(22):5172-6.

Avery L, Horvitz HR. (1987) A cell that dies during wild-type C. elegans development can function as a neuron in a ced-3 mutant *Cell.* **51**(6):1071-8.

Avery L, Horvitz HR. (1989) Pharyngeal pumping continues after laser killing of the pharyngeal nervous system of C. elegans. Neuron. **3**(4):473-85

Avery, L. Horvitz, HR. (1990) Effects of starvation and neuroactive drugs on feeding in *Caenorhabditis elegans*. *J. Exp Zoology*. **253**, 263-270.

Avery, L (1993) The genetics of feeding in Caenorhabditis elegans. *Genetics*. **133**, 897-917.

Avery, L. (1993) Motor neuron M3 controls pharyngeal muscle relaxation timing in *C.elegans. Journal of Experimental Biology*. **175**, 283-297.

Avery L, Raizen D, Lockery S. (1995) Electrophysiological Methods. *Methods in Cell Biology*. **48**, 257-269.

Avery, L (1997) Feeding and defecation. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, Riddle DL, Blumenthal T, Meyer BJ, Priess JR, editors. *C.elegans* II. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, p 679–716.

Avery L, Shtonda BB. (2003) Food transport in the C. elegans pharynx. *J Exp Biol*. **206**(Pt 14):2441-57

Azkue J. J., Mateos J. M., Elezgarai I., Benitex R., Osorio A., Diez J., Bilbao A., Bidaurrazaga A., Grandes P. (2000) The metabotropic glutamate receptor subtype mGluR 2/3 is located at extrasynaptic loci in rat spinal dorsal horn synapses. *Neurosci. Lett.* **287**, 236–238.

Bahn, S., Volk, B. and Wisden, W. (1994) Kainate receptor gene expression in the developing rat brain. *Journal of Neuroscience*. **14**, 5525-5547.

Bamber BA, Beg AA, Twyman RE, Jorgensen EM. (1999) The caenorhabditis elegans unc-49 locus encodes multiple subunits of a heteromultimeric GABA receptor. *The Journal of Neuroscience*. **19**, 5348-59.

Baran R, Aronoff R, Garriga G. (1999) The *C.elegans* homeodomain gene unc-42 regulates chemosensory and glutamate receptor expression. *Development.* . **126**, 2241-2251

Bargmann CI, Hartwieg E, Horvitz HR. (1993) Odorant selective genes and neurons mediate olfaction in *C. elegans. Cell.* **74**, 515-527.

Bargmann CI, Kaplan JM. (1998) Signalling transduction in the *C.elegans* nervous system. *Neuroscience.* **21**, 279-308.

Bargman, CI (1998) Neurobiology of the Caenorhabditis elegans Genome. *Science.* **282**, 2028-2033

Bargmann CI, Horvitz HR. (1991) Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in C. elegans. *Neuron.* **7**, 729-742.

de Bartolomeis A, Iasevoli F. (2003) The Homer family and the signal transduction system at glutamatergic postsynaptic density: potential role in behavior and pharmacotherapy.*Psychopharmacol Bull.* **37**, 51-83.

Batchelor AM, Madge DJ, Garthwaite J. (1994) Synaptic activation of metabotropic glutamate receptors in the parallel fibre-Purkinje cell pathway in rat cerebellar slices. *Neuroscience.* **63**, 911-915.

Becamel C, Figge A, Poliak S, Dumuis A, Peles E, Bockaert J, Lubbert H, Ullmer C. (2001) Interaction of serotonin 5-hydroxytryptamine type 2C receptors with PDZ10 of the multi-PDZ domain protein MUPP1. *The Journal of Biological Chemistry*. **276**, 12974-12982.

Becamel C, Alonso G, Galeotti N, Demey E, Jouin P, Ullmer C, Dumuis A, Bockaert J, Marin P. (2002) Synaptic multiprotein complexes associated with 5-HT(2C) receptors: a proteomic approach. *EMBO J.* **21**(10):2332-42.

Beg AA, Jorgensen EM (2003) EXP-1 is an excitatory GABA-gated cation channel. *Nature Neuroscience.* **6**, 1145-52.

Bespalov A, Dumpis M, Piotrovsky L, Zvartau E (1994) Excitatory amino acid receptor antagonist kynurenic acid attenuates rewarding potential of morphine. *European Journal of Pharmacology.* **264**, 233-9.

R Baran, R Aronoff, and G Garriga (1999)The C. elegans homeodomain gene unc-42 regulates chemosensory and glutamate receptor expression. *Development.* **126**, 2241-2251.

Bimboim, HC and Doly, J. (1979) A rapid alkaline lysis procedure for screening recombinant plasmid DNA. *Nucleic Acid Research.* **7**, 1513-1522.

Blumcke I, Behle K, Malitschek B, Kuhn R, Knopfel T, Wolf HK, Wiestler OD. (1996) Immunohistochemical distribution of metabotropic glutamate receptor subtypes mGluR1b, mGluR2/3, mGluR4a and mGluR5 in human hippocampus *Brain Research.* **736**, 217-226.

Bockaert, J. Pin, J-P. (1999) Molecular tinkering of G-protein-coupled receptors: an evolutionary success. *The EMBO Journal.* **18**, 1723-1729.

Bockaert, J. Marin, P. Dumuis, A. Fagni, L. (2003) The 'magic tail' of G-protein-coupled receptors: anchorage for functional protein networks. *FEBS Letters.* **546**, 65-72

Boeckers, T.M (2002) ProSAP/Shank proteins- a family of higher order organizing molecules of the postsynaptic density with am emerging role in human neurological disease. *Journal of Neurochemistry.* **81**, 903-910

Bogdanik, L. et al (2004) The Drosophila metabotropic glutamate receptor DmGluRA regulates activity-dependent synaptic facilitation and fine synaptic morphology. *Journal of Neuroscience.* **24**. 9105-16.

Bolton, MM et al (2000) Localization and Stabilization of Ionotropic Glutamate Receptors at Synapses. *Cellular and Molecular Life Sciences.* . **57**, 1517-1525.

Borg, J-P et al (1998) Identification of an evolutionary conserved heterotrimeric protein complex involved in protein targeting. *Journal of Biological Chemistry.* **273**, 31633-31636.

Bortolotto, ZA et al (1999) Roles of metabotropic glutamate receptors in LTP and LTD in the hippocampus. *Current Opinion in Neurobiology.* **9**, 299-304.

Bottai D, Guzowski JF, Schwarz MK, Kang SH, Xiao B, Lanahan A, Worley PF, Seeburg PH. (2002) Synaptic activity-induced conversion of intronic to exonic sequence in Homer 1 immediate early gene expression. *The Journal of Neuroscience.* **22**, 167-75.

Boulter, J., Hollmann, M., O'Shea-Greenfeld, A., Hartley, M., Deneris, E., Maron, C. and Heinemann, S. (1990) Molecular cloning and functional expression of glutamate receptor subunit genes. *Science.* **249**, 1033-1037.

Bradke, F. et al (1999) Sorting out receptor trafficking. *Neuron.* **22**, 411-417.

Bradke F, Dotti CG (2000) Differentiated neurons retain the capacity to generate axons from dendrites. *Current Biology.* **10**(22):1467-70.

Brakeman PR, Lanahan AA, O'Brien R, Roche K, Barnes CA, Huganir RL, Worley PF. (1997) Homer: a protein that selectively binds metabotropic glutamate receptors *Nature.* **386**, 284-288.

Bredt, DS. et al (1998) Sorting out genes that regulate epithelial and neuronal polarity. *Cell.* **94**, 691-4.

Bredt, DS. et al (2003) Assembly and plasticity of the glutamatergic postsynaptic specialization. *Currrent Opinion in Neurobiology.* **13**, 111-118.

Brenner S. (1974) The genetics of Caenorhabditis elegans. *Genetics.* **77**(1):71-94.

Brockie, P. et al (2001) Differential expression of glutamate receptor subunits in the nervous system of *C.elegans* and their regulation by the homeodomain protein UNC-42. *The Journal of Neuroscience.* **21**, 1510-1522.

Brockie, P. et al (2001) The *C.elegans* glutamate receptor subunit NMR-1 is required for slow NMDA activated currents that regulate reversal frequency during locomotion. *Neuron.* **31**, 617-630.

Brody T, Cravchik A (2000) *Drosophila melanogaster* G protein coupled receptors. *Journal of Cell Biology.* **150**, F83–F88.

Brown, T.A. (2002) Chapter 9 in the book Genomes. Second edition. BIOS Scientific Publishers Ltd.

Brownlee, D. Fairweather, I (1999) Exploring the neurotransmitter labyrinth in nematodes. *TINS.* **22**, 16-24.

Bruno V, Battaglia G, Copani A, Giffard RG, Raciti G, Raffaele R, Shinozaki H, Nicoletti F. (1995) Activation of class II or III metabotropic glutamate receptors protects cultured cortical neurons against excitotoxic degeneration. *European Journal of Neuroscience.* **7**, 1906-1913.

Burbea, M et al (2002) Ubiquitin and AP180 regulate the abundance of GLR-1 glutamate receptors at postsynaptic elements in *C.elegans. Neuron.* **35**, 107-120.

Cai Z, Saugstad JA, Sorensen SD, Ciombor KJ, Zhang C, Schaffhauser H, Hubalek F, Pohl J, Duvoisin RM, Conn PJ. (2001) Cyclic AMP-dependent protein kinase phosphorylates group III metabotropic glutamate receptors and inhibits their function as presynaptic receptors. *Journal of Neurochemistry.* **78**, 756-66.

Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A.* **99**(12):7877-82.

Canepari, M. et al (2003) Evidence for protein tyrosine phosphatase, tyrosine kinase and G-protein regulation of the parallel fiber metabotropic slow EPSC of rat cerebellar purkinge neurons. *The Journal of Neuorscience.* **23**, 4066-4071.

Cao, T. et al (1999) A kinase regulated PDZ-domain interaction controls endocytic sorting of the β2-adrenergic receptor. *Nature.* **401**, 286-290.

Castillo PE, Malenka RC, Nicoll RA. (1997) Kainate receptors mediate a slow postsynaptic current in hippocampal CA3 neurons. *Nature* **388**, 182-6.

Chalfie M, Sulston JE, White JG, Southgate E, Thomson JN, Brenner S. (1985) The neural circuit for touch sensitivity in Caenorhabditis elegans. *Journal of Neuroscience.* **5**, 956-964.

Chaudhari, N et al (2000) A metabotropic glutamate receptor variant functions as a taste receptor. *Nature Neuroscience.* **3**, 113-119.

Chiamulera, C et al (2001) Reinforcing and locomotor stimulant effects of cocaine are absent in mGluR5 null mutant mice. *Nature Neuroscience.* **4**, 873-4.

Chittajallu, R et al (1996) Regulation of glutamate release by presynaptic kainate receptors in the hippocampus. *Nature.* **379**, 78-81.

Cheng, J. et al (1998) Homer binds a novel proline rich motif and links group I metabotropic glutamate receptors with IP3 receptors. *Neuron.* **21**, 717-726.

Chien, CT et al (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proceeding of the National Academy of Science USA.* **88**, 9578-9582.

Cho, K. Bashir, Z. (2002) Cooperation between mgl receptors: a depressing mechanism. *Trends in Neuroscience.* **25**, 405-411

Christensen, JK. Et al (2004) A mosaic of functional kainate receptors in hippocampal interneurons. *Journal of Neuroscience.* **24**, 8986-8993.

Chung, J-J. Shikano, S. Hanyu, Y. Li, M. (2002) Functional diversity of protein C-terminin: more than zipcoding? *TRENDS in Cell Biology.* **12**, 146-150

Ciruela F, Robbins MJ, Willis AC, McIlhinney RA. (1999) Interactions of the C terminus of metabotropic glutamate receptor type 1alpha with rat brain proteins: evidence for a direct interaction with tubulin.. *The Journal of Neurochemistry.* **72**, 346-354.

Ciruela, F. et al (2000) Homer-1c/vest-1L modulates the cell surface targetting metabotropic glutamate receptor type 1□: Evidence for an anchoring function. *Molecular and Cellular Neuroscience.* **15**, 36-50.

Ciruela, F. McIlhinney RA. (2001) Metabotropic glutamate receptor type I alpha and tubulin assemble into dynamic interacting complexes. *Journal of Neurochemistry.* **76**. 750-757.

Clapham, DE and Neer, EJ (1997) G-protein beta, gamma subunits. *Annu. Rev. Pharmacol. Toxicol.* **37**, 167-203.

Collingridge GL, Bliss TV. (1995) Memories of NMDA receptors and LTP. *Trends Neurosci.* **18**, 54-6.

Conn, P.J. et al (1997) Pharmacology and function of metabotropic glutamate receptors. Annual *Review of Pharmacology and Toxicology.* **37**, 205-237.

Conquet, F et al. (1994) Motor deficit and impairment of synaptic plasiticity in mice lacking mGluR1. *Nature.* **372**, 237-243.

Conrad R, Thomas J, Spieth J, Blumenthal T. (1991) Insertion of part of an intron into the 5' untranslated region of a Caenorhabditis elegans gene converts it into a trans-spliced gene. *Mol Cell Biol.* **11**, 1921-6.

Contractor A, Swanson G, Heinemann SF. (2001). Kainate receptors are involved in short- and long-term plasticity at mossy fiber synapses in the hippocampus. *Neuron.* **29**, 209-16

Contractor A, Sailer AW, Darstein M, Maron C, Xu J, Swanson GT, Heinemann SF. (2003) Loss of kainate receptor-mediated heterosynaptic facilitation of mossy-fiber synapses in KA2-/- mice. *Journal of Neuroscience.* **23**, 422-9.

Cook A, Christopher J. Franks and Lindy Holden-Dye Electrophysiological recordings from the pharynx. (May 17 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.7.1, http://www.wormbook.org.

Craig, A.M. Boudin, H. (2001) Molecular heterogeneity of central synapses: afferent and target regulation. *Nature Neuroscience.* **4**, 569-578

Cully, DF et al (1994) Cloning of an avermectin-sensitive glutmate gated chloride channel from *Caenorhabditis elegans. Nature.* **371**, 707-710.

Davies AG, Pierce-Shimomura JT, Kim H, VanHoven MK, Thiele TR, Bonci A, Bargmann CI, McIntire SL. (2003) A central role of the BK potassium channel in behavioral responses to ethanol in C. elegans. *Cell.* **115**(6):655-66.

Davies AG, McIntire SL. (2004) Using C. elegans to screen for targets of ethanol and behavior-altering drugs. *Biol Proced Online.* **6**:113-119.

Davis MW, Fleischhauer R, Dent JA, Joho RH, Avery L. (1999) A mutation in the C. elegans EXP-2 potassium channel that alters feeding behavior. *Science.* **286**(5449):2501-4

Davis RE. (1996) Spliced leader RNA trans-splicing in metazoa. *Parasitol Today.* **12**(1):33-40

De Blassi, A. Conn, P.J. Pin, J-P. Nicoletti, F. (2001) Molecular determinants of metabotropic glutamate receptor signaling. *Trends in Pharmacological Sciences.* **22**, 114-120

de Bono M, Bargmann CI (1998) Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. *Cell.* **94**, 679-689.

de Bono. M (2003) Molecular approaches to Aggregation Behaviour and Social Attachment. *Journal of Neurobiology.* **54**, 78-92.

Dempsey CM, Mackenzie SM, Gargus A, Blanco G, Sze JY (2005) Serotonin (5HT), fluoxetine, imipramine and dopamine target distinct 5HT receptor signaling to modulate Caenorhabditis elegans egg-laying behavior. *Genetics.* **169**(3):1425-36.

Dent, J.A. Davis, M. Avery, L. (1997) avr-15 encodes a chloride channel subunit that mediates inhibitory glutamatergic neurotransmission and ivermectin sensitivity in Caenorhabditis elegans. *EMBO J.* **16**, 5867-5879

Dent JA, Smith MM, Vassilatis DK, Avery L. (2000)\The genetics of ivermectin resistance in Caenorhabditis elegans. *Proceedings of the National Academy of Science USA.* **97**, 2674-2679.

De Vry J, Horvath E, Schreiber R. (2001) Neuroprotective and behavioral effects of the selective metabotropic glutamate mGlu(1) receptor antagonist BAY 36-7620. *Eur J Pharmacol.* **428**, 203-14.

Dhingra, A et al (2004) A Retinal-Specific Regulator of G-protein Signaling Interacts with Gl lo and Accelerates an Expressed Metabotropic Glutamate Receptor 6 Cascade. *The Journal of Neuroscience.* **24**, 5684-5693.

Dietrich D, Kral T, Clusmann H, Friedl M, Schramm J.(2002) Presynaptic group II metabotropic glutamate receptors reduce stimulated and spontaneous transmitter release in human dentate gyrus. *Neuropharmacology.* **42**, 297-305.

Dotti, CG. Sullivan, CA. Banker, GA. (1988) The establishment of polarity by hippocampal neurons in culture. *Journal of Neuroscience.* **8**. 1454-68.

Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ *Cell.* **85**, 1067-1076.

Driscoll M, Dean E, Reilly E, Bergholz E, Chalfie M. (1989) Genetic and molecular analysis of a Caenorhabditis elegans beta-tubulin that conveys benzimidazole sensitivity. *J Cell Biol.* **109**(6 Pt 1):2993-3003

Driscoll, M. Kaplan, J. (1997) Mechanotransduction. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, Riddle DL, Blumenthal T, Meyer BJ, Priess JR, editors. *C. elegans* II. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, p 645–678.

Duerr JS, Frisby DL, Gaskin J, Duke A, Asermely K, Huddleston D, Eiden LE, Rand JB. (1999) The cat-1 gene of Caenorhabditis elegans encodes a vesicular monoamine transporter required for specific monoamine-dependent behaviors. *J Neurosci.* **19**(1):72-84.

Dupuy D, Li QR, Deplancke B, Boxem M, Hao T, Lamesch P, Sequerra R, Bosak S, Doucette-Stamm L, Hope IA, Hill DE, Walhout AJ, Vidal M. (2004) A first version of the Caenorhabditis elegans Promoterome. *Genome Res.* **14**(10B):2169-75.

Dwyer, ND et al (1998) Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell.* **93**, 455-66.

Ebralidze, A. K., Rossi, D. J., Tonegawa, S. and Slater, N.T. (1996) Modification of NMDA receptor channels and synap-tic transmission by targeted disruption of the NR2C gene. *Journal of Neuroscience.* **16**, 5014-5025.

Edgley ML, Riddle DL. (2001) LG II balancer chromosomes in Caenorhabditis elegans: mT1(II;III) and the mIn1 set of dominantly and recessively marked inversions. *Mol Genet Genomics.* **266**(3):385-95

Egebjerg, J. and Heinemann, S. F. (1993) Ca2+ permeability of unedited and edited versions of the kainate selective glutamate receptor GluR6. *Proc. natn. Acad. Sci. U. S. A.* **90**, 755-759.

Ehlers MD. (1999) Synapse structure: glutamate receptors connected by the shanks. *Curr Biol.* **9**(22):R848-50.

Ehrhard KN, Jacoby JJ, Fu XY, Jahn R, Dohlman HG. (2000) Use of G-protein fusions to monitor integral membrane protein-protein interactions in yeast. *Nat Biotechnol.* **18**(10):1075-9.

El Far O, Airas J, Wischmeyer E, Nehring RB, Karschin A, Betz H. (2000) Interaction of the C-terminal tail region of the metabotropic glutamate receptor 7 with the protein kinase C substrate PICK1 *The European Journal of Neuroscience.* **12**, 1-9.

El Far O, Bofill-Cardona E, Airas JM, O'Connor V, Boehm S, Freissmuth M, Nanoff C, Betz H. (2001) Mapping of calmodulin and Gbetagamma binding domains within the C-terminal region of the metabotropic glutamate receptor 7A. *The Journal of Biological Chemistry.* **276**. 30662-30669.

EL Far, O. Betz, H. (2002) G-protein-coupled receptors for neurotransmitter amino acids: C-terminal tails, crowded signalosomes. *Biochem J.* **365**, 329-336.

Enz R. (2002) The actin-binding protein Filamin-A interacts with the metabotropic glutamate receptor type 7. *FEBS Letters.* **514**, 184-188.

Enz, R et al. (2003) Different binding motifs in metabotropic glutamate receptor type 7b for filamin A, protein phosphatase 1C, protein interacting with protein kinase C (PICK) 1 and syntenin allow the formation of multimeric protein complexes *Biocehmical Journal.* **372**.

Eroglu, C. et al. (2003) Glutamate binding affinity of drosophila metabotropic glutamate receptor is modulated by association with lipid rafts. *PNAS.* **100**, 10219-10224.

Estojak J, Brent R, Golemis EA. (1995) Correlation of two-hybrid affinity data with in vitro measurements. *Mol Cell Biol.* **15**(10):5820-9.

Fagni L, Ango F, Perroy J, Bockaert J. (2004) Identification and functional roles of metabotropic glutamate receptor-interacting proteins. *Cell and Developmental Biology.* **15**, 289-298.

Farr CD, Gafken PR, Norbeck AD, Doneanu CE, Stapels MD, Barofsky DF, Minami M, Saugstad JA. (2004) Proteomic analysis of native metabotropic glutamate receptor 5 protein complexes reveals novel molecular constituents. *J Neurochem.* **91**(2):438-50

Farrens, DL. Et al (1996) Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science.* **274**, 768-770.

Fehr, C et al (2002) Congenic Mapping of Alcohol and Pentobarbital Withdrawal Liability Loci to a <1 Centimorgan Interval of Murine Chromosome 4: Identification of *Mpdz* as a Candidate Gene. *The Journal of Neuroscience.* **22**, 3730-3738.

Feng, W et al. (2003) Tandem PDZ repeats in glutamate binding proteins have a novel mode of PDZ domain-mediated target binding. *Nature Structural Biology.* **10**, 972-978.

Ferguson, EL and Horvitz HR (1985) Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode Caenorhabditis elegans *Genetics.* **110**, 17-72.

Fire A, Seydoux GC, Xu SQ. (1995). EFFICIENT GENE-EXPRESSION: REQUIRES PUNCTUATION!. International C. elegans Meeting Abstract. *Wormbase* **WBPaper00021766**

Flajolet, M et al (2003) Protein phosphatase 2C binds selectively to and dephosphorylates metabotropic glutamate receptor 3. *Proceedings of the National Academy of Science USA.* **100**, 16006-16011.

Fleming, JT et al (1997) *Caenorhabditis elegans* Levamisole Resistance Genes *lev-1*, *unc-29*, and *unc-38* Encode Functional Nicotinic Acetylcholine Receptor Subunits. *The Journal of Neuroscience.* **17**, 5843–5857.

Flor PJ, Lukic S, Ruegg D, Leonhardt T, Knopfel T, Kuhn R. (1995) Molecular cloning, functional expression and pharmacological characterization of the human metabotropic glutamate receptor type 4. *Neuropharmacology.* **34**(2):149-55.

Flor PJ, Gomeza J, Tones MA, Kuhn R, Pin JP, Knopfel T. (1996) The C-terminal domain of the mGluR1 metabotropic glutamate receptor affects sensitivity to agonists. *Journal of Neurochemistry.* **67**, 58-63.

Forrest, D., Yuzaki, M., Soares, H. D., Ng, L., Luk, D. C., Sheng, M., Stewart, C. L., Morgan, J. I., Connor, J. A. and Curran, T. (1994) Targeted disruption of NMDA receptor 1 gene abolishes NMDA response and results in neonatal death. *Neuron.* **13**, 325-338.

Foster, A. C., Mena, E. E., Monaghan, D. T. and Cotman, C.W. (1981) Synaptic localization of kainic acid binding sites. *Nature.* **289**, 73-75.

Francesconi, A. Duvoisin, RM. (1998) Role of the second and third intracellular loops of metabotropic glutamate receptors in mediating dual signal transduction and activation. *Journal of Biological Chemistry.* **273**. 5615-5624.

Francesconi, A. Duvoisin, R.M. (2000) Opposing effects of protein kinase C and protein kinase A on metabotropic glutamate receptor signalling: selective desensitization of the inositol triphosphate/Ca^{2+} pathway by phosphorylation of the receptor G-protein coupling domain. *Proc Natl Acad Sci USA.* **97**. 6185-6190.

Francesconi, A et al (2002) Alternative splicing unmasks dendritic and axonal targeting signals in metabotropic glutamate receptor 1. *Journal of Neuroscience.* **22**, 2196-2205.

Francis, M. (2003) Bridging the gap between genes and behaviour: recent advances in the electrophysiological analysis of neural function in Caenorhabditis elegans. *Trends in Neurosciences.* **26**, 90-99.

Franks CJ, Pemberton D, Vinogradova I, Cook A, Walker RJ, Holden-Dye L. (2002) Ionic basis of the resting membrane potential and action potential in the pharyngeal muscle of Caenorhabditis elegans. *J Neurophysiol.* **87**(2):954-61

Franks CJ, Holden-Dye L, Bull K, Luedtke S, Walker RJ. (2006) Anatomy, physiology and pharmacology of Caenorhabditis elegans pharynx: a model to define gene function in a simple neural system. *Invert Neurosci.* **6**(3):105-22.

Freienkamp, H (2002) Organisation of G-protein-coupled receptor signalling complexes by scaffolding proteins. *Current Opinion in Pharmacology.* **2**, 581-586

Fukada, Y., Takao, T., Ohguro, H., Yoshizawa, T., Akino, T., et al. (1990) Farnesylated gamma-subunit of photoreceptor G protein indispensable for GTP-binding. *Nature* **346**:658-60

Funada M, Yasuo S, Yoshimura T, Ebihara S, Sasagawa H, Kitagawa Y, Kadowaki T. (2004) Characterization of the two distinct subtypes of metabotropic glutamate receptors from honeybee, Apis mellifera. *Neurosci Lett.* **359**(3):190-4

Furukawa H, Gouaux E. (2003) Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligandbinding core. *EMBO J.* **22**:2873–2885.

Galvez T, Prezeau L, Milioti G, Franek M, Joly C, Froestl W, Bettler B, Bertrand HO, Blahos J, Pin JP. (2000)Mapping the agonist-binding site of GABAB type 1 subunit sheds light on the activation process of GABAB receptors *The Journal of Biological Chemistry.* **275**, 41166-71174.

Galvez T, Duthey B, Kniazeff J, Blahos J, Rovelli G, Bettler B, Prezeau L, Pin JP. (2001) Allosteric interactions between GB1 and GB2 subunits are required for optimal GABA(B) receptor function. *EMBO Journal.* **20**, 2152-2159.

Garner, C. Nash, J. (2001) Chemical Synapses. In: *Nature Encyclopedia of Life Sciences.* London: Nature Publishing Group. http://www.els.net/ [doi:10.1038/npg.els.0000037]

Gereau, R.W. 4th, Heinemann SF (1998) Role of protein kinase C phosphorylation in rapid desensitization of metabotropic glutamate receptor 5. *Neuron.* **20**(1):143-51.

Gilman, AG (1987) G proteins: transducers of receptor-generated signals. *Annu. Rev. Biochem.* **56**, 615-649.

Gomex, M. et al (2003) Signalling via the neuronal calcium sensor ncs-1 regulates associative learning and memory in *C.elegans. Neuron.* **30**, 241-248.

Gomeza, J. et al (1996) The second intracellular loop of metabotropic glutamate receptor 1 cooperates with other intracellular domains to control coupling to G-proteins. *The Journal of Biological Chemistry.* **271**, 2199-2205.

Gonzalez MI, Robinson MB. (2004) Neurotransmitter transporters: why dance with so many partners? Curr Opin Pharmacol. 4(1):30-5.

Gower, NJ et al (2001) Dissection of the promoter region of the inositol 1,4,5-trisphosphate receptor gene, itr-1, in C. elegans: a molecular basis for cell-specific expression of IP3R isoforms. *The Journal of Molecular Biology.* **306**, 145-157.

Grant, S.G.N, Hussi, H. (2001) Proteomics of multiprotein complexes: answering fundamental questions in neuroscience. *TRENDS in Biotechnology.* **19**. S49-S54.

Greenamyre, J.T. Porter R.H. (1994) Anatomy and physiology of glutamate in the CNS. *Neurology.* **44**:S7-13.

Grootjans JJ, Zimmermann P, Reekmans G, Smets A, Degeest G, Durr J, David G. (1997) Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. *Proc Natl Acad Sci U S A.* **94**(25):13683-8

Grunwald, ME et al (2004) Clathrin-mediated endocytosis is required for compensatiry regulation of GLR-1 glutamate receptors after activity blockade. *Proceeding of the National Academy of Science USA.* **101**, 3190-3195.

Gu MX, York JD, Warshawsky I, Majerus PW. (1991) Identification, cloning, and expression of a cytosolic megakaryocyte protein-tyrosine-phosphatase with sequence homology to cytoskeletal protein 4.1. *Proc Natl Acad Sci U S A*. **88**(13):5867-71.

Gudermann, T et al (1997) Functional and structural complexity of signal transduction via G-protein coupled receptors. *Annual Rev. Neurosci*. **20**, 399-427.

Gustafson TA, He W, Craparo A, Schaub CD, O'Neill TJ. (1995) Phosphotyrosine-dependent interaction of SHC and insulin receptor substrate 1 with the NPEY motif of the insulin receptor via a novel non-SH2 domain *Molecular Cell Biology*. **15**, 2500-2508.

Hall, R. et al (1998) The β2-adrenergic receptor interacts with the Na+/H+ exchange regulatory factor to control Na+/H+ exchange. *Nature*. **392**, 626-630.

Hallam, SJ et al (2002) SYD-1, a presynaptic protein with PDZ, C2 and rhoGAP-like domains, specifies axon identity in C.elegans. *Nature Neuroscience*. **5**, 1137-46.

Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, Jennings EG, Zeitlinger J, Pokholok DK, Kellis M, Rolfe PA, Takusagawa KT, Lander ES, Gifford DK, Fraenkel E, Young RA. (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature*. **431**(7004):99-104

Harris, B.Z. Lim, W.A. (2001) Mechanism and role of PDZ domains in signalling complex assembly. *Journal of Cell Science*. **114**, 3219-3231.

Harris, B. et al (2002) Co-ordinated folding and association of the Lin-2, -7 (L27) domain. *The Journal of Biological Chemistry*. **277**, 34902-34908.

Hart, AC et al. (1995) Synaptic code for sensory modalities revealed by *C.elegans* GLR-1 glutamate receptor. *Nature*. **378**, 82-85.

Hawasli AH, Saifee O, Liu C, Nonet ML, Crowder CM. (2004) Resistance to volatile anesthetics by mutations enhancing excitatory neurotransmitter release in Caenorhabditis elegans. *Genetics*. **168**(2):831-43

Hedgecock, EM et al (1985) Axonal guidance mutants of Caenorhabditis elegans identified by filling sensory neurons with fluorescein dyes. *Developmental Biology*. **111**, 158-170.

Heidinger V, Manzerra P, Wang XQ, Strasser U, Yu SP, Choi DW, Behrens MM (2002) Metabotropic glutamate receptor 1-induced upregulation of NMDA receptor current: mediation through the Pyk2/Src-family kinase pathway in cortical neurons. *J Neurosci*. **22**(13):5452-61.

Heim, R and Tsien, RY. (1996) Engineering green flourescent protein for improved brightness, longer wavelengths and flourescence resonance energy transfer. *Current Biology*. **6**, 178-182.

Herb, A., Burnashev, N., Werner, P, Sakmann, B., Wisden, W. and Seeburg, P. H. (1992) The KA-2 subunit of excitatory amino acid receptors shows widespread expression in brain and forms ion channels with distantly related subunits. *Neuron* **8**, 775-785.

Hermans, H. Saunders, R. Selkirk, JV. et al. (2000) Complex involvement of pertusis toxin-sensitive G-proteins in the regulation of type I alpha metabotropic glutamate receptor signalling in baby hamster kidney cells. *Molecular Pharmacology*. **58**. 352-360.

Hermans, E. Challiss RA (2001) Structural signalling and regulatory properties of the group I metabotropic glutamate receptors. *Biochemistry*. **359**, 465-484.

Heuss, T et al. (2000) *Neuroscience*. **23**, 469-475.

Heyi, L. et al (1997) Identification of chemical synapses in the pharynx of *C.elegans*. *Neurobiology*. **94**, 5912-5916.

Heynen AJ, Quinlan EM, Bae DC, Bear MF. (2000) Bidirectional, activity-dependent regulation of glutamate receptors in the adult hippocampus in vivo. *Neuron*. **28**, 527-36.

Hill RJ, Sternberg PW. (1992) The gene lin-3 encodes an inductive signal for vulval development in C. elegans. *Nature*. **358**, 470-476.

Hirbec H, Perestenko O, Nishimune A, Meyer G, Nakanishi S, Henley JM, Dev KK. (2002) The PDZ proteins PICK1, GRIP, and syntenin bind multiple glutamate receptor subtypes. Analysis of PDZ binding motifs. *The Journal of Biological Chemistry*. **277**, 15221-4.

Hironaka K, Umemori H, Tezuka T, Mishina M, Yamamoto T. (2000) The protein-tyrosine phosphatase PTPMEG interacts with glutamate receptor delta 2 and epsilon subunits. *J Biol Chem*. **275**(21):16167-73.

Hobert, O (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. *Biotechniques*. **32**(4):728-30

Hobert, O. (2003). Behavioural pasticity in *C.elegans*: Paradigms, circuits and genes. *Journal of Neurobiology*. **54**, 203-233

Hobson, RJ et al (2003) SER-7b, a constitutively active Galphas coupled 5-HT7-like receptor expressed in the Caenorhabditis elegans M4 pharyngeal motorneuron. *Journal of Neurochemistry*. **87**, 22-29.

Hodgkin, A.L., Huxley, A.F. (1952) Propagation of electrical signals along giant nerve fibers. *Proc R Soc Lond B Biol Sci*. **140**(899):177-83.

Hollmann M, Maron C, Heinemann S. (1994) N-glycosylation site tagging suggests a three transmembrane domain topology for the glutamate receptor GluR1. *Neuron*. **3**. 1331-43.

Homayoun H, Stefani MR, Adams BW, Tamagan GD, Moghaddam B. (2004) Functional Interaction Between NMDA and mGlu5 Receptors: Effects on Working Memory, Instrumental Learning, Motor Behaviors, and Dopamine Release. *Neuropsychopharmacology*. **29**(7):1259-69.

Hope IA. Related Articles, Links (1991) 'Promoter trapping' in Caenorhabditis elegans. *Development*. **113**(2):399-408

Hope IA, Arnold JM, McCarroll D, Jun G, Krupa AP, Herbert R. (1998) Promoter trapping identifies real genes in C. elegans. *Mol Gen Genet*. **260**(2-3):300-8

Horoszok, L et al. (2001) GLC-3: a novel fipronil and BIDN-sensitive, but picrotoxinin-insensitive, L-glutamate-gated chloride channel subunit from Caenorhabditis elegans. *British Journal of Pharmacology*. **132**. 1247-1254.

Hoskins R, Hajnal AF, Harp SA, Kim SK (1996) The C. elegans vulval induction gene lin-2 encodes a member of the MAGUK family of cell junction proteins. *Development*. **122**, 97-111.

Huang XY, Hirsh D. (1989) A second trans-spliced RNA leader sequence in the nematode Caenorhabditis elegans. *Proc Natl Acad Sci U S A*. **86**, 8640-4.

Huber, A (2001) Scaffolding protein organize multimolecular protein complexes for sensory signal transduction. *European Journal of Neuroscience*. **14**, 769-776.

Hung, A.Y., Sheng, M (2002) PDZ Domains: Structural Modules for Protein Complex Assembly. *The Journal of Biological Chemistry*. **277**, 5699-5702.

Hung TJ, Kemphues KJ Hung TJ, Kemphues KJ (1999) PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in Caenorhabditis elegans embryos. *Development.* **126**(1):127-35.

Husi H, Ward MA, Choudhary JS, Blackstock WP, Grant SG. (2000) Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. *Nat Neurosci.* 3(7):661-9

Hwang SY, Wei J, Westhoff JH, Duncan RS, Ozawa F, Volpe P, Inokuchi K, Koulen P. (2003) Differential functional interaction of two Vesl/Homer protein isoforms with ryanodine receptor type 1: a novel mechanism for control of intracellular calcium signaling. *Cell Calcium.* **34**, 177-84.

Heynen, AJ et al (2000) Bidirectional, activity-dependent regulation of glutamate receptors in the adult hippocampus in vivo. *Neuron.* **28**, 527-536.

Ichise T, Kano M, Hashimoto K, Yanagihara D, Nakao K, Shigemoto R, Katsuki M, Aiba A. (2000)mGluR1 in cerebellar Purkinje cells essential for long-term depression, synapse elimination, and motor coordination. *Science.* **288**, 1832-1835.

Ikeda K., Nagasawa M., Mori H., Araki K., Sakimura K., Watanabe M., Inoue Y. and Mishina M. (1992) Cloning and expression of the epsilon 4 subunit of the NMDA receptor channel. *FEBS Lett.* **313**, 3438.

Ikeda, K., Araki, K., Takayama, C., Inoue, Y., Yagi, T., Aizawa,S. and Mishina, M. (1995) Reduced spontaneous activity of mice defective in the E4 subunit of the NMDA receptor channel. *Molecular Brain Research.* **33**, 61-71.

Im YJ, Lee JH, Park SH, Park SJ, Rho SH, Kang GB, Kim E, Eom SH. (2003) Crystal structure of the Shank PDZ-ligand complex reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. *J Biol Chem.* **278**, 48099-104.

Ireland,D.R.; Guevremont,D.; Williams,J.M.; Abraham,W.C. (2004) Metabotropic glutamate receptor-mediated depression of the slow afterhyperpolarization is gated by tyrosine phosphatases in hippocampal CA1 pyramidal neurons. *Journal of Neurophysiology.* **92**, 2811-2819.

Ishihara, T. Katsura, I (1996) Metabotropic glutamate receptor in C.elegans. *Worm Breeder's Gazette* **14**(3): 40

Ishihara, T. Francesconi, A. Duvoisin, R. Katsura, I. (1997) Metabotropic glutamate receptors in C.elegans. *International Worm Meeting Abstract* **265**

Ishihara, T. et al (2002) HEN-1, a secretory protein with an LDAL receptor mkotif, regulates sensory integration and learning in Caenorhabditis elegans. *Cell.* **109**, 639-649.

Ishii M, Kurachi Y. (2003) Physiological actions of regulators of G-protein signaling (RGS) proteins. *Life Sci.* **74**(2-3):163-71.

Ishikawa K, Nash SR, Nishimune A, Neki A, Kaneko S, Nakanishi S (1999) Competitive interaction of seven in absentia homolog-1A and Ca2+/calmodulin with the cytoplasmic tail of group 1 metabotropic glutamate receptors. *Genes Cells.* **4**, 381-90.

Jansen, G et al (1999) The complete family of genes encoding G-proteins of *Caenorhabditis elegans. Nature Genetics.* **21**, 414-419.

Jee C, Lee J, Lee JI, Lee WH, Park BJ, Yu JR, Park E, Kim E, Ahnn J. (2004) SHN-1, a Shank homologue in C. elegans, affects defecation rhythm via the inositol-1,4,5-trisphosphate receptor. *FEBS Lett.* **561**(1-3):29-36.

Jingami, H. et al (2003) Structure of the metabotropic glutamate receptor. *Current Opinion in Neurobiology.* **13**, 271-278.

Jonas, P., Racca, C., Sakmann, B., Seeburg, P. and Monyer, H. (1994) Differences in Ca2+ permeability of AMPA-type glutamate receptor channels in neocortical neurones caused by differential GluR-B subunit expression. *Neuron.* **12**, 1281-1289.

Jonas, P. and Burnashev, N. (1995) Molecular mechanisms controlling calcium entry through AMPA-type glutamate receptor channels. *Neuron.* **15**, 987-990.

Kadotani, H., Hirano, T., Masugi, M., Nakamura, K., Nakao, K., Katsuki, M. and Nakanishi, S. (1996) Motor discoordination results from combined gene disruption of the NMDA receptor NR2A and NR2C subunit, but not from single disruption of the NR2A or NR2C subunit. *Journal of Neuroscience.* **16**, 7859-7867.

Kaech SM, Whitfield CW, Kim SK. (1998) The LIN-2/LIN-7/LIN-10 complex mediates basolateral membrane localization of the C. elegans EGF receptor LET-23 in vulval epithelial cells *Cell.* **94**. 761-771.

Kaletta T, Hengartner MO. (2006) Finding function in novel targets: C. elegans as a model organism. *Nat Rev Drug Discov* **5**(5):387-98.

Kamath, RS. Fraser, AG. Dong, Y. et al (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature.* **421**. 231-7.

Kammermeier, PJ et al (1999) Expression of RGS2 Alters the Coupling of Metabotropic Glutamate receptor 1a to M-Type K^+ and N-Type Ca^{2+} Channels. *Neuron.* **22**, 819-829.

Kaplan, J. Horvitz, H. (1993) A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. *Proceedings of the National Academy of Science USA.* **90**, 2227-2231.

Kato, A et al. (1997) vesl, a gene encoding VASP/Ena family related protein, is upregulated during seizure, long-term potentiation and synaptogenesis. *FEBS Letters.* **412**, 183-189.

Kato A, Ozawa F, Saitoh Y, Fukazawa Y, Sugiyama H, Inokuchi K (1998) Novel members of the Vesl/Homer family of PDZ proteins that bind metabotropic glutamate receptors. *The Journal of Biological Chemistry.* **273**, 23969-23975.

Katz, B. Miledi, R. (1965) Release of acetylcholine from a nerve terminal by electric pulses of variable strength and duration. *Nature.* **207**. 1097-8.

Kaziro Y, Itoh H, Kozasa T, Nakafuku M, Satoh T. (1991) Structure and function of signal-transducing GTP-binding proteins. *Annu Rev Biochem.* **60**:349-400.

Keinanen, K., Wisden, W., Sommer, B., Werner, P., Herb, A.,Verdoorn, T. A., Sakmann, B. and Seeburg, P. H. (1990) A family of AMPA-selective glutamate receptors. *Science.* **249**, 556-560.

Kenny, PJ et al (2003) Group II metabotropic and alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA)/kainate glutamate receptors regulate the deficit in brain reward function associated with nicotine withdrawal in rats. *Journal of Pharmacology and Experimental Therapeutics.* **306**, 1068-1076.

Kenny, PJ et al (2004) The ups and downs of addiction: role of metabotropic glutamate receptors. *Trends in Pharmacological Science.* **25**, 265-272.

Kew, J. et al (2001) Activity dependent presynaptic autoinhibition by group II metabotropic glutamate receptors at the perforant path inputs to the dentate gyrus and CA1. *Neuropharmacology.* **40**, 20-27.

Kim CH, Takamiya K, Petralia RS, Sattler R, Yu S, Zhou W, Kalb R, Wenthold R, Huganir R. (2005) Persistent hippocampal CA1 LTP in mice lacking the C-terminal PDZ ligand of GluR1. *Nat Neurosci.* **8**(8):985-7.

Kim, SJ (2003) Activation of the TRPC1 cation channel by metabotropic glutamate receptor mGluR1. *Nature.* **426**, 285-291.

Kim, SK . (1997) Polarized signaling: basolateral receptor localization in epithelial cells by PDZ-containing proteins. *Current Opinion in Cell Biology.* **9**, 853-859.

Kitano J, Kimura K, Yamazaki Y, Soda T, Shigemoto R, Nakajima Y, Nakanishi S. (2002) Tamalin, a PDZ domain-containing protein, links a protein complex formation of group 1 metabotropic glutamate receptors and the guanine nucleotide exchange factor cytohesins. *The Journal of Neuroscience.* **22**, 1280-1289.

Kitano J, Nishida M, Itsukaichi Y, Minami I, Ogawa M, Hirano T, Mori Y, Nakanishi S. (2003) Direct interaction and functional coupling between metabotropic glutamate receptor subtype 1 and voltage-sensitive Cav2.1 Ca2+ channel. *J Biol Chem.* **278**, 25101-8.

Kniazeff J, Bessis AS, Maurel D, Ansanay H, Prezeau L, Pin JP. (2004) Closed state of both binding domains of homodimeric mGlu receptors is required for full activity. *Nat Struct Mol Biol.* **11**(8):706-13

Knoflach F, Kemp JA. (1998) Metabotropic glutamate group II receptors activate a G protein-coupled inwardly rectifying K+ current in neurones of the rat cerebellum. *Journal of Physiology.* **509**, 347-354.

Kornau, HC et al (1995) Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science.* **269**, 1737-1740.

Krause M, Hirsh D. (1987) A trans-spliced leader sequence on actin mRNA in C. elegans. *Cell.* **49**(6):753-61

Krupnick JG, Benovic JL. (1998) The role of receptor kinases and arrestins in G protein-coupled receptor regulation. *Annu Rev Pharmacol Toxicol.* **38**, 289-319.

Kunishima N, Shimada Y, Tsuji Y, Sato T, Yamamoto M, Kumasaka T, Nakanishi S, Jingami H, Morikawa K. (2000) Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. *Nature.* **407**, 971-7.

Kunishima, N. et al (2002) Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. *Nature.* **407**, 971-977.

Kutsuwada, T., Kashiwabuchi, N., Mori, H., Sakimura, K.,Kushiya, E., Araki, K., Meguro, H., Masaki, H., Kumanishi,T., Arakawa, M. and Mishina, M. (1992) Molecular diversity of the NMDA receptor channel. *Nature.* **358**, 36-41.

Kutsuwada, T., Sakimura, K., Manabe, T., Takayama, C., Katakura, N., Kushiya, E., Natsume, R., Watanabe, M., Inoue, Y., Yagi, T., Aizawa, S., Arakawa, M., Takahashi, T., Nakamura, Y., Mori, H. and Mishina, M. (1996) Impairment of suckling response, trigeminal neuronal pattern formation, and hippocampal LTD in NMDA receptor E2 subunit mutant mice. *Neuron.* **16**, 333-344.

Lackner MR, Kindt RM, Carroll PM, Brown K, Cancilla MR, Chen C, de Silva H, Franke Y, Guan B, Heuer T, Hung T, Keegan K, Lee JM, Manne V, O'Brien C, Parry D, Perez-Villar JJ, Reddy RK, Xiao H, Zhan H, Cockett M, Plowman G, Fitzgerald K, Costa M, Ross-Macdonald P. (2005) Chemical genetics identifies Rab geranylgeranyl transferase as an apoptotic target of farnesyl transferase inhibitors. *Cancer Cell.* **7**(4):325-36

Larkman AU. (1991) Dendritic morphology of pyramidal neurones of the visual cortex of the rat: I. Branching patterns. *J Comp Neurol.* **306**(2):307-19.

Leanna CA, Hannink M. (1996) The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. *Nucleic Acids Res.* **24**(17):3341-7

Lee RY, Sawin ER, Chalfie M, Horvitz HR, Avery L. (1999) EAT-4, a homolog of a mammalian sodium-dependent inorganic phosphate cotransporter, is necessary for glutamatergic neurotransmission in caenorhabditis elegans. *The Journal of Neuroscience.* 159-167.

Lee TI, Young RA. (2000) Transcription of eukaryotic protein-coding genes. *Annu Rev Genet.* 34:77-137

Li Lin, Vera M Hapiak, Hong Xiao, Katherine Smith, Richard W Komuniecki. (2004) The Caenorhabditis elegans multiple PDZ domain containing protein, MPZ-1 is involved in the assembly/localization of signaling complexes that include GPCRs and DAF-18 (PTEN). West Coast Worm Meeting 135 **WBPaper00023955**

Lorez, M et al (2003) Group III metabotropic glutamate receptors as autoreceptors in the cerebellar cortex. *British Journal of Pharmacology.* **138**. 614-625.

Linden, A-M et al (2002) Increased anxiety related behavior in mice deficient for metabotropic glutamate 8 (mGlu8) receptor. *Neuropharmacology.* **43**, 251-259.

Lominac KD, Kapasova Z, Hannun RA, Patterson C, Middaugh LD, Szumlinski KK. (2006) Behavioral and neurochemical interactions between Group 1 mGluR antagonists and ethanol: Potential insight into their anti-addictive properties. Drug Alcohol Depend. May 10

Lu, Y-M et al. (1997) *The Journal of Neuroscience.* **17**, 5196-5205.

Lujan, R. et al (1996) Perisynaptic location of metabotropic glutamate reeptors mGluR1 and mGluR5 on dendrites and dendritic spines in the rat hippocampus. *European Journal of Neuroscience.* **8**, 1488-1500.

Lujan R, Roberts JD, Shigemoto R, Ohishi H, Somogyi P. (1997) Differential plasma membrane distribution of metabotropic glutamate receptors mGluR1 alpha, mGluR2 and mGluR5, relative to neurotransmitter release sites. *Journal of Chemical Neuroanatomy.* **13**, 219-241.

Mahon, M. et al (2002) Na+/H+ echanger regulatory factor 2 directs parathyroid hormone I receptor signalling. *Nature.* **417**, 858-860.

Malenka, RC. (2003) Synaptic plasticity and AMPA receptor trafficking. *Annals New York Academy of Science.* **1003**. 1-11.

Man, HY et al (2000) Intracellular trafficking of AMPA receptors in synaptic plasticity. *Cellular and Molecular Life Sciences.* **57**, 1526-1534.

Maragos, W. F., Penney, J. B. and Young, A. B. (1988) Anatomic correlation of NMDA and 3H-TCP-labeled receptors in rat brain. *Journal of Neuroscience.* **8**, 49 -501.

Marchand S, Cartaud J. (2002) Targeted trafficking of neurotransmitter receptors to synaptic sites. *Mol Neurobiol.* **26**(1):117-35.

Margeta-Mitrovic M, Jan YN, Jan LY. (2000) A trafficking checkpoint controls GABA(B) receptor heterodimerization. *Neuron.* **27**, 97-106.

Maricq, AV. Et al (1995) Mechanosensory signalling in C. elegans mediated by the GLR-1 glutamate receptor. *Nature.* **378**. 78-81.

Maroney, PA et al (1995) Most mRNAs in the nematode Ascaris lumbricoides are trans-spliced: a role for spliced leader addition in translational efficiency. *RNA.* **1**, 714-723.

Martin S, Henley JM. (2004) Activity-dependent endocytic sorting of kainate receptors to recycling or degradation pathways. *EMBO J.* **23**(24):4749-59.

Martin-Ruiz, M. et al (2001) Control of serotonergic function in medial prefrontal cortex by serotonin-2A receptors through a glutamate dependent mechanism. *The Journal of Neuroscience.* **21**, 9856-9866.

Mary S, Gomeza J, Prezeau L, Bockaert J, Pin JP. (1998) A cluster of basic residues in the carboxyl-terminal tail of the short metabotropic glutamate receptor 1 variants impairs their coupling to phospholipase C. *The Journal of Biological Chemistry.* **273**, 425-32.

Masu, M et al. (1995) Synaptic deficit of the ON response in Visual Transmission by Targeted Disruption of the mGluR6 Gene. *Cell.* **80**, 757-765.

Masugi, M et al (1999) Metabotropic glutamate receptor subtype 7 ablation causes deficit in fear response and conditioned taste aversion. *The Journal of Neuroscience.* **19**, 955-963.

Matsui T, Kita H. (2003) Activation of group III metabotropic glutamate receptors presynaptically reduces both GABAergic and glutamatergic transmission in the rat globus pallidus. *Neuroscience.* **122**, 727-732.

Mayer, M. L., Westbrook, G. L and Guthrie, P. B. (1984) Voltage-dependent block by Mg2+ of NMDA responses in spinal cord neurones. *Nature.* **309**, 261-263.

Mayer, M. L. and Westbrook, G. L. (1987) Permeation and block of N-methyl-D-aspartic acid receptor channels by divalent cations in mouse cultured central neurones. *Journal of Physiology.* **394**, 501-527.

McClintock TS, Sammeta N. (2003) Trafficking prerogatives of olfactory receptors. *Neuroreport.* **14**(12):1547-52.

McCool, BA et al (1998) Rat group I metabotropic glutamate receptors inhibit neuronal Ca2+ channels via multiple signal transduction pathways in HEK 293 cells. *Journal of Neurophysiology.* **79**, 379-391.

McKay JP, Raizen DM, Gottschalk A, Schafer WR, Avery L (2004) eat-2 and eat-18 are required for nicotinic neurotransmission in the Caenorhabditis elegans pharynx. *Genetics.* **166**(1):161-9.

McGee, AW et al (2003) Assembly and plasiticity of the glutamatergic postsynaptic specialization. *Current Opinion in Neurobiology.* **13**, 111-118.

McIlhinney RA, Molnar E, Atack JR, Whiting PJ. (1996) Cell surface expression of the human N-methyl-D-aspartate receptor subunit 1a requires the co-expression of the NR2A subunit in transfected cells. *Neuroscience* **70**, 989–97.

Meldrum, BS. et al (1990) Excitatory amino acid neurotoxicity and neurodegenerative disease. *Trends in Pharmacological Science.* **11**(9). 379-87.

Meldrum, BS. (2000) Glutamate as a neurotransmitter in the brain: review of physiology and pathology. *Journal of Nutrition.* **130**(4S Suppl). 1007S-15S.

Mellem, J. et al (2002) Decoding of polymodal sensory stimuli by postsynaptic glutamate receptors in C.elegans. *Neuron.* **36**, 933-944.

Mendelsohn, AR and Brent, R. (1994) Biotechnology applications of interaction traps/two-hybrid systems. *Current Opinion in Biotechnology.* **5**, 482-486.

Mertens I, Vandingenen A, Clynen E, Nachman RJ, De Loof A, Schoofs L. (2005) Characterization of an RFamide-related peptide orphan GPCR in C. elegans. *Ann N Y Acad Sci.* **1040**:410-2.

Mertens I, Clinckspoor I, Janssen T, Nachman R, Schoofs L. (2006) FMRFamide related peptide ligands activate the Caenorhabditis elegans orphan GPCR Y59H11AL.1. *Peptides.* **27**(6):1291-6.

Meyer G, Varoqueaux F, Neeb A, Oschlies M, Brose N. (2004) The complexity of PDZ domain-mediated interactions at glutamatergic synapses: a case study on neuroligin. *Neuropharmacology.* **47**(5):724-33.

Miller KG, Alfonso A, Nguyen M, Crowell JA, Johnson CD, Rand JB (1996) A genetic selection for Caenorhabditis elegans synaptic transmission mutants. *Proc Natl Acad Sci USA* **93**: 12593-12598.

Mitchell, SJ et al (2000) Glutamate spillover suppresses inhibition by activating presynaptic mGluRs. *Nature.* **404**, 498-502.

Mitri C, Parmentier ML, Pin JP, Bockaert J, Grau Y. (2004) Divergent evolution in metabotropic glutamate receptors. A new receptor activated by an endogenous ligand different from glutamate in insects. *J Biol Chem.* **279**(10):9313-20.

Monaghan, D. T. and Cotman, C. W. (1982) The distribution of [3H]kainic acid binding sites in rat CNS as determined by auto-radiography. *Brain Research.* **252**, 91-100.

Monaghan DT, Yao D, Cotman CW. (1984) Distribution of [3H]AMPA binding sites in rat brain as determined by quantitative autoradiography. *Brain Res.* **324**(1):160-4

Morrison, GE. Van der Kooy, D. (2001) A mutation in the AMPA-type glutamate receptor, glr-1, blocks olfactory associative and nonassociative learning in Caenorhabditis elegans. *Behavioural Neuroscience.* **115**(3). 640-9.

Mori, H. Mishina, M. (1995) Structure and function of the NMDA receptor channel. *Neuropharmacology.* **34**, 1219-1237.

Moriyoshi K., Masu M., Ishii T., Shigemoto R., Mizuno N. and Nakanishi S. (1991) Molecular cloning and characterization of the rat NMDA receptor. *Nature* **354**: 31-37.

Mostov, KE et al (1995) Regulation of protein traffic in polarized epithelial cells. *Bioessays.* **17**, 129-138.

Mulle C, Sailer A, Swanson GT, Brana C, O'Gorman S, Bettler B, Heinemann SF. (2000) Subunit composition of kainate receptors in hippocampal interneurons. *Neuron.* **28**, 475-84.

Naisbitt S, Kim E, Tu JC, Xiao B, Sala C, Valtschanoff J, Weinberg RJ, Worley PF, Sheng M. (1999) Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin *Neuron.* **23**, 569-582.

Nakajima Y, Yamamoto T, Nakayama T, Nakanishi S. (1999) A relationship between protein kinase C phosphorylation and calmodulin binding to the metabotropic glutamate receptor subtype 7 *The Journal of Biological Chemistry.* **274**, 27573-27577.

Nakanishi S. (1992) Molecular diversity of glutamate receptors and implications for brain function *Science.* **258**. 597-603.

Nakanishi, S. et al (1998) Glutamate receptors: brain function and signal transduction. *Brain Res Brain Res Rev.* **26**(2-3):230-5.

Nance J. (2005) PAR proteins and the establishment of cell polarity during C. elegans development. *Bioessays.* **27**(2):126-35

Nawy S. (1999) The metabotropic receptor mGluR6 may signal through G(o), but not phosphodiesterase, in retinal bipolar cells. *J Neurosci.* **19**(8):2938-44

Niacaris, T et al. (2003) Serotonin regulates repolarization of the *C.elegans* pharyngeal muscle. *The Journal of Experimental Biology.* **206**, 223-231.

Nielsen, H et al (1999) Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Eng.* **12**, 3-9.

Nguyen M, Alfonso A, Johnson CD, Rand JB. (1995) Caenorhabditis elegans mutants resistant to inhibitors of acetylcholinesterase. *Genetics*. **140**(2):527-35

Nilsen, TW et al (1993) TRAN-Splicing of nematode premessenger *RNA*. *Annual Review of Microbiology*. **47**, 413-440.

Nusser, Z et al (1994) Subsynaptic segregation of metabotropic and ionotropic glutamate receptors as revealed by immunogold localization. *Neuroscience*. **61**, 421-427.

O'Connor V, El Far O, Bofill-Cardona E, Nanoff C, Freissmuth M, Karschin A, Airas JM, Betz H, Boehm S. (1999)Calmodulin dependence of presynaptic metabotropic glutamate receptor signaling. *Science*. **286**, 1180-84.

Offermans, S. (2003) G-proteins as transducers in transmembrane signalling. *Progress in Biophysics and Molecular Biology*. **83**, 101-130.

Okamoto N, Hori S, Akazawa C, Hayashi Y, Shigemoto R, Mizuno N, Nakanishi S. (1994) Molecular characterization of a new metabotropic glutamate receptor mGluR7 coupled to inhibitory cyclic AMP signal transduction. *The Journal of Biological Chemistry*. **269**, 1231-1236.

Okamoto, t et al (2000) Ultraviolet absorbance at 260 and 280 nm in RNA measurement is dependent on measurement solution. *International Journal of Molecular Medicine*. **5**, 657-9.

Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. (1993) Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics*. **135**(2):385-404.

Ollis, D. et al (1985) The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. *Nature*. **313**, 763-766.

Osborne MA, Dalton S, Kochan JP. (1995) The yeast tribrid system--genetic detection of trans-phosphorylated ITAM-SH2-interactions. *Biotechnology (N Y)*. **13**(13):1474-8

Ozawa S, Kamiya H, Tsuzuki K (1998) Glutamate receptors in the mammalian central nervous system *Progress in Neurobiology*. **54**. 581-618.

Papaioannou S, Marsden D, Franks CJ, Walker RJ, Holden-Dye L. (2005) Role of a FMRFamide-like family of neuropeptides in the pharyngeal nervous system of Caenorhabditis elegans. *J Neurobiol*. **65**(3):304-19

Palmer CL, Cotton L, Henley JM. (2005) The molecular pharmacology and cell biology of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors. *Pharmacol Rev*. **57**(2):253-77.

Park, Y-S et al (2003) Characterization of GAR-2, a Novel G Protein-Linked Acetylcholine Receptor from *Caenorhabditis elegans*. *Journal of Neurochemistry*. **75**, 1800–1809.

Parker LL, Backstrom JR, Sanders-Bush E, Shieh BH. (2003) Agonist-induced phosphorylation of the serotonin 5-HT2C receptor regulates its interaction with multiple PDZ protein 1 *Journal of Biological Chemistry*. **278**, 21576-21583.

Parmentier, ML et al (1996) Cloning and functional expression of a Drosophila metabotropic glutamate receptor expressed in the embryonic CNS. *The Journal of Neuroscience*. **16**, 6687-6694.

Parmentier ML, Galvez T, Acher F, Peyre B, Pellicciari R, Grau Y, Bockaert J, Pin JP. (2000) Conservation of the ligand recognition site of metabotropic glutamate receptors during evolution. *Neuropharmacology*. **39**(7):1119-31

Peters A, Palay SL. (1996) The morphology of synapses. *J Neurocytol*. **25**(12):687-700.

Pallay, SL. (1958) The morphology of synapses in the central nervous system. *Experimental Cell Research*. **14**. 275-93.