Investigations into the function and regulation of the C-terminal binding protein (CTBP-1) in *C. elegans*

Duygu Yücel

THE UNIVERSITY OF SYDNEY

A thesis submitted for the degree of Doctor of Philosophy,

University of Sydney

School of Molecular Bioscience

University of Sydney

January 2012

Table of contents

List of illustrations

Statement of Originality

The work described in this thesis was performed between March 2008 and January 2012 in the School of Molecular Bioscience at the University of Sydney. The experiments were performed by myself unless stated otherwise. None of this material has been previously presented for the purpose of obtaining any other degree.

Duygu Yücel

Acknowledgments

These 169 pages tell my unwritten story of being a guest in the Land of Oz. Many people made this 3 year and 10 month-long journey an enjoyable one. To start with the journey of discoveries, I would like to express my gratitude to my supervisor Hannah Nicholas for her support and guidance. Hannah provided me with opportunities which are not available to most graduate students, and I am indebted to her. I also thank Hannah for her patience during the preparation of this thesis.

I thank members of Nicholas Lab for being a great company: Sashi, Slavica, Anna, Estelle, Yee Lian, Hong, Aaron Rhodes and Aaron Lun. Special thanks to Anna and Yee Lian for proofreading first two chapters of my thesis, and to Slavica and Estelle for making the office a peaceful environment. I enjoyed sharing the office with them.

I had the privilege to benefit from experience and knowledge of two more labs: Crossley Lab which now is at UNSW, Sydney-AU and Miller Lab at the Vanderbilt University, Nashville-USA. I thank to past and present members of the Crossley Lab: Richard, Ali, Robert, Crisbel, Cassie, Thanh, Vitri, Feyza, Stella, Jon, Laura, Briony and Noelia. I am grateful for support of my co-supervisor Merlin Crossley at most critical times of my PhD. Merlin's optimism has been inspirational for me. I would like to thank Miller Lab members Clay, Rebecca, Rachel, Sarah, Tim, Kathy and Cody for their help and good discussion, and David Miller for his guidance on cell-specific microarray experiments and for sharing his expertise and reagents. I especially thank Clay for helping me with the analysis of the microarray data and Rebecca for finalising the last steps of the microarrays. In addition, I would like to thank Douglas Chappell for his assistance in setting up the program for my next-generation sequencing results, without his help I would not be able to identify the mysterious mutation.

I would like to thank the Sydney University International Scholarship Scheme (UsydIS) for the financial support. I thank Sashi Kant and Dale Hancock for giving me the opportunity to teach for undergraduate labs, and in particular I thank Sashi for her generosity and for sharing her finest selection of wines with me. Throughout 3 years and 10 months, I have been lucky to meet good friends. I especially thank Shajeaa, Luke and Torsten for being supportive and understanding. I have enjoyed debating the politics of the world at nocturnal times with them and saving the world by the time sun rises. I will miss discussing "Why?" and "How?" of everything around us. I also thank my housemates (n=4) for being bearable people to share the house with and also for washing my dishes once in a while whilst I was writing my thesis. I thank friends from Australian Turkish Music Ensemble for being supportive especially when I first moved in Sydney and friends from tango for making me forget about the whole PhD for at least one song length at stressful times.

As I have come from other part of the world (Kayseri, Turkey), I would like to thank my friends from various places in Turkey for being caring and supportive. Lastly, I thank dearest members of my family for their continual support and encouragement and for always being there for me despite the time zone difference. It will never be enough no matter how much I thank to my mother Ayfer, father Haluk and brother Semih. I can not wait to re-unite with my family and friends.

Duygu Yücel

Publications arising from this thesis

Journal articles

Yucel, D., Llamosas E., Crossley, M., Nicholas, H. *C. elegans* CTBP-1 functions in specification of a subset of cholinergic motor neurons. *Manuscript in Preparation*

Conference abstracts

Yucel, D., Llamosas E., McWhirter, R., Miller[III], D., Crossley, M., Nicholas, H. (2011). Deciphering neuronal functions of C-terminal Binding Protein using *C. elegans* as a model. In: *Sydney University Neuroscience and Mental Health Theme, Sydney, Australia*

Yucel, D., McWhirter, R., Miller[III], D., Crossley, M., Nicholas, H. (2011). The role of C-terminal Binding Protein (CtBP) in the gene regulatory network. In: *58th Annual Genetics Society of AustralAsia Conference, Melbourne, Australia.*

Yucel, D., Crossley, M., Nicholas, H. (2010). Insights into C-terminal Binding Protein via a mutagenesis screen in *C. elegans.* In: *OzBio 2010, Melbourne, Australia*

Yucel, D., Crossley, M., Nicholas, H. (2010). Insights into C-terminal Binding Protein via a mutagenesis screen in *C. elegans.* In: *EMBO Conference Series C. elegans Development and Gene Expression, Heidelberg, Germany.*

Yucel, D., Crossley, M., Nicholas, H. (2010). Deciphering functions of C-terminal Binding Protein. In: *The 31st Annual Lorne Genome Conference, Melbourne, Australia.*

Yucel, D., Crossley, M., Nicholas, H. (2010). Insights into C-terminal Binding Protein via a mutagenesis screen in *C. elegans*. In: *16th Cell and Developmental Biology Meeting, Sydney, Australia.*

Yucel, D., Wu, T., Crossley, M., Nicholas, H. (2009). Insights into the functions of C-terminal Binding Protein. In: *17th International C. elegans Meeting, Los Angeles, USA.*

Yucel, D., Sibbritt, T., Bert, S., Crossley, M., Nicholas, H. (2009). Using *C. elegans* to screen for genes that regulate fat. In: *Obesity Forum, Sydney, Australia.*

Yucel, D., Sibbritt, T., Bert, S., Crossley, M., Nicholas, H. (2009). The Mystery of C-terminal Binding Protein: *The C. elegans experience. Australian Worm Meeting, Kiolao, Australia.*

Yucel, D., Wu, T., Nicholas, H. (2009). Analysis of C-terminal Binding Protein expression in *C. elegans*. . In: *The 30th Annual Lorne Genome Conference, Melbourne, Australia.*

Yucel, D., Wu, T., Nicholas, H. (2008). Analysis of C-terminal Binding Protein expression in *C. elegans*. . In: *ComBio 2010, Canberra, Australia.*

Abstract

C-terminal Binding Protein (CtBP) is a transcriptional co-repressor which plays roles in development and apoptosis. There are two highly related CtBP genes in vertebrates (CtBP1 and CtBP2). Knock-out of both *CtBP1* and *CtBP2* results in lethality early in embryogenesis. Recently, we cloned a single CtBP in *Caenorhabditis elegans* (*C. elegans*). Like its mammalian counterparts, the *C. elegans* CtBP, called CTBP-1, functions as a transcriptional co-repressor and docks onto transcription factors containing an amino acid motif of the form PXDLS. The CtBP family of proteins are mechanistically well characterised, however, there is limited information on their *in vivo* functions. In this thesis, we explored how CTBP-1 functions using *C. elegans* as a model. First, we found that CTBP-1 is predominantly expressed in the nervous system. With mutant analysis we identified a role for CTBP-1 in specification of a subset of cholinergic neurons, and our data suggests that this may occur through interaction with neuron-specific transcription factor ZAG-1. To identify genes regulated by CTBP-1 for correct specification of neurons, cell-specific microarray experiments were employed where a comparison of gene expression profiles of all neurons from *ctbp-1(ok498)* mutants and wild type animals identified putative CTBP-1 targets, some of which display neuronal roles.

In addition to targets of CTBP-1 regulation, regulators of CTBP-1 were investigated through a mutagenesis screen which led to the isolation of mutants with changes in CTBP-1 expression. This approach revealed a new regulator of CTBP-1 which was identified by next-generation sequencing. We confirmed the identification of this gene by successfully rescuing a CTBP-1 over-expression phenotype in epidermal cells caused by a nonsense mutation in this novel regulator. Bioinformatic analysis has shown that this novel regulator displays homology to human Ino80d, a subunit of an ATP-dependent chromatin remodelling complex called Ino80.

Taken together, these results indicate a role for CTBP-1 in the nervous system and propose a possible mechanism by which CTBP-1 expression is regulated. Given the evolutionary conservation of CtBP from nematodes to mammals, our results may ultimately facilitate the understanding of CtBP function in neuronal regulation in higher organisms including humans, and identify novel components of the gene regulatory networks in which CtBP is involved.

Abbreviations

AD	activation domain
AGRF	Australian Genome Research Facility
aa	amino acid(s)
ALS	Amyotrophic Lateral Sclerosis
BLAST	Basic Local Alignment Search Tool
bp	base pair
cDNA	complementary DNA
CTBP-1	C-terminal binding protein 1
ChIP	Chromatin immunoprecipitation
Hipk	homeodomain-interacting protein kinase
CtBP	C-terminal binding protein
COE	Collier/Olf1/EBF
CoREST	co-supressor of REST
CNS	central nervous system
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA-binding protein
DIC	differential interference contrast
dNTP	deoxynucleotide triphosphate
E. coli	*Escherichia coli*
EMS	ethyl methanesulfonate
EMT	epithelial to mesenchymal transition
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
GFP	green fluorescent protein

HAT	histone acetyltransferase
HDAC	histone deacetylase
HMT	histone methyltransferase
kb	kilobase pairs (1000 bp)
kDa	kilodalton
KO	Knock-out
LB	Luria-Bertani
MQW	Milli-Q® Water
mRNA	messenger RNA
NAD(H)	nicotinamide adenine dinucleotide
NGM	nematode growth medium
OP50	*Escherichia coli* strain OP50
PAG-3	PAttern of Gene expression abnormal-3
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PI	propidium iodide
PTC	premature termination codon
RNA	ribonucleic acid
RNAi	RNA interference
RT-PCR	Reverse transcriptase-PCR reaction
SD.	standard deviation
SUMO	small ubiquitin-like modifier
SWI/SNF	switching mating type/sucrose non-fermenting
TF	transcription factor
THAP	Thanatos-associated protein

UNC	uncoordinated
ZAG-1	Zinc finger involved in Axon Guidance-1

Note: The abbreviation for *C. elegans* protein names is all in capital letters e.g., a homologue of mammalian CtBP is CTBP-1. Genes that are not characterised are designated with their cosmid names e.g., C34E10.8 and T23E7.2.

CHAPTER 1

INTRODUCTION

1.1 Chromatin and transcription

Throughout development a diverse range of cell types are generated with tight control of gene expression. The regulation of gene expression occurs at multiple levels ranging from transcription of a gene and translation of its encoded protein, to post-translational modifications which tag proteins with signals for subcellular localisation, activation or degradation.

An average eukaryotic cell has approximately 2 meters of DNA packed into a 5 micrometer nucleus. This is achieved by multiple levels of compaction of the DNA. Chromatin, the state in which DNA is organised in the nucleus, is composed of structural subunits called nucleosomes. The nucleosome consists of 146 bp DNA wrapped around an octamer of proteins, called histones (Cairns, 2009). Histones are evolutionarily conserved proteins with a C-terminal domain essential for formation of nucleosomes, and an unstructured N-terminal domain called a histone tail. Chromatin structure is altered by decoration of histone tails with post-translational modifications via histone modifying enzymes such as histone acetyl transferases (HATs) and histone methyl transferases (HMTs) (Kouzarides, 2007). Histone modification enzymes are also part of chromatin modifying complexes such as the ATP-dependent chromatin remodelling complex (ADCR) which alters chromatin structure by mobilisation of nucleosomes or histone subunits (Flaus and Owen-Hughes, 2004; Wu and Grunstein, 2000).

ADCRs regulate chromatin by sliding nucleosomes or replacing core histones with their variants in an ATP-dependent manner. ADCRs can be divided into four families based on the structure of their ATPase domain: SWI/SNF, ISWI, CHD and INO80. These four families are also characterised by additional features. For instance, ATPases of the SWI/SNF class contain

a bromodomain which mediates binding to acetylated histone tails, the ISWI family is characterised by a SANT domain, and CHD family members contain a chromodomain and PHD domain, which mediates interactions with methylated histone tail residues. The INO80 family is distinguished from the other classes of ATPases by the presence of a short spacer sequence that separates the ATPase domain into two (Bao and Shen, 2007). That is, the INO80 class is characterised by its split ATPase domain as opposed to the intact ATPase domains of other ADCRs. The INO80 complex has recently been characterised and relatively less is known about this remodelling complex compared with other ADCRs.

Taken together, the homoeostasis of gene expression is maintained through coordinated regulation of genes by chromatin regulatory factors, transcription factors and histone modifying enzymes. Histone modifying enzymes are in part recruited to promoters by co-regulators which form a bridge between transcription factors and histone modifiers. One of these co-regulators is the co-repressor C-terminal Binding Protein (CtBP).

1.2 The CtBP family

The CtBP family of proteins are transcriptional co-repressors that are evolutionarily conserved within the animal kingdom. In vertebrates, CtBP is encoded by two separate genes, CtBP1 and CtBP2, as opposed to invertebrates, which have a single CtBP gene. In vertebrates, further complexity in the CtBP family arises from the expression of two CtBP isoforms from both the CtBP1 and CtBP2 loci. CtBP1 and CtBP1/BARS are produced from the CtBP1 locus via alternative splicing and CtBP2 and RIBEYE are produced from the CtBP2 locus, with the latter being generated with the use of an alternative promoter. Insights into the functions of these proteins were gained with Knock-Out (KO) studies in mice, where CtBP1 and CtBP2 double KO animals died at an early embryonic stage (E10.5) suggestive of

vital roles for CtBP1 and CtBP2 during development (Hildebrand and Soriano, 2002). Furthermore, the phenotypes of single and double KO mice provided hints on distinct and redundant roles of the CtBP proteins that are consistent with the expression patterns displayed by these proteins. Both CtBP1 and CtBP2 are expressed ubiquitously; however CtBP2 shows higher levels of expression during the embryonic stage and is enriched in the ear, eye and broadly in the nervous system. On the other hand, CtBP1 is enriched in the neural epithelium. Consistent with these observations, double KO mice display delayed development of the neural tube and malformation in multiple tissues. Interestingly, RIBEYE which is encoded by the CtBP2 locus is expressed exclusively in the synaptic ribbons in the retina and inner ear (Schmitz et al., 2000) whereas CtBP1/BARS is expressed ubiquitously and functions in membrane fission, particularly in Golgi maintenance (Bonazzi et al., 2005; Nardini et al., 2003; Weigert et al., 1999).

Due to lethality and redundancy, it has been challenging to decipher *in vivo* roles of CtBPs in mammalian systems. However, cellular assays performed on mammalian cell lines have provided insights into the modes of action of CtBP in repression of gene expression. CtBP is recruited to the promoters via interaction with a specific amino acid sequence on transcription factors which was first demonstrated in its identification as a binding partner of the adenoviral E1A oncoprotein. CtBP was reported to interact with a short amino acid sequence, PXDLS, in E1A (Boyd et al., 1993; Schaeper et al., 1995). The CtBP family of proteins also interact with transcription factors, such as the zinc-finger protein ZNF217, by other means including a recently-identified RRX motif (Quinlan et al., 2006).

1.3 The CtBP repression complex

CtBP acts as an adapter between DNA-binding transcription factors and a group of histone modifiers. The first evidence for the association of the CtBP co-repressor with histone modifying proteins was demonstrated by yeast-two hybrid experiments where class II Histone Deacetylases (HDACs) were found to interact directly with CtBP1 through the PXDLS motif (Schaeper et al., 1995). CtBP also associates with class I HDACs, however this interaction is indirect and it is thought that the co-repressor CoREST (co-repressor of RE-1 Silencing transcription factor REST) is the bridging molecule between CtBP and class I HDACs (HDAC1 and HDAC2) (Kuppuswamy et al., 2008).

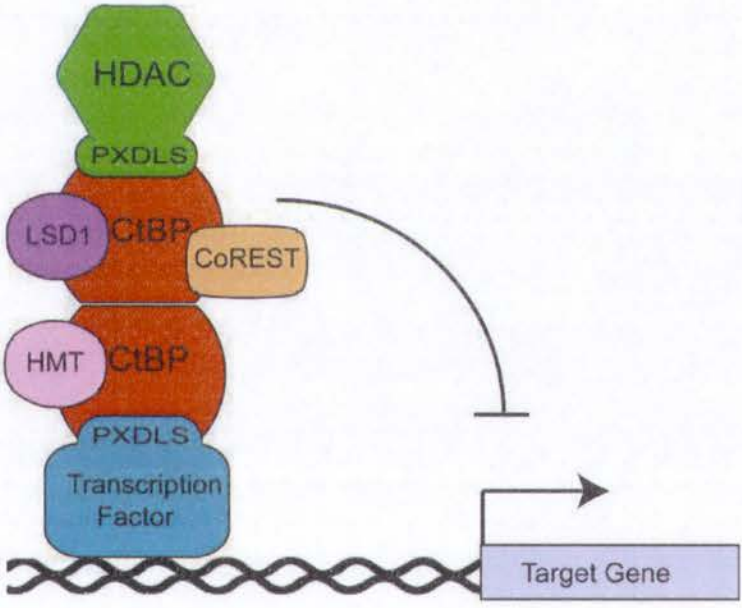


Figure 1.1: The CtBP repression complex. CtBP docks onto transcription factors containing the amino acid motif PXDLS and recruits histone modifying enzymes and other co-repressors.

The repression mechanism of CtBP was further elucidated with the identification of the CtBP nuclear complex via immunoprecipitation and mass spectrometry experiments by Shi and coworkers (Shi et al., 2003). In this study, CtBP was found in a multiprotein complex with histone modifying proteins and transcription factors. In addition to HDACs, CtBP was also shown to associate with other chromatin modifiers such as HMTs (G9a and GLP) and Lysine Specific Demethylase 1 (LSD1) (Fig. 1.1).

1.4 Regulation of CtBP by phosphorylation

CtBP was described as a phosphoprotein at the time of its identification as an interacting partner of the E1A adenoviral protein (Kajimura et al., 2008). The first evidence for how this phosphorylation might occur came from Yeast-Two Hybrid (Y2H) studies where p21-activated kinase-1 (Pak-1) was found to interact with CtBP (Barnes et al., 2003). Pak-1 belongs to the Ste-20 like serine/threonine protein kinase family. It phosphorylates CtBP1 on Ser-158 and thereby modulates its subcellular localisation. As a transcriptional co-regulator, CtBP1 functions in the nucleus. However, phosphorylation by Pak-1 changes its intracellular localisation from the nucleus to the cytoplasm, resulting in the relief of repression of CtBP1 target genes (Barnes et al., 2003).

Another kinase that phosphorylates CtBP is the Homeodomain Interacting Protein Kinase 2 (Hipk2) (Zhang et al., 2003). Similar to Pak-1, the interaction between Hipk2 and CtBP was identified through Y2H assays. Hipk2 is also a Serine/Threonine kinase and phosphorylates CtBP1 on Ser-422. Members of the Hipk family of proteins play roles in cell proliferation and apoptosis (Isono et al., 2006). One mechanism through which this is achieved is by the Hipk2-mediated phosphorylation of CtBP upon UV irradiation which targets CtBP for degradation and promotes cell death (Zhang et al., 2003; Zhang et al., 2005). Absence of CtBP leads to

increased expression of proapoptotic proteins including PERP, Noxa and Bax (Grooteclaes et al., 2003). HIPK2-mediated control of CtBP stability has also recently been shown to be important in cellular differentiation. In neural progenitor cells called neural crest cells, knockdown of Hipk2 was shown to stabilise CtBP resulting in repression of genes that promote melanocyte differentiation. By inference, under normal conditions, these progenitor cells acquire a gene expression profile favouring a melanocytic lineage during the differentiation process by degradation of CtBP via Hipk2 (Liang et al., 2011).

In summary, phosphorylation of CtBP at Ser-158 by Pak-1 results in a change in subcellular localisation and compromised repressive activity whereas phosphorylation at Ser-422 by Hipk2 induces the degradation of CtBP relieving apoptotic resistance and driving melanocyte differentiation depending on the context.

1.5 Regulation of CtBP by SUMO and nNOS

The Small Ubiquitin-like Modifier (SUMO) family of proteins are small molecules that function in targeting proteins for proteosomal degradation. They modulate various cellular processes depending on their targets (Dohmen, 2004). It has been established that CtBP1 is modified by SUMO-1 at Lys-428 and that mutation of this residue changes the subcellular localisation of CtBP1 from the nucleus to the cytoplasm (Feng et al., 2003; Lin et al., 2003). In accordance with this observation, cellular reporter assays have shown that this mutant form of CtBP1 can not repress its canonical target, E-cadherin, as efficiently as wildtype CtBP1. These results suggest that SUMOylation is required for optimal CtBP repression activity.

Another means by which the subcellular localisation of CtBP is regulated is through the interaction with a neuron specific enzyme called neuronal nitric-oxide synthase (nNOS). This

enzyme is involved in the production of nitric oxide which is a cell signalling molecule with important functions in the central nervous system. nNOS contains a type III PDZ domain (PSD-95, discs-large and zona occludens-1). Such domains have been shown to serve as protein-protein interaction interfaces. The nNOS PDZ domain binds CtBP and this interaction shifts CtBP from the nucleus to the cytoplasm (Riefler and Firestein, 2001). By inference, this would lead to relief of repression mediated by CtBP, however, the biological significance of regulatory function of nNOS in terms of its effect on the expression of CtBP target genes remains to be investigated.

1.6 Regulation of CtBP activity by NADH

In addition to the described mechanisms which affect the stability or subcellular localisation of CtBP, one further mechanism by which CtBP activity is controlled is through the interaction with nicotinamide adenine dinucleotides NAD(H). NAD(H) is a cellular signalling molecule with roles in metabolism, ageing and transcriptional regulation (Lin and Guarente, 2003). CtBP preferentially binds to NADH and the functional importance of the NADH-binding ability of CtBP was assessed by Y2H experiments and luciferase assays (Fjeld et al., 2003; Thio et al., 2004; Zhang et al., 2002). The CtBP family of proteins works as homo or heterodimers and these experiments demonstrated that NADH-binding is dispensable for dimerisation. However, binding of NADH to CtBP enhances the repression activity of CtBP. It was shown that mutant CtBP defective in NADH-binding retains its repressive activity, however, it is not as efficient as wild type CtBP in repressing E-cadherin expression through interaction with the transcription factor ZEB at the promoter. Therefore, CtBP can still repress transcription without its nucleotide binding domain but NADH-binding significantly increases its repressive activity. Taken together, interaction of CtBP with NADH might have a role in recruitment of additional factors to the target promoters. Therefore, it is possible that NADH-

binding acts as an aid for tethering CtBP to the multiprotein complexes and chromatin modifiers.

1.7 *C. elegans* CtBP vs human CtBP1

Our group has recently identified the sole CtBP homologue in *C. elegans*, called CTBP-1, and has shown that *C. elegans* CTBP-1 is functionally homologous to its mammalian counterparts in that it mediates repression of gene expression by binding to transcription factors containing the PXDLS motif (Nicholas et al., 2008). In addition, *C. elegans* CTBP-1 shows a high homology to mammalian homologues in terms of amino acid sequence. Figure 1.2 depicts human CtBP1 and *C. elegans* CTBP-1 containing several conserved domains. *C. elegans* CTBP-1 shows a 55 % amino acid identity and 74 % similarity within the dehydrogenase domain as indicated in yellow (Fig. 1.2) (Nicholas et al., 2008). The dehydrogenase domain,

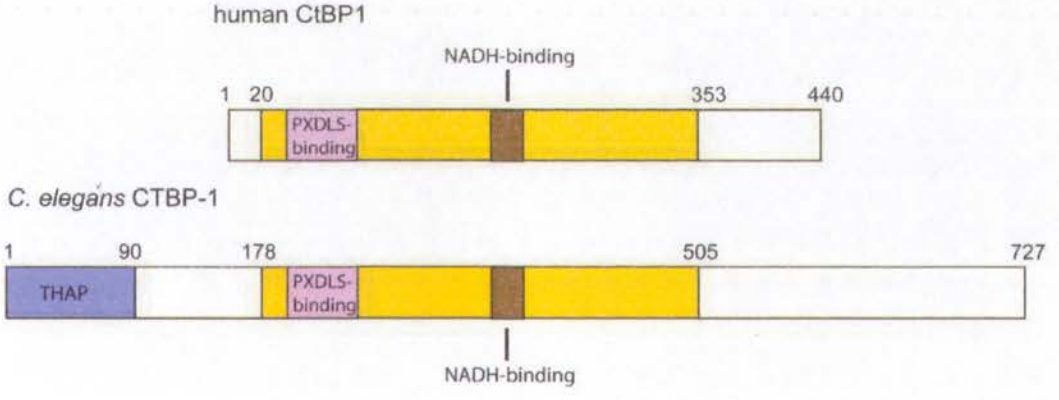


Figure 1.2: Schematic representations of the human CtBP1 and *C. elegans* CTBP-1 proteins. *C. elegans* CTBP-1 shows 55% identity and 74% homology to human CtBP1 within the d-isomer-specific 2-hydroxyacid dehydrogenase domain (yellow). This region includes NADH-binding domain and PXDLS-binding cleft. *C. elegans* CTBP-1 has an additional domain at the N-terminus called a THAP domain.

which is similar to d-isomer-specific 2-hydroxyacid dehydrogenase domain, contains the NADH-binding domain and PXDLS-binding cleft. However, *C. elegans* CTBP-1 contains an additional domain called Thanotas-Associated Protein (THAP) which is a recently-characterised DNA-binding domain (Roussigne et al., 2003).

Due to high sequence similarity and functional homology of *C. elegans* CTBP-1 to its mammalian counterparts, *C. elegans* serves as an excellent model in which to investigate the *in vivo* roles of CtBP. The *C. elegans* CTBP-1 has recently been reported to play a role in regulation of life span of the nematode. Mutants animals with an NADH-binding deficit in CTBP-1 (designated as *ok498* allele, revised in Chapter 3, Section 3.3) have extended life span and reduced fat levels (Chen et al., 2009). As elaborated above, CtBP proteins are NADH-regulated and thus implicated as a metabolic sensors. The study reported by Chen *et al.* demonstrates a link between the NADH-sensing ability of CtBP and its role in metabolism.

1.8 The model organism *C. elegans*

The round worm *C. elegans* serves as a great model organism with its simple anatomy and fully characterised, invariant cell lineage. *C. elegans* is a genetically tractable nematode with a short life cycle. It has a genome size of 97Mbp that is predicted to encode 19000 proteins, half of which have mammalian homologues (Consortium, 1998). Since its proposal as a model organism in which to study animal development and behavior, *C. elegans* has been used as a valuable tool for genetic studies. With the advantage of its characteristics that are elaborated below, studies in *C. elegans* have provided many important insights into eukaryotic biology. For instance, RNA interference was discovered in *C. elegans* and is now widely used for targeted knockdown in higher organisms (Fire et al., 1998). In addition, the first micro RNA was reported in *C. elegans*, bringing a new understanding to post-transcriptional regulation of gene expression (Lee et al., 1993). In addition to these

discoveries relating to gene regulation, important advances relating to understanding the molecular mechanisms of programmed cell death were made in *C. elegans* with key regulators later found to be conserved in higher organisms (Sulston and Horvitz 1977; Kimble and Hirsh 1979; Sulston et al. 1983).

C. elegans is cultivated easily on a nematode growth medium supplemented with bacteria as a food source. *C. elegans* is a self-fertilising, sexually dimorphic (hermaphrodite) nematode which produces both oocytes and sperm. However, fertilisation can also occur through mating with males that are generated at low incidence by chromosomal non-disjunction.

At the time of hatching, the *C. elegans* hermaphrodite has 558 cells and throughout development more cells are generated with further divisions, some of which undergo apoptosis. Between hatching and adulthood, *C. elegans* goes through four larval stages and an adult hermaphrodite has 959 somatic cells after apoptotic events. The speed of the growth from egg to adult depends on the temperature at which worms are cultivated. For instance, the generation time at 25°C (from egg to adult) is approximately 2.5 days.

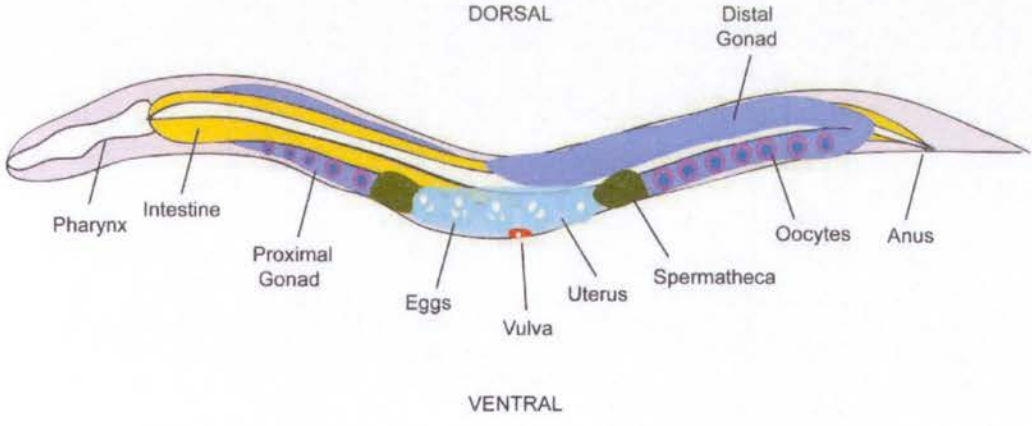


Figure 1.3: Basic anatomy of *C. elegans* at adult stage. Adapted from Varki *et al.* (Varki et al., 2009.)

The *C. elegans* body is formed of two tubes with the space between these tubes called the pseudocoelom. The inner tube is formed of pharynx, intestine and gonad (in adult stage), whereas the outer tube is formed of cuticle, the underlying epidermal tissue called hypodermis, neurons and muscles. The basic anatomy of *C. elegans* is illustrated in Figure 1.3. Of particular interest to this work are the development and anatomy of the nematode vulva and nervous system, which are detailed below.

1.9 The *C. elegans* vulva

The vulva is the egg-laying apparatus in the hermaphrodite *C. elegans*. It is a tube-like structure that connects the internal copulation organ to the external body (Sharma-Kishore et al., 1999). Vulval morphogenesis continues until fourth larval stage through a number of divisions that in the end forms the tube-like structure. At hatching, six out of twelve P cells form the vulval equivalence group and hold the potential to give rise to vulval tissue depending on the signal from a gonadal cell called the Anchor Cell (AC). The AC cells secrets the Epidermal Growth Factor (EGF) homolog LIN-3 which specifically induces P5-6-7 to acquire a vulval cell fate while P3-4 and P8 further divide and take the default path of

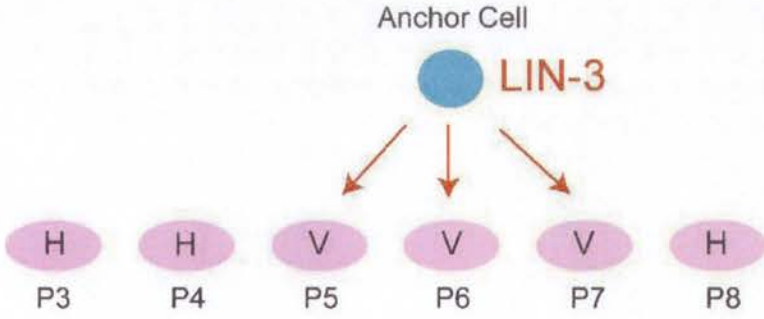


Figure 1.4: Model of vulval induction in *C. elegans*. In a wild type animal, the anchor cell (AC) secrets LIN-3/EGF which induces vulval induction in P5, P6 and P7 blast cells, and P3, P4 and P8 acquire hypodermal cell fate and fuses with the hypodermal syncytium.

fusing with the hypodermal cells (Fig. 1.4) (Cui et al., 2006b). The development of *C. elegans* vulva is controlled by various factors, many of which function in chromatin regulation in higher organism (Andersen, 2007; Ceol and Horvitz, 2004).

1.10 Anatomy of the *C. elegans* nervous system

The *C. elegans* nervous system is composed of 302 neurons. The resolution of the *C. elegans* nervous system is at electron microscopic level (White et al., 1986). The wiring network and the cell lineage of all 302 neurons are intricately described. The 302 neurons comprising the nervous system of a hermaphrodite worm at adult stage are classified in 118 different groups according to their function (i.e. sensory neurons, interneurons and motorneurons), neurotransmitter identity (i.e. cholinergic, GABAergic, glutamergic, serotonergic) and morphology. It is predicted that there are 5000 chemical synapses, 700 gap junctions and 2000 neuromuscular junctions connecting the entire nervous system (White et al., 1988). The *C. elegans* nervous system is largely bilaterally symmetrical with 198 out of 302 neurons being presented in pairs (Hobert et al., 2002). Most of these bilaterally symmetrical neurons are composed of sensory neurons and interneurons. Interestingly, motor neurons, most of which reside in the ventral midline of the body, do not have symmetrical pairs.

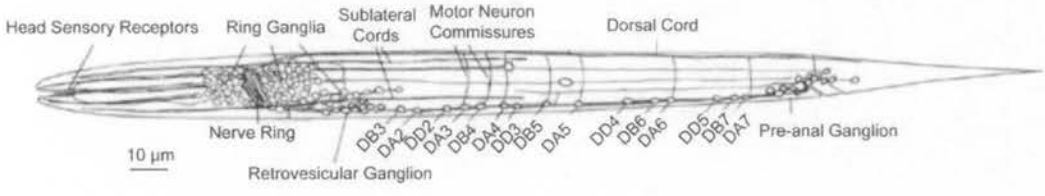


Figure 1.5: The nervous system of *C. elegans* at first larval stage (L1) is depicted. Adapted from Antebi *et al.* (Antebi et al., 1997).

At hatching, the nervous system of the worm is composed of 222 neurons (Fig. 1.5). By adult stage more neurons are added to the circuitry to reach the adult total of 302 described above. Of these, 75 lie on the ventral midline forming the Ventral Nerve Cord (VNC), a network of motor neurons resembling the vertebrate axial nerve cord (White et al., 1976). At the time of hatching, there are three classes of VNC motor neurons, namely DA, DB and DD (Fig, 1.5). Later in the development, 5 different classes of neurons are added to the VNC motor neurons. DA and DB neurons are cholinergic and DD neurons are GABAergic. These embryonic VNC motor neurons have longitudinal and circumferential axons innervating the body wall muscles dorsally. Whereas postembryonic neurons, namely VA, VB and VC send axons innervating the body wall muscles on the ventral side. These different classes of VNC motor neurons coordinate body movement of the nematode.

Taken together, with its well defined nervous system, *C. elegans* serves as an excellent model in which to investigate how neurons are generated and specified.

1.11 Aim of this project

The CtBP family of proteins are mechanistically well characterised, however, relatively little is known about their *in vivo* roles. Thus far, various roles for CtBP proteins have been reported by virtue of their interaction with several transcription factors in mammalian cell lines. However, the biological significance of these interactions has not been elucidated. The research detailed in this thesis aimed to understand how CtBP functions using *C. elegans* as a model. Towards achieving this goal, the gene regulatory network in which CTBP-1 is involved and factors that regulate CTBP-1 were explored. An interest was developed in the nervous system of the nematode due to pre-dominant neuronal expression of CTBP-1. Phenotypic analysis of animals with reduced CTBP-1 activity and cell-specific microarray profiling on wild type and CTBP-1-deficient *C. elegans* neurons was employed to investigate CTBP-1 targets. Furthermore, with a forward genetics approach, a mutagenesis screen was carried out to identify regulators of CTBP-1.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals and Reagents

- Acrylamide/bis-acrylamide 30% solution electrophoresis reagent (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Agar bacteriological (Amyl Media Pty. Ltd., Dandenong, Australia)
- Agarose (DNA grade) (Applichem, Darmstadt, Germany)
- Ampicillin sodium salt (Roche Molecular Biochemicals, Mannheim, Germany)
- Bovine Serum Albumin (BSA) (100x) (New England Biolabs, Ipswich, USA)
- Calcium chloride (Ajax Laboratory Chemicals, Taren Point, NSW, Australia)
- Chloroform (Biolab Scientific, Clayton, VIC, Austrlia)
- Chloroform:isoamyl alcohol 24:1 (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Cholesterol (Sigma Chemical Company, Castle Hill, NSW, Australia)
- DAPI (Abacus ALS, Australia)
- Deoxynucleotide triphosphates (dNTPs) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Diethyl pyrocarbonate (DEPC) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Dithiothreitol (DTT) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- DNA loading dye (6x) (Quantum Scientific, Fermentas Life Sciences, Murarrie, Australia)
- Dulbecco's Modified Eagle Medium (DMEM) (low glucose) (Gibco-BRL Life Technologies, Grand Island, New York, USA)
- EDTA di-sodium salt (Ethylenediaminetetra-acetic acid di-sodium salt) (Ajax

Finechem Pty. Ltd., Taren Point, Australia)

- Ethanol (APS Finechem, Seven Hills, Australia)
- Ethidium bromide (Roche Molecular Biochemicals)
- Foetal Calf Serum (FCS) (Commonwealth Serum Laboratories, Parkville, VIC, Australia)
- FuGENE® 6 transfection reagent (Roche Molecular Biochemicals)
- Full range Rainbow™ Recombinant Protein Molecular Weight Marker (Amersham™ GE Healthcare, Little Chalfont, UK)
- GeneRuler™ DNA ladder mix (Fermentas Life Sciences, Ontario, Canada)
- Glycerol (Sigma Chemical Company, Castle Hill, NSW, Australia)
- HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Instant skim milk powder (Fonterra™, Mount Waverley, Australia)
- Isopropanol (BDH Chemicals, Port Fairy, VIC, Australia)
- L-15 medium (Gibco)
- Magnesium chloride (Biolab Scientific)
- Magnesium sulphate (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- β-mercaptoethanol (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Methanol (EM Science, Gibbstown)
- MOPS SDS Running Buffer (20x) (Invitrogen™ Carlsbad, USA)
- NEBuffer 2 (New England Biolabs, Ipswich, USA)
- Penicillin, Streptomycin and Glutamine solution (1%) (Gibco-BRL Life Technologies)
- Phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma Chemical Company, Castle Hill, NSW, Australia)

- Phosphate Buffered Saline (PBS) tablets (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Phenylmethylsulfonide (PMSF) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Polyoxyethylenesorbitanmonolaurate (Tween™-20) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Potassium dihydrogen orthophosphate (KH_2PO_4) (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Di-potassium hydrogen orthophosphate (K_2HPO_4) (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Sodium chloride (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Sodium dodecyl sulphate (lauryl sulphate sodium salt) (SDS) (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Sodium hydroxide (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Di-sodium hydrogen orthophosphate (Na_2HPO_4) (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Sodium hypochlorite 12.5% w/v solution (Ajax Finechem Pty. Ltd., Taren Point, Australia)
- Tris-hydroxymethyl-methylamine (Tris) (Ajax Laboratory Chemicals, Taren Point, NSW, Australia)
- Tetramisole hydrochloride (Sigma Chemical Company, Castle Hill, NSW, Australia)
- T-octylphenoxypolyethoxyethanol (Triton® X-100) (Sigma Chemical Company)
- TRI reagent (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Tryptone (peptone) (Amyl Media, Dandenong, VIC, Australia)
- Yeast extract (Amyl Media Pty. Ltd., Dandenong, Australia)

2.1.2 Antibodies

Primary antibody

Anti-Gal4DBD polyclonal antibody (Santa Cruz Biotechnology)

Secondary antibody

Horseradish peroxidase-linked anti-rabbit antibody (Amersham Biosciences)

2.1.3 Commercial reagents and kits

- Wizard® SV Gel and PCR clean-up system (Promega, Madison, USA)
- NuPAGE® 10% Bis-Tris Gel (Invitrogen™, Carlsbad, USA)
- Millipore Immobilon™ Western Chemiluminescent HRP Substrate (Millipore Corporation, Billerica, USA)
  - o HRP Substrate Peroxide Solution
  - o HRP Substrate Luminol Reagent
- Jetstar 2.0 Plasmid Purification MIDI Kit (Astral Scientific, Genomed, Gymea, Australia)
- RNeasy® Mini Kit (Qiagen, Doncaster, Australia)
- GoTaq® Green Master Mix Sample (Promega Corporation, Madison, USA)
- Dual Luciferase Reporter Assay System (Promega Corporation, Madison, USA)
- FastPlasmid Mini kit (Eppendorf South Pacific, North Ryde, NSW, Australia)
- SuperScript™ III First-Strand cDNA synthesis Kit (Invitrogen Australia Pty. Ltd., Mount Waverley, VIC, Australia)

2.1.4 Equipment

- Nanodrop1000 Spectrophotometer (Thermo Scientific)
- TD-20/20 Luminometer Turner Designs (DLReady™)

- CyberScan510 pH meter
- XCell II Blot Module Novex® Mini-Cell (Invitrogen™ Carlsbad, USA)
- Life Sciences BioTrace™ Pure Nitrocellulose Blotting Membrane (Pall Corporation, Pensacola, USA)
- Kodak X-Omatic Cassette
- Kodak Medical X-tray Film General Purpose Blue (Carestream Health Inc., Rochester, New York, USA)
- 6-well cell culture plate (Cellstar®, Greiner bio-one)
- 10 cm tissue culture dishes (Cellstar®, Greiner bio-one)
- Spin-X® centrifuge tube filters (Trace Biosciences, Castle Hill, NSW, Australia)

2.1.5 External Procedures

- Next-Generation Sequencing: Australian Genome Research Facility (AGRF) (Gehrmann Laboratories, The University of Queensland, Research Road, St Lucia, Brisbane, Queensland, Australia)
- Isolation of neurons via FACS: Vanderbilt Flow Cytometry Core Lab (Vanderbilt University , 1211 Medical Center Drive, Vanderbilt University, Nashville-Tennessee, USA)
- Tiling Arrays : Genome Sciences Resource (GSR) Microarrays (Vanderbilt University, 465 21st Avenue South, Nashville-Tennessee, USA)

2.1.6 Enzymes

- T4 DNA Ligase (poly(deoxyribonucleotide):poly(deoxyribonucleotide)) ligase (AMP-forming) (New England Biolabs, Ipswich, USA)
- *Bam*HI, *Hind*III, *Eco*RI, *Eco*RV (Type II site-specific deoxyribonucleases, E.C. 3.1.21.4)

(NewEngland Biolabs, Ipswich, USA)

- Chitinase (Sigma)
- SSIII® Reverse Transcriptase (DNA nucleotidyltransferase (RNA directed)) (Invitrogen Pty. Ltd.)
- RNAse H (Endoribonuclease H, E.C. 3.1.26.4) (Invitrogen Pty. Ltd.)
- AmpliTaq® Gold DNA Polymerase – SYBR green reagent (deoxynucleosidetriphosphate: DNA deoxynucleotidyltransferase (DNA directed), E.C..7.7.7) (Applied Biosystems)
- RNAse-free DNAse (deoxyribonuclease I, E.C. 3.1.21.1) (Qiagen Pty. Ltd.)
- RNAse A (pancreatic ribonuclease, E.C. 3.1.27.5) (Gentra Systems, Inc., Minneapolis, MN, USA 18)
- Proteinase K (peptidase K, E.C. 3.4.21.64) (Boehringer Ingelheim Pty. Ltd., North Ryde, NSW, Australia)
- RedTaq® DNA Polymerase (Sigma Chemical Company, Castle Hill, NSW, Australia)
- Platinum Taq DNA polymerase (deoxynucleotide-triphosphate: DNA deoxynucleotidyltransferase (DNA directed))
- Proteinase inhibitors: Aprotinin (1 mg/mL), Leupeptin (1 mg/mL), PMSF (Sigma Chemical Company, Castle Hill, NSW, Australia)

2.2 Methods

2.2.1 Nematode stocks and maintenance

The strains used in this study were cultured as previously described (Brenner, 1974). The Bristol isolate N2 was used as the wild type strain. Animals were grown on Nematode Growth Medium (NGM) (3 g/L Sodium Chloride, 2.5 g/L Bacto peptone, 17 g/L Bacteriological agar). seeded with OP50 strain of *E. coli*. Occasionally, cultures were cleaned with spot

bleaching to eliminate mould contamination. Depending on the experimental conditions nematodes were grown at different temperatures (15 °C, 20 °C and 25 °C).

The following strains were used for investigation of *ctbp-1* function:

N2 (Bristol), LGX: HRN017 *ctbp-1(ok498)*, LGI: HRN054 *pctbp-1::ctbp-1::gfp(ausIs1)*; LGII: HRN119 *unc-25::gfp(juIs76)* (Huang et al., 2002); LGIV: *zag-1(rh315);* LGV: *lag-2::gfp(qIs56)* (Blelloch et al., 1999); linkage unknown: NC300 *unc-4::gfp(wdIs5)* (Lickteig et al., 2001), NW1229 *F25B3.3::gfp(evIs111)* (Altun-Gultekin et al., 2001) (J Culotti unpublished), NC575 *acr-5::yfp(wdIs23)* (Christensen et al., 2002), VH37 *unc-3::gfp(rhIs11)* (Wacker et al., 2003a).

The following strains were used for investigation of C34E10.8 function:

HRN099 and HRN109 *pctbp-1::ctbp-1::gfp(ausIs1);* ExC34E10.8*(ausEx23); aus3*

HRN101 and HRN102 *pctbp-1::ctbp-1::gfp(ausIs1); rol-6(su1006)(ausEx24); aus3*

HRN110 *pC34E10.8::C34E10.8::gfp(ausEx25)*

HRN096 *lin-15AB(n765)* (Clark et al., 1994)

HRN125 *lin-15AB(n765); aus3*

2.2.2 Germline transformation

For generation of a CTBP-1::GFP reporter construct, a PCR fusion approach was used, which employs two rounds of amplification. In the first round of amplification, the full-length *ctbp-1* with its 2kb 5′ flanking region was amplified from N2 genomic DNA using Expand Long Template kit (Roche) according to manufacturer's protocol, and *gfp* was amplified from pPD95.75 (Addgene) (Hobert, 2002). The reverse primer used for amplification of the *ctbp-1* amplicon contained a 24bp-long complementary sequence to the *gfp* sequence, therefore in the second round of amplification a *ctbp-1::gfp* fusion product was produced with the aid of this 24bp-long overlap sequence. The fusion template was injected at a concentration of 50

ng/µl to the gonad of the N2 hermaphrodites as previously described (Fisher et al., 2010). Transformants were selected using a fluorescence dissecting microscope. Three independent transgenic lines carrying extrachromosomal arrays of this *pctbp-1::ctbp-1::gfp* fusion product were analysed and, since they showed equivalent expression patterns (data not shown), one array was selected for integration by gamma irradiation.

Primer details:

pPD9575gfpf (#182): AGCTTGCATGCCTGCAGGTCG

pPD9575gfpr (#183): AAGGGCCCGTACGGCCGACTA

nestedgfprev (#210): GGAAACAGTTATGTTTGGTATA

ctbp proA (#226): GATAACTCTCTTGTTGTATTTGG

ctbpoverlap(#228):

AGTCGACCTGCAGGCATGCAAGCTTGTGGCCAATGGTTGCTCATC

ctbp pro nest (#227): AGCATTACTCGAAAGTTGTCG

2.2.3 Generation of integrated transgenic strain carrying pctbp-1::ctbp-1::gfp reporter

Extrachromosomal CTBP-1::GFP array was integrated using gamma irradiation (Koelle, 1994). First, animals carrying extrachromosomal array were selected and allowed to lay eggs. When the progeny reached L4 stage, they were exposed to gamma irradiation at a dose of 3600 Radiation Absorbed Dose (Rad). Mutagenised population gave rise to F1 progeny and 500 F1s were singled on individual plates to produce F2 generation to score for integrants. From 500 plates containing progeny of F1, i.e. F2 generation, 2 animals were selected and singled on individual plates. Based on the probability of an F1 animal being heterozygous for the integration; the lucky plate containing a heterozygous F1 will produce F2 progeny quarter of which will be homozygous for the integration. In the end, 1000 plates were screened for the

integration and one out of thousand plates contained transgenic animals all of which express GFP. This strain was outcrossed 6 times to remove background mutations and was given the allele name of *ausIs1*.

2.2.4 Reporter marker scoring:

Strains with UNC-4::GFP, UNC-25::GFP, ACR-5::YFP, F25B3.3::GFP, UNC-3::GFP markers were crossed into *ctbp-1(ok498)* mutants. Gravid hermaphrodites of mutants and wild-types were treated with BS2X (0.5 N NaOH, 20 % (v/v) Clorox) to isolate the eggs. For synchronisation, eggs were incubated in M9 buffer (6 g Na_2HPO_4, 3 g KH_2PO_4, 5 g NaCl, 0.25 g $MgSO_4 \cdot 7H_2O$ per liter) overnight (18-22 hours) and starved L1s were plated on nematode growth medium seeded with OP50 bacteria. They were allowed to recover for approximately 2 hours and then prepared for microscopy by anesthetising in 20 mM NaN_3 or 1% tetramisole. Anesthetised worms were mounted onto 2% agarose pads for scoring. The confocal images were taken with an Olympus FV1000 microscope at the Australian Centre for Microscopy and Microanalysis (ACMM), University of Sydney.

2.2.5 Genotyping

The genotype of the strains for the *ctbp-1(ok498)* mutation was confirmed via multiplex PCR using the following primers:

ok498flankr (#138): GCTCATTGTTCATTGACCG

CtBPintron6rev (#170): GCCACAGCTTCATTTAACACC

ok498flankl (#137): AGCGTAATTGCAGAGTACC

A single worm was picked into 3µl of lysis mix (50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM $MgCl_2$, 0.45 % NP-40, 0.45% tween-20, 0.01 % gelatine and 0.5 mg/ml proteinase K) and frozen at -80°C. The frozen pellet was incubated at 65°C for 1 hour followed with 15 minutes of 95°C incubation to inactivate Proteinase K. PCR mix (0.25 µM dNTPs, 1X reaction buffer containing $MgCl_2$, 0.025 units/µl *Taq* DNA polymerase) was prepared using reverse primers (ok498flankr and CtBPintron6rev) at a final concentration of 0.2 µM and common forward primer (ok498flankl) was used at a final concentration of 2 µM. RedTaq DNA polymerase (Sigma-Aldrich) was used and amplicons were run on a 1.5 % agarose gel to determine the genotype of the strains. Mutant animals were distinguished from wild type animals on the basis of product size differences, i.e. 588bp amplicon for wild types and 236 bp amplicon for mutants. Thermal cycling conditions used for this genotyping protocol are as follows: 94 °C 2 min, 94 °C 20sec, 54 °C 40sec, 72 °C 1min (30 cycles of step 2-step 4), 72 °C 5min.

For genotyping of *aus3* mutants, the same conditions as above for PCR and thermal cycling were used, together with the following primers at a final concentration of 0.2 µM. These primers produce three bands at 177 bp, 335 bp and 512 bp in a wild type and only one band at 335 bp in *aus3* mutants.

657: CTTCTCAAACTTTTCGATCTCAAG

658: GAACAGCTTGAGGGCTCTG

659: TGATGCCAGTAGCAATGAATT

660: TAAATACAAACCCCGAAGCC

The genotyping conditions for the *lin-15AB(n765)* allele was the same as above for both the PCR reaction mix and the thermal cycling except the annealing temperature, which was 48 °C. Primers targeting lin-15AB locus:

648: CATTGTGTGTAGGTAAATTTAGCG

649: TTTCATTGTCAATGTAGTGGAGG

650: CATTGTGTGTCCAAACGAAG

2.2.6 Yeast-two hybrid assays

Preparation and transformation of competent yeast was performed according to the MATCHMAKER Two-Hybrid System protocol (Clontech). In order to test interaction of ZAG-1 and CTBP-1, Yeast Two Hybrid method was used according to the manufacturer's instructions. Briefly, the HF7c yeast strain was transformed with constructs containing the gene of interest fused to GAL4-AD or GAL-DBD. Tina Wu (Wu, 2006) generated constructs by subcloning wild type and mutant proteins into pGAD10 (Clontech) and pGBT9 (new) as described previously (Hampsey, 1998; Nicholas et al., 2008). Mutation of the ZAG-1 PXDLS motif and the CTBP-1 PXDLS binding cleft was performed by site-directed mutagenesis as described by Nicholas *et al.* (Nicholas et al., 2008). Yeast transformations were performed by Estelle Llamosas at the University of Sydney.

2.2.7 Luciferase assays

These experiments were carried out by Tina Wu at the University of Sydney (Wu, 2006). COS-1 cells were cultured and transfected with Luciferase reporter constructs as previously described (Hampsey, 1998; Nicholas et al., 2008). Following the transfection of COS-1 cells with the mammalian expression vector pcDNA3 (Invitrogen) containing the relevant cDNA (i.e., ZAG-1 and CTBP-1 and mutated forms of these proteins), the Dual Luciferase Reporter Assay System (Promega) was used to measure the expression of the firefly and *Renilla* luciferase genes according to the manufacturer's instructions. The measurements were detected using a Turner Designs model TD 20/20 luminometer and Firefly Luciferase

expression intensity for each sample was normalised against the Renilla Luciferase expression intensity.

2.2.8 Western blot

Western Blot was performed using nuclear extracts from mammalian cells as described previously (Hampsey, 1998). *C. elegans* CTBP-1 was detected using an anti-GAL4DBD antibody as CTBP-1 constructs introduced into mammalian cells (i.e. COS cells) encode a GAL4DBD-CTBP-1 fusion protein. 20 μg rabbit anti-Gal4DBD antibody was diluted in 10 mL TBST (50 mM Tris, pH7.4, 150 mM NaCl, 0.05% Tween-20) and the blot was incubated with diluted primary antibody for 1 hour at room temperature. The membrane was washed > 2 times with TBST and was incubated for another hour with secondary antibody (1 μL horseradish peroxidase (HRP) conjugated anti-rabbit antibody diluted in 10 mL of TBST). HRP substrate peroxide solution and luminol reagent (1:1) (Millipore) were added to the membrane. The membrane was developed with a Kodak X-Omatic Cassette and Kodak Medical X-ray Film. For size estimation, rainbow molecular weight markers (GE Healthcare) were used.

2.2.9 Isolation of embryonic cells

Strains containing the pan-neural marker (F25B3.3::GFP) in wild type and *ctbp-1(ok498)* mutant background were cultured on 60 mm NGM/OP50 plates until starved and chunked onto 150 mm plates containing 8-P media (2.5% (w/v) agar, 50mM NaCl, 2.5% (w/v) peptone) seeded with Na22 strain of *E. coli*. In this way, semi-synchronous gravid adults were obtained and eggs were isolated by bleaching animals with BS2X (0.5 M NaOH, 20% (v/v) hypochlorite solution). The reaction was stopped by adding egg buffer (118 mM NaCl, 48 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 25 mM Hepes, pH 7.3, 340 mOsm) and bleach solution

was removed by several washing steps with the egg buffer. Embryos were separated from carcasses and debris by floatation on 30% sucrose. Embryos floating on top were collected into a separate tube and washed with egg buffer in order to remove the sucrose. After washing with egg buffer, embryos were pelleted and treated with the enzyme chitinase (0.5U/ml in egg buffer) for 45 minutes. To stop the reaction, L-15 medium (Gibco 21083-027) supplemented with 10% FBS and penicillin/streptomycin (referred to as L15-10) was added to the enzyme suspension (2:1 L15-10/chitinase solution). The dissociated embryos were filtered through a 0.5 µm syringe filter (Durapore). Cells were plated at a density of ~10^7 cells/ml on single-well chambered slides (Nalge Nunc International) that are coated with poly-L-lysine (0.01 %, Sigma). Following plating, cells were incubated in a humidified chamber at 25 °C overnight. A wild type N2 strain was used as a control for autofluorescence for FACS step and treated using identical conditions.

2.2.10 FACS analysis

To isolate GFP-expressing neurons, first the supernatant from an overnight embryonic culture was discarded and cells dislodged using 3 ml of egg buffer. This was filtered with a 0.5 µm syringe filter and collected in 5 mL polystyrene round-bottom tubes (BD Falcon). Immediately before sorting, Propidium Iodide (PI) was added to the tube at a final concentration of 1 µg/ml in order to label the dead cells.

Sorting experiments were performed on a FACSAria flow cytometer (75 µm nozzle, ~10.000-15.000 events/second) (Becton Dickinson, San Jose, CA) at the Vanderbilt Flow Cytometry Core Lab, Vanderbilt University. Autofluorescence levels were established using non-GFP cells from wild type (N2) strain. The PI levels used were adjusted to define the sorting gate for damaged cells. FACS gates for size and granularity were adjusted to achieve ~90 % purity

which means 90 percent of the cell population marked with GFP which comprise the neuronal cell population, thereby excluding cell clumps and debris (Fox et al., 2005). The fraction of isolated neurons (~90%) was determined by fluorescence microscopy 4-6 hours after plating on 4-well chamber slide coated with peanut lectin (Sigma). In addition to GFP-tagged neuronal cells, all embryonic cells were also collected via FACS as a control for data analysis. This ensured that both neuron samples and all cell samples were treated in the same way. The yield for both F25B3.3 in wild type F25B3.3 in *ctbp-1(ok498)* mutant samples ranged from 25.000-100.000 cells. 3 independent cultures were prepared for each condition.

2.2.11 RNA isolation, amplification and hybridization

Sorted neuronal and control embryonic cells were collected in a tube containing Trizol LS (Invitrogen) and total RNA was isolated as described by Fox *et al.* (Fox et al., 2005). The samples were extracted with chloroform and RNA was precipitated with isopropanol. The pellet was washed twice with 75% ethanol and resuspended in RNAase-free water. RNA was purified using an RNA purification kit (Zymo Research) according to the manufacturer's protocol. The quality and yield of the purified RNA was assessed by a Bioanalyser (Agilent). The total RNA isolated ranged from 300 pg to 4 ng.

The following steps after RNA isolation was performed by Rebecca McWhirter at the Vanderbilt University. The RNA was amplified using WT-Ovation Pico kit (NuGEN Technologies, Inc) according to the manufacturer's instructions. The starting material for amplification and generation of single-stranded cDNA ranged from 1.5 ng to 20 ng RNA. 3 µg of cDNA from each amplified sample was used for generation of double-stranded cDNA, with WT-Ovation Exon Module (NuGEN Technologies, Inc). 5 µg of double-stranded cDNA was fragmented, labeled using Encore Biotin Module (NuGEN Technologies, Inc) and

hybridized to an Affymetrix Microarray- GeneChip® *C. elegans* Tiling 1.0R Array. The *C. elegans* 1.0R Array is a single array containing over 3 million perfect match (PM)/mismatch (MM) probe pairs tiled throughout the complete non-repetitive *C. elegans* genome.

2.2.12 Analysis of data

The following analysis was done by Clay Spencer at the Vanderbilt University. Hybridisation intensities for each data set were quantile-normalized using robust multiarray analysis (RMA) (Bolstad et al., 2003; Fox et al., 2005). Expressed transcripts were defined using Affymetrix MAS 5.0 and the expression intensity for any given transcribed region was further normalised to reduce probe sequence bias as described previously (Georg Zeller, 2008). To detect neuronal targets of CTBP-1, normalised intensities for the pan-neural data set in *ctbp-1(ok498)* mutants versus wild types were statistically analysed using Significance Analysis of Microarray software (SAM, Stanford) (Irizarry et al., 2003). A two-class unpaired analysis of the data was performed to identify genes in *ctbp-1(ok498)* mutants that differ by $\geq$ 2-fold from the wild type neuron data sets at a p-value of $\leq$ 0.05. These genes were considered significantly upregulated in *ctbp-1(ok498)* neurons compared with the wild type neurons (Appendix Table 2). Analysis of control data from all embryonic cells was performed using the same protocol (Appendix Table 3).

To validate microarray results, pan-neural enrichment was compared with the published pan-neurally enriched gene list (Von Stetina et al., 2007). Normalised intensities in wild type neurons versus all wild type cells and *ctbp-1(ok498)* neurons versus all *ctbp-1(ok498)* cells were statistically analysed as above. For wild type data sets, 1445 pan-neurally enriched transcripts were detected, and for *ctbp-1(ok498)* mutant data set 1407 pan-neurally enriched

transcripts were detected, compared to the published data in which 1637 pan-neurally enriched transcripts were detected (Von Stetina et al., 2007).

2.2.13 Mutagenesis screen

A non-clonal F2 mutagenesis screen was carried out by mutating *ctbp-1::gfp (ausIs1)* transgenic animals with Ethyl Methane Sulfonate (EMS). First, worms were washed off plates with M9 (1 mM MgSO4; 22 mM KH2PO4; 42 mM Na2HPO4; 86 mM NaCl) then washed several times and the worm pellet was resuspended in 1 ml M9 buffer after the last wash. In a fume hood, 3 ml EMS (47 mM) solution was added to the 1 ml suspension at a concentration of 35 mM. Nematodes were incubated for 4 hours at room temperature for mutagenesis to take place; during this time the mixture was swirled every half an hour to avoid hypoxic conditions. After 4-hour incubation, worms were washed three times with M9 buffer to remove EMS solution. They were resuspended in a minimal volume, then distributed onto NGM plates containing OP50 bacteria and kept at room temperature for 2 hours for recovery. Later, L4 animals were picked onto new plates and aged 24 hours at 25 °C to remove early progeny that may not have been effectively mutagenised. Next day, 4 gravid adults were picked onto 20 plates. After 4 hours parents were removed onto new plates and further incubated for 4 hours at the same temperature. This step was repeated once more such that in the end 60 plates containing mutagenised embryos were accumulated. Approximately 3000 F1 progeny was grown till adulthood and were allowed to produce F2 offspring. Each set of plates from 4-hour incubations were staggered in order to ensure careful and efficient scoring of the F2 progeny for any change in the CTBP-1::GFP expression. Last batch of plates were scored first for 2 consecutive days which allowed for observation of progeny at different developmental stages. First and second batch of plates were also scored for two days and were staggered as depicted in Figure 2.1. Candidate F2 mutants were selfed to check whether the

phenotype transmitted to the next generation. All observations were done under dissecting microscope.

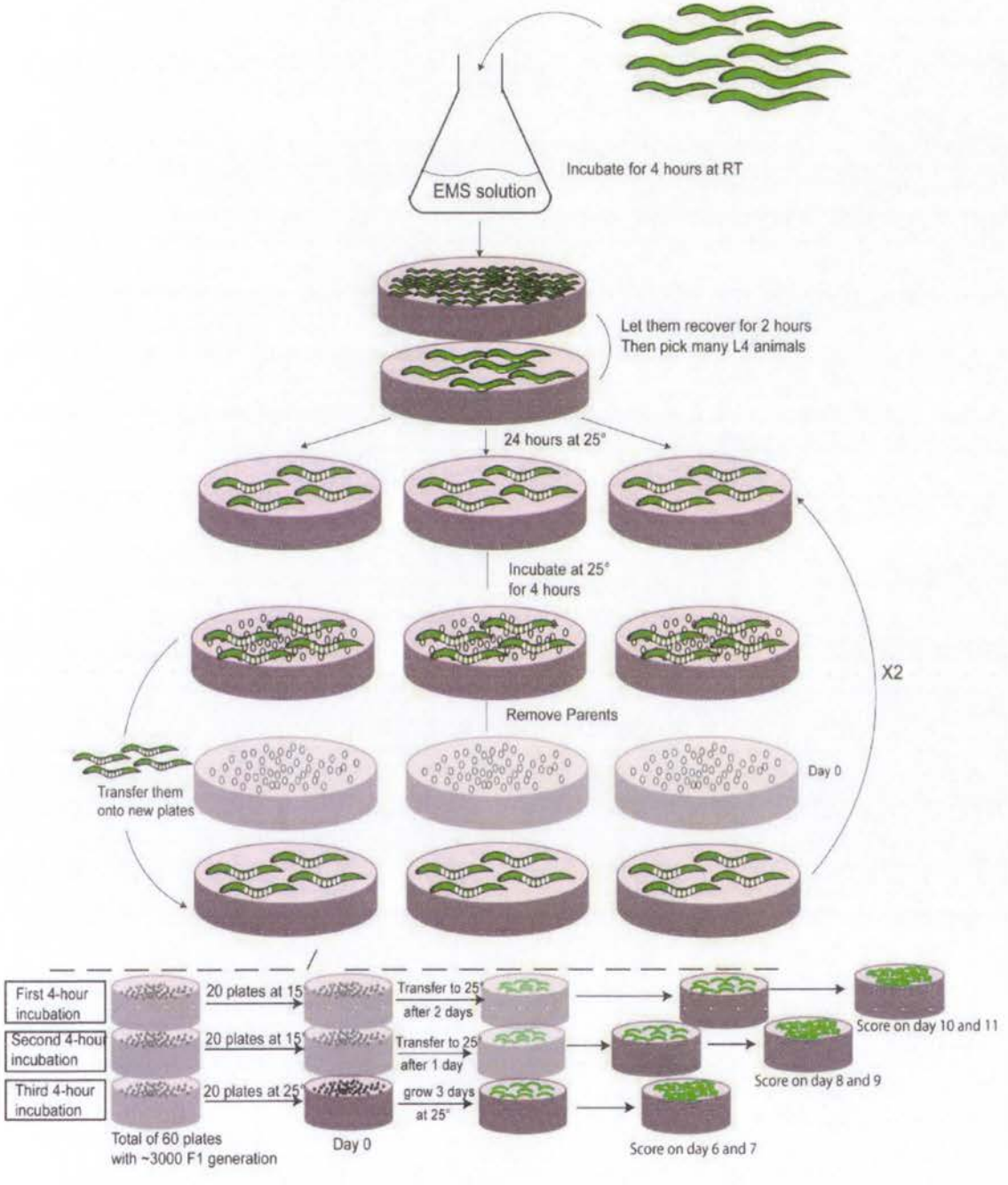


Figure 2.1: Mutagenesis screen for the identification of mutants with altered expression of CTBP-1::GFP. Transgenic animals expressing CTBP-1::GFP were treated with Ethyl Methane Sulfonate (EMS) for 4 hours and after 2-hour recovery on nematode plates (NGM seeded with OP50), L4 stage animals were picked onto new plates and aged for 24 hours at 25 °C to obtain gravid adults. On the next day, 4 gravid adults were picked onto each of 20 new plates and incubated for 4 hours at 25 °C then this step was repeated twice more. In 12 hours, three batches of plates (3x20 plates) were accumulated which contained eggs laid by EMS-treated parents. The growth rate of progeny on these plates was staggered by culturing the plates at different temperatures to enable efficient scoring. Plates accumulated from first and second set of incubations were kept at 15 °C whereas the plates from third set of incubation were kept at 25 °C. After 3 days, the third batch of plates were populated with gravid F1 generation and after 3-4 more days this batch of plates were scored for any difference in the expression level or pattern of CTBP-1::GFP. The day where all plates had just F1 eggs is called Day 0. By taking this day as reference, in Day 1, the second batch of plates were shifted to 25 °C and scored on Day 8 and 9. The first batch of plates were shifted to 25 °C in Day 2 and scored in Day 10 and 11. This way, the F2 generation was scored on Day 6-7 for third batch of plates, on Day 8-9 for second batch of plates and Day 10-11 for first batch of plates.

2.2.14 Snip-SNP mapping

Conventional snip-SNP mapping was used to locate the mutation of interest on a particular chromosome. The methodology described by Davis *et al.* was followed to amplify and digest a short polymorphic sequence from both N2 and Hawaiian isolates to identify the origin of mutation according to linkage analysis (Davis et al., 2005). Primer details can be found in the list provided by Davis *et. al.* in Table 1.

2.2.15 ImageJ quantification

The phenotype of *aus3* mutants was measured using software Image J. To measure the GFP intensity in the hypodermal cells where CTBP-1::GFP intensity is higher in *aus3* mutants (*pctbp-1::ctbp-1::gfp (ausIs1); aus3)* compared with the wild type (*pctbp-1::ctbp-1::gfp(ausIs1))*, a particular area of hypodermal cells was selected. Therefore, the anterior region containing the first two pairs of intestinal cells was selected and CTBP-1::GFP

intensity in this area of hypodermal cells was quantified using the threshold function of ImageJ. The Olympus FV1000 confocal laser scanning microscope equipped with an EGFP filter (488nm excitation) was used to acquire images at the Australian Centre for Microscopy and Microanalysis (ACMM), University of Sydney. For both strains, 10 animals were imaged with the same settings (i.e., same laser intensity and same voltage). The hypodermal CTBP-1::GFP intensity in *pctbp-1::ctbp-1::gfp(ausIs1); aus3* mutants was normalised against the *pctbp-1::ctbp-1::gfp(ausIs1)* strain. Student's t-Test was used for statistical analysis.

2.2.16 RNA extraction from whole worms and cDNA synthesis

Worms were synchronised and grown on NGM/OP50 plates until the young adult stage. They were then washed off the plates with M9 and collected into a 15 mL tube. In order to avoid contamination with OP50, worms were washed with M9 several times until a clear supernatant was observed. The worm pellet was resuspended in TRI Reagent® in a 2:1 volume ratio and snap frozen in liquid nitrogen. The frozen pellet was thawed and this freeze-thaw procedure was repeated a further two times. The pellet was supplemented with TRI Reagent® (4:3 volume ratio) and mixed vigorously. Chloroform:isoamylalcohol (24:1) was added to the worm lysate (5:1 volume ratio) to extract RNA and 1 volume of 70% ethanol was added to the aqueous layer and transferred to an RNeasy column (Qiagen) to purify RNA according to the manufacturer's instructions. RNA was eluted off the column and treated with a DNase kit (Ambion) to eliminate contamination by genomic DNA. cDNA was synthesized using SuperScript® VILO™ cDNA Synthesis kit (Invitrogen) as outlined in the manual provided by the manufacturer.

2.2.17 Real-time RT-PCR

In order to measure levels of C34E10.8 mRNA in *aus3* mutants and in N2, primers were designed targeting a sequence before and after the premature stop codon. The NCBI primer-blast tool was used to design the following primers. After Premature Stop Codon:

C34E10RT-F_AD (#624): AGCTGTTCGGCGACAAACTCCAG

C34E10RT-R_AD (#625): GTATGCTGGGAGCCCGAAGCC

Before Premature Stop Codon:

C34E10RT-F_BC (#626): TCCACAGCCTCGTTTCCCAGT

C3410RT-R_BC (#627): TGGAACACCGCGTCTCGATGG

Primers for reference gene *ubc-2:*

ubc-2L (#247): AGGGAGGTGTCTTCTTCCTCAC

ubc-2R (#248): CGGATTTGGATCACAGAGCAGC

The reactions were prepared using SYBR® Green Mastermix (Roche) and primers listed were used at a concentration of 125 nM in a final volume of 20 µl. The reactions were processed by the 7500 Fast Real-time PCR System (Applied Biosystems) under the default thermal cycling conditions provided by the Applied Biosystems. The results were exported onto a Microsoft Excel spreadsheet and analysed according to the delta-delta Ct Method (Livak et al., 2001). The C34E10.8 mRNA levels were normalised against the levels of *ubc-2* mRNA. Student's t test was used for statistical analysis.

2.2.18 Isolation of genomic DNA for next-generation sequencing

The genomic DNA was isolated from *aus3* mutants and the Nicholas Lab reference strain according to protocol described by Eric Jorgensen's Laboratory, University of Utah. Worms

were grown on five NGM/OP50 plates (9 cm) and washed off the plates with M9 and collected in a 15 ml tube. The tube containing the mixed population of *aus3* mutants or reference strain in M9 was spun for 30 seconds at 500 g to pellet the worms. Supernatant was removed and pellet was washed in 4ml of TEN solution (20 mM Tris, 50 mM EDTA, 100 mM NaCl, pH 7.5)., The supernatant was removed and the worm pellet was resuspended in 500 µl TEN and frozen at -20 °C overnight. The frozen pellet was thawed and transferred to a 1.5 ml tube. After addition of SDS (0.5 %) and Proteinase K (0.1 mg/ml) the worm pellet was incubated at 55 °C for 1 hour and resuspended at 10, 20 and 30 minutes. The solution was supplemented with more Proteinase K (0.2 mg/ml) after 1 hour and further incubated for 1 hour at 55°C., the solution was then extracted twice with Phenol/Chloroform at a 1:1 volume ratio. Then, Chloroform/Isoamylalcohol (24:1) extraction was performed and DNA was precipitated in 3 M NaOAc and ethanol. The DNA pellet was washed in ice cold 70% ethanol and air dried for 15 minutes. The air dried pellet was resuspended in 500 µl TEN and treated with RNase A (40 mg/ml) for one hour at 37 °C. The solution was extracted with Phenol/Chloroform and Chloroform/Isoamylalcohol (24:1) as described earlier. Genomic DNA was precipitated and washed with ethanol, and then air dried. Finally, the pellet was resuspended in 100 µl sterile distilled water and stored at -20 °C.

2.2.19 Next-generation sequencing and analysis of data

Genomic DNA from *aus3* mutants and the Nicholas Laboratory reference strain were processed by Illumina's GAII Sequencing Service provided by the Australian Genome Research Faciltiy (AGRF). Nicholas Laboratory reference strain was an additional mutant that was isolated from the screen (low expression of *pctbp-1::ctbp-1::gfp* array). This mutant was used as a laboratory specific reference strain for the analysis of next-generation sequencing results. Analysis of the results was performed by the software MAQGene (Bigelow et al.,

2009). MAQGene was set up by Dr Doug Chappell, University of Sydney and results were analysed according to instructions by Bigelow *et al.*, (Bigelow et al., 2009). Briefly, results for *aus3* mutant and Nicholas Lab reference strain were exported as a Microsoft Excel file. The output for MAQGene processed sequencing data yields a list of differences (or variants) between the MAQGene reference sequence (i.e. N2 sequencing data) and the query strain. The mutation was identified by a 75 bp single end run (1 lane) for both *aus3* mutant and our reference strain. The variants of *aus3* mutant on Chromosome III and variants of Nicholas reference strain on Chromosome III were compared and identical variants were eliminated to obtain the variant causative of the *aus3* phenotype. This approach produced 14 candidate genes with different type of mutations (Chapter 5, Table 5.3). The variants identified through MAQGene were confirmed by Sanger sequencing allowing MAQGene sequencing errors to be identified.

2.2.20 Rescue of overexpression phenotype in aus3 mutants

For rescue of CTBP-1::GFP overexpression phenotype in *aus3* mutants, *C34E10.8* was amplified from N2 genomic DNA using primers 595 (forward) and 596 (reverse).

C34E10.8-proF1(#595): GAAAGTTACCTGAACAAATGTCG

C34E10.8-proR (#596): CCTGTTACGTAACACACAAGAGC

The Expand Long Template kit from Roche was used for amplification of C34E10.8. The amplified product was introduced into the gonad of the *aus3* mutant with plasmid pRF4 containing the *rol-6(su1006)* as a co-transformation marker. For control, pRF4 (rol-6) plasmid alone (i.e. without C34E10.8 template) was injected into the *aus3* mutants. 2 lines were generated for both *aus3; C34E10.8*; *rol-6(su1006)* and *aus3; rol-6(su1006)*. The concentration of the templates is 50 ng/µl in all cases.

In order to generate C34E10.8::GFP fusion for expression analysis different primer sets were used:

C34E10.8-proF2 (#597): GACCCACACAACAAGCTACG

hase-1overlap(#688):

AGTCGACCTGCAGGCATGCAAGCTAAGTTCCTCCTCATATTCCTC

hase-1 nested (#689): ACGATCACAGATGAACAG

nestedgfprev (#210): GGAAACAGTTATGTTTGGTATA

GFP was amplified from pPD95.75:

pPD9575gfpf (#182): AGCTTGCATGCCTGCAGGTCG

pPD9575gfpr (#183): AAGGGCCCGTACGGCCGACTA

The fusion product was introduced to N2 strain as described above. Concentration of the template used for injection is 50 ng/µl.

2.2.21 DAPI staining of whole worms

A small plate (5cm) of transgenic worms carrying *pctbp-1::ctbp-1::gfp* array were washed with M9 buffer into a 2 ml microfuge tube. Worms were allowed to settle on ice for 5 min and supernatant was discarded leaving minimum amount of M9 buffer. DAPI was added to the worm pellet at a concentration of 150 ng/ml and worms were incubated at room temperature for 30 minutes. DAPI was removed at the end of incubation and worms were washed with M9 buffer for 20 minutes. The washing step was repeated leaving worm pellet with residual M9 buffer. 10 µl of this pellet was mounted on agarose pads for microscopy.

2.2.22 Oligonucleotide synthesis

Oligonucleotides used in this study were synthesised by Sigma-Aldrich. (http://www.sigmaaldrich.com/configurator/servlet/DesignCenter)

2.2.23 Standard molecular biology techniques

Standard molecular biology techniques were performed according to protocols provided by Sambrook *et al.* (Sambrook, 1989). These techniques include PCR, restriction digestion, agarose and polyacrylamide gel electrophoresis, phenol/chloroform extraction, mini-preparations of plasmid DNA, nuclear extraction.

CHAPTER 3

C. ELEGANS CTBP-1 FUNCTIONS IN THE SPECIFICATION OF A SUBSET OF CHOLINERGIC MOTOR NEURONS

3.1 Introduction

The CtBP family of proteins are transcriptional co-repressors that are conserved from nematodes to mammals. As reviewed in Chapter 1, most of the information on CtBP function has been obtained through experiments on mammalian cell lines. These experiments have shown that CtBPs play roles in a number of biological processes such as the epithelial to mesenchymal transition (EMT), p53-independent apoptosis, and fat metabolism (Grooteclaes and Frisch, 2000, Grooteclaes et al., 2003, Zhang et al., 2003, Kajimura et al., 2008). Deciphering *in vivo* functions of CtBP has been challenging due to redundancy between family members and early embryonic lethality of knockouts in vertebrate systems. As a result of this, much remains to be discovered about the *in vivo* functions of CtBP proteins. In this chapter, we explored the functions of the nematode CTBP-1.

3.2 CTBP-1 is pre-dominantly expressed in the neurons

Towards an understanding of CTBP-1 function, its expression pattern was assessed using a fluorescent reporter construct. To generate this reporter construct, a PCR fusion approach was used. The entire coding sequence of *ctbp-1* together with 2kb upstream of the start site of translation (incorporating the presumptive promoter sequence) was fused to the coding sequence for the green fluorescent protein (gfp) followed by the 3′ untranslated region (UTR) of the nematode *unc-54* gene (Fig. 3.1A). The nematode specific *unc-54* 3′ UTR was used in order to aid the expression of the *pctbp-1::ctbp-1::gfp* transgene. This is a standard approach, with the *unc-54* 3′ UTR being routinely used in nematode transgenes as it is permissive for expression in all somatic cells (Hunt-Newbury et al., 2007). However, since the endogenous 3′ UTR of the *ctbp-1* gene was not included in the construct, micro RNA regulation of *ctbp-1* will not be reflected with this reporter construct. The *pctbp-1::ctbp-1::gfp* fusion product was injected into the gonad of wild type worms and three transgenic lines established. This

method of generating transgenic animals produces strains which contain multiple copies of the injected PCR product as an extrachromosomal array. A preliminary analysis of the three strains expressing carrying the *pctbp-1::ctbp-1::gfp* transgene as an extrachromosomal array found the expression pattern to be equivalent in all three strains (data not shown). Although relatively stable, such extrachromosomal arrays can be lost through both meiosis and mitosis, with the latter resulting in a mosaic expression pattern which can confound analysis. In order to overcome this, one of the extrachromosomal arrays, *ausEx21[pctbp-1::ctbp-1::gfp]* was selected for integration into the genome. Animals carrying *ausEx21[pctbp-1::ctbp-1::gfp]* were exposed to gamma irradiation and 1000 of the second generation self progeny of these animals were screened for homozygous integrants. Through this process, one line carrying the integrated transgene *ausIs1[pctbp-1::ctbp-1::gfp]* was identified.

Analysis of the stable, chromosomally-integrated transgenic strain expressing CTBP-1::GFP (*ausIs1*) has shown that CTBP-1::GFP expression starts at the mid-embryonic stage (Fig 3.1 B). From the three-fold stage of embryogenesis (Fig. 3.1 B), through larval development (Fig. 3.1 C) and into adulthood (Fig. 3.1 E), CTBP-1 expression can be seen in both the head and the tail of the worm, and in a row of cells along the ventral side of the worm. The pattern of expression mirrors the distribution of neuronal cell bodies, which show a clustering in several ganglia in the head and tail, and in the ventral nerve cord (VNC).

We focussed on the VNC and identified the cells expressing CTBP-1 in this region. In newly hatched larvae the VNC is composed of 22 motorneurons. These motor neurons are classified according to the type of neurotransmitter they use for synaptic transmission, with cholinergic neurons named DA and DB, while GABAergic neurons are named DD. There are 9 DA, 7 DB and 6 DD neurons in the VNC at the first larval (L1) stage (White et al., 1986, White et al., 1976). At the L1 stage, robust expression of the CTBP-1::GFP fusion protein was detected in

all cholinergic motor neurons in the VNC (DA and DB, marked in Fig. 3.1C) while low levels of expression were also observed in the GABAergic neurons (DD, not shown).

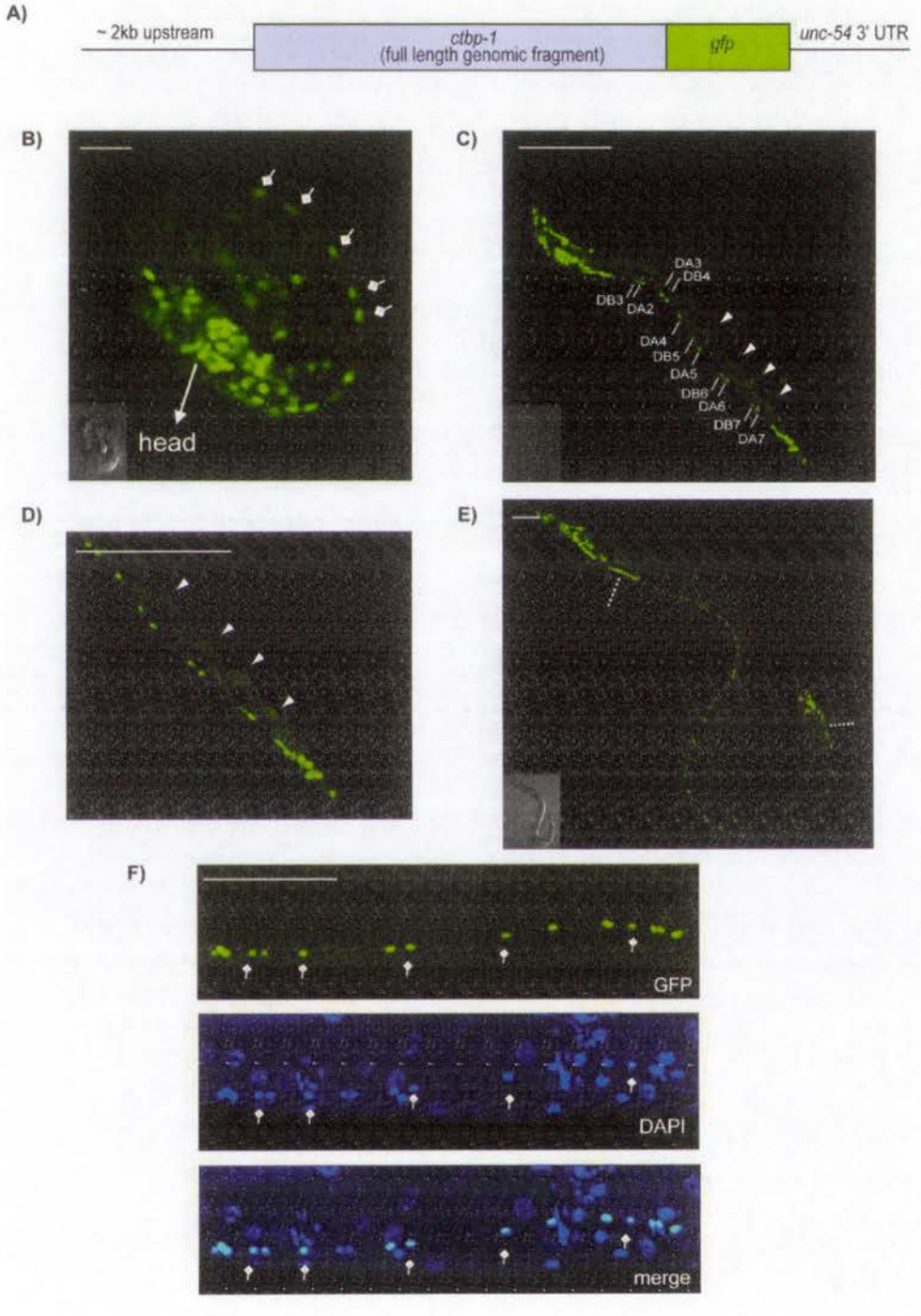


Figure 3.1: CTBP-1 is expressed pre-dominantly in the neurons. A) Schematic representation of the reporter construct used to examine the expression pattern of CTBP-1. The whole coding sequence of *ctbp-1* was amplified using wild type genomic DNA and joined to *gfp* by PCR fusion. Since *ctbp-1* is amplified from genomic DNA the introns are also included in the reporter gene fusion. The *ctbp-1::gfp* fusion is under the control of approximately 2 kb of *ctbp-1* promoter and the *unc-54* 3' UTR. **B)** At the 3-fold stage of embryogenesis expression of the CTBP-1::GFP reporter is observed in many neurons. A subset of VNC neurons are marked (➡) **C)** L1 stage worm expressing CTBP-1::GFP. Cholinergic VNC cells are labelled. Faint expression of CTBP-1::GFP is detected in the hypodermal cells some of which are indicated with arrowheads. **D)** The zoomed view of the worm at L1 stage in panel C corresponding to the posterior region. For clarity, hypodermal cells which are also indicated in panel C are shown (arrow heads). **E)** CTBP-1::GFP is expressed predominantly in the neurons at young adult stage. CTBP-1::GFP is expressed in neurons in the head, tail and Ventral Nerve Cord. The region between the retrovesicular and the pre-anal ganglia is indicated with dashed lines. Note that hypodermal cells are not in focus. **F)** Representative images show subcellular localisation of CTBP-1::GFP. DAPI panel shows nuclear staining and cyan color in merged image indicates co-localisation of CTBP-1::GFP and DAPI. Some of the VNC neurons are marked (➡) in all three images. The inset for panels B, C, and D shows DIC images for the respective images. Scale bar: 10 μm for panel B and 50 μm for panels C, D, E and F. Anterior is to the left.

Later in development, 53 postembryonic motor neurons are added to the VNC circuitry bringing the total number of neurons to 75. With the generation of post-embryonic neurons, the VNC is composed of 8 different classes of motor neurons: embryonic DA, DB and DD motor neurons (22 neurons in total), and post-embryonic VA, VB, VC, VD and AS motor neurons (53 neurons in total) (Von Stetina et al., 2005, White et al., 1976). Out of 75 motor neurons, 59 of them reside between the retrovesicular and pre-anal ganglia in a worm at the adult stage. Our analysis based on counting the number of CTBP-1::GFP expressing cells between the retrovesicular and pre-anal ganglia suggests that CTBP-1 is likely to be expressed in all VNC motor neurons in this region (Fig. 3.1 E). In addition, expression of CTBP-1 in the neurons in the head and in the tail persists through the adult stage as shown in Figure 3.1 E.

Outside the nervous system, a low level of CTBP-1 expression was also detected in hypodermal cells, some of which are indicated in Figure 3.1 C and shown at a larger scale in Figure 3.1 D.

The pattern of CTBP-1 expression suggested that the CTBP-1 protein might be localised to the nucleus. Worms carrying the *ausIs1(pctbp-1::ctbp-1::gfp)* transgene were stained with the DNA-binding dye DAPI and colocalisation of the GFP and DAPI signals detected (Fig. 3.1 F), confirming nuclear localisation of the CTBP-1 protein.

3.3 Mutant form of CTBP-1 has reduced repression activity

In order to investigate the roles of CTBP-1 within the nervous system, we acquired a strain carrying a mutant allele called *ok498* that had been generated by the *C. elegans* gene knockout consortium. The strain contains a 1629 base pair deletion in the *ctbp-1* gene (Fig. 3.2 A and B). Since both deletion endpoints lie within introns, this is an in-frame deletion, and the mutant gene is predicted to encode a truncated protein. While this truncated protein lacks the majority of the dehydrogenase-like domain, including the NAD(H) binding site, the N-terminal THAP domain and C-terminal region remain intact (Fig. 3.2 C and D).

To examine whether the *ok498* allele would be useful for assessing *in vivo* functions of CTBP-1 we investigated whether the repression activity of the truncated form of CTBP-1 encoded by this allele is compromised. It has previously been demonstrated that when fused to the DNA binding domain of the yeast transcription factor Gal4 (Gal4DBD), the wild type form of nematode CTBP-1 can repress the expression of a Gal4-responsive reporter (Nicholas et al., 2008). We therefore tested the capacity of the truncated CTBP-1 encoded by the *ok498* allele to repress reporter gene expression as follows.

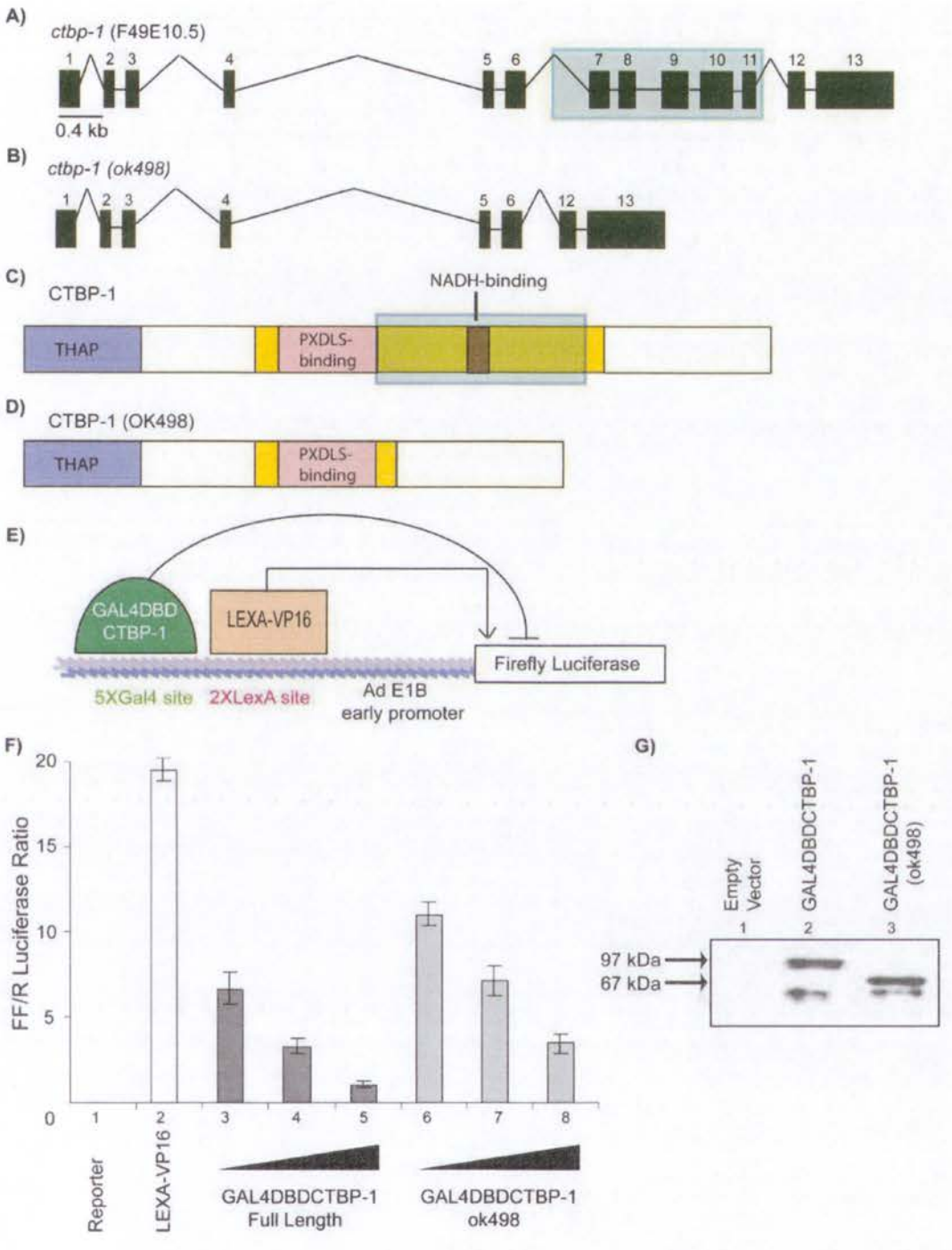


Figure 3.2: Mutation in *ctbp-1* locus leads to reduction in the repression activity of CTBP-1. A) Wild type genomic locus of *ctbp-1* is depicted. The shaded region indicates the portion of the *ctbp-1* gene that is deleted in the *ok498* allele. **B)** The mutant allele of *ctbp-1 (ok498)* is shorter than the wild copy due to the deletion of exons

7-11 and half of 6th intron. **C)** Wild type CTBP-1 protein has 4 domains: THAP, PXDLS-binding, NADH-binding and dehydrogenase (shown in yellow) domains. Deleted region of the protein in *ctbp-1 (ok498)* mutant is highlighted. **D)** Mutant CTBP-1 lacks the NADH-binding domain and a large portion of the dehydrogenase domain. **E)** Schematic representation of the luciferase reporter gene used to assay CTBP-1 repression activity. The firefly luciferase gene is under the control of the Adenoviral E1B early promoter. The regulatory region also includes LEXA binding sites and GAL4DBD binding sites. LEXA-VP16 activates expression of Firefly luciferase and the GAL4DBD enables recruitment of the fused CTBP-1 protein to the promoter. **F)** Gal4DBD-CTBP-1 and Gal4DBD-CTBP-1(ok498) fusion constructs were tested for their ability to repress LEXA–VP16-activated expression of a luciferase reporter in COS-1 cells. COS cells in 6 wells plates were transfected using FuGENE® 6 with 3 μg of (Gal4)5(LexA)2-E1B-Luc reporter plasmid (columns 1-8), 1 μg LexA-VP16 expression plasmid (columns 2-8) and 0.25 μg (column 3), 0.5 μg (column 4) and 1 μg (column 5) of pcDNA3-Gal4DBD-CTBP-1 plasmid and finally 0.25 μg (column 6), 0.5 μg (column 7) and 1 μg (column 8) of pcDNA3-Gal4DBD-CTBP-1(ok498) plasmid. pcDNA3 empty vector was added to make the amount of DNA in each transfection equal. 10 ng of a plasmid expressing Renilla Luciferase (pRL-Luc) was used in each transfection to control for transfection efficiency. Cells were incubated for 48 hours following the transfection. Cells were then harvested and dual luciferase assays were performed to measure the amount of Firefly (FF) and Renilla (R) luciferase expression in each case. The FF luciferase measurement was normalised using the R luciferase measurement and FF/R luciferase ratio is shown. The results shown are biological replicates (n=3) and error bars indicate ± 1 SD for a representative experiment. **G)** Western blots were performed to confirm equivalent expression of the GAL4DBD-CTBP-1 and GAL4DBD-CTBP-1(ok498) proteins. 10 cm petri dishes of COS cells were transfected using FuGENE® 6 with 4 μg pcDNA3 alone (lane 1), 4 μg pcDNA3-Gal4DBD-CTBP-1 (lane 2) and 4 μg pcDNA3-Gal4DBD-CTBP-1(ok498) (lane 3). Cells were incubated for 48 hours following the transfection before cells were harvested and nuclear extracts prepared. Equal amounts of each nuclear extract was run on a 12% SDS PAGE gel, the proteins were transferred to nitrocellulose membrane by Western blot and the GAL4 fusion proteins were visualised using a rabbit anti-Gal4DBD polyclonal antibody. The bands corresponding in size to GAL4DBD-CTBP-1 (97 kDa) and to GAL4DBD-CTBP-1(ok498) (67kDa) are indicated.

A luciferase reporter under the control of a promoter containing two LexA binding sites and five Gal4 binding sites upstream of the adenoviral E1B promoter was activated by expression

of LEXA-VP16 fusion proteins and the effect of expression of a Gal4DBD-CTBP-1(ok498) fusion was measured (Fig. 3.2 E). The Gal4DBD-CTBP-1(ok498) fusion protein repressed expression of the luciferase reporter in a dose-dependent manner. However, compared with the repression achieved by the wild type Gal4DBD-CTBP-1 fusion protein, the repression activity of this truncated version is reduced ~2.5 fold (Fig. 3.2 F). Western blots confirmed equivalent expression of the wild type and mutant Gal4DBD-CTBP-1 fusion proteins (Fig. 3.2 G). These experiments indicate that the repression activity of CTBP-1 is compromised by the deletion that defines the *ok498* allele. Worms carrying the *ctbp-1(ok498)* allele can therefore be used to examine the *in vivo* effects of reduction of CTBP-1 function.

3.4 *ctbp-1(ok498)* mutants show neuronal defects

Given the findings that CTBP-1 is broadly expressed in the nervous system and that CTBP-1 repression activity is compromised by the *ok498* mutation, we used *ctbp-1(ok498)* mutants to investigate whether reduced CTBP-1 activity leads to abnormalities in neuronal development. To this end we first looked at any gross phenotypes that might be indicative of neuronal defects and noted that the *ctbp-1(ok498)* mutants display mildly uncoordinated backward movement. In *C. elegans*, locomotion is coordinated by several classes of neurons including the embryonic motor neurons of the ventral nerve cord (DA, DB, DD). Since CTBP-1 is expressed in these embryonic motor neurons and since *ctbp-1(ok498)* mutants are defective in their movement, we investigated whether these cells are appropriately specified in the *ctbp-1(ok498)* mutant worms.

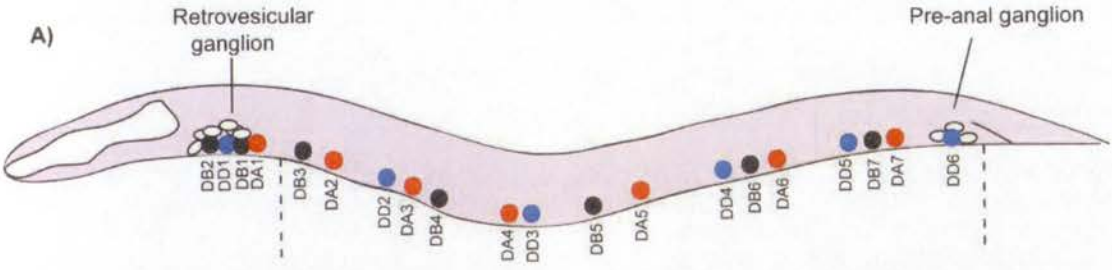
The specification of the subset of embryonic motorneurons which lie between the retrovesicular and pre-anal ganglia was examined. At hatching, 15 motorneurons are aligned in this region: five DB type neurons (DB3-DB7), five DD type neurons (DD2-DD6), and five DA type neurons (DA2-DA7) (Fig. 3.3 A). These three types of neurons are characterised by expression

of ACR-5, UNC-25 and UNC-4, respectively. We therefore used fluorescent reporters, ACR-5::YFP, UNC-25::GFP and UNC-4::GFP, to assess the specification of the embryonic motor neurons in wild type and *ctbp-1(ok498)* mutant worms.

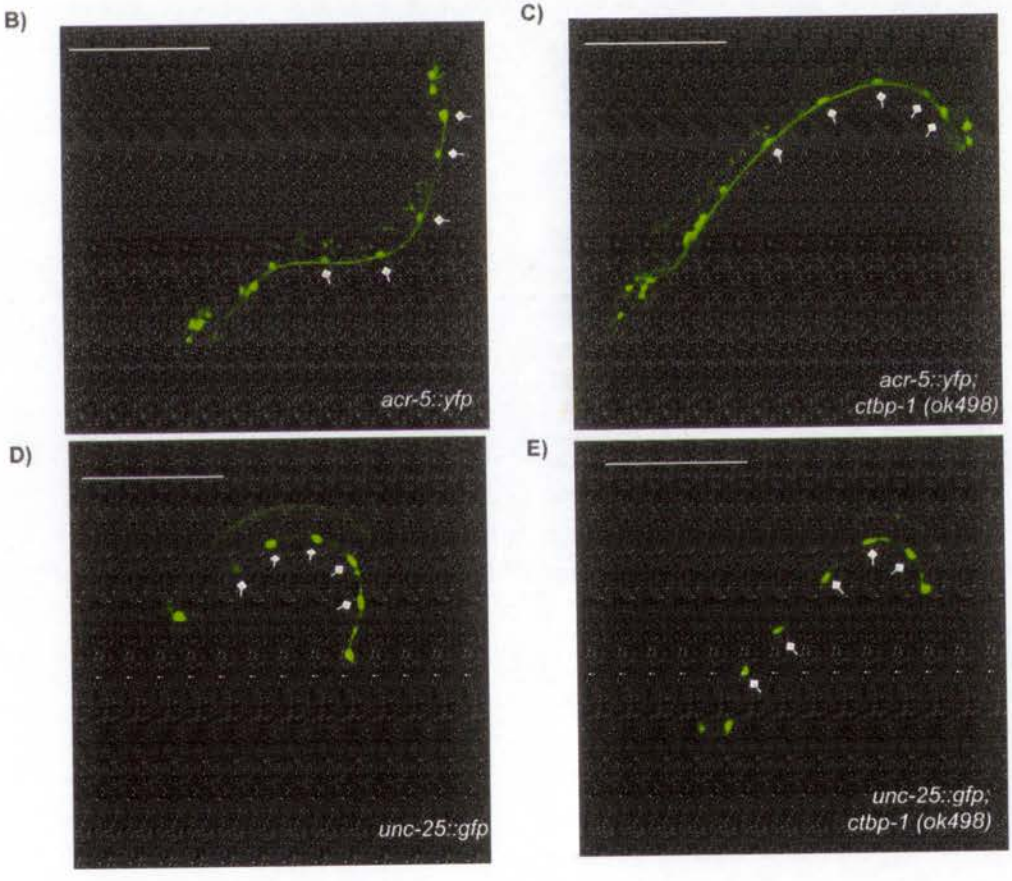
The DB type cholinergic neurons were examined using the ACR-5::YFP reporter (Christensen et al., 2002). In all wild type (n=50) and *ctbp-1(ok498)* (n=50) mutant animals examined, five *acr-5::yfp*-expressing cells were identified (Fig. 3.3 B and C). That is, reduction of CTBP-1 function does not affect expression of the DB marker ACR-5.

UNC-25::GFP marks the DD type GABAergic cells (Jin et al., 1999). For analysis of UNC-25::GFP expression, DD6 which resides in the pre-anal ganglion was included. In 50 *ctbp-1(ok498)* mutant animals scored, all showed the same expression pattern as wild type (n=50). That is, each had five UNC-25::GFP-expressing cells (Fig. 3.3 D and E). This indicates that full CTBP-1 function is not required for the correct expression of the DD marker UNC-25.

The DA type cholinergic neurons of the VNC were examined using the UNC-4::GFP reporter. In wild type animals five or six UNC-4::GFP-expressing cells were observed in 53% of the examined population (n=105) (Fig. 3D). The remainder of the population showed expression of UNC-4::GFP in fewer cells with the following frequencies: 4 cells (29%), 3 cells (12%), 2 cells (3%) and 1 cell (3%). In a *ctbp-1(ok498)* mutant population, a significant reduction in the number of UNC-4::GFP-expressing cells was observed; only 21% of examined animals (n=132) showed five or six UNC-4::GFP-expressing cells while 31% of animals showed expression in 4 cells, 21% in 3 cells, 10% in 2 cells, 12% in 1 cell, and 5% showed no GFP expression in the examined region (Fig. 3.3 F and G).



DB3-DB7	ACR-5::YFP
DD2-DD6	UNC-25::GFP
DA2-DA7	UNC-4:::GFP



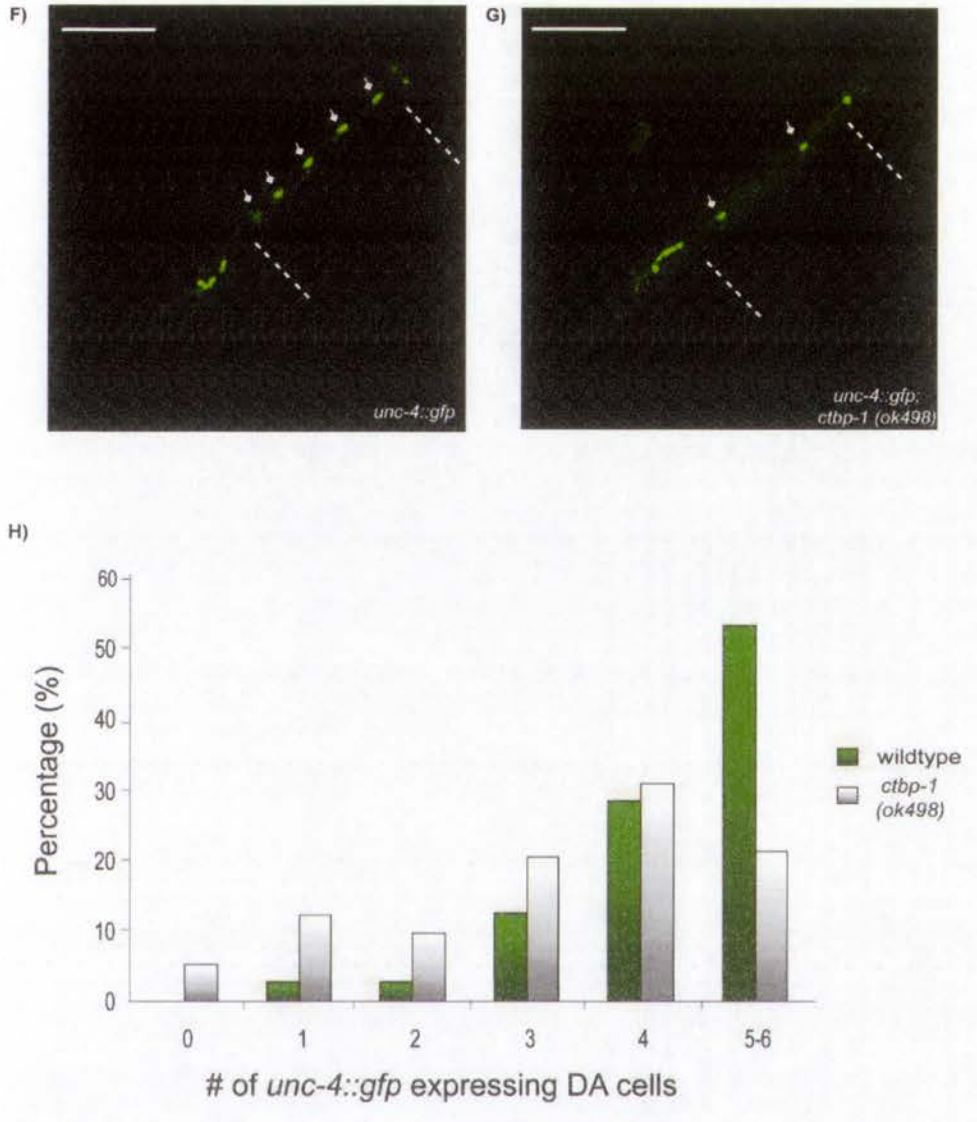


Figure 3.3: *ctbp-1(ok498)* mutants show defects in DA class neurons. A) The schematic shows Ventral Nerve Cord motor neurons of a hermaphrodite at the L1 stage. (Adapted from Sulston *et al.*, 1983). The analysed region is indicated with dashed lines. For markers ACR-5::YFP and UNC-4::GFP the analysed region resides between the neurons in the head, termed the retrovesicular ganglion, to the neurons in the tail termed the pre-anal ganglion. However, for the UNC-25::GFP marker, the DD6 cell in the pre-anal ganglion is included in the analysis **B)** Representative images show a hermaphrodite at the L1 stage expressing the ACR-5::YFP marker in wild type **C)** and *ctbp-1(ok498)* mutant background. Neurons DB3-DB7 are marked (➸). **D)** Representative images show a hermaphrodite at L1 stage expressing UNC-25::GFP in wild type **E)** and *ctbp-1(ok498)* mutants background. Neurons corresponding to DD2-DD6 are marked (➸). n=50 for both ACR-5::YFP and UNC-

25::GFP expressing wild type and mutant animals. **F)** Representative images showing the UNC-4::GFP marker in wild type, and **G)** *ctbp-1(ok498)* mutant background. The region analysed is labelled with dashed lines and DA cells are individually marked (➡◆). **H)** The plot is a graph of number of animals in percentage versus number of fluorescent nuclei expressing UNC-4::GFP in the specified number of cells in the analysed region (DA2-DA7) in wild type (n=105) and *ctbp-1(ok498)* mutants (n=132). Scale bar is 50μm for all panels. Anterior is to the left.

Having identified a defect in the expression of the DA marker, *unc-4::gfp*, in *ctbp-1(ok498)* mutant worms, we investigated whether the observed reduction in the number of cells expressing UNC-4::GFP is due to a genuine reduction in the number of DA cells, or whether the DA cells are present but do not express the DA marker *unc-4::gfp*. To distinguish between these two possibilities the development of all VNC cells was examined using the pan-neuronal reporter *F25B3.3::gfp* (Altun-Gultekin et al., 2001). Scoring 50 animals at L1 developmental stage, 15 cells in the VNC were apparent in both wild type and *ctbp-1(ok498)* mutant backgrounds (Fig. 3.4 A and B). This suggests that the VNC cells, including the DA neurons, are intact in the *ctbp-1(ok498)* mutant but fail to express UNC-4::GFP.

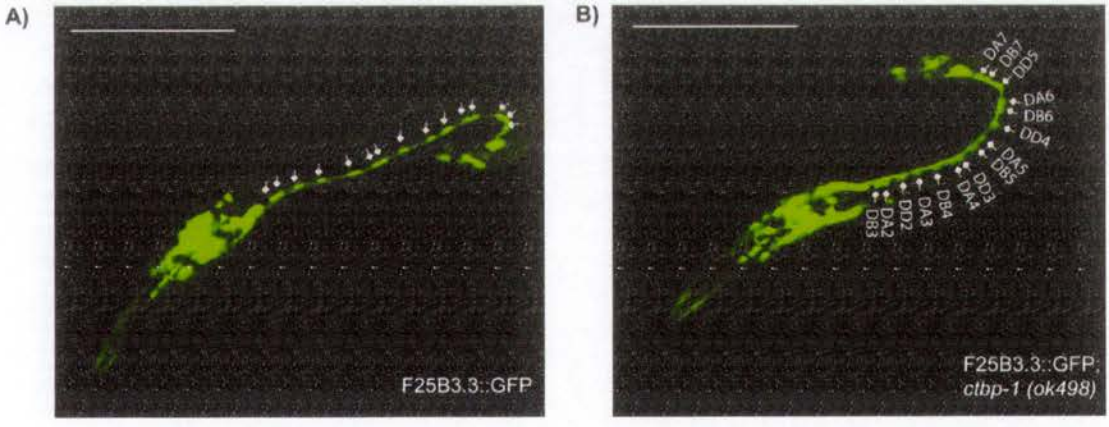


Figure 3.4: The *ctbp-1(ok498)* mutants do not lack embryonic Ventral Nerve Cord neurons. Representative images of worms at the L1 stage expressing **A)** a pan-neural marker (F25B3.3::GFP) in wild type **B)** and *ctbp-1*

mutant background. VNC motorneurons between the retrovesicular and pre-anal ganglion are labelled. The *ctbp-1* mutants express F25B3.3::GFP pan-neural marker in all VNC neurons like the wild type animal. n= 50 for both panels. Scale bar is 50 µm. Anterior is to the left.

This raised the question of whether the observed defect in the *ctbp-1(ok498)* mutant was specific to the UNC-4::GFP marker. Therefore, an UNC-3::GFP marker, which is expressed in both DA and DB neurons, was used to further examine the integrity of the DA neurons in the *ctbp-1(ok498)* mutants. In 50 animals at L1 developmental stage, nine UNC-3::GFP-expressing cells in the examined region of the VNC were apparent in both wild type and *ctbp-1(ok498)* mutant animals (Fig. 3.5 A and B). As the UNC-3::GFP marker is expressed in DB as well as DA neurons, this observation supports the earlier conclusion that the DB neurons are intact in the *ctbp-1(ok498)* mutants as indicated by wild type-like expression of the ACR-5::YFP reporter. More importantly, this observation suggests that although one characteristic of the DA neurons (UNC-4 expression) is affected by reduction of CTBP-1 function, at least one other characteristic of these neurons (UNC-3 expression) is not dependent on full CTBP-1 activity.

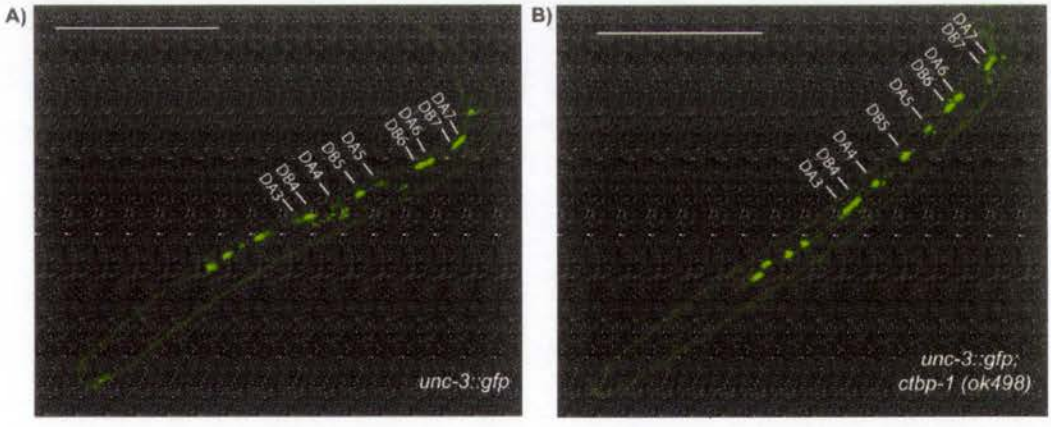


Figure 3.5: The *ctbp-1(ok498)* mutants do not show defects in expression of UNC-3::GFP reporter which marks DA and DB neurons at the L1 stage. Representative images of worms at the L1 stage expressing the

UNC-3::GFP reporter **A)** in wild type **B)** and *ctbp-1(ok498)* mutant background. UNC-3::GFP marks both DA and DB type neurons in the VNC where *ctbp-1(ok498)* mutants do not display a defect in neither DA nor DB neurons with this marker. The region analysed spans from DA3 to DA7 since in wild type animals UNC-3:GFP expression is not always detected in DA2. n= 50 for both panels. Scale bar is 50 μm. Anterior is to the left.

In summary, reduction of CTBP-1 activity was not observed to affect the specification of the DB or DD motor neurons, but was found to alter the expression of a DA neuron-specific marker, suggesting that CTBP-1 may be required to define class-specific traits of the DA type embryonic motor neurons.

3.5 CTBP-1 interacts with neuronal transcription factor ZAG-1

As CTBP-1 is a transcriptional co-repressor protein, we next sought to indentify transcription factors that CTBP-1 might work with to control the differentiation of DA motor neurons. To this end yeast two hybrid screens were performed in our laboratory (P-S Khoo, unpublished) using the full length CTBP-1 as a bait. In addition to screening against a nematode mixed-stage cDNA library, a screen was also performed against a murine erythroleukaemia cDNA library. Given the high degree of similarity between the nematode CtBP and murine CtBPs, and, in particular, the conservation of a key interaction surface (the PXDLS-binding pocket) it was reasoned that screening a murine library might enable the identification of additional relevant interactors. From these screens, 8 positive clones were identified (4 from the *C. elegans* library screen and 4 from the murine library screen) and the cDNA inserts were sequenced. One of those from the murine screen was identified as a partial cDNA encoding residues 580-747 of murine δEF1. δEF is a transcription factor that contains three PXDLS motifs and has been shown to interact with murine CtBP (Furusawa et al., 1999, Postigo and Dean, 1999). The partial clone identified from the yeast two hybrid screen contains all three of these PXDLS motifs (Fig. 3.6 A).

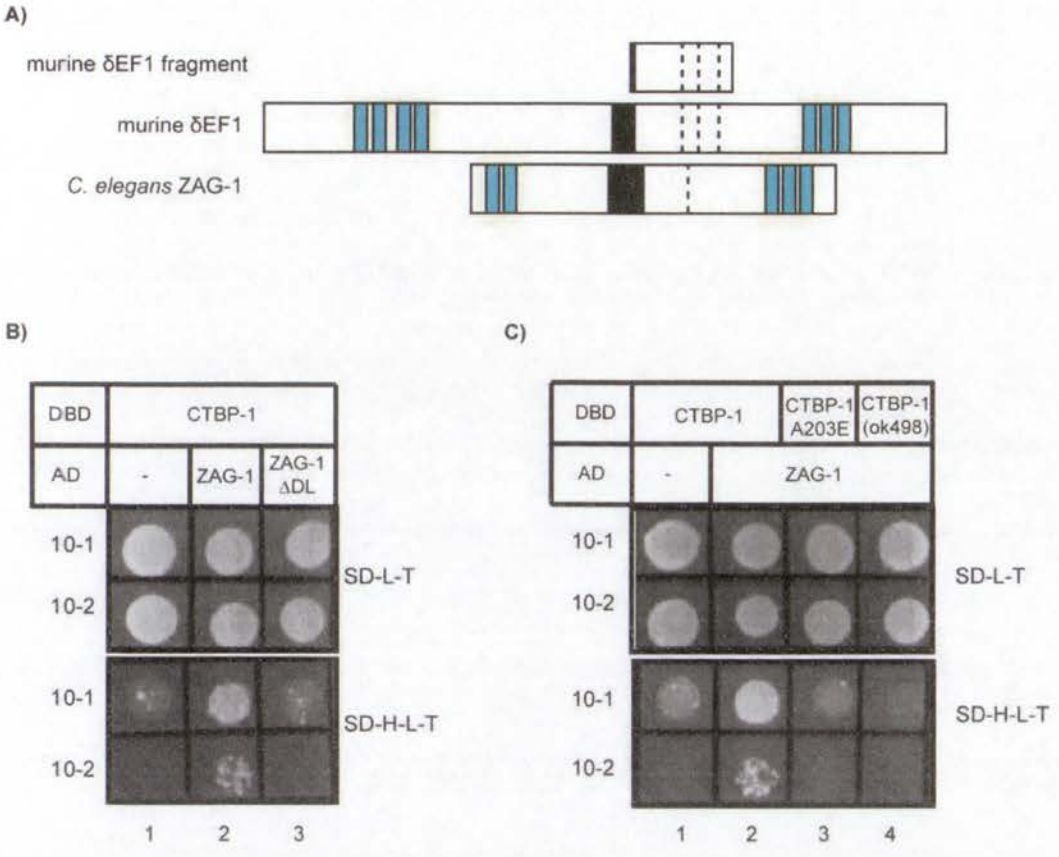


Figure 3.6: CTBP-1 interacts with neuronal transcription factor ZAG-1. A) A yeast two hybrid screen for proteins that interact with nematode CTBP-1 identified a partial cDNA encoding residues 580-747 of murine δEF1 (upper and middle panels). The nematode homologue of murine δEF1 is ZAG-1 (lower panel). Zinc finger domains are shown in blue, homeodomains are shown in black, and putative CtBP interaction motifs are shown by dashed lines. **B and C)** Yeast two-hybrid assays were performed to examine the interaction of CTBP-1 with ZAG-1. The assays were performed by fusing the test proteins to the C-terminus of either the Gal4 DNA binding domain (DBD) or the Gal4 activation domain (AD). Transformants were selected on SD-L-T plates. Colonies were picked from these plates and incubated overnight in liquid SD-L-T media. The cultures were normalised to an OD600 of 1 and serial dilutions were made. 10 ul was then pipetted on to SD-L-T plates and SD-H-L-T, which were incubated at 30 °C for 72 hours. Growth on SD-H-L-T plates is indicative of a physical interaction between the test proteins. The interaction of CTBP-1 and ZAG-1 was assessed (B column 2). The PXDLS-like motif in ZAG-1 was mutated and the mutant protein (ZAG-1 ΔDL) was tested for its capacity to interact with CTBP-1 (B column 3). The dependence of the interaction between CTBP-1 and ZAG-1 on the PXDLS-binding cleft of CTBP-1 was assessed by mutating one residue within the putative PXDLS-binding cleft and testing the

capacity of the mutant protein (CTBP-1 A203E) to interact with ZAG-1 (C column 3). The interaction between the mutant version of CTBP-1 encoded by the *ok498* allele (CTBP-1(ok498)) and ZAG-1 was also assessed (C column 4).

A nematode homologue of δEF1 has been identified called ZAG-1 (Fig. 3.6 A). This nematode protein has one PXDLS motif and has previously been hypothesised to be an interacting partner of CtBP (Wacker et al., 2003b, Clark and Chiu, 2003). We therefore tested whether CTBP-1 can indeed interact with ZAG-1 using a yeast two-hybrid assay. As shown in Figure 3.6 B column 2, CTBP-1 was confirmed to interact with ZAG-1 in this assay.

Having demonstrated that ZAG-1 and CTBP-1 are physical interacting partners, we next sought to assess whether this interaction is dependent on the PXDLS motif of ZAG-1. The PXDLS motif was mutated and the capacity of the mutant form of ZAG-1, called ZAG-1ΔDL, to interact with CTBP-1 was assessed. No interaction was detected (Fig. 3.6 B, column 3). To corroborate this finding, interaction of the wild type ZAG-1 protein with a mutant version of CTBP-1 in which the PXDLS binding pocket is disrupted was tested. No interaction was observed in this assay (Fig. 3.6 C, column 3), confirming the dependence of the ZAG-1/CTBP-1 interaction of the PXDLS motif. Together these data indicate that ZAG-1 and CTBP-1 could be functional partners in *C. elegans*.

If indeed ZAG-1 and CTBP-1 are functionally interacting partners in the nematode, it was of interest to assess whether this interaction would be disrupted in the ctbp-*1(ok498)* mutant. As described, animals carrying this mutant allele are predicted to express a truncated form of CTBP-1 lacking amino acids 217-486. Although most of the critical residues of the PXDLS binding pocket are still present in this truncated form of CTBP-1 (amino acids 180, 182, 193, 201, 203, 204, 205 and 214), one of the interaction residues (217) is deleted. Furthermore,

comparison with the structural elements surrounding this pocket (from mammalian structures) indicates that the structural scaffold that supports the PXDLS binding pocket is likely to be disrupted in the *ok498* truncated protein (Nardini et al., 2003). Consistent with this, the truncated version does not interact with ZAG-1 (Fig. 3.6 C, column 4). Importantly, this truncated form of CTBP-1 retains interaction with other non-PXDLS-dependent interactors (P-S Khoo, unpublished) indicating that the fusion protein is expressed and stable, and confirming that the failure of the ZAG-1 interaction is indeed due to removal of a key interaction surface. Since the interaction of ZAG-1 with the mutant version of CTBP-1 is compromised, we conclude that control of gene expression by ZAG-1 could be compromised in the *ctbp-1(ok498)* mutant strain.

3.6 *ctbp-1* and *zag-1* show expression in similar type of neurons

If CTBP-1 and ZAG-1 physically interact to regulate gene expression, these proteins must be present in the same cells. Previous reporter gene analysis has shown that ZAG-1 is expressed predominantly in neurons starting at the mid embryonic stage (Wacker et al., 2003a). To complement this published work, we used an expression profile database which catalogues expression profiles of *C. elegans* genes (Spencer et al., 2011) and compared the expression pattern of *ctbp-1* and *zag-1* (Fig. 3.7). The expression profiles are derived from cell type-specific microarray experiments and show for each gene the level of expression in various tissues at different developmental stages. Tissue specific expression is achieved through promoters that drive expression of *gfp* in the pertinent tissue types. For instance, the expression profile of LE GABAergic neurons was determined by extracting RNA from cells expressing the *punc-25::gfp(juIs56)* reporter, while the expression profile of LE A-class neurons on the basis of the promoter used, that is *punc-25::gfp(juIs56)* reporter for the formerwas determined by extracting RNA from cells expressing the *punc-4::unc-*

4::gfp(wdIs5) reporter (Spencer et al., 2011). The expression intensity for each cell/tissue type is normalised against the relevant developmental reference which in the end produces a pattern of expression showing peaks in some tissues indicating higher levels of expression and troughs for lower levels of expression. In Figure 3.7, the graph illustrates the relative expression intensity of *ctbp-1* and *zag-1* in various tissue types. *zag-1* and *ctbp-1* show a similar trend of expression during development from late embryonic (LE) to young adult (YA) stage as depicted at the right end of the graph (Fig. 3.7). Although, *ctbp-1* and *zag-1* are

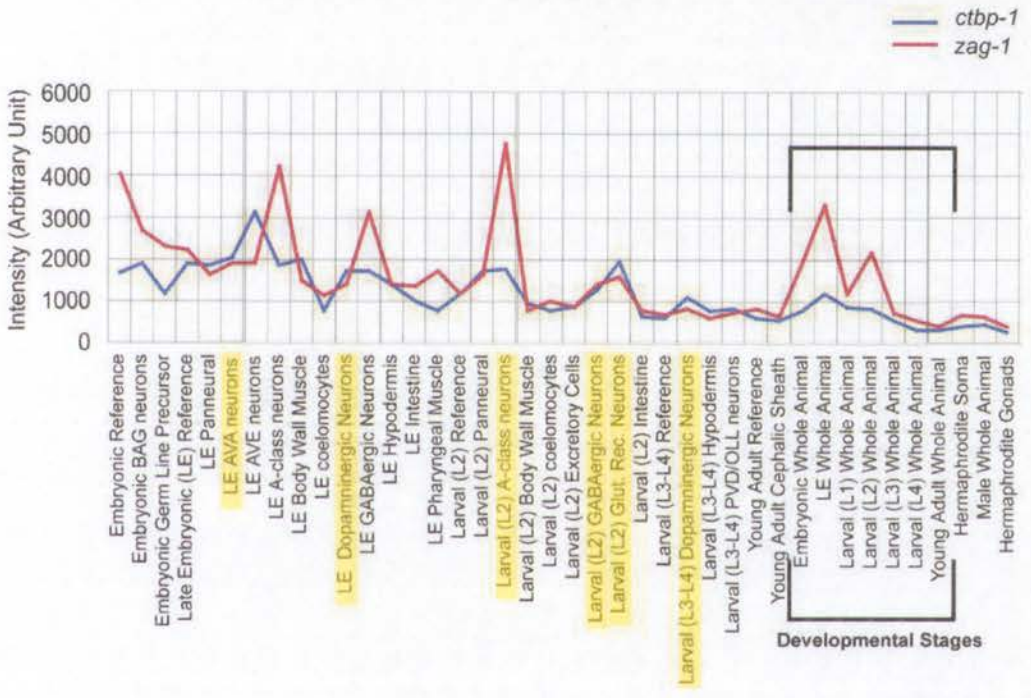


Figure 3.7: *ctbp-1* and *zag-1* show a similar expression pattern throughout development and enrichment in the same type of tissues. The graph shows relative expression intensity of *ctbp-1* and *zag-1* plotted against the tissue types and developmental stages. The data are obtained from a website Wormviz which includes results from a recently published paper on a *C. elegans* gene expression map (Spencer et al., 2011). The intensity of expression is determined by normalisation of expression levels of each gene in the relevant tissue against an appropriate reference samples. For instance, for embryonic stage the reference is Embryonic Reference. Embryonic neurons that are cultured overnight are called late embryonic and the reference for these data sets is Late Embryonic

Reference. A similar referencing system is used for all later developmental stages. The tissue types where both *ctbp-1* and *zag-1* are highly expressed are shaded in yellow. These tissues are AVA and dopaminergic neurons at late embryonic stage, A-class, GABAergic and glutamate receptor neurons at L2 stage and dopaminergic neurons at L3-L4 stage. Abbreviations as follows: LE: Late Embryonic, YA: Young Adult and Glut. Rec. Neurons: Glutamate receptor neurons.

co-expressed in some of the tissue types according to the cell-specific microarray data (Fig. 3.7, yellow-highlighted tissues), the expression profiles do not match in all cell types analysed. Cell types in which both genes are relatively highly expressed include embryonic GABAergic neurons, and postembryonic Glutamergic and Dopaminergic neurons as well as A-class neurons (DA and VA). Of particular interest to this study is the observation that *zag-1* and *ctbp-1* are both detected in larval A-class neurons which include the DA neurons, making it possible for these two proteins to work together to regulate gene expression in DA type neurons. That *zag-1* and *ctbp-1* are also co-expressed in other neurons suggests that these two proteins may also act together to control the development or function of additional neurons beyond those that have been investigated in this study. It should be noted, however, that these data represent only expression of the *zag-1* and *ctbp-1* transcripts and the relative expression levels may not correlate with the level of ZAG-1 and CTBP-1 proteins expressed by the examined cell types. In particular, both show a peak of expression during late embryogenesis (LE whole animal in Figure 3.7). In terms of specific tissues, both genes are relatively highly expressed in embryonic GABAergic neurons, and postembryonic Glutamergic and Dopaminergic neurons as well as A-class neurons (DA and VA). In particular interest for this study, *zag-1* and *ctbp-1* are both detected in larval A-class neurons which include the DA neurons, making it possible for these two proteins to work together to regulate gene expression in DA type neurons. That *zag-1* and *ctbp-1* are also co-expressed in other neurons suggests that these two proteins may also act together to control the development or function

of additional neurons beyond those that have been investigated in this study. However, since the graph represents expression patterns at transcriptional level, it is of question whether the mRNA expression pattern correlates with the expression pattern at the protein level.

3.7 Reduction of ZAG-1 function affects expression of the DA marker UNC-4::GFP

Expression of *zag-1* and *ctbp-1* in the same type of neurons according to cell-specific microarray experiments and lack of interaction between ZAG-1 and CTBP-1(ok498) mutant protein have suggested that *ctbp-1(ok498)* mutants and *zag-1* mutants might have similar defects. Indeed, at the gross phenotypic level, both *zag-1* and *ctbp-1(ok498)* mutant animals display a backwards kinker uncoordinated phenotype. Interestingly, similar to *ctbp-1(ok498)* mutants, zag-*1(rh315)* mutants were reported to display a defect in their DA neurons (Wacker et al., 2003a). In that report, the same DA marker (UNC-4::GFP) that has been used in the current study was used but, in contrast to our analysis which scored DA2-DA7, neurons corresponding to DA3-DA9 were scored. By this method, Wacker *et al.* found that *zag-1(rh315)* mutants had fewer UNC-4::GFP expressing DA cells than wild type worms. We took an initiative to re-analyse the neurons in *zag-1(rh315)* mutants using the scoring system that had been applied to the *ctbp-1(ok498)* mutants. That is, expressing of UNC-4::GFP in DA2-DA7 was examined. In this analysis, 34 *zag-1(rh315)* mutant animals expressing UNC-4::GFP were scored in which 15% of the population showed UNC-4::GFP expression in 5-6 DA cells (Fig. 3.8 C) as opposed to 53% of the population in wild type animals expressing UNC-4::GFP (see Section 3.4 and Appendix Table 1). Moreover, 23% of the *zag-1(rh315)* population had no UNC-4::GFP expressing DA cells in the analysed region while none of the wild type animals lacked all UNC-4::GFP expressing DA cells. The remainder of the *zag-1(rh315)* mutant population showed the following percentages for the indicated number of UNC-4::GFP expressing DA cells: 6% of the population had 4, 12% had 3, 23% had 2 UNC-4::GFP expressing DA cells and finally 21% had 1 UNC-4::GFP expressing DA cell.

Representative images in Figure 3.8 A and B show wild type and *zag-1(rh315)* mutant animals expressing UNC-4::GFP, respectively.

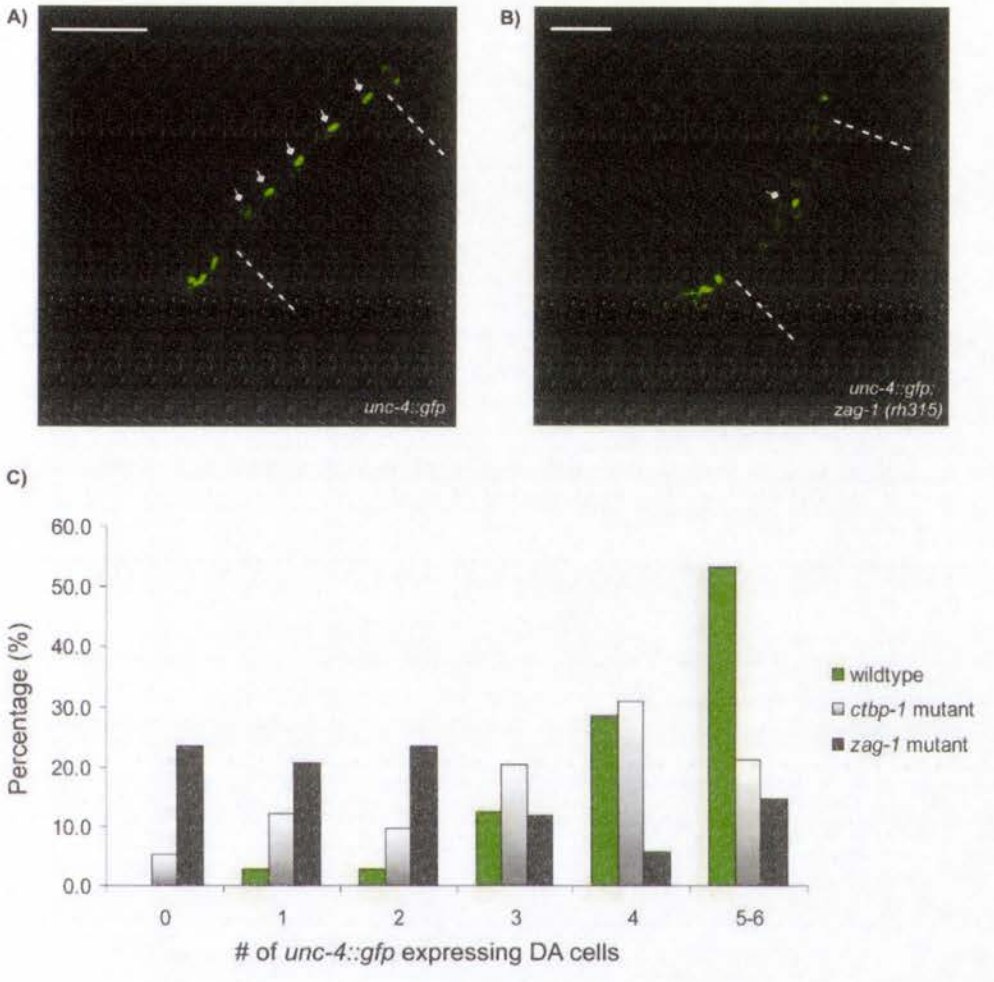


Figure 3.8: The *zag-1(rh315)* mutants show a similar DA defect observed in *ctbp-1(ok498)* mutants. Representative images of **A)** wild type **B)** and *zag-1(rh315)* mutant animals expressing UNC-4::GFP. Animals are at the L1 stage. The region analysed is indicated by dashed lines. **C)** The plot is a graph of number of animals in percentage versus number of fluorescent nuclei expressing UNC-4::GFP in the analysed region (DA2-DA7) in wild type, *zag-1(rh315)* and *ctbp-1(ok498)* mutants. To generate this graph, the data of wild type and *ctbp-1(ok498)* mutants from graph in Figure 3.3 H were used. The sample size for *zag-1* mutants is 34. Scale bar is 50 µm. Anterior is to the left.

The above analysis shows that, similar to *ctbp-1(ok498)* mutants, *zag-1(rh315)* mutants have fewer UNC-4::GFP expressing DA neurons than wild type. While 70% of the *zag-1(rh315)* mutant animals and 30% of the *ctbp-1(ok498)* mutants showed UNC-4::GFP expression in two or fewer DA cells, only 6% of the wild type population shows this pattern of UNC-4::GFP expression. Also similar to *ctbp-1(ok498)* mutants, *zag-1(rh315)* mutants have been reported to show wild-type like expression of a pan-neural marker at the L1 stage (Wacker et al., 2003a), suggesting that the alteration in expression of UNC-4::GFP is not a result of absence of DA cells. Like in the case of CTBP-1, these observations suggest that full ZAG-1 activity is not required for the generation of the DA neurons, but is required for the correct acquisition of at least one of the class-specific traits of the DA type embryonic motor neurons.

3.8 Discussion

Results detailed in this chapter have shown that *C. elegans* CTBP-1 is pre-dominantly expressed in the nervous system and that CTBP-1 activity is required for correct expression of a cell-specific reporter in DA-type cholinergic neurons. Furthermore, it has been demonstrated that CTBP-1 physically interacts with the neuron-specific transcription factor ZAG-1, and that *zag-1(rh315)* mutant animals display a defect in the DA cells that is similar to that of *ctbp-1(ok498)* mutants. These findings suggest that CTBP-1 acts as a co-repressor for the ZAG-1 transcription factor and that together these proteins control gene expression to confer class-specific traits on the DA-type cholinergic motor neurons of *C. elegans*. As detailed in Chapter 1, other members of the CtBP family of proteins have previously been reported to have neuronal functions as inferred from developmental defects in knockout mice and flies. Additionally, neuronal roles for the vertebrate homologues of ZAG-1 have also been described. For instance, like the CtBP KO mouse which is defective in neural tube development, mice carrying a null mutation in the ZAG-1 homologue SIP1/δEF2/ZEB2 show

a neural tube closure defect (Hildebrand and Soriano, 2002; Miyoshi et al., 2006; Van de Putte et al., 2003). Moreover, in Xenopus, SIP1/δEF2/ZEB2 represses BMP4 (Bone Morphogenetic Protein 4) through interaction with XCtBP thereby inducing neuronal differentiation in Xenopus (Nitta et al., 2007; van Grunsven et al., 2007). The findings presented in this chapter thus provide further support for an evolutionarily conserved role for the CtBP proteins in neuronal fate specification.

Many factors involved in the development and connectivity of neurons are evolutionarily conserved. With its well-defined nervous system, studies in *C. elegans* have provided fruitful insights into how neurons are generated and specified (Hobert, 2010). The first step towards generation of a neuron is a cell fate decision which allows for selection of neuronal vs non-neuronal fate. Neurons then terminally differentiate, acquiring class-specific characteristics that confer functionality. Some features that characterise specific neurons are their cell body placement, synaptic connectivity and axodendritic morphology (Hobert, 2010).

The DA motor neurons which were the focus of the investigations reported here are terminally differentiated neurons with several characterised traits including that they express the neurotransmitter acetylcholine and that they send commissures from the ventral nerve cord, where the cell bodies reside, to the dorsal nerve cord, from where they innervate the body wall muscles (White et al., 1976). These two features are shared with the DB motor neurons, however, the DA and DB neurons also each possess distinct characteristics. For instance, while the axons of the DA-type neurons turn anteriorly after reaching the dorsal side, those of the DB-type neurons turn posteriorly (White et al., 1976).

Several factors have been identified which are involved in the specification of DA cells, conferring on them the properties that are different from DB-type cells. One of these factors is

the homeodomain transcription factor UNC-4 (Miller et al., 1993). In DA motor neurons, UNC-4 works with the Groucho homolog UNC-37 to negatively regulate certain genes that are characteristic of DB-type neurons, such as *acr-5* (encoding an acetylcholine receptor) and *del-1*(encoding a DEG-type ion channel) (Pflugrad et al., 1997; Winnier et al., 1999). On the other hand, in DB cells, *unc-4* expression is repressed by the nematode eve homolog VAB-7, indirectly leading to the activation of *acr-5* (Esmaeili et al., 2002).

Another factor that is important for the specification of the DA and DB motor neurons is the COE transcription factor UNC-3. This protein acts as a terminal selector for A and B type cholinergic motor neurons including embryonic DA and DB type cells (Prasad et al., 1998; Von Stetina et al., 2005). Recently, UNC-3 was reported to function upstream of UNC-4 and VAB-7 and to activate these subtype-specific transcriptional repressors for diversification of

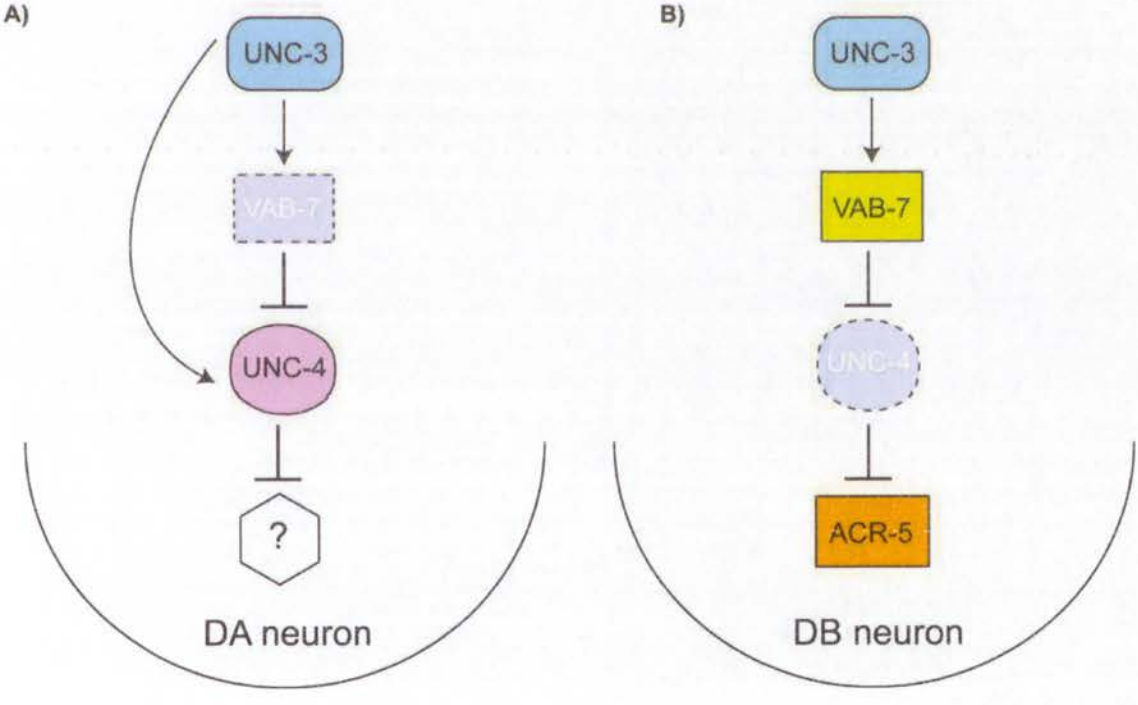


Figure 3.9: The UNC-3 transcription factor works as a terminal selector in embryonically derived cholinergic DA and DB neurons. A) UNC-3 activates UNC-4 and indirectly represses DB type traits in DA

neurons. The terminal differentiation genes that are regulated in this pathway are yet to be identified. **B)** UNC-3 activates DB-specific VAB-7 which represses UNC-4 and this way prevents DB cells from acquiring a DA fate. This way, one of the terminal differentiation genes, *acr-5*, that in DA cells is repressed by UNC-4 is activated, allowing DB cells to adopt DB cell characteristics.

DA and DB type neurons (Kratsios et al., 2011). According to the model proposed by Kratsios *et al.*, UNC-3 activates UNC-4 in DA cells and thereby induces repression of a subset of terminal differentiation genes such as *acr-5* and others that are yet to be identified (Fig. 3.9 A). The model further suggests that in DB cells, UNC-3 activates VAB-7 which in turn suppresses UNC-4 and enables expression of the DB-specific terminal differentiation gene *acr-5* (Fig. 3.9 B).

In this chapter, we demonstrated that *ctbp-1(ok498)* mutants have fewer UNC-4::GFP expressing DA cells than wild type, indicating that CTBP-1 activity is required for the correct expression of *unc-4* in DA cells. In contrast, we found that UNC-3::GFP expression is not affected in the absence of fully functioning CTBP-1. These observations suggest that CTBP-1 most likely acts downstream of UNC-3 and upstream of UNC-4. Given that CTBP-1 functions as a transcriptional repressor, genes that are directly regulated by CTBP-1 would be expected to show increased expression when CTBP-1 activity is compromised as in the *ctbp-1(ok498)* mutant. Since instead we observed an absence of UNC-4::GFP expression in a subset of DA neurons in *ctbp-1(ok498)* mutants, UNC-4 is not likely to be directly regulated by CTBP-1. It is possible, however, that UNC-4 is an indirect target of CTBP-1. For instance, CTBP-1 might function to repress the expression of some other factor which in turn represses UNC-4 expression in DA cells (Fig. 3.10). In the case where CTBP-1 activity is compromised, as in the *ctbp-1(ok498)* mutant, repression of this putative repressor of UNC-4 would be relieved, leading to reduced levels of UNC-4 expression in DA cells. Based on the

model proposed by Kratsios *et al.*, the putative repressor of UNC-4 expression that is invoked in our model of CTBP-1 function described above could be VAB-7. If CTBP-1 usually functions to repress VAB-7 in the DA cells, reduction of CTBP-1 function would lead to upregulation of VAB-7, which in turn would repress expression of UNC-4. Further work will be required to assess the validity of the above model and to understand the precise role played by CTBP-1 in the acquisition of DA-type motor neuron characteristics. For instance, the prediction that arises from our model is that expression of VAB-7 would be upregulated when CTBP-1 function is reduced. This could be tested by introducing a *vab-7* reporter such as *stIs10140[pvab-7::mcherry]* into the *ctbp-1(ok498)* mutant background. Normally expressed in the DB neurons, ectopic expression of the *vab-7* reporter in the DA neurons of *ctbp-1(ok498)* mutants would support the proposed model. In addition, comparison of *vab-7* transcript levels in *ctbp-1(ok498)* mutant and wild type neurons, particularly in DA neurons, might provide further insight into whether CTBP-1 targets *vab-7*. Moreover, chromatin immunoprecipitation (ChIP) analysis will provide an answer to whether *vab-7* is directly targeted by CTBP-1. It will also be of great interest to identify downstream effects of reduced UNC-4 levels (i.e., which terminal differentiation genes are affected). In addition, given the centrality of UNC-3 to the acquisition of both DA and DB-type characteristics, the possibility that UNC-3 might control the expression of *ctbp-1* should also be investigated. This could readily be achieved by introducing the CTBP-1::GFP reporter into *unc-3* mutants to see whether its expression level is reduced compared with wild type.

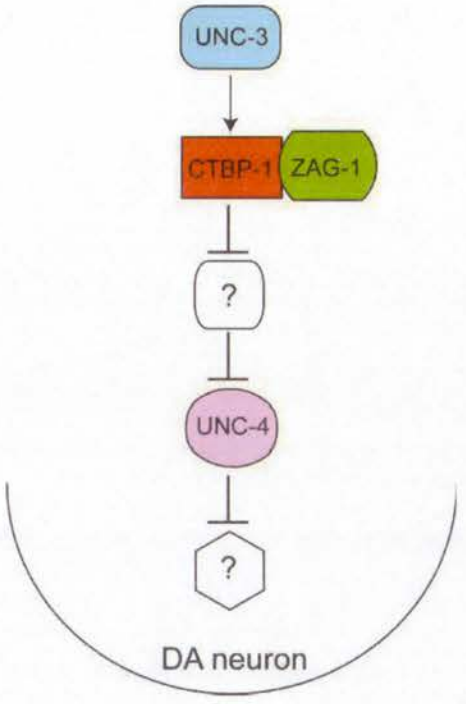


Figure 3.10: A model of how CTBP-1 and ZAG-1 might regulate specification of DA neurons. CTBP-1 might be activated by UNC-3 and by some mechanism activate UNC-4 activities in the DA neurons. This could be via an adaptor protein that works as an UNC-4 repressor. In this scenario, CTBP-1 inhibits expression of a novel UNC-4 repressor via interaction with ZAG-1 and this way indirectly activates UNC-4. UNC-4 activity is required for correct expression of terminal differentiation genes. CTBP-1 and ZAG-1 might therefore indirectly regulate expression of these genes.

Although the precise contribution made by CTBP-1 to the acquisition of DA-type characteristics is as yet unresolved, the data presented in this chapter suggest that CTBP-1 works together with the transcription factor ZAG-1 to perform this role. However, it is interesting that the DA defect displayed by the *zag-1(rh315)* mutants is more pronounced compared with the *ctbp-1(ok498)* mutants. These differences may be attributable to differences in the severity of each of the mutant alleles tested, since it is likely that neither of these alleles is null. In the case of *ctbp-1(ok498)*, our repression assays suggest that the truncated CTBP-1 protein encoded by the *ok498* allele retains significant repression activity

(Fig. 3.2 F). Although recruitment of CtBPs to promoters is usually dependent on interaction with PXDLS-containing transcription factors and our yeast two hybrid assays indicated that interaction with the PXDLS-containing transcription factor ZAG-1 is abolished by the *ok498* mutation (Fig. 3.6 C), the mutant version of CTBP-1 may still be recruited to target promoters through the THAP domain which is unaffected by the *ok498* allele. The THAP domain is a recently-characterised sequence-specific DNA binding domain (Roussigne et al., 2003). It is possible that a *ctbp-1* mutant animal with a stronger allele might exhibit more penetrant specification defects in DA neurons. The *zag-1(rh315)* allele is also not a null allele, since null alleles of *zag-1* cause early larval lethality but animals that are homozygous for the *rh315* allele develop to adulthood (Wacker et al., 2003a). Indeed, the mutant version of ZAG-1 encoded by the *rh315* allele has an intact PXDLS sequence (Wacker et al., 2003a), suggesting that this truncated version of ZAG-1 may retain its ability to recruit CTBP-1.

In addition to the DA defect reported here, if CTBP-1 and ZAG-1 are working together as proposed above, the *ctbp-1(ok498)* and *zag-1(rh315)* mutants might display additional similar defects. Wacker *et al.* (2003a) reported that *zag-1(rh315)* mutants show a subtle defect in expression of UNC-3::GFP in their DA and DB cells where there are fewer UNC-3::GFP-expressing DA and DB neurons in *zag-1(rh315)* mutants (a shift from 14-15 UNC-3::GFP-marked cells to 11-13 in *zag-1(rh315)* mutants). However, *ctbp-1(ok498)* mutants did not show a defect in the expression of UNC-3::GFP. It is possible that UNC-3::GFP expression might have been affected with a stronger *ctbp-1* allele. On the other hand, it is also possible that ZAG-1 has CTBP-1-independent activities similar to the vertebrate homologue of ZAG-1, δEF1, which is reported to have CtBP-independent repression activities (Sekido et al., 1997). In addition to the neuronal specification defects described above, *zag-1* mutants are also reported to have axonal guidance defects (Clark and Chiu, 2003; Wacker et al., 2003a). It

would therefore be of interest to explore whether CTBP-1 also plays a role in regulating axonal guidance.

In addition to the interaction with ZAG-1 reported here, our group has shown that CTBP-1 interacts with two other neuronal transcription factors, zinc-finger protein PAG-3 (Nicholas et al., 2008) and Sox domain transcription factor EGL-13 (P Khoo, unpublished). While the former is involved in regulation of larval motor neuron fate, the latter is reported to play roles in cell fate determination of BAG sensory neurons (Cameron et al., 2002; Cinar et al., 2003; Jia et al., 1996; Jia et al., 1997; Petersen Gramstrup, 2011). Given these interactions, and the broad neuronal expression of CTBP-1 described here, it is likely that CTBP-1 functions in additional aspects of neuronal regulation that are yet to be explored.

In summary, our results define a role for CtBP in the specification of terminally differentiated neurons. Specifically, we propose that CTBP-1 interacts with ZAG-1 to define class-specific traits of DA motor neurons. Target gene analysis will elucidate precisely how CTBP-1 functions to ensure the correct specification of DA neurons.

CHAPTER 4

IDENTIFICATION OF NEURONAL TARGETS OF CTBP-1

4.1 Introduction

In order to understand better the roles of the nematode CtBP, it was of interest to uncover those genes that are regulated by this transcriptional co-repressor. Previous studies both in our laboratory (T. Sibbritt and A. Reid unpublished) and by Chen et al. (Chen et al., 2009), have used microarray analysis of transcripts derived from wild type and *ctbp-1* mutant nematodes at the young adult stage to identify genes whose correct regulation is dependent on *ctbp-1*. Although these experiments have identified numerous genes that are misregulated when CTBP-1 activity is compromised, there are several limitations to these studies. For instance, by analysing transcripts from whole animals many of the changes in gene expression that were detected are likely to arise indirectly as a downstream result of changes in expression of genes whose promoters are directly regulated by CTBP-1. Furthermore, since these experiments were performed at the young adult stage they would not have detected any changes in transcripts that are regulated in a CTBP-1-dependent manner during development. As presented in the preceding chapter, CTBP-1 was found to be expressed pre-dominantly in many neurons and *ctbp-1(ok498)* mutants were shown to display a specification defect in a subset of cholinergic motor neurons in the VNC. In light of these observations, we sought to identify genes that are regulated in a CTBP-1-dependent manner in the neurons alone using a method called Microarray Profiling of *C. elegans* Cells (MAPCeL) (Fox et al., 2005). While the specific goal of these studies was to uncover the genes that are regulated by CTBP-1 in the cholinergic motor neurons, given the broad neuronal expression of CTBP-1, these experiments were designed to identify the entire neuronal complement of CTBP-1-regulated genes. In this way, not only did we aim to understand the requirement for CTBP-1 for the correct specification of the cholinergic motor neurons, but also to gain insights into additional roles of CTBP-1 in other neurons.

4.2 Embryonic neurons were isolated from *ctbp-1* mutant and wild type animals using MAPCeL Method

Due to its small size (1mm at adult stage), it has been challenging to perform cell-specific gene expression profiling in *C. elegans* until the recent development of the MAPCeL method (Fox et al., 2005). This method involves tagging specific cells of interest with a fluorescent marker, then isolating these cells via Fluorescence Activated Cell Sorting (FACS). Transcripts from these sorted cells are then analysed using tiling microarrays (Fig 4.1).

We used this strategy to identify neuron-specific target genes of CTBP-1. Embryonic neurons were isolated from wild type (Chapter 3 Fig 3.4 A) and *ctbp-1(ok498)* mutant (Chapter 3 Fig 3.4 B) animals each expressing the pan-neural marker F25B3.3::GFP. To do this, embryos were isolated from gravid adults by bleaching and embryonic cells were dissociated via disruption of the egg shell. The dissociated embryonic cells were then cultured overnight to grow them to an appropriate stage for the intended analysis. Then, F25B3.3::GFP marked embryonic neurons were isolated via FACS and collected for isolation and amplification of total RNA. As a control, all embryonic cells (i.e., GFP positive and non-GFP cells) were also collected for the purpose of normalisation in the analysis of differentially expressed transcripts in the neurons. That is, the level of expression of a transcript in the neuronal samples was normalised against its level of expression in all embryonic cells to enable the identification of neuronally-enriched transcripts (see Chapter 2, Section 2.2.12). In addition, the data set obtained from all wild type and *ctbp-1(ok498)* mutant embryonic cells could be used for identification of genes that are regulated in a CTBP-1-dependent manner in the entire embryo, adding to our understanding of CTBP-1-dependent control of gene regulation both in the nervous system and in other tissues.

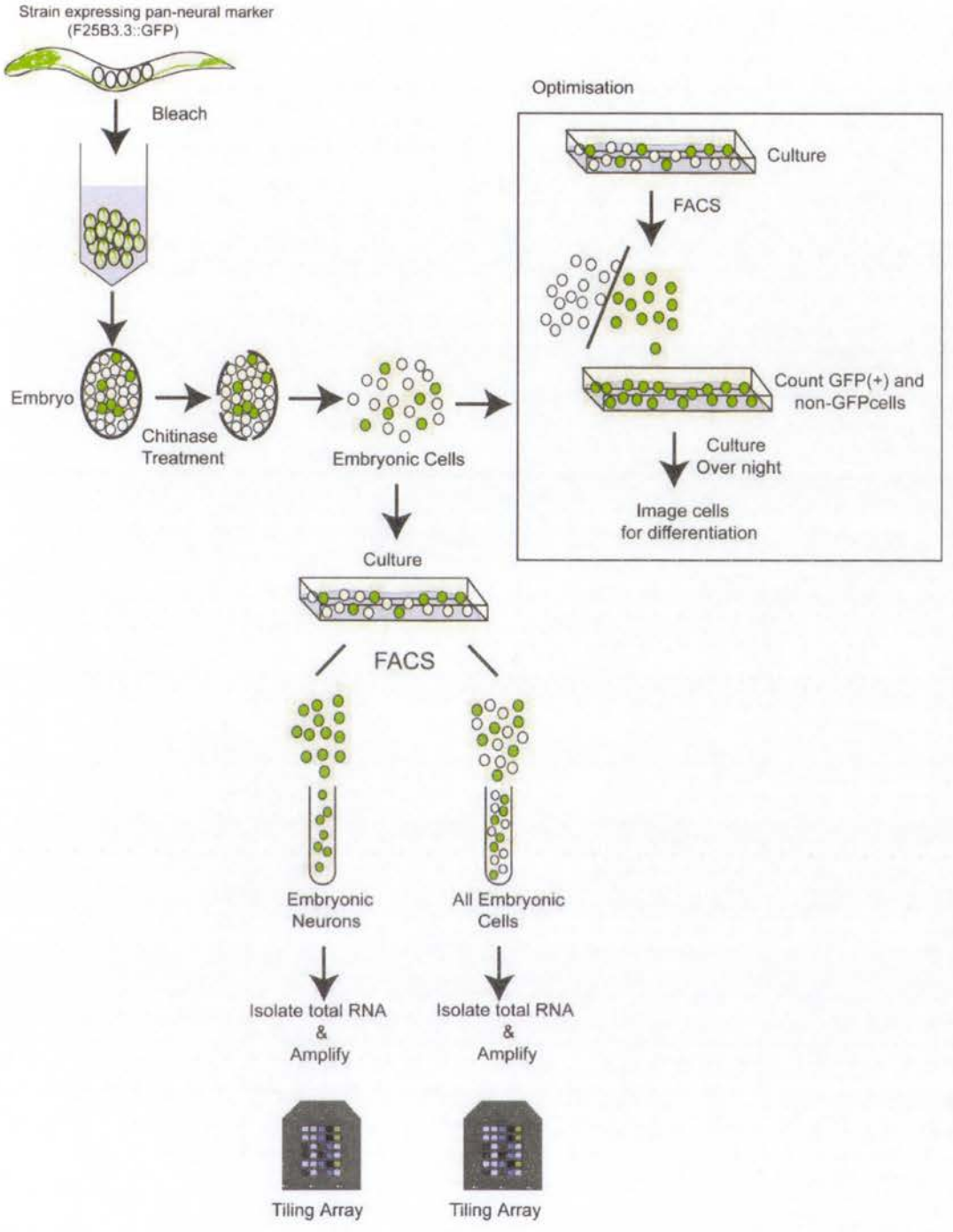


Figure 4.1: Microarray Profiling of C. elegans Cells (MAPCeL) method allows for identification of differentially expressed transcripts in a tissue-specific manner. MAPCeL makes use of transgenic strains expressing GFP in particular type of cells and isolation of these cells via FACS. Here the MAPCeL strategy for neurons is depicted showing isolation of embryonic neurons from transgenic animals expressing the pan-neural

marker F25B3.3::GFP. The method was first optimised as outlined where optimum FACS parameters were set for isolation of GFP positive cells. Gravid adults were bleached to obtain embryos. Then embryonic cells were dissociated by disruption of the egg shell with the help of the enzyme chitinase. Dissociated embryonic cells were cultured overnight to allow for expression of genes turned on later in embryogenesis. GFP-expressing cells were collected via FACS and further cultured to count GFP positive cells and to check that cells differentiate into neurons. With the optimised FACS parameters, three independent worm cultures for both wild type and *ctbp-1(ok498)* mutants were used for isolation of neurons. After bleaching, dissociated embryonic cells were cultured overnight and neurons were isolated via FACS with the set parameters. All embryonic cells were also processed through the FACS to ensure that pan-neural and all cells samples were treated in the same way. The illustration shows only the preparation of samples from the wild type strain expressing the pan-neural GFP marker, however, samples from *ctbp-1(ok498)* mutants expressing the pan-neural GFP marker were processed and collected at the same time. Total RNA was extracted from neurons and all embryonic cells, amplified, and labelled for hybridisation to tiling arrays.

In MAPCeL, isolation of GFP-marked cells via FACS is technically challenging; non-GFP cells occasionally adhere to the GFP cells thus decreasing the purity of the sample. Therefore, before preparing the experimental samples that would be used for microarray analysis, the conditions for FACS were first optimised by adjusting the gate to exclude cell clumps. To assess the purity of the resultant sample, the isolated cells were collected in a culture dish and examined by fluorescence microscopy as shown in Figure 4.1. With the gate adjusted as shown in Figure 4.2 A, 90% of the cell population isolated by FACS from both wild type and *ctbp-1(ok498)* mutants were found to be GFP positive (Fig 4.2 B). The isolated cells were further tested to assess whether they differentiate into neurons after overnight culture. As indicated by the growth of processes that are characteristic of neurons (Figure 4.2 C, D and E), differentiation of the FACS-sorted cells into neurons was confirmed.

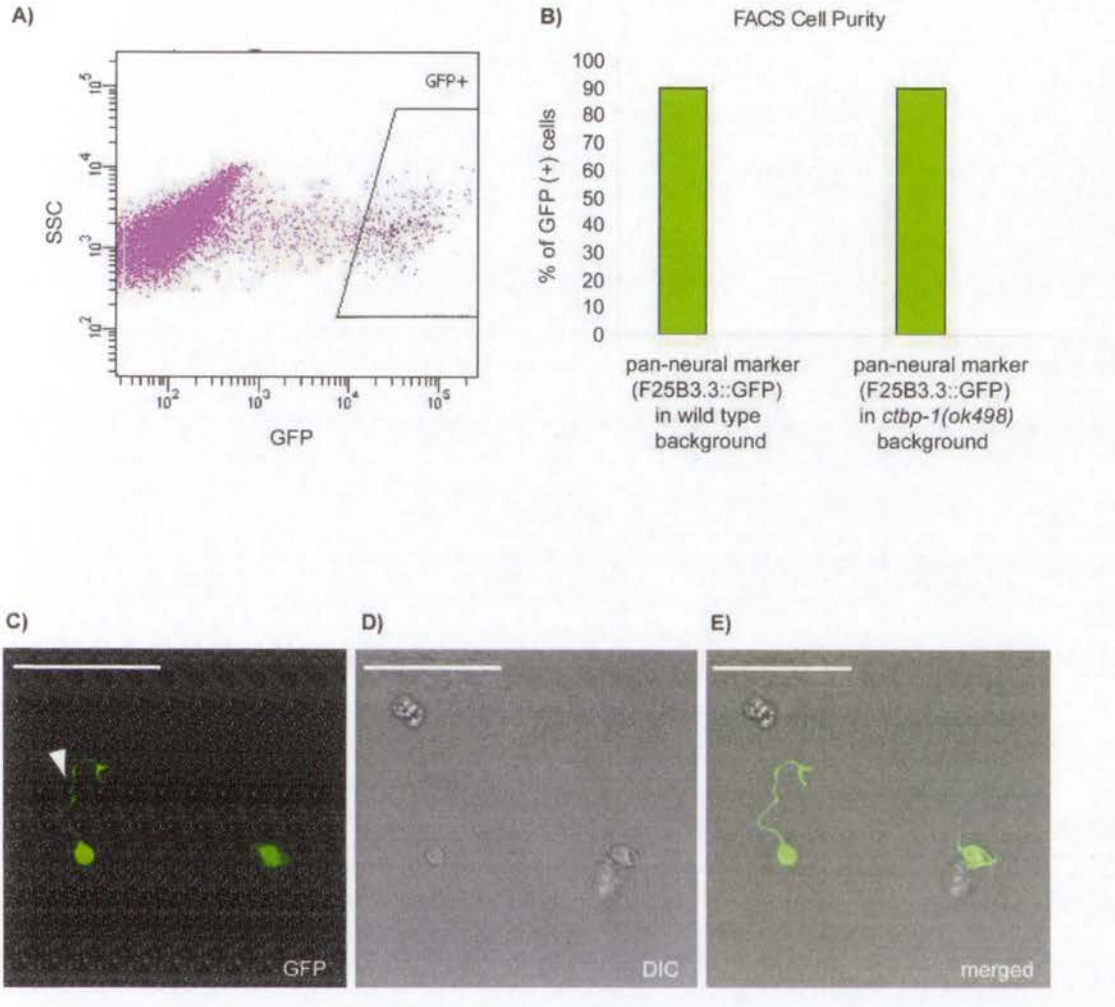


Figure 4.2: Neurons are isolated with high purity. A) The graph is a plot of Side Scatter (SSC) versus log of GFP intensity showing the width of the gate chosen for FACS. **B)** With the gate adjusted as in panel A, the percentage of cells expressing GFP in both *ctbp-1(ok498)* mutants and wild type strains was 90%. Total number of cells scored was 370 for wild type and 426 for *ctbp-1(ok498)* mutants. **C)** Cells isolated via optimised FACS conditions are allowed to differentiate over night. Isolated cells acquire neuron-like characteristics as shown with the processes (labelled with arrow head) sent out of by the cell on the left. **D)** DIC image **E)** and merged image of cells are shown. Scale bar indicates 7.5 μm.

After assessing the purity of the sample and confirming that isolated cells exhibit neuron-like features, three independent worm cultures were used for isolation of wild type and *ctbp-1(ok498)* mutant neurons with these adjusted parameters. As shown in Figure 4.1, for these experimental samples, the isolated cells were collected in a tube for extraction and amplification of RNA followed with labelling and hybridisation to Affymetrix Tiling arrays. In addition to neurons, all embryonic cells were collected through FACS to ensure that both neurons and all cells were treated in the same way. Altogether, with triplicates of neurons and all cells from wild types and *ctbp-1(ok498)* mutants, 12 samples were analysed.

4.3 Microarray analysis confirms successful isolation of neurons from wild type and *ctbp-1(ok498)* mutant embryos

MAPCeL has been extensively used for the profiling of tissue-specific gene expression patterns in *C. elegans* allowing a spatial and temporal map of gene expression to be constructed through analysis of various tissues at different developmental stages. As part of these earlier studies, transcripts that are enriched in the nervous system of *C. elegans* have been published previously (Spencer et al., 2011; Von Stetina et al., 2007). Therefore, before embarking on a detailed analysis of genes that are misregulated in the neurons as a result of reduction of *ctbp-1* function, we first identified transcripts that are enriched in the neurons of wild type and *ctbp-1(ok498)* mutants and compared these with the published list of neuronally-enriched transcripts. For identification of pan-neurally-enriched transcripts in our data sets, we considered those that show a $\geq$2 fold change compared with the expression level in all embryonic cells and required that this fold change have an adjusted p-value of ≤ 0.05. With this statistical approach, in the wild type data set 1445 pan-neurally-enriched transcripts were detected and in the *ctbp-1(ok498)* mutant data set 1407 pan-neurally-enriched transcripts were detected.

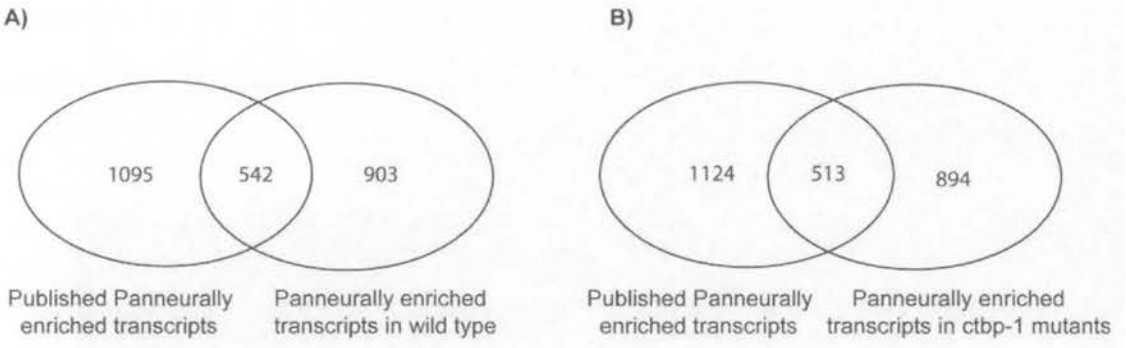


Figure 4.3: Pan-neurally enriched transcripts correlate with the published list. Pan-neurally expressed genes with fold change of ≥ 2 and adjusted p-value of ≤ 0.05 were compared with the published pan-neurally enriched data. The pan-neurally enriched transcripts in wild type samples (A) and in *ctbp-1(ok498)* mutant samples (B) show considerable overlap with the published (Von Stetina et al., 2007) pan-neurally enriched transcripts.

These lists were then compared with the published list of 1637 pan-neurally enriched transcripts in a wild type animal expressing the pan-neural marker F25B3.3::GFP (Von Stetina et al., 2007). The enrichment of pan-neurally expressed transcripts in our experiment correlates with the published results as shown in Figure 4.3. Out of 1637 published pan-neurally enriched genes, 542 of them overlap with our wild type pan-neurally enriched data (Fig. 4.3 A) and 513 of them overlap with the *ctbp-1(ok498)* mutant pan-neurally enriched data (Fig. 4.3 B). Overall, the pan-neurally enriched genes in our experiments constitute approximately 33% of the published pan-neurally enriched genes.

4.4 Levels of *ctbp-1* transcript corresponding to deleted exons in *ctbp-1(ok498)* mutants is 8 times lower than the levels in wild type samples

As another means of validating the microarray data, the level of the *ctbp-1* transcript in wild type and *ctbp-1(ok498)* mutant samples was compared. As reviewed in Chapter 3 (Section 3.3

and Fig. 3.2 A and B) *ctbp-1(ok498)* mutants harbor a deletion that spans a large portion of the *ctbp-1* locus including exons 7, 8, 9, 10 and 11. Therefore, the level of *ctbp-1* signal for the deleted exons in *ctbp-1(ok498)* mutant samples was compared with wild type samples for both the neuron only and all embryonic cells data sets.

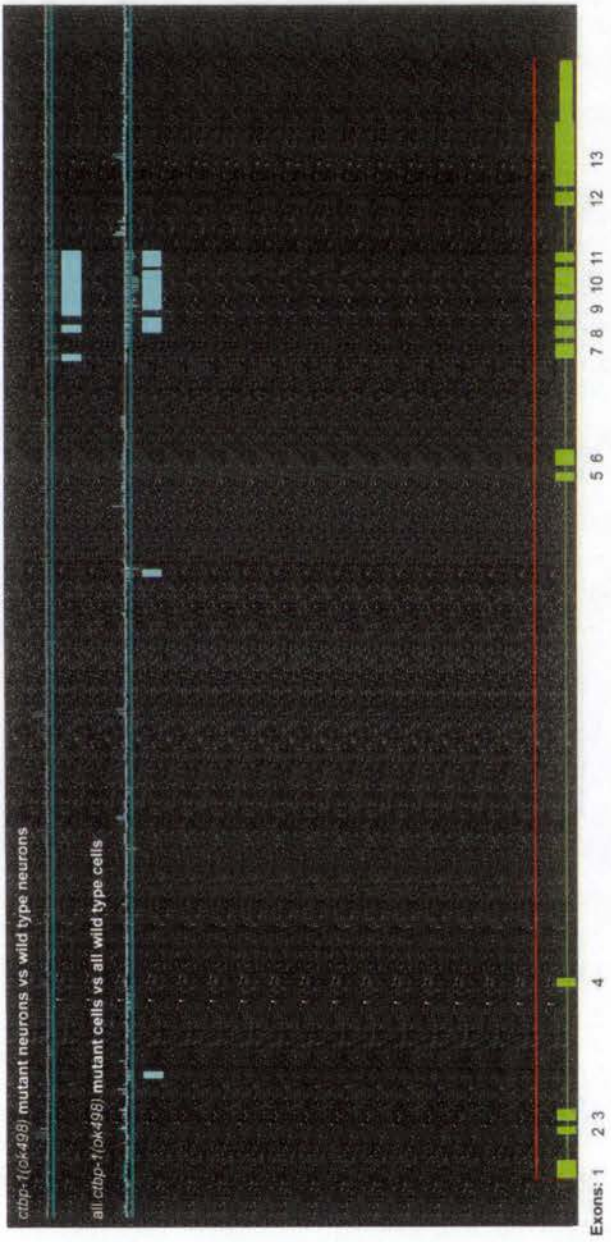


Figure 4.4: In *ctbp-1(ok498)* mutant samples the intensity of *ctbp-1* expression corresponding to deleted exons is much lower than wild type expression levels. The *ok498* allele is a deletion allele covering exons 7, 8, 9, 10, 11. Blue bars indicate an 8-fold decrease in *ctbp-1* expression levels.

As expected, in *ctbp-1(ok498)* mutant samples, the signals for these exons were found to be lower than the signals in the corresponding wild type samples (Fig. 4.4). As a control, the expression levels of exons of the *ubc-2* gene were compared.

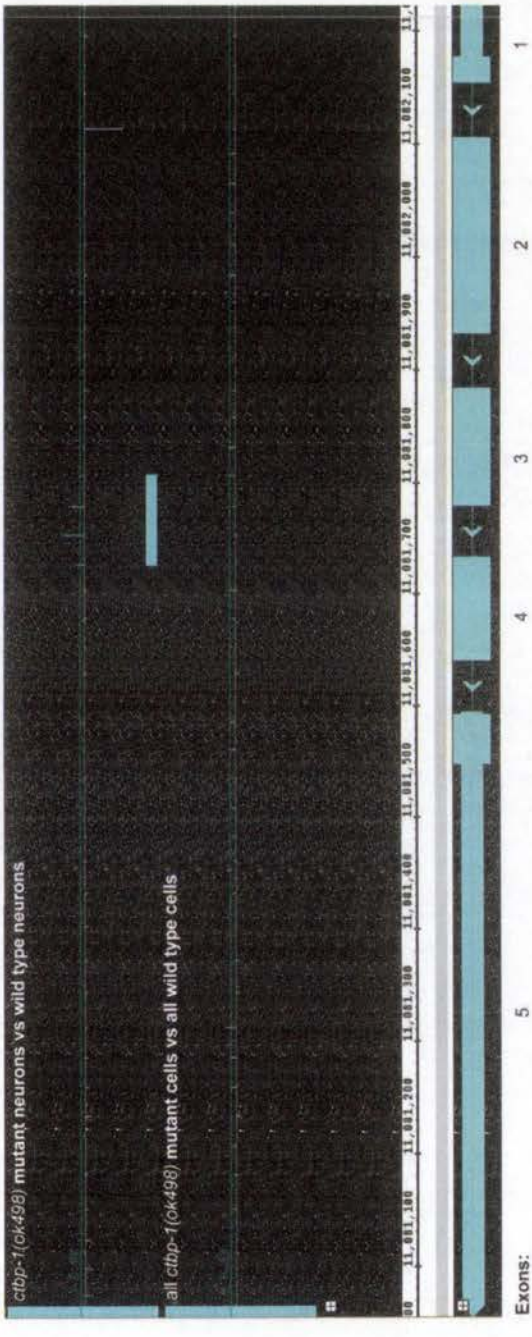


Figure 4.5: In *ctbp-1(ok498)* mutant samples the intensity of *ubc-2* expression corresponding to the exons is same as wild type expression levels. The *ubc-2* gene is shown in 3' to 5' direction at the bottom of the screenshot. Exons are shown with blue boxes and introns with arrowheads. The blue bar which resides in the center of third intron is a non-specific signal detected in the *ctbp-1* mutant vs wild type neurons. The expression intensity corresponding to the exons and introns is same in all *ctbp-1* mutant cells vs all wild type cells.

The *ubc-2* gene was chosen as a control since it was used as a reference gene for Real-time RT-PCR experiments (see section 2.2.17); and the expression level of *ubc-2* is predicted not to vary between *ctbp-1* mutant and wild type samples. As shown in Fig. 4.5, the signals for *ubc-2* exons are unchanged in *ctbp-1* mutant samples compared with the wild type samples. (The blue bar which corresponds to the center of third intron is a non-specific signal).

4.5 Microarray profiling of neurons identified putative targets of CTBP-1

Having completed the global analysis described above, attention was focused on the upregulated gene set. In this case more stringent criteria were applied to include in the analysis only those transcripts with the most robust expression changes between the wild type and *ctbp-1(ok498)* mutant samples. That is, only transcripts with a fold change ≥ 2 and p-value ≤ 0.05 were selected. 41 genes were found to meet these criteria (Appendix Table 2). The majority of these are uncharacterised. For this reason, a more refined gene list was constructed by selecting from among the 41 genes only those with assigned gene names, indicative of some level of prior characterisation. In this way, 15 putative CTBP-1 target genes were identified (Table 4.1). Despite the assignment of formal gene names to each of these 15 genes, most of these have not been previously studied in great depth. Therefore, to assess whether these genes might represent genuine neuronal targets of CTBP-1-mediated control, information on the pattern of expression of these genes was gathered from the literature and from expression pattern databases (Clay Spencer, 2011; Efimenko et al., 2005; Nathoo et al., 2001; Todd Starich, 2001; Vogel and Hedgecock, 2001). One common feature of these genes is that they show a range of neuronal expression from pan-neural to expression in specific classes of neurons, supporting the assertion that they might be neuronal target genes of CTBP-1.

Table 4.1: Putative neuron-specific targets of CTBP-1 identified by MAPCeL. Significantly upregulated genes in *ctbp-1(ok498)* mutant neurons compared with wild type neurons are listed (fold change ≥ 2 and p-value ≤ 0.05). The genes listed in this table are the ones which are assigned with a gene name. Refer to Appendix Table 1 for a detailed gene list. Description, domain and function of these genes are listed along with their expression pattern. Expression pattern information was obtained from the website Wormviz. Expression analysis based on reporter analysis is shown in red (see Discussion for details).

Gene Name	Description of gene name	Domain/Motif	Function	Expression pattern	Fold Change	P-value
ddl-3	daf-16-dependent longevity	tetratrico-peptide repeat (TPR)	implicated in life span	BAG, AVE and dopaminergic neurons, body wall muscles	3.6	0.038
ddo-2	D-aspartate oxidase	D-aspartate oxidase and D-Amino acid oxidase	required for viability in mass RNAi screens	BAG, AVA, GABAergic neurons, coelomocytes and intestine	2.7	0.022
ttr-3	transthyretin related family domain	predicted transthyretin-like domain	uncharacterised nematode specific protein	Pan-neural expression in addition to hypodermal and germline precursor cells	2.7	0.012
nspd-1	nematode specific peptide family, group D	uncharacterised	uncharacterised	Pan-neural expression in addition to expression in coelomocytes	2.6	0.013
dylt-2	dynein light chain	dynein light chain (Tctex-1 domain)	ciliogenesis	BAG neurons, dopaminergic neurons, amphids, phasmids	2.6	0.012
srd-16	serpentine receptor class D	Chemoreceptor/7TM receptor	uncharacterised	AVA and dopaminergic neurons	2.4	0.001
nlp-19	neuropeptide-like protein	uncharacterised	involved in response to food signal	AVE, GABAergic neurons, body wall muscles, pharyngeal muscle and pharyngeal neurons and ASH neurons	2.4	0.007
cnc-10	caenacin(Caenorhabditis bacteriocin)	uncharacterised	uncharacterised	A class neurons, coelomocytes and intestine	2.2	0.025
fkb-2	FK506-binding protein family	peptidylprolyl cis/trans isomerase	uncharacterised	AVE, dopaminergic neurons, intestine	2.1	0.030
sdz-19	SKN-1 dependent zygotic transcript	Tam3-transposase (Ac family)	uncharacterised	germline precursors, panneural, hypodermal	2.1	0.019
sdz-5	SKN-1 dependent zygotic transcript	predicted transposase	uncharacterised	BAG and A class neurons, dopaminergic, hypodermal	2.1	0.019
fipr-3	fungus induced protein related	uncharacterised	uncharacterised	AVE and GABAergic neurons, body wall muscles	2.0	0.048
arrd-3	arrestin domain protein	predicted arrestin	uncharacterised	panneural, GABAergic, intestine	2.0	0.022
glb-15	globin related	Globin and related hemoprotein	uncharacterised	A class neurons, dopaminergic and AVA neurons	2.0	0.003
ins-37	insulin related	insulin like peptide	uncharacterised	AVA and dopaminergic neurons	2.0	0.004

4.6 Microarray profiling of all embryonic cells identified additional putative targets of CTBP-1

Since RNA samples from all embryonic cells had been prepared and analysed in order to normalise the neuronal data set as described above, these all cell data sets were also compared with the view to identifying additional putative CTBP-1 target genes. With the same approach taken for the identification of neuron-specific target genes, from a comparison of all *ctbp-1(ok498)* mutant and all wild type embryonic cells, transcripts with a fold change ≥ 2 and p-value ≤ 0.05 were selected (Appendix Table 3). This approach identified 58 genes which are potentially regulated by CTBP-1. Similar to the neuron-specific CTBP-1 target gene list, the majority of genes listed in Appendix Table 3 are not characterised. Therefore, a more refined list of putative CTBP-1 target genes was constructed by selecting only those genes with assigned gene names (Table 4.2). Again, the expression pattern of these genes was obtained from the Wormviz website as well as from the literature. Interestingly, one of the neuron-specific CTBP-1 target genes *ttr-3* (Table 4.1) is also detected in the all cells data set (Table 4.2). Furthermore, although the genes identified as potential CTBP-1 target genes through this analysis show a range of patterns of expression, each of the ten genes is expressed in either the nervous system or the hypodermis. Since our earlier analysis of the pattern of CTBP-1 expression found that CTBP-1 is expressed in both these tissues, all of these genes have the potential to be directly regulated by CTBP-1.

Table 4.2: Putative CTBP-1 target genes identified by microarray analysis of all embryonic cells Significantly upregulated genes in all *ctbp-1(ok498)* mutant cells compared with all wild type cells are listed (fold change ≥ 2 and p-value ≤ 0.05). The genes listed in this table are the ones which are assigned with a gene name. Refer to Appendix Table 2 for a detailed gene list. Description, domain and function of these genes are listed along with their expression pattern. Expression pattern information was obtained from the website Wormviz. Expression analysis based on reporter analysis is shown in red (see Discussion for details).

Gene Name	Description of gene name	Domain/Motif	Function	Expression pattern	Fold Change	P-value
zyx-1	zyxin	LIM domain	component of nematode cytoskeleton, actin-adaptor protein	AVE, body wall muscles, dopaminergic and GABAergic neurons	3.0	0.006
ttr-3	transthyretin related family domain	predicted transthyretin-like domain	uncharacterised nematode specific protein	Pan-neural expression in addition to hypodermal and germline precursor cells	2.4	0.021
nhr-144	nuclear hormone receptor family	uncharacterised	uncharacterised	hypodermis and intestine	2.3	0.034
him-4	high incidence of males	Immunoglobin related	epithelial attachment	GABAergic neurons, distal body wall muscles	2.3	0.010
inx-13	innexin	Innexin-type	formation of gap junctions	AVA, A class and dopaminergic neurons, hypodermis, excretory cells	2.2	0.034
tdp-1	tar DNA-binding protein homolog	RNA recognition domain (RRM) domain	implicated in neurodegeneration	BAG, GABAergic and A class neurons, coelomocytes, intestine	2.1	0.003
fkb-8	FK506-binding protein family	peptidyl-prolyl cis/trans isomerase	uncharacterised	AVA, coelomocytes, intestine	2.1	0.005
lev-11	levamisole resistant	Actin filament-coating protein tropomyosin	component of muscle assembly	body wall muscles, pharyngael muscles, intestine, germline, AVE and dopaminergic neurons	2.1	0.010
mec-2	mechanosensory abnormality	PID (proliferation, ion and death) superfamily	touch response	touch neurons, VNC neurons and neurons in the head and tail, BAG, dopaminergic and A class neurons	2.0	0.012
cat-4	abnormal catecholamine distribution	GTP cyclohydrolase I	serotonin and dopamine biosynthesis	serotenergic and dopmainergic neurons, muscles, hypodermis	2.0	0.022

4.7 Discussion

With the ultimate goal of identifying neuron-specific target genes of the co-repressor CTBP-1, the microarray based method MAPCeL was employed to detect differentially expressed genes in *ctbp-1* mutant neurons compared with wild type neurons. In chapter 3, we reported a neuronal role for CTBP-1 in DA type embryonic neurons in the VNC and have shown that CTBP-1 is expressed broadly in the nervous system. In the light of these observations we sought to identify neuronal targets of CTBP-1 and found various genes with neuronal functions.

Despite the publication of several reports on neuron-specific transcription factors that interact with CtBP, to date there is little information on neuron-specific CtBP target genes (Melhuish and Wotton, 2000; Postigo and Dean, 1999). The only reported neuronal CtBP target in mammalian systems is the epiloptegenesis-related gene BDNF (Garriga-Canut et al., 2006). Other reported CtBP targets have roles in UV-induced apoptosis (Grooteclaes et al., 2003) and the epithelial to mesenchymal transition (EMT) (Grooteclaes and Frisch, 2000). Due to technical challenges associated with redundancy of the two CtBPs in the mammalian systems and early embryonic lethality of double knockouts, these CtBP target genes have been identified in experiments conducted on mammalian cell lines rather than whole animals. In contrast, in *C. elegans*, experiments have been performed to identify CtBP target genes *in vivo*. In *C. elegans*, microarray experiments have been performed on whole animals at the young adult stage comparing transcripts from wild type and *ctbp-1(ok498)* mutant worms. Although these experiments identified 213 genes that are upregulated in the mutant samples, representing potential CTBP-1 target genes, there was no obvious enrichment of neuronal genes among these as might be expected for direct targets of a primarily neuronally-expressed co-repressor (Chen et al., 2009). As an alternative approach to identifying CTBP-1 target

genes, we analysed transcripts from embryonic neurons. From these analyses, a group of genes were identified which are robustly upregulated when CTBP-1 function is reduced. Several of these display neuronal functions and may be *bona fide* CTBP-1 targets.

As noted earlier, one specific goal of the experiments described here was to identify genes that are controlled by CTBP-1 in DA cells, in order to better understand the role played by CTBP-1 in defining the characteristics of these motor neurons. To more directly address this question, rather than performing MAPCeL on all neurons as we have done, it would have been preferable to perform the transcriptional profiling specifically on the DA neurons. To achieve this, a *gfp* marker that is specific to the DA neurons would need to be used to sort the relevant cell population. Since the best characterised marker of the DA neurons is UNC-4::GFP, and since we have noted loss of expression of this marker from the DA neurons of *ctbp-1(ok498)* mutant animals, it was not possible to perform this experiment. Nonetheless, DA-specific targets of CTBP-1 may be found in the set of upregulated genes that were identified in the pan-neuronal samples that were analysed here. Direct analysis of the expression pattern of the candidate target genes in both wild type and *ctbp-1 (ok498)* mutants using fluorescent reporters will be required to assess this.

Although additional experiments will be required to assess whether the genes upregulated in the *ctbp-1(ok498)* mutant neurons are direct targets of CTBP-1 and to determine which of these are specific to DA cells, it is of interest to consider the potential functions of the putative target genes and how they might relate to the activities that we and others have ascribed to CTBP-1.

Of the putative neuronal target genes identified in this study, three are relatively well characterised: *nlp-19, dylt-2* and *ddl-3*.The first is a member of the neuropeptide like family,

of which there are 32 in *C. elegans*. Most of *nlp* genes are expressed in neurons and it has been reported that *nlp-19* is expressed in pharyngeal neurons and sensory neurons (ASH) (Nathoo et al., 2001). Using the expression information available on the website Wormviz we have found that *nlp-19* is also expressed in AVE and GABAergic neurons in addition to muscles (Table 4.1) (Clay Spencer, 2011). Studies on neuropeptide-like proteins in *C. elegans* have shown that, similar to their mammalian orthologs, *nlp*-type genes are involved in behavioural responses of the worm (Harris et al., 2010). Consistent with this feature of the *nlp* family, *nlp-19* has been reported to function in response to food signal where worms fed with *nlp-19* RNAi avoided food (Harris G P, 2010). Interestingly, a behavioural role for CTBP-1 is suggested by the observation that mutants carrying a reduction of function allele of *ctbp-1* fail to adapt to alcohol (J. Bettinger, personal communication) (Bettinger, 2011). Moreover, in our laboratory, the same reduction of function allele described above was observed to confer on animals an unusual feeding behaviour (A. Reid, unpublished) suggesting that CTBP-1 might be targeting genes involved in behavioural response of the nematode.

Another putative CTBP-1 target identified in the current work is *dylt-2* which encodes dynein light chain and is expressed in many neurons including ciliated sensory neurons (i.e., amphids and phasmids) according to DYLT-2::GFP reporter analysis (Efimenko et al., 2005). DYLT-2 is thought to have a role in assembly of the dynein and is implicated in intraflagellar transport (IFT).

The third relatively characterised CTBP-1 target gene, *ddl-3*, has been identified through an RNAi screen looking for regulators of life span (Hansen et al., 2005). A number of genetic pathways and processes regulate life span in *C. elegans*. Dietary restriction, ablation of the germline, and mutations in the Insulin/IGF-1 pathway extend the life span of the worm

through different mechanisms (Apfeld and Kenyon, 1998; Dorman et al., 1995). The signalling cascade that is downstream of insulin/IGF-1 involves the FOXO transcription factor DAF-16 which is required for life span extension of animals that are deficient in insulin/IGF1 signalling (Lin et al., 2001). However, life span extension through dietary restriction does not require DAF-16, suggesting distinct mechanisms for life extension depending on the context. For identification of novel components that are involved in life extension, a large-scale RNAi screen was employed by Hansen *et. al.* in which a number of genes, including *ddl-3*, were reported to extend life span in a *daf-16*-dependent manner (Hansen et al., 2005). *ddl-3* is also identified in our microarray analysis as a putative CTBP-1 target. Given the role of CTBP-1 in regulation of life span (Chen et al., 2009), *ddl-3* might be one of the CTBP-1 targets that contributes to correct regulation of life span, however the mechanism of this regulation needs further exploration.

In addition to the neuron-specific CTBP-1 targets, we also identified those genes which are significantly upregulated in the RNA samples derived from all embryonic cells of *ctbp-1(ok498)* mutants. Several of these have been studied in some detail, including *tdp-1*, *mec-2*, *cat-4*, *zyx-1*, *him-4*, *lev-11* and *inx-13*. These candidate CTBP-1 target genes will be described in turn.

Out of putative CTBP-1 target genes identified in all cells data set, one of the interesting genes is *tdp-1*, which encodes a protein homologous to mammalian TDP-43. TDP-43 is associated with neurodegenerative diseases such as ALS where deposition of human TDP-43 in the nervous system causes neurotoxicity (Del Bo et al., 2009). Overexpression of *tdp-1* in the nervous system of *C. elegans* results in neurotoxicity which in turn causes uncoordinated locomotion (Ash et al., 2010). This observation recapitulates the neurotoxicity caused by

deposition of human TDP-43. Moreover, recently TDP-1 was reported to have a role in life span, where loss of TDP-1 extends life span of the worm (Tao Zhang et al., 2012). It is of great interest to assess whether *tdp-1* is indeed a direct target of CTBP-1 and to determine whether the 2-fold overexpression of *tdp-1* that we have observed in embryonic cells of *ctbp-1(ok498)* mutants has any phenotypic consequences.

Another putative CTBP-1 target that was identified in the all cells data set is *mec-2*. The transmembrane protein MEC-2 is responsible for transduction of mechanosensory signal in touch receptor neurons (Brown et al., 2008). Interestingly, expression of *mec-2* is regulated by the transcription factor PAG-3, which is an interacting partner of CTBP-1 (Nicholas et al., 2008); *pag-3* mutants show ectopic expression of a *mec-2::gfp* reporter in BDU interneurons (Jia et al., 1997). Given that PAG-3 physically interacts with CTBP-1, it is possible that CTBP-1 acts as a co-repressor for PAG-3 to repress *mec-2* expression in BDU interneurons. This notion could be tested by examining whether a *mec-2::gfp* reporter shows ectopic expression in *ctbp-1(ok498)* mutants.

Another putative CTBP-1 target gene identified in this analysis is *cat-4*. CAT-4 has a role in dopamine and serotonin biosynthesis and *cat-4* mutants are defective in forward movement due to a defect in serotonin signalling (Wakabayashi et al., 2005). This defect in serotonin signalling also manifests itself in other contexts such as food sensing; *cat-4* mutants have impaired food-odour sensing (Nuttley et al., 2002).

zyx-1, another putative CTBP-1 target gene, encodes a homolog of vertebrate Zyxin, which functions as a focal adhesion protein (Smith et al., 2002). Similar to its vertebrate counterpart, *C. elegans* ZYX-1 acts as a cytoskeletal adaptor protein and is implicated in muscle function

through interaction with another protein called DYC-1, which is crucial for muscle function (Lecroisey et al., 2008). Another putative target, *him-4*, also functions in muscles and is expressed in distal body wall muscles according to reporter gene analysis (Vogel and Hedgecock, 2001). It encodes a novel Extracellular Matrix (ECM) protein, hemicentin, and functions in tissue/cell maintenance (Michaux et al., 2001). Another CTBP-1 target with characterised functions in muscle cells is *lev-11* (*lev*amisole resistant) which encodes the protein tropomyosin (Anyanful et al., 2001; Kagawa et al., 1995). Levamisole is a strong agonist of the neurotransmitter acetylcholine and *lev-11* mutants show a mild resistance to levamisole, a phenotype observed in mutants with defective acetylcholine receptors (Lewis et al., 1980). Finally, *inx-13* belongs to the family of innexins which are homologs of vertebrate connexins. Similar to connexins, innexins form gap junctions. INX-13 is one of many nematode innexin genes and is expressed in excretory cells and in the membrane of hypodermal cells. (Todd Starich, 2001). Mutant analysis has shown that *inx-13* is required for viability at an early larval stage L1. It is predicted that *inx-13* mutant animals show larval lethality due disruption of osmoregulation within the hypodermal cells and the body cavity.

It is of great interest to determine whether the genes described here and listed in Appendix Table 1 and 2 are *bona fide* target genes of CTBP-1. As a first step towards this goal, ChIP could be employed to check whether CTBP-1 is actually localised at the relevant promoters. Another means of verifying CTBP-1 targets could be through the analysis of transcriptional reporters of pertinent genes. In this case, transgenes consisting of the promoter of the gene of interest fused to *gfp* would be introduced into the *ctbp-1(ok498)* mutant background and analysed for changes in the pattern or intensity of GFP expression compared with wild type. GFP would be expected to be upregulated in *ctbp-1* mutants if the reporter gene is a *bona fide* direct CTBP-1 target gene. Since microarrays have limitations in distinguishing direct targets

from indirect targets, ChIP-seq can be used as an alternative approach to assist in identification of direct CTBP-1 targets.

In summary, in this study various putative CTBP-1 target genes with distinct roles have been identified. These findings provide the starting point for future investigations aimed at understanding the *in vivo* significance of correct regulation of these genes.

CHAPTER 5

A MUTAGENESIS SCREEN CARRIED OUT FOR THE INVESTIGATION OF CTBP-1 REGULATORS

5.1 Introduction

Members of the CtBP family of co-repressors regulate transcription by acting as bridging molecules between DNA-binding transcription factors and histone modifying proteins. Although the modes of action of CtBP proteins have been well-investigated, less attention has been focussed on how the CtBP proteins are regulated. There is only little information on how their expression or repressive activity is modulated. So far the reported regulatory mechanisms that influence CtBP activity involve post-translational modifications via kinases and via signalling molecules as revised in the general introduction (Chapter 1). For instance, CtBP1 is phosphorylated by a homeodomain interacting protein kinase Hipk2 which induces degradation of CtBP1 and by p21-activated kinase-1 (Pak-1) which changes the subcellular distribution of CtBP1 (Barnes et al., 2003; Zhang et al., 2003). Another regulator of CtBP1 that changes its subcellular localisation is nNOS; upon binding, CtBP1 shifts from the nucleus to the cytoplasm (Riefler and Firestein, 2001). And finally, repressive activity of CtBP proteins is enhanced by NADH binding (Zhang et al., 2002). In this chapter, we sought to discover regulators of *C. elegans* CTBP-1 by exploiting the genetic tractability of *C. elegans*. With the goal of discovering novel regulators of CtBP, we carried out a mutagenesis screen using the transgenic strain expressing CTBP-1::GFP fusion protein under the control *ctbp-1* promoter (Chapter 3, Section 3.2). The experiments described in this chapter seek to identify and characterise one of the putative CTBP-1 regulators isolated through this screen.

5.2 A mutagenesis screen identified several mutants with altered CTBP-1::GFP expression

In a wild type animal, a CTBP-1 reporter gene is expressed at high levels in the nervous system and at lower levels in the hypodermal cells (see Chapter 3 Figure 3.1 D). With the aim of discovering novel genes that are involved in regulation of CTBP-1, a mutagenesis screen

was performed in which mutations that altered the intensity or pattern of expression of this *ctbp-1::gfp* reporter gene were sought. To this end, a population of animals expressing the CTBP-1::GFP fusion protein under the control of the *ctbp-1* promoter (integrated transgene *ausIs1*) were treated with the mutagen Ethyl Methane Sulfonate (EMS) as depicted in Figure 5.1.

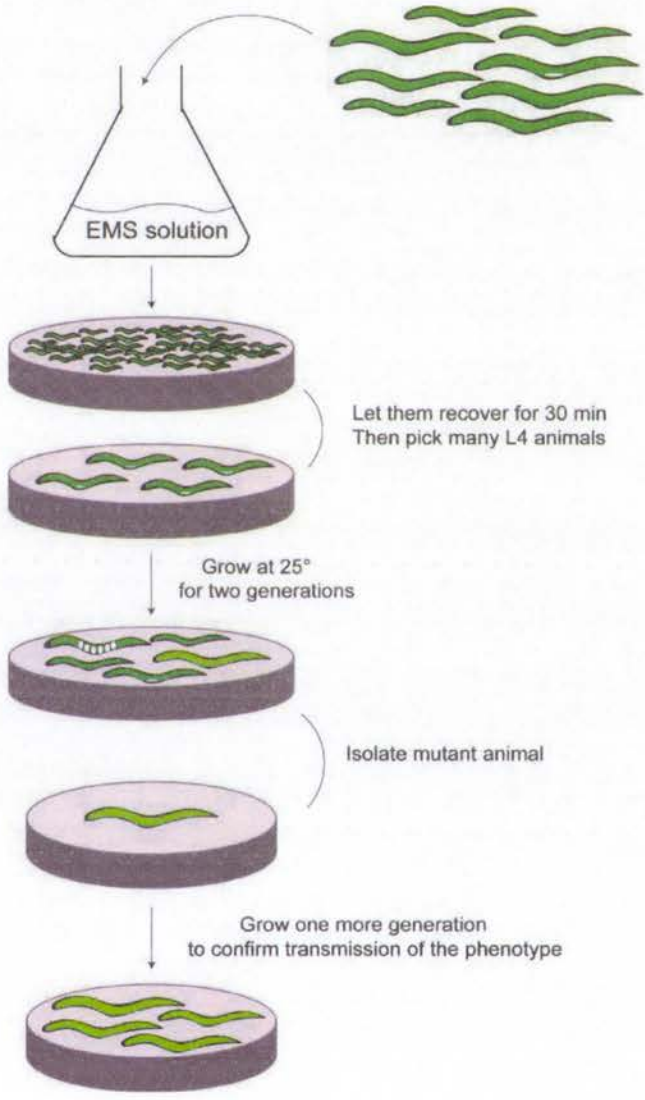


Figure 5.1: Mutagenesis of *ctbp-1::gfp* transgenic animals. A mixed population of animals was treated with Ethyl Methane Sulfonate (EMS), washed and allowed to recover for half an hour. After mutagenesis, animals at

the L4 stage were selected and many L4 animals were singled on individual plates then grown for two generations. Within the resultant populations, individual mutants with alternations to the pattern and level of CTBP-1::GFP expression were sought. Mutants that were isolated in this way were allowed to grow for one more generation in order to confirm the transmission of the phenotype.

The second generation self-progeny of these mutagenised animals were then scored for any change in the level or pattern of CTBP-1::GFP expression. The self-progeny of the mutants that had been identified through this preliminary screening were scored again and only those that showed the same phenotype in this third generation were retained. 6000 haploid genomes were screened in this way and 8 mutants were isolated (Table 5.1). These 8 mutants show a range of alterations to CTBP-1::GFP expression. These include both increases and decreases in expression levels of the CTBP-1::GFP reporter in particular cells or tissues as noted in Table 5.1.

Table 5.1: Eight mutants with altered expression of a *ctbp-1::gfp* transgene were identified. The strain name, genotype and phenotype of the mutants are described.

Strain name	Genotype	*ctbp-1::gfp* expression	other phenotypes
HRN060	*ausIs1 (ctbp-1::gfp); aus2*	increased expression of the *ctbp-1::gfp* array in the Ventral Nerve Cord	uncoordinated
HRN061	*ausIs1 (ctbp-1::gfp); aus3*	increased expression of the *ctbp-1::gfp* array in the hypodermis	not detected
HRN062	*ausIs1 (ctbp-1::gfp); aus4*	ectopic expression of the *ctbp-1::gfp* array along the body	not detected
HRN063	*ausIs1 (ctbp-1::gfp); aus5*	decreased expression of the *ctbp-1::gfp* array both in the nervous system and hypodermis	mildy uncoordinated
HRN064	*ausIs1 (ctbp-1::gfp); aus6*	fewer number of *ctbp-1::gfp* expressing tail cells	slow growing
HRN065	*ausIs1 (ctbp-1::gfp); aus7*	decreased expression of the *ctbp-1::gfp* array both in the nervous system and hypodermis	very sick
HRN066	*ausIs1 (ctbp-1::gfp); aus8*	extra cells expressing *ctbp-1::gfp* array near the pharynx	not detected
HRN067	*ausIs1 (ctbp-1::gfp); aus9*	extra cells expressing *ctbp-1::gfp* array near the tail	not detected

One mutant, *ausIs1 (ctbp-1::gfp); aus3,* was selected for further study. This mutant was selected as it showed a robust change in *ctbp-1::gfp* expression that was readily observable by dissecting microscope. Furthermore, the mutant allele was found to be recessive, which would facilitate identification of the phenotype-causing mutation. Before embarking on characterisation of this mutant, it was first outcrossed 6 times (that is, mated with wild type males) in order to remove background mutations.

5.3 The *aus3* allele alters CTBP-1::GFP expression in the hypodermis

Compared with wild type animals, which show only low levels of CTBP-1::GFP expression in the hypodermis (Fig. 5.2 A), the *ausIs1 (ctbp-1::gfp); aus3* mutant shows increased levels of CTBP-1::GFP in this tissue (Fig. 5.2 B). Interestingly, there is no detectable difference in the level of CTBP-1::GFP expression in the nervous system as assessed by fluorescence microscopy. That is, the effect of the *aus3* mutation on CTBP-1::GFP expression appears to be specific to the hypodermis. Although the previous focus of this work had been on the neuronal functions of CTBP-1, the most robust mutant phenotype observed was in the hypodermis, therefore hypodermal expression was followed further.

In *C. elegans* adult hermaphrodites the hypodermis is composed of 11 hypodermal cells abbreviated as Hyp1-11. One unusual characteristic of the hypodermal cells is their multinucleate nature. For instance, Hyp7 has more than a hundred nuclei and covers almost the entire body. Levels of CTBP-1::GFP expression in the hypodermis were measured using ImageJ software to quantify pixel intensity in a defined area of the anterior portion of the hypodermis. Confocal images of 10 animals per strain were captured and the intensity of CTBP-1::GFP expression in the hypodermal cells was calculated. This has shown that the

level of CTBP-1::GFP is 1.5 times higher in *aus3* mutants than in wildtype animals (Fig. 5.2 C). (**, $P < 0.01$, Student's *t* test).

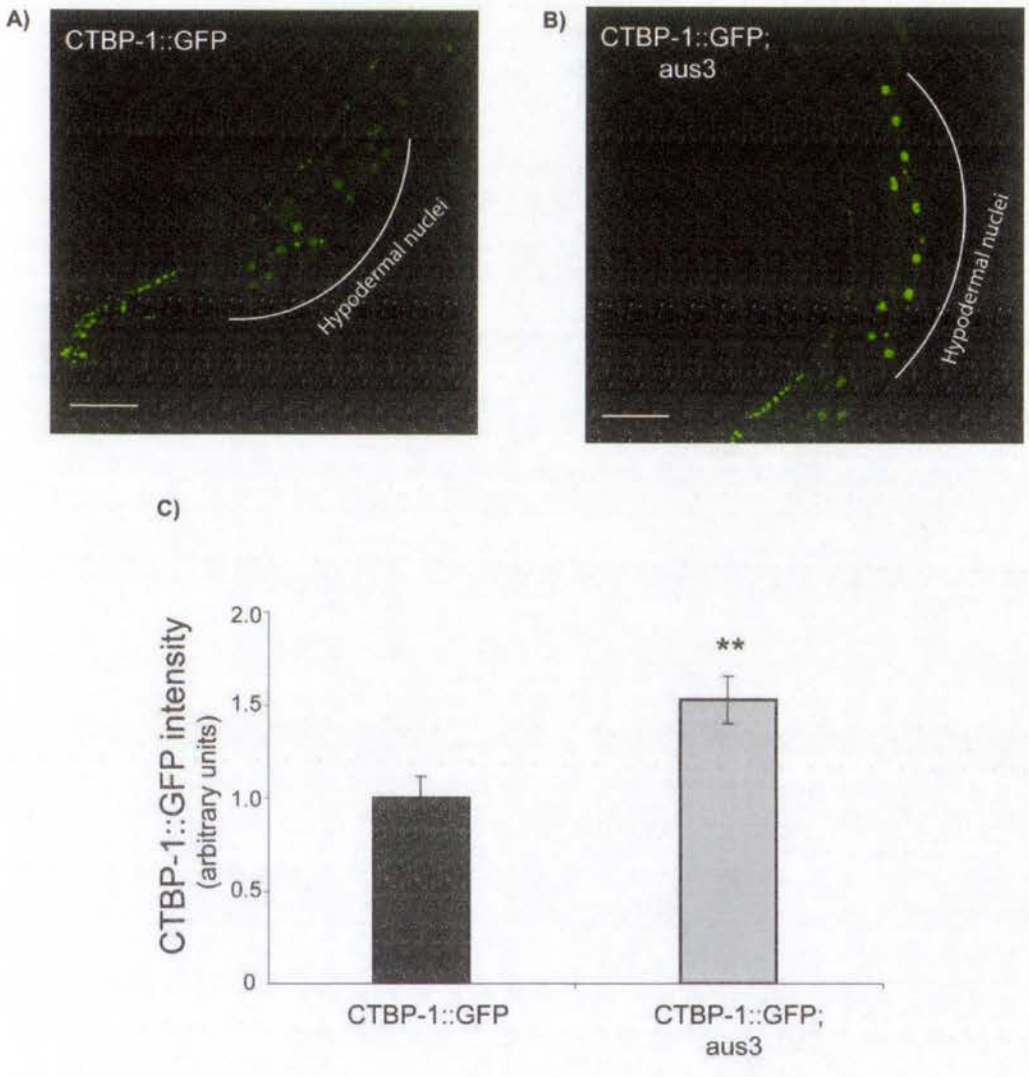


Figure 5.2: Mutant animals carrying the *aus3* allele showed elevated levels of CTBP-1::GFP expression in the hypodermis. Representative images of **A)** CTBP-1::GFP expression in wildtype background **B)** and CTBP-1::GFP expression in *aus3* mutant background are shown. Hypodermal nuclei are indicated. Anterior is to the left, scale bar shows 50 µm. **C)** Graph comparing CTBP-1::GFP expression intensity in the hypodermis of wild type and *aus3* mutant animals. The anterior region of transgenic animals in wild type and *aus3* mutant background were imaged and the expression level of CTBP-1::GFP was quantified using ImageJ software with the same capture settings for all conditions (i.e, laser intensity, voltage and gain power). The graph presents

average intensity of CTBP-1::GFP expression in wild type and *aus3* mutants. Error bars indicate SEM of 10 independent measurements of CTBP-1::GFP expression intensity for both wild type and *aus3* mutants. (**, $P < 0.01$, Student's *t* test).

5.4 The *aus3* mutation maps to the centre of Chromosome III

Having confirmed that the *aus3* mutation leads to a robust upregulation of CTBP-1::GFP expression specifically in the hypodermal nuclei, linkage analysis was used to determine the approximate genomic location of the mutation. In *C. elegans* this can be achieved by snipSNP mapping. That is, mapping using single nucleotide polymorphisms (SNPs) which alter a restriction site (Davis et al., 2005). Many such SNPs have been identified between two wild *C. elegans* isolates, namely the N2 and Hawaiian (HA) strains. A set of 6 of these snipSNPs, which fall close to the middle of each of the six nematode chromosomes were used to locate the *aus3* mutation.

The strain carrying the *aus3* mutation is in the N2 background. Hermaphrodites of this strain were therefore crossed with males of the HA strain and cross progeny (which are heterozygous for the N2 and the HA snipSNPs) were allowed to produce progeny by self fertilisation. Animals showing the *aus3* mutant phenotype (that is, increased hypodermal expression of the *ctbp-1::gfp* transgene) were picked from these progeny and assayed for the snipSNPs of interest. For snipSNPs that are unlinked to the *aus3* mutation, one quarter of the assayed animals will be homozygous for the N2 SNP, one quarter will be homozygous for the HA SNP, and the remaining half will be heterozygous for the N2 and HA SNPs (ie ~75% of assayed animals will have at least one copy of the N2 SNP). In contrast, for snipSNPs that are linked to the *aus3* mutation, the frequency of observation of the N2 SNP will be increased (Fig. 5.3). It should be noted that because the phenotype of the *aus3* mutation can only be detected in the presence of the *ctbp-1::gfp* transgene, it was predicted that a high incidence of

the N2 SNP would be observed on chromosome I (the location of the *ausIs1(ctbp-1::gfp)* transgene) in addition to the chromosomal location of the *aus3* mutation.

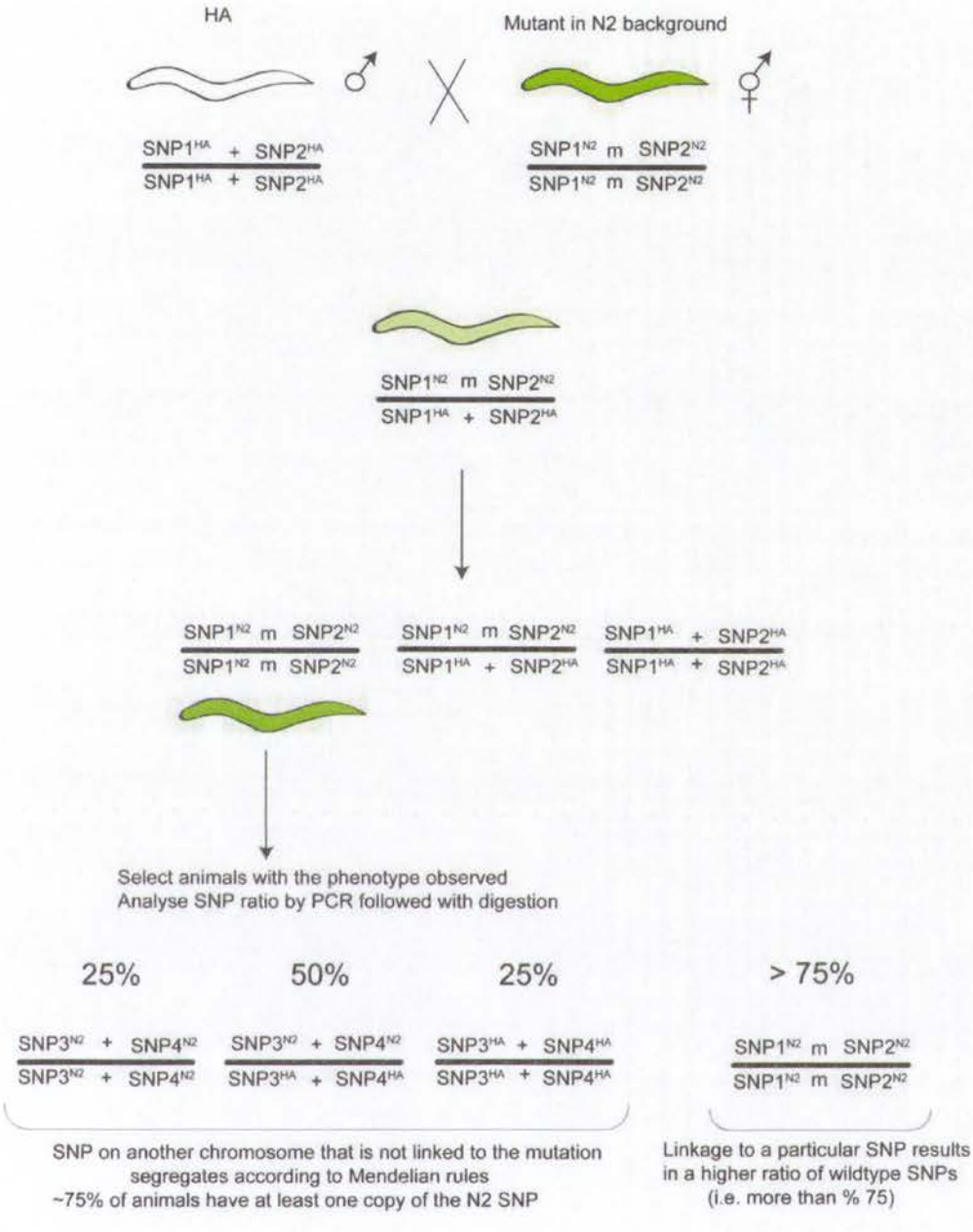


Figure 5.3: snipSNP mapping was used to identify the chromosomal location of *aus3*. Using two wild isolates of *C. elegans* (N2 and HA) and polymorphic differences between them that alter restriction sites, the *aus3* allele was mapped using snipSNP mapping. First, *ausIs1 (ctbp-1::gfp); aus3* mutants were mated with HA males. The heterozygous progeny were allowed to self fertilise and animals displaying the phenotype of the *aus3*

mutant, that is, increased CTBP-1::GFP expression in the hypodermis were selected from among the progeny for linkage analysis. Originally, the *ausIs1 (ctbp-1::gfp); aus3* mutant is in N2 background therefore if the *aus3* mutation is linked to a particular chromosome, the N2 SNPs will be observed at higher frequency on that chromosome. In the unlinked condition, SNPs will segregate according to the Mendelian ratio thus ~75% of animals will have at least one copy of the N2 SNP.

The initial linkage analysis using snipSNPs on the centre of each chromosome suggested that the *aus3* mutation is on chromosome III. Then, in order to locate *aus3* mutation to a more refined position on chromosome III, snipSNPs corresponding to the left end and right end of this chromosome were assayed. As indicated by the high percentage of N2 snipSNPs detected at the centre of chromosome III compared with either of the arms, it was concluded that *aus3* mutation maps on chromosome III close to the centre of the chromosome (Table 5.2). It should be noted that the number of animals assayed for each snipSNP was small, thus these mapping experiments provided only a tentative assignment of the genomic location of *aus3*.

Table 5.2: Linkage analysis of *aus3* mutant has shown that *aus3* allele is on Chromosome III, close to position -1. The percentage of N2 type SNPs are indicated for each chromosomal location. (The number of animals assayed for each location is as follows: ChI, -1 n=6, ChII, 1 n=3, ChIII, -25 n=6, ChIII, -1 n=6, ChIII, 21 n=8, ChIV, 1 n=7, ChV, 1 n=7, ChX,2 n=3)

Linkage Group, Position	% of Mutants with wild type SNP
I, -1	83 %
II, 1	33 %
III, -25	17 %
III, -1	100 %
III, 21	75 %
IV, 1	14 %
V, 1	29 %
X, 2	0 %

5.5 Fourteen candidate genes were identified via Next-Generation Sequencing

C. elegans holds the distinction of being the first multicellular organism to have its genome fully sequenced (Consortium, 1998). Recently, efficient and fast whole genome sequencing technologies called next-generation sequencing have been developed and these technologies have been applied in *C. elegans* where whole genome sequencing was used for mutant identification (Bigelow et al., 2009). In the first publication of this technique, variants in a genome of interest were identified using the published sequence of the N2 strain as a reference. However, it is now recommended that a laboratory-specific reference strain be used due to the recent finding that two "wildtype" strains cultivated in different laboratories have thousands of variants (Flibotte et al., 2010).

In order to identify the mutation that causes the overexpression phenotype in *aus3* mutants, the Illumina sequencing platform was used to sequence both our lab-specific reference strain and the *aus3* mutant. As stated earlier, the *ausIs1 (ctbp-1::gfp); aus3* mutant was outcrossed 6 times with wildtype (N2) males in order to minimise background mutations that are not causative of the *ctbp-1::gfp* overexpression phenotype. For sequence analysis, genomic DNA was extracted from this outcrossed strain. *C. elegans* has a genome size of 100,281,426 bases and Illumina sequencing platform allowed for up to 12x (1 lane) average coverage which covers up to 90% of the genome (Bigelow et al., 2009).

Despite the fact that the *aus3* mutant was outcrossed prior to whole genome sequencing, a comparison of the *aus3* genome sequence with the 1998 N2 reference sequence using MAQGene software revealed 12424 variants (Appendix Table 4). Since linkage analysis had localised the *aus3* mutation to chromosome III, only the variants on that chromosome were considered further, of which there were 474. To exclude those that are unrelated to the *aus3*

mutation and are simply sequence changes present in our laboratory N2 strain, any sequence variations which also appeared in the sequence obtained from our lab-specific reference strain were removed. After this subtraction, 14 variants remained on chromosome III which were considered further. Each of these 14 variants were then checked by standard Sanger sequencing, which confirmed the mutation in 10 cases but returned wild type sequences in the remaining 4 instances (Table 5.3). Thus 10 candidate mutations which could be causative of the *aus3* mutant phenotype remained.

Table 5.3: Sequence variants were identified on chromosome III between the *aus3* mutant and reference strain. Type of mutation, cosmid name and assigned gene names if any are indicated along with the map positions for each candidate gene. Sanger sequencing was used to identify any sequencing errors.This secondary sequencing demonstrated that the variants identified in those candidate genes that are highlighted in yellow were sequencing errors.

DNA	Start	Reference Base	Sample Base	Variant Class	Description	Cosmid Name	Gene name	Map position (cM)
III	5250546	G	A	premature stop	CAG->TAG[Gln->stop]	C34E10.8	NA	-1.86
III	5683535	A	T	missense	ATG->AAG[Met->Lys]	F54E7.7	rcn-1	-1.45
III	6322252	G	T	missense	ACA->AAA[Thr->Lys]	C56G2.5	NA	-1.4
III	6376018	G	A	missense	GAA->AAA[Glu->Lys]	C16A3.7	tag-182	-1.32
III	6990359	T	G	splice donor	none	C06E8.3b	prk-1	-0.85
III	7021100	A	T	five prime UTR	none	F11H8.1.1	rfl-1	-0.85
III	7357296	C	T	five prime UTR	none	F08F8.10a	NA	-0.75
III	7363343	A	G	missense	TTT->CTT[Phe->Leu]	F08F8.5	numr-1	-0.075
III	7608528	T	C	missense	AAG->AGG[Lys->Arg]	T04A6.3	NA	-0.66
III	8552913	T	G	missense	GTG->GGG[Val->Gly]	C02D5.3	gsto-2	-0.25
III	8552916	C	G	missense	GCT->GGT[Ala->Gly]	C02D5.3	gsto-2	-0.25
III	10498116	C	A	missense	CAG->AAG[Gln->Lys]	F43D9.1	NA	2.58
III	10498117	A	G	missense	CAG->CGG[Gln->Arg]	F43D9.1	NA	2.58
III	10498118	G	C	missense	CAG->CAC[Gln->His]	F43D9.1	NA	2.58
III	10553060	G	A	missense	CCA->TCA[Pro->Ser]	T21C12.4	NA	3.62
III	11871798	A	G	missense	GTG->GCG[Val->Ala]	Y56A3A.6	ntl-3	13.27
III	13074298	T	C	three prime UTR	none	Y39E4B.6	NA	20.75

5.6 C34E10.8 rescues the overexpression phenotype of *aus3* mutants

In order to identify the phenotype-causing mutation among the 10 identified candidates, rescue experiments were performed. That is, a wild type copy of the candidate gene was introduced into the *aus3* mutant strain, and its capacity to ameliorate the mutant phenotype was assessed.

As stated earlier, the *aus3* mutation was mapped close to the centre of chromosome III. In this region, a nonsense mutation in the gene C34E10.8 (map position chromosome III, -1.86) was selected as the best candidate for the phenotype-causing mutation. The entire coding region of C34E10.8 along with 5kb upstream of the start site of translation and 325 bases downstream of the stop codon was amplified by PCR from wild type N2 genomic DNA. Stable transgenic lines carrying this PCR product along with a co-transformation marker plasmid pRF4 (containing the rol-6*(su1006)* allele which confers a dominant Roller phenotype) were generated by microinjection. Lines carrying only the co-transformation marker were also generated as a control. In both the test and the control cases, the animals carrying the extrachromosomal arrays also carried the *ausIs1 (ctbp-1::gfp)* transgene and the *aus3* mutation. As shown in Figure 5.4, introduction of the co-transformation marker and the wild type C34E10.8 genomic region reduced the expression of CTBP-1::GFP in the hypodermis of *aus3* mutants, while introduction of the co-transformation marker alone did not alter CTBP-1::GFP expression. The observed reduction of expression in animals carrying the wild type C34E10.8 genomic region suggests that the nonsense mutation that had been identified in C34E10.8 is indeed causative of the increased levels of CTBP-1::GFP in the hypodermis in the *aus3* mutant.

Two additional interesting observations were made through these rescue experiments. First, animals carrying the wild type C34E10.8 genomic region showed barely detectable levels of expression of CTBP-1::GFP in the hypodermis (Figure 5.4 B). That is, the levels of hypodermal GFP in these animals is less than the low level that is observed in wild type animals. Since extrachromosomal arrays like the *ausEx23* (C34E10.8 + *rol-6*) array contain many copies of the transgene, it is likely that the level of expression of C34E10.8 in this strain is greater than in wild type animals. The reduction of expression of CTBP-1::GFP in this

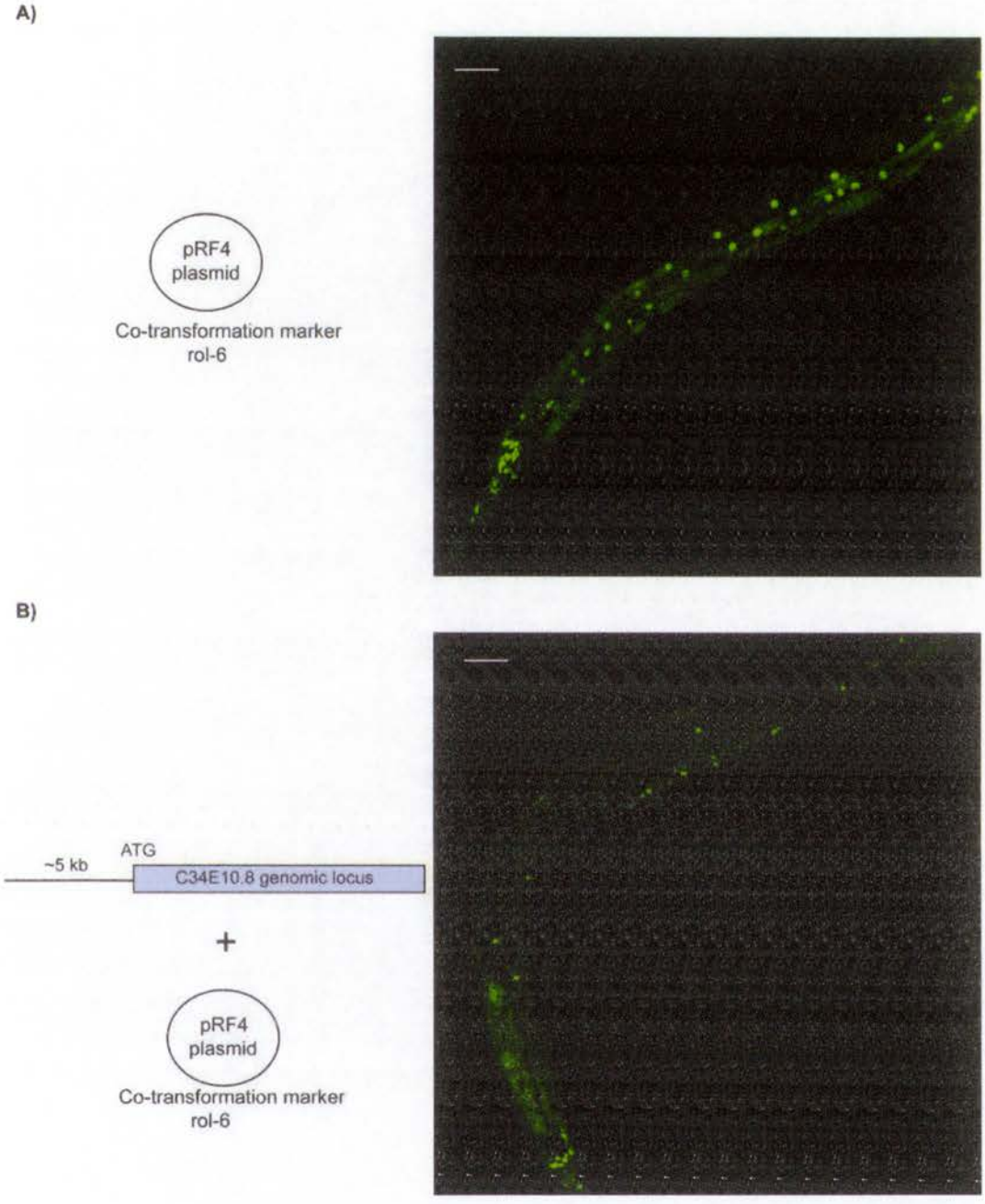


Figure 5.4: C34E10.8 rescues the CTBP-1::GFP overexpression phenotype of *aus3* mutants. Representative images of **A)** the control strain carrying *aus3* mutation and co-transformation marker *rol-6 (su1006)* **B)** and rescued strain carrying *aus3* mutation, *rol-6 (su1006)* marker and wild type copy of genomic C34E10.8. Scale bar indicates 50μm.

strain below the levels observed in wild type is consistent with C34E10.8 acting to repress gene expression.

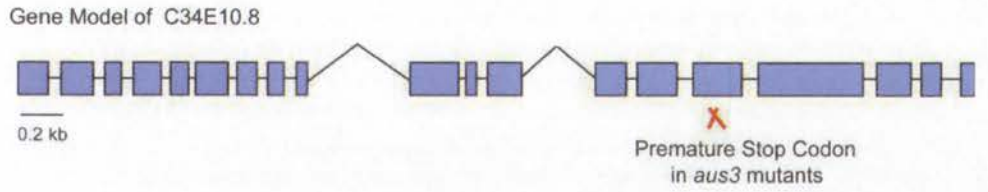
Second, no discernable difference was observed in the expression of CTBP-1::GFP in the nervous system in the strain carrying the C34E10.8 transgene array. Together with the observed lack of upregulation of CTBP-1::GFP in the nervous system of *aus3* mutants (section 5.3), this failure of the C34E10.8 transgene array to downregulate CTBP-1::GFP expression in the nervous system suggests that C34E10.8 regulates the *ctbp-1::gfp* transgene specifically in the hypodermis.

5.7 C34E10.8 might be degraded by NMD in *aus3* mutants

The mutation that was identified in C34E10.8 is a nonsense mutation [Gln->stop]. This mutation is therefore predicted to lead to the production of a truncated C34E10.8 protein (Fig 5.5 A); instead of the 1024 aa wildtype C34E10.8, a mutant protein of 680aa would be produced (Fig. 5.5 B). However, the mRNA with a premature stop codon could also be subject to nonsense mediated decay (NMD) (Chang et al., 2007). In order to test this, Real-Time RT PCR was used to quantify C34E10.8 mRNA levels in *aus3* mutant.

First, a strain carrying only the *aus3* mutation was generated (that is, the *ctbp-1::gfp* transgene was removed). Real-time RT PCR was then used to quantify *C34E10.8* transcript levels relative to the levels of a control transcript (*ubc-2*) in wild type and *aus3* mutant animals. Two regions of the *C34E10.8* transcript were assayed corresponding to nucleotide positions 1556-1700 and 2057-2203 of the *C34E10.8* cDNA, that is, one before and one after the premature stop codon. Reduced levels of *C34E10.8* mRNA were found in *aus3* mutants compared with

A)



B)

MKQTTRPPAQYIVAPGTRFRLVANEYDKNKYGQCNYA
SYRTLVRCKQIRSKEELAKHGGRCEEHVEFSKTLENNH
KKEVMRCHAENDSKMQRRRFDPWIASNEYISDDDDYL
QAAQTVPQRLPDVANDDILDNNSLRYAEYYTDKDILNI
KMDLVQKDIDDLIEFKELVTSQAQKEHELLGNDEEEDY
PTDMAQRRIFKASTKYSRNDYLTLTTIDPVFHQCCVGP
DMDDSLVIVHTMHSILDKIDNFEPIDSEKQCNKPALHLS
KFCFDHIILDRSQKMFDVCNACGLTAIGGVDPKCSFHIK
SSAIAETTSCPCNRCVQPGEHASPKDEKNSITCYLNSSD
DDEPTLGNLSRIESMVSPMQQFSNQSNTLTAPLPRRYQ
GPPAQVLRPPQMGPPPGINQVPYQPKANRTPPMTSQQL
HEQQKLKMQEEEMMSQTCASDFRVRPIDASQFGGGKK
KQRLPPRRSPSFGTSPNSYQFHQQSQKKMPSIISTAYNS
SPGKMNFQGWKNQSTSSATRPLPQPRFPVHAARSQQPK
MIPLEQTQDSIEDDIGPSPMFQGPEPSRRGVPYYKNAYR
RTELPSRHAQHSPLTPSTSTSSSQLLAPPKSPQPGTSSQT
FRSQASRLPIAPHRAIAAGLNPADVGTRPAYRSQMAGQ
RPGMTPSAQQGSPQLISPPRQGSMMPVAMNSTOPSPQAV
RRQTPVPPYRLMGPQRVTTSYTVVRSGSSSSVAGPSRS
SVASGSQHTALDTVQHDPRLANINVRTFLSIGNRDLSTL
TQDEIDLLMAGNSPEKGGRKAGAPGAKESSKAAGGAQ
KGTSAASTSVPEPTKSSESSVDPQSDVSFSNPSPAPEVIE
KVAPAAMTITSNKRKIDETLASESTSSEATLIHDTTSSSS
AETVSGEPPAKKSSDVSAPVPSPEKEKEKIDRPKTPKSS
TKRTTPTPSGRTPRAAAIAANQAISHSKPNVPSASTSSS
AASTDQENPLDLLAELSVAAAAEEQQQAIGSTSKNGGS
TKKTQRKSPSRSSIGKKRENSEEYEEELStop

C)

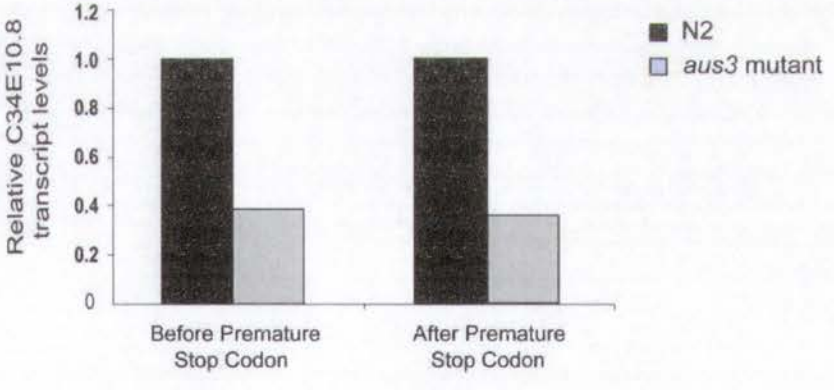


Figure 5.5: The premature stop codon in *aus3* mutants might lead to production of a truncated protein and degradation of the C34E10.8 transcript via non-sense mediated decay (NMD). A) The gene model for C34E10.8 and the position of the premature stop codon in *aus3* mutants is indicated. The nonsense mutation in the *aus3* allele is in the middle of the 15th exon and it is 109 nucleotides apart from the 14/15 exon-exon junction. Exons are shown as blue boxes and introns are indicated as black lines. **B)** The amino

acid sequence of C34E10.8 is shown, with the premature stop codon identified by whole genome sequencing in the *aus3* mutant shown in red. **C)** C34E10.8 transcript levels were quantified by real-time RT-PCR with primer sets corresponding to exon-exon junctions before and after the premature stop codon. Gene expression was normalised to *ubc-2* levels and then normalised again to the wild type values. The values are average of triplicate assays of cDNA derived from one population of worms.

wild type animals, suggesting that the mutant *C34E10.8* transcript in *aus3* mutants may be partially degraded by nonsense mediated decay (Fig. 5.5 C).

The finding of reduced *C34E10.8* transcripts in the *aus3* mutant, together with the prediction that remaining transcripts would encode a truncated protein, suggests that *aus3* may be a hypomorphic allele of *C34E10.8* rather than a null allele; the expression level and activity of the C34E10.8 protein is likely compromised but not entirely abolished.

5.8 C34E10.8 is expressed in the somatic cells

As suggested by the absence of a formal gene name, C34E10.8 has not been the focus of any previous studies. Therefore, in order to discover more about C34E10.8 and, in particular, to clarify how the *C34E10.8(aus3)* mutation affects expression of the *ctbp-1::gfp* transgene, a characterisation of this novel gene and its encoded protein was embarked upon. As a first step, the pattern of expression of C34E10.8 was assessed. For this purpose, transgenic animals carrying a *C34E10.8::gfp* reporter construct were generated by microinjection. The reporter construct was generated by PCR fusion and is composed of the entire coding sequence of C34E10.8 fused to the *gfp* coding sequence, under the control of the putative C34E10.8 promoter (~ 2kb upstream of the start site of translation) (Fig. 5.6 A).

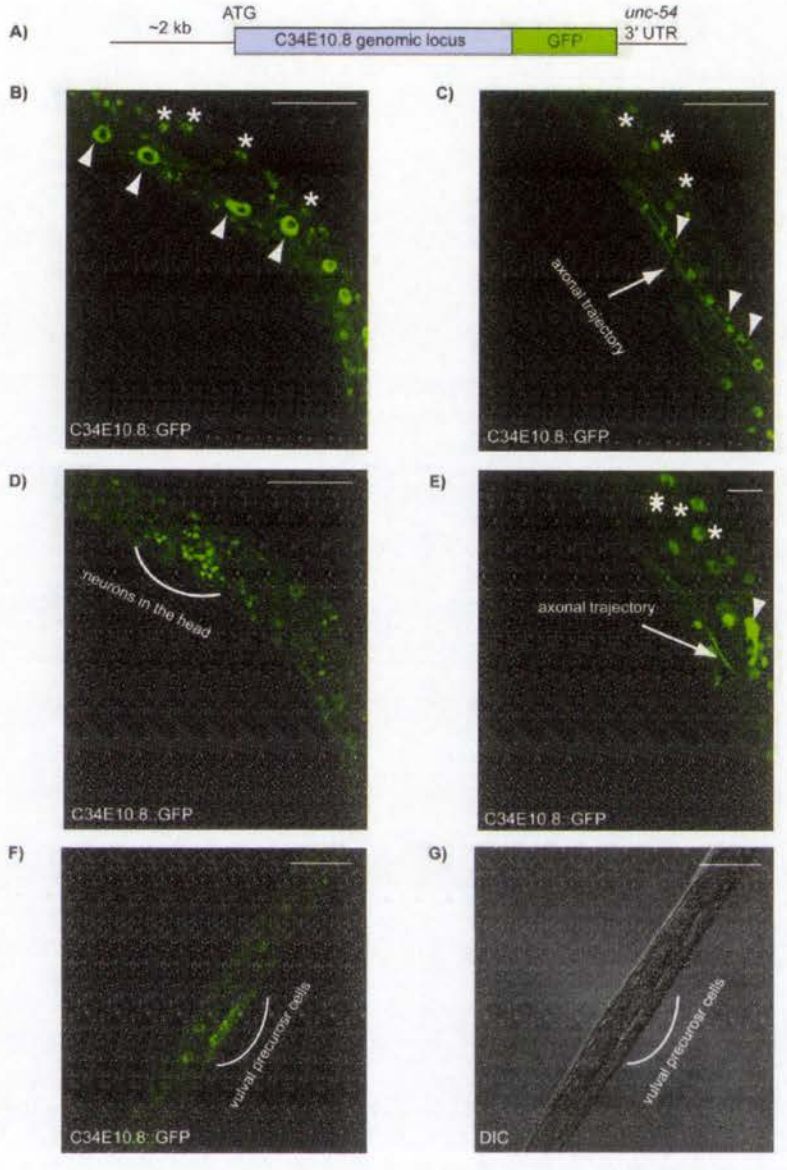


Figure 5.6: C34E10.8 is expressed in many somatic cells. A) The reporter construct carrying the entire coding sequence of full length genomic C34E10.8 with its 2kb promoter was used to generate a transgenic animal expressing C34E10.8::GFP fusion protein for expression pattern analysis. To aid expression of the fusion template the *unc-54* 3′ UTR was used. Analysis of transgenic animals at adult stage has shown that C34E10.8::GFP expression is detected in various tissues. Identification of cells is based on the anatomy of the nematode, morphology of the cells and the size of the nucleolus. **B)** Expression of C34E10.8::GFP is observed in hypodermal cells (indicated as *) and in the intestinal cells (indicated with arrow heads). **C)** Expression of

C34E10.8::GFP is also detected in neurons in the Ventral Nerve Cord some of which are indicated with arrow heads. Hypodermal cell are also in focus and some of them are labelled with *. Arrow indicates axonal trajectory expressing C34E10.8::GFP. **D)** C34E10.8::GFP is expressed in neurons in the head. **E)** C34E10.8::GFP is expressed in the interneurons PVQ (only one is in focus that is indicated with arrow head) and the axonal trajectory sent by these neurons (indicated with arrowhead). Again, some of the hypodermal cells are marked with *. **F)** C34E10.8::GFP is expressed in vulval precursor cells as indicated. **G)** DIC image showing vulval precursor cells. Anterior is to the left, scale bar indicates 50μm.

Transgenic animals carrying the *C34E10.8::gfp* reporter show broad expression of the C34E10.8::GFP fusion protein in many somatic cells throughout all larval stages and into adulthood. In line with standard practice for *C. elegans* reporters, the C34E10.8 reporter construct carries the *unc-54* 3'UTR, rather than the endogenous C34E10.8 3'UTR. As a result, any micro RNA-mediated regulation of C34E10.8 will not take place in the context of this construct.

The expression pattern of C34E10.8::GFP in different tissues has been analysed on the basis of the anatomy of the nematode, morphology of the cells, and the size of nucleolus. C34E10.8::GFP is expressed in the hypodermal cells, some of which are indicated in Figure 5.6 B. Intestinal expression of C34E10.8::GFP is also detected and indicated. C34E10.8::GFP expression is detected in other cell types such as Ventral Nerve Cord neurons (Fig. 5.6 C), neurons in the head (Fig. 5.6 D) as well as in vulval precursor cells (Fig. 5.6 F and G). In the majority of cells, GFP fluorescence is localised to the nucleus as demonstrated by DAPI staining (data not shown). Taken together, according to reporter gene analysis C34E10.8 is expressed in various different tissues and largely localised to the nucleus. Hypodermal expression of C34E10.8 suggests that it might directly regulate CTBP-1 in these cells.

5.9 C34E10.8 shows some homology to chromatin regulator Ino80d

Since C34E10.8 was uncharacterised, to gain an insight into the roles that this nematode protein might play as a broadly expressed nuclear factor, mammalian homologs were sought using BLAST searches. However, initial high stringency searches showed no obvious homology to mammalian proteins. Relaxation of the search parameters allowed the identification of a small segment of similarity (32 amino acids) with a mouse protein called Ino80d (d subunit of multiprotein Ino80 complex) (Fig. 5.7 A).

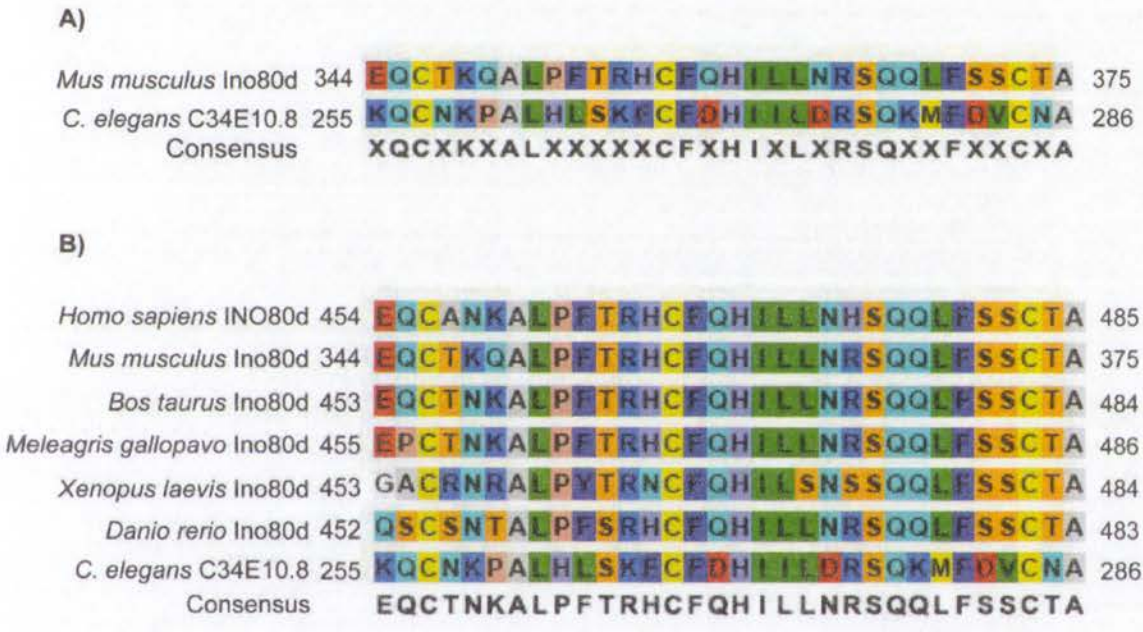


Figure 5.7: C34E10.8 shows a 32-amino-acid homology to Ino80d. A) C34E10.8 is homologous to mouse Ino80d. **B)** C34E10.8 also shows homology to other mammalian and non-mammalian Ino80d proteins. The BLAST searches using Ino80d homologs in *Meleagris gallopavo* (wild turkey), *Xenopus laevis* (frog) and *Danio rerio* (zebrafish) against the *C. elegans* genome have identified the same 32-amino acid sequence as in the case of alignment of mammalian homologs with C34E10.8. Amino acid residues are presented with colors where similar amino acids have same color codes. For instance, Isoleucine (I) and Leucine (L) are both colored in green

since both amino acids are nonpolar and neutral whereas polar and neutral amino acids Threonine (T) and Serine (S) are colored in orange.

These 32 amino acid residues are also largely conserved in Ino80d homologues from other organisms. Figure 5.7 B shows the conserved sequence that is produced by alignment of Ino80d from wild turkey *Meleagris gallopavo*, zebrafish *Danio reiro* and the frog *Xenopus laevis* Ino80d aligned with *C. elegans* C34E10.8 and human and mouse Ino80d. Overall, alignment of Ino80d from different species identified a small segment of sequence conserved across these species.

5.10 C34E10.8 is largely conserved among nematodes

Following on the evolutionary perspective, alignment of C34E10.8 against other nematodes such as *C. remanei* and *C. briggsae* genomes have identified highly conserved hypothetical proteins. C34E10.8 is largely conserved in these nematodes especially at the N terminus (Appendix Figure 1). *C. remanei* homolog CRE_29326 is 49% identical and 61% similar whereas *C. briggsae* homolog CBG_21184 is 59% identical and 70% similar to C34E10.8. Both of these hypothetical proteins have domains identical to C34E10.8 one of which is RING finger domain. At the C-terminus C34E10.8 seems mostly unstructured and interestingly contains a sequence characteristic of a viral core nucleocapsid protein called Gagp24 (Appendix Figure 2). More importantly, the 32 amino acid sequence that was found to be conserved between the putative mammalian Ino80d proteins and C34E10.8 is also largely conserved among the nematodes (Fig. 5.8).

Figure 5.8: C34E10.8 is largely conserved among other nematodes. The 32-amino-acid-long sequence is conserved in other nematodes *C. remanei* and *C. briggsae*. The hypothetical Ino80d proteins in these nematodes (CRE_29326 and CBG_21184) are annotated.

5.11 C34E10.8 and hINO80d both contain a RING finger domain

To further investigate the homology between C34E10.8 and Ino80d, the domain structures of C34E10.8 and the human protein hINO80d were analysed with the use of a software package called Workbench. This analysis revealed that like C34E10.8, hINO80d contains a RING finger (Appendix Figure 2 and 3). The RING (Really Interesting New Gene) finger is a type of zinc finger domain and is composed of 40-60 amino acids containing three cysteine, one histidine and four cysteine residues ($Cys_3HisCys_4$). RING finger domains can function in protein-protein interactions as well as protein-DNA interactions (Freemont, 2000). Figure 5.9 A shows the alignment between the characterised RING domain of the protein human RING1 and putative RING domains of hINO80d and C34E10.8. It should be noted that the distribution of the Cys and His residues in the putative RING finger domains of hINO80d and C34E10.8 does not exactly match that of human RING1 which has cysteine residues that are separated by two amino acids. However, although cysteine residues are not spaced as evenly as in human RING1, the number of cysteine residues (i.e. three cysteine, followed by 4 more cysteine residues after the histidine) and the close proximity of the third cysteine to the

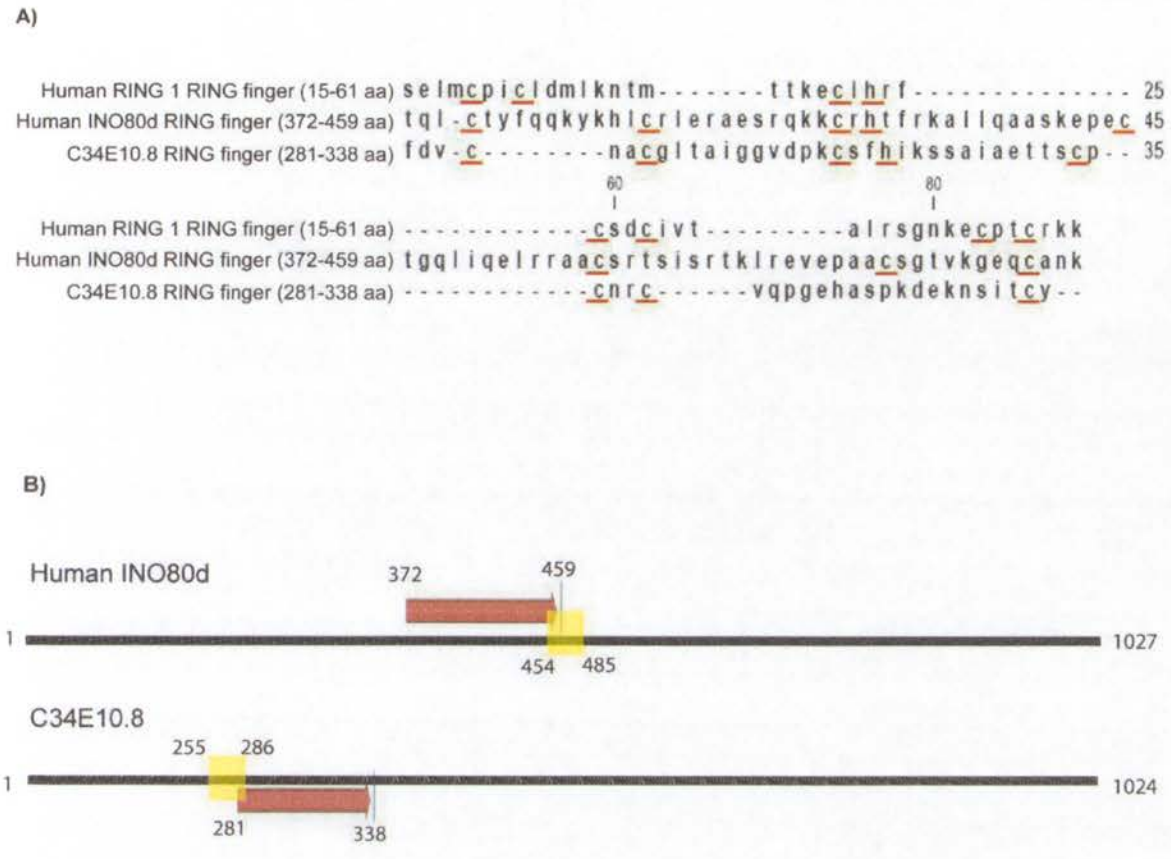


Figure 5.9: C34E10.8 and Ino80d are novel RING finger proteins. A) The alignment of the RING domain of human RING1 with the putative RING finger domains of C34E10.8 and human Ino80d. The Cys and His residues that characterise the RING finger domain are underlined in red. **B)** Overview of C34E10.8 and human INO80d proteins. The arrows indicate putative RING finger domains and the region highlighted in yellow corresponds to the 32-amino-acid conserved sequence. The numbers indicate the residues corresponding to start and end of the RING domain as well as the conserved region.

histidine residue in all three proteins argue that C34E10.8 and hINO80d might present a novel class of RING finger proteins. Indeed, the RING finger domain is rather loosely defined, as exemplified by another *C. elegans* protein called TAM-1, which has been described as a RING finger protein. In this case, the RING domain spans a region of 133 amino acids and the two cysteine residues after the histidine are separated by 85 amino acids from the last two

cysteine residues (Hsieh et al., 1999). Therefore, the arrangement of cysteine residues after the histidine seems largely flexible. As shown in Figure 5.9 B, the putative RING finger domain of C34E10.8 immediately follows the conserved 32 amino acid sequence whereas in hINO80d, the RING finger immediately precedes the conserved region.

5.12 The *C. elegans* genome has homologues of Ino80 complex subunits

As described above, analysis of the C34E10.8 sequence indicated that the closest homologues in non-nematodes are members of the Ino80d family. Since Ino80d is a component of the Ino80 complex and since this complex has not previously been characterised in *C. elegans*, we searched for other nematode components of the Ino80 complex using BLAST. From these searches, a presumptive *C. elegans* INO80.com containing 13 subunits that are either homologous to human or yeast INO80.com subunits was identified (Table 5.4). Human INO80.com has 14 subunits, and 11 putative homologues were found in *C. elegans*. Putative nematode homologues of two components of the yeast Ino80 complex which are not conserved in humans (Taf14 and Nhp10) were also identified. While some of the subunits such as *ruvb-1*, *ruvb-2*, *swsn-6* and *ssl-1* are highly similar to their homologues as indicated with the low E-values, others among the putative nematode homologues of Ino80 complex components show only weak similarity and further investigation is required to determine whether these are indeed functionally homologous to their mammalian counterparts.

Two of the identified genes, namely *ruvb-1* and *ruvb-2* are known homologues of Rvb1 and Rvb2, respectively but have not been previously considered in the context of the Ino80 complex in *C. elegans* (Updike and Mango, 2007). Similarly, *ssl-1* has previously been identified as a nematode homologue of p400 SWI/SNF ATPase (Ceol et al., 2004).

Table 5.4: Identification of Ino80-like complex in *C. elegans*. The previously-characterised subunits of yeast and mammalian Ino80 complex are listed along with a brief description. E-value indicates the homology between the human or yeast Ino80 complex components and the putative *C. elegans* homologues identified in this study.

Yeast INO80	Human Ino80	Putative *C.elegans* homolog	Brief Description	E-value
Ino80	Ino80	ssl-1	p400/SWR1/INO80 ATPase	6.00E-92
Arp8	Arp8	arx-2	ARp2/3 complex component family member	3.00E-12
Arp5	Arp5	act-2	Actin family member	1.00E-19
Arp4	BAF53	swsn-6	Actin-related protein -	3.00E-129
Act1	Actin	actin	Actin family	------
Rvb1	RUVBL1	ruvb-1	RUVB (recombination protein) homolog family member ATPase	0.00E+00
Rvb2	RUVBL2	ruvb-2	RUVB (recombination protein) homolog family member ATPase	0.00E+00
Ies2	Ino80b	T23B5.4	Uncharacterised protein	0.032
Ies6	Ino80c	C17E4.6	YL-1 protein	0.085
not conserved	Ino80d	C34E10.8	Uncharacterised protein	9.00E-04
not conserved	Ino80e	not identified	NA	------
not conserved	NFRKB	not identified	NA	------
not conserved	TCF3 fusion partner	not identified	NA	------
not conserved	MCRS1	H28O16.2	FHA domain	2.00E-33
Taf14	not conserved	Y105E8B.7	YEATS domain	1.00E-08
Nhp10	not conserved	hmg-3	Nucleosome-binding HMG family member	7.00E-10
Ies1	not conserved	not identified	NA	------
Ies3	not conserved	not identified	NA	------
Ies4	not conserved	not identified	NA	------
Ies5	not conserved	not identified	NA	------

In summary, these results suggest the presence of an INO80-like complex in *C. elegans* which includes the Ino80d-like protein C34E10.8.

5.13 C34E10.8 functions as a SynMuv suppressor

We surveyed the literature to learn more about C34E10.8 and found that it had been identified in a large-scale RNAi screen for suppressors of the SynMuv pathway which affects vulval development (Cui et al., 2006b). As reviewed in Chapter 1, the *C. elegans* vulva develops from cells of the ventral hypodermis called the vulval precursor cells (VPCs), of which there

are six. During the L3 stage, the gonadal anchor cell expresses the epidermal growth factor LIN-3, which instructs the central three VPCs to execute vulval fates. The remaining three VPCs instead fuse with the surrounding hypodermis (Fig. 5.10 A). If the LIN-3 growth factor

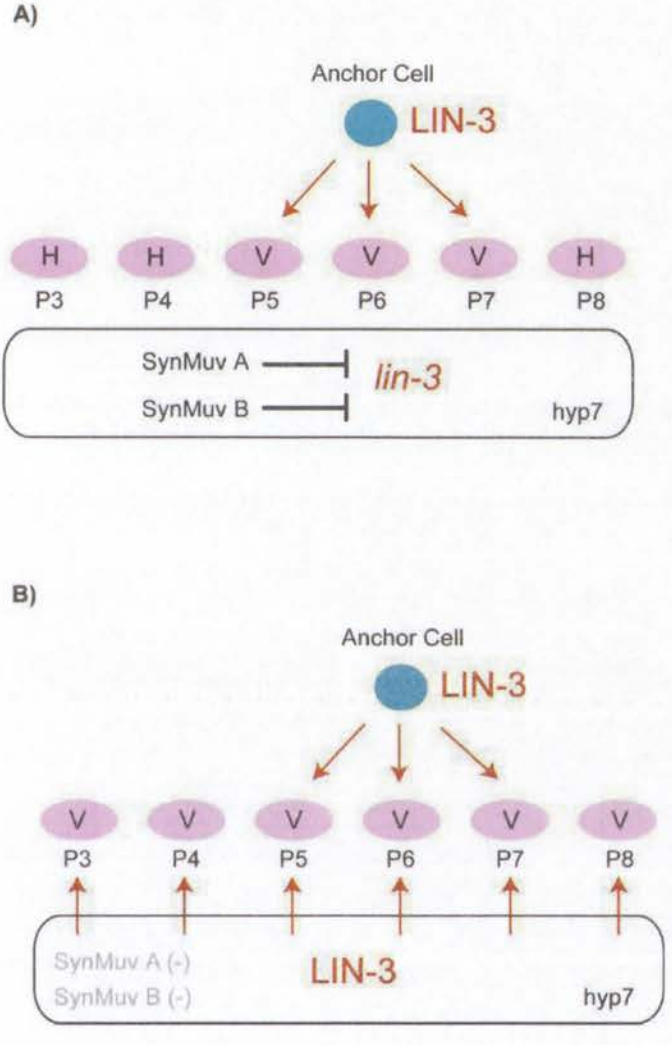


Figure 5.10: The model illustrating the mechanism by which SynMuv genes regulate vulval development. **A)** In wild type animals, anchor cell (AC) induces vulval induction of P5, P6 and P7 blast cells and P3, P4 and P8 acquire hypodermal cell fate and fuses with the hyp7 syncytium. **B)** In animals with mutations in both SynMuv A and SynMuv B class of genes, the hyp7 cell ectopically secretes LIN-3 resulting in induction of all P cells for vulval cell fate. Adapted from Cui *et al.* (Cui et al., 2006b)

is inappropriately expressed in cells other than the anchor cell, additional VPCs execute vulval fates, giving rise to a multiple vulva (Muv) phenotype (Fig. 5.10 B). Inappropriate expression of LIN-3 in the syncytial hypodermal cell hyp7 is prevented by proteins encoded by the synthetic multivulva (SynMuv) genes. These genes are grouped into several classes including SynMuvA and B, which act redundantly to repress *lin-3* expression. Single mutation of a class A SynMuv gene or a class B SynMuv gene does not cause a multivulva phenotype, but the simultaneous mutation of both a SynMuvA gene and a SynMuv B gene, relieves repression of *lin-3* in hyp7, resulting in a SynMuv phenotype.

The screen performed by Cui *et al.* (2006) set out to identify RNAi treatments which suppressed the SynMuv phenotype of a *lin-15AB(n765)* mutant (which carries the mutant allele *n765* that disrupts the function of both the SynMuv A gene *lin-15A* and the SynMuv B gene *lin-15B*) . Through this screen 32 SynMuv suppressor genes were identified, one of which is C34E10.8.

It was therefore of great interest to assess whether the *aus3* mutation in C34E10.8 could also suppress the SynMuv phenotype of *lin-15AB(n765)*, to confirm the designation of *C34E10.8* as a SynMuv suppressor gene. To this end, *C34E10.8(aus3)* mutants were crossed with *lin-15AB(n765)* mutants and a strain that is homozygous for both the *aus3* mutation and the *lin-15AB(n765)* mutation was generated. Animals of this strain were then assessed for the multivulva phenotype. Only 14% of the *lin-15AB(n765); C34E10.8(aus3)* mutants displayed a multivulva phenotype while 100% of *lin-15AB(n765)* mutants show this phenotype (Figure 5.11 A). The representative images show *lin-15AB(n765)* mutants with ectopic vulva

indicated and *lin-15AB(n765); C34E10.8(aus3)* mutants with wild type-like vulva (Figure 5.11 B and C). These observations confirm that *C34E10.8* is a SynMuv suppressor gene.

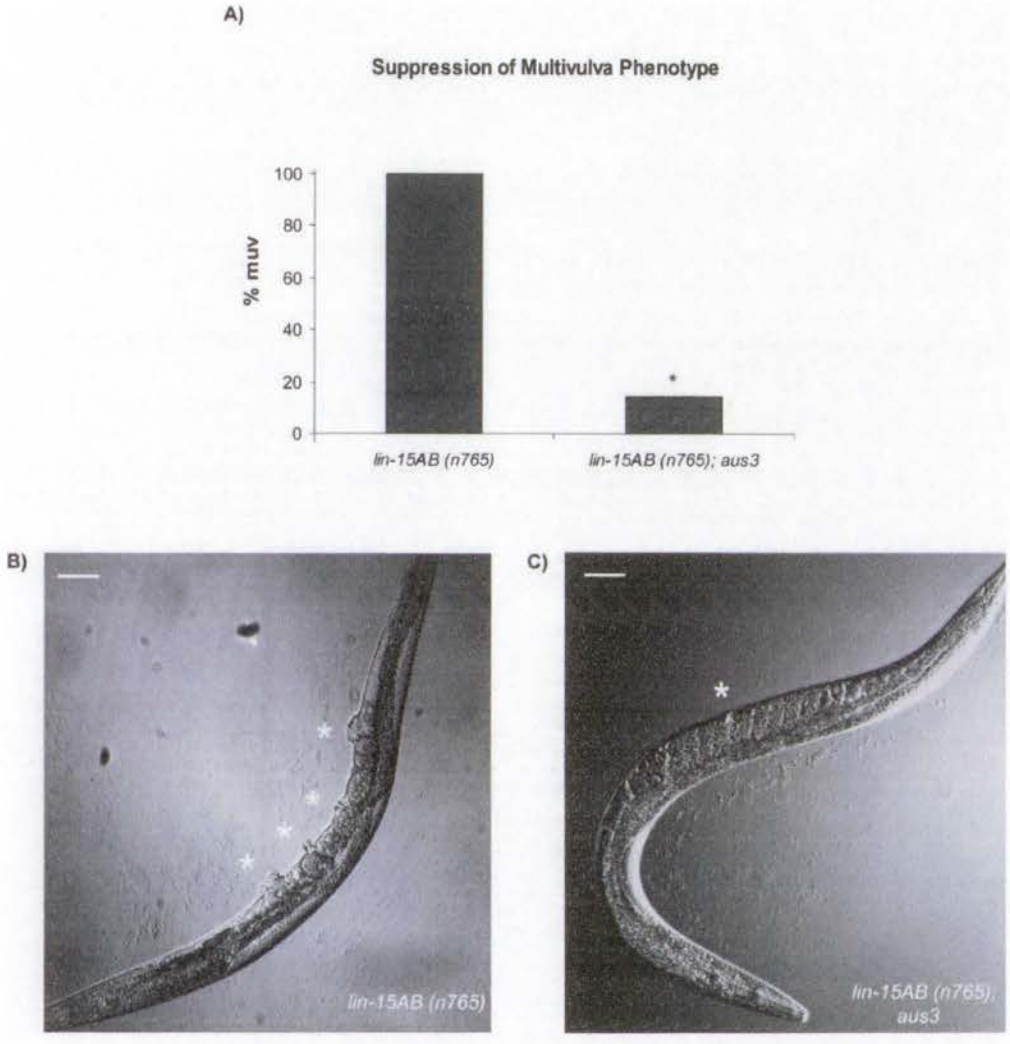


Figure 5.11: The *aus3* mutation suppresses the multivulva phenotype of *lin-15AB(n765)* mutants. A) The graph indicates suppression of multivulva phenotype in *aus3* mutants. n=297 for *lin-15AB(n765)* and n=242 for *lin-15AB(n765); aus3.* The significance test was employed on the average of percentage multivulva animals obtained from two independent plates of worms (*, $P < 0.05$, Student's *t* test). Representative images of **B)** *lin-15AB(n765)* mutant **C)** and *lin-15AB(n765)* in *aus3* mutant background. Scale bar indicates 50µm.

Interestingly the suppression of the *lin-15AB(n765)* multiple vulva phenotype by the *C34E10.8(aus3)* mutation is not fully penetrant; 14% of animals are still Muv. While RNAi of C34E10.8 also causes incomplete suppression, with 3% of the RNAi-treated population still displaying the Muv phenotype(Cui et al., 2006a), the suppression achieved by the RNAi treatment is greater than that produced by the *aus3* mutation. This supports the earlier assertion that *aus3* may be a hypomorphic allele of *C34E10.8*.

5.14 Discussion

In this chapter, we used a mutagenesis screen to look for factors that regulate CTBP-1 and identified a mutant with increased levels of CTBP-1::GFP expression in the hypodermal cells. Through next generation sequencing, we identified a number of candidate mutations which might account for the overexpression phenotype. We then confirmed the causative mutation by introducing wild type copies of the candidate gene *C34E10.8*, which rescued the overexpression phenotype.

In this mutagenesis screen, it was expected that the likely changes to CTBP-1::GFP expression that would be detected would be either at the transcriptional level, affecting expression of the reporter gene, or at the protein level, affecting the stability of the CTBP-1::GFP fusion protein. Although we have not directly demonstrated that C34E10.8 regulates transcription of the *ctbp-1::gfp* transgene, the information gathered on C34E10.8 is suggestive of a role in regulation of the transgene at the transcriptional level. First, C34E10.8 encodes a protein with weak similarity to a subunit of ATP-dependent chromatin remodelling complex Ino80, Ino80d. Second, C34E10.8 has been shown to physically associate with DPY-30 which is a component of COMPASS chromatin remodelling complex (Pferdehirt et al., 2011). Since chromatin remodelling activities contribute to the transcriptional control of gene expression,

these data linking C34E10.8 to the chromatin remodelling machinery support a role for C34E10.8 in the transcriptional regulation of *ctbp-1::gfp* and are considered in more detail below.

5.14.1 C34E10.8 might be the nematode Ino80d

The human INO80 complex is composed of 14 subunits, 8 of which are conserved from yeast to man. On the other hand, 5 of these subunits, including Ino80d, are metazoan-specific and are not required for ATP-dependent nucleosome remodelling activity of human Ino80 complex (Chen et al., 2011). On the basis of their domain structure, these metazoan-specific subunits are proposed to play roles in mediating DNA-binding or acting as adaptor proteins between the main Ino80 scaffold and other subunits (Jin et al., 2005). In particular, Ino80d is thought to mediate protein-DNA or protein-protein interactions on the basis of a characteristic pattern of histidine and cysteine residues. The human Ino80 complex has roles in transcription and DNA repair (Cai et al., 2007; Conaway and Conaway, 2009).

C34E10.8 shows a homology to mammalian Ino80d through a 32-amino-acid conserved sequence. Interestingly, both C34E10.8 and human Ino80d have putative RING finger domains therefore, it is likely that C34E10.8 and Ino80d mediate protein-protein or protein-DNA interactions. In addition, both proteins have distinct predicted domains (Appendix Figure 2 and 3). In the light of similar and distinct domains, it is possible that mammalian and nematode Ino80d have gained or lost domains during evolution, conferring new functions to these proteins. Most of the components of yeast and human chromatin factors have been assigned as homologues due to their similar domains rather than similarity in their entire amino acid sequence. Moreover, there are a number of examples of proteins in the nematode that show no or little sequence homology to their mammalian counterparts but are

functionally homologous. For instance, the gap junction proteins in *C. elegans* called innexins, show no sequence homology to vertebrate homolog connexins (Starich et al., 2003). However, connexins and innexins display the same function of connecting cells for intercellular transfer of small molecules. Taken together, C34E10.8 holds the potential to be a functional homolog of Ino80d through conservation of a short amino acid sequence.

5.14.2 How might C34E10.8 function as a chromatin regulator?

If indeed C34E10.8 is the nematode homologue of Ino80d, it is expected to function as part of an Ino80-like chromatin remodelling complex. In this study, BLAST searches enabled the identification of nematode homologues of several additional components of the mammalian Ino80 complex (Table 5.4), supporting the existence of a nematode chromatin remodelling complex analogous to the mammalian Ino80 complex.

Although this putative nematode Ino80 complex has not previously been reported, a related chromatin remodelling complex called TIP60 has been described in *C. elegans* (Ceol and Horvitz, 2004). Both mammalian and nematode TIP60 complexes incorporate two functional units, the SWR1-like ATP-dependent chromatin remodelling complex and the NuA4 histone acetylation complex (Fig. 5.12). Like the catalytic subunit of the mammalian INO80 complex, called INO80, the catalytic subunit of the TIP60 complex, p400, is characterised by a split-ATPase domain (Ho and Crabtree, 2010). Beyond this similarity, the Ino80 and TIP60 complexes are related by the sharing of several components (eg BAF53, actin, RUVBL1/2) (Figure 5.12).

Interestingly, it appears that the nematode genome encodes only a single homologue of the split-ATPases INO80 and p400, called SSL-1. Thus, the putative INO80 complex that we

have described contains the same catalytic subunit as the previously-characterised nematode TIP60 complex. We propose that SSL-1 may associate with distinct accessory proteins to assemble either a TIP60-like complex or an INO80-like complex. C34E10.8 may then function as part of this INO80-like complex to regulate chromatin.

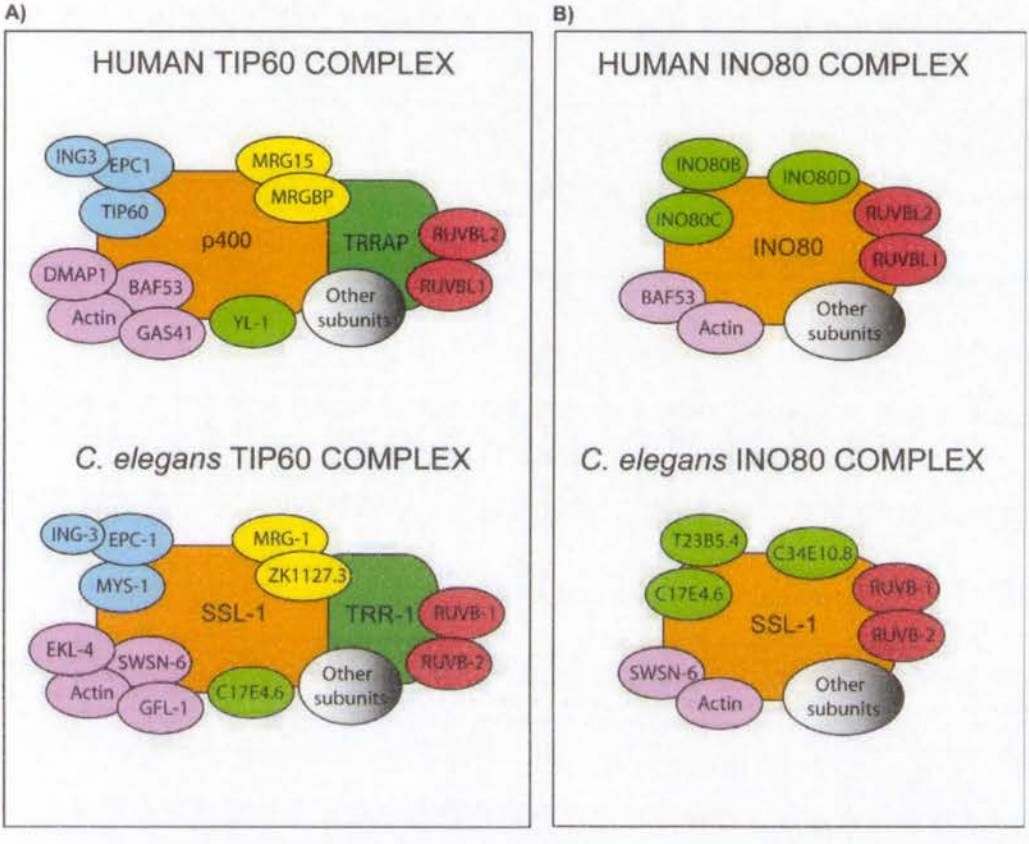


Figure 5.12: The main assembly of Human TIP60 and INO80 complexes and predicted *C. elegans* TIP60 and INO80 complexes are illustrated. A) Human TIP60 subunits assemble around the catalytic subunits p400 and TRRAP. Only subunits with *C. elegans* homologues are included. The p400 and subunits associated with it (colored in blue, yellow and pink) are homologues to yeast SWR1 complex subunits. The TRRAP and subunits RUVBL1-2 resemble the NuA4 complex in yeast. Subunits colored in pink are shared in yeast SWR1 and NuA4 complexes **B)** Human INO80 complex assembles around the catalytic subunit INO80 ATPase. Predicted *C. elegans* INO80 complex has the same catalytic subunit as in TIP60 complex. Only human INO80 subunits with *C. elegans* homologues are included. The color-coding indicates the homology between the subunits in human and *C. elegans* TIP60 and INO80 complexes.

One other mechanism by which C34E10.8 might participate in the regulation of chromatin is through its interaction with DPY-30. In a large scale yeast-two hybrid screen, C34E10.8 was reported to physically interact with DPY-30 (Simonis et al., 2009). In *C. elegans*, DPY-30 serves at least two functions. First, it regulates gene expression on the X chromosome through the dosage compensation complex (Hsu et al., 1995). DPY-30 also functions outside of dosage compensation as a component of the MLL/COMPASS histone methyl transferase complex (Pferdehirt et al., 2011). Interestingly, DPY-30 was identified in the same large scale RNAi screen looking for modulators of vulval development that identified C34E10.8 as an uncharacterised SynMuv suppressor (Cui et al., 2006a). Moreover, the same group also reported a SynMuv suppressor role for another uncharacterised protein called F54D11.2 which is also reported to be an interacting partner of DPY-30 (Cui et al., 2006a; Simonis et al., 2009). Therefore, it is possible that C34E10.8, DPY-30 and F54D11.2 act together to suppress the SynMuv pathway, perhaps through the regulation of chromatin.

5.14.3 Does the aus3 mutation have an effect on endogenous ctbp-1?

The outstanding question of this piece of work is whether C34E10.8 regulates the endogenous *ctbp-1* locus in the same way that it regulates the *pctbp-1::ctbp-1::gfp* transgene. One possibility is that C34E10.8 is involved in silencing of transgenes rather than silencing the *ctbp-1* promoter in particular. We tested this hypothesis by using a different transgene, *lag-2::gfp(qIs56)*, and looked at whether the expression of LAG-2::GFP changes in *aus3* mutants. Unlike the *pctbp-1::ctbp-1::gfp* transgene, the *lag-2::gfp* transgene did not show a detectable change in the expression levels in *aus3* mutants compared to the wild types (data not shown). With the elimination of C34E10.8 acting as a silencer of transgene arrays, several attempts

were made to address the question of whether C34E10.8 regulates the endogenous *ctbp-1*. However, none of these yielded conclusive outcomes.

First, an anti-CTBP-1 antibody was obtained with the view to using Western blots of whole worm protein lysates as well as immunohistochemistry to assess whether levels of the endogenous CTBP-1 protein are increased in the *C34E10.8(aus3)* mutant strain. To date, however, we have not been able to detect endogenous CTBP-1 with this antibody and so have not been able to make the necessary comparisons.

Additionally, real-time RT PCR was used to assess whether levels of the endogenous *ctbp-1* transcript are increased in the *C34E10.8(aus3)* mutant strain. Assays of RNA extracted from animals at the young adult stage showed no difference in *ctbp-1* transcript levels between wild type and *C34E10.8(aus3)* mutants (data not shown). Although this may suggest that C34E10.8 does not regulate expression of endogenous *ctbp-1* it is also possible that we were simply unable to detect the small increase in transcripts that is predicted based on experiments with the *pctbp-1::ctbp-1::gfp* transgene. That is, since only a 1.5 fold increase in CTBP-1::GFP fluorescence in the hypodermis had been observed in the *C34E10.8(aus3)* mutant and no difference in neuronal expression had been observed in this strain, the net change in endogenous *ctbp-1* expression in a whole animal is expected to be less than 1.5 fold and therefore likely to be undetectable with the whole animal real-time RT PCR method employed here.

Interestingly, there is no discernable change in the level of expression of CTBP-1:GFP in the neuronal tissue in *aus3* mutants as compared with wild type. This may be attributable to the high intensity of the GFP signal in the wild type neurons which might compromise our

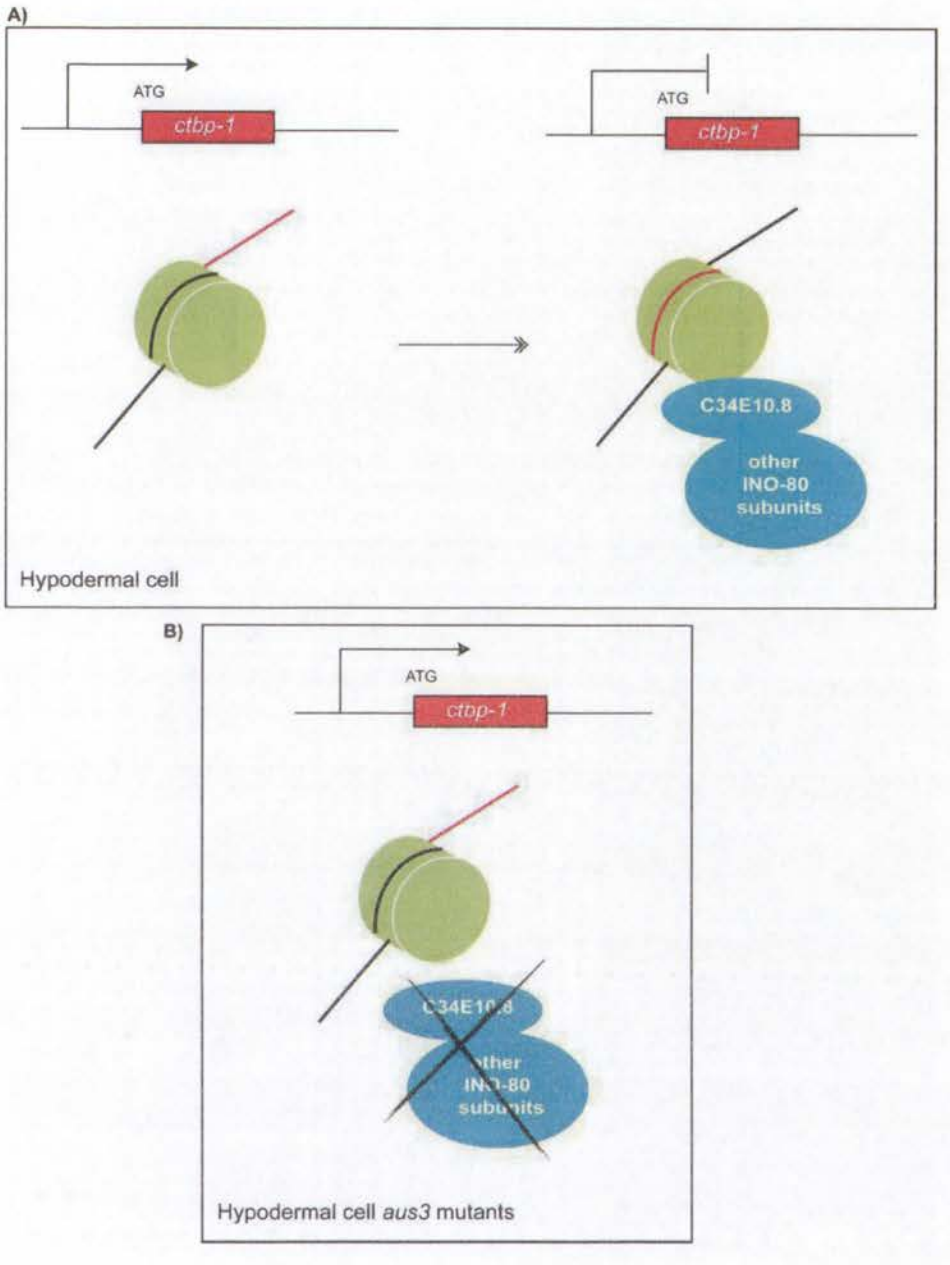


Figure 5.13: Model illustrating regulation of *ctbp-1* by C34E10.8 in hypodermal cells. A) In wild type hypodermal cells, C34E10.8 represses expression of *ctbp-1* by associating with INO80 complex. **B)** In *aus3* mutant hypodermal cells, C34E10.8 can no longer mediate repression of *ctbp-1*.

capacity to resolve a difference in the expression levels in the *aus3* mutant neurons. It is also possible that C34E10.8 selectively regulates CTBP-1 in the hypodermal cell but not in the

neurons. This cell type-specific activity of C34E10.8 is supported by the observation that although the *pctbp-1::ctbp-1::gfp(ausIs1); aus3* strain carrying wild type copies of C34E10.8 shows significantly reduced expression of CTBP-1::GFP in the hypodermis, it does not show detectable reduction of expression in neuronal CTBP-1:GFP expression by fluorescence microscopy (see Figure 5.4). Taken together, C34E10.8 is likely to be involved in context-dependent regulation of transcription of *ctbp*-1, repressing *ctbp-1* specifically in the hypodermal cells (Fig. 5.13).

Given the limitations described above, alternative approaches will be required to conclusively determine whether C34E10.8 is a genuine regulator of *ctbp-1* expression in the hypodermis of wild type worms. It is nonetheless of interest to consider the functional importance of such regulation, in particular with reference to the development of the vulva.

5.14.4 The potential roles of C34E10.8 and ctbp-1 in vulval development

As elaborated in Results Section 5.13, SynMuv genes regulate vulval development by acting in hyp7 to inhibit expression of lin-3/EGF in these cells, thereby preventing additional Vulval Precursor Cells (VPCs) from acquiring vulval cell fates (Cui et al., 2006b). In this way, SynMuv proteins ensure development of a single vulva. Upon reduction in SynMuv gene activities, LIN-3 expression is elevated in hyp7, inducing additional VPCs to take on vulval fates (see Fig. 5.14 A). Recently, another group of genes that influence vulval development have been identified, called SynMuv suppressors. Reduced SynMuv suppressor activity in SynMuv A and B-depleted hyp7 cells leads to suppression of the ectopic vulval induction (Fig. 5.14 B). One mechanism by which this could be achieved is via repression of LIN-3 expression in hyp7 cells. In the case of the SynMuv suppressor that has been studied here, C34E10.8, it is possible that reduction of *C34E10.8* function leads to repression of LIN-3

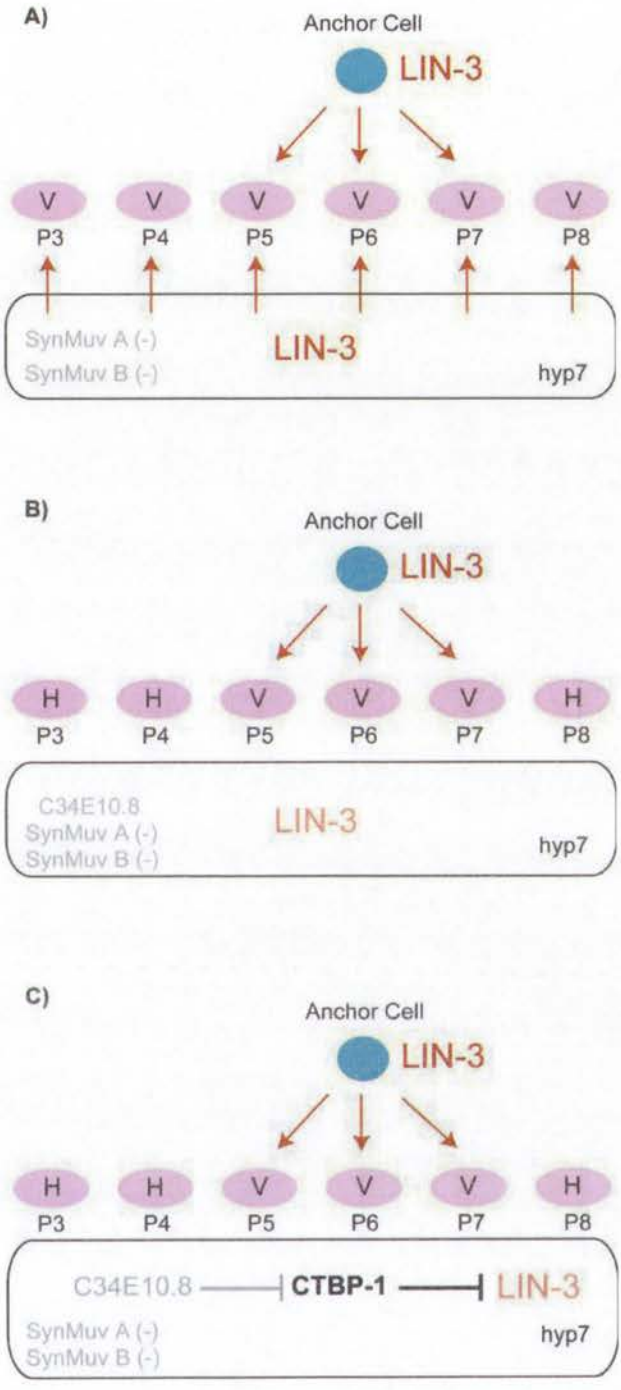


Figure 5.14: The proposed mechanism by which C34E 10.8 regulates vulval cell fate via downregulating *ctbp-1*. **A)** In an animal with mutations in SynMuv A and B class of genes, LIN-3 is ectopically expressed in hyp7 syncytium which results in induction of all Vulval Precursor Cells (P3-8) to acquire a vulval fate (indicated as V). **B)** SynMuv suppressors antogonise the SynMuv gene activities, therefore in SynMuv A and B mutant

animal, reduced function of a SynMuv suppressor (i.e., C34E10.8) leads to re-instatement of repression of *lin-3* expression in the hyp7 and correct induction of only P5-P7. **C)** Since CTBP-1 is a repressor protein, it is possible that mutation of *C34E10.8* suppresses the SynMuv phenotype by increasing the expression of endogenous CTBP-1, which in turn represses *lin-3* expression in hyp7. The proposed mechanism is based on the model suggested by Cui *et al.* (Cui et al., 2006b)

expression in hyp7 via CTBP-1. In this scenario, CTBP-1 is upregulated in hyp7 by loss of *C34E10.8* function and CTBP-1 then directly represses lin-3/EGF expression (Fig. 5.14 C). That is, CTBP-1 can associate with and repress the *lin-3* promoter, with increased levels of the CTBP-1 co-repressor compensating for the loss of SynMuv A and SynMuv B repressor activities.

The model proposed above suggests that *ctbp-1* itself might be a SynMuv gene, with CTBP-1 functioning in wild type animals to repress expression of *lin-3* in hyp7. Alternatively, CTBP-1 may not normally target the *lin-3* promoter but, when overexpressed as observed in the case of the *C34E10.8(aus3)* mutation, this repressor may be incorrectly recruited to the *lin-3* promoter, again compensating for the absence of the usual repressors of this promoter. To test whether *ctbp-1* is a SynMuv gene, strains carrying both the *ctbp-1(ok498)* mutation and either a mutation in a class A SynMuv gene (*lin-8*) or a class B SynMuv gene (*lin-36*) were generated and scored for the presence of the multiple vulva phenotype (data not shown). Although *lin-8(n111); lin-36(n766)* double mutants display the SynMuv phenotype (Thomas and Horvitz, 1999), neither *lin-8(n111); ctbp-1(ok498)* double mutants nor *lin-36(n766); ctbp-1(ok498)* double mutants showed this phenotype. Although these preliminary observations showed no synthetic interaction between the *ctbp-1(ok498)* mutation and either a SynMuv A or a SynMuv B mutation, and therefore do not support the notion that *ctbp-1* is a SynMuv gene, further investigations will be required to resolve this. As suggested in

preceding chapters, the *ok498* allele of *ctbp-1* is likely not to be null. Therefore, a stronger *ctbp-1* allele might yet reveal SynMuv activity for *ctbp-1*. This and other aspects of the proposed model await further investigation.

In summary, although we can not yet ascribe a specific role for C34E10.8 in chromatin regulation, with its homology to Ino80d and interaction with chromatin regulator DPY-30, we conclude that C34E10.8 holds the potential to regulate gene expression through the modification or remodeling of chromatin. We have confirmed that *C34E10.8* is a SynMuv suppressor gene, and proposed a model by which suppression of the SynMuv phenotype may be achieved through *C34E10.8*-dependent regulation of *ctbp-1*expression.

CHAPTER 6

GENERAL DISCUSSION

6.1 Summary

The aim of this thesis was to bring an understanding to the roles and regulation of CtBP proteins using *C. elegans* as a model. A role for *C. elegans* CTBP-1 in the specification of a subset of cholinergic neurons has been revealed and a neuronal transcription factor (ZAG-1) with which CTBP-1 works to perform this role has been identified. Furthermore, putative CTBP-1 target genes with distinct roles in the nervous system ranging from the behavioural response of the nematode to neurodegeneration have been found. In addition, the regulation of CTBP-1 was explored using a mutagenesis screen, and a novel regulator of gene expression called C34E10.8 was identified and characterised. In this way, several important insights into the *in vivo* functions and regulation of a member of the CtBP family of co-repressors have been obtained.

6.2 CtBP family of proteins play distinct neuronal roles

The CtBP family of proteins are vital for mammalian development as suggested by the embryonic lethality of the double KO mice (Hildebrand and Soriano, 2002). Due to the redundancy and lethality, the function of CtBPs has largely been explored using mammalian cell lines, which provided evidence of a role for the CtBP proteins in a number of biological processes such as the epithelial to mesenchymal transition (EMT), apoptosis, and metabolism (Grooteclaes et al., 2003; Grooteclaes and Frisch, 2000; Kajimura et al., 2008; Zhang et al., 2003). Moreover, the CtBP family of proteins are reported to play roles in the nervous system via their interaction with neuronal factors. For instance, CtBP is implicated in proper patterning of the cerebral cortex in humans via interaction with BMP signalling component TGIF (Melhuish and Wotton, 2000) and in metabolic regulation of gene expression via its interaction with neuronal transcription factor NRSF (neural silencing factor) in epileptogenesis (Garriga-Canut et al., 2006). In addition, RIBEYE, an isoform of CtBP2, is

found in abundance at the synaptic ribbons which are unusual structures found uniquely in the vertebrate sensory organs of retinal neurons. RIBEYE is implicated in the formation of synaptic ribbons thus having a role outside the nucleus, however, how it functions molecularly to regulate synaptic ribbon assembly is unclear (Schmitz et al., 2000). Overall, CtBP family of proteins play distinct roles in the nervous system depending on the context. However, these experiments do not bring an understanding on how CtBP is involved in neuronal regulation *in vivo*. Here, we described a precise role for *C. elegans* CTBP-1 in the cholinergic neurons, and have shown a unique feature for *C. elegans* CTBP-1 with its predominantly neuronal expression pattern as opposed to the ubiquitous expression pattern of vertebrate CtBPs. Moreover, *C. elegans* CTBP-1 shows a nuclear localisation whereas vertebrate CtBPs are found both in the nucleus and cytoplasm and display cytoplasmic functions as detailed above. These observations suggest that unlike its vertebrate counterparts, *C. elegans* CTBP-1 may function solely in gene regulation.

6.3 CTBP-1 has a role in neuron specificity and targets neuronal genes with various functions

C. elegans neuronal circuitry is formed of a diverse range of neurons with various functions in chemosensation, locomotion, feeding and egglaying. The nervous system of *C. elegans* is comprised of 118 different classes of neurons which are regulated in a combinatorial fashion for correct specification (Hobert, 2010). In this study, we found a role for CTBP-1 in the specification of DA neurons possibly through indirect regulation of UNC-4. Transcription factors UNC-3, UNC-4 and VAB-7 act in combination to regulate diversification of subgroups of cholinergic neurons (namely VA, DA and VB, DB) (Kratsios et al., 2011; Esmaeili et al., 2002; Von Stetina and Miller, 2002). Further investigation is required for deciphering the mechanism by which CTBP-1 regulates DA specification in combination with

the aforementioned transcription factors.

In this study, a neuronal transcription factor ZAG-1 was shown to interact with CTBP-1 and the phenotypic analysis of *ctbp-1* and *zag-1* mutants suggested these two proteins might act together to regulate specification of DA neurons. Further insights into the neuronal role of CTBP-1 were brought from a comparison of gene expression profiles of *ctbp-1* mutant and wild type neurons. A number of transcripts were identified in this analysis and, as suggested by the neuronal expression pattern of these putative targets, CTBP-1 might directly target these genes to regulate aspects of neuronal specification or development. ChIP and reporter gene analysis will elucidate whether these genes identified with increased levels in *ctbp-1* mutant neurons are *bona fide* CTBP-1 targets. Moreover, it will be interesting to identify shared targets of ZAG-1 and CTBP-1. To date, targets of ZAG-1 have not been reported. The genes identified as CTBP-1 targets could be tested for ZAG-1-dependent regulation using reporter gene analysis.

Given the broad and continual expression of CTBP-1 in the nervous system, CTBP-1 is likely to regulate additional aspects of neuronal development or function. Further investigation will enlighten such roles of CTBP-1 in the nervous system.

6.4 The novel protein C34E10.8 might act in a chromatin regulatory complex

In addition to the described neuronal roles of CTBP-1, reporter gene analysis indicated a low level of CTBP-1 expression in the hypodermis, suggesting roles for CTBP-1 in gene regulation outside of the nervous system. Subsequently, in a screen for regulators of CTBP-1, a mutant (*aus3)* with elevated levels of CTBP-1::GFP in the hypodermal cells was identified. Through whole genome sequencing the mutated gene was found to be *C34E10.8*, which

encodes a protein that displays a weak homology to human INO80d, a subunit of an ATP-dependent chromatin remodelling complex INO80. In light of a previous report suggesting a role for C34E10.8 in the regulation of vulval cell fate, *aus3* mutants were analysed and C34E10.8 was confirmed to act in vulval development as a SynMuv suppressor gene. Numerous other SynMuv suppressor genes encode proteins with known roles in chromatin regulation, for example *isw-1*, which is homologous to the catalytic subunit of ATP-dependent chromatin remodelling complex ISWI (Andersen et al., 2006). Together with the similarity of C34E10.8 to INO80d, this phenotypic evidence suggests that C34E10.8 might be a novel regulator of chromatin. Further support for this hypothesis comes from the finding that C34E10.8 physically interacts with DPY-30, which is a component of COMPASS methylation complex (Hsu et al., 1995) and is also reported to have a SynMuv suppressor role (Cui et al., 2006).

An INO80-like complex has not previously been reported in *C. elegans*, however a TIP60 complex has been described (Ceol and Horvitz, 2004). Interestingly, some of components of the TIP60 complex in humans are shared with the human INO80 complex. Similarly, subunits of the predicted *C. elegans* INO80-like complex proposed in this study are shared with the *C. elegans* TIP60 complex. Several of these, like C34E10.8, are implicated in vulval development. It will be of great interest to find out whether an INO80 complex is present in *C. elegans* and whether C34E10.8 is indeed a component of this complex. Since the vertebrate INO80 complex is relatively uncharacterised, these investigations stand to shed important light on both the *in vivo* activities of the complex and on the molecular functions of the components of the complex including Ino80d.

6.5 Final summary

The results detailed in this thesis provide several insights into the function and regulation of the transcriptional co-repressor protein CTBP-1. These include the demonstration of an *in vivo* role for CTBP-1 in the nervous system. The physical interaction between ZAG-1 and CTBP-1, together with similar neuronal defects in *zag-1* and *ctbp-1* mutant animals, suggest a role for these proteins in specification of cholinergic DA neurons. This is the first demonstration of a specific role for a member of the CtBP family in the acquisition of cell-type specific characteristics within the nervous system and raises the question of whether CtBPs might play a similar role in the mammalian nervous system. Moreover, through a mutagenesis screen, C34E10.8 was identified as a putative regulator of CTBP-1, and through sequence analysis, this protein was found to share a small region of homology with INO80d, a chromatin regulatory factor. Through mutant analysis, C34E10.8 was found to play a role in the regulation of vulval development. Thus, through these investigations on CTBP-1, a putative novel chromatin regulatory factor has been characterised.

REFERENCES

Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001). A regulatory cascade of three homeobox genes, ceh-10 ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C-elegans. *Development* **128**, 1951-1969.

Andersen, E. C., Lu, X. W. and Horvitz, H. R. (2006). C-elegans ISWI and NURF301 antagonize an Rb-like pathway in the determination of multiple cell fates. *Development* **133**, 2695-2704.

Antebi A., N. C. R. a. H. E. M. (1997). Cell and growth cone migrations in C. elegans II: Cold Spring Harbor Laboratory Press.

Anyanful, A., Sakube, Y., Takuwa, K. and Kagawa, H. (2001). The third and fourth tropomyosin isoforms of Caenorhabditis elegans are expressed in the pharynx and intestines and are essential for development and morphology. *Journal of Molecular Biology* **313**, 525-537.

Apfeld, J. and Kenyon, C. (1998). Cell nonautonomy of C-elegans daf-2 function in the regulation of diapause and life span. *Cell* **95**, 199-210.

Ash, P. E. A., Zhang, Y. J., Roberts, C. M., Saldi, T., Hutter, H., Buratti, E., Petrucelli, L. and Link, C. D. (2010). Neurotoxic effects of TDP-43 overexpression in C-elegans. *Human Molecular Genetics* **19**, 3206-3218.

Bao, Y. H. and Shen, X. T. (2007). INO80 subfamily of chromatin remodeling complexes. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **618**, 18-29.

Barnes, C. J., Vadlamudi, R. K., Mishra, S. K., Jacobson, R. H., Li, F. and Kumar, R. (2003). Functional inactivation of a transcriptional corepressor by a signaling kinase. *Nature Structural Biology* **10**, 622-628.

Bettinger, J. C., Bolling, Mia H., Leung, Kapo, Alaimo, Joseph T., Davies, Andrew G. (2011). Membrane lipid environment is important for the development of acute functional tolerance to ethanol. In *International Worm Meeting*, Los Angeles, CA.

Bigelow, H., Doitsidou, M., Sarin, S. and Hobert, O. (2009). MAQGene: software to facilitate C. elegans mutant genome sequence analysis. *Nature Methods* **6**, 549-549.

Blelloch, R., Anna-Arriola, S. S., Gao, D., Li, Y., Hodgkin, J. and Kimble, J. (1999). The gon-1 gene is required for gonadal morphogenesis in *Caenorhabditis elegans*. *Developmental Biology* **216**, 382-393.

Bolstad, B. M., Irizarry, R. A., Astrand, M. and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193.

Bonazzi, M., Spano, S., Turacchio, G., Cericola, C., Valente, C., Colanzi, A., Kweon, H. S., Hsu, V. W., Polishchuck, E. V., Polishchuck, R. S. et al. (2005). CtBP3/BARS drives membrane fission in dynamin-independent transport pathways. *Nature Cell Biology* **7**, 570-580.

Boyd, K. M., Subramanian, T., Schaeper, U., Laregina, M., Bayley, S. and Chinnadurai G. (1993). A region in the C-terminus of Adenovirus-phosphoprotein is required for assocation with a cellular phosphoprotein and important for the negative modulation of T24-Ras mediated transformtation, tumorigenesis and metastasis. *Embo Journal* **12**, 469-478.

Brenner, S. (1974).Genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.

Brown, A. L., Liao, Z. W. and Goodman, M. B. (2008). MEC-2 and MEC-6 in the Caenorhabditis elegans sensory mechanotransduction complex: Auxiliary Subunits that enable channel activity. *Journal of General Physiology* **131**, 605-616.

Cai, Y., Jin, J., Yao, T., Gottschalk, A. J., Swanson, S. K., Wu, S., Shi, Y., Washburn, M. P., Florens, L., Conaway, R. C., Conaway, J. W. (2007). YY1 functions with INO80 to activate transcription. *Nature Structural & Molecular Biology* **14**, 872 – 874.

Cairns, B. R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* **461**, 193-198.

Cameron, S., Clark, S. G., McDermott, J. B., Aamodt, E. and Horvitz, H. R. (2002). PAG-3, a Zn-finger transcription factor, determines neuroblast fate in C-elegans. *Development* **129**, 1763-1774.

Ceol, C. J. and Horvitz, H. R. (2004). A new class of C-elegans synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of ras signaling. *Developmental Cell* **6**, 563-576.

Chang, Y. F., Imam, J. S. and Wilkinson, M. E. (2007). The nonsense-mediated decay RNA surveillance pathway. *Annual Review of Biochemistry* **76**, 51-74.

Chen, L., Cai, Y., Jin, J. J., Florens, L., Swanson, S. K., Washburn, M. P., Conaway, J. W. and Conaway, R. C. (2011). Subunit Organization of the Human INO80 Chromatin Remodeling Complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodelling. *Journal of Biological Chemistry* **286**, 11283-11289.

Chen, S., Whetstine, J. R., Ghosh, S., Hanover, J. A., Gali, R. R., Grosu, P. and Shi, Y. (2009). The conserved NAD(H)-dependent corepressor CTBP-1 regulates Caenorhabditis elegans life span. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1496-1501.

Christensen, M., Estevez, A., Yin, X. Y., Fox, R., Morrison, R., McDonnell, M., Gleason, C., Miller, D. M. and Strange, K. (2002). A primary culture system for functional analysis of C-elegans neurons and muscle cells. *Neuron* **33**, 503-514.

Cinar, H. N., Richards, K. L., Oommen, K. S. and Newman, A. P. (2003). The EGL-13 SOX domain transcription factor affects the uterine pi cell lineages in Caenorhabditis elegans. *Genetics* **165**, 1623-1628.

Clark, S. G. and Chiu, C. (2003). C-elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. *Development* **130**, 3781-3794.

Clark, S. G., Lu, X. W. and Horvitz, H. R. (1994).The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine kinas signaling pathway, encodes 2 different proteins. *Genetics* **137**, 987-997.

Clay Spencer, D. M., III. (2011).Wormviz. http://jsp.weigelworld.org/wormviz/tileviz.jsp?experiment=wormviz&normalization=absolute&probesetcsv=&action=Run.

Conaway, R. C. and Conaway, J. W. (2009). The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends in Biochemical Sciences* **34**, 71-77.

Consortium, C. e. S. (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012-2018.

Cui, M., Kim, E. B. and Han, M. (2006a). Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in *C. elegans*. *Plos Genetics* **2**, 719-732.

Cui, M. X., Chen, J., Myers, T. R., Hwang, B. J., Sternberg, P. W., Greenwald, I. and Han, M. (2006b). SynMuv genes redundantly inhibit lin-3/EGF expression to prevent inappropriate vulval induction in *C. elegans*. *Developmental Cell* **10**, 667-672.

Davis, M. W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S. and Jorgensen, E. M. (2005). Rapid single nucleotide polymorphism mapping in *C. elegans*. *BMC Genomics* **6**, 118.

Del Bo, R., Ghezzi, S., Corti, S., Pandolfo, M., Ranieri, M., Santoro, D., Ghione, I., Prelle, A., Orsetti, V., Mancuso, M. et al. (2009). TARDBP (TDP-43) sequence analysis in patients with familial and sporadic ALS: identification of two novel mutations. *European Journal of Neurology* **16**, 727-732.

Dorman, J. B., Albinder, B., Shroyer, T. and Kenyon, C. (1995). The *age-1* and *daf-2* genes function in a common pathway to control the life span of *Caenorhabditis elegans*. *Genetics* **141**, 1399-1406.

Dohmen, R. J. (2004). SUMO protein modification. *Biochimica Et Biophysica Acta-Molecular Cell Research* **1695**, 113-131.

Efimenko, E., Bubb, K., Mak, H. Y., Holzman, T., Leroux, M. R., Ruvkun, G., Thomas, J. H. and Swoboda, P. (2005). Analysis of xbx genes in *C. elegans*. *Development* **132**, 1923-1934.

Esmaeili, B., Ross, J. M., Neades, C., Miller, D. M. and Ahringer, J. (2002). The *C. elegans* even-skipped homologue, vab-7, specifies DB motoneurone identity and axon trajectory. *Development* **129**, 853-862.

Feng, X. H., Sun, B. H., Liang, M., Brunicardi, F. C., Melchior, F. and Lin, X. (2003). Differential regulation of CtBP copressor function by SUMOylation. *Faseb Journal* **17**, A995-A995.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S.E., Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391,** 806–811.

Fisher, K., Southall, S. M., Wilson, J. R. and Poulin, G. B. (2010). Methylation and demethylation activities of a *C. elegans* MLL-like complex attenuate RAS signalling. *Developmental Biology* **341**, 142-153.

Fjeld, C. C., Birdsong, W. T. and Goodman, R. H. (2003). Differential binding of NAD(+) and NADH allows the transcriptional corepressor carboxyl-terminal binding protein to serve as a metabolic sensor. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9202-9207.

Flaus, A. and Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? *Current Opinion in Genetics & Development* **14**, 165-173.

Flibotte, T., Edgley, M. L., Chaudhry, I., Taylor, J., Neil, S. E., Rogula, A., Zapf, R., Hirst, M., Butterfield, Y., Jones, S. J., Marra, M., A., Barstead, R. J., Moerman, D. G. (2010). Whole-Genome Profiling of Mutagenesis in *Caenorhabditis elegans*. *Genetics* **185**, 431–441.

Fox, R. M., Von Stetina, S. E., Barlow, S. J., Shaffer, C., Olszewski, K. L., Moore, J. H., Dupuy, D., Vidal, M. and Miller, D. M. (2005). A gene expression fingerprint of C-elegans embryonic motor neurons. *BMC Genomics* **6**, 42.

Freemont, P. S. (2000). Ubiquitination: RING for destruction? *Current Biology* **10**, 84-87.

Furusawa, T., Moribe, H., Kondoh, H. & Higashi, Y. (1999) Identification of CtBP1and CtBP2 as corepressors of zinc finger-homeodomain factor delta EF1. *Molecular and Cellular Biology.* **12**, 8581–8585.

Garriga-Canut, M., Schoenike, B., Qazi, R., Bergendahl, K., Daley, T. J., Pfender, R. M., Morrison, J. F., Ockuly, J., Stafstrom, C., Sutula, T. et al. (2006). 2-deoxy-D-glucose reduces epilepsy progression by NRSF-CtBP-dependent metabolic regulation of chromatin structure. *Nature Neuroscience* **9**, 1382-1387.

Georg Zeller, S. R. H., Sascha Laubinger, Detlef Weigel, Gunnar Rätsch (2008). Transcript normalization and segmentation of tiling array data. *Pacific Symposium on Biocomputing* **2008;:527-538**.

Grooteclaes, M. L. and Frisch, S. M. (2000). Evidence for a function of CtBP in epithelial gene regulation and anoikis. *Oncogene* **19**, 3823-3828.

Grooteclaes, M., Deveraux, Q., Hildebrand, J., Zhang, Q. H., Goodman, R. H. and Frisch, S. M. (2003). C-terminal-binding protein corepresses epithelial and proapoptotic gene expression programs. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4568-4573.

Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews* **62**, 465-503.

Hansen, M., Hsu, A. L., Dillin, A. and Kenyon, C. (2005). New genes tied to endocrine, metabolic, and dietary regulation of lifespan from a *Caenorhabditis elegans* genomic RNAi screen. *Plos Genetics* **1**, 119-128.

Harris, G., Mills, H., Wragg, R., Hapiak, V., Castelletto, M., Korchnak, A. and Komuniecki, R. W. (2010). The Monoaminergic Modulation of Sensory-Mediated Aversive Responses in *Caenorhabditis elegans* Requires Glutamatergic/Peptidergic Cotransmission. *Journal of Neuroscience* **30**, 7889-7899.

Harris G P, K. A., Castelletto M, Mills H J, Komuniecki R W. (2010). Complex Food-Signal Modulates Aversive Behavior In *C. elegans*. In *Neuronal Development, Synaptic Function and Behavior*, Madison, WI.

Hildebrand, J. D. and Soriano, P. (2002). Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. In *Molecular and Cellular Biology* **22,** 5296-5307.

Ho, L. and Crabtree, G. R. (2010). Chromatin remodelling during development. *Nature* **463**, 474-484.

Hobert, O. (2002). PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C-elegans. *Biotechniques* **32**, 728-730.

Hobert, O., Johnston, R. J. and Chang, S. (2002). Left-right asymmetry in the nervous system: The *Caenorhabditis elegans* model. *Nature Reviews Neuroscience* **3**, 629-640.

Hobert, O. (2010). Neurogenesis in the nematode *Caenorhabditis elegans*. In *WormBook*, (ed. J. M. K. Eric M. Jorgensen): The *C. elegans* Research Community.

Hsieh, J., Liu, J., Kostas, S. A., Chang, C., Sternberg, P. W. and Fire, A. (1999). The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. *Genes & Development* **13**, 2958-2970.

Hsu, D. R., Chuang, P. T. and Meyer, B. J. (1995). DPY-30, a nuclear protein essential early in embryogenesis for *Caenorhabditis elegans* dosage compensation. *Development* **121**, 3323-3334.

Huang, X., Cheng, H. J., Tessier-Lavigne, M. and Jin, Y. S. (2002). MAX-1, a novel PH/MyTH4/FERM domain cytoplasmic protein implicated in netrin-mediated axon repulsion. *Neuron* **34**, 563-576.

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H. M., Pan, J., Schulz, A. K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel, R., Jones, S. J., Marra, M. A., Baillie, D. L., Moerman, D. G. (2007) High-Throughput *In Vivo* Analysis of Gene Expression in *Caenorhabditis elegans*. *PLOS Biology* **5**, 1981-1997.

Isono, K., Nemoto, K., Li, Y. Y., Takada, Y., Suzuki, R., Katsuki, M., Nakagawara, A. and Koseki, H. (2006). Overlapping roles for homeodomain-interacting protein kinases

Hipk1 and Hipk2 in the mediation of cell growth in response to morphogenetic and genotoxic signals. *Molecular and Cellular Biology* **26**, 2758-2771.

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. and Speed, T. P. (2003). Summaries of affymetrix GeneChip probe level data. *Nucleic Acids Research* **31**.

Jia, Y. W., Xie, G. F. and Aamodt, E. (1996). pag-3, a *Caenorhabditis elegans* gene involved in touch neuron gene expression and coordinated movement. *Genetics* **142**, 141-147.

Jia, Y. W., Xie, G. F., McDermott, J. B. and Aamodt, E. (1997). The *C. elegans* gene *pag-3* is homologous to the zinc finger proto-oncogene *gfi-1*. *Development* **124**, 2063-2073.

Jin, J. Y., Cai, Y., Yao, T., Gottschalk, A. J., Florens, L., Swanson, S. K., Gutierrez, J. L., Coleman, M. K., Workman, J. L., Mushegian, A. et al. (2005). A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *Journal of Biological Chemistry* **280**, 41207-41212.

Jin, Y., Jorgensen, E., Hartwieg, E., Horvitz, H. R. (1999). The *Caenorhabditis elegans* gene *unc-25* encodes glutamic acid decarboxylase and is required for synaptic transmission but not synaptic development. *The Journal of Neuroscience* **19**, 539–548.

Kagawa, H., Sugimoto, K., Matsumoto, H., Inoue, T., Imadzu, H., Takuwa, K. and Sakube, Y. (1995). Genome structure, mapping and expression of the tropomyosin gene *tmy-1* of *Caenorhabditis elegans*. *Journal of Molecular Biology* **251**, 603-613.

Kajimura, S., Seale, P., Tomaru, T., Erdjument-Bromage, H., Cooper, M. P., Ruas, J. L., Chin, S., Tempst, P., Lazar, M. A. and Spiegelman, B. M. (2008). Regulation of the brown and white fat gene programs through a PRDM16/CtBP transcriptional complex. *Genes & Development* **22**, 1397-1409.

Kimble, J., Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Developmental Biology* **70**, 396–417.

Koelle, M. (1994). Integrating extrachromosomal arrays into the *C. elegans* chromosomes.

Kratsios, P., Stolfi, A., Levine, M., Hobert, O. (2011). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nature Neuroscience* **15**, 205-214.

Kuppuswamy, M., Vijayalingam, S., Zhao, L. J., Zhou, Y., Subramanian, T., Ryerse, J. and Chinnadurai, G. (2008). Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. *Molecular and Cellular Biology* **28**, 269-281.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.

Lickteig, K. M., Duerr, J. S., Frisby, D. L., Hall, D. H., Rand, J. B. and Miller, D. M. (2001). Regulation of neurotransmitter vesicles by the homeodomain protein UNC-4 and its

transcriptional corepressor UNC-37/Groucho in *Caenorhabditis elegans* cholinergic motor neurons. *Journal of Neuroscience* **21**, 2001-2014.

Lecroisey, C., Martin, E., Mariol, M.-C., Granger, L., Schwab, Y., Labouesse, M., Segalat, L. and Gieseler, K. (2008). DYC-1, a protein functionally linked to dystrophin in *Caenorhabditis elegans* is associated with the dense body, where it interacts with the muscle LIM domain protein ZYX-1. *Molecular Biology of the Cell* **19**, 785-796.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.

Lewis, J. A., Wu, C. H., Berg, H. and Levine, J. H. (1980). The Genetics of levamisole resistance in the nematode *Caenorhabditis elegans*. *Genetics* **95**, 905-928.

Lin, K., Hsin, H., Libina, N. and Kenyon, C. (2001). Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nature Genetics* **28**, 139-145.

Liang, H. Z., Fekete, D. M. and Andrisani, O. M. (2011). CtBP2 Downregulation during Neural Crest Specification Induces Expression of Mitf and REST, Resulting in Melanocyte Differentiation and Sympathoadrenal Lineage Suppression. *Molecular and Cellular Biology* **31**, 955-970.

Lin, S. J. and Guarente, L. (2003). Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease. *Current Opinion in Cell Biology* **15**, 241-246.

Lin, X., Sun, B. H., Liang, M., Liang, Y. Y., Gast, A., Hildebrand, J., Brunicardi, F. C., Melchior, F. and Feng, X. H. (2003). Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding. *Molecular Cell* **11**, 1389-1396.

Livak, K. J., Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**, 402-408.

Melhuish, T. A. and Wotton, D. (2000). The interaction of the carboxyl terminus-binding protein with the Smad corepressor TGIF is disrupted by a holoprosencephaly mutation in TGIF. *Journal of Biological Chemistry* **275**, 39762-39766.

Michaux, G., Legouis, R. and Labouesse, M. (2001). Epithelial biology: lessons from *Caenorhabditis elegans*. *Gene* **277**, 83-100.

Miller, D. M., Niemeyer, C. J. and Chitkara, P. (1993). Dominant *unc-37* mutations suppress the movement defect of a homeodomain mutation in *unc-4*, a neural specificity gene in *Caenorhabditis elegans*. *Genetics* **135**, 741-753.

Miyoshi, T., Maruhashi, M., Van De Putte, T., Kondoh, H., Huylebroeck, D. and Higashi, Y. (2006). Complementary expression pattern of Zfhx1 genes Sip1 and delta EF1 in the mouse embryo and their genetic interaction revealed by compound mutants. *Developmental Dynamics* **235**, 1941-1952.

Nardini, M., Spanò, S., Cericola, C., Pesce, A., Massaro, A., Millo, E., Luini, A., Corda, D., Bolognesi, M. (2003). CtBP/BARS: a dual-function protein involved in transcription co-repression, and Golgi membrane fission. *The EMBO Journal* **22,** 3122-3130.

Nathoo, A. N., Moeller, R. A., Westlund, B. A. and Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14000-14005.

Nicholas, H. R., Lowry, J. A., Wu, T. and Crossley, M. (2008). The *Caenorhabditis elegans* protein CTBP-1 defines a new group of THAP domain-containing CtBP corepressors. *Journal of Molecular Biology* **375**, 1-11.

Nitta, K. R., Takahashi, S., Haramoto, Y., Fukuda, M., Tanegashima, K., Onuma, Y. and Asashima, M. (2007). The N-terminus zinc finger domain of Xenopus SIP1 is important for neural induction, but not for suppression of Xbra expression. In *International Journal of Developmental Biology*, **51**, 321-325.

Nuttley, W. M., Atkinson-Leadbeater, K. P. and van der Kooy, D. (2002). Serotonin mediates food-odor associative learning in the nematode *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12449-12454.

Petersen Gramstrup, R. J. P. (2011). EGL-13, a SOX transcription factor, regulates neuronal cell fate determination of the BAG neurons. In *International Worm Meeting*. Los Angeles, CA.

Pflugrad, A., Meir, J. Y. J., Barnes, T. M. and Miller, D. M. (1997). The Groucho-like transcription factor UNC-37 functions with the neural specificity gene *unc-4* to govern motor neuron identity in *C. elegans*. *Development* **124**, 1699-1709.

Prasad, B. C., Ye, B., Zackhary, R., Schrader, K., Seydoux, G. and Reed, R. R. (1998). unc-3, a gene required for axonal guidance in *Caenorhabditis elegans*, encodes a member of the O/E family of transcription factors. *Development* **125**, 1561-1568.

Pferdehirt, R. R., Kruesi, W. S. and Meyer, B. J. (2011). An MLL/COMPASS subunit functions in the *C. elegans* dosage compensation complex to target X chromosomes for transcriptional regulation of gene expression. *Genes & Development* **25**, 499-515.

Postigo, A. A. and Dean, D. C. (1999). ZEB represses transcription through interaction with the corepressor CtBP. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6683-6688.

Quinlan, K. G. R., Nardini, M., Verger, A., Francescato, P., Yaswen, P., Corda, D., Bolognesi, M. and Crossley, M. (2006). Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins. *Molecular and Cellular Biology* **26**, 8159-8172.

Riefler, G. M. and Firestein, B. L. (2001). Binding of neuronal nitric-oxide synthase (nNOS) to carboxyl-terminal-binding protein (CtBP) changes the localization of CtBP from the nucleus to the cytosol - A novel function for targeting by the PDZ domain of nNOS. *Journal of Biological Chemistry* **276**, 48262-48268.

Roussigne, M., Kossida, S., Lavigne, A. C., Clouaire, T., Ecochard, V., Glories, A., Amalric, F. and Girard, J. P. (2003). The THAP domain: a novel protein motif with similarity to the DNA-binding domain of P element transposase. *Trends in Biochemical Sciences* **28**, 66-69.

Sambrook, J., Fritsch, E.K., and Maniatis, T. (1989). Molecular Cloning. A laboratory manual: Cold Spring Harbour Laboratory Press.

Schaeper, U., Boyd, J. M., Verma, S., Uhlmann, E., Subramanian, T. and Chinadurai, G. (1995). Molecular cloning and characterisation of a cellular phosphoprotein that interacts with a conserved C-terminal domain of Adenovirus E1A involved in negative modulation of oncogenic transformation. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10467-10471.

Schmitz, F., Konigstorfer, A. and Sudhof, T. C. (2000). RIBEYE, a component of synaptic ribbons: A protein's journey through evolution provides insight into synaptic ribbon function. *Neuron* **28**, 857-872.

Sekido, R., Murai, K., Kamachi, Y. and Kondoh, H. (1997). Two mechanisms in the action of repressor delta EF1: binding site competition with an activator and active repression. *Genes to Cells* **2**, 771-783.

Sharma-Kishore, R., White, J. G., Southgate, E. and Podbilewicz, B. (1999). Formation of the vulva in *Caenorhabditis elegans*: a paradigm for organogenesis. *Development* **126**, 691-699.

Shi, Y. J., Sawada, J., Sui, G. C., Affar, E. B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P. S., Nakatani, Y. and Shi, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* **422**, 735-738.

Simonis, N., Rual, J. F., Carvunis, A. R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Sahalie, J. M., Venkatesan, K., Gebreab, F. et al. (2009). Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. *Nature Methods* **6**, 47-54.

Smith, P., Leung-Chiu, W. M., Montgomery, R., Orsborn, A., Kuznicki, K., Gressman-Coberly, E., Mutapcic, L. and Bennett, K. (2002). The GLH proteins, *Caenorhabditis elegans* P granule components, associate with CSN-5 and KGB-1, proteins necessary for fertility, and with ZYX-1, a predicted cytoskeletal protein. *Developmental Biology* **251**, 333-347.

Spencer, W. C., Zeller, G., Watson, J. D., Henz, S. R., Watkins, K. L., McWhirter, R. D., Petersen, S., Sreedharan, V. T., Widmer, C., Jo, J. et al. (2011). A spatial and temporal map of *C. elegans* gene expression. *Genome Research* **21**, 325-341.

Starich, T. A., Miller, A., Nguyen, R. L., Hall, D. H. and Shaw, J. E. (2003). The *Caenorhabditis elegans* innexin INX-3 is localized to gap junctions and is essential for embryonic development. *Developmental Biology* **256**, 403-417.

Sulston, J. E., Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Developmental Biology* **56**, 110–156.

Sulston, J. E., Schierenberg, E., White, J. G., Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology* **100**, 64–119.

Tao Zhang, H.-Y. H., Haiping Hao, Conover Talbot Jr., Jiou Wang. (2012). *Caenorhabditis elegans* RNA-processing Protein TDP-1 Regulates Protein Homeostasis and Lifespan. *The Journal of Biological Chemistry*. **287**, 8371-8382.

Todd Starich, J. S. (2001). Gap junction protein INX-13 is essential for excretory cell function. In *International C. elegans Meeting*, Los Angeles, CA.

Thomas, J. H. and Horvitz, H. R. (1999). The *C. elegans* gene lin-36 acts cell autonomously in the lin-35 Rb pathway. *Development* **126**, 3449-3459.

Thio, S. S. C., Bonventre, J. V. and Hsu, S. I. H. (2004). The CtBP2 co-repressor is regulated by NADH-dependent dimerization and possesses a novel N-terminal repression domain. *Nucleic Acids Research* **32**, 1836-1847.

Updike, D. L. and Mango, S. E. (2007). Genetic suppressors of *Caenorhabditis elegans* pha-4/FoxA identify the predicted AAA helicase ruub-1/RuvB. *Genetics* **177**, 819-833.

van Grunsven, L. A., Taelman, V., Michiels, C., Verstappen, G., Souopgui, J., Nichane, M., Moens, E., Opdecamp, K., Vanhomwegen, J., Kricha, S. et al. (2007). XSip1 neuralizing activity involves the co-repressor CtBP and occurs through BMP dependent and independent mechanisms. In *Developmental Biology*, **306**, 34-49.

Van de Putte, T., Maruhashi, M., Francis, A., Nelles, L., Kondoh, H., Huylebroeck, D. and Higashi, Y. (2003). Mice lacking Zfhx1b, the gene that codes for Smad- interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. *American Journal of Human Genetics* **72**, 465-470.

Varki A, C. R., Esko JD. (2009). Essentials of Glycobiology. NY: Cold Spring Harbor Laboratory Press.

Vogel, B. E. and Hedgecock, E. M. (2001). Hemicentin, a conserved extracellular member of the immunoglobulin superfamily, organizes epithelial and other cell attachments into oriented line-shaped junctions. *Development* **128**, 883-894.

Von Stetina, S. E., Treinin, M. and Miller, D. M. (2005). The motor circuit. *Neurobiology of C. elegans* **69**, 125.

Von Stetina, S. E., Watson, J. D., Fox, R. M., Olszewski, K. L., Spencer, W. C., Roy, P. J. and Miller, D. M. (2007). Cell-specific microarray profiling experiments reveal a comprehensive picture of gene expression in the *C. elegans* nervous system. *Genome Biology* **8**, 135.

Wacker, I., Schwarz, V., Hedgecock, E. M. and Hutter, H. (2003a). zag-1, a Zn-finger homeodomain transcription factor controlling neuronal differentiation and axon outgrowth in *C. elegans*. *Development* **130**, 3795-3805.

Wacker, I., Schwarz, V., Hedgecock, E. M. and Hutter, H. (2003b). Zag-1, a zinc-finger-homeobox transcription factor controlling neuronal differentiation in *C. elegans*. In *European Journal of Cell Biology*, **82**, 134-134.

Wakabayashi, T., Osada, T. and Shingai, R. (2005). Serotonin deficiency shortens the duration of forward movement in *Caenorhabditis elegans*. *Bioscience Biotechnology and Biochemistry* **69**, 1767-1770.

Weigert, R., Silletta, M. G., Spano, S., Turacchio, G., Cericola, C., Colanzi, A., Senatore, S., Mancini, R., Polishchuk, E. V., Salmona, M. et al. (1999). CtBP/BARS induces fission of Golgi membranes by acylating lysophosphatidic acid. *Nature* **402**, 429-433.

White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1976). Structure of the Ventral Nerve Cord of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **275**, 327.

White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1986). Structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **314**, 1-340.

White, J., Southgate, E. and Durbin, R. (1988). The nematode *Caenorhabditis elegans*. Neuroanatomy. *Cold Spring Harbor Monograph Series* **17**, 433-455.

Winnier, A. R., Meir, J. Y. J., Ross, J. M., Tavernarakis, N., Driscoll, M., Ishihara, T., Katsura, I. and Miller, D. M. (1999). UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in *Caenorhabditis elegans*. *Genes & Development* **13**, 2774-2786.

Wu, J. S. and Grunstein, M. (2000). 25 years after the nucleosome model: chromatin modifications. *Trends in Biochemical Sciences* **25**, 619-623.

Wu, T. T. L. (2006). Elucidating the Role of C-Terminal Binding Protein (CtBP) in *Caenorhabditis elegans*. In *School of Molecular Bioscience*, Bachelor of Science (Honours), Sydney: University of Sydney.

Zhang, Q. H., Piston, D. W. and Goodman, R. H. (2002). Regulation of corepressor function by nuclear NADH. *Science* **295**, 1895-1897.

Zhang, O. H., Yoshimatsu, Y., Hildebrand, J., Frisch, S. M. and Goodman, R. H. (2003). Homeodomain interacting protein kinase 2 promotes apoptosis by downregulating the transcriptional corepressor CtBP. *Cell* **115**, 177-186.

Zhang, Q. H., Nottke, A. and Goodman, R. H. (2005). Homeodomain-interacting protein kinase-2 mediates CtBP phosphorylation and degradation in UV-triggered apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2802-2807.

APPENDIX

Appendix Table 1: Mutants of *ctbp-1* and *zag-1* show similar defects in their DA cells. 102 wildtype, 132 *ctbp-1(ok498)* mutant and 34 *zag-1(rh315)* mutant animals expressing UNC-4::GFP were examined at the L1 stage. The number of UNC-4::GFP marked DA cells per animal is shown as the percentage of the population size.

GFP marker	Neurons Analysed	Number of Cells Expressing GFP	Phenotype in wild type background (n=105)	Phenotype in *ctbp-1 (ok498)* background (n=132)	Phenotype in *zag-1 (rh315)* background (n=34)
unc-4::gfp	DA2-7	6-5	53.3 %	21.2 %	14.7 %
		4	28.6 %	31.1 %	5.9 %
		3	12.4 %	20.5 %	11.8 %
		2	2.9 %	9.8 %	23.5 %
		1	2.9 %	12.1 %	20.6 %
		0	0 %	5.3 %	23.5 %

Appendix Table 2: Neuron-specific CTBP-1 targets that meet the statistical criteria of fold change ≥ 2 and p-value ≤ 0.05 are listed. Genes that do not have a formal gene name are indicated with a cosmid name.

Cosmid or Gene Name	Fold Change	P-Value
T23E7.2	4.41	0.002
ddl-3	3.59	0.038
B0207.10	2.77	0.009
C26C6.6	2.77	0.001
ddo-2	2.74	0.022
ttr-3	2.71	0.012
nspd-1	2.62	0.013
dylt-2	2.59	0.012
W09C3.1	2.49	0.047
Y60A3A.23	2.48	0.021
F53A9.2	2.44	0.044
srd-16	2.43	0.001
nlp-19	2.43	0.007
Y73C8B.5	2.40	0.007
F46C8.3	2.38	0.015
F59G1.8	2.29	0.023
F55C9.11	2.29	0.043

R09H10.3	2.28	0.022
C04G6.10	2.27	0.033
cnc-10	2.19	0.025
F22E5.6	2.17	0.040
F46C8.1	2.17	0.016
EEED8.4	2.13	0.016
C47B2.9	2.11	0.009
fkb-2	2.08	0.030
F02D8.5	2.08	0.003
sdz-19	2.07	0.019
F15G9.2	2.07	0.014
T13C5.4	2.07	0.015
sdz-5	2.06	0.019
fipr-3	2.05	0.048
arrd-3	2.04	0.022
glb-15	2.03	0.003
ZC443.1	2.03	0.015
ZK512.8	2.03	0.002
Y57G11B.97	2.02	0.009
F20B6.4	2.02	0.022
C04E6.13	2.02	0.015
E02H9.3	2.01	0.032
ins-37	2.01	0.004
F40G12.7	2.00	0.033

Appendix Table 3: CTBP-1 targets identified through analysis of all cells data set which meet the statistical criteria of fold change ≥ 2 and p-value ≤ 0.05 are listed. Genes that do not have a formal gene name are indicated with a cosmid name.

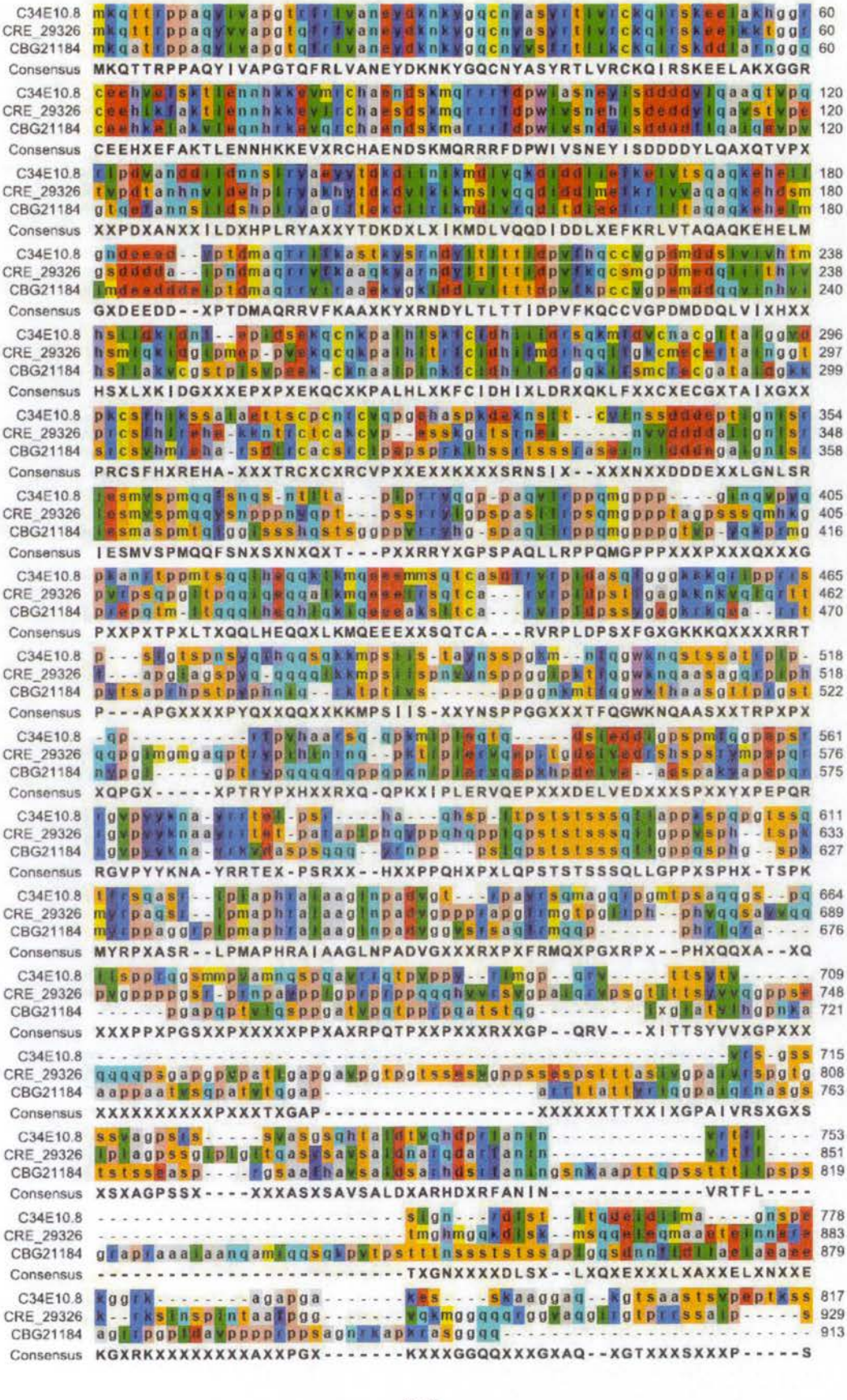
Cosmid or Gene Name	Fold Change	P-Value
F57C9.3	3.07	0.003
zyx-1	2.97	0.006
W04D12.1	2.92	0.130
Y54G2A.41	2.85	0.007
C35D6.3	2.72	0.061
F59B8.1	2.66	0.004
F36G3.3	2.59	0.036

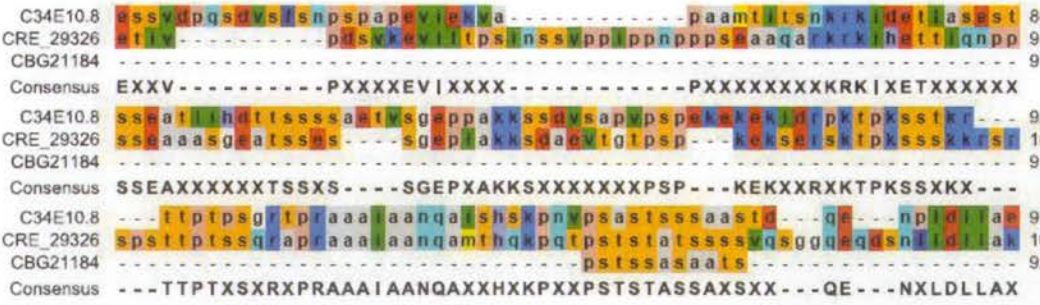
Y26D4A.5	2.51	0.006
acbp-3	2.46	0.149
F53G12.13	2.45	0.048
ttr-3	2.44	0.021
Y104H12D.2	2.39	0.026
T28A11.3	2.36	0.055
F21C10.5	2.33	0.074
F41D9.2	2.33	0.007
nhr-144	2.33	0.034
F22B3.5	2.31	0.149
F26F2.5	2.30	0.084
him-4	2.30	0.010
C04F6.2	2.27	0.081
Y97E10AR.8	2.25	0.069
Y47H9C.12	2.25	0.034
ttr-5	2.22	0.102
inx-13	2.22	0.034
Y6G8.5	2.21	0.071
C33D9.10	2.21	0.096
B0273.1	2.20	0.001
R07C12.4	2.19	0.033
C03B8.3	2.18	0.069
T01B7.9	2.17	0.013
F48F5.3	2.17	0.262
T23G11.11	2.15	0.037
Y37A1B.7	2.14	0.012
K10B3.1	2.13	0.168
T06D10.3	2.12	0.074
Y48G1BL.5	2.12	0.006
grl-24	2.11	0.069
C04H5.1	2.11	0.005
C37A5.11	2.11	0.034
Y41D4A.7	2.10	0.024
H42K12.3	2.10	0.174
tdp-1	2.07	0.003
gpc-1	2.07	0.080
M110.8	2.07	0.029
F49E8.2	2.06	0.073
fkb-8	2.06	0.005
C09B9.8	2.06	0.062
lev-11	2.05	0.010
F41B4.2	2.04	0.011
T20F10.5	2.03	0.025

glrx-10	2.03	0.083
Y37E11AL.4	2.02	0.028
2RSSE.2	2.02	0.016
mec-2	2.02	0.012
T01B4.3	2.01	0.008
nspe-3	2.01	0.057
F41G3.19	2.01	0.010
cat-4	2.00	0.022

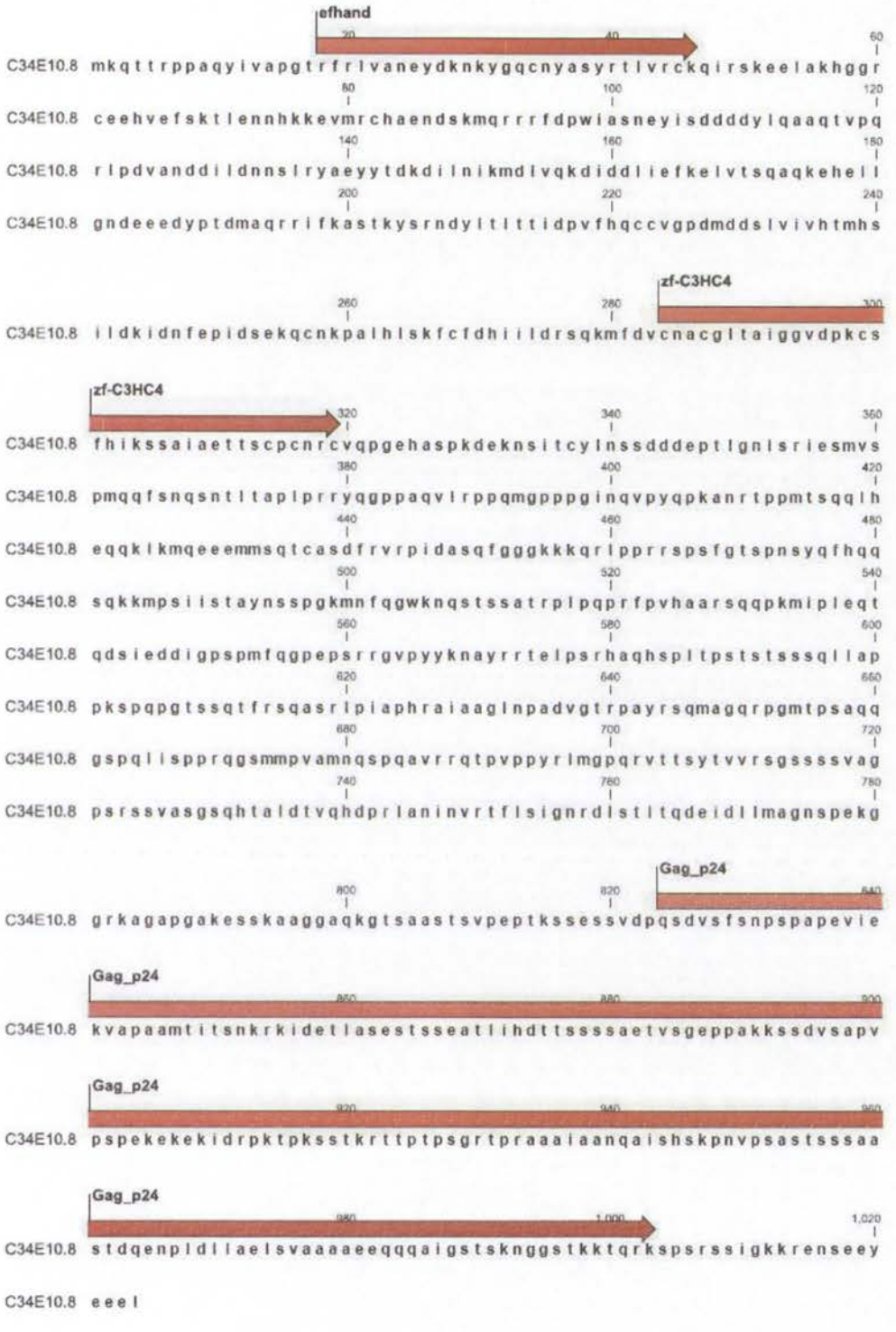
Appendix Table 4: The variants identified through next generation sequencing are listed. The type of variants and number of each type variant across the genome of *aus3* mutant animal and the lab reference strain are listed.

Class	Variant # in laboratory reference	Variant # in aus3 mutant
Five prime UTR	74	109
Frameshift	31	20
Inframe	2	7
Missense	449	882
ncRNA	47	62
Nongenic	7758	10304
Non-start	2	2
Premature Stop	10	15
Readthrough	4	7
Silent	251	428
SNP	538	573
Splice acceptor	2	5
Splice donor	3	10

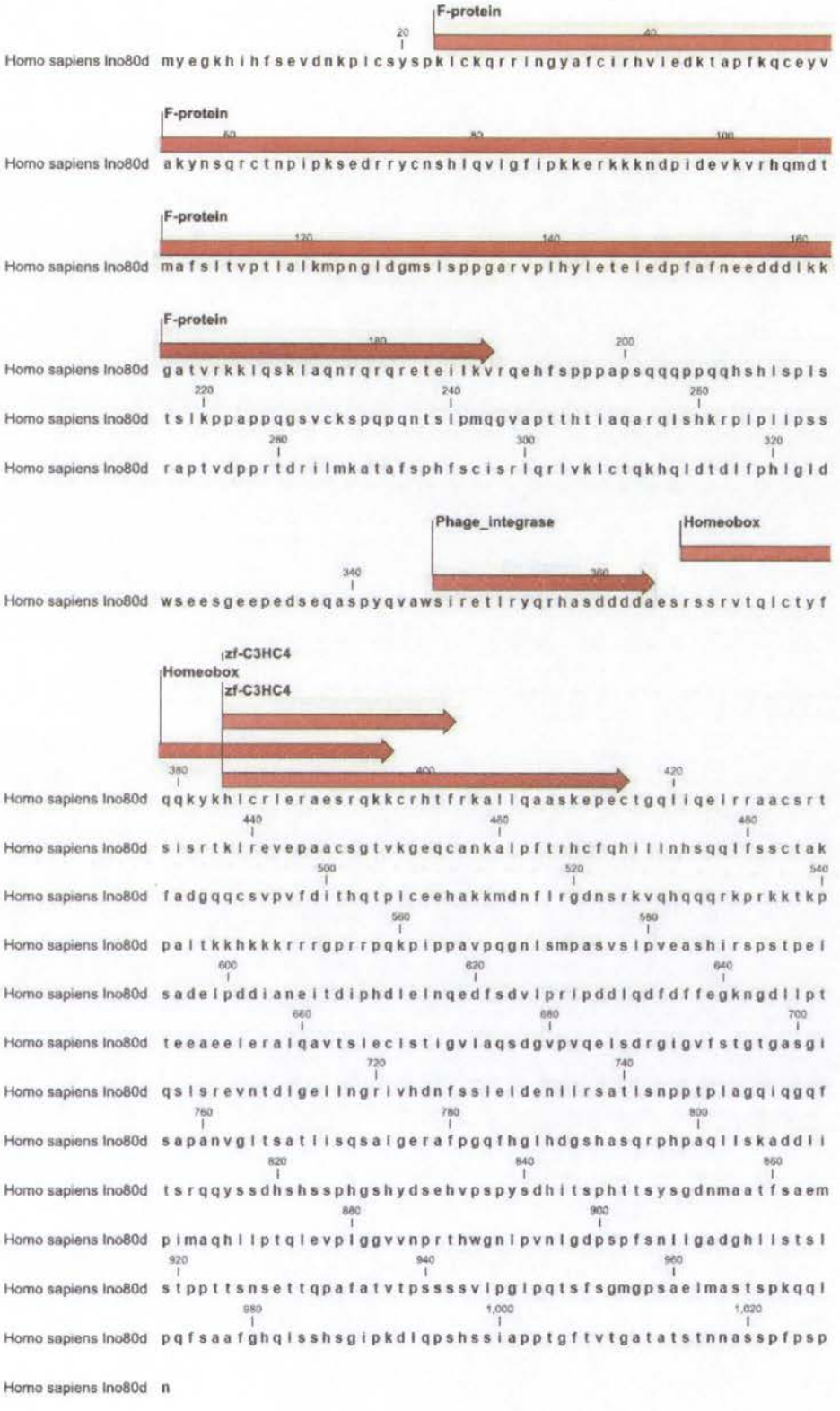

C34E10.8
CRE_29326
CBG21184
Consensus



Appendix Figure 1: Alignment of orthologs of C34E10.8 in *Caenorhabditis briggsae* (CBG_21184) and in *Caenorhabditis remanei* (CRE_29326). C34E10.8 is largely conserved at the N-terminus and unstructured at the C-terminus as indicated with the alignment of it orthologs in other nematodes.



Appendix Figure 2: The predicted domain structure of C34E10.8. The predicted domains were obtained through the software Workbench. According to this prediction, C34E10.8 has a putative EF-hand domain, a RING finger domain (indicated as zf-C3HC4) and a viral Gag_p24 domain.



Appendix Figure 3: The predicted domain structure of human INO80d. The predicted domains were obtained through the software Workbench. According to this prediction, human INO80d has a putative F-protein

domain (F-box), RING finger domains (indicated as zf-C3HC4), a homeobox and finally a phage integrase domain