UNIVERSITY OF CALGARY

gob-1 is a Trehalose-6-Phosphate Phosphatase Required for Intestine Development in the Nematode *Caenorhabditis elegans*

by

Jay Dene Kormish

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

CALGARY, ALBERTA

DECEMBER 20, 2004

Abstract

A genetic screen was developed with the intention of finding *C. elegans* genes that are required for intestine development or differentiation. The screen was based on the gut obstructed phenotype of the *elt-2* null mutant. *elt-2* is a GATA transcription factor believed to be the major transcription factor driving intestine differentiation. It was anticipated that screening for a similar gut obstructed mutant phenotype could identify genes that were functioning redundantly to *elt-2* or downstream of *elt-2*. One of the mutations identified in this screen, *gob-1*(*ca17*), was mapped to the right-hand side of the X chromosome. Positional cloning of the *gob-1*(*ca17*) allele showed that it was a deletion of approximately 15 open reading frames. RNA-mediated interference to a gene in this region, H13N06.3, phenocopied the early larval arrest and gut obstructed phenotype of *gob-1*(*ca17*) and was thus identified as the gene responsible for this phenotype. The *gob-1* RNAi mutant causes an early defect in the intestine. The *gob-1* gene and GOB-1 protein are expressed specifically in the intestine during its early development. GOB-1 is a member of the Haloacid dehalogenase (HAD)-like hydrolase superfamily and is the first identified trehalose-6-phosphate phosphatase in the nematode phylum. GOB-1's function in trehalose metabolism is currently unclear but mounting evidence suggests that GOB-1 may have a function required for viability of the worm independent of trehalose synthesis. ELT-2 is sufficient but not necessary for early embryonic expression of GOB-1. The ability of ELT-2 to drive *gob-1* expression suggests that loss of *gob-1*

function may be a component of the later *elt-2* loss of function defect in the larval intestine.

Acknowledgements

I would like to thank the following people. First and foremost, I would like to thank my thesis supervisor Dr. James McGhee. Thank you for guiding me in the right direction. I would like to also thank my graduate committee members Dr. Paul Mains, Dr. William Brook and Dr. Sarah McFarlane for their advice during my degree and Dr. Dave Hansen and Dr. Joseph Culotti for agreeing to be a part of my defense committee. I would like to thank Mr. W. Dong for assistance with the EM and Dr. G. Moorhead for assistance with the phosphorylase a phosphatase assay. I would like to thank past and present members of the McGhee lab for critical discussions and technical assistance.

I would like to thank my family for their love and support. I would especially like to thank my husband Scott MacNamara for his love, support and patience and for his help with the editing of this thesis.

To My Husband

Scott MacNamara

and

My Parents

Walter and Lavina Kormish

Table of Contents

Page

CHAPTER TWO - Materials and Methods

List of Tables

List of Figures

List of Abbreviations

AA – amino acid(s)

DNA – deoxyribonucleic acid

EM – electron microscopy

DAPI – 4,6-diamindino-2-phenyl-indole

GFP – green fluorescent protein

Gob – gut obstructed

GST – glutathione-S-transferase

HAD – haloacid dehalogenase

MBD – maltose binding protein

PAGE – polyacrylamide gel electrophoresis

PFD – post first division

pNPP – para-nitrophenyl phosphate

RNA – ribonucleic acid

RNAi – double stranded-mediated interference

RT-PCR – reverse transcriptase PCR

PCR – polymerase chain reaction

SDS – sodium dodecyl sulfate

SNP – single nucleotide polymorphism

T6P – trehalose-6-phosphate

TPP – trehalose phosphate phosphatase

TPS – trehalose phosphate synthase

CHAPTER ONE – Introduction

Introduction

The intestine is vital for life. The ability to absorb nutrients from the environment is necessary for the survival of all organisms and is likely one of the most ancient functions that organisms possess. Despite the importance of this organ system, we only know a small fraction of what is needed to truly understand it. How does the digestive tract develop? What transcription factors or signaling pathways are driving its formation in the early embryo? What genes are expressed in the tissues of these organs to make them functional? Although we are beginning to understand some of these processes, much of the truth is yet to be discovered.

In order to gain some insight into the development of the digestive tract, many researchers have opted to use model organisms. The ease of manipulation of these organisms can allow for a deeper understanding of certain processes. As these organisms are studied, common themes of development become obvious and we can potentially gain a better understanding of these processes in ourselves. In this thesis, a study of the *C. elegans* intestine will be presented. The ease of studying this organ in this particular organism makes it ideal for this sort of study. Background leading up to this study and a genetic screen devised to find new genes involved in the development of the intestine will

be presented. The identity and potential function of one of the genes identified will be described.

General Description of *Caenorhabditis elegans*

Caenorhabditis elegans is a free-living soil round worm belonging to the phylum Nematoda (Wood, 1988). This organism was published as a model system by Sydney Brenner in 1974 and has since been well characterized with regards to development, genetics and cell biology (Riddle et al. 1997). *C. elegans* is ideal for the study of organ development. The adult is approximately one millimeter in length and is easily viewed using a dissecting microscope. The animals are transparent at all developmental stages allowing for easy viewing of its major organ systems. The generation period of *C. elegans* is short, about three days from the time an egg is laid to the reaching of reproductive maturity. This allows for developmental experiments to be completed in a relatively short period of time when compared to other model systems.

Embryogenesis occurs within a chitinized egg shell. Embryos are staged by two criteria – the time in minutes after the first cell division (termed post-first cell division/PFD) at 22˚C and the morphological stage reached during development. Embryonic time points that will be commonly referred to in this thesis will be the 8E cell stage (~170 min PFD), the bean stage (~280 min PFD), the comma stage (~400 min PFD), the 1½-fold stage (~420 min PFD), the 2-fold stage (~450 min PFD) and the pretzel stage (~520 min PFD)

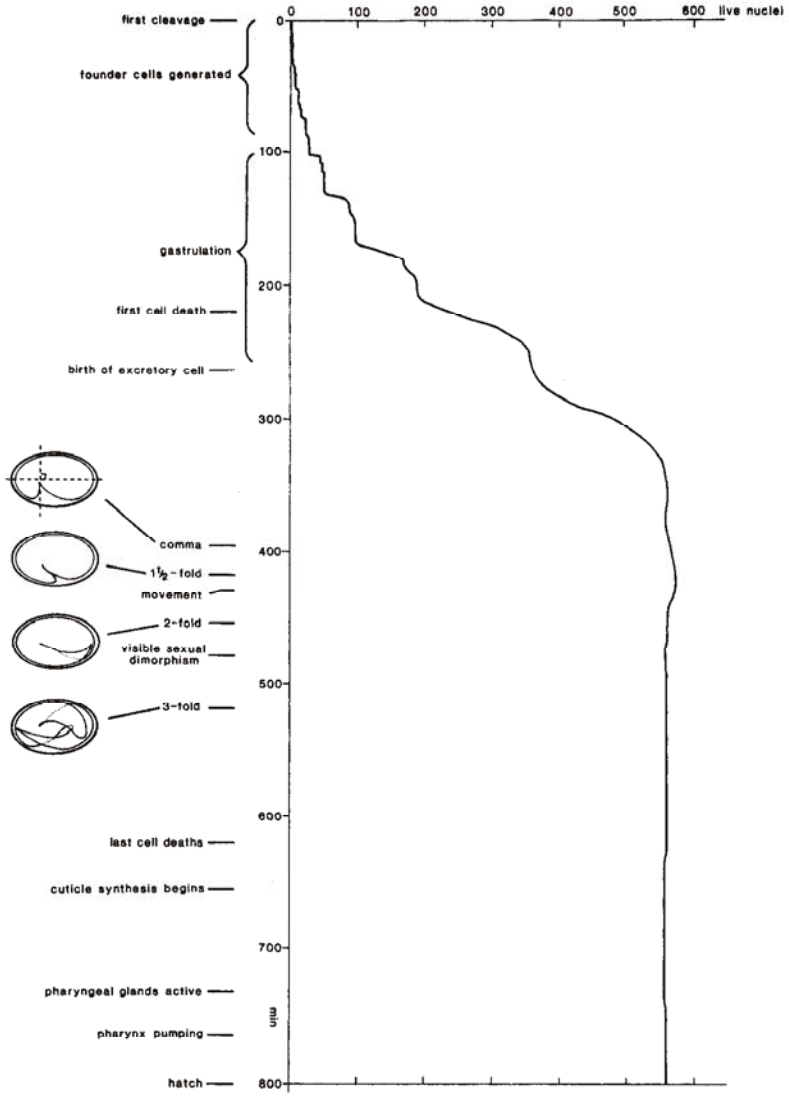


Figure 1: *C. elegans* embryo stages. The time in minutes after the first cell division at 22˚C vs. the number of nuclei present in the embryo at that time. Key morphological stages are diagrammed. Taken from Sulston et al. (1983).

(Sulston et al., 1983) **(see Figure 1)**. Upon hatching, a nearly fully developed larva grows in size and undergoes a series of four molts. The stage before each of these four molts has been termed L1, L2, L3 and L4 respectively. Characteristic cell divisions, programmed cell deaths, neuronal migrations and gonad maturation steps define each larval stage.

C. elegans has six chromosomes (five autosomes and one sex chromosome). These chromosomes are estimated to contain 97 x 10^9 base pairs of DNA and 19,099 predicted genes (Consortium, 1998). There are two sexes within the species: the hermaphrodite, which contains two X chromosomes, and the male, which contains only one X chromosome. The hermaphrodite is self-fertilizing and has the ability to produce oocytes and sperm at separate times. The male only produces sperm and this sperm is used to preferentially fertilize a hermaphrodite's oocytes upon copulation.

After fertilization, the *C. elegans* embryo undergoes a series of cell divisions and cell movements that are largely invariant from animal to animal (Sulston et al., 1983). The lineage of these cell divisions and cell specifications have been well characterized and allow for the origins of all 959 somatic nuclei within the adult hermaphrodite (1031 in the male) to be precisely traced to their origin in the early embryo.

The adult worm has five major organ systems. Surrounding the worm is the cuticle-secreting hypodermis. Centrally located is the digestive tract including the pharynx, the intestine proper and the hindgut. Four muscle quadrants juxtaposed to the hypodermis run along the length of the worm. The animal has

a nervous system consisting of a ventral nerve tract and a ring-like nerve network at the base of the pharynx. The reproductive organ in the hermaphrodite is a bilobed organ. The two lobes of the gonad meet at the vulva opening, through which copulation and egg-laying occur **(see Figure 2 and 3)**.

The Intestine

The digestive tract of the worm is fully developed and functioning by the time the animal hatches. The adult intestine consists of 20 cells, which form a series of nine rings termed Int 1 through 9. Each ring consists of two cells with the exception of the most anterior ring, Int 1, which is made up of four cells **(see Figure 4)**. After embryogenesis, the intestinal cells undergo rounds of endoreduplication leading to approximately 32 copies of the haploid genome in each nucleus (Hedgecock and White, 1985). The 20 cells of the adult gut are clonally derived from a single cell, the E cell that is formed in the eight cell embryo. The series of cell divisions, cell movements and polarization events that occur to form the functioning intestine have been well described (Leung et al., 1999; Sulston et al., 1983) **(see Figure 5)**. After the first division of the E cell, gastrulation occurs and the intestine precursors move into the interior of the embryo. The second division is along the left/right axis and is the only division in this axis. The E cells then undergo a series of skewed anterior/posterior divisions to form two rows of five dorsal and three ventral cells. It is at this 16 E cell stage that signs of epithelial polarization occur. At this point, the midline of

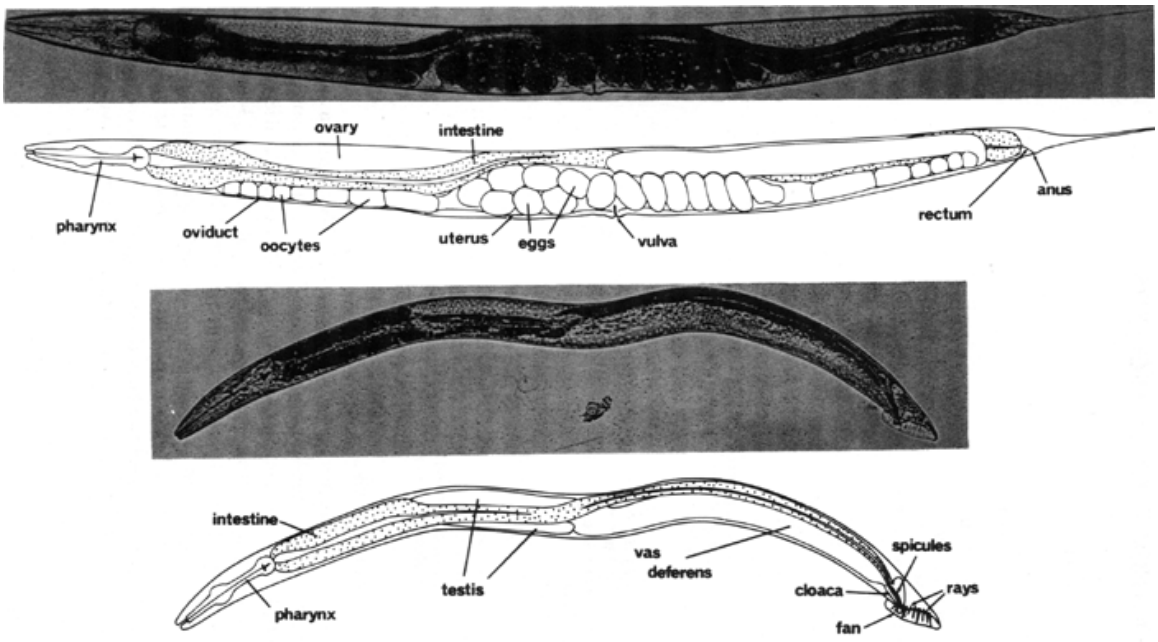


Figure 2: Anatomy of *C. elegans*. The hermaphrodite (top) and the male (bottom) anatomy are diagrammed. Figure taken from Wood (1988).

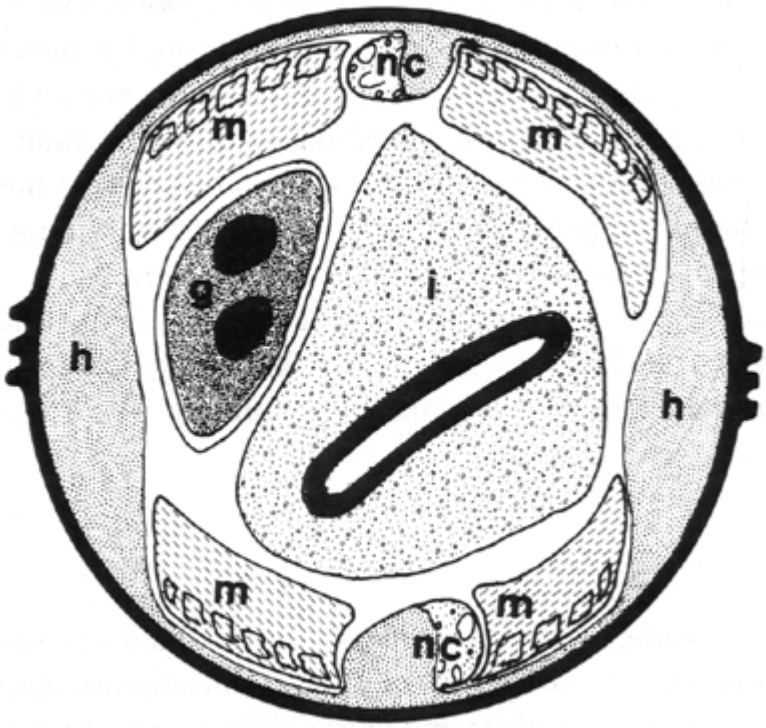


Figure 3: Cross-section of the adult worm. The five major organ systems are diagrammed: the hypodermis (h), the intestine (i), the body wall muscle (m), the nerve cord (nc) and the gonad (g). Figure is taken from Wood (1988).

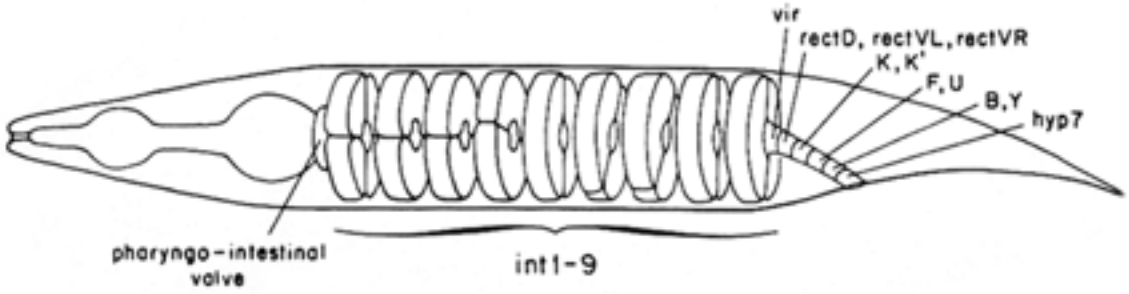
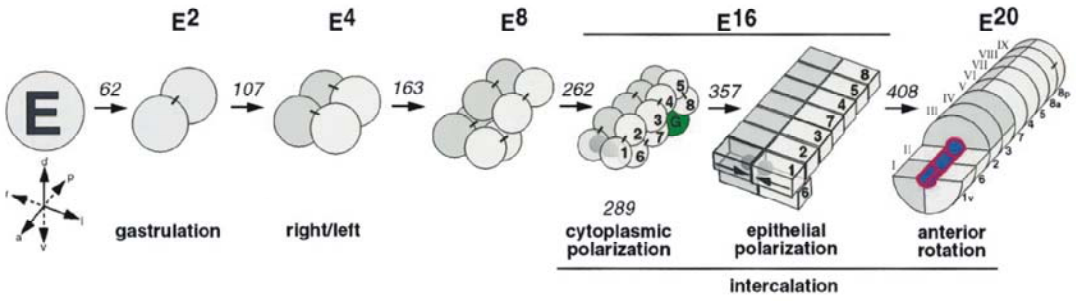


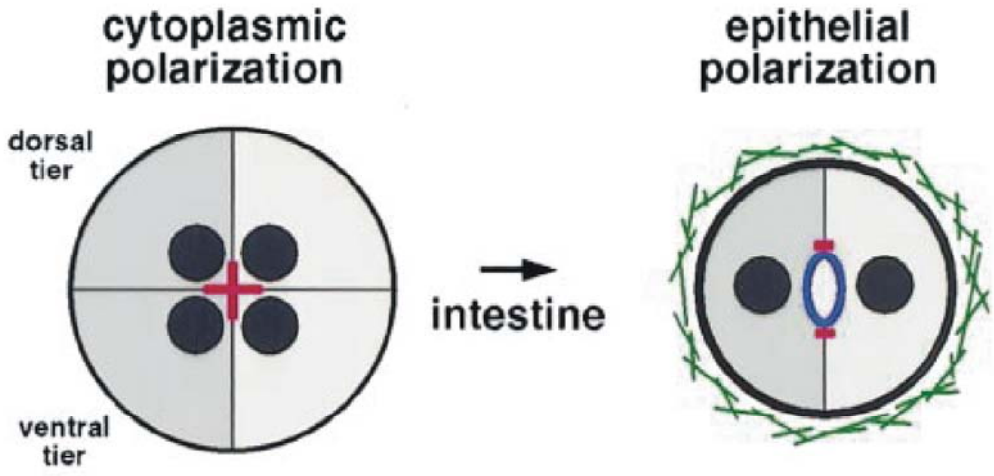
Figure 4: Digestive tract of the worm. The nine Int rings of the intestine are diagrammed. Int 1 is composed of four cells. The remaining eight Int rings are composed of two cells. Figure is taken from Schedin et al. (1991).

Figure 5: Cell divisions, cell rearrangements and epithelial polarization events that are characteristic of the *C. elegans* intestine. (A) The intestine originates from the single E cell. The E cell undergoes a series of stereotyped divisions and cell movements. At the 16 E cell stage, cellular polarization begins. (B) A cross-section of the polarizing intestine. The midline (left diagram) is defined by the points of cell contact (highlighted in red). As polarization proceeds, the intestine matures into an epithelium (right diagram). The formation of adherens junctions (red), lumenal surface (blue) and a basement membrane (green) are required for proper epithelial function. Figure taken from Leung et al. (1999).

A


E
E2
E4
E8
E16
E20
62
107
163
262
357
408
289
gastrulation
right/left
cytoplasmic polarization
epithelial polarization
anterior rotation
intercalation

B


cytoplasmic polarization
epithelial polarization
dorsal tier
ventral tier
intestine

the gut becomes obvious and is defined by the point of contact between the two rows of left and right cells. The onset of gut cell polarization is marked by relocation of the nuclear-associated centrosomes to the midline and a concentration of microtubules at the midline cortex. This event is followed by the movement of the nuclei within each row of cells to the midline. With the movement of the nuclei to the midline, there is a concomitant movement of cytoplasmic components away from the midline. Subsequent to this cytoplasmic polarization, the cells begin to take on an epithelium-like structure and elongate along their apical/basal axis. The apical surface forms at the point of midline cell contacts and is the future site of the intestine lumen. During epithelialization, there is a directional movement of cell components (e.g., PKC-3 (Atypical Protein Kinase C)) to the apical surface. As polarization proceeds, the lumen starts to form as the cell surfaces at the midline start transiently to lose adhesiveness. The exact nature of this loss of adhesion at this particular site is unknown but it is associated with the deposition of small cytoplasmic bodies called “apical vesicles” at the apical surface. Separation begins before the formation of adherens junction-like structures even though there is apical location of known components (AJM-1 and HMR-1) (Segbert et al., 2004). As development proceeds, the cells intercalate to form a tube-like structure. Interestingly, the cells intercalate in a way that preserves the points of lumen separation that have begun to form.

The ability to polarize appears to be a property intrinsic to the E cell lineage. Cell/cell contacts are necessary and cells can somehow distinguish

between left/right pairs of cells and cells within the same row. Contacts from outside the gut also appear to be important to bias the left/right axis and allow for bilateral symmetry of the organ. Later in embryogenesis, the first and last pair of cells undergo a second division to form the full complement of 20 cells.

Intestinal Epithelium

The functional intestine is a polarized epithelial tissue. There are three cellular events that the intestinal cells undergo as they mature into a functional epithelium. The first event is one of polarization. As stated above, polarization requires left/right cell contacts. The PKC-3 protein localizes to the apical membrane early during the polarization of the intestinal lineage (Leung et al., 1999). PKC-3 was originally defined by its requirement for the establishment of anterior/posterior polarity in the one cell embryo (Macara, 2004; Ohno, 2001; Wodarz, 2002). In the one cell embryo, PKC-3 is a component of the PAR-3/PAR-6/PKC-3 complex that is localized to the anterior pole of the cell. Each component of this ternary complex is required for the proper localization of the other components. Furthermore, the posterior localization of other proteins such as PAR-1 and PAR-2 are dependent on the correct localization of this complex. The early lethality caused by the knockdowns of *par-3*, *par-6*, *pkc-3* and *par-1* has hampered a detailed study of these proteins during intestine polarization. Homologues of these proteins in other organisms have been found to be a contributing factor in establishing the apical/basal polarity of epithelial cells **(see**

Figure 6). During Drosophila neuroectoderm development, proteins with sequence similarity to PAR-3 (Bazooka), PAR-6 (DmPAR-6) and PKC-3 (DaPKC) localize to the apical surface of the epithelium. Similarly, in cultured mammalian epithelial cells PAR-6 (PAR-6), aPKC (PKC-3) and ASIP (atypical protein kinase C-specific interacting protein/PAR-3) are localized to the apical surface of the epithelium. In both the neuroectoderm in flies and the epithelium in mammals, the components of the PAR-3/PAR-6/PKC-3-like complex depend on each other for the complex's stable localization. At least two additional features of mammalian epithelium development are similar to the polarization of the *C. elegans* one cell embryos. A PAR-1 (EMK/MARK/mPAR-1)-like protein is localized to the lateral membranes. A small GTPase, Cdc42, is activated and localized to the apical surface by cell/cell contacts. Cdc42 in turn recruits aPKC and PAR-6 to the apical surface.

Considering the conservation of polarity establishment in other systems and the localization of PKC-3 to the apical surface of intestinal cells, one could envisage a potential mode of polarization for the intestinal epithelium. At the point of left and right E cell contacts, a polarization signal, perhaps mediated by CDC-42, activates cortical rearrangement and microtubule reorganization at the apical surface. With this rearrangement of cellular components, the PAR-3/PAR-6/PKC-3 complex accumulates at the apical surface. This complex then could actively exclude basolateral markers, potentially PAR-1, through the kinase activity of PKC-3. Although appealing, this model is based solely on the

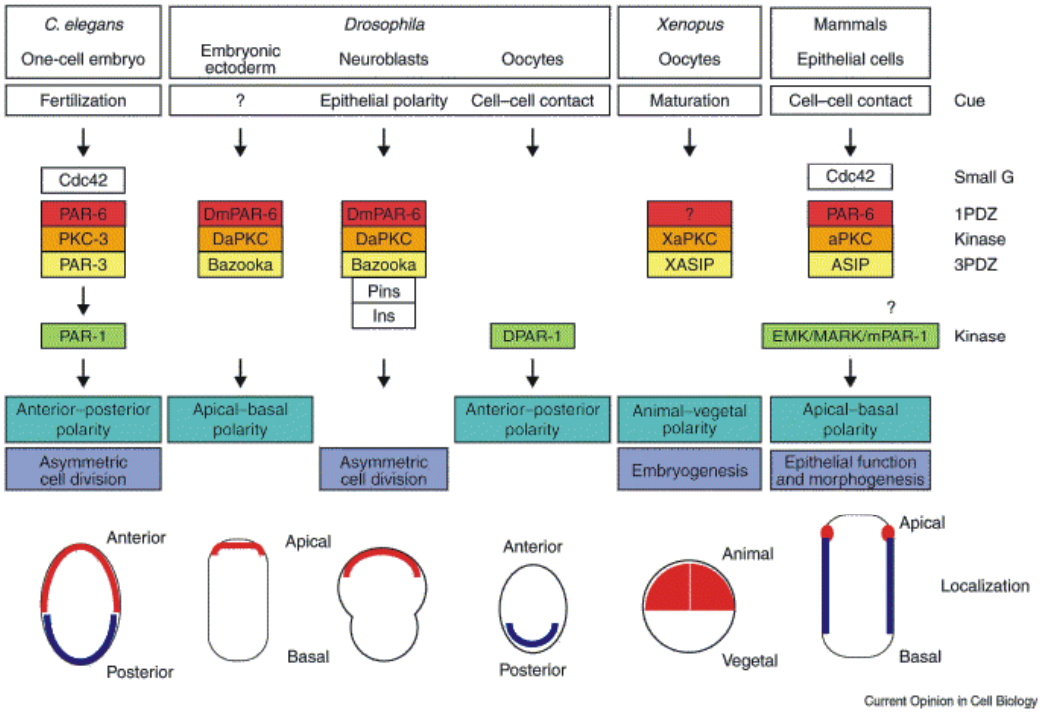


Figure 6: Summary of the conservation of proteins required for polarization. In the *C. elegans* one cell embryo, CDC-42 is required to signal for the localization of the PAR-3/PAR-6/PKC-3 complex to the posterior of the cell. PAR-1 is concomitantly localized to the anterior of the embryo. A similar group of related proteins are required for mammalian epithelium polarization. Cdc42 is localized to the apical surface (where cells are in contact) and signals for the localization of the ASIP/PAR-6/aPKC complex. The EMK/MARK/mPAR-1 protein is localized to the basolateral surface of the cell. Figure taken from Ohno (2001).

functional conservation of components involved in cell polarity and still needs to be experimentally tested.

The second cellular event required for epithelial development is the formation of the *C. elegans* equivalent of the adherens junctions. The formation of the adherens junctions plays two important functions in the epithelium. Adherens junctions are required for a tight adhesion between a single layer of cells. This tight adhesion creates a physical barrier between two different organ compartments. With the differential expression of molecule transporters on the apical vs. basal surface of the epithelium, this physical barrier allows for the directional flow of molecules. Adherens junctions are also required for the formation of a cortical boundary between the apical region and the basolateral region of the cell. The components of the adherens junctions and how they influence each other's localization has been well studied (Bossinger et al., 2004; Segbert et al., 2004). It seems that *C. elegans* has many of the same proteins found in adherens junctions of other species, but some of these proteins have changed their position within the structure and their relationship to other components. The core of the *C. elegans* apical junction is a cellular belt of proteins made up of an apical unit and a basal unit. The apical belt is composed of HMP-2, HMP-1 and HMR-1, which bear similarity to β-catenin, α-catenin and E-cadherin proteins respectively **(see Figure 7)**. The basal unit is composed of the AJM-1, DLG-1 and IFC-2 proteins. AJM-1 is a novel coil-coil protein, DLG-1 is a Discs large related MAGUK domain containing scaffold protein and IFC-2 is an intermediate filament (Bossinger et al., 2004; Bossinger et al., 2001; Firestein

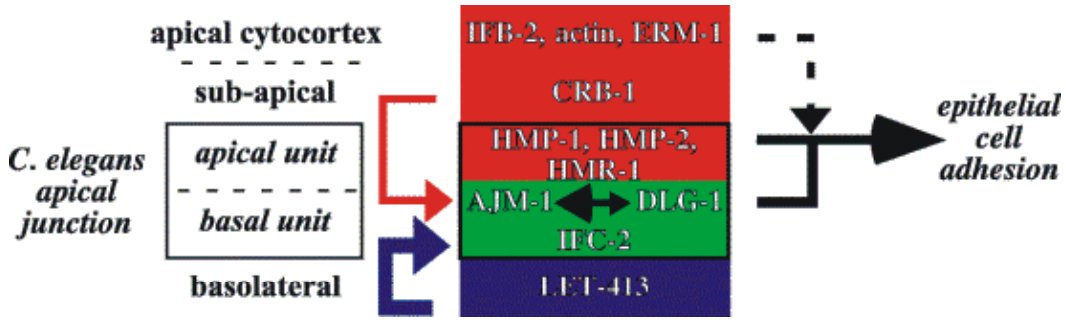


Figure 7: Key proteins involved in adherens junction formation. The apical belt is composed of HMP-2 (β-catenin), HMP-1 (α-catenin) and HMR-1 (E-cadherin). The basal unit is composed of the AJM-1 (coil-coil protein), DLG-1 (Discs large-like protein), and IFC-2 (intermediate filament). Both units act redundantly to maintain the adherens junction. LET-413 is required for proper localization of DLG-1 (Bossinger et al., 2001; McMahon et al., 2001). IFB-2 (intermediate filament), actin and ERM-1 (Ezrin-Radixin-Moesin-like protein) are localized to the apical cortex (Bossinger et al., 2004; Gobel et al., 2004). Figure taken from Segbert et al. (2004).

and Rongo, 2001; Koppen et al., 2001). In addition to the core of the apical junction, the cell is divided into three domains. CRB-1, a Crumbs related protein, forms a slightly more apical belt with regards to the core unit. In Drosophila, Crumbs is a transmembrane protein that binds to MAGUK domain containing proteins (Segbert et al., 2004). A basolateral domain is defined by the basolateral distribution of LET-413, which is a Drosophila Scribble-like homologue and a PDZ domain containing protein (Bossinger et al., 2004; McMahon et al., 2001a). Along the apical cortex, several proteins are found such as actin, IFB-2 (an intermediate filament), and ERM-1, (an Ezrin-Radixin-Moesin-like protein) (Bossinger et al., 2004; Gobel et al., 2004). Members of the core apical unit and the basal unit act redundantly to maintain the integrity of the adherens structure. LET-413 is required for the correct localization of the complex and directly affects the localization of DLG-1 and AJM-1 (Bossinger et al., 2001; McMahon et al., 2001a). CRB-1 is not necessary for junction formation but seems to provide a redundant assembly cue for DLG-1 in the absence of the catenin/cadherin and LET-413 proteins.

The third cellular event required for epithelial development is the maintenance of the adherens structure. LET-413 is required for this maintenance phase, as proteins of the apical cortex such as IFB-2 move to more basolateral positions if this protein is absent (Bossinger et al., 2004).

E Cell Specification

As stated above, the intestine is derived from a single cell in the eight cell embryo **(see Figure 8)**. The early divisions that give rise to the E blastomere have been well described (Sulston et al., 1983). Two AB daughter cells, the P2 cell and the EMS cell are present in the four cell embryo. EMS divides to form the anterior MS cell, founder of posterior pharynx and muscle, and the posterior E cell, founder of the intestine lineage. In order for the E cell to be properly specified, the EMS cell must undergo a polarization event and an asymmetric division of determinants to its daughter cells. Several efforts have been made to elucidate the signaling events in generating this EMS polarization.

Early experiments established the importance of cell contact and signaling between the P2 and EMS cell (Goldstein, 1992; Goldstein, 1993; Goldstein, 1995a; Goldstein, 1995b; Schierenberg, 1987). Contact with the P2 cell allowed for segregation of intestine differentiating potential to the posterior EMS daughter and reoriented the cell division plane to be perpendicular to the site of cell contact. Without this P2/EMS cell contact, the intestine was found never to form (Schierenberg, 1987; Goldstein, 1992). Blastomere isolation and culturing experiments defined three key properties of the P2/EMS signaling event. Any side of the EMS cell could respond to the signaling event. Once the signaling event had been established, a second site could not be induced. The EMS had a short time period in which it could respond to the P1 signal; in contrast, the P2 cell and descendants could produce the polarization signal for up to three cell

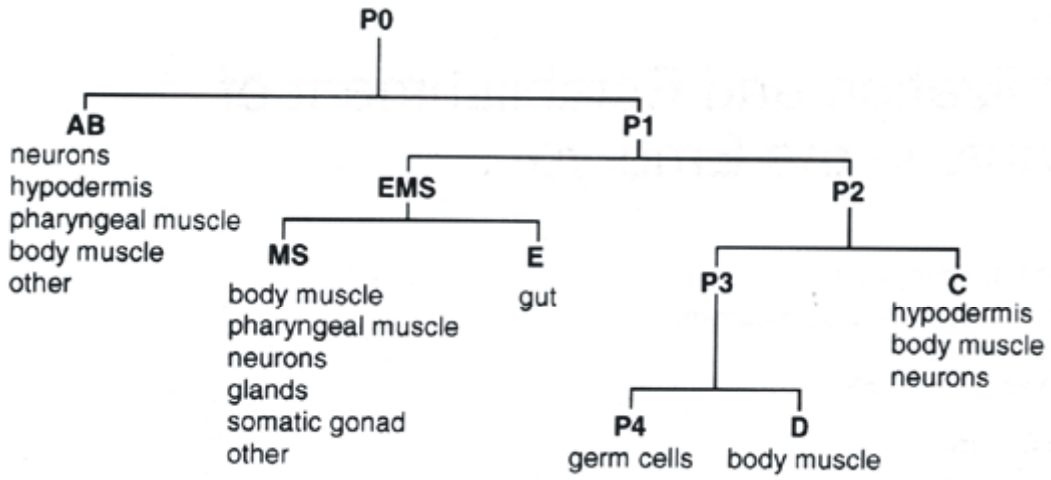


Figure 8: The intestine arises from a single cell in the early embryo. Diagrammed above are the series of early cell divisions of the embryo and the tissues to which these cells eventually give rise. The E cell is the only cell that gives rise to an exclusive tissue type. Figure taken from Riddle et al. (1997).

cycles.

It was later determined that at least three pathways were involved in the P2 to EMS signaling event: the Wingless (Wnt) signaling cascade, a Nemo-like kinase (MAPK) signaling cascade, and a Src-like mediated signaling cascade (Bei et al., 2002; Korswagen, 2002; Thorpe et al., 2000). Genetic screens and reverse genetics were initially used to identify the following members of the Wnt signaling pathway involved in E cell determination: *mom-1* (Porcupine), *mom-2* (Wnt), *mom-5* (Frizzled), *sgg-1*/*gsk-3* (GSK-3β), *wrm-1* (β-catenin), *apr-1* (APC), and *pop-1* (TCF/LEF1) (Lin et al., 1995a; Rocheleau et al., 1997; Schlesinger et al., 1999; Thorpe et al., 1997). *pop-1* was determined to negatively regulate E cell specification; in its absence, both cells took on an E cell fate. Other members of the Wnt pathway were required to positively specify the E cell fate; in their absence, both daughter cells acquired the default MS cell fate. Genetic experiments in which loss of function *pop-1* alleles were combined with loss of function alleles of other members of the Wnt pathway suggested that Wnt signaling was required to down-regulate *pop-1* activity. It was later discovered that POP-1 down-regulation is occurring by the active export of the protein from the nucleus of the E cell (Lo et al., 2004). POP-1 is related to the family of TCF/LEF1 transcription factors. Differential modulation of POP-1 transcriptional activity in the MS vs. the E cell seems to be the major result of the Wnt-mediated signaling during endoderm specification (Lin et al., 1995).

A Nemo-like kinase pathway has been found to converge on the Wnt pathway to positively regulate E cell specification. *mom-4*, a TAK-1 (TGF-β

activated kinase)-like protein, and its co-activator TAP-1 (Tab-1-like protein), are the likely phosphorylators/activators of the LIT-1 protein, a Nemo-like kinase and member of the MAPK-like kinases (Kaletta et al., 1997; Shin et al., 1999; Smit et al., 2004). *In vitro* studies have shown that the phosphorylated LIT-1 protein forms a complex with WRM-1 (β-catenin) protein and in turn phosphorylates POP-1. Phosphorylation of POP-1 then signals for its removal from the E cell nucleus. *In vitro* assays and genetic results have converged to suggest that the *mom-4* cascade is activated by the Wnt pathway. Therefore the Wnt signal converges on both the members of the Wnt pathway and members of the MAPK kinase pathway to down-regulate POP-1.

The two tyrosine kinases, *src-1* (a c-Src related protein) and *mes-1*, are involved in a signaling pathway parallel to the Wnt pathway during E cell specification (Bei et al., 2002). *src-1* and *mes-1* were proposed to act in a parallel pathway based on the presence of genetic synergy with members of the Wnt signaling pathway. Activation of the Src pathway relies on the P2/EMS cell contact and results in the accumulation of phosphorylated tyrosine at the site of cell/cell contact. Despite having an activation signal independent from the Wnt ligand, the Src pathway seems to converge on the regulation of WRM-1 and LIT-1 and therefore, theoretically, POP-1

In addition to their role in specifying endoderm, the Src pathway and part of the Wnt signaling pathway are required for controlling the orientation of the plane of division (Bei et al., 2002; Schlesinger et al., 1999). How *src-1* and *mes-1* are controlling spindle orientation is currently unknown. Upstream components

of the Wnt pathway, up to and including *sgg-1*, are required for proper spindle orientation (Schlesinger et al., 1999). The components downstream of *sgg-1* required for spindle orientation do not require active transcription and this suggests that they may have a direct interaction with cellular components such as microtubules and the actin cytoskeleton.

In many ways, the Wnt signaling pathway required for endoderm specification in the worm has deviated from the canonical pathway established in flies, mammals and other Wnt-like pathways in the worm (reviewed by Korswagen, 2002). In the canonical pathway, the Wnt signal acts to stabilize β-catenin by preventing its targeted degradation by Axin/APC/GSK-3β (Bienz and Clevers, 2000; Cadigan and Nusse, 1997 and references therein). With its stabilization, β-catenin is able to enter the nucleus and convert the TCF/LEF1 related transcription factor from a repressor to an activator. Conventional β-catenin homologues have a strong trans-activation domain. It is believed that this trans-activation domain and competition with co-repressors such as Groucho and histone deacetylases allow for its repressor-to-activator switch (Billin et al., 2000; Brantjes et al., 2001; Hecht et al., 2000; Levanon et al., 1998; Roose et al., 1998). In the canonical pathway, the Nemo-like kinase signal acts as a negative feedback to the Wnt signaling. Upon Wnt activation of the Nemo-like kinase pathway, TCF/LEF1 proteins are phosphorylated and inhibited in their activation function (Ishitani et al., 1999; Smit et al., 2004; Thorpe and Moon, 2004). In *C. elegans,* the endoderm specification pathway is unusual in that WRM-1 cannot form a stable complex with POP-1 such as seen with vertebrate β-catenin and

TCF proteins (Korswagen et al., 2000). Secondly, WRM-1 is missing the GSK-3β phosphorylation site for degradation. Instead, SGG-1 along with APR-1 (APC) and PRY-1 (Axin) are activated by the Wnt signal and complex with WRM-1 to sequester POP-1 from the nucleus (reviewed by Thorpe et al., 2000). The GSK-3β sequestering activity is enhanced by the Nemo-like kinase phosphorylation of POP-1 and POP-1 exclusion from the nucleus (Smit et al., 2004).

Transcription Factor Network

The transcription factor network leading up to and required for endoderm specification has been studied in detail **(see Figure 9)**. At least three transcription factors are required for establishment of the EMS cell fate. One maternally contributed transcription factor, SKN-1, has its localization restricted to the EMS and P2 cell of the four cell embryo (Bowerman et al., 1993). *skn-1* is absolutely required for MS cell fate (posterior pharynx and muscle) and partially required for E cell fate (intestine proper) (Bowerman et al., 1992). SKN-1 is a member of the basic leucine zipper (bZIP) family of transcription factors (Blackwell et al., 1994). It seems likely that SKN-1's main function is the activation of two nearly identical GATA-like transcription factors *med-1* and *med-2* (Maduro et al., 2002; Maduro and Rothman, 2002). SKN-1 activation of *med-1*

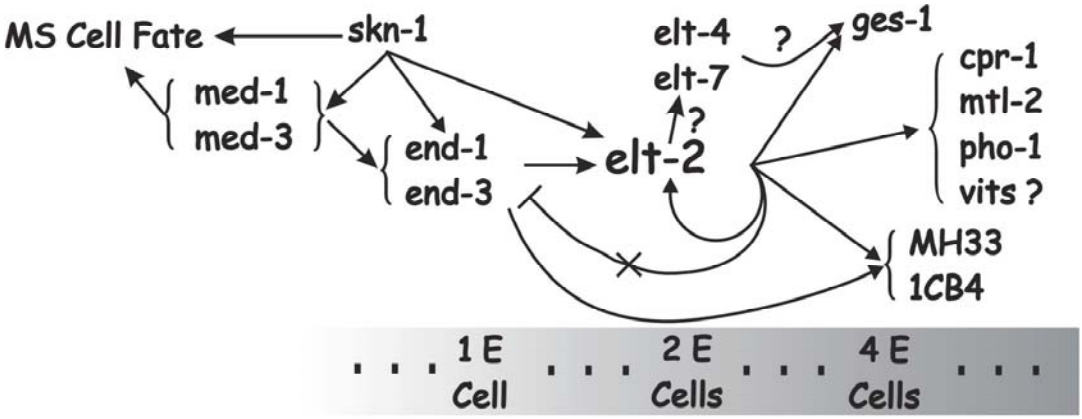


Figure 9: The network of known transcription factors driving intestine differentiation. *elt-2* is believed to be the main driving force behind intestine differentiation (Fukushige et al., 1998). ELT-2 is sufficient to drive expression of most early intestine markers tested. ELT-2 is not necessary for expression of early intestine markers and therefore unidentified redundancy is evident in the above pathway.

and *med-2* is one of the first known switches from maternal to zygotically expressed transcription factors in the early embryo. Double *med-1* and *med-2* RNAi experiments show that these two proteins are absolutely required for specification of MS derived posterior pharynx and muscle and partially required for specification of E derived intestine. In the absence of *med-1/2*, the EMS blastomere takes on the fate of the C cell. The C cell arises from the asymmetric division of the germline precursor, P2, and gives rise to hypodermis and a distinct subset of body wall muscles. MED-1 over-expression is sufficient to cause ectopic expression of MS and E cell fates. The model where SKN-1's main function is to activate *med-1* and *med-2* expression has been supported by the four following experiments. Ectopic expression of SKN-1 can ectopically express *med-1* and *med-2* reporters. The SKN-1 protein has also been shown to be able to directly bind to the promoter of *med-1*. The over-expression of MED-1 and MED-2 can overcome the requirement for SKN-1. Most convincingly, *med-1* and *med-2* are required for the effect of SKN-1 over-expression.

The first transcription factors to be expressed only in the E cell lineage are the two GATA-like transcription factors *end-1* and *end-3* (Zhu et al., 1998; Zhu et al., 1997). *end-1* mRNA is first detected in the 1 E cell but disappears by the 4 E cell stage. Ectopic expression of END-1 has been shown to be sufficient for ectopic intestine formation (Zhu et al., 1998). Although *end-1* is not necessary for intestine formation, the double knockdown of *end-1* and a closely related GATA factor *end-3* results in a penetrant loss of intestine (unpublished results cited by Maduro and Rothman, 2002). It has been postulated that the main

output of the *skn-1, med-1* and *med-2* transcription factors, along with the Wnt signaling cascade, is the activation of *end-1* in the E cell. This model is supported by the finding that over-expression of END-1 can overcome the requirement for *skn-1* and *wrm-1* in intestine differentiation (Zhu et al., 1998). Detection of the *in vivo* binding of a tagged MED-1 protein to transgenic arrays of the *end-1* promoter, suggests that MED-1 binds directly to the *end-1* promoter in both MS and the E cell (Maduro et al., 2002). The similar detection of POP-1 binding to the *end-1* promoter in the MS but not the E cell suggests that POP-1 negatively regulates *end-1* expression in the MS cell and possibly allows for *med-1*-driven expression of distinct differentiation pathways in the MS vs. the E cell.

The main output of these early transcription and signaling events is believed to be the up-regulation of the GATA-like transcription factor ELT-2. Upon its expression, ELT-2 is proposed to be the main transcription factor driving intestine differentiation. Antibodies generated against ELT-2 detect the protein starting at the 2 E cell stage (Fukushige et al., 1998). ELT-2 is then expressed specifically in the intestine throughout the life of the worm. This continued expression of ELT-2 even after the disappearance of END-1 is likely due to the ability of the ELT-2 protein to auto-regulate its expression. Ectopic expression of END-1 and SKN-1 causes ELT-2 to be ectopically expressed, while ectopic expression of ELT-2 is able to drive expression of intestine-specific markers such as *ges-1*, *ifb-2*, and gut granules (Fukushige et al., 1998; Zhu et al., 1998). Homozygous *elt-2* null mutants arrest as early larvae and possess a severely

malformed intestine, supporting the model that *elt-2* is necessary for proper development of the intestine. *elt-2* is also likely to have a necessary function in the larval and adult intestine. Although *elt-2* is not necessary for expression of early intestine markers, *elt-2* is necessary for expression of genes, such as *pho-1*, that are expressed during later intestine development (T. Fukushige and J. McGhee personal communication).

Significant redundancy appears to exist within the current network of known transcription factors. During the initial specification of the EMS and E cell fate, knockdown of *skn-1*, *med-1*/*2* and *end-1*/*3* have not been able to completely eliminate formation of the intestine. This is despite the observation that some mutations in members of the Wnt signaling pathway, alone or in combination, have resulted in a complete loss of intestine (i.e., the *lit-1* mutant alone (Kaletta et al., 1997), the *wrm-1*(RNAi) mutant alone, the *apr-1*(RNAi) *mom-5* double mutant (Rocheleau et al., 1997), the mo*m-4 mom-2* double mutant, the *mom-5*(RNAi) *mom-4* double mutant and the *apr-1*(RNAi) *mom-4* double mutant (Thorpe et al., 1997; Shin et al., 1999). This scenario suggests that elimination of intestine formation is possible and that the Wnt signaling pathway is regulating expression (directly or indirectly) of undiscovered transcription factors.

Furthermore, although *elt-2* is sufficient to drive expression of gut specific genes, it is not necessary for expression of several intestinal markers such as *ges-1*, *ifb-2* and gut granules. To date only one intestine specific gene, *pho-1*, absolutely requires *elt-2* for its expression (T. Fukushige and J. McGhee personal communication).

Attempts at finding other GATA factors that could be acting redundantly with the currently known GATA transcription factors have provided few answers. For instance, a small GATA factor, *elt-4*, located directly upstream of *elt-2* and likely the result of a recent duplication of *elt-2*, has no obvious function alone or in combination with *elt-2* (Fukushige et al., 2003). A second GATA factor, *elt-7*, that is most closely related to *elt-2* and *elt-4* in sequence and expression pattern, has been found to have a variable penetrance with regards to its effect on the intestine. Loss of *elt-7* function alone or in combination with loss of *elt-2* function has not resulted in loss of expression of intestine specific markers (personal communication K. Strohmaier, J. Rothman and J. McGhee).

Some non-GATA transcription factors have been found to play a role in intestine differentiation or maintenance and therefore provide promising candidates for redundant factors. *skn-1*, the maternal bZIP transcription factor required for EMS specification, also has a zygotic function in intestine maintenance (An and Blackwell, 2003; Bowerman et al., 1992). Mutants that remove zygotic function of *skn-1* cause larval lethality due to a progressive degeneration of the intestine (Bowerman et al., 1992). *skn-1* is also expressed at the right time and place to be involved in intestine function. Anti-SKN-1 antibodies detect SKN-1 protein in the hypodermis and intestine of the elongating embryo (Bowerman et al., 1993). More recently, a SKN-1::GFP translational reporter is able to detect SKN-1 translation in the intestine as early as the 4 E cell stage and this expression remains strong in the intestine until late

embryogenesis (An and Blackwell, 2003). Weaker SKN-1::GFP expression is detected in the intestines of larvae and young adults.

SKN-1's role in maintenance of the gut may involve the regulation of genes required for resistance to oxidative stress (An and Blackwell, 2003). Exposure to oxidative stress causes a variable increase in expression of SKN-1::GFP in the intestinal anterior and posterior nuclei.

An and Blackwell (2003) have suggested that expression of detoxifying enzymes and the transcription factors that regulate them in endodermally derived organs have been evolutionarily conserved. SKN-1 is unusual compared to other bZIP family members in that it has a highly divergent DNA binding domain (Blackwell et al., 1994). Nonetheless, SKN-1 possesses a highly conserved DIDLID transactivation domain that is present in mammalian Nrf related proteins. Nrf proteins, in animals as diverse as yeast and vertebrates, have been found to be functionally involved in regulation of detoxification genes (An and Blackwell, 2003; Walker et al., 2000). For instance, in the mouse, *Nrf2* is expressed in liver, stomach and small intestine and is required for the detoxification response in these organs (McMahon et al., 2001b).

Reverse genetics has been used to find other transcription factors involved in intestine development in the worm. Two genes *odd-1* and *odd-2* were identified based on their sequence similarity to *odd-skipped* homologues in Drosophila (Buckley et al., 2004). *odd-1* and *odd-2* GFP transcriptional reporters were used to define the expression pattern of these genes. Both genes are expressed in the intestine just prior to elongation and this intestine expression

continues into the adult. Expression in the embryo was also observed in the head and in cells surrounding the intestine. RNAi to *odd-2* caused an early larval arrest, a decrease in pharynx pumping, a pharynx/intestinal valve defect and a severely malformed intestine. *odd-1* RNAi had a similar phenotype with the exception of additional cavities being detected around the pharynx. Double *odd-1*/*2* RNAi resulted in similar defects to either single mutant with the exception of an additive valve phenotype. Together these results suggest that *odd-1* and *odd-2* may act in a similar pathway for intestine and pharynx development but have redundant function in pharynx/intestinal valve formation. It has been predicted that non-endodermal functions of the odd-skipped genes is a derived state and that their ancestral function is in endoderm development. *drumstick* and *bowl* are Drosophila *odd-skipped*-like genes with roles in gut development and terminal end specification respectively. The expression of the human *odd-skipped*-like gene *OSR-2* in the colon and small intestine and its over-expression in pancreatic and intestinal cancers has suggested an important intestinal function for this gene (Katoh, 2002). Although intriguing, more research needs to be done in this area in order to strengthen the conservation argument.

The gene *rnt-1* was isolated based on its similarity to *runt*-like genes in Drosophila and was found to have a role in intestine and hypodermis development in the worm (Lee et al., 2004; Nam et al., 2002). A RNT-1::GFP translational reporter detected protein expression in the intestine in late embryonic stages through to the L3 larval stage. RNT-1 was also detected earlier in the hypodermis and persisted in the hypodermis until the L3 larval

stage. RNAi to the *rnt-1* resulted in a moderately penetrant embryonic lethality and larval arrest. The arrested larvae displayed severe hypodermal defects and malformed intestines. This supported a role for *rnt-1* in hypodermis and intestine development. A mammalian *runt* related gene *Runx3* is expressed in gastrointestinal epithelial cells and is involved in controlling cell proliferation of these cells (Li et al., 2002). Whether there is a direct evolutionary link between these two genes is unclear. Investigation of the role of *runt*/*RUNX*-like genes in other organisms may help resolve this ambiguity.

Conservations Between *C. elegans* and Other Model Systems

One goal in understanding intestine development in the worm is to be able to gain a better understating of the development of more complex organisms such as humans. The relative ease in studying *C. elegans* can allow for a much more thorough and in-depth study of organs such as the intestine. Potentially, findings in worms can then be applied to the understanding of intestine development in other systems. The validity of this approach relies on the assumption that regulatory networks involved in intestine development in the worm are similar to those in vertebrates. Progress in the understanding of worm, fly and vertebrate development is starting to provide evidence that this may be the case and that some transcription factor networks are conserved across phyla. Some of the similarities and dissimilarities that have been discovered to exist between various model systems are discussed below.

In general, it seems that very different strategies have evolved to establish germ layers in various organisms. However, once germ layers have been established, the transcription factors that drive differentiation of organs have been conserved. For instance, the Nodal (TGFβ) signaling and antagonism by BMPs (Bone Morphogenetic Proteins) required for endoderm formation in vertebrates does not seem to be active in endoderm formation in the worm (Shivdasani, 2002; Stainier, 2002). Furthermore, the Notch signaling required to separate endoderm from mesoderm and the requirement for T-box related transcription factors (VegT in *Xenopus laevis*, Eomesodermin in mouse and Spadetail in zebrafish) in endoderm specification is likely to be specific to vertebrates. Only the Wnt signaling pathway and their downstream effectors, the TCF/LEF1 family of transcription factors, could be a potentially conserved mode of germ layer and endoderm specification. Five major groups of transcription factors are involved in vertebrate endoderm differentiation: the Fork head-like transcription factors, the Mix family of paired-class homeodomain transcription factors, the GATA-like transcription factors, the Sox/HMG family of transcription factors, and the HNF4α nuclear hormone receptor family of transcription factors (reviewed in Roberts, 2000; Shivdasani, 2002; Simon and Gordon, 1995; Stainier, 2002; Zaret, 1999). Certainly two (and potentially two more) of these transcription factor families are involved in worm intestine development.

No obvious Mix-like transcription factors have been identified in the worm genome sequence nor have any diverged family members been shown to be involved in endoderm differentiation. Sox17 in Xenopus and related proteins

(i.e., Casanova in zebrafish and Sox17 in mice) are members of the HMG family of transcription factors and play pivotal roles in the differentiation of endoderm in these organisms. No obvious homologues have been identified in the *C. elegans* genome. However, despite the inability to find obvious homologues of Sox17, other Sox domain containing proteins have been found in the *C. elegans* genome. At least one of these genes, *cog-2*/*egl-13*, has been found to be expressed in the intestine and therefore may have a role in endoderm development (Hanna-Rose and Han, 1999).

It is not currently evident whether the function of the Wnt signaling pathway in vertebrates is a conserved or independently evolved mechanism of endoderm specification. β-catenin is positively required for endoderm specification in sea urchins and ascidians and Wnt signaling is required for germlayer patterning in vertebrates. Wnt signaling is also required for endoderm formation in worms (see above) but has the opposite affect of removing TCF and β-catenin from the nucleus. Clearly, if Wnt signaling has been conserved during endoderm specification, the components of the pathway and its net effect have diverged during evolution.

The function of nuclear hormones receptors during intestine differentiation has been proposed to be conserved between worms and mice (Maduro and Rothman, 2002). *end-2* later renamed *dpr-1* (for dauer phermone responsive) was identified along with *end-1* and *end-3* to be one of the genes responsible for the gutless phenotype of a large deletion termed the “endoderm-determining region” (EDR) (Zhu et al., 1997). Ectopic expression of END-2 was able to

partially rescue endoderm formation in the EDR homozygous strain. It has been suggested by Maduro and Rothman (2002) that this nuclear hormone receptor may be homologous to the *HNF4α* gene in mice. *HNF4α* has been found to be expressed in the definitive endoderm of the mouse embryo and is required for development of hepatic epithelial structures (Duncan et al., 1994; Taraviras et al., 1994; Parviz et al., 2003). Since the function of the *end-1* gene has not been studied in detail, it is difficult to determine if this is a true case of homology.

The GATA and Fork head related transcription factors have been most highly conserved during metazoan digestive tract development. *pha-4*, a *fork head* related transcription factor, has been found to play a critical role in foregut and hindgut development in *C. elegans* (Kalb et al., 1998; Mango et al., 1994). *pha-4* is expressed in all pharynx cells and in their precursors, and this expression occurs despite the diverse lineal origin of the pharyngeal cells. To date most, if not all, genes that have been found to be expressed in the pharynx have PHA-4 regulatory sites in their promoters (Gaudet and Mango, 2002). The above properties of *pha-4* have lead to it being defined as an organ identity factor. Fork head transcription factor regulation of foregut and hindgut specification has been highly conserved throughout metazoan evolution. The Drosophila gene *fork head* is expressed in the terminal regions of the developing embryo and has been found to be necessary for foregut and hindgut development (Weigel et al., 1990; Weigel et al., 1989). A related family of transcription factors, HNF3α, β and γ in the mouse, has also been shown to be

involved in anterior and posterior endoderm development (reviewed in Zaret, 1999). All three *HNF3* isoforms are expressed during endoderm development but *HNF3β* has the most striking requirement by being necessary for proper foregut formation. Similar genes, *fkd1*,*2*,*4* and *7* in zebrafish, are also expressed in the anterior and posterior endoderm but the consequence of knockdown of these genes in the developing endoderm has not yet been determined (Odenthal and Nusslein-Volhard, 1998). It is difficult to draw direct parallels between development of the foregut and hindgut in invertebrates and vertebrates. The origins of these tissues in each organism are very different (i.e., mesodermal in *C. elegans*, ectodermal in flies and endodermal in vertebrates). The similarities are nonetheless intriguing and suggest that *fork head*-like transcription factors have been ancestrally involved in foregut and hind gut development.

Transcription factors of the GATA family also have a remarkable conservation during endoderm development. As reviewed earlier, there is a cascade of GATA transcription factors driving mesendoderm (*med-1*/*2*) and endoderm (*end-1*/*3* and *elt-2*) development in the worm. Although none have been found to be necessary for endoderm formation, all are sufficient to drive expression of intestine markers. *elt-2* is the only GATA factor expressed in both embryonic and adult intestine. Out of all the GATA factors detected in *C. elegans*, *elt-2* is most closely related to the *serpent* GATA factor in flies and the *GATA-4/5/6* GATA factors in vertebrates. *serpent* is necessary for midgut development in Drosophila (Reuter, 1994). A second Drosophila GATA factor, *grain*/*dGATAc*, is expressed in the anterior and posterior midgut primordia (Lin et

al., 1995b). The role of *grain* in midgut is unclear but *grain* is required for epithelial migration/invagination in other tissues (Brown and Castelli-Gair Hombria, 2000). The *GATA-4*/*5*/*6* genes are expressed in the endoderm of the mouse (Laverriere et al., 1994). *GATA-4* is necessary for foregut and hindgut development in mice and is required for proper differentiation of the gastric epithelium (Narita et al., 1997; Jacobsen et al., 2002). A *gata5* homologue in zebrafish is necessary for endoderm development. In the zebrafish, Gata5 is sufficient to drive expression of endodermal genes and is necessary for gut tube and endoderm derived organ development (Reiter et al., 2001). The role of *GATA-6* has not yet been determined.

Both Fork head and GATA transcription factors have been suggested to be “potentiators” of endoderm differentiation (Bossard and Zaret, 1998; Zaret, 1999). Both families are expressed early during the formation of the endoderm-derived organs. Furthermore, Fork head and GATA transcription factors are the earliest transcription factors shown to bind the promoters of endodermally expressed genes and have been found to occupy these promoter regions prior to the onset of differentiation. The pioneering function of Fork head and GATA transcription factors, and their functional and sequence conservation across metazoa, suggest that these two families may be ancestral drivers of digestive tract development. Some of the strongest evidence supporting this evolutionary conservation is the ability of END-1 to ectopically express endoderm markers in Xenopus embryos (Shoichet et al., 2000). In similar experiments, an inhibitory END-1 construct has been shown to disrupt endogenous endoderm

development. These results show that despite considerable divergence, *C. elegans* GATA factors are still able to bind and drive expression of endoderm specific genes in vertebrates. This suggests conservation of GATA factor driven endoderm development at two levels. First, GATA factors have been conserved as endoderm potentiators in invertebrates and vertebrates. Secondly, there has been a considerable retention of GATA factor binding sites in endodermally expressed genes.

Overall it appears that there has been a considerable conservation of genes required for endoderm development and function across phyla. Fork head and GATA transcription factors provide the strongest evidence for conservation of pathways driving establishment of the digestive tract. Genes required for later digestive tract function such as the *Nrf*, *odd-skipped* and *runt*/*RUNX*-like genes also have potentially conserved roles in worms and vertebrates. These conservations define *C. elegans* as a good organism for finding evolutionarily retained genes involved in endoderm development and differentiation. This conservation is perhaps not that surprising given the context of the *C. elegans* and human genome. Humans have 26,588 predicted proteins in their genome while *C. elegans* has 19,099 (Consortium, 1998; Consortium, 2001; Venter et al., 2001). This suggests that, at least, at the protein coding level, *C. elegans* and humans have a similar degree of complexity. In fact, up to 74% of studied human proteins (4,979 considered) have related proteins in C. elegans (Consortium, 1998). More conservative estimates predict 2031 orthologs exist between worms and humans (Venter et al., 2001). Many of the differences

between worm and humans have been attributed to expansion of acquired immune functions, neuronal functions, hemostasis, apoptosis and signaling pathways during development and homeostasis in humans. This leaves the possibility for potential conservation of transcription factors involved in digestive tract development.

Thesis Overview

The founding purpose of this thesis was to identify additional genes involved in intestine development and differentiation in *C. elegans*. Based on identified conservation, as discussed above, it seemed reasonable that genes found to be important for worm intestine development/function could also have a function in vertebrate intestine development/function. Past research has identified a network of GATA factors to be at the core of worm intestine specification but considerable redundancy is evident within this pathway. In the search for new factors involved in intestine development, I developed a genetic screen for intestine defects. This screen was based on the phenotype of the *elt-2* null mutant where an intestine is formed but is morphologically abnormal (Fukushige et al., 1998). This defect was termed the gut obstructed (Gob) phenotype because the resulting intestine defects caused an inability for worms to pass material through their digestive tract. To enhance the likelihood of finding genes specifically involved in intestine development, I applied strict screening criteria. Other screens looking for defects in endoderm development had

focused on early phenotypes and the formation of the E lineage (Rocheleau et al., 1997; Thorpe et al., 1997; Zhu et al., 1997). A deficiency screen covering 74% of the genome looking for morphogenesis defects identified relatively few regions that are required for intestine development (Labouesse, 1997). The failure to identify such regions in this latter screen could be attributed to the early lethality caused by homozygous deletions. The gut obstructed screen was the first of its type looking for early larval intestine phenotypes. All evidence, at the onset of development, suggested that it would be a good method for identifying new genes involved in intestine development.

Early in the progress of the screen, a mutant candidate *gob-1*(*ca17*) was positionally cloned and characterized with regards to its role in intestine development. *gob-1* turns out to be a gene expressed in the intestine during embryonic development. The protein is a new member of the HAD (Haloacid dehalogenase)-like hydrolase super family and has a robust and specific phosphatase activity for the substrate trehalose-6-phosphate. Trehalose-6-phosphate is an intermediate in the synthesis of the disaccharide trehalose. Trehalose is an important sugar in *C. elegans* with obvious functions in metabolism and stress tolerance. Despite GOB-1's defined activity as a trehalose-6-phosphate phosphatase, a direct relationship of intestine formation to trehalose metabolism has been difficult to find. Mounting evidence suggests that *gob-1* may have a role independent of trehalose synthesis. The Gob-1 screen, the characterization of *in vivo gob-1* expression and function of the GOB-1 protein is the topic of this thesis.

CHAPTER TWO - Materials and Methods

Strains

The following strains and mutant alleles were used: N2 Bristol strain, Hawaiian CB4856 CB subclone of HA-8, *eat-13*(*ad522*), *eat-15*(*ad602*), *eat-17*(*ad707*), *pha-2*(*ad472*), *aex-1*(*sa9*), *unc-104*(*rh43*), *gsa-1*(*pk75*), *hmp-2*(*zu364*), *let-80*(*s96*), *snb-1*(*js124*), *let-23*(*sy17*), *unc-7*(*e5*), *elt-2*(*ca15*), + / *szT1*[*lon-2*(*e678*)] I; + / *szT1* X, *lon-2*(*e678*), *tra-2*(*q276*), *mnDp1* (X;V) / +; *mnDf8* X, *mnDp1* (X;V) / +; *mnDf43* X, *mnDp1* (X;V) / +; *mnDf20* X, *mnDp1* (X;V) / +; *mnDf13 X*, *lin-15*(*n765ts*), *dpy-5*(*e61*), *unc-13*(*e1091*), *mnDp1* (X;V) / +; *gob-1*(*ca17*) *lin-15*(*n765ts*), *rrf-3*(*pk1426*), *dpy-8*(*e130*), *unc-9*(*e101*), *lin-2*(*e1309*), *lin-14*(*n179*) and *unc-3*(*e151*).

Culturing Techniques

Worms were cultured as previously described (Wood 1988).

EMS Mutagenesis

The mutagenesis protocol was slightly modified from Brenner (1974). Healthy N2 L4s were washed with M9 buffer (42.2 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 85.6 mM NaCl and 1.0 mM $MgSO_4$) and incubated, with gentle rocking, for three to five hours in a 15 mL solution of M9 buffer and 25 mM EMS (ethyl

methanesulfonate). Worms were washed three times and placed on a seeded plate. L4s were allowed to lay eggs and their F1 progeny were isolated to individual plates.

Worm Lysis for PCR

The worm lysis procedure used to extract DNA from worms for PCR was modified from Truett et al. (2000). One to five individual worms were placed into 2 µL of alkaline lysis solution (25 mM NaOH and 0.2 mM disodium ethylene diamine tetraacetic acid (EDTA), covered with 6 µL of paraffin oil, heated at 95°C for 20 minutes and then cooled to 4°C. Two µL of neutralizing solution (40 mM Tris-HCl) and then two µL of 3X worm lysis solution (150 mM KCl, 7.5 mM $MgCl_2$, 30 mM Tris-HCl pH 8.3, 1.35 % Tween 20, 1.35% NP40, 0.003% gelatin and 1.25 mg/mL Proteinase K) was added to the mixture. The contents were heated at 60°C for 60 minutes, 95°C for 15 minutes and then cooled to 4°C.

PCR Detection of the *elt-2* Wild-type vs. Deletion Allele

The primers HawN ctttgtagggtaattgaggttc, HawB gtagggtacacatgtttg and HawR gtctgatcgcctgttttatctac were used in the PCR detection of the wild-type *elt-2* allele and the *elt-2*(*ca15*) deletion allele (Fukushige et al., 1998). The wild-type *elt-2* allele would result in a PCR fragment of 3240 bp in size with the HawN/HawB primers and a 2910 bp fragment with the HawR/HawB primers.

The *elt-2*(*ca15*) allele would produce a band of 1012 bp with the HawN/HawB primers and a 682 bp fragment with the HawR/HawB primers.

Bead Feeding Assay

The fluorescent microspheres used in the bead feeding assay were purchased from Bangs Beads (Fishers, IN). The beads were manufactured from a uniformly dyed polymer and were carboxyl group modified. They had a mean diameter of 0.19 μm and had an excitation and emission wave length of 460 and 480 nm respectively. A mixture containing 40 μL of a 1% bead stock solution (sterilized by dialysis against ethanol), 10 μL of an overnight OP50 bacteria culture and 50 μL of M9 buffer was spread onto a five cm NGM agar plate. Young adults were allowed to lay eggs on the plate overnight and removed the next morning. The F1 progeny were grown to L4/young adults and removed with M9 onto bead-free plates. Plates were scanned 12 to 18 hours later under a dissecting scope fitted with epifluorescence for animals with strong fluorescence in the intestine.

snip-SNPs Used for the Positional Cloning of *gob-1*(*ca17*)

The following primers and restriction sites were used for each single nucleotode polymorphism (snip-SNP) (Wicks et al., 2001): C33A11 1933 gttttcactggcgctactcc, cggatatcgacctgaaagatg cut with *Hae*III; F01G12

tagatatcgtggaacccc, cctgggaatccgttttctcc cut with *Hind*II; C23H4 ccttcaatcgcttttgtagg, gatttacgtaacggctgagg cut with *Bst*BI; C05E7 ggctctgagaaaccaacaag, tgtttgcgatgacgtgtcag cut with *Sau*3AI and R01E6 gccctgtgacgtcaattagtag, cttcttccaacaaccgattg cut with *Sau*3AI. The following three polymorphisms were recorded in the Wicks et al. (2001) database. The predicted base changes were confirmed and the following primers and restriction enzymes were used for each snip-SNP: K09E9 8901 tacgaatcacaaaattctttccc, ggttcctcgaaaaaacaacataag (*Hinc*II), F59D12 15190 gcaaaaaatattgccaataaactg, cgaaatgagaaactgctgatagtc (*Mse*I) and C02C6 26821 gagtacggataggatagcgaggag, ttcaacttatcggagtcacaacg (*Acc*II).

Candidate Gene RNAi

Double-stranded RNA was generated from a DNA template. For each gene targeted, nested primer PCR was used to amplify the region from genomic DNA. The inside primers used for the second step of amplification had T7 polymerase promoters engineered into their 5' sequence. The average length of the PCR products was 1000 base pairs and for each gene the primers were positioned so as to minimize the incorporation of intron sequence. Each dsRNA mixture was injected at a concentration of 1000 µg/mL. dsRNA was generated using the following protocol modified from the manufacture Promega. 1.5 µg of template DNA was incubated for one hour to overnight at 37°C in 100 µL of a solution containing 1X Transcription buffer (Promega), 10 mM dithiothreitol

(DTT), 35 units of RNAguard (Amersham Pharmacia), 0.5 mM of rATP, rUTP, rCTP and rGTP, and 40 – 60 units of T7 RNA polymerase (Promega). Once the transcription reaction was complete, the reaction was incubated with two units of DNase I (Amersham Pharmacia) for 15 minutes at 37°C. Sodium acetate (diethyl pyrocarbonate (DEPC) treated) was added to the reaction to a final concentration of 0.3 M. The dsRNA was purified with one phenol extraction, three chloroform extractions and then precipitated with a 2.5-fold volume of 95% ethanol. The dsRNA was resuspended in a 10 mM sodium phosphate (pH 7.5) and 1 mM EDTA solution (DEPC treated).

For each targeted gene, ~15 N2 hermaphrodites were each injected with the dsRNA solution into a single gonad. After injection, the hermaphrodites were allowed to lay eggs overnight and removed from the plate the next morning. These isolated F1 progeny were scored approximately 24 hours later for larval arrest and general appearance. For the repeated experiment of H13N06.3 and H13N06.4 knockdown, *rrf*-3 hermaphrodites were injected but otherwise experimental conditions were kept constant (Simmer et al., 2002). The following genes were targeted for RNAi. The nested primers used to generate the DNA template for RNA synthesis follow each gene. C02C6.2 cggaagaaaaagcccaaaatc, agaaaggagcttggcacaatg, taatacgactcactatagggagaccacgatcgagtgcagaaaactgtcc and taatacgactcactatagggagaccacatgaagcttgatgcaccactg. C02C6.3 tcgagaataaactaatttgcgctg, tgtggagtatgtttgagaaaatcg, taatacgactcactatagggagaccactggaaaattagccgactacgc and

taatacgactcactatagggagaccacttattggaaatgtccaaaacttgc. R03A10.4
ttagctgctgaaaccaaagcc, tccagcgtctgggatgatag,
taatacgactcactatagggagaccactcaacgtcggacaaggatttc and
taatacgactcactatagggagaccacccttcctccagcatttttgc. H30A04.1
caccagtgactttcacggctac, ttctttcctcgtcgctctcc,
taatacgactcactatagggagaccactcacacagccaactgtgactg and
taatacgactcactatagggagaccacccttcctccagcatttttgc. H13N06.6
cactcaattctacattgccgc, gacaatatccattcatggccc,
taatacgactcactatagggagaccacttcgggataagagatgggttg and
taatacgactcactatagggagaccacggtggtactgaatttgcatcg. H13N06.2
ttggtgctcctggaatagtcc, tcgatgaacttttgccatttg,
taatacgactcactatagggagaccacaaacctccctaccgtcatgc and
taatacgactcactatagggagaccaccttctgccttgttctgccag. H13N06.5
cccacgtcactttattgcctac, caccaccaccatgagaatgtc,
taatacgactcactatagggagaccacttgcctacctctcctctcctg and
taatacgactcactatagggagaccacgtccatgagaatgcgagtgc. H13N06.4
ttcgacgacaagatcaagcac, tgccaggcattatgaatcaaac,
taatacgactcactatagggagaccacttgaaggagcagatgttgatcc and
taatacgactcactatagggagaccacccacgaacattccaaataccac. H13N06.3
tctggattctcgtctgctgc, tgcgatctacctggactgttg,
taatacgactcactatagggagaccacctgcgacacaaacttttccaac and
taatacgactcactatagggagaccacttgtggacgatgaaattcgg. F39D8.2
tgaagtcgatgagctgatgg, agtttggtggaagttgagttgg,

taatacgactcactatagggagaccacatggagacggcgatttg and taatacgactcactatagggagaccacggaagttgagttggaaatccac. C33A11.4 ggccatttgtggacgatg, gcacctggtccgaacttg, taatacgactcactatagggagaccacgaaattgcgacaaccagtgc and taatacgactcactatagggagaccaccggataaagaccagtttttgggg. C33A11.2 tcgcagtgctcaaacacg, agcagcagaagcaaatcacg, taatacgactcactatagggagaccacgcagtgctcaaacacgacg and taatacgactcactatagggagaccacggtgtgactcggtgcgtaag.

Transgenic Rescue / Cosmid Preparation

Protocols used for transformation and cosmid preparation were previously published (Mello and Fire, 1995). In each transformation experiment, one of the three folowing transformation marker plasmids was used at a concentration of 50 µg/mL: the *ttx-3* promoter-driven GFP reporter (Hobert et al., 1997), the *unc-119* promoter-driven GFP reporter (gift from M. Maduro and J. Rothman) and pRF4 [*rol-6*(*su1006*)] reporter (gift from A. Fire)

The following genomic regions (~20 Kbp in length) were PCR-amplified using the Expand Long Template PCR protocol (Roche). The primers used to amplify it follow each region. C33A11 7944 – F47C8 5287 ttttgttgatctcaaaaccaccc, tccttctgatcacatttacacgc; F47C8 3108 – F47C8 20132 ttggcgttggtttcctattatttc, ttttgtcaaaactgcccaatactg; F39D8 16798 – R03A10 6140 gcttgtcggattaggttccatatc, tcgtaagaatgaaggcttcgtg; R03A10 2461 – R03A10 18882

tccgacgcttgagagctg, agttccaaattgacaagtcaggg; R03A10 15660 – R03A10 38578 tgtcttctctagatggctttgtcg, tagcaacgagacacctgaaaatg; R03A10 22639 – C46E1 5247 atcccactcaaataaattgcgag, tctcatttggcactgatgatagg; C46E1 2413 – H13N06 12557 tcacatcaaattggtctcaatcg, aacaagcacagactgaataacgg; H13N06 10519 – H13N06 17013 ttcccagtggtcaagtgagtagac, tctccaatttcttctctgacatcg; H13N06 10942 - H30A04 2734 ggtttcatgatctcaatttgattc, tcaaagttcgatgaagtagatgac; H30A04 682 – H30A04 15019 acttcacccaactctttccaaac, agtcagtggtctgcaatttgaac and C02C6 2628 – C02C6 19288 caccaacatttttttaataccgcc, aattttccactgacgctacttcac. The two primers aacagcacgacttacatcaaacc and ttgattgagaaatgacatccacc were used to detect the *lin-15*(*n765ts*) lesion (Clark et al., 1994).

<u>*gob-1*(*ca17*) Deletion PCR</u>

The following primers were used to define the *gob-1*(*ca17*) deleted region. C33A11.2 tcgcagtgctcaaacacg, agcagcagaagcaaatcacg; C33A11.4 ggccatttgtggacgatg, gcacctggtccgaacttg; F39D8.1 caacaacaaggtcaacagggtc, ttagtcgtcgttggtgtcgtc; R03A10.4 ttagctgctgaaaccaaagcc, tccagcgtctgggatgatag; H13N06.2 ttggtgctcctggaatagtcc, tcgatgaacttttttgccatttg; H13N06.3 gaaacggcacattaaacaggc, ttcatagttccatcccaatcagtg and tctggattctcgtctgctgc, tgcgatctacctggactgttg; H13N06.4 ttcgacgacaagatcaagcac, tgccaggcattatgaatcaaac; H13N06.5 cccacgtcactttattgcctac, caccaccaccatgagaatgtc; H30A04.1 caccagtgactttcacggctac, ttctttcctcgtcgctctcc;

C02C6.2 cggaagaaaaagcccaaaatc, agaaaggagcttggcacaatg and C02C6.3 tcgagaataaactaatttgcgctg, tgtggagtatgtttgagaaaatcg.

gob-1 cDNA Confirmation

The Titan One Tube RT-PCR protocol (Roche) was used to reverse-transcribe and PCR amplify *gob-1* cDNAs. The SL1 primer ggtttaattacccaagtttgag and the *gob-1* primer ttcatagttccatcccaatcagtg were used in each step of the protocol (Krause 1995). Amplification products were cloned into pCR 2.1-Topo (Invitrogen) and sequenced using the M13 Universal (cgccagggttttcccagtcacgac) and M13 Reverse (agcggataacaatttcacacagga) primers.

Detection of Early *gob-1* RNAi Phenotype

For detection of the early *gob*-1 RNAi phenotype, N2 animals were injected with *gob-1* dsRNA (2000 µg/mL). Freshly hatched animals that were laid 6, 11 and 15 hours post-injection were isolated for analysis. In a second experiment, hermaphrodites were injected with *gob*-1 dsRNA (3000 – 3500 µg/mL) and F1 progeny were collected 12 hours post-injection. For the bead feeding assay, hermaphrodites were injected with *gob-1* dsRNA (3000 – 3500 µg/mL). F1 progeny that were laid 10 to 18 hours after injection were isolated

after hatching. These larvae were allowed to feed on a mixture of beads and *E. coli* for four hours, removed and examined for the Gob phenotype.

IFB-2 (MH33) and AJM-1 (MH27) Immunohistochemistry

Young hermaphrodites were injected with *gob-1* dsRNA (1000 µg/mL) and allowed to lay eggs overnight. The next day, embryos were collected from the plate and the hermaphrodites were dissected to collect internal embryos. Staining methods were followed as published by Bossinger et al. (2004) with the following changes. Slides were subbed in a solution 0.2% gelatin, 0.02% Chrome Alum, 1 mM sodium azide and 1 mg/mL of polylysine (A. Skop and B. Meyer personal communication). The primary antibodies were used at a 1/3 dilution and the secondary antibodies (donkey anti-mouse conjugated Alexa 594 (Molecular Probes)) were used at a 1/100 dilution.

Electron Microscopy

Hermaphrodites were injected with *gob-1* dsRNA (1000 µg/mL) into a single gonad. Freshly hatched larvae that were laid 8 to 12 hours after injection were allowed to feed on a mixture of beads and *E. coli* for four hours. Larvae with an obvious gut obstruction were isolated for processing. Similarly, freshly hatched N2 larvae were allowed to feed for four hours on a beaded plate and then isolated for processing. Embedding of the *gob-1*(RNAi) mutant and N2 L1

larvae was performed at 25°C unless otherwise noted. The larvae isolated above were immersed into a solution of 0.1 M Hepes and 2.7% glutaraldehyde and had their head and tail cut off with a 25 gauge needle. The larvae were fixed overnight at 4°C. The next morning the fixed larvae were transferred to a 1% agarose pad, aligned and covered with 1% agarose. The resulting agarose block (two to three mm in thickness) was placed into a two mL eppendorf tube. The following staining and embedding protocol was performed on a rotating wheel to maximize infiltration of each step. Each agarose block was washed for one hour in 0.1 M Hepes and stained in a solution of 0.1 M Hepes and 1% Osmium tetroxide overnight. After staining, the agarose block was dehydrated with a series of hour-long ethanol washes (30%, 50%, 70% and 90% ethanol by volume). A final 100% ethanol wash was done overnight. The following infiltration steps were performed over the next three days: a 50% ethanol and 50% n-BGE (n-Butyl Glycidyl Ether) infiltration for two hours, a 100% n-BGE infiltration overnight, a 50% Quetol embedding mixture (34% Quetol 651, 53% NSA (Nonenyl Succinic Anhydride), 11% NMA (Nadic Methyl Anhydride) and 2% DMP-20 (2,4,6 – Tri (Dimethylaminomethyl) phenol) and 50% nBGE infiltration overnight and a 100% Quetol infiltration overnight. Once the infiltration was complete, the block was baked at 60°C for 24 hours. The sectioning and the EM photography of the blocks was performed by Mr. Wei Dong (EM Technician at the University of Calgary).

gob-1 Transcriptional Reporter

2323 bp of the *gob-1* promoter was used to create a *gob-1* promoter-driven GFP/*lacZ* reporter. The sequence used started at cosmid position H13N06 11302 (148 bp downstream of the upstream gene's (H13N06.2) polyadenylation signal) and included consecutive sequence until the ATG of H13N06.3b (coordinate H13N06 13625). The Hercules Enhanced polymerase (Stratagene) was used to PCR-amplify the sequence from genomic DNA. PCR primers with *Bam*HI ggtggtgggtcgacgaaaattgtaaaattgcaaatttc and *Sal*I ggtggtggggatccatttgtgattctgaaaattgcca cut sites engineered into their 5' sequence were used for amplification of the insert. This genomic sequence was inserted in frame into the pPD96.04 *Bam*HI *Sal*I restriction sites. pPD96.04 is a promoterless vector with a nuclear localization sequence, a GFP/*lacZ* coding region, an *unc-54* 3' UTR sequence and a 5' decoy sequence (www.ciwemb.edu/pages/firelab.html). The final construct was sequenced across the insertion sites using the GFP (gtgcccattaacatcaccatc) and the M13 Reverse (agcggataacaatttcacacagga) primers. Three clones were isolated that were generated from independent PCR amplification products. DNA from these three clones were pooled and injected simultaneously at a final concentration of 45 µg/mL (15 µg/mL each). Three clones were used to prevent a potential PCR generated mutation from comprising the majority of the injected pool. The transformation reporter, pRF4 [*rol-6*(*su1006*)], was co-injected at 40 µg/mL. Heritable transgenic lines were obtained (six were studied in total, two of these

were studied in detail). The transgenic larvae were permeablized and stained with a solution of M9 buffer containing 0.2% phenoxypropanol and 0.1 µg/mL of DAPI (4,6-diamindino-2-phenyl-indole). To increase DAPI penetration the mixture was briefly heated over a flame. Animals were allowed to stain for five to ten minutes and then immediately mounted on a 2% agarose pad for viewing.

β-galactosidase and *ges-1* Staining

A modified protocol from Edgar and McGhee (1986) was used. Embryos were washed from a feeding plate and dissected from gravid hermaphrodites. Embryos were transferred to a well slide containing 180 µL of M9 buffer. Twenty µL of 6% hypochlorite was added to the M9 solution and mixed until embryos and bacteria were separated. Embryos were immediately transferred to 1% bovine serum albumin (BSA) and washed in 1% BSA two more times. Embryos were washed twice with M9 and transferred onto polylysine/gelatin subbed slides with rails spaced at 22 mm with double stick tape (described above in IFB-2 (MH33) and AJM-1 (MH27) Immunohistochemistry). The slide was over-laid with a 22 mm by 50 mm coverslip and rinsed through with five washes of M9 and six washes of *ges-1* fixative (5 mM NaOH, 0.125 M phosphate buffer (36 mM KH_2PO_4 + 89 mM Na_2HPO_4), 22.5 mg/mL paraformaldehyde and 0.1% glutaraldehyde). With each wash the liquid was removed with filter paper and the new wash was applied to the other side of the slide with a pipette. Embryos were permeablized with gentle pushing and incubated on ice for three minutes. The

slide was washed with five washes of M9 and then with five washes of β-galactosidase stain (0.4 M phosphate buffer pH 7.5, 2 mM $MgCl_2$, 0.004% SDS, Redox Solution (5 mM Ferrocyandate and 5 mM Ferricyanide) and 0.05% X-Gal (5-bromo-4-chloro-3-indoyl β-galactopyranoside)) or *ges-1* stain. For *ges-1* stain recipe see Edgar and McGhee (1986). β-galactosidase staining was done overnight at 37°C and *ges-1* staining was performed at 4°C for one hour.

N-GOB-1::GST Fusion Protein

Three regions of the GOB-1 protein predicted to be highly antigenic were used to generate antibodies (described in results). The Hopp-Woods method using a 17 amino acid (AA) interval was used to predict antigenicity (Hopp and Woods, 1981). The Kyte-Doolittle method using a 17 AA interval gave similar results (Kyte and Doolittle, 1982). Pfx Platinum polymerase (Invitrogen) was used to amplify a region in the N-terminus (N-GOB-1) from the yk562f10 *gob-1* cDNA clone. The following primers, with *Bam*HI and *Xho*I digest sites incorporated into the 5' end of each sequence respectively, were used to amplify N-GOB-1: ggatccccatcgccagtcagtctatcgaag and ctcgagaacatttgcaaatgagttgttctc. The PCR fragments were cloned into the pCR-2.1 Topo vector (Invitrogen) and later cut with *Bam*HI and *Xho*I. The pGEX-5X-3 vector (Amersham) was also cut by *Bam*HI and *Xho*I and the insert was directionally cloned into the vector. The insert and insertion site were sequenced using the 5' pGEX primer ggctggcaagccacgtttggt. The vector was maintained in the JM109 bacterial strain

and introduced into the BL21 Star (DE3) bacterial strain (Invitrogen) for expression. The protein was expressed and purified using the published method in Ausubel et al. (1994). The following conditions were used for N-GOB-1::GST expression. BL21 (DE3) cultures were grown overnight. Fresh 14 mL cultures were inoculated the next day and grown to an optical density (OD) $_{600}$ reading of 0.5 to 0.7. The cultures were induced for three hours with 1 mM isopropyl-β-D-thiogalactoside (IPTG). The eluted protein was dialysed in 1X phosphate buffered solution (PBS) and protein concentration was determined using the Bradford Assay. The N-GOB-1::GST peptide was predicted to be 39.3 kDa. Cleavage of N-GOB-1::GST with Factor Xa was predicted to remove the majority of the GST fusion and to result in the N-GOB-1 (13.0 kDa) and the GST (26.3 kDa) peptides **(see Figure 24)**. Factor Xa (New England Biolabs) cleavage was performed as described in Ausubel et al., (1994).

Cross-linking of the 12 AA Peptide to BSA

The CGGKKAADENPENVR peptide (Macromolecular Resources, USA) was cross-linked to BSA using the following method. All steps were performed at 25°C. Ten mg of BSA was dissolved in 500 µL of a solution containing 0.1M sodium phosphate (pH 6.8) and 0.1% SDS. Twenty µL of this solution was saved for gel analysis. The BSA solution was stirred and a total of 250 µL of a 25 mg/mL MBS (m-maleimidobenzoyl-N-hydroxysuccinimide ester in N,N-dimethyl formamide) solution was added. The MBS solution above was added in

50 µL steps every ten minutes. The MBS/BSA solution was dialyzed overnight against a solution of 0.1 M sodium phosphate pH 6.8 and 0.1% SDS. Twenty µL of the dialysed solution was saved for gel analysis. Five mg of the 12 AA peptide was dissolved in 1000 µL of the 0.1 M sodium phosphate 0.1% SDS solution and 100 µL of 250 mM EDTA was added to this solution. The dialyzed MBS/BSA solution was added to the 12 AA peptide solution and incubated overnight. The following day, the 12 AA/MBS/BSA solution was dialyzed against the 0.1 M sodium phosphate pH 6.8 and 0.1% SDS solution. Twenty µL of this solution was saved for gel analysis. One and five µL of the following sample dilutions were run on an 8% PAGE gel to analyze the success of the linkage reaction. The BSA solution was diluted 1 in 40 and the MBS/BSA and 12 AA/MBS/BSA solutions were diluted 1 in 20 **(see Figure 25)**.

Generation of Anti-GOB-1 Antibodies

The peptide CGGKKAADENPENVR was cross-linked to keyhole limpet hemocyanin (KLH) for inoculation into two rabbits. Inoculation and generation of the peptide and linkage to KLH was done by Macromolecular Resources (Colorado, USA). Final bleeds were taken at day 98 and 154. Two rabbits were injected with the N-GOB-1::GST peptide isolated from a polyacrylamide gel and boosted every two weeks. Macromolecular Resources also did inoculation of these rabbits. Final bleeds were taken at day 98. For both the 12 AA peptide and the N-GOB-1::GST protein, bleeds were taken at two week intervals and

tested for reactivity on Western blots. The 12 AA bleeds were tested by Western for the detection of the small peptide cross-linked to BSA. The N-GOB-1 bleeds were tested for detection of the N-terminal GOB-1::GST fusion protein and the N-GOB-1::GST protein cleaved with Factor Xa. Four guinea pigs were boosted with 20 µg of N-GOB-1::GST peptide isolated from a polyacrylamide gel. Gel slices were liquified in 100 µL of PBS and emulsified with an equal volume of Freund's complete adjuvant. The guinea pigs were boosted every two weeks and terminated after 107 days (SACRC Hybridoma Facility, University of Calgary).

An N-terminal GOB-1 peptide fused to the maltose binding domain (MBD) peptide was used for antibody purification. The N-GOB-1::MBD was expressed and purified as previously published (Ausubel et al., 1994). Two primers atttcagaattcggatccatcgccagtcagtctatcgaag and ggcggcggccaagcttgctcaaacatttgcaaatgagttgttctc, with *Bam*HI and *Hind*III restriction sites engineered in respectively, were used to amplify the N-GOB-1 fragment from the yk562f10 cDNA clone. The pMal-c2 vector (New England Biolabs) was linearized with *Bam*HI and *Hind*III and the amplified insert was directionally cloned into pMal-c2 vector. The M13 Universal (cgccagggttttcccagtcacgac), the M13 Reverse (agcggataacaatttcacacagga) and the malE (ggtcgtcagactgtcgatgaagcc) primers were used to sequence the insert and the *malE* start site. Protein expression was performed as described above. The vector was transformed into BL21 (DE3), grown to an OD_{600} of 0.4 to 0.5 and induced with 0.3 mM IPTG for three hours. The purified N-GOB-1::MBD

protein was of the expected size (53.8 kDa). The Factor Xa digest of the protein resulted in two fragments of the predicted size. Both the full-length peptide and the cleaved product reacted with MBD mono-colonal antibody (New England Biolabs) in Westerns **(see Figure 26)**.

Five mg of the 12 AA peptide was used for cross-linking to a column provided by the Pierce ImmunoPure AG/AB Immobilization kit #2. Two mg of the N-GOB-1::GST protein was cross-linked to the same brand of column as per manufacturer's instructions. Terminal bleeds were individually passed through each respective column as specified by the Pierce protocol. In addition to the Pierce protocol, a modified elution protocol was used to increase the isolation of high affinity antibodies. Eight mL of a 0.1 M glycine solution (pH 2.8) was first passed through the two mL bed volume column. 465 µL fractions were collected into tubes containing 35 µL of 1 M Tris pH 8.0. The column was eluted a second time with eight mL of 0.1 M glycine solution (pH 1.8) and 400 µL fractions were collected into tubes containing 100 µL of 1 M Tris pH 8.0. An OD_{280} reading of each fraction was taken to estimate the concentration of protein in each fraction and the three fractions with the highest concentration of protein were pooled. The protein was dialysed against a 1X PBS (pH 7.0) solution and then a 1X PBS and 50% glycerol solution. Fractions collected in early experiments were run on a SDS polyacrylamide gel to confirm that proteins of the correct size to be IgG antibodies were the major component of the eluted proteins. In these experiments three bands were obvious, the strongest band being approximately the predicted size of an IgG protein (160 kDa) and the two weaker bands,

approximately the size predicted for the heavy (55 kDa) and light chain (25 kDa) of the IgG protein (Harlow and Lane 1998).

The pre-bleeds and the antibodies purified from the terminal bleeds of the N-GOB-1::GST inoculated animals were tested for affinity to the full length and Factor Xa cleaved N-GOB-1::MBD protein. The pre-bleeds and the antibodies isolated from terminal bleeds of the 12 AA peptide-boosted animals were tested for affinity to the BSA-cross-linked peptide and the 12 AA peptide alone. For the N-GOB-1::GST inoculated animals, all four guinea pigs gave a clean detection of the N-GOB-1 peptide with no cross reactivity to MBD protein. Similar results were seen with one of the rabbits. For the 12 AA peptide-inoculated animals, the antibodies isolated from both rabbits only reacted with the peptide or peptide-BSA cross-linked protein. In all Westerns a 1/1000 dilution of the primary antibody was used. For the secondary antibody, a 1/2500 dilution of alkaline phosphatase conjugated secondary antibody (Jackson Immuno Research Laboratories Inc.) or a 1/1000 dilution of a horse radish peroxidase conjugated secondary antibody (Jackson Immuno Research Laboratories Inc.) was used.

GOB-1 Over-expression Experiments

The vector pPD 95.11 was cut with *Sma*I and *Stu*I and religated to remove the *lacZ* sequence (www.ciwemb.edu/pages/firelab.html). This allowed for the generation of vector that only contained a decoy sequence. Two primers gtgttgtgtgccatttgtcg and agtttgttgaactgacgccg were used to amplify a 5831 bp

long piece of the *gob-1* gene from genomic DNA. The fragment was cloned into pCR Topo 2.1 for maintenance. Both the vector and the pCR Topo 2.1 vector were cut with *Bam*HI and *Apa*I. The insert was then directionally cloned into the modified pPD 95.11 vector. The entire coding region of *gob-1* and the 5' insertion site was sequenced with the following primers: agtttgttgaactgacgccg, gcagaggtcgtagagaaaacgag, tctggattctcgtctgctgc, tgccgattaacttacgatgtcttc, catcgagttgatcccaacagtc, ttcatagttccatcccaatcagtg, tgcgatctacctggactgttg, ctgactggcgatagtcatttgtg, gaaacggcacattaaacaggc, gccgttattcagtctgtgcttg, tgccgattaacttacgatgtcttc and agcggataacaatttcacacagga). A *ttx-3* promoter-driven transformation reporter (50 µg/mL) (Hobert et al., 1997) was injected along with the purified plasmid (50 µg/mL). Three independently derived and heritable transgenic strains (non-integrated) were obtained from two different injection experiments (i.e., six lines were studied in total). The purified antibodies from a single guinea pig, GP2, were used for the immunohistochemical staining results. The optimal conditions for GOB-1 immunohistochemical staining were determined to be a 1/100 dilution of the primary antibody with overnight incubation at 4°C and a 1/1000 dilution of an Alexa 594 labelled donkey anti-guinea pig secondary antibody (Molecular Probes) with a three hour incubation at 25°C.

Protein Alignment

The BLASTP default settings were used to find proteins with similar sequences to the full length GOB-1 protein. The following five proteins and resulting E values (the probability that the alignment would occur by chance) were identified : TPS-2/F19H8.1 (1261 AA) over a 426 AA interval had 27% identity and 47% similarity ($e = 2\ e^{-41}$) to GOB-1, TPS-1/ZK54.2 (1308 AA) over a 362 AA interval had 29% identity and 47% similarity ($e = 2\ e^{-34}$) to GOB-1, *Chlorobium tepidium* NP_662202 (416 AA) over the entire protein had 25% identity and 48% similarity ($e = 6\ e^{-24}$) to GOB-1 and *Caenorhabditis briggsae* CBG007712 (434 AA) over the entire protein had 77% identity and 85% similarity ($e = 3.9\ e^{-177}$) to GOB-1.

The PSI-BLAST default settings were used to search for similarity in more distantly related sequences. TPS-1, TPS-2, GOB-1 and NP_662202 were used to create a PSI-BLAST weighted matrix. The third reiteration found similarity to a family of hypothetical prokaryotic proteins. The top five matches are listed below with their respective E value: NP_645331 *Staphylococcus aureus subsp. aureus MW2* (Interpro predicted HAD hydrolase) $e = 5\ e^{-26}$, NP_241363 *Bacillus halodurans* (Interpro predicted HAD-hydrolase) $e = 2\ e^{-23}$, NP_693457 *Oceanobacillus iheyensis* HTE831 (Interpro predicted HAD-hydrolase) $e = 3\ e^{-09}$, NP_228460 *Thermotoga maritime* (Interpro predicted HAD hydrolase of Cof IIB family) $e = 4\ e^{-09}$ and NP_622154 *Thermoanaerobacter tengcongensis* (Interpro predicted HAD hydrolase of Cof IIB sub-family) $e = 9\ e^{-07}$. Several of these

proteins were predicted to be members of the HAD-like hydrolase subfamily IIB. This family primarily functions as sucrose phosphatases. The Cof IIB subfamily of HAD hydrolases is restricted almost exclusively to eubacteria and archeabacteria.

GOB-1 (*C. elegans*), CBG007712 (*C. briggsae*), TPS-1 (*C. elegans*) TPS-2 (*C. elegans*) and NP_662202 (*Chlorobium tepidium*) were aligned using T-Coffee software (http: //igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi and Notredame et al., 2000) and Boxshade (http: //www.ch.embnet.org/software/BOX_doc.html - authors Kay Hofmann and Michael D. Baron) was used to highlight amino acid similarities. Select representatives of HAD-like hydrolase proteins defined by Koonin and Tatusov (1994) and Aravind et al. (1998) were used as templates to define the relationship of GOB-1 to other members of the family. For each protein conserved domains I, II and III as defined by Aravind et al. (1998) were isolated and aligned using T-Coffee. Similar amino acids were highlighted using Boxshade. Proximity to one another within the vertical alignment is representative of relatedness without a root.

GOB-1::MBD

The primers ggcggcggcggatccatgaactgtgaaaaaggtagg and ggcggcggccaagcttgtcatccagagaactctgaaat (with 5' *Bam*HI and *Hind*III sites respectively) were used to amplify the full coding region of *gob-1* from a *C.*

elegans cDNA library. Herculase Enhanced polymerase (Stratagene) was used in the PCR reaction following manufacture's instruction. The resulting PCR product was isolated from a 1% agarose gel and digested with *Hind*III and *Bam*HI. The pMal-c2 (New England Biolabs) vector was also cut with *Hind*III and *Bam*HI and the PCR product was directionally cloned into the vector. The insert was sequenced with the following primers: tgcgatctacctggactgttg, M13 Universal (cgccagggttttcccagtcacgac), M13 Reverse (agcggataacaatttcacacagga) and malE (ggtcgtcagactgtcgatgaagcc). A mutation was detected within the insertion site resulting in a Glu to Gly substitution at amino acid six of GOB-1. Since this area was not well conserved in the GOB-1 protein and the substitution was conservative, this construct was used for protein production. The plasmid was transformed into BL21 (DE3) and the protein was expressed and purified similar to the methods above with the exception that protein expression was induced with 0.5 mM IPTG for three hours. Protein eluted from the amylose resin column was dialyzed in 10 mM PIPES (pH 7.0), 100 mM KCl and 50% glycerol and stored at -20°C. The concentration of the protein was measured using the Bradford Assay. The GOB-1::MBD fusion protein was predicted to have an average molecular weight of 91.8 kDa. The isolated protein and protein cut with Factor Xa were run on a SDS PAGE gel and Western analysis was performed using anti-MBD antibodies (data not shown).

pNPP Phosphatase Assay

The Protein Folding Optimization Kit (Protein Foldase TM, Geno Technology Inc.) was used in an attempt to ensure that GOB-1::MBD was in an active configuration. GOB-1::MBD was brought to a concentration of 500 µg/mL. Twenty µL of this protein solution was added to a final volume of 200 µL for each Foldase reaction solution. The final Foldase 1,2,3 and 4 reactions were dialysed against 50 mM PIPES and 10% glycerol (pH 7.0) and the entire reaction was used in the phosphatase assay. For a dialysis contamination control, elution buffer only was dialysed in parallel and tested in parallel in the phosphatase assay. Two µg of protein was used in each reaction and an equal volume (2.5 µL) of control dialysis was used.

The phosphatase reaction conditions were modified from Yeo et al. (2003). The reaction buffer used was a solution of 20 mM p-nitrophenyl phosphate (pNPP), 50 mM PIPES (pH = 5.0, 6.0, 7.0 or 8.0), 50 mM $MgCl_2$, 10% glycerol and 0.5 mM DTT. A final reaction volume of 200 µL was incubated at 30°C. Forty-five µL of each reaction was taken for each time point and quenched with 180 µL of a solution of 0.25 M NaOH and 50 mM EDTA. An optical density reading at 410 nm was taken shortly after quenching. A water/reaction solution control was used as blank and incubated with the reactions. The molar absorptivity of pNPP is 18.8 X 10^{-3} (OD_{404} pH 10.15) (McComb et al., 1979). This constant along with the predicted molecular weight of GOB-1:MBD (MW = 91805.87) was used to calculate the amount of phosphate released per mole of

protein. The catalytic constant (kcat) was calculated as the V_{max} / [Enzyme] (Dixon et al. 1979).

Phosphorylase a Phosphatase Assay

The protocol followed was as previously described (Mackintosh and Moorhead, 1999; Stubbs et al., 2001)

Malachite Green Assay for Phosphatase Activity

The malachite green assay was modified from Yeo et al. (2003) and Cogan et al. (1999). The phosphatase reactions were performed in a 250 µL solution of 50 mM bis/tris (pH 7.0), 10 mM $MgCl_2$, 10% glycerol and 0.5 mM DTT. 100 µL of this reaction was quenched by the addition of 30 µL of ammonium heptamolybdate (6 mM final) and H_2S0_4 (0.45 M final). Thirty µL of a solution of malachite green oxalate (120 µM final) and polyvinyl alcohol (0.056% final) was added to the stopped reaction and incubated at room temperature for 20 minutes. For each experiment the reactions were performed in 1.5 mL eppendorf tubes. OD_{610} readings were measured in 96 well flat bottom plates. One 100 µL aliquot of each reaction was stopped at ten minutes and a second 100 µL aliquot was stopped at ten minutes and ten seconds. For each plate reading four columns of phosphate concentration controls were included (two columns at the beginning and two columns at the end of the plate). The controls consisted of

two independently diluted series of the NaH_2PO_4 salt (0 mM, 150 mM, 225 mM, 300 mM, 375 mM, 450 mM, 525 mM and 600 mM). 12.5 µL of each concentration of salt was added to the reaction buffer to give a final reaction buffer concentration equivalent to all of the reactions. The readings from these standards were used to generate a graph of PO_4 concentration vs. OD_{610} reading. Within this phosphate concentration range, the relationship between $[PO_4]$ and OD_{610} was roughly linear. For each set of controls, a linear regression equation was calculated and used to predict the amount of free phosphate ion liberated in each reaction.

pH and $MgCl_2$ dependence

The pH for each reaction was changed using the following buffers: acetic acid/sodium acetate pH (4.0 and pH 5.0), bis-tris (pH 6.0 and 7.0), tricine (pH 8.0) and CHES (pH 9.0 and 10.0). All buffers were used at a final concentration of 50 mM except for the acetic acid/sodium acetate solution, which was used at a final concentration of 20 mM. The entire reaction mixture, lacking the protein and substrate, had its final pH measured to confirm the pH of the reaction.

Substrate specificity

For each substrate reaction, a substrate only control was included in parallel in order to detect any free phosphate present due to substrate

degradation. Some substrates, notably O-phospho-L-serine, glucose-6-phosphate and mannose-6-phosphate were prone to degradation. In order to keep the free phosphate detection levels within the range of the assay, a lower concentration of these substrates was used in the assay. Readings obtained from the substrate only readings were subtracted from those incubated with GOB-1.

Km determination

In order to determine the Km of the GOB-1 phosphatase activity, reactions were performed with a range of substrate concentrations. T6P final concentrations of 0.1 mM, 0.5 mM, 1.0 mM, 5.0 mM and 10.0 mM were used with three different protein concentrations of 0.05 µg, 0.10 µg and 0.15 µg per reaction. For each T6P and protein combination, liberation of free phosphate was detected as described above. In these experiments, an additional control of 10 mM T6P was added to calculate free phosphate liberated by T6P background degradation. The OD_{610} reading from substrate degradation was used to correct each reaction reading according to the concentration of T6P added to the reaction. Each reaction reading was blanked against protein only control for each protein concentration. The Km was calculated using a Lineweaver-Burk plot (Dixon and Webb, 1979). The inverse of substrate concentration (1/[S]) was plotted against the inverse of the velocity of each reaction (1/V). With this plotting method the y intercept of $1/V_{max}$ and x intercept of 1/Km could be used to

calculate the V_{max} and Km. The catalytic constant (kcat) was determined as the maximum velocity of the reaction divided by enzyme concentration (V_{max} / [E]).

RNAi to Trehalose Pathway / Trehalose Supplementation

dsRNAs to *tps-2*/F19H8.1 and *tps-1*/ZK54.2 were made using the L4440 RNAi feeding vectors constructed by Kamath and Ahringer (2003). The vector specific primers L4440 F (catgttctttcctgcgttatc) and L4440 R (ctgcaaggcgattaagttg) were used to amplify the PCR target. T7 polymerase was used to synthesize RNA from the template. Hermaphrodites were injected and allowed to lay eggs overnight. Twelve to 16 hours later, the hermaphrodites were removed. Plates were examined 24 hrs later for dead eggs and arrested L1 larvae. 48 hours later plates were re-examined for arrested L1 and L2/L3 larvae. Similar quantification was performed for the trehalose supplementation experiments. Standard NGM plates (Wood, 1988) were supplemented with 0.1 M final concentrations of trehalose and sucrose. Plates were seeded with the OP50 strain of bacteria and allowed to grow overnight at room temperature. Less culture was plated on the supplemented plates in an attempt to keep lawn sizes similar.

Lectin Staining Experiments

FITC conjugated lectins were purchased from Polysciences Inc. Lectin staining was performed with acetone fixation as previously described (Borgonie et al., 1994; Borgonie et al., 1997).

Intestine Rotation Experiments

The *elt-2* promoter-driven GFP reporter used was previously described (Fukushige et al., 1999). Injections and examination of the progeny were timed so that comma-staged embryos on the plate would have been laid 8½ hours or later after the mother was injected with dsRNA. An uninjected control strain was grown in parallel. Embryos were collected from each population and examined for GFP expression. Z-stacks of 15 to 20 images of each embryo were taken. The stacks were analyzed using ImageJ software to determine the relative position of each nucleus within the intestine.

elt-2 Knockdown

Hermaphrodites were injected with *elt-2* dsRNA (1000 µg/mL), allowed to lay eggs overnight and dissected 12 to 16 hours later. F1 embryos were then stained for β-galactosidase activity. Embryos were randomly chosen and categorized as to whether staining was present in the intestine or not.

ELT-2 and ELT-1 Over-expression

The heat-shock driven ELT-2 and ELT-1 integrated transgenic lines used were previously described (Fukushige et al., 1998; Gilleard and McGhee, 2001). Embryos were isolated at the two to four cell stage, incubated at 20°C and then heat shocked at 30°C. For ideal ELT-2 conditions, embryos were incubated at 20°C for 90 minutes and then heat-shocked at 30°C for 30 minutes. Ideal conditions for ELT-1 over-expression included incubation at 20°C for 60 minutes and 30°C for 40 minutes. Under both conditions, heat-shocked embryos were allowed to grow overnight at 20°C. Constructs were exposed to both conditions and were stained for either *ges-1* or β-galactosidase activity. *ges-1* activity was quantified based on the intensity and the proportion of the embryos that stained. β-galactosidase staining was quantified based on the number of nuclei staining in each embryo.

CHAPTER THREE - The Gut Obstructed Screen and Identification of *gob-1*

elt-2(*ca15*) Suppressor Screen

elt-2 is a *C. elegans* GATA transcription factor that has been shown to have a pivotal role in intestine specification and differentiation (Fukushige et al., 1998). An *elt-2* suppressor screen was developed in order to find genes that were genetically interacting with *elt-2* function. For this *elt-2* suppression experiment, the *elt-2(ca15)* null allele was balanced by the *szT1* chromosomal rearrangement (McKim et al., 1988). The *szT1* chromosome rearrangement is a reciprocal translocation between chromosomes X and I. During meiosis, the fused X^RI^R chromosome tends to segregate with the normal I chromosome, while the fused X^LI^L chromosome tends to segregate with the normal X chromosome. A double dose of the *szT1* chromosome is lethal due to a mutation present on the I^L arm of the balancer. With this balanced strain, the only F1 progeny that are expected to be viable are those that carry the parental genotype, *lon-2* males that result from the non-disjunction of the normal X chromosome and progeny that carry two copies of the full X and I chromosomes concomitant with a mutation that suppresses the homozygous *elt-2*(c*a15*) lethal phenotype. A strain was constructed so that the *elt-2(ca15)* null allele and the *dpy-5 unc-13* alleles were pseudolinked on the I chromosome. Thus a rescued *elt-2* homozygote would also display a Dumpy & Uncoordinated phenotype **(see Figure 10)**.

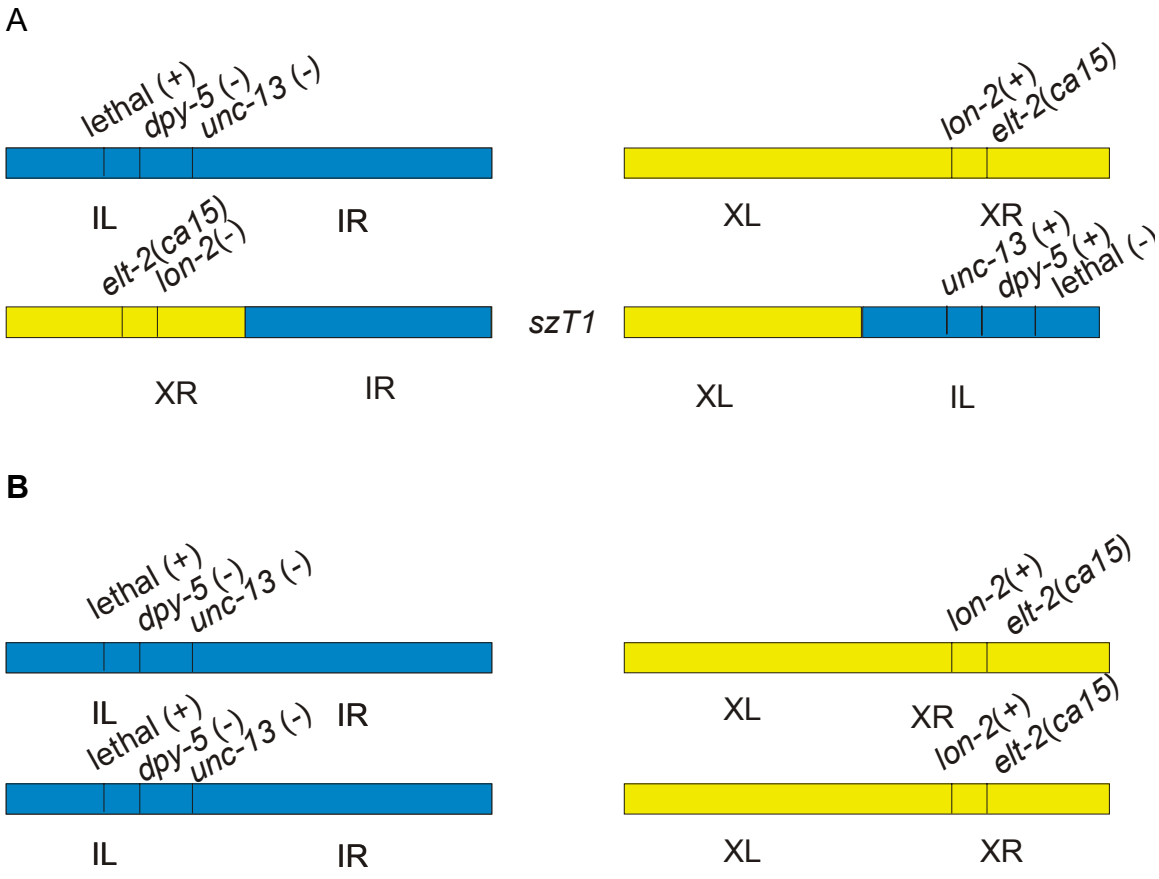


Figure 10: The predominant viable progeny expected for the suppressor screen. (A) Genotype of the *elt-2*(*ca15*) *szT1* balanced parental strain. *szT1* is a reciprocal translocation between the X and I chromosomes. XR and IL were inverted during the translocation event. A lethal mutation designated lethal (-) is carried on the IL portion of the *szT1* balancer. (B) One of the possible segregants that will only be viable if a suppressor for the *elt-2*(*ca15*) homozygous mutant is present in the genome.

L4 larvae of the *dpy-5 unc-13* /*szT1*[*lon-2*(*e678*)] I; *elt-2*(*ca15*) /*szT1* X heterozygous strain were exposed to the mutagen EMS. In total, 9192 haploid genomes were screened. Due to the average frequency of 1/2000 for EMS causing a loss of function mutation in any single gene, this number of haploid genomes was deemed to be more than adequate to ensure the mutation of any gene that could suppress the *elt-2* phenotype (Brenner, 1974; Gengyo-Ando and Mitani, 2000). During this screen, nine lines resulted in viable Dpy Unc progeny. After out-crossing and PCR detection of the wild-type *elt-2* allele in all nine of these strains, it was concluded that the phenotype was not due to a suppression of the *elt-2* null allele but likely due to *szT1* breakdown. In retrospect, the failure of this screen to isolate suppressors of the *elt-2* null allele is not surprising. ELT-2 is thought to be required for the expression of many genes (Fukushige et al., 1998). This suggests that *elt-2* is not involved in a simple linear pathway and with its loss, many gene functions may be comprimised. The vast transcriptional function of ELT-2 may make the complete rescue of the *elt-2* null phenotype and restored expression of all the genes under its influence difficult to obtain.

The Gob (Gut Obstructed) Screen

elt-2(*ca15*) homozygotes, presumed genetic nulls for the *elt-2* gene, arrest as early larvae and show severe intestine defects. It has been previously shown that these animals have a closed lumen and when fed fluorescent beads are unable to pass them into the lumen of the intestine (Fukushige et al., 1998). The

pharynges of these animals are largely normal and with continued feeding and pharyngeal pumping the beads build up in the terminal bulb and anterior intestine. We have termed this phenotype the Gut Obstructed (Gob) phenotype. The basic premise of the gut obstructed screen was that the Gob phenotype could be used to isolate other genes necessary for intestine development. Mutated genes that displayed the Gob phenotype could be the downstream effectors of *elt-2* or potentially uncharacterized genes in a redundant pathway.

Using the gut obstructed phenotype, a method was developed using fluorescent polystyrene microspheres to assay for the degree of gut malformation. As a part of the screening protocol, the microspheres were mixed into a lawn of bacteria and hatched worms were allowed to feed on the sphere/bacteria mixture. These same larvae were later removed to a bead free plate and allowed to defecate.

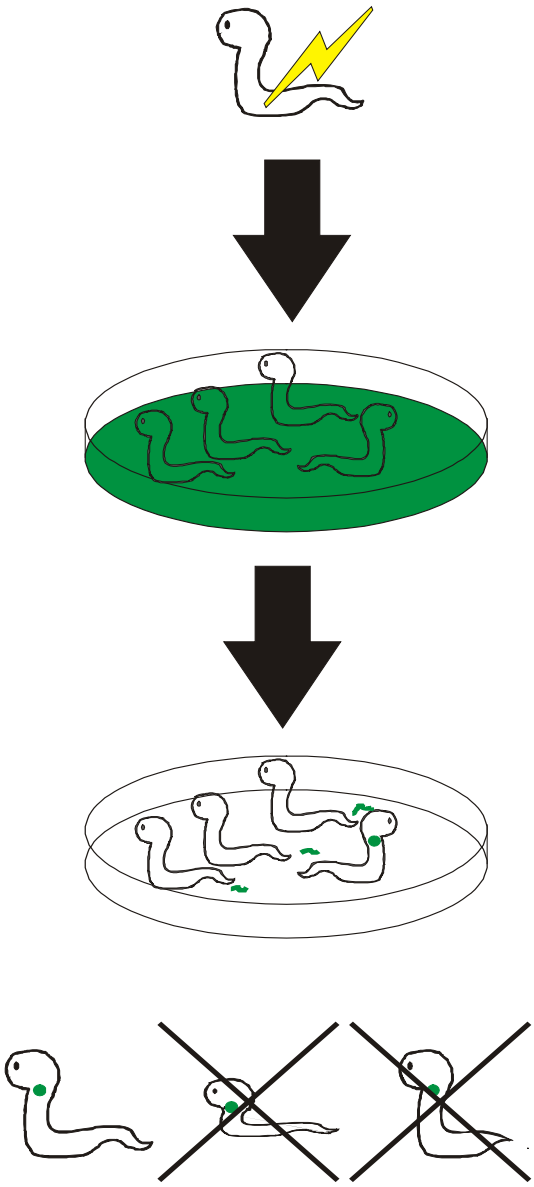
The *elt-2* Gob phenotype is readily distinguishable from other digestive tract defects examined. Several mutant strains of pharynx defective (*eat-13*, *eat-15*, *eat-17* and *pha-2*), constipated (*aex-1*), body wall muscle defective (*unc-54)*, uncoordinated (*unc-104*) and L1 lethal (*gsa-1*, *hmp-2*, *let-80*, *snb-1* and *let-23*) mutants were used as a control for specificity of the screen (Adachi and Kagawa, 2003; Avery, 1993; Korswagen et al., 1997; Otsuka et al., 1991; Rocheleau et al., 1997; Rose and Baillie, 1980; Simske and Kim, 1995; Thomas, 1990; Waterston et al., 1980). The following criteria served to increase the chances of isolating genes that had a specific function in intestine development. Animals were required to hatch and have a normal morphology outside the intestine.

Furthermore, a structurally normal and pumping pharynx would be required for feeding. All of these criteria would select against the detection of genes with ubiquitous or general house-keeping functions.

Similar to the *elt-2* suppressor screen, L4 larvae were exposed to the EMS mutagen. The F1 progeny of mutagenized worms were allowed to feed on beads overnight and were then transferred to a bead-free plate the next day. These same F1 progeny were examined for bead location at zero hours and one hour after transfer to the bead free plate **(see Figure 11)**. In total, 1758 haploid genomes were screened and of these, 68 isolates displayed a Gob-like defect. These 68 isolates were then re-examined for phenotypes using the following criteria: general morphology, stage of arrest, gross hypodermal defects, fertility, gonad morphology, general gut morphology, lumen morphology, location of obstruction, constipation, pharynx morphology and pharynx pumping. Eighteen out of the 68 original isolates were out-crossed and re-examined for the above criteria. Animals with persistent low fertility and gross morphology defects were rejected. Three strains were considered good candidates with predominantly gut specific defects. One strain, *gob-1*(*ca17*), was chosen for further study as it showed a high degree of penetrance for larval arrest and the Gob phenotype. Other gross defects in the *gob-1*(*ca17*) strain were not obvious.

The isolation of mutations that resulted in largely gut specific defects confirmed that the screen could be successful. Based upon the average EMS induced forward mutation rate of 1 in 2000 and the failure to isolate *elt-2* alleles, the screen was far from saturation and could, in the future, be used to identify

Figure 11: Summary of Gob screening strategy. L4 larvae are exposed to the EMS mutagen. F1 progeny, presumably heterozygous for induced mutations, are isolated and allowed to self-fertilize on a plate seeded with bacteria and beads. F2 progeny are allowed to hatch and feed on the bacteria/bead mixture. The entire brood is washed from the plate and plated onto a bead free plate. One hour later the progeny are examined for the Gob phenotype. Strains with sterility and severe developmental defects are rejected.



many more genes that cause the Gob phenotype (Brenner, 1974; Gengyo-Ando and Mitani, 2000)

Characterization of *gob-1*(*ca17*)

The *gob-1*(*ca17*) heterozygous mutant, isolated from the Gob screen, was out-crossed four times. This out-crossed strain was re-examined for the penetrance of lethality and the Gob phenotype. The F1 progeny (n=1806) of the *gob-1*(*ca17*) heterozygous strain had 16% of its population arrest as early larvae and 4% of its population arrest as dead eggs. Homozygous *gob-1*(*ca17*) animals could not be isolated, suggesting that the remaining 5% of expected homozygous mutants were arresting as later larvae or sterile adults. The detected arrested larvae seem to fall into two phenotypic groups. Approximately 50% of the arrested L1s had a starved appearance (i.e., the intestine cells were lighter in color and had fewer lipid droplets), dark crystal-like accumulations in the intestinal cells, and a lumen with an abnormal and stuffed appearance **(see Figure 12)**. This phenotype most closely resembled the *elt-2* null phenotype but was not as severe with regards to the degree of gut obstruction. The other 50% of the arrested larvae had a clearly open lumen but the intestine still had a starved appearance and dark crystal-like accumulations. The penetrant larval arrest and *elt-2* like phenotype of *gob-1*(*ca17*) homozygotes prompted the decision to proceed with the positional cloning and identification of the gene causing the observed defect.

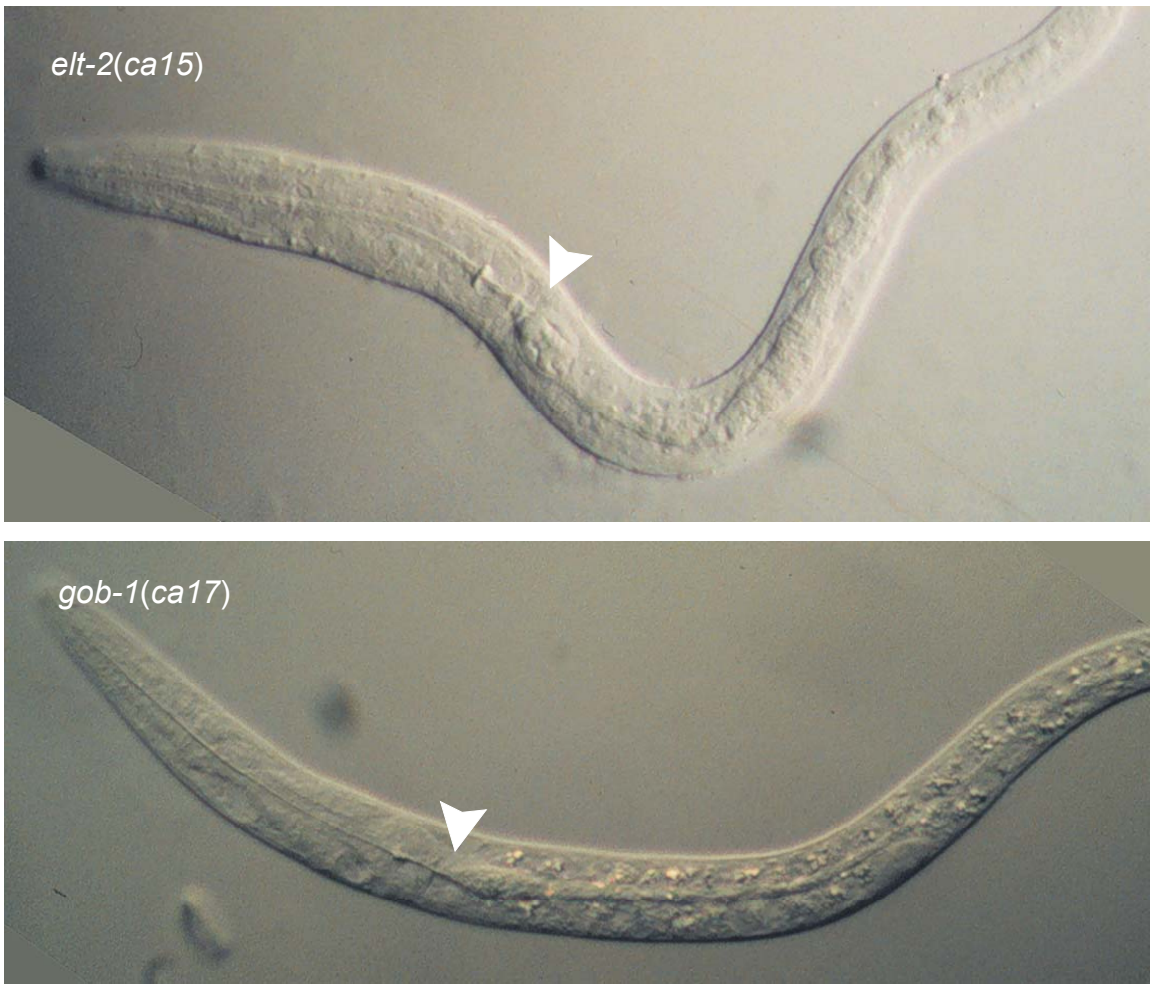


Figure 12: Differential interference contrast image of the *elt-2*(*ca15*) homozygous mutant (top) and the *gob-1*(*ca17*) homozygous mutant (bottom). *gob-1*(*ca17*) mutants arrest as early larvae. Gut obstructions (white▼) in *gob-1*(*ca17*) are obvious, although not as severe as those seen in the *elt-2* mutant. *gob-1*(*ca17*) displays an increase in birefringent/crystal like structures in the gut which is a phenotype not seen in *elt-2*(*ca15*).

Linkage of *gob-1*(*ca17*) to the X

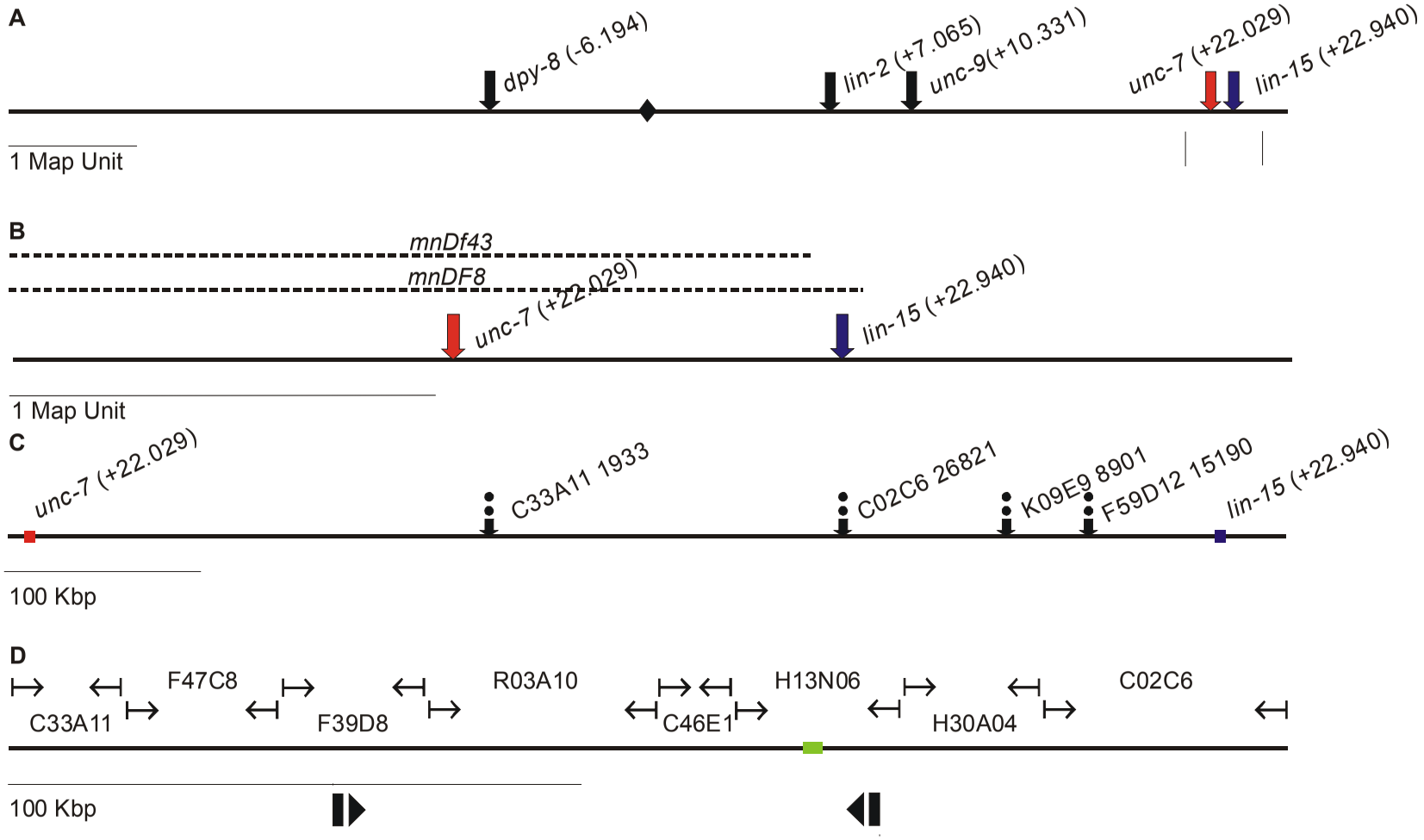
Genetic crosses were undertaken to determine the chromosomal linkage of *gob-1*(*ca17*). To test for linkage of *gob-1*(*ca17*) on the X, *lon-2* / *gob-1*(*ca17*) heterozygous animals were generated. The heterozygotes were allowed to self-fertilize and Lon-2 progeny were isolated and examined for the co-segregation of *gob-1*(*ca17*). 375 Lon-2 animals were genotyped and 179 (47.7%) were found to also carry *gob-1*(*ca17*). These results did not initially suggest a strong linkage to the X, since independent segregation would result in an expected ratio of 50%. Similar crosses were done to test *gob-1*(*ca17*) linkage to the other five chromosomes but independent segregation was also detected for these chromosomes. Additional genetic experiments were undertaken to detect *gob-1*(*ca17*) linkage on the X. In earlier experiments *gob-1*(*ca17*) was determined to be homozygous recessive for a lethal phenotype. Following this, *lon-2* / *gob-1*(*ca17*) heterozygotes were mated to wild-type males and the phenotype of the males was tested. If *gob-1* was on the X, one would expect only Lon-2 males to survive. If *gob-1* was not on the X, one would expect approximately equal ratios of Lon-2 and wild-type males. From the above mating events, 712 F1 progeny were categorized. Three hundred and ninety-five of the F1s were wild-type hermaphrodites, 0 were Lon-2 hermaphrodites (an indication of lack of self-fertilization), 42 were wild-type males and 134 were Lon-2 males. The strong tendency towards the absence of wild-type males suggested that *gob-1*(*ca17*) was probably on the X, but a fair genetic distance away from *lon-2*. The

considerable genetic distance between *gob-1*(ca17) and *lon-2* could explain the detected independent segregation between the genes in earlier experiments.

Three Factor Mapping of *gob-1*(*ca17*)

Two X-linked genetic markers, *dpy-8* and *unc-9*, were used to genetically map *gob-1*(*ca17*) **(see Figure 13 A)**. Homozygous *tra-2*(*q276*) males were used to generate males with two X chromosomes (Wood, 1988). This was done to keep *gob-1*(*ca17*) in a heterozygous state and allow for viable males and mating. A strain was generated where *gob-1*(*ca17*) was located in *trans* to a *dpy-8 unc-9* double mutant. Recombinant Dpy non-Unc and Unc non-Dpy progeny were isolated, allowed to self fertilize and tested for the presence of *gob-1*(*ca17*) based on larval arrest. 189 Dpy non-Unc animals were isolated of which 186 were *dpy-8 gob-1*(*ca17*) and three were *dpy-8 gob-1*(+) in genotype. 120 Unc non-Dpy animals were isolated and of those 8 were *unc-9 gob-1*(*ca17*) and 112 were *unc-9 gob-1*(+) in genotype. It is assumed that double crossover events are highly suppressed on the X chromosome and thus rare (Hodgkin et al., 1979). With this in consideration, these data placed *gob-1*(*ca17*) close to the left hand side of *unc-9*.

Figure 13: Positional cloning of *gob-1*. (A) A scale drawing of the X chromosome and the genetic markers *dpy-8*, *lin-2* and *unc-8* used to genetically map *gob-1*(*ca17*). (B) A scale drawing of the *mnDf43* and *mnDf8* deletions used in non-complementation experiments to define the genetic position of *gob-1*(*ca17*). (C) A scale drawing of the SNPs used to define the physical location of *gob-1*(*ca17*). (D) Scale drawing of the cosmids spanning the genomic region defined by the SNP physical mapping. Approximate deletion defining the *gob-1*(*ca17*) lesion is marked with brackets (▮▸). The green block represents the position of the *gob-1* gene coding region.


A
dpy-8 (-6.194)
lin-2 (+7.065)
unc-9(+10.331)
unc-7 (+22.029)
lin-15 (+22.940)
1 Map Unit
B
mnDf43
mnDF8
unc-7 (+22.029)
lin-15 (+22.940)
1 Map Unit
C
unc-7 (+22.029)
C33A11 1933
C02C6 26821
K09E9 8901
F59D12 15190
lin-15 (+22.940)
100 Kbp
D
C33A11
F47C8
F39D8
R03A10
C46E1
H13N06
H30A04
C02C6
100 Kbp

Deletion and Single Nucleotide Polymorphisim Physical Mapping of *gob-1*(*ca17*)

Deletion mapping and single nucleotide polymorphisms (SNPs) mapping were used to position *gob-1*(*ca17*) to the right end of the X. *tra-2* (II); *gob-1*(*ca17*) / *lin-14 unc-3* (X) males were crossed to hermaphrodites carrying deletions in this region (Wood, 1988). *mnDf43* (X), *mnDf20* (X) and *mnDf13* (X) complemented *gob-1*(*ca17*) while *mnDf8* (X) did not complement *gob-1*(*ca17*). This suggested that *gob-1*(*ca17*) was located between the break points of *mnDf43* and *mnDf8* **(Figure 13 B and Table 1)**.

The genome sequence of the Hawaiian isolate of *C. elegans*, when compared to the Bristol strain, has been found to have a high density of single nucleotide polymorphisms (SNPs) (Wicks et al., 2001). Some of these substitutions have resulted in a change in a restriction enzyme cleavage site (referred to as snip-SNPs) and can be used as convenient physical markers for the positional cloning of genes. *lin-2 gob-1*(*ca17*) / Hawaiian heterozygous animals were created and allowed to self-fertilize. Lin-2 non-Gob-1 crossover events were obtained (n = 24), isolated as homozygotes, and tested for the presence of the N2 vs. Hawaiian version of a particular SNP. Similar experiments were done with *unc-9 gob-1*(*ca17*) double mutants (n = 27). Confirmed snip-SNP markers and their estimated genetic map position were used to further define the position of *gob-1*(*ca17*). SNPs for the cosmids C0E57 (map position = 10.09), R01E6 (13.32), C33A11 (20.79), and F01G12 (23.73)

Table 1: *tra-2*(II); *gob-1*(*ca17*) / *lin-14 unc-3* (X) males were crossed to strains carrying deletions in this region. All deletions were homozygous and balanced by the translocated duplication *mnDp1* (X:V). The deletions *mnDf43* X, *mnDf20* X and *mnDf13* X complemented *gob-1*(*ca17*) while the deletion *mnDf8* did not (data shown for *mnDf43* and *mnDf8*). Two copies of the *mnDp1* duplication result in a lethal/sterile phenotype. Unbalanced deletions result in lethal defects ranging from embryonic lethality (dead eggs) to early larval arrest. With successful mating, either *gob-1* or *lin-14 unc-3* would be in *trans* to the tested deletion. *unc-3* was not complemented by either deletion so the presence of Unc progeny was used as a means of calculating mating efficiency. Wild-type F2, presumably carrying the *mnDp*1 duplication or the result of self-fertilization (SF), were allowed to self-fertilize and were genotyped based on the phenotypes of their progeny. The non-complementation by *mnDf8* does not become obvious until the F2 generation. The low ratio of *gob-1* larvae (0.04 when expected is 0.25) in the F1 progeny could be attributed to an increase in the dead egg phenotype in the *gob-1*(*ca17*) / *mnDf8* heterozygote.

The *mnDf8* deletion does not complement the *gob-1*(*ca17*) mutation while *mnDf43* does complement *gob-1*(*ca17*)

Deletion Complementation Results for SP269 (*mnDf8*) F1 Generation

	WT Adults		Unc Adults	Arrested L1 Gob	Arrested L1 non-Gob	Dead Eggs
	Observed	Corrected for Mating Efficiency				
594	429	213 (0.66)	95 (0.30)	13 (0.04)	14	43

Expected Ratios					
Complementation	0.50	0.25	0.25	SF *	UK †
Non-complementation	0.75	0.25	-	SF *	UK †

SF * = result from the self-fertilization of the *mnDp1* (X;V) / +; *mnDf8* X hermaphrodite
UK † = unknown origin

Segregation From Wild-type F1(*mnDf8*)

tra-2 carriers			non-*tra-2* carriers	Dead/ low fertility	Total Progeny
Unc-3 non-gob	1/8 gob (V+X carrier)	¼ gob (≥3 worms)			
33 (0.56)	25 (0.42)	1	52	7	118

Expected Ratios						
Complementation	0.33	0.33	0.33	SF *	UK †	-
Non-complementation	0.50	0.50	-	SF *	UK †	-

SF * = progeny of the F1s that resulted from the self-fertilization of the *mnDp1* (X;V) / +; *mnDf8* X hermaphrodite
UK † = unknown origin

Deletion Complementation Results for SP511 (*mnDf43*) F1 generation

Total Progeny	WT Adults		Unc Adults	Arrested L1 Gob	Arrested L1 non-Gob	Dead Eggs
	Observed	Corrected for Mating Efficiency				
417	319	135 (0.77)	40[67] (0.23)	0 (0.0)	8	23

Expected Ratios

Complementation	0.50	0.25	0.25	SF *	UK †
Non-complementation	0.75	0.25	-	SF *	UK †

SF * = result from the self-fertilization of the *mnDp1* (X;V) / +; *mnDf43* X hermaphrodite
UK † = unknown origin

Segregation From Wild-type F1(*mnDf43*)

tra-2 carriers			non-*tra-2* carriers	Dead/ low fertility	Total Progeny
Unc-3 non-gob	1/8 gob (V+X carrier)	¼ gob (≥3 worms)			
34 (0.60)	18 (0.31)	5 (0.09)	35	17	109

Expected Ratios

Complementation	0.33	0.33	0.33	SF *	UK †	-
Non-complementation	0.50	0.50	-	SF *	UK †	-

SF * = progeny of the F1s that resulted from the self-fertilization of the *mnDp1* (X;V) / +; *mnDf43* X hermaphrodite
UK † = unknown origin

were used to determine that *gob-1*(*ca17*) mapped close to cosmid C33A11 **(see Table 2)**. Two additional genetic markers, *lin-15* and *unc-7,* located in the vicinity of C33A11 were used to further map *gob-1*(*ca17*). Similar to above, the *lin-15 gob-1*(*ca17*) and *unc-7 gob-1*(*ca17*) double mutants were placed in *trans* to the Hawaiian chromosome and crossover events were tested for the presence of the Hawaiian or the N2 version of a particular SNP. Polymorphisms in cosmid K09E9, F59D12, C02C6 and C33A11 were used to determine that *gob-1*(*ca17*) mapped between C02C6 and C33A11 **(Figure 13 C and Table 2)**.

The physical position of *gob-1*(*ca17*) determined by the physical markers came as an initial surprise since the three factor genetic cross predicted that *gob-1*(*ca17*) was close to *unc-9*. Previous reports have suggested that double recombination events are largely suppressed across the X chromosome (Hodgkin et al., 1979). One way to reconcile the polymorphism data would be to assume the rare crossover species detected, *unc-9 gob-1*(*ca17*) and *dpy-8 gob-1*(+), were due to double crossover events. Perhaps the ends of the X chromosome are more susceptible than the bulk of the chromosome to double recombination events. Using closer genetic markers could prevent future errors.

Transgenic Rescue of *gob-1*(*ca17*)

The viability of *gob-1*(*ca17*) / + heterozygotes defined *gob-1*(*ca17*) as a recessive lethal mutation. This suggested that the lethality of the *gob-1*(*ca17*) homozygote could be rescued by the presence of a wild-type copy of the gene.

Table 2: *lin-2*, *unc-9*, *unc-7* or *lin-15* genetic markers were used to create double mutants with *gob-1*(*ca17*). The doubles mutants were placed in *trans* to the Hawaiian chromosome. Recombination events between the genetic marker and *gob-1*(*ca17*) were isolated and homozygotes of the recombinant chromosomes were tested for the presence of the Hawaiian or N2 version of the polymorphism. *lin-2*, *unc-9* and *unc-7* are to the left of *gob-1*(*ca17*) and *lin-15* are located to the right. Recombination events detected between the tested SNP and *gob-1* (*ca17*) and total recombinants tested for each genetic marker are presented in the table.

Recombination events detected between Single Nucleotide Polymorphisms and *gob-1*(ca17)

Genetic Marker (Genetic Location *)	Cosmid Location of SNP	Recombination Events Detected Between the SNP and *gob-1*(*ca17*) (Total number of recombinants screened) NA (Not Available)
lin-2 (7.065)	C0E57	21 (24)
lin-2 (7.065)	R01E6	14 (24)
lin-2 (7.065)	F01G12	4 (24)
lin-2 (7.065)	C33A11	0 (24)
unc-9 (10.331)	C0E57	27 (27)
unc-9 (10.331)	R01E6	17 (27)
unc-9 (10.331)	F01G12	4 (27)
unc-9 (10.331)	C33A11	0 (27)
unc-7 (22.029)	C33A11	1 (19)
lin-15 (22.940)	F59D12	11 (16)
lin-15 (22.940)	K09E9	7(16)
lin-15 (22.940)	C02C6	1 (16)

* genetic location refers to the position of the gene, in map units, to the right of centre of the X chromosome

Segments of wild-type dsDNA were injected into the balanced *mnDp1* (X;V); *gob-1*(*ca17*) *lin-15* X strain in an attempt to isolate *lin-15 gob-1*(*ca17*) rescued animals. *mnDp1* has been defined as a duplication of a part of the X chromosome translocated to the V chromosome and complements *lin-15* and *gob-1*(*ca17*) mutants (Wood 1988). This duplication is lethal in the homozygous state and is segregated at the frequency of the V chromosome during meiosis. Cosmids ZK662, C03H12, ZK662/C03H12, H08J19/F59D12, and K09E9/K09A9 were injected but did not yield *gob-1*(*ca17*) rescue **(see Table 3)**. Ten overlapping PCR products spanning the mapped interval, each approximately 20 Kbp in size, were also used in similar rescue experiments. However, they too were not able to rescue the *gob-1*(*ca17*) homozygous mutant **(see Table 3)**. Potential rescued candidates were tested for the presence of the tightly linked *lin-15*(*n765ts*) allele but all candidates were found to have a copy of the wild-type *lin-15* allele and therefore predicted to also have a wild-type copy of the *gob-1* gene.

Candidate RNAi

With the inability to obtain transgenic rescue of the *gob-1*(*ca17*) mutant, a second strategy was undertaken to identify the gene responsible for the *gob-1*(*ca17*) phenotype. Predicted genes identified within the mapped interval were targeted for dsRNA degradation. With this approach, it was anticipated that the Gob phenotype could be assigned to one of these genes. The first genes to be

Table 3: The *mnDp1* translocated duplication is lost at the frequency of normal chromosome segregation. It is expected that in the *mnDp1* (X;V) / +; *gob-1* (*ca17*) *lin-15* X parental strain, 25% of the F1 progeny will arrest as *gob-1*(*ca17*) *lin-15* homozygous larvae. Transgenic rescue of the *gob-1*(*ca17*) allele would potentially allow for the *gob-1 lin-15* homozygotes to reach early adulthood and be detected by the Lin-15 multi-vulva phenotype.

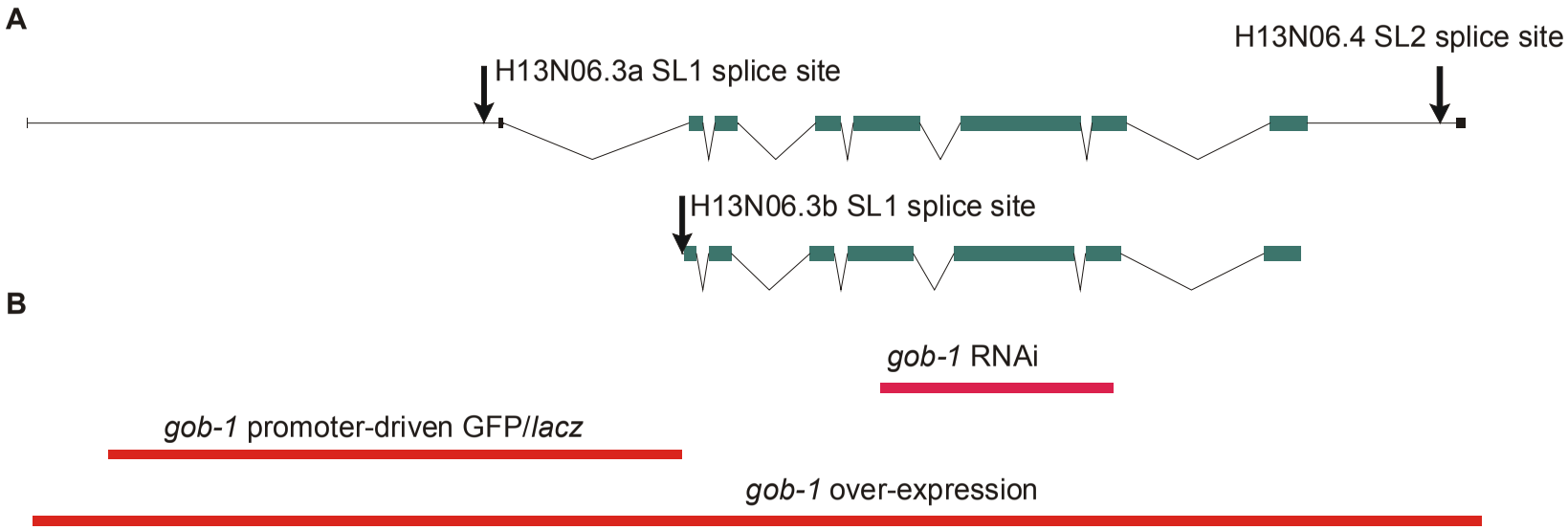
DNA fragments injected into *mnDp1* (X;V) / +; *gob-1 lin-15* X during transgenic rescue experiments of *gob-1*(*ca17*)

Area Targeted	# of injected animals (Concentration of injected DNA µg/mL)	Total # of Transgenics (# of Genomic Carriers)
PCR fragments		
C33A11 7944 – F47C8 5287	15 (2)	NA (F1s immediately screened for rescue)
F47C8 3108 – F47C8 20132	15 (2)	NA
F39D8 16798 – R03A10 6140	16 (2)	NA
R03A10 2461 – R03A10 18882	16 (2)	NA
R03A10 15660 – R03A10 38578	12 (10)	NA
R03A10 22639 – C46E1 5247	13 (10)	NA
C46E1 2413 – H13N06 12557	15 (50) 12 (10) 12 (2) 15 (0.4)	23 (4) 27 (4) 52 (10) 42 (1)
H13N06 10519 – H13N06 17013	11 (20) 15 (10) 15 (5) 15 (2.5)	53 (1) 48 (3) 111 (2) 115(2)
H13N06 10942 - H30A04 2734	15 (50) 11 (10) 13 (2) 12 (0.4)	13 (0) 10 (0) 64 (7) 24 (1)
H30A04 682 – H30A04 15019	7 (10) 13 (2) 12 (0.4)	6 (0) 40 (6) 7 (2)
C02C6 2628 – C02C6 19288	15 (2)	NA
Cosmids		
ZK662	18 (50)	13 (10)
C03H12	9 (50)	8 (7)
ZK662/ C03H12	17 (50)	14 (15)
H08J19 / F59D12	16 (50)	14 (14)
K09E9/K09A9	14 (50)	14 (4)

targeted for degradation were those with expression tags. The following predicted open reading frames were targeted for degradation: C02C6.2, C02C6.3, R03A10.4, H30A04.1, H13N06.6, H13N06.2, H13N06.5, H13N06.4, H13N06.3, F39D8.2, C33A11.4 and C33A11.2. RNAi to C02C6.2, H13N06.2 and H13N06.5 gave a penetrant dead egg and larval arrest phenotype. RNAi to H13N06.4 and C33A11.4 resulted in low fertility and small brood sizes. Only RNAi to H13N06.3 gave both a strong larval arrest and the Gob phenotype. H13N06.3 has now been re-named *gob-1*.

H13N06.3 is predicted to be upstream of H13N06.4 as a part of a two gene operon (Blumenthal et al., 2002). Since RNAi to H13N06.3 could potentially also knockdown H13N06.4 expression, both genes were targeted with 1000 µg/mL of dsRNA in an *rrf-3* RNAi sensitive background (Simmer et al., 2002). The aim was to determine if H13N06.4 could also be responsible for the Gob phenotype. H13N06.3 dsRNA, which targeted 738 bp of H13N06.3 cDNA sequence, was synthesized **(see Figure 14 B)**. Of the H13N06.3 RNAi mutant F1 progeny (n = 45), 40% were sterile or fertile adults, 13% were arrested L1 larvae and 47% were arrested L2/L3 larvae. H13N06.4 dsRNA that targeted 643 bps of H13N06.4 cDNA sequence was synthesized and of the F1s analyzed (n = 45) 91% were fertile or sterile adults, 4% were arrested L1 larvae and 4% were arrested L2/L3 larvae. Overall, RNAi to H13N06.4 did not produce a Gob-like phenotype or a penetrant early larval arrest although gonad migration defects were obvious in ten of the sterile adults. These results were consistent with the loss of H13N06.4 function not being responsible for a major portion of

Figure 14: Cartoon of the *gob-1* gene (drawn to scale). (A) A representation of the *gob-1* coding region as determined by cDNA sequence. H13N06.3a is longer than H13N06.3b by an extra 5' exon (eight extra amino acids). The SL1/SL2 splice sites for H13N06.3a, H13N06.3b and H13N06.4 are marked. (B) A scale representation of the genomic regions (red lines) that were used for the *gob-1* RNAi injection, the *gob-1* transcriptional reporter and the GOB-1 over-expression construct.


A
H13N06.4 SL2 splice site
H13N06.3a SL1 splice site
H13N06.3b SL1 splice site
B
gob-1 RNAi
gob-1 promoter-driven GFP/lacz
gob-1 over-expression

the H13N06.3 defect.

gob-1(*ca17*) is a Deletion of Approximately 15 Open Reading Frames

Once H13N06.3 was identified, an attempt was made to sequence this gene in the *gob-1*(*ca17*) strain and thus identify a potential loss of function lesion. Attempts to PCR-amplify the H13N06.3 region in *gob-1*(*ca17*) homozygotes failed despite the fact that wild-type PCR controls were able to amplify the targeted sequence and the *lin-15* wild-type allele could be detected in similarly prepared *gob-1*(*ca17*) homozygote samples. In an attempt to resolve this amplification problem, PCR targets surrounding the H13N06.3 gene were also targeted for amplification. PCR amplification could not be obtained from the *gob-1*(*ca17*) mutant using oligonucleotides lying between the genes F39D8.1 and H13N06.5 (21 attempts using 8 different pairs of oligonucleotides) despite amplification being successful outside this region. In addition to these PCR results, a large recombination free area between cosmid C33A11 and C02C6 was detected during the positional cloning of *gob-1*(*ca17*). These two results converge to suggest that *gob-1*(*ca17*) is, in fact, a deletion of at least 15 predicted open reading frames **(see Figure 13 D)**. In retrospect, this deletion is the reason why transgenic rescue could not be obtained. Several of the genes predicted to be deleted in the *gob-1*(*ca17*) lesion resulted in lethality during the candidate RNAi screen. Rescuing all of these lethal phenotypes would likely require larger fragments of exogenous DNA than what was used.

gob-1(*ca17*) cDNA Confirmation

gob-1 cDNAs were sequenced in order to confirm the predicted exons displayed in the WormBase/Acedb (Stein et al., 2001; Walsh et al., 1998). Two reported cDNAs yk452a9 and yk562f10 were sequenced and were able to confirm the predicted 3' sequence (Reboul et al., 2001). To obtain additional 5' cDNA sequence missing in these clones, single tube RT PCR and cDNA amplification was performed on a mixed stage worm cDNA library. H13N06.3 was predicted to be in an operon and thus, like many *C. elegans* transcripts, was likely to be trans-spliced with an SL1 splice leader sequence (Blumental et al., 2002; Wood, 1988). For cDNA amplification, an SL1 sequence-specific oligonucleotide and an internal *gob-1* cDNA oligonucleotide were used. The resulting amplification products were cloned and sequenced. From these sequences, the predicted internal *gob-1* splice sites in WormBase were confirmed. Two alternatively spliced products were confirmed but the longer product H13N06.3a was missing the first predicted exon **(see Figure 14 A)**. H13N06.3a has H13N06 12787 as the cosmid position of its first ATG and is thus predicted to give rise to a 439 AA protein. H13N06.3b has its starting ATG at the H13N06 13625 cosmid position and is predicted to give rise to a 431 AA protein. H13N06.3b differs from H13N06.3a only by missing the first eight N-terminal amino acids. Both share the same stop codon at cosmid position H13N06 16120 – 16122 and polyadenylation splice site at H13N06 16662. The polyadenylation

splice is further down stream than predicted by WormBase/Acedb (Stein et al., 2001; Walsh et al., 1998).

In conclusion, the gut obstructed screen is able to identify mutations that cause gut-specific defects. This screen could be used in the future to identify other genes required for intestine function. *gob-1*(*ca17*) was identified in this screen and is positioned on the X. *gob-1*(*ca17*) is a deletion of ~15 open reading frames but, of the genes in this region, only RNAi to H13N06.3 phenocopies the larval arrest and the intestine defect. The coding region of H13N06.3 is confirmed and has been renamed *gob-1*.

CHAPTER FOUR - Characterization of the *gob-1* Loss of Function Phenotype

gob-1 Loss of Function Defect

gob-1 targeted RNAi was used to define the *gob-1* loss of function phenotype. The *gob-1*(*ca17*) strain was no longer used to define the *gob-1* mutant phenotype in order to avoid additional phenotypes caused by the deletion of other genes. The RNAi sensitive strain *rrf-3* was injected with 1000 µg/mL of *gob-1* dsRNA (Simmer et al., 2002) and the F1 (n = 45) progeny were studied for larval arrest **(see Figure 14 B)**. Thirteen percent of the F1 animals arrested as L1 larvae, 47% arrested as L2 or L3 larvae and 11% arrested as L4s or sterile adults. These *gob-1* loss of function experiments confirmed that *gob-1* is necessary for survival of the animal past the early larval stages.

gob-1 loss of function was predicted to have an early defect on intestine formation. In order to address this, wild-type animals were examined for the degree of intestine malformation in the *gob-1* loss of function mutant. Hermaphrodites were injected with 2000 µg/mL of *gob-1* dsRNA. Freshly hatched mutant F1 animals were isolated (n = 166) and were categorized with respect to the degree of intestinal defect using Nomarski optics. Forty-six percent had an open lumen, 23% had a closed lumen and 31% had an intermediate phenotype **(see Figure 15)**. Higher concentrations of dsRNA were able to increase the fraction of F1s with the closed lumen phenotype. With the

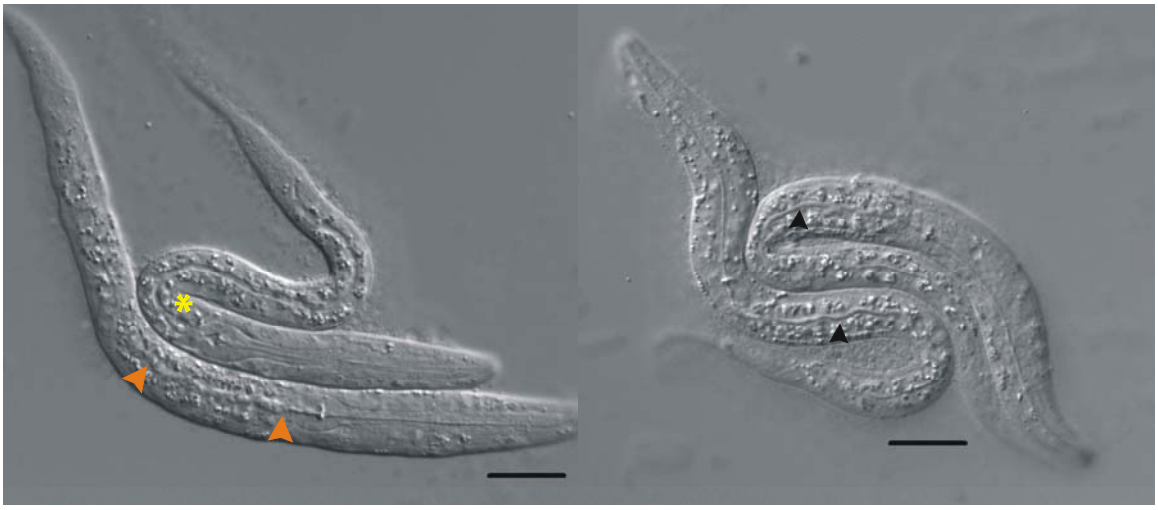


Figure 15: The Gob phenotype can be detected early in the development of the *gob*-1(RNAi) mutant. Hermaphrodites were injected with *gob-1* dsRNA (2000 µg/mL). F1 progeny were isolated within 15 minutes of hatching. Examples of the three most common phenotypes are shown above. A complete lumen closure is seen in the left picture (lumenal area defined by the orange arrowheads). Two examples of partial lumen closure are seen on the right. Areas of the lumen that appear to be partially opened are marked with black arrowheads. A *gob-1* mutant animal with an open lumen is seen in the left picture. The open lumen is marked with a yellow asterisk. All mutants have presumed lipid particles in their head region and higher than normal levels of birefringent granules. Scale bar 20 µm.

injection of 3000 to 3500 µg/mL of dsRNA, 9% of F1 progeny (n = 54) had an open lumen, 56% had a closed lumen and 35% had an intermediate phenotype. The bead feeding assay was used to determine if defects detected under Nomarski were actually resulting in a gut obstruction. In a separate but similar experiment to the above (using 3000 to 3500 µg/mL of dsRNA), freshly hatched F1 animals (n = 139) were allowed to feed on a mixture of beads and *E. coli* for four hours and were then examined for the location of the beads. Twenty-four percent of the F1 progeny had beads located only in the anterior digestive tract, 6% had beads located in the anterior to mid digestive tract and 57% had beads located throughout the entire digestive tract. 14% could not be categorized because they were too sick to feed during the allotted time period **(Figure 16)**. The penetrant localization of the beads only in the anterior digestive tract confirmed the findings that a strong lumen malformation was present at hatching. From these results, it was concluded that the gut obstruction observed in the *gob-1* loss of function mutant is likely present during embryo development and is not the result of subsequent degeneration once feeding has commenced.

Aside from the gut obstruction, other defects were detected in the *gob-1* RNAi animals. All of the F1 animals examined had dark crystal-like structures in their intestine. RNAi F1 animals also had lipid droplets located in the head prior to hatching. The number of these lipid droplets in their head decreased after the animals hatched and attempted to feed. There was also a penetrant presence of large vacuoles in the head. Such vacuoles have at least two possible origins. Head vacuoles are often seen in ageing animals and are a sign of necrosis and

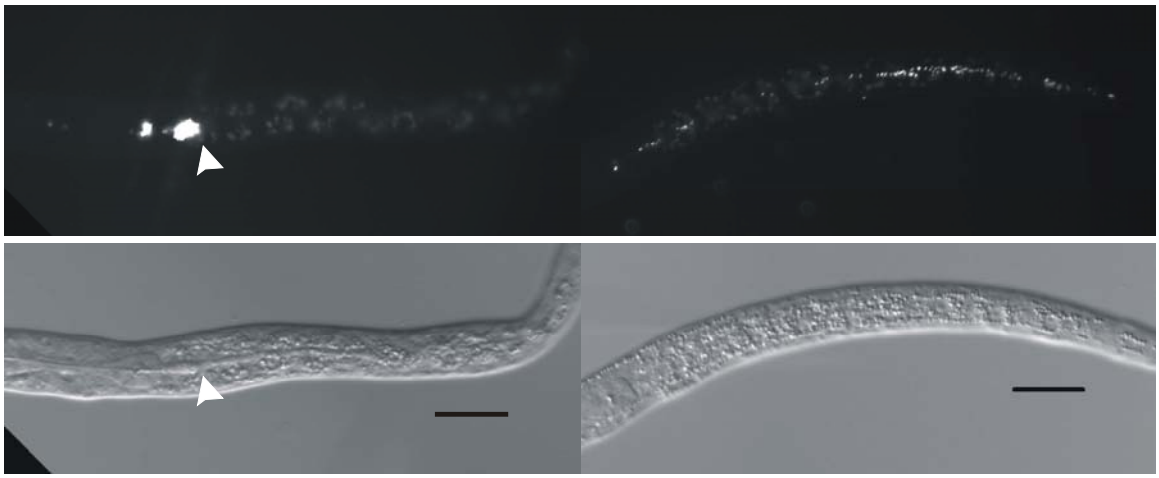


Figure 16: Freshly hatched *gob-1*(RNAi) mutants display a strong gut obstruction. Freshly hatched *gob-1*(RNAi) mutants were allowed to feed on a mixture of beads and bacteria for four hours. The location of the beads (top) and the Nomarski phenotype (bottom) are shown for two different worms. An example of the complete lumen closure phenotype (point of obstruction indicated with white arrowhead) can be seen on the left. An example of an open lumen and beads located throughout the digestive tract can be seen on the right. Scale bar 20 μm.

death (Herndon et al., 2002). Vacuoles in such a position have also been an indication of excretory valve failure (Fujita et al., 2003). Neither the lipid droplets nor the head vacuoles were studied in detail since both could be indicators of starvation and impending death.

AJM-1 and IFB-2 Distribution Patterns in *gob-1* Mutants

One of the potential sources for the *gob-1*(RNAi) lumen defect could be a defect in the establishment or maintenance of epithelial polarity. Two markers, AJM-1 and IFB-2, were used to assay for epithelial integrity. AJM-1 is a novel coil-coil protein that forms a core component of the adherens junction-like structure of the worm (Segbert et al., 2004). Just prior to the bean stage (~240 minutes PFD), a punctate pattern forms along the midline of the intestine. As elongation proceeds, the AJM-1 pattern refines into a ladder-like pattern and concentrates at apical foci where two cells are in contact. IFB-2 is an intermediate filament that is specifically localized to the apical cortex of the intestinal cell and eventually becomes a component of the intestinal terminal web underlying the brush border (Bossinger et al., 2004). In the bean staged embryo (~280 minutes PFD), IBF-2 is predominantly cytoplasmic but begins to accumulate in a punctuate pattern at the apical surface of the intestinal cells. As elongation proceeds, IBF-2 becomes less apparent in the cytoplasm and more defined as an apical structure. Disruption of either AJM-1 or IFB-2 patterns has been observed in mutants defective in apical junction formation and maintenance

of epithelial integrity (Bossinger et al., 2004; Segbert et al., 2004). Worms were subjected to *gob-1* RNAi and F1 embryos were processed for IFB-2 and AJM-1 protein detection. Similarly staged wild-type embryos were processed and the localization of the AJM-1 and IFB-2 proteins were compared to the localization patterns of these proteins in the *gob-1*(RNAi) mutants. Overall, no disruption in AJM-1 protein localization was detected in the *gob-1* mutant at any stage (data not shown). This finding suggests that adherens junctions are intact in the *gob-1* mutant. IFB-2 protein localization was for the most part normal, although subtle defects could be detected. When compared to wild-type, the initial localization of IFB-2 in the gob-1 mutant was slightly disorganized. In later embryos, IFB-2 localization was comparatively uneven and patchy in the mutant **(see Figure 17)**. Due to the dynamic nature of IFB-2 localization during development, the true extent of mislocalization was difficult to quantify. Overall, the subtle defect seen could account for a breakdown of the terminal web-like structure of the intestine and could explain part of the lumenal phenotype seen in *gob-1* mutant embryos. The fact that IFB-2 is still localized to the apical surface suggests that apical/basal sorting within the mutant intestine is largely normal.

gob-1 Mutant Ultra-structural Defects

To elucidate the nature of the *gob-1* loss of function defect at the cellular level, electron microscopy photographs were taken of cross-sections of the mutant larvae. Hermaphrodites were injected with 1000 µg/mL of *gob-1* dsRNA

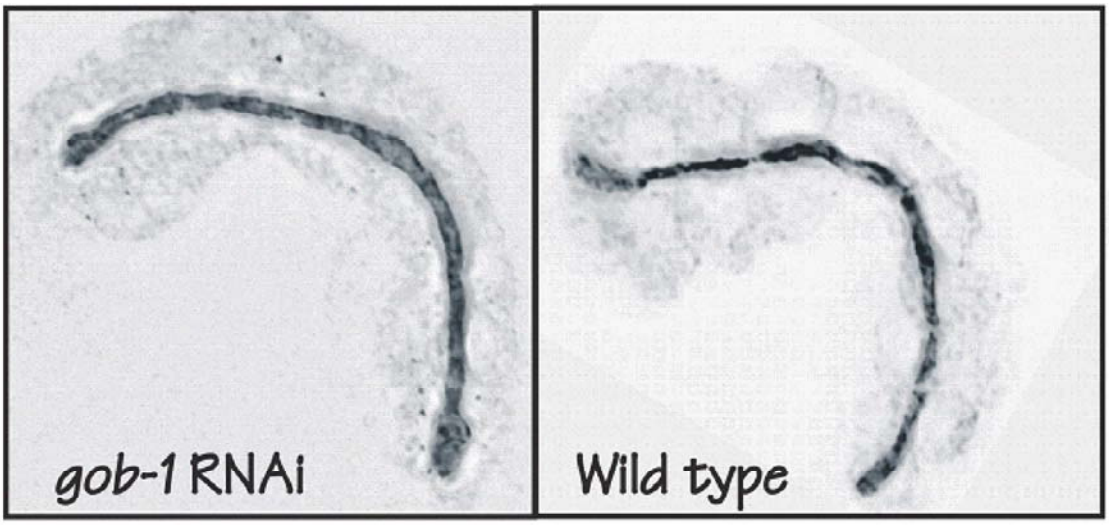


Figure 17: IFB-2 immunohistochemical staining of the *gob-1*(RNAi) mutant. In the pictures above, very little difference can be detected in protein localization pattern between the wild-type and *gob-1*(RNAi) embryos. In other samples, subtle differences in protein localization can be detected.

and F1 eggs were collected at 8 to 12 hours after injection. The collected F1s were allowed to hatch and feed on a mixture of beads and bacteria. L1 larvae with obvious gut obstructions were isolated and embedded for sectioning. Eight *gob-1* RNAi L1 larvae and nine wild-type L1 larvae were embedded and cross-sectioned. Each sample had cross-sections collected at 10 to 15 µm intervals along the length of the worm. Each of these cross-sections, five cross-sections for each condition, had EM photographs taken. Major defects, such as the complete malformation of lumen structure seen in *elt-2* mutants (Fukushige et al., 1998), were not detected in *gob-1* RNAi mutants **(see Figure 18)**. This defect was not seen despite the isolation of larvae with obvious gut obstructions. This suggests that the obstructions seen in *gob-1* mutants are transient along the length of the lumen and not as severe as initially predicted by Nomarski optics. Lumen size seemed to be normal in the *gob-1* RNAi mutant. Villus length was on average shorter and the villi were less dense in the mutant **(see Table 4)**. The electron dense endotube (a stress-bearing network located just under the microvilli) frequently appeared to be thick and irregular in the *gob-1* mutant **(see Figure 19)**. For each cross-section the endotube defect was quantified based on the degree of the defect. In the wild-type cross-sections, only 1/23 had a moderate endotube defect, while 22/23 had a thin, continuous and well-defined electron dense structure. In the *gob-1* RNAi cross-sections, 10/26 had an endotube structure that was irregular in thickness and continuity, 8/26 had a moderate endotube defect and 8/26 had an endotube that was similar to wild-type. IFB-2 is a component of the endotube, and this endotube phenotype may

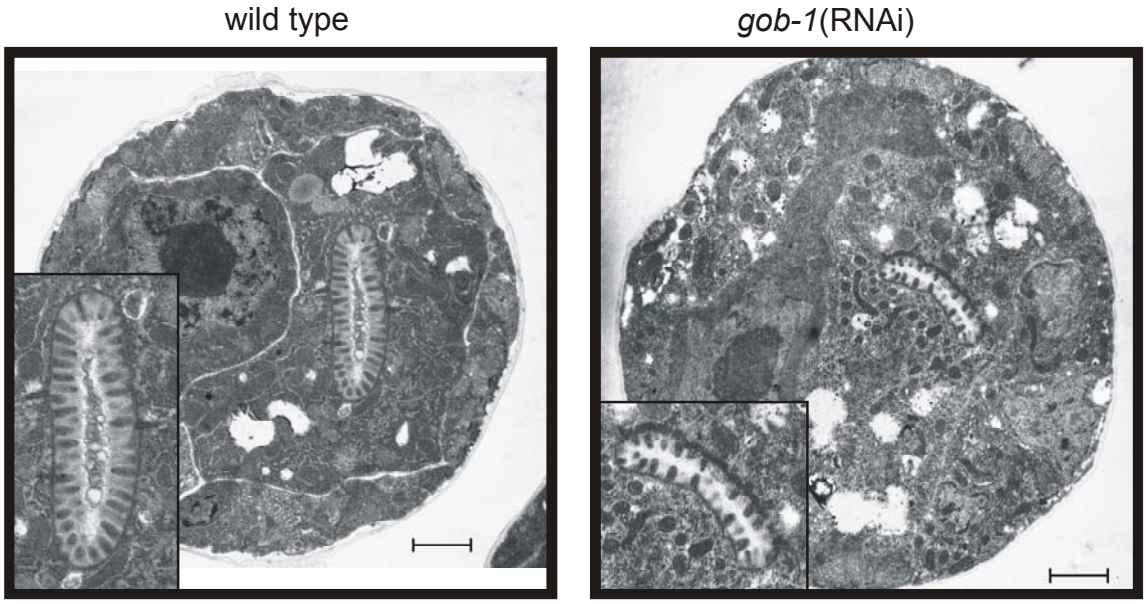


Figure 18: The *gob-1*(RNAi) mutant does not display a complete loss of lumenal structures. Electron microscopy cross-sections, typical of those seen for the wild-type and *gob-1* mutant larvae, are shown above. Both are cross-sections through the gonad and are at similar locations along the anterior/posterior axis of each worm. When the two conditions are compared, no consistent defects are seen outside the lumen. In the *gob-1* mutant, the lumen (right inset) appears to be open. This suggests that the gut obstructed phenotype seen in the mutant may be due to transient closures or surface adhesions present along the length of the lumen. Scale bar 10.58 nm.

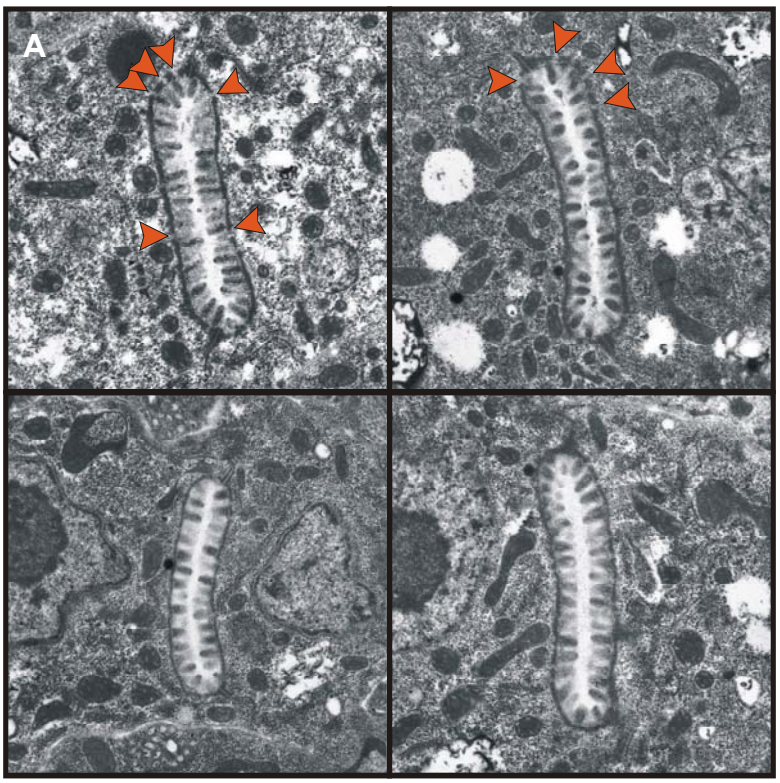
***gob-1*(RNAi) mutant microvilli are shorter and less dense than the microvilli present in wild-type larvae**

	N2 non-gonad	N2 gonad	*gob-1* RNAi non-gonad	*gob-1* RNAi gonad
Lumenal Circumference / Worm Circumference	0.27 (0.03)	0.27 (0.02)	0.25 (0.05)	0.22 (0.04)
Lumenal Circumference / Intestinal Circumference	0.35 (0.04)	0.35 (0.02)	0.32 (0.05)	0.29 (0.03)
Villi Length	0.032 nm (0.007)	0.027 nm (0.004)	0.022 nm (0.004)	0.021 nm (0.003)
Villi Density (Villi # / Lumen circumference)	61.50 villi/nm (4.56)	63.70 villi/nm (2.02)	56.10 villi/nm (5.03)	53.95 villi/nm (8.02)

Table 4: For the *gob-1*(RNAi) mutant and N2 larvae, lumenal circumference, microvilli length and microvilli density were compared. For N2 larvae, 5 cross-sections through the gonad and 18 cross-sections outside the gonad were compared to 4 gonad region cross-sections and 22 non-gonad cross-sections of *gob-1*(RNAi) larvae. There appears to be little difference between the lumen circumferences of the wild-type larvae vs. the mutant larvae. In the *gob-1* RNAi larvae, both the length and the density of the microvilli are lower than wild-type. For each calculation the average measurement is displayed with the standard deviation in brackets.

Figure 19: *gob-1*(RNAi) mutants may have defects in their endotube. (A) Selected cross-sections of four different *gob-1*(RNAi) mutants. The top two panels are examples of severe endotube defects. The endotube is the electron dense structure located below the lumen surface. Gaps in the structure are pointed out with red arrowheads. The endotubes also appear to be abnormally thick in some of the mutants. The bottom right figure is an example of a moderate endotube defect. The endotube appears thick but no discontinuities are obvious. The bottom left figure is an example of a normal endotube in a mutant. The endotube appears to be thin and well defined. (B) Cross-sections of four different N2 larvae. The endotube appears to be thin and continuous in all examples. Microvilli are longer and more dense than those seen in the mutant (a single microvillus is indicated with the yellow bracket in the bottom right hand picture). Each diagram is 0.42 nm in width.

gob-1(RNAi)



wild-type

be able to explain the detected mild IFB-2 localization defect in the *gob-1* mutant. The endotube defect may also explain the lumen defect seen in the Normarski pictures of *gob-1* RNAi larvae.

gob-1 is Expressed at the Right Time and Place to be Involved in Intestine Development

gob-1 loss of function results in an early and penetrant intestine defect. To determine if *gob-1* is being expressed at the right time and place to be involved in intestine development, the *gob-1* gene and GOB-1 protein expression was examined. A *gob-1* promoter driven transcriptional reporter was used to track *gob-1* gene expression. A fragment containing 2323 bps of the *gob-1* promoter was used to create a GFP/*lacZ* transcriptional reporter **(see Figure 14 B)**. Multiple independently derived strains were obtained that were heritably transgenic for this reporter. In live embryos, the earliest GFP expression was detected at the 8 E cell stage and was exclusive to the intestine (data not shown). Strong GFP expression was seen in the intestine from this early time point through to and including adulthood. Starting at the comma stage, additional GFP expression is seen in the head and hypodermis **(see Figure 20)**. By the two-fold stage and continuing into adulthood, expression is almost ubiquitous with strong staining in the intestine, pharynx, body wall muscle, tail and unidentified head neurons **(see Figure 21)**. Similar results were seen in embryos stained for β-galactosidase activity **(see Figure 22)**. The above

Figure 20: Live embryos expressing the *gob-1*::GFP/*lacZ* transcriptional reporter. (A) A bean staged embryo showing strong GFP expression in the intestine. The observed hypodermal staining is likely to be due to reporter background (hypodermal cells displaying GFP expression are indicated with yellow arrowheads). (B) A comma staged embryo displaying a strong GFP expression in the intestine. Weak expression is beginning to appear in the head and in the hypodermis (head cells expressing GFP indicated with green arrows). (C) An embryo approaching the 1½-fold stage with strong GFP expression in the intestine. A weak staining is detectable in pharynx precursors (pharynx precursors expressing GFP indicated with green arrows). Scale bar 10 μm. Embryos are shown with the anterior to the left.

Nomarski Green Fluorescent Protein

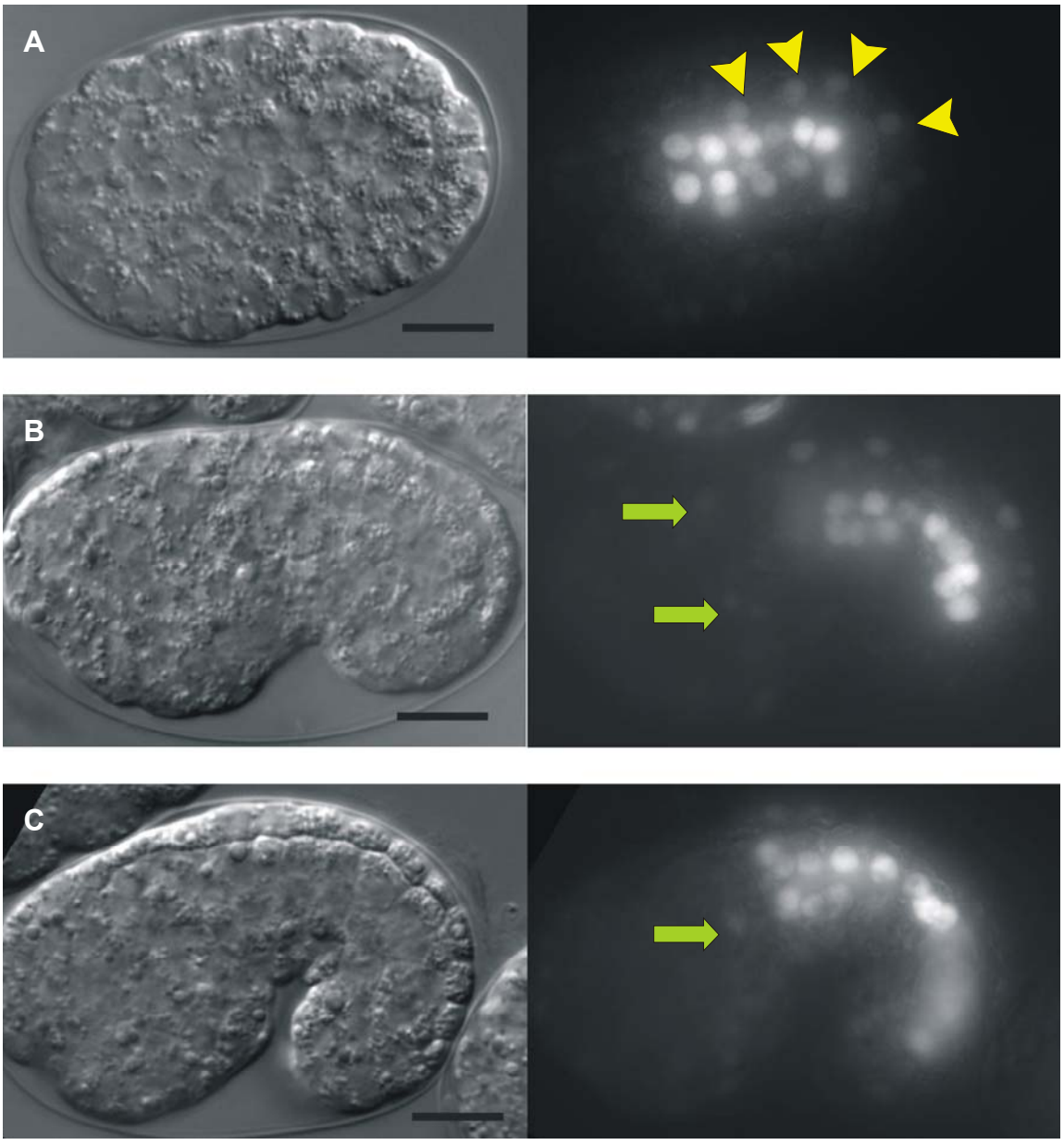


Figure 21: The *gob-1* transcription reporter is expressed in many tissues in the L2 larva. Images are shown of detected GFP expression in the head region (top figure), mid-region (middle figure) and posterior region (bottom figure) of L2 larvae. The upper panel of each pair displays GFP and the lower panel of each pair displays DAPI staining of the same larva. The top figure shows strong staining in the pharynx (defined by white dotted line) and head neurons (yellow arrowheads). The middle figure displays strong GFP expression in the intestinal cells (yellow asterisks). In other samples expression is also seen in the muscle cells. The bottom pair of figures displays strong staining in the posterior intestine and a moderate expression in tail cells (white bracket). Scale bar 10 μm.

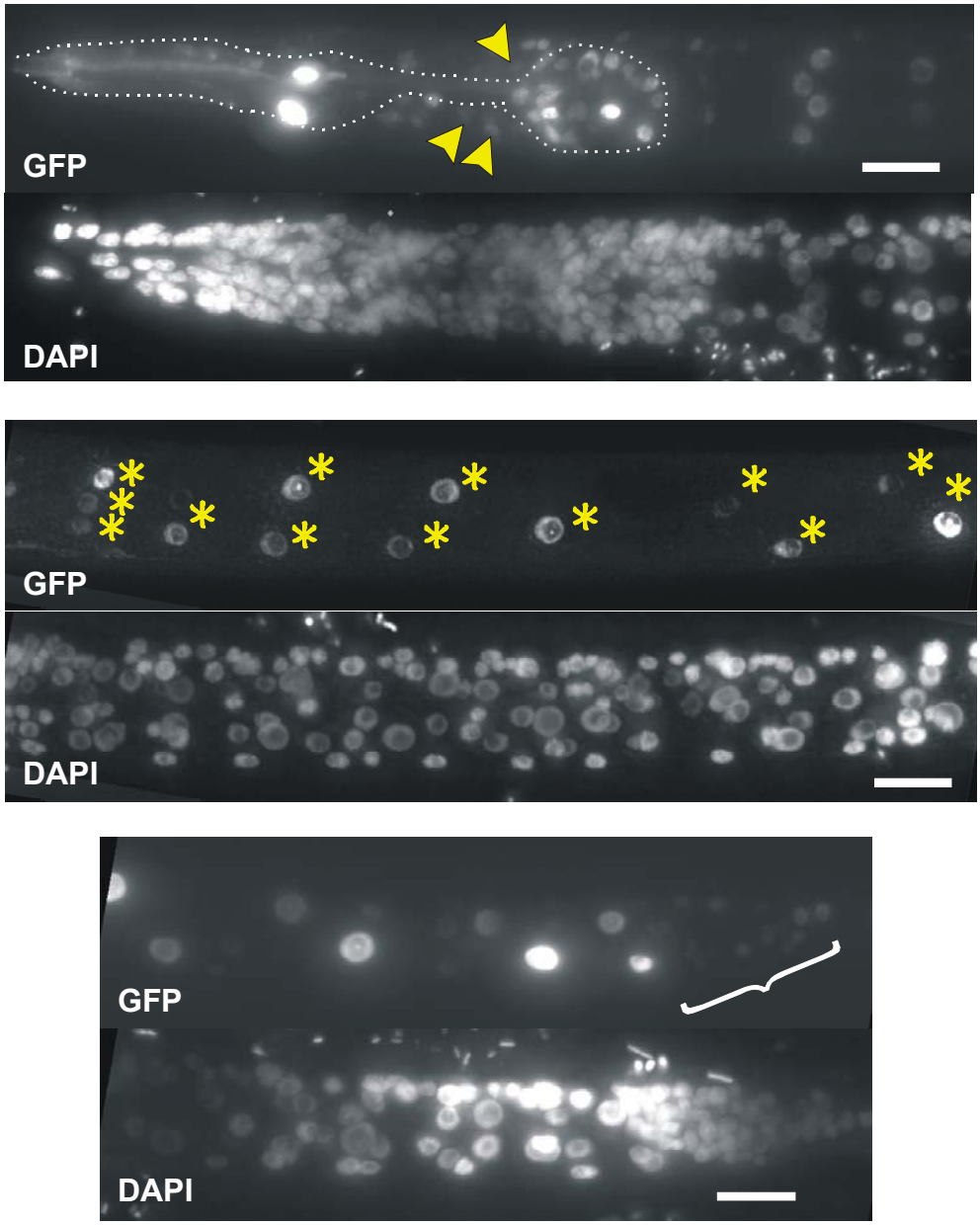
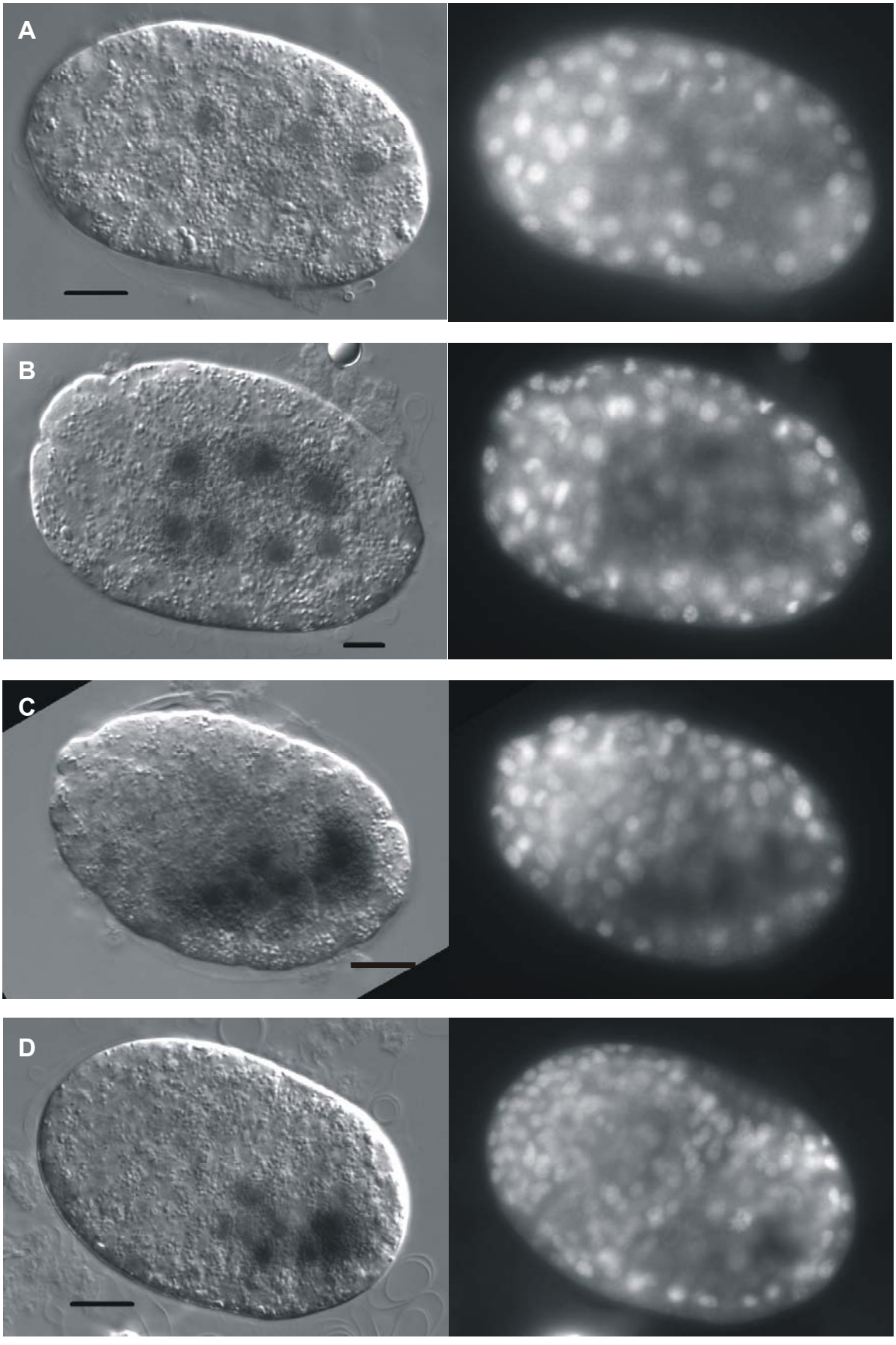

GFP
DAPI
GFP
DAPI
GFP
DAPI

Figure 22: β-galactosidase staining produced by the *gob-1* transcriptional reporter is expressed in the intestine. Embryos transgenic for the *gob-1* GFP/*lacZ* transcriptional reporter were stained for β-galactosidase activity. Pairs of images of β-galactosidase staining (left) and DAPI staining (right) of representative embryos are presented. The earliest appearance of *gob-1* expression is at the 8 E cell stage (~170 total cells (A) and ~230 cells (B)). Weak staining outside the intestine (in the head and hypodermis) is evident at the 12 E cell stage (~250 cells (C)) and the 14 E cell stage (~270 cells (D)). Scale bar is 10 μm. Embryos are shown with their anterior to the left.


A
B
C
D

experiments demonstrated that *gob-1* was being transcribed in the intestine during its early development.

In addition to *gob-1* transcriptional reporters, attempts were made to detect the endogenous GOB-1 protein in the developing embryo. Three regions of the GOB-1 protein that were predicted to be highly antigenic were used to generate antibodies against GOB-1 (Kyte and Doolittle, 1982). The amino acid (AA) sequences used for antibody generation were the following: a N-terminal 102 amino acid glutathione-S-transferase (GST) fusion protein located from amino acid position 11 to 113 (N-GOB-1::GST), a centrally located 87 amino acid GST fusion protein located from amino acid position 239 to 326, and a C-terminal 12 amino acid (12 AA) peptide located from position 382 to 393 **(see Figure 23)**. An antigenic response during immunohistochemistry of fixed embryos was never obtained with the 12 AA peptide and the centrally located 87 AA GST fusion peptide was never injected due to an ambiguity in identifying the expressed protein. The identity of the expressed N-GOB-1::GST was confirmed using Factor Xa cleavage and trypsin cleaved mass spectrophotometry **(see Figure 24)**. Two rabbits were inoculated with the 12 AA peptide and the terminal bleeds were purified using a column to which the peptide was cross-linked **(see Figure 25)**. Two rabbits and four guinea pigs were inoculated with N-GOB-1::GST and and terminal bleeds were affinity purified using a column to which the N-GOB-1::MBD fusion protein was cross-linked **(see Figure 26 and 27)**. In addition to a traditional elution protocol, a more stringent elution protocol was used in an attempt to isolate high affinity antibodies (B. Burke personal communication).

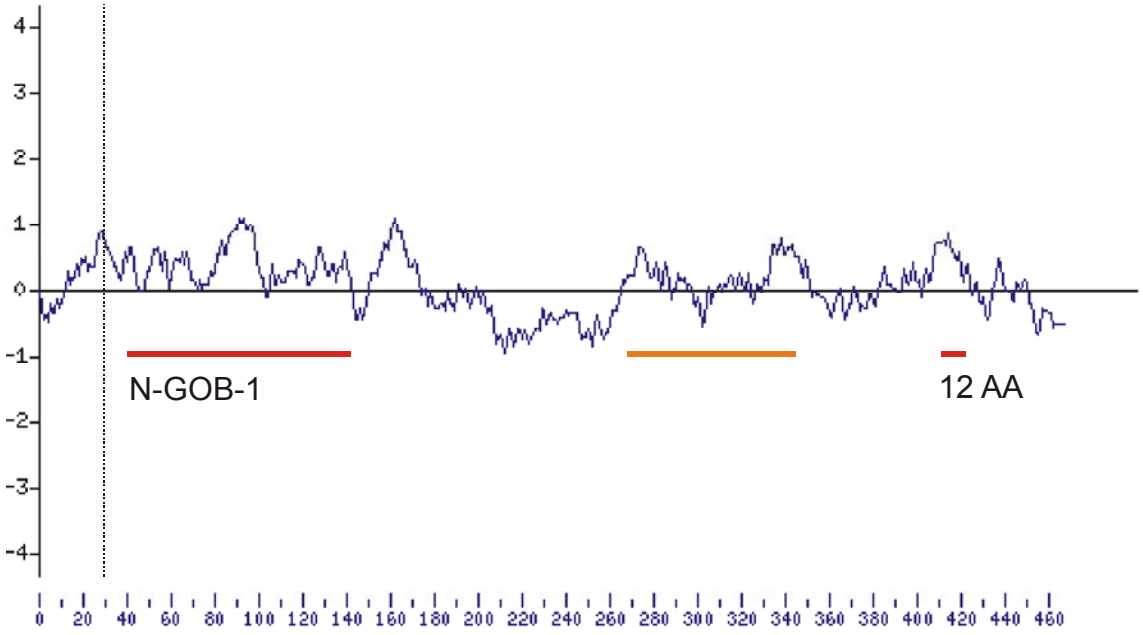


Figure 23: Three peptide regions were considered for generation of anti-GOB-1 antibodies. The antigenicity of the GOB-1 protein was predicted using the Hopp-Woods method (Hopp and Woods, 1981). Default parameters were used with an interval of 17 AA. The actual protein was later confirmed to be 29 AA shorter than predicted in Acedb (black dotted line defines actual length). The intervals used for generation of antibodies are shown with red lines. The names given to each are located below each region. The interval represented with the orange line was also considered but proved to be difficult to express in bacteria.

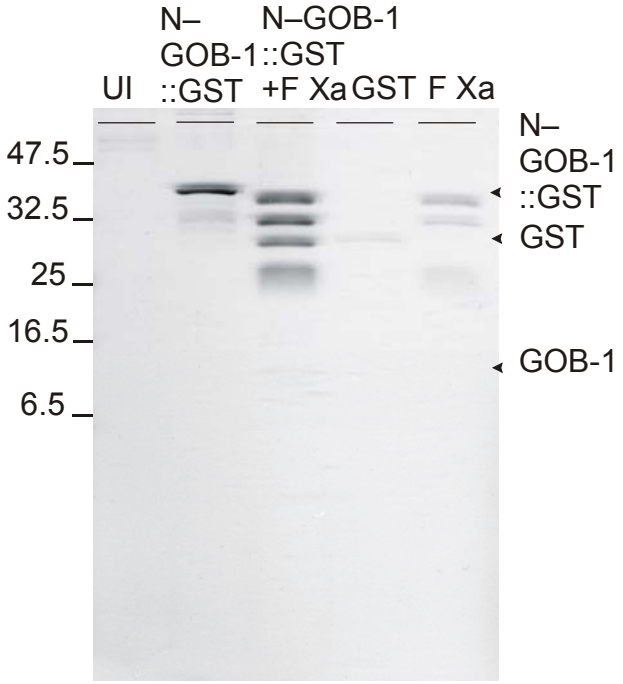
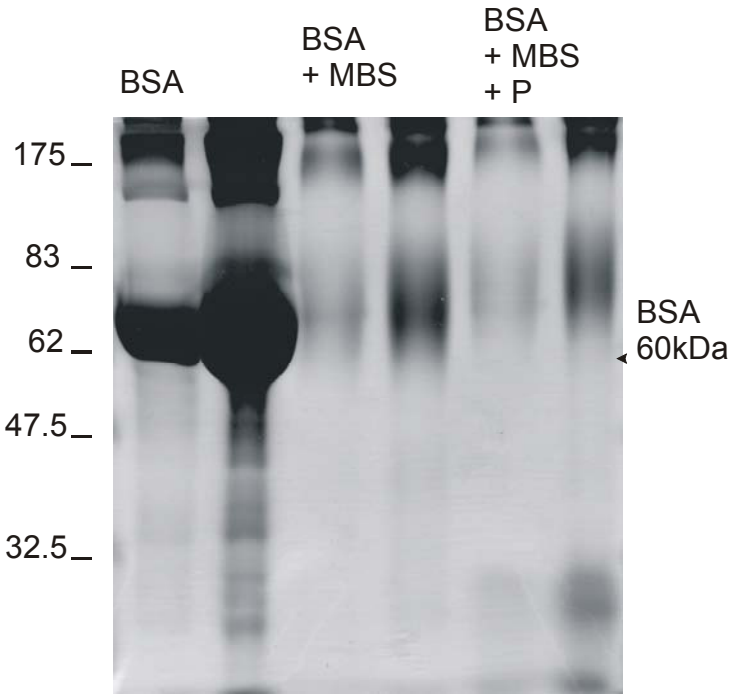


Figure 24: N-GOB-1::GST gives the correct Factor Xa cleavage results. The following proteins (left to right) were run on a 15% PAGE gel and stained with coomassie blue dye: 15 µL of the mock-eluted product of an uninduced culture (UI), 375 ng of the eluted N-GOB-1::GST protein (39.3 KDa), ~2.7 µg of the Factor Xa (F Xa) cleaved N-GOB-1::GST (N-GOB-1 (13.0 KDa) and GST (26.3 KDa)), 50 ng of GST control protein, and 500 ng of the Factor Xa enzyme.

Figure 25: The antibodies present in final bleed serums of the 12 AA inoculated rabbits 1111 and 1114 are specific for the 12 AA peptide. (A) The peptide was cross-linked to BSA. The cross-linking intermediates and final product are displayed on the coomasie stained gel (see methods). The first two lanes show two dilutions of the starting BSA solution. The next two lanes show two dilutions of the BSA protein exposed to the cross-linker MBS. The last two columns show two dilutions of the BSA cross-linked to the 12 AA peptide. The increase in the molecular weight of the BSA protein at each linkage step suggests that successful cross-linking has occurred. (B) PBS, BSA, GST, the 12 AA peptide (P) and the 12 AA peptide linked to BSA (BSA + P) were spotted on filter paper and reacted with either the pre-immune serums or with the final bleeds of the two inoculated rabbits. Rabbit 1114 gave the most specific detection of the 12 AA peptide. Rabbit 1111 gave a preferential detection of the 12 AA peptide over other proteins tested.

A


BSA
BSA
+ MBS
BSA
+ MBS
+ P
175
83
62
47.5
32.5
BSA
60kDa

B

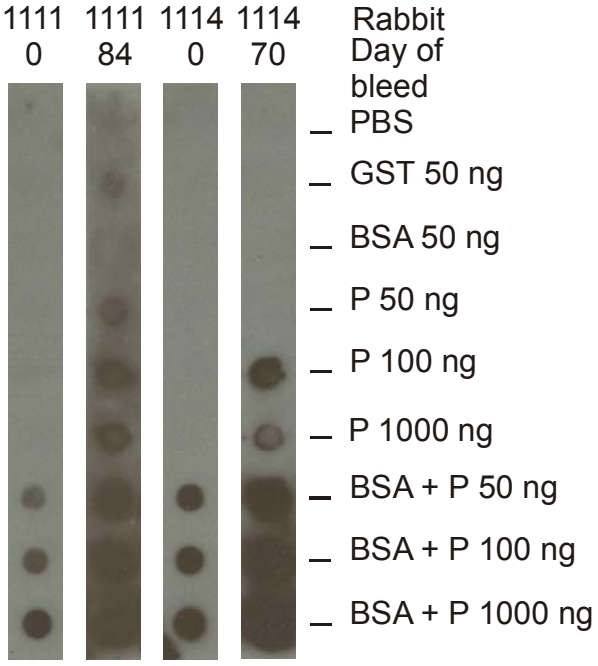
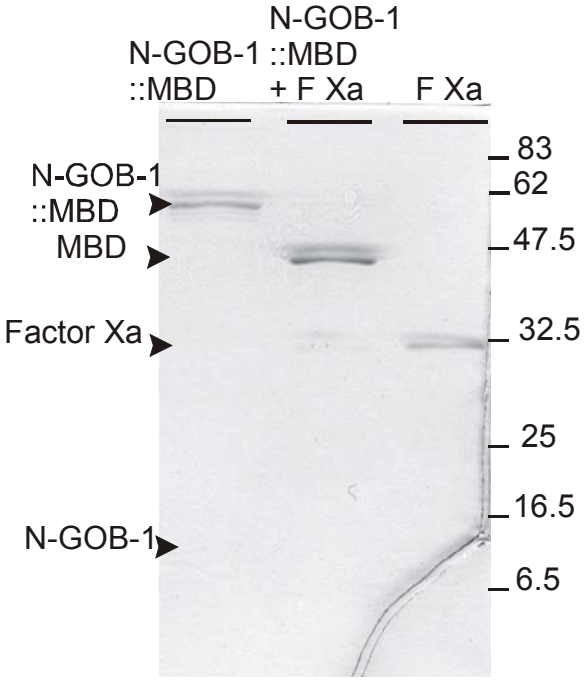

1111
0
1111
84
1114
0
1114
70
Rabbit
Day of
bleed
PBS
GST 50 ng
BSA 50 ng
P 50 ng
P 100 ng
P 1000 ng
BSA + P 50 ng
BSA + P 100 ng
BSA + P 1000 ng

Figure 26: The N-GOB-1::MBD fusion protein gives the expected products when cleaved with Factor Xa. (A) 12% coomasie stained PAGE gel of the following proteins: 500 ng of the eluted N-GOB-1::MBD protein (53.8 kDa), 820 ng of Factor Xa (F Xa) cleaved N-GOB-1::MBD protein (N-GOB-1 13.0 kDa and MBD 40.8 kDa) and 369 ng of the Factor Xa enzyme. Parallel samples were run on the same gel and used for Western blotting using an anti-MBD antibody. (B) Western blot of the same proteins as in A. The anti-MBD monoclonal antibody was used at a 1/5000 dilution and the alkaline phosphatase secondary antibody was used at a dilution of 1/2500.

A


N-GOB-1 ::MBD
N-GOB-1 ::MBD + F Xa
F Xa
83
62
47.5
32.5
25
16.5
6.5
N-GOB-1 ::MBD
MBD
Factor Xa
N-GOB-1

B

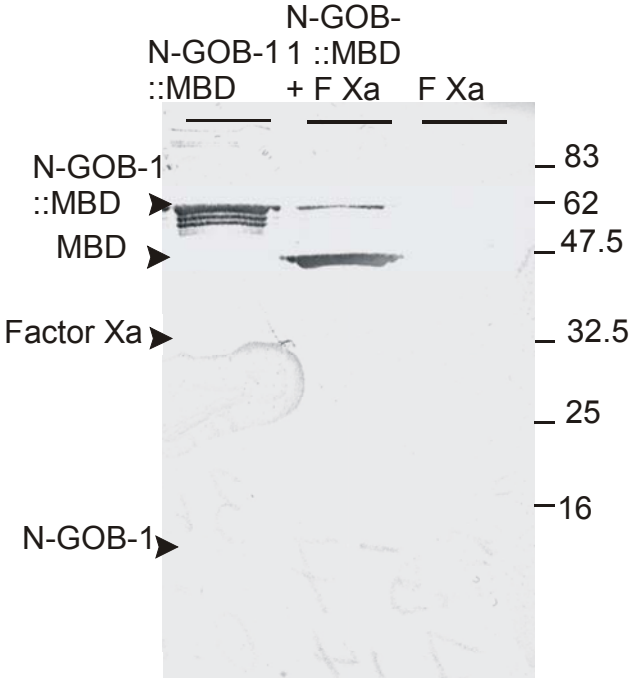
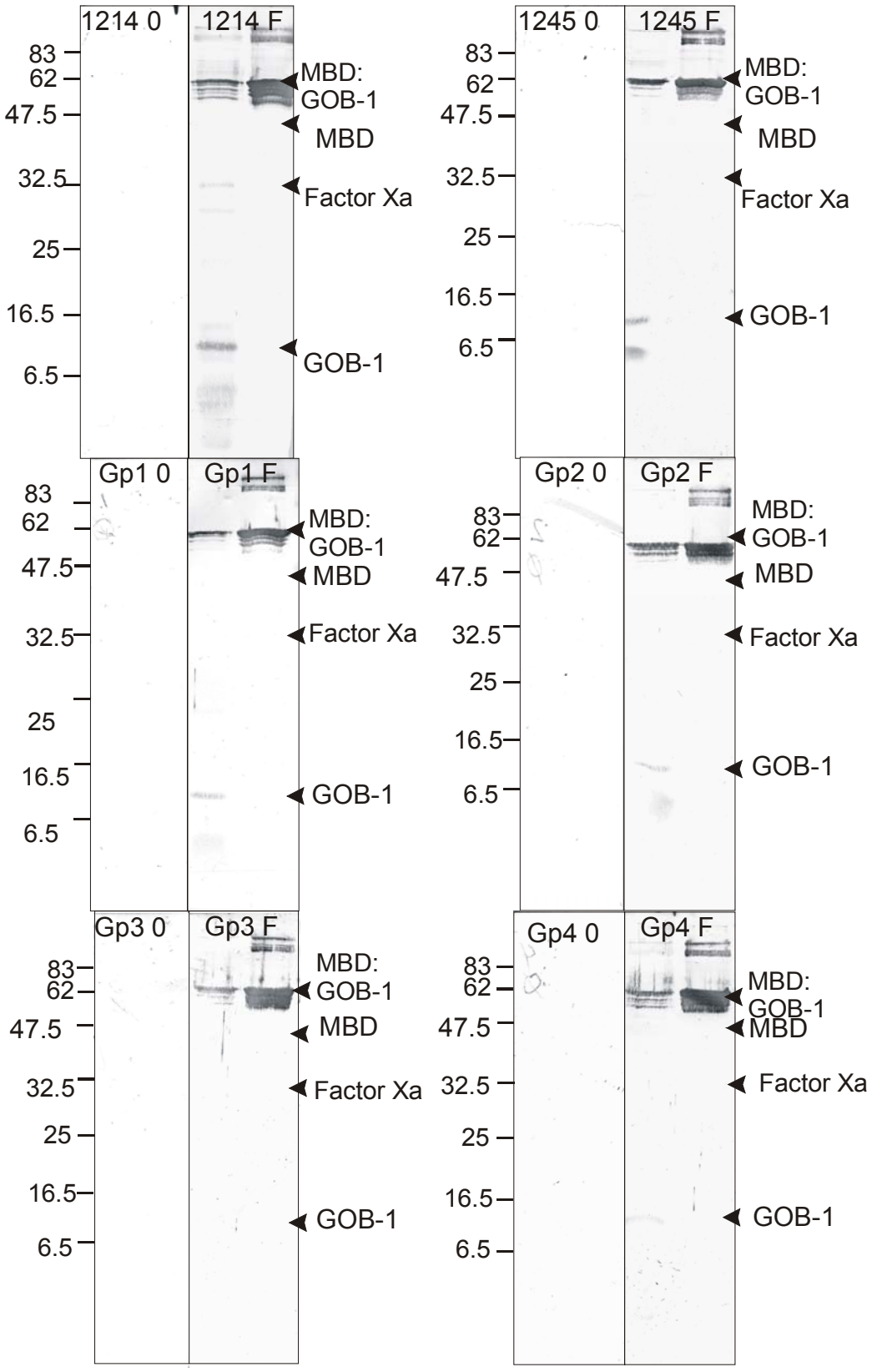

N-GOB-1 ::MBD
N-GOB-1 ::MBD + F Xa
F Xa
83
62
47.5
32.5
25
16
N-GOB-1 ::MBD
MBD
Factor Xa
N-GOB-1

Figure 27: The antibodies obtained from most of the animals inoculated with the N-GOB-1::GST protein specifically detect the N-GOB-1 portion of the N-GOB-1::MBD protein. Twelve Western blots are shown of the two rabbits (top row) and four guinea pigs (bottom two rows) inoculated with the N-GOB-1::GST protein. For each animal, the pre-bleed serum is on the left and the terminal bleed serum is on the right. For each blot, 820 ng of the Factor Xa cleaved N-GOB-1::MBD protein is in the first lane and 500 ng of the full length N-GOB-1::MBD protein is in the second lane. The proteins were run on a 12% PAGE gel and transferred for Western analysis. Each blot had a 1/1000 dilution of serum and a 1/1000 dilution of the alkaline phosphatase conjugated secondary antibody. For all animals except rabbit 1214, the antibodies in the terminal bleeds specifically recognize the N-GOB-1 peptide and the full length N-GOB-1::MBD protein.


1214 0
1214 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1
1245 0
1245 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1
Gp1 0
Gp1 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1
Gp2 0
Gp2 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1
Gp3 0
Gp3 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1
Gp4 0
Gp4 F
83
62
47.5
32.5
25
16.5
6.5
MBD:
GOB-1
MBD
Factor Xa
GOB-1

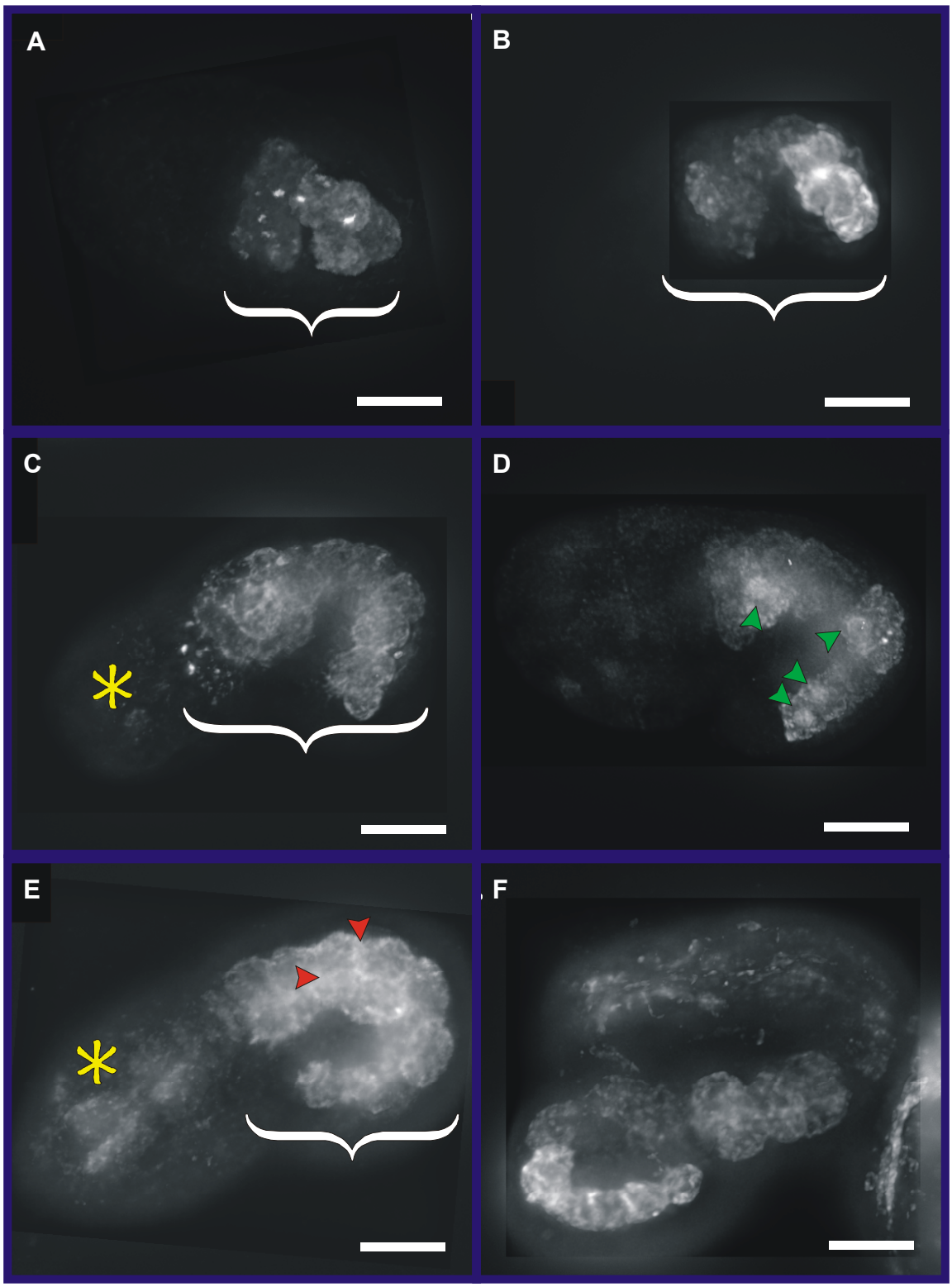
Despite extensive efforts, the GOB-1 protein could not be detected in wild-type animals. All of the antibodies generated against the N-GOB-1::GST fusion protein (primary antibody dilution range of 1/3 to 1/1000 and secondary antibody dilution range of 1/100 to 1/1000) were used in an attempt to refine GOB-1 detection patterns. In an additional effort to detect the endogenous GOB-1 protein, a biotin conjugated secondary antibody (1/3 to 1/1000 dilution) with a streptavadin conjugated tertiary antibody (1/1000 to 1/2500 dilution) were used in an attempt to detect a potentially weak signal. In an attempt to reduce non-specific back ground, primary and secondary antibodies were pre-absorbed to *gob-1*(RNAi) arrested larvae but also did not improve GOB-1 detection. The failure to detect GOB-1 protein in wild-type animals was attributed to three possibilities. First, the N-GOB-1::GST fusion protein used for antibody generation might not represent an *in vivo* protein conformation. Second, the GOB-1 protein may be masked by other proteins or expressed at extremely low levels *in vivo* and could not be detected using conventional immunohistochemistry. Third, the acetone/methanol fixation protocol used during the immunohistochemical staining protocol was not optimal to maintain the GOB-1 protein in a detectable conformation.

In order to address the scenario of GOB-1 under expression, a strain was constructed that over-expressed the wild-type GOB-1 protein. A 5831 bp long fragment of *gob-1* genomic sequence was inserted into a modified version of the pPD 95.11 vector **(see Figure 14 B)**. The *gob-1* genomic sequence inserted into the vector included 1885 bp of upstream *gob-1* sequence (24 bp 3' to the stop

codon of the upstream gene) and 674 bp of the downstream *gob-1* sequence (134 bp down stream of the *gob-1* polyadenylation splice site). The inclusion of this entire genomic region was thought to be potentially large enough to reflect endogenous GOB-1 expression. The premise behind this experiment was that the incorporation of this genomic region into a transgenic array, and thus retention of multiple copies of this region within the animal, would create a situation where GOB-1 would be expressed at a high enough level for immunohistochemical detection. Three independently derived and heritable transgenic strains (non-integrated) were obtained from two different injection experiments (i.e., six lines were studied in total). All six strains gave identical staining results. Furthermore, all antibodies generated against N-GOB-1::GST gave similar staining patterns with only subtle differences detected between each animal. The purified antibodies from a single guinea pig that gave the most consistent staining pattern were used in the remainder of the experiments.

The immunohistochemical staining pattern seen in the GOB-1 over-expressing strains was similar to those predicted by the transcriptional reporter **(see Figure 28)**. The earliest stage at which GOB-1 could be detected was the 8 E cell stage and at this time point was intestine specific. Intestine specific detection continued until the comma stage, when staining could also be detected in the head. By the two fold stage, strong staining in the head, pharynx and intestine was detected. In the pretzel, several tissues showed GOB-1 staining such as the pharynx, intestine, muscle and hind gut. The protein appeared consistently cytoplasmic throughout all of the stages examined. Cytoplasmic

Figure 28: GOB-1 protein localization in the developing embryo. Using transgenic animals over-expressing GOB-1, the GOB-1 protein can be detected in the cytoplasm of intestinal cells. (A) At the 8 E cell stage, GOB-1 can be detected in the intestine (intestine is indicated with a white bracket). GOB-1 is only expressed in the intestine at this stage. (B) GOB-1 staining is detected in the intestine of the a comma stage embryo (intestine is indicated with a white bracket). In some comma embryos, weak GOB-1 expression is seen in the head. (C) GOB-1 expression is evident in the intestine (intestine is indicated with a white bracket) and head region (head staining indicated with yellow asterisk) of the 1½-fold embryo. (D) GOB-1 can be detected in the intestinal nuclei of some embryos (green arrowheads point to stained nuclei). (E) Extensive staining can be detected in the pharynx (pharynx staining indicated with yellow asterisk) of the 2-fold embryo. In this particular picture GOB-1 appears to be localized at the lumenal surface and points of intestine cell contact (the intestine is indicated with a white bracket and red arrowheads indicate sites of cortical enrichment). (F) By the pretzel stage, GOB-1 is nearly ubiquitously expressed. Scale bar 10 μm.


A
B
C
D
E
F

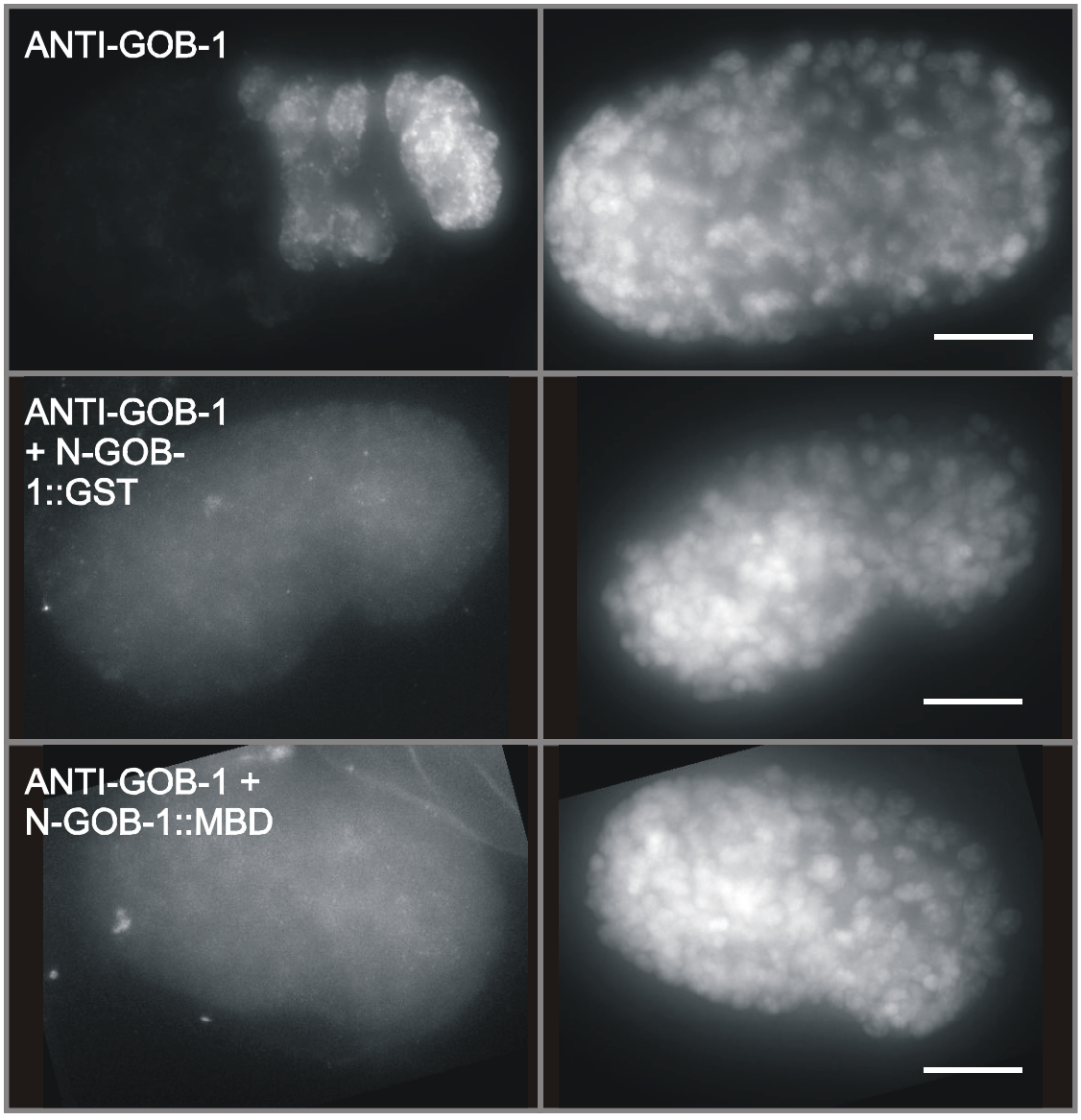
staining was excluded from undefined cytoplasmic bodies and appeared to be stronger at the cell cortex and gut lumen of some embryos. Curiously, a strong staining in gut nuclei was detected in a subset of the population of embryos at all stages. The proportion of the embryo population that displayed nuclear staining varied for each transgenic line (i.e., 19/27 in one line and 1/23 in another) and the proportion of the population that showed nuclear staining decreased to undetectable levels as each strain was passaged. Nonetheless, a certain degree of nuclear staining was detected in all of the lines studied. The consistent detection of cytoplasmic staining from experiment to experiment suggests that this may reflect the endogenous GOB-1 expression pattern. It presently cannot be determined if the nuclear detection of GOB-1 represents its true localization pattern because of the inconsistencies in its detection. The nuclear detection may well be an artefact of over-expression. This possibility will be difficult to resolve without the detection of the GOB-1 protein in the wild-type animal.

To identify an embryo as transgenic, detection of GOB-1 protein and the expression of the co-injected transgenic reporter (a *ttx-3* promoter driven GFP) was examined. In these experiments, there was a strong coincidence between the presence of GFP in neuron precursors in the head and the detection of GOB-1 staining (i.e., 79.6% (n = 54) and 86.7% (n = 45) in two independent lines). Animals that had lost the transgenic array had no detectable staining above background and this was similar to the results obtained with the staining of wild-type animals. Mosaic detection of the over-expression array was expected due

to the loss of the non-integrated array. The absence of staining concomitant with the loss of the transgenic array suggests that the antibody is indeed specific for the over-expressed GOB-1 protein. As a further specificity control, the antibodies used for GOB-1 detection were pre-incubated with the following proteins: the N-GOB-1::GST protein used to generate the anti-GOB-1 antibodies, the N-GOB-1::MBD protein used to purify the anti-GOB-1 antibodies and the full length GOB-1::MBD fusion protein **(see Figure 29)**. The full length GOB-1::MBD protein has been found to have robust enzymatic activity and is therefore likely to represent the configuration of the endogenous protein (see CHAPTER FIVE - Characterization of GOB-1 Function). In all pre-incubation experiments, antibody staining was eliminated in all examined embryos, i.e., all detected staining could be competed away by the GOB-1 protein. This has lead to the conclusion that the antibodies used in the above experiments were specifically detecting the over-expressed GOB-1 protein and may reflect an endogenous temporal and tissue-specific expression pattern.

In summary, the loss of *gob-1* function results in a penentrant intestine defect and early larval lethality. The intestine defect is obvious at hatching and suggests that the intestine defect is manifested during embryogenesis. The reason for the gut obstructions seen in the *gob-1* mutants is not entirely obvious but may be related to defects seen in the endotube structure of the mutant. The *gob-1* gene and GOB-1 protein are expressed during the development of the embryonic gut and therefore are expressed at the right time and place to be required for intestine development.

Figure 29: The immunohistochemical detection of the over-expressed GOB-1 protein is eliminated when the anti-GOB-1 antibody is pre-incubated with the GOB-1 fusion proteins. Anti-GOB-1 staining pictures are on the left and the same embryos DAPI stained are on the right. All embryos are approximately at the comma stage and are representative of all of the embryos for each experiment. The immunohistochemical signal seen in the top figure is completely eliminated when the antibody is pre-incubated with the N-GOB-1::GST protein (the protein used to generate the antibody) or the N-GOB-1::MBD protein (the protein used to purify the antibodies). A similar result (not shown) was obtained when the anti-GOB-1 antibody was pre-incubated with the full length (catalytically active) GOB-1::MBD protein. Scale bars are 10 μm.


ANTI-GOB-1
ANTI-GOB-1
+ N-GOB-
1::GST
ANTI-GOB-1 +
N-GOB-1::MBD

CHAPTER FIVE - Characterization of GOB-1 Function

GOB-1 is a member of the HAD-like hydrolase family

With the *gob-1* gene being identified as having a critical function in the *C. elegans* embryonic intestine, attempts were made to identify the molecular function of the GOB-1 protein. Searches for GOB-1 sequence similarity to other genes were used as a method to identify a potential GOB-1 function. A standard BLASTP search using the GOB-1 protein as the query identified four similar proteins (National Center for Biotechnology Information http://www.ncbi.nlm.nih.gov; Altschul et al., 1990). One of the proteins identified was a predicted protein for *Caenorhabditis briggsae*, a closely related roundworm. The *C. briggsae* protein is of a similar size to the GOB-1 protein and shows 77% amino acid identity to GOB-1 across the entire protein sequence. With this degree of identity, this protein is likely to be a *C. briggsae gob-1* ortholog. Two other larger *C. elegans* proteins, TPS-1 and TPS-2 (Trehalose Phosphate Synthase), also show a significant similarity to GOB-1 in their C-terminal portions (29% identity and 47% similarity, and 27% identity and 47% similarity respectively). The final detected protein was found in the genome of *Chlorobium tepidium* a green sulphur bacterium. This latter protein was predicted to exist and showed nearly as much identity as the two other *C. elegans* proteins to GOB-1 with a 25% identity and 48% similarity across the entire length of the protein.

Clues to GOB-1 function were not initially obvious based on the above comparisons. The prokaryotic gene is only a hypothetical protein and has not been characterized. TPS-1 and TPS-2 proteins are known to be involved in trehalose metabolism but the domain related to their known catalytic activity, the glycosyltransferase 20 domain, is located outside the shared domain. Initial clues to GOB-1's molecular activity were obtained utilizing a PSI-BLAST search in which TPS-1, TPS-2, GOB-1 and the *Chlorobium tepidium* protein sequence was used to create a PSI-BLAST weighted matrix (Altschul et al., 1997). The third reiteration of the search detected GOB-1 similarity to a large family of hypothetical prokaryotic proteins. The top five matches gave E values ranging from $5e^{-26}$ to $9e^{-07}$. Although none of the identified proteins had a described function, all were predicted to be members of the family of the HAD (Haloacid dehalogenase)-like hydrolase superfamily. Indeed, the highly conserved residues of the HAD hydrolase catalytic domains are absolutely conserved in the GOB-1 protein (Aravind et al., 1998; Koonin and Tatusov, 1994) **(see Figure 30)**. The most highly conserved domain of GOB-1, domain I containing the DXDGT domain, shows the closest similarity to members of the HAD family that are sugar phosphatases **(see Figure 31)**.

GOB-1 Phosphatase Assay

In November of 2003, three papers were published in Nature describing a novel role for members of the Eyes absent family of proteins (Li et al., 2003;

Figure 30: GOB-1 and closely related proteins possess all of the conserved domains required to be a HAD-like hydrolase. The T-coffee alignment of GOB-1, the *C. briggsae* GOB-1 ortholog, the *C. elegans* TPS-1 and TPS-2 proteins and the *C. tepidium* predicted protein is shown. The domains that are highly conserved in HAD-like hydrolase family members are displayed in blue above the alignment. The domains that are often conserved are displayed in red above the alignment. The presence of an “h” represents the requirement for a hydrophobic residue.

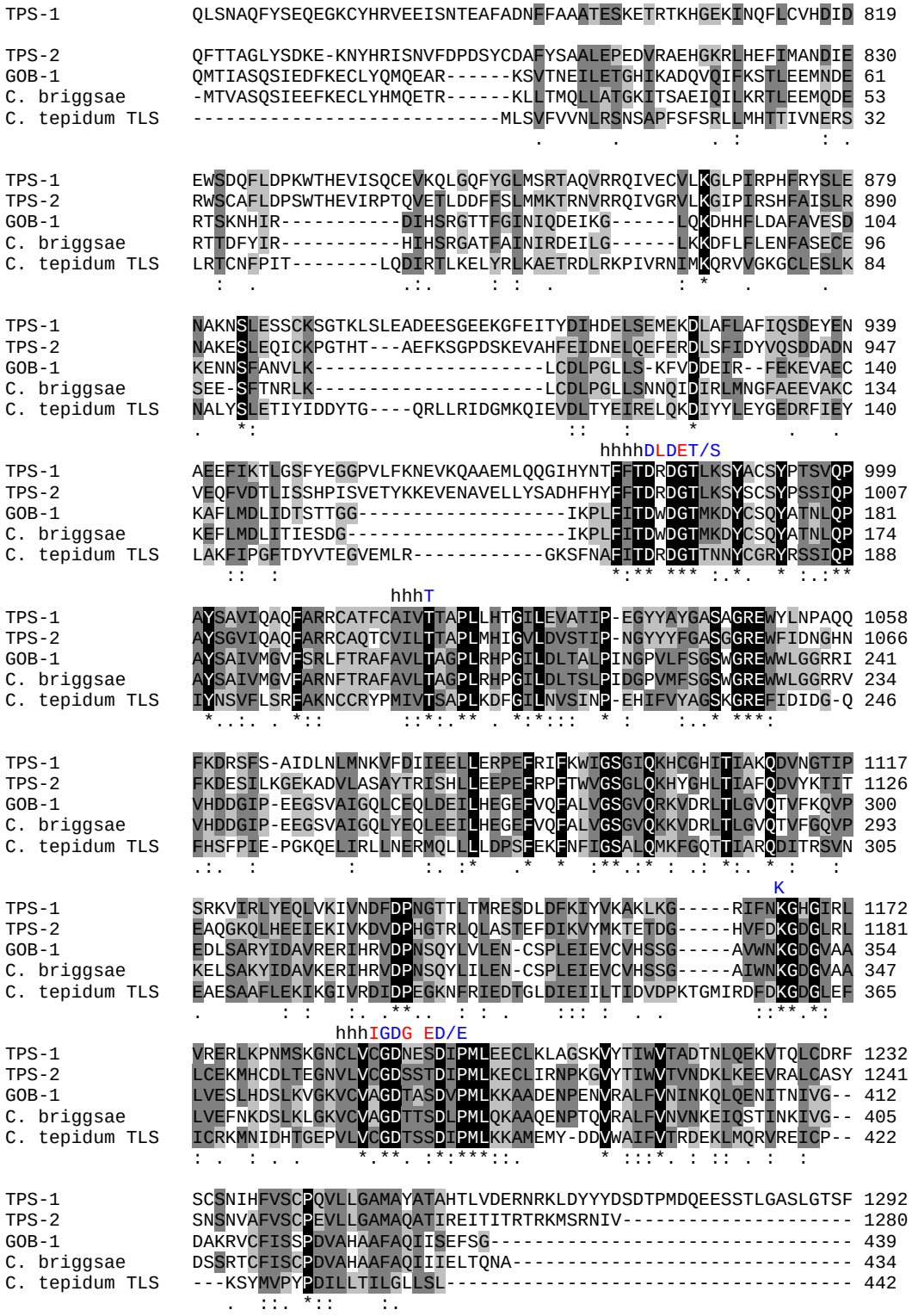
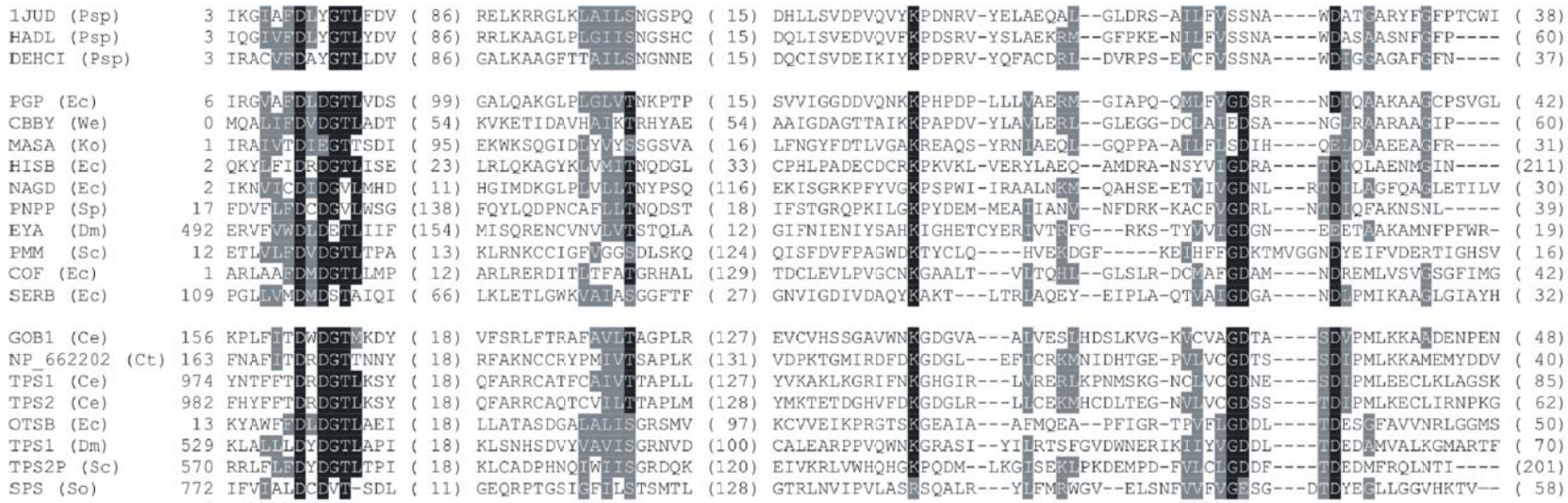

TPS-1 QLSNAQFYSEQEGKCYHRVEEISNTEAFADNFFAAATESKETRTKHGEKINQFLCVHDID 819
TPS-2 QFTTAGLYSDKE-KNYHRISNVFDPDSYCDAFYSAALEPEDVRAEHGKRLHEFIMANDIE 830
GOB-1 QMTIASQSIEDFKECLYQMQEAR------KSVTNEILETGHIKADQVQIFKSTLEEMNDE 61
C. briggsae -MTVASQSIEEFKECLYHMQETR------KLLTMQLLATGKITSAEIQILKRTLEEMQDE 53
C. tepidum TLS ----------MLSVFVVNLRSNSAPFSFSRLLMHTTIVNERS 32
TPS-1 EWSDQFLDPKWTHEVISQCEVKQLGQFYGLMSRTAQVRRQIVECVLKGLPIRPHFRYSLE 879
TPS-2 RWSCAFLDPSWTHEVIRPTQVETLDDFFSLMMKTRNVRRQIVGRVLKGIPIRSHFAISLR 890
GOB-1 RTSKNHIR----------DIHSRGTTFGINIQDEIKG------LQKDHHFLDAFAVESD 104
C. briggsae RTTDFYIR----------HIHSRGATFAINIRDEILG------LKKDFLFLENFASECE 96
C. tepidum TLS LRTCNFPIT--------LQDIRTLKELYRLKAETRDLRKPIVRNIMKQRVVGKGCLESLK 84
TPS-1 NAKNSLESSCKSGTKLSLEADEESGEEKGFEITYDIHDELSEMEKDLAFLAFIQSDEYEN 939
TPS-2 NAKESLEQICKPGTHT---AEFKSGPDSKEVAHFEIDNELQEFERDLSFIDYVQSDDADN 947
GOB-1 KENNSFANVLK----------LCDLPGLLS-KFVDDEIR--FEKEVAEC 140
C. briggsae SEE-SFTNRLK----------LCDLPGLLSNNQIDIRLMNGFAEEVAKC 134
C. tepidum TLS NALYSLETIYIDDYTG----QRLLRIDGMKQIEVDLTYEIRELQKDIYYLEYGEDRFIEY 140
hhhhDLDET/S
TPS-1 AEEFIKTLGSFYEGGPVLFKNEVKQAAEMLQQGIHYNTFFTDRDGTLKSYACSYPTSVQP 999
TPS-2 VEQFVDTLISSHPISVETYKKEVENAVELLYSADHFHYFFTDRDGTLKSYSCSYPSSIQP 1007
GOB-1 KAFLMDLIDTSTTGG----------IKPLFITDWDGTMKDYCSQYATNLQP 181
C. briggsae KEFLMDLITIESDG----------IKPLFITDWDGTMKDYCSQYATNLQP 174
C. tepidum TLS LAKFIPGFTDYVTEGVEMLR----------GKSFNAFITDRDGTTNNYCGRYRSSIQP 188
hhhT
TPS-1 AYSAVIQAQFARRCATFCAIVTTAPLLHTGILEVATIP-EGYYAYGASAGREWYLNPAQQ 1058
TPS-2 AYSGVIQAQFARRCAQTCVILTTAPLMHIGVLDVSTIP-NGYYYFGASGGREWFIDNGHN 1066
GOB-1 AYSAIVMGVFSRLFTRAFAVLTAGPLRHPGILDLTALPINGPVLFSGSWGREWWLGGRRI 241
C. briggsae AYSAIVMGVFARNFTRAFAVLTAGPLRHPGILDLTSLPIDGPVMFSGSWGREWWLGGRRV 234
C. tepidum TLS IYNSVFLSRFAKNCCRYPMIVTSAPLKDFGILNVSINP-EHIFVYAGSKGREFIDIDG-Q 246
TPS-1 FKDRSFS-AIDLNLMNKVFDIIEELLERPEFRIFKWIGSGIQKHCGHITIAKQDVNGTIP 1117
TPS-2 FKDESILKGEKADVLASAYTRISHLLEEPEFRPFTWVGSGLQKHYGHLTIAFQDVYKTIT 1126
GOB-1 VHDDGIP-EEGSVAIGQLCEQLDEILHEGEFVQFALVGSGVQRKVDRLTLGVQTVFKQVP 300
C. briggsae VHDDGIP-EEGSVAIGQLYEQLEEILHEGEFVQFALVGSGVQKKVDRLTLGVQTVFGQVP 293
C. tepidum TLS FHSFPIE-PGKQELIRLLNERMQLLLLDPSFEKFNFIGSALQMKFGQTTIARQDITRSVN 305
K
TPS-1 SRKVIRLYEQLVKIVNDFDPNGTTLTMRESDLDFKIYVKAKLKG-----RIFNKGHGIRL 1172
TPS-2 EAQGKQLHEEIEKIVKDVDPHGTRLQLASTEFDIKVYMKTETDG-----HVFDKGDGLRL 1181
GOB-1 EDLSARYIDAVRERIHRVDPNSQYLVLEN-CSPLEIEVCVHSSG-----AVWNKGDGVAA 354
C. briggsae KELSAKYIDAVKERIHRVDPNSQYLILEN-CSPLEIEVCVHSSG-----AIWNKGDGVAA 347
C. tepidum TLS EAESAAFLEKIKGIVRDIDPEGKNFRIEDTGLDIEIILTIDVDPKTGMIRDFDKGDGLEF 365
hhhIGDG ED/E
TPS-1 VRERLKPNMSKGNCLVCGDNESDIPMLEECLKLAGSKVYTIWVTADTNLQEKVTQLCDRF 1232
TPS-2 LCEKMHCDLTEGNVLVCGDSSTDIPMLKECLIRNPKGVYTIWVTVNDKLKEEVRALCASY 1241
GOB-1 LVESLHDSLKVGKVCVAGDTASDVPMLKKAADENPENVRALFVNINKQLQENITNIVG-- 412
C. briggsae LVEFNKDSLKLGKVCVAGDTTSDLPMLQKAAQENPTQVRALFVNVNKEIQSTINKIVG-- 405
C. tepidum TLS ICRKMNIDHTGEPVLVCGDTSSDIPMLKKAMEMY-DDVWAIFVTRDEKLMQRVREICP-- 422
TPS-1 SCSNIHFVSCPQVLLGAMAYATAHTLVDERNRKLDYYYDSDTPMDQEESSTLGASLGTSF 1292
TPS-2 SNSNVAFVSCPEVLLGAMAQATIREITITRTRKMSRNIV---------- 1280
GOB-1 DAKRVCFISSPDVAHAAFAQIISEFSG---------- 439
C. briggsae DSSRTCFISCPDVAHAAFAQIIIELTQNA---------- 434
C. tepidum TLS ---KSYMVPYPDILLTILGLLSL---------- 442

Figure 31: Alignment of the GOB-1 Domain I, Domain II and Domain III sequence (Aravind et al., 1998), with similar domains of the other HAD-like hydrolases suggest that GOB-1 is most closely related to the sugar phosphatases or the serine phosphate phosphatases. Enzymatic activity and species from which they were isolated for each protein are given below. 1JUD = L-2-haloacid dehalogenase II (*Pseudomonas sp. YL*), HADL = L-2 haloalkanoic acid dehalogenase (*Pseudomonas putida*), DEHCI = (S)-2-haloacid dehalogenase I (*Pseudomonas sp. CBS3*), PGP = phosphoglycolate phosphatase (*Escherichia coli*), CBBY = pentose-5-phosphate 3-epimerase (*Wautersia eutropha*), MASA = E-1 enzyme enolase-phosphatase (*Klebsiella oxytoca*), HISB = histidinol-phosphatase (*Escherichia coli O6*), NAGD = n-acetylglucosamine (*Escherichia coli*), EYA = eyes absent protein phosphatase (*Drosophila melanogaster*), PMM = phosphomannomutase (*Saccharomyces cerevisiae*), COF = Cof protein (*Escherichia coli*), SERB = phosphoserine phosphatase (*Escherichia coli*), GOB-1= trehalose phosphate phosphatase (*Caenorhabditis elegans*), NP_662202 = hypothetical protein (*Chlorobium tepidium*), TPS-1 = predicted trehalose phosphate phosphatase (*Caenorhabditis elegans*), TPS-2 = predicted trehalose phosphate phosphatase (*Caenorhabditis elegans*), OTSB trehalose phosphate phosphatase (*Escherichia coli*), TPS1 = trehalose phosphate phosphatase (*Drosophila melanogaster*), TPS2 = trehalose phosphate phosphatase (*Saccharomyces cerevisiae*) and SPS = sucrose phosphate phosphatase (*Spinacia oleracea*).


1JUD (Psp) 3 IKGIAFDLYGTLFDV (86) RELKRRGLKLAILSNGSPQ (15) DHLLSVDPVQVYKPDNRV-YELAEQAH--GLDRS-AILFVSSNA----WDATGARYFGFPTCWI (38)
HADL (Psp) 3 IQGIVFDLYGTLYDV (86) RRLKAAGLPLGIISNGSHC (15) DQLISVEDVQVFKPDSRV-YSLAEKRM--GFPKE-NILFVSSNA----WDASAASNFGFP---- (60)
DEHCI (Psp) 3 IRACVFDAYGTLLDV (86) GALKAAGFTTAILSNGNNE (15) DQCISVDEIKIYKPDPRV-YQFACDRI--DVRPS-EVCFVSSNA----WDIGGAGAFGFN---- (37)
PGP (Ec) 6 IRGVAFDLDGTLVDS (99) GALQAKGLPLGLVTNKPTP (15) SVVIGGDDVQNKKPHPDP-LLLVAERM--GIAPQ-QMLFVGDSR----NDIQAAKAAGCPSVGL (42)
CBBY (We) 0 MQALIFDVDGTLADT (54) KVKETIDAVHAIKTRHYAE (54) AAIGDAGTTAIKKPAPDV-YLAVLERI--GLEGG-DCLAIEDSA----NGIRAARAAGIP---- (60)
MASA (Ko) 1 IRAIVTDIEGTTSDI (95) EKWKSQGIDLYVYSSGSVA (16) LFNGYFDTLVGAKREAQS-YRNIAEQL--GQPPA-AILFLSDIH----QELDAAEEAGFR---- (31)
HISB (Ec) 2 QKYLFIDRDGTLISE (23) LRLQKAGYKLVMITNQDGL (33) CPHLPADECDCRKPKVKL-VERYLAEQ--AMDRA-NSYVLGDRA----TDIQLAENMGIN---- (211)
NAGD (Ec) 2 IKNVICDIDGVLMHD (11) HGIMDKGLPLVLLTNYPSQ (116) EKISGRKPFYVGKPSPWI-IRAALNKM--QAHSE-ETVIVGDNL---RTDILAGFQAGLETILV (30)
PNPP (Sp) 17 FDVFLFDCDGVLWSG (138) FQYLQDPNCAFLLTNQDST (18) IFSTGRQPKILGKPYDEM-MEAIIANV--NFDRK-KACFVGDRL---NIDIQFAKNSNL----- (39)
EYA (Dm) 492 ERVFVWDLDETLIIF (154) MISQRENCVNVLVTSTQLA (12) GIFNIENIYSAHKIGHETCYERIVTRFG---RKS-TYVVIGDGN----EEETAAKAMNFPFWR- (19)
PMM (Sc) 12 ETLVLFDVDGTLTPA (13) KLRNKCCIGFVGGSDLSKQ (124) QISFDVFPAGWDKTYCLQ----HVEKDGF-----KEIHFFGDKTMVGGNDYEIFVDERTIGHSV (16)
COF (Ec) 1 ARLAAFDMDGTLLMP (12) ARLRERDITLTFATGRHAL (129) TDCLEVLPVGCNKGAALT---VLTQHL--GLSLR-DCMAFGDAM----NDREMLVSVGSGFIMG (42)
SERB (Ec) 109 PGLLVMDMDSTAIQI (66) LKLETLGWKVAIASGGFTF (27) GNVIGDIVDAQYKAKT---LTRLAQEY--EIPLA-QTVAIGDGA----NDLPMIKAAGLGIAYH (32)
GOB1 (Ce) 156 KPLFITDWDGTMKDY (18) VFSRLFTRAFAVLTAGPLR (127) EVCVHSSGAVWNKGDGVA---AIVESFHDSLKVG-KVCVAGDTA----SDVPMLKKAGDENPEN (48)
NP_662202 (Ct) 163 FNAFITDRDGTTNNY (18) RFAKNCCRYPMIVTSAPLK (131) VDPKTGMIRDFDKGDGL---EFICRKMNIDHTGE-PVLVCGDTS----SDIPMLKKAMEMYDDV (40)
TPS1 (Ce) 974 YNTFFTDRDGTLKSY (18) QFARRCATFCAIVTTAPLL (127) YVKAKLKGRIFNKGHGIR---LVRERLKPNMSKG-NCLVCGDNE----SDIPMLEECLKLAGSK (85)
TPS2 (Ce) 982 FHYFFTDRDGTLKSY (18) QFARRCAQTCVILTTAPLM (128) YMKTETDGHVFDKGDGLR---LLCEKMHCDLTEG-NVLVCGDSS----TDIPMLKECLIRNPKG (62)
OTSB (Ec) 13 KYAWFFDLDGTLAEI (18) LLATASDGALALISGRSMV (97) KCVVEIKPRGTSKGEAIA---AFMQEA--PFIGR-TPVFLGDDL----IDESGFAVVNRLGGMS (50)
TPS1 (Dm) 529 KLALLLDYDGTLAPI (18) KLSNHSDVYVAVISGRNVD (100) CALEARPPVQWNKGRASI--YILRTSFGVDWNERIKIIYVGDDL----TDEDAMVALKGMARTF (70)
TPS2P (Sc) 570 RRLFLFDYDGTLTPI (18) KLCADPHNQIWIISGRDQK (120) EIVKRLVWHQHGKPQDM--LKGISEKLPKDEMPD-FVLCLGDDF----TDEDMFRQLNTI---- (201)
SPS (So) 772 IFVLALDCDVT-SDL (11) GEQRPTGSISFILSTSMTL (128) GTRLNVIPVLASRSQALR---YLFMRWGV--ELSNFVVFVGESG---DIDYEGLLGGVHKTV-- (58)
consensus 1 hhhhDXDXT hhhT K hhhhGD D
S E

Rayapureddi et al., 2003; Tootle et al., 2003). In both flies and mammals, the respective Eyes absent (Eya) and Eya3 proteins were shown to have HAD-like hydrolase domains, which were demonstrated to have protein phosphatase activity. Since this was the only known case of a HAD-like hydrolase domain being involved in organ development, attempts were made to find a similar phosphatase activity for the GOB-1 protein. A full length GOB-1::MBD protein was expressed in bacteria, purified and tested for phosphatase activity on a variety of substrates. Initial *in vitro* assays tested for the ability of GOB-1 to remove the phosphate group from the p-nitrophenyl phosphate (pNPP) substrate. Under alkaline conditions, the concentration of p-nitrophenol and therefore phosphate release can be determined by the OD reading of the quenched reaction solution at a wavelength of 410 nm (McComb et al., 1979). Two µg of GOB-1::MBD protein was tested for pNPP phosphatase activity and phosphate release was measured after 0.5, 1 and 3 hours. Three different pH conditions were used (5.0, 6.0 and 7.0), as were four different protein re-folding controls. Only a very modest phosphatase activity was detected using the pNPP substrate and the above reaction conditions. The highest activity detected was for one of the refolded proteins and the pH conditions of 5.0. For these conditions, a kcat of 2.40×10^{-4} min^{-1} was calculated. The original protein directly purified from the amylose column showed a very similar activity under the same pH conditions (kcat of 3.02×10^{-4} min^{-1}). This kcat defined a weak GOB-1 phosphatase activity for the pNPP substrate, but was similar to the activity obtained (kcat of 0.017 min^{-1}) by Tootle et al. (2003) for the mouse Eya3 protein also using the pNPP

substrate.

In an attempt to increase the detected phosphatase activity of GOB-1, several different substrates were tested. A P^{32} labelled phosphorylase a protein was used as a substrate to test for GOB-1 protein phosphatase activity. 17.8 µg of the GOB-1::MBD protein and 1/10, 1/100,1/1000 and 1/10000 serial dilutions of this protein solution were tested for phosphatase activity in a standard phosphorylase a phosphatase protocol (Mackintosh and Moorhead, 1999; Stubbs et al., 2001). Under all conditions, no phosphate release was detected above background (data not shown), therefore phosphorylase a was not a GOB-1 substrate.

The conserved sequence relationship of GOB-1 to the C-terminal portions of the *C. elegans* TPS-1 and TPS-2 proteins prompted the testing of phosphorylated sugars as substrates. Although the C-terminal domains of TPS-1 and TPS-2 were not predicted to be trehalose-6-phosphate phosphatases (TPP), several pieces of evidence suggest that this in fact may be the case. First, the inability to detect obvious TPP domains in the worm (Pellerone et al., 2003) and the observation that the C-terminal domains of these proteins were likely HAD-like hydrolase superfamily members, suggested that these domains may be a diverged TPP domain. Second, the components of the trehalose synthesis pathway in yeast and flies tend to be associated in a tight protein complex or located on the same protein respectively (Bell et al., 1998; Chen et al., 2002; Reinders et al., 1997; Singer and Lindquist, 1998; Soto et al., 2002; Vandercammen et al., 1989). The location of the HAD-like domain adjacent to

the trehalose synthase domain suggested that the proximity of these domains might have a functional purpose. A modified malachite green assay was used to detect the release of free phosphate from various sugar phosphates (Cogan et al., 1999). GOB-1 was found to have a robust phosphatase activity with the trehalose-6-phosphate as substrate (i.e., Km 136 μM and kcat 7.77 s^{-1}) **(see Figure 32)**. The optimal pH conditions of the *in vitro* reaction was 7.0 and the removal of Mg^{2+} ions from the reaction almost completely eliminated the GOB-1 phosphatase activity **(see Figure 33)**. The requirement of Mg^{2+} or a similarly charged ion for catalytic activity is a property common to most HAD-like hydrolases (Cho et al., 2001). Boiling the GOB-1 protein prior to the reaction was able to reduce but not completely eliminate enzymatic activity (data not shown). This may not be that unusual in that the strong protein stability of other TPP proteins under conditions of high heat and detergent has been detected (Klutts et al., 2003). Other substrates were tested for phosphate removal including the phosphorylated amino acids O-phospho-L- serine and O-phospho-L-tyrosine, the phosphorylated disaccharide sucrose-6-phosphate, the phosphorylated monosaccharides D-glucose-6-phosphate, D-mannose-6-phosphate and the phosphorylated pentose D-ribose-5-phosphate **(see Figure 34)**. None of the above mentioned substrates resulted in a detectable phosphate release when incubated with GOB-1 under optimal reaction conditions. From these *in vitro* results, it can be concluded that GOB-1 is a robust trehalose-6-phosphate phosphatase and this substrate is the only one found thus far with a

Figure 32: GOB-1 is a robust trehalose-6-phosphate phosphatase. Five different concentrations of T6P were used with three different concentrations of GOB-1::MBD to determine the velocity of the reaction under each condition. Two trials were performed with replicate readings giving a total of four readings for each condition. The resulting four velocity calculations were averaged and the error bars represent the standard deviation of each condition. The average Km was estimated to be 136 μM. The V_{max} was 0.0177, 0.0314 and 0.0524 μM/s for 0.05, 010 and 0.15 μg of protein respectively. Average K_{cat} was 7.77 s^{-1}.

Lineweaver-Burk Plot
1/ [trehalose-6-phosphate] vs. 1 / Velocity of Formation of Free Phosphate
Three Different Concentrations of GOB-1::MBD Protein

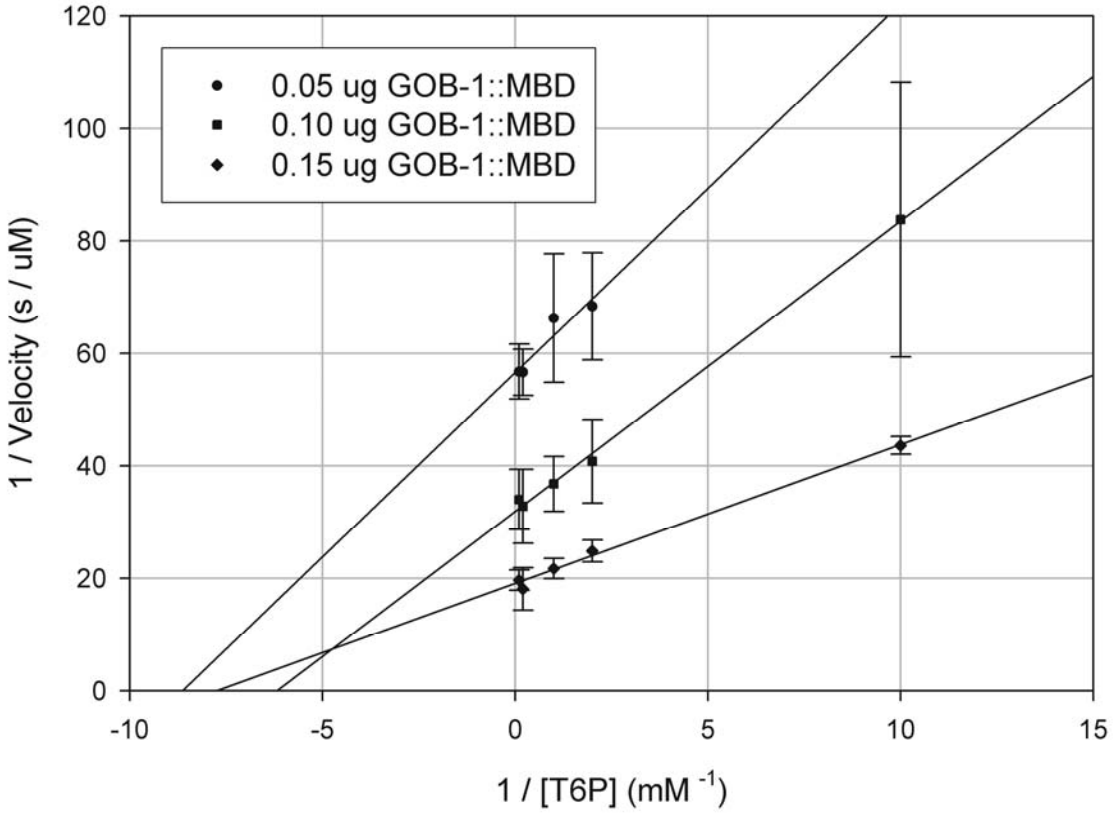


Figure 33: pH and Mg^{2+} dependence of GOB-1 phosphatase activity. 0.10 µg of GOB-1::MBD protein and 2.5 mM of T6P substrate was used in the malachite green assay. The liberation of free phosphate was detected after the reaction was allowed to proceed for ten minutes. Two trails were done with a replicate of each reaction. Error bars represent standard deviation of the mean of the four readings taken.

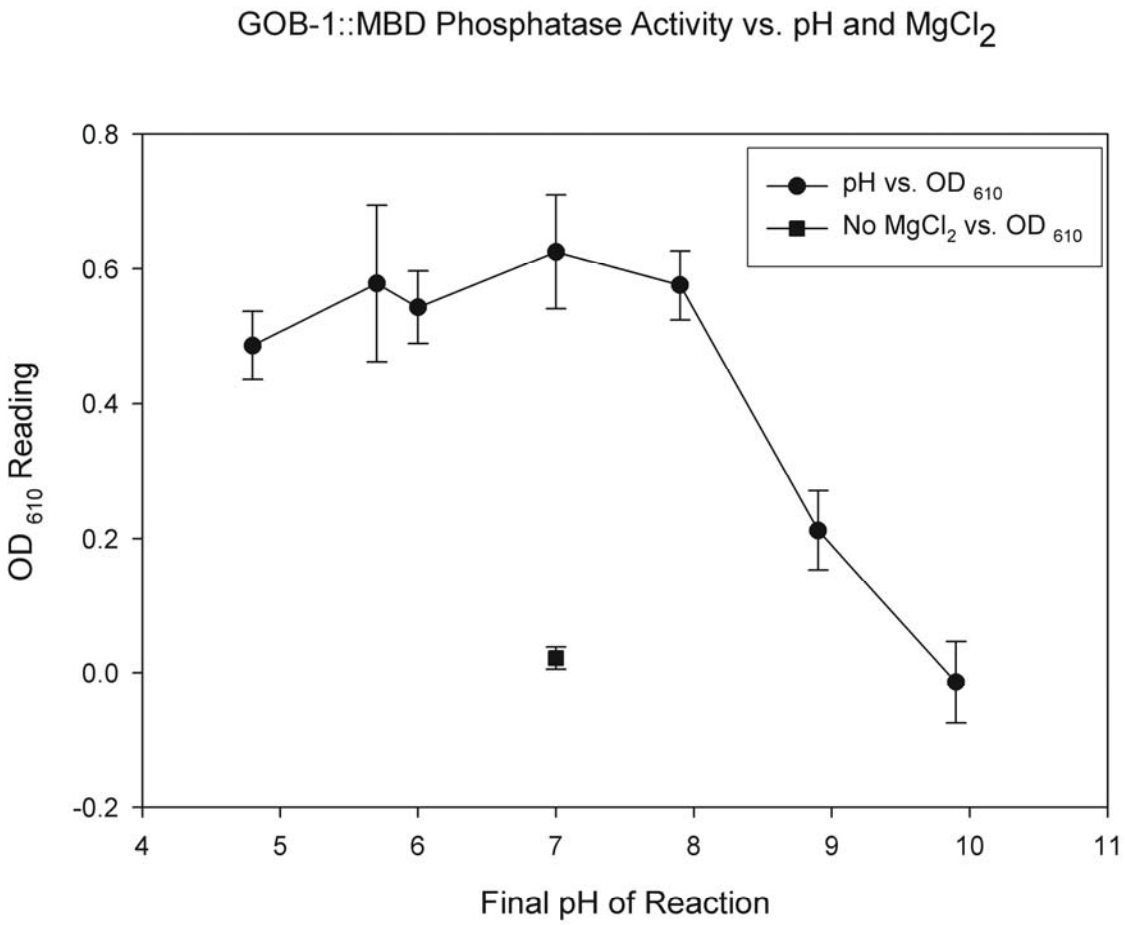
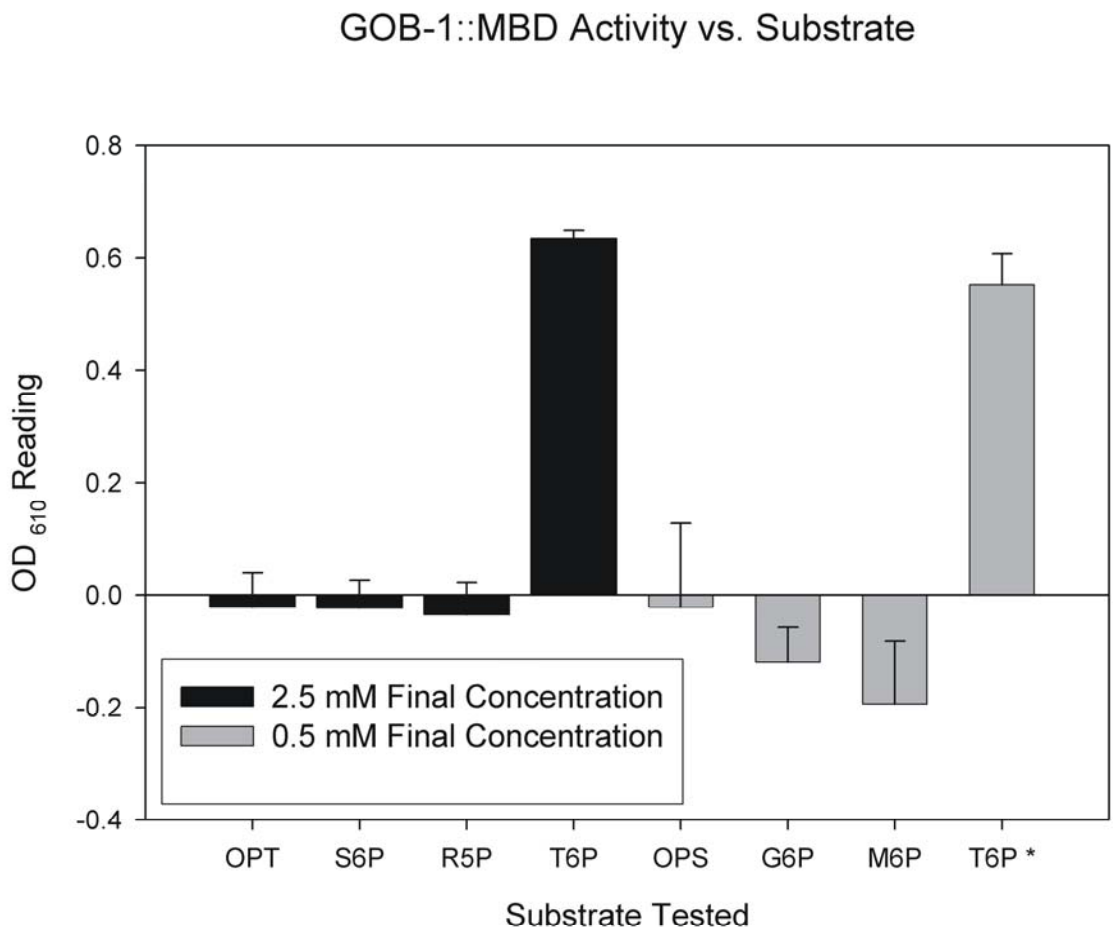

GOB-1::MBD Phosphatase Activity vs. pH and MgCl2
pH vs. OD 610
No MgCl2 vs. OD 610
OD 610 Reading
Final pH of Reaction
0.8
0.6
0.4
0.2
0.0
-0.2
4
5
6
7
8
9
10
11

Figure 34: GOB-1::MBD phosphatase activity is specific for trehalose-6-phosphate. Six other phosphorylated substrates were tested for the ability of GOB-1 to remove their phosphates. Two trials were done in replicate. The readings were averaged and the error bars represent the standard deviation. The following substrates were tested. OPT (O-phospho-L-tyrosine), S6P (sucrose-6-phosphate), R5P (ribose-5-phosphate), T6P (trehalose-6-phosphate 2.5 mM), OPS (O-phospho-L-serine), G6P (glucose-6-phosphate), M6P (maltose-6-phosphate) and T6P* (trehalose-6-phosphate 0.5 mM).


GOB-1::MBD Activity vs. Substrate
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
OD 610 Reading
2.5 mM Final Concentration
0.5 mM Final Concentration
OPT
S6P
R5P
T6P
OPS
G6P
M6P
T6P *
Substrate Tested

significant level of activity.

gob-1 Function in the Trehalose Synthesis Pathway

The amino acid similarity of GOB-1 to the C-terminal domains of the *C. elegans* TPS-1 and TPS-2 proteins and its substrate preference for trehalose-6-phosphate suggested that GOB-1 might have a role in trehalose synthesis. Previous attempts at knocking down *tps-1* and *tps-2* function with RNAi feeding, either individually or simultaneously, produced no obvious phenotypic effect (Pellerone et al., 2003), despite trehalose levels being dramatically reduced (lowered to 7% of wild-type levels) in the *tps-1* and *tps-2* double mutant (Pellerone et al., 2003). The above RNAi experiments were performed via RNA feeding and thus the failure of being able to detect a phenotype could be due to an incomplete knockdown of *tps-1* and *tps-2* activity (Kamath et al., 2003). RNAi knockdown of *tps-1* and *tps-2* was repeated using injection of dsRNA into the gonad. With the injection technique rather than the feeding method, a more robust loss of function effect could be achieved. When *gob-1* dsRNA was injected alone, 47.6% of the F1s (n = 396) arrested as larvae **(see Table 5)**. When *tps-1* and *tps-2* dsRNA were injected in combination only 8.3% of the F1 population (n = 682) arrested as larvae. When *gob-1*, *tps-1* and *tps-2* dsRNAs were injected together, 28.6% of the F1s (n = 538) arrested as larvae. From these results it was concluded that a stronger reduction in function of *tps-1* and *tps-2* did not result in a lethal phenotype. It could further be concluded that the

RNAi targeted knockdown of members of the trehalose synthesis pathway

	24 hrs			48 hrs		
	Total F1	Arrested L1	Dead Eggs	Adults and L4 larvae	Arrested L1	Arrested L2/L3
tps-2 (1000 µg/mL)	551	54 (9.8%)	25 (4.5%)	444 (81.9%)	29 (5.4%)	69 (12.7%)
tps-1 (1000 µg/mL)	345	55 (15.9%)	49 (14.2%)	634 (83.7%)	57 (7.5%)	65 (8.6%)
tps-1 and *tps-2* (500 µg/mL each)	288	24 (8.3%)	5 (1.7%)	682 (91.8%)	16(2.2%)	45 (6.1%)
gob-1 (500 µg/mL)	698	99 (14.2%)	37 (5.3%)	396 (52.4%)	53 (7.0%)	307 (40.6%)
gob-1, *tps-1* and *tps-2* (500 µg/mL each)	747	110 (14.7%)	29 (3.9%)	538 (71.4%)	66(8.8%)	149 (19.8%)

Table 5: *gob-1*, *tps-1* and *tps-2* were targeted for RNAi knockdown. Hermaphrodites were injected with dsRNA and their F1 progeny were assayed for larval arrest and embryonic arrest at 24 hours and 48 hours after being laid. The triple mutant does not rescue the *gob-1* loss of function larval arrest phenotype.

triple gene knockdown did not result in an enhanced *gob-1* phenotype (see CHAPTER SIX - Discussion for pathway implications of these results). With the triple RNAi, *gob-1* loss of function lethality was reduced but it is difficult to conclude if this is a real genetic interaction or a reduced loss of function effect due to the attempt of knocking down three genes simultaneously (Kamath et al., 2003). Genetic analysis with null mutations of these genes would allow for a firmer conclusion about their possible genetic interactions.

To examine further the role of *gob-1* in trehalose metabolism, attempts were made to see if trehalose supplemented media could rescue the *gob-1* mutant phenotype. Hermaphrodites were injected with *gob-1* dsRNA (1000 µg/mL) and ~12 hours later were allowed to lay eggs on standard NGM feeding plates, 0.1 M sucrose supplemented plates and 0.1 M trehalose supplemented plates. F1 progeny were examined for dead eggs and the stage of larval arrest at 24 hours and 48 hours after being laid **(see Table 6)**. Overall, trehalose was not able to rescue the lethality of the *gob-1* RNAi mutants. There did appear to be a shift from L1 arrest to L2 arrest on the trehalose supplemented plate compared to the unsupplemented and sucrose supplemented plates but it was difficult to conclude that this shift in arrest stage was trehalose specific. The bacterial lawns on the trehalose supplemented plates grew much faster than the lawns on the other two conditions. With the thicker lawns, arrested L1s were more difficult to detect. The increased nutritional state of the bacterial lawn could potentially allow for a second molt and this change in nutritional state would be unrelated to any requirement for trehalose. There was the additional

Trehalose supplemented medium does not rescue the *gob-1* loss of function early larval arrest

	Total F1	Adults	L4/L3	L2	L1	Dead Eggs
Uninjected N2 NGM 24 hours	1058	-	-	-	21 (0.02)	20 (0.02)
NGM 24 hours	561	-	87 (15.5)	269 (48.0)	56 (10.0)	-
NGM 48 hours		81 (14.4)	62 (11.1)	253 (45.1)	48 (8.6)	83 (14.8)
NGM + Trehalose 24 hours	389	-	63 (16.2)	275 (70.7)	4 (1.0)	24 (6.2)
NGM + Trehalose 48 hours		62 (15.9)	81 (20.8)	254 (65.3)	4 (1.0)	-
NGM + Sucrose 24 hours	659	-	67 (10.2)	312 (47.3)	90 (13.7)	112 (17.0)
NGM + Sucrose 48 hours		68 (10.3)	21 (3.2)	341 (51.7)	54 (8.2)	-

Table 6: Hermaphrodites were injected with *gob-1* dsRNA. F1 progeny were laid on NGM plates supplemented with trehalose or sucrose and assayed for stage of larval arrest and embryonic arrest at 24 hours and 48 hours after being laid. The uninjected strain was also tested for background lethality. The number of progeny listed in each category is followed by the percent of their brood population that they represent. The presence of excess trehalose does not rescue the larval lethality caused by *gob-1* loss of function mutant.

consideration that it is currently unknown if trehalose can be directly absorbed by the worm. Since the *gob-1* mutant larvae still had an early larval arrest phenotype, the trehalose supplementation experiments were not pursued further.

Lectin Experiments

The intestinal glycocalyx is a network of protein-bound sugar chains located on the apical surface of endodermal epithelial cells (Gebhard and Gebert, 1999). The glycocalyx in the mouse intestine has been shown to mediate intestine cell interactions with substances located in the intestine lumen. A similar glycocalyx structure has been detected in *C. elegans* (Leung et al., 1999). It was postulated that since GOB-1 had a potential role in trehalose metabolism it might have a role in synthesis of sugar-based structures in the cell. Defects in the construction of the *gob-1* mutant glycocalyx could result in a loss in glycocalyx-mediated signalling and other lumen surface properties such as protection of the apical membrane or regulation of apical surface adhesion. Such mentioned defects could explain the lumen defects seen in *gob-1* mutants.

Lectins are proteins that are able to bind with high specificity and affinity to various carbohydrate based structures. Fluorescently labelled lectins have been used to define specific carbohydrate deposition patterns in various tissues of the worm (Borgonie et al., 1994; Borgonie et al., 1997). A series of 10 FITC conjugated lectins (BPA, Con A, DBA, GSI, GSII, MPA, PNA, SBA, UEAI and

WGA) were used to detect carbohydrate distribution patterns in early larvae. Five of these lectins, (SBA, DBA, PNA, GSI and BPA), were found to detect carbohydrates in the intestine **(see Figure 35)**. Control wild-type larvae and larvae exposed to *gob-1* RNAi were stained in parallel and compared for differences in lectin distribution. No differences between control and *gob-1* RNAi lectin staining patterns were detected. Thus it appears that *gob-1* is not necessary for at least a subset of carbohydrate-derived intestine structures or the gross maintenance of lumenal carbohydrates. However, from these pilot experiments, it cannot be ruled out that *gob-1* is not necessary for the function of an undetected carbohydrate structure.

gob-1's Relationship to *lin-12*

Animals defective in the LIN-12/Notch signalling pathway are defective in the rotation of rings Int II through IV during intestine development (Hermann et al., 2000). Because predominant intestinal kinks were seen in a subset of *gob-1* mutant animals, it was examined whether these twists were due to the failure of intestine rotation in *gob-1* mutant animals. *lin-12* dependent intestine rotation is almost complete at the 1½-fold stage of embryogenesis. This stage was used to quantify the degree of intestine rotation in wild-type (n = 12) and *gob-1* (n = 11) mutant animals. An integrated *elt-2* promoter driven GFP construct was used to visualize the intestine nuclei in the wild-type and mutant animals (Fukushige et al. 1999). In retrospect, intestinal nuclear position alone was not the best

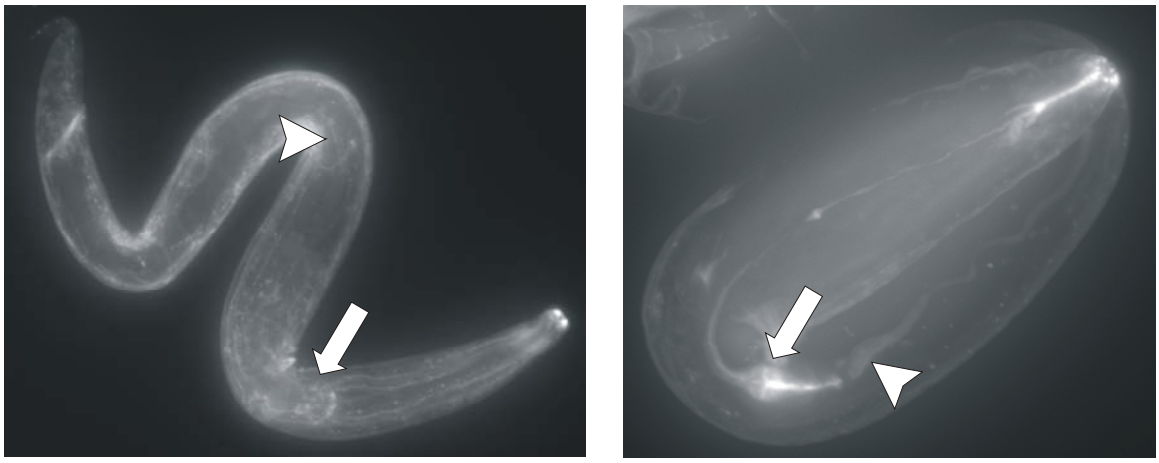


Figure 35: Lectin staining. Typical staining patterns for the PNA FITC conjugated (left) and the GSI FITC conjugated (right) lectins in the wild-type pretzel embryo. The PNA-FITC lectin is able to detect a carbohydrate that lines the pharynx (white arrow) and intestine lumen (white arrowhead). Staining was also found in the head and along the body wall. The GSI-FITC lectin detects a carbohydrate lining the pharynx (white arrow) and intestine lumen (white arrowhead). No significant changes in any lectin staining pattern was detected in the *gob-1*(RNAi) embryos (data not shown).

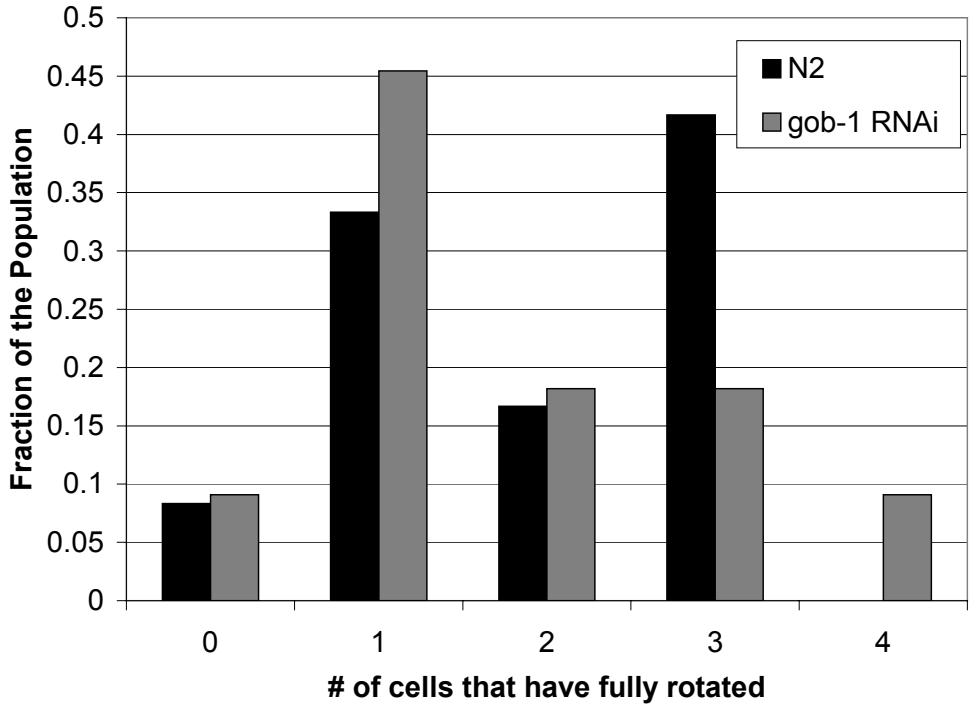
indicator of ring rotation. Nuclear position seemed to lag behind the cellular movements and due to this lag considerable variability in the extent of rotation was seen in both wild-type and the *gob-1* mutants. However, despite the detected variability within each group, no obvious difference in intestine rotation was seen between the two groups **(see Figure 36)**. Therefore, the *lin-12* associated intestine rotation was not responsible for the lumenal twists seen in the *gob-1* loss of function mutant.

elt-2 is Sufficient But Not Necessary for *gob-1* Expression

The basic premise behind the Gob screen and the identification of genes that have a necessary function in intestine development was to find genes redundant with *elt-2* or downstream effectors of *elt-2*. Three experiments were done to test the relationship between *elt-2* and *gob-1*. The first experiment tested whether *elt-2* was necessary for *gob-1* expression. The *gob-1* transcriptional reporter strains were stained for β-galactosidase expression in either a wild-type or *elt-2* RNAi mutant background. Under both conditions, embryos of all stages were examined for the presence of β-galactosidase staining in the intestine. In one reporter strain, 67% of wild-type embryos had intestine staining (n = 152), while 72% of the *elt-2* RNAi affected animals (n = 170) also had intestine staining. Similar numbers were found for an independently derived transgenic strain. Since the transcriptional reporter used above was present as an extra-chromosomal array, ratios of staining vs.

Figure 36: Hermaphrodites with an integrated *elt-2*-promoter-driven GFP, were injected with *gob-1* dsRNA (1000 µg/mL). Two pools of *gob-1*(RNAi) mutant 1½-fold F1 embryos were collected. The first pool was predicted to be laid at about 8½ hours after injection and the second pool was predicted to be laid about 14 hours after injection. A wild-type control was done in parallel. Live embryos were mounted for microscopy and pictures were taken. Z projections of stacks were used to assess the degree of twist progression. Nuclei that were positioned at a 90˚ rotation with reference to Int 1 and Int V were scored as having completed their rotation.

The lumenal twists seen in *gob-1* loss of function mutants are not due to a defect in *lin-12* signalling in the intestine



non-staining animals were similar to those expected for loss of the transgenic array. From these experiments, it can be concluded that *elt-2* is not necessary for early *gob-1* reporter expression in the intestine of embryos.

The second experiment tested whether ELT-2 was sufficient to drive expression of *gob-1*. Double transgenic lines were created that carried either an integrated heat-shock promoter-driven ELT-2 or ELT-1 protein together with the extra-chromosomal *gob-1* promoter-driven GFP/*lacZ* reporter (Fukushige et al., 1998; Gilleard and McGhee, 2001). ELT-1 is a *C. elegans* GATA transcription factor expressed in the hypodermis and was used as a control for GATA factor specificity following over-expression. The doubly transgenic strains were exposed to heat-shock conditions and assayed for the number of nuclei with β-galactosidase activity. Optimal conditions for both ELT-1 and ELT-2 over-expression were used. The ability to drive ectopic expression of the transcriptional reporter was quantified via counting of the nuclei in the embryo that had β-galactosidase activity present. On average, ELT-2 ectopic expression caused more embryonic nuclei to express the transcriptional reporter than did the ELT-1 control, even though embryos from both treatments arrested at approximately the same stage **(see Figure 37 and 38)**. Heat-shocked induced over-expression of ELT-2 caused embryos to arrest with an average of 241 cells (standard deviation ± 31.9) and heat shock driven over-expression of ELT-1 caused embryos to arrest with an average of 234.5 cells (standard deviation ± 35.4). Under each of the heat-shock conditions, the degree of intestine differentiation was monitored using *ges-1* staining **(see Figure 39)**. Over-

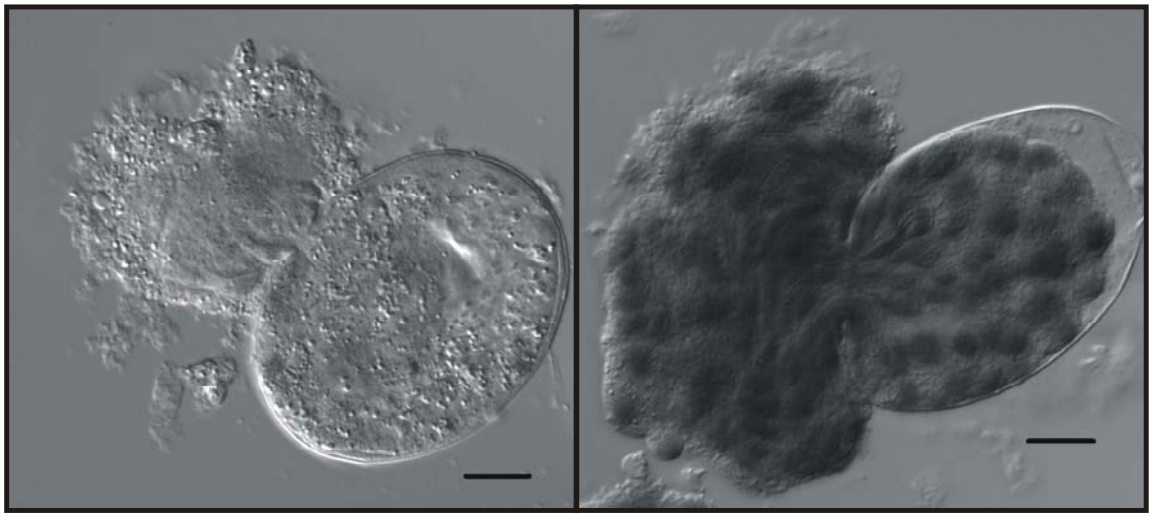


Figure 37: Ectopic expression of ELT-2 can drive ectopic expression of *gob-1*. Embryos possessing a *gob-1*-driven transcriptional reporter were exposed to ectopic expression of ELT-2 or a second GATA transcription factor ELT-1. Representatives of typical embryos over-expressing ELT-1 (left) and ELT-2 (right) are shown above. Ectopic ELT-2 expression is able to force expression of the *gob-1* transcriptional reporter in most cells of the embryo. Ectopic ELT-1 expression is unable to reproduce this effect. Scale bar 10 μm.

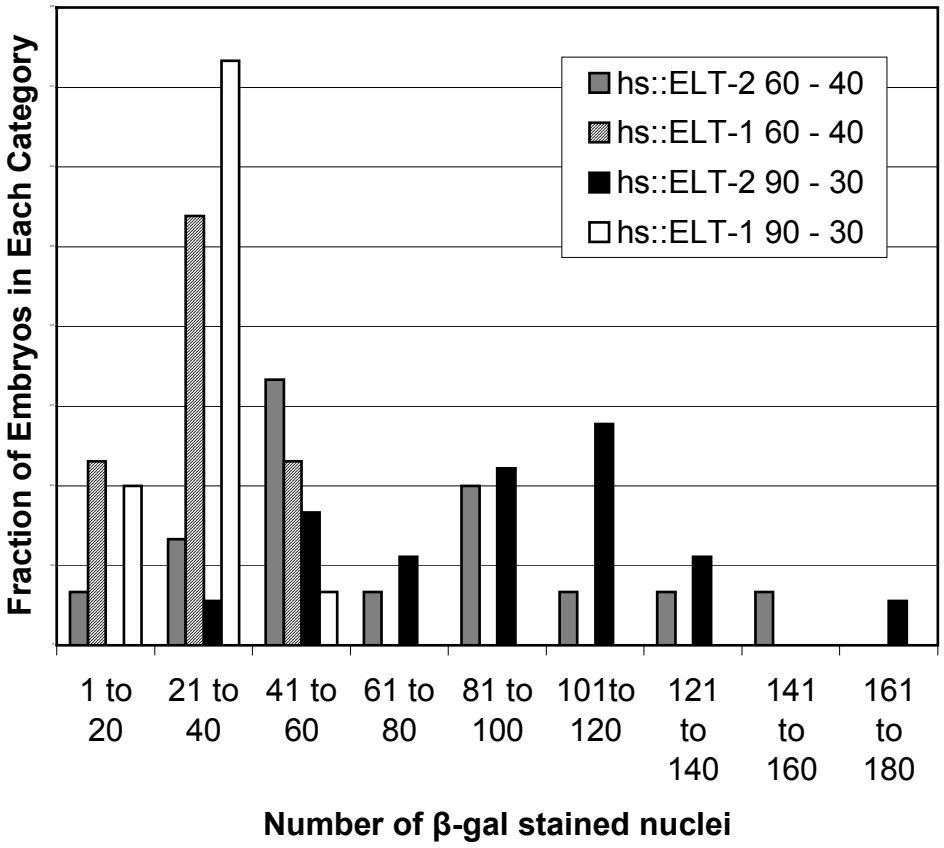
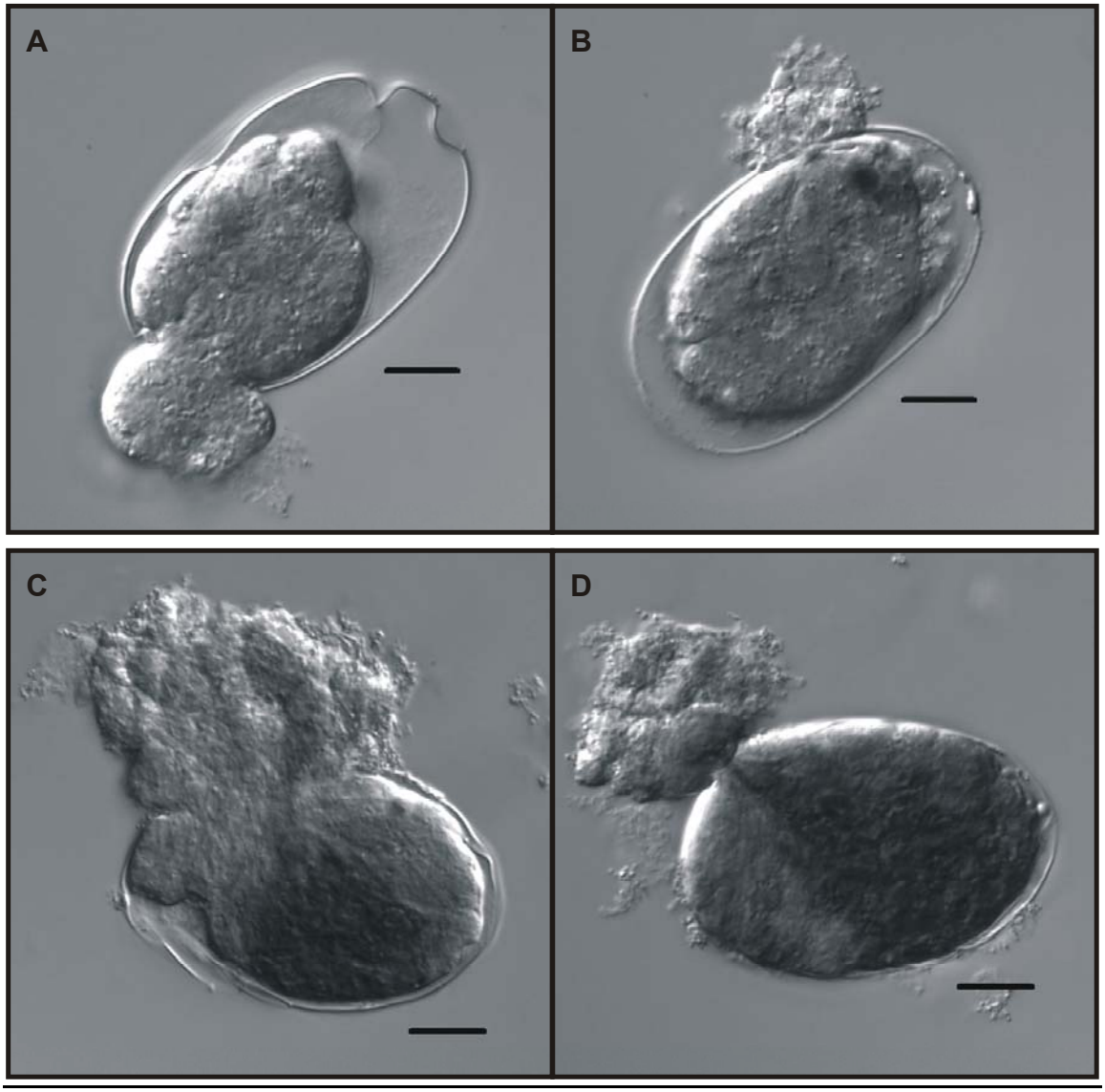


Figure 38: Ectopic ELT-2 expression causes more cells to express *gob-1* than ELT-1 controls. Animals that were transgenic for the heat-shock promoter-driven ELT-2 or ELT-1 protein and the *gob-1* promoter-driven GFP/*lacZ* reporter, were exposed to heat-shock conditions and assayed for the number of nuclei with β-galactosidase activity. Optimal conditions for both ELT-1 (60 minutes at 20˚C and 40 minutes at 30˚C) and ELT-2 (90 minutes at 20˚C and 30 minutes at 30˚C) were used. Results from two independently derived transgenic lines were used in the data presented.

Figure 39: ELT-2 ectopic expression results in ectopic differentiation of intestine. Staining for *ges-1* activity was done on ELT-1 and ELT-2 over-expressing embryos to assess the degree of intestine differentiation under these conditions. (A) A representative of an ELT-1 over-expressing embryo with weak generalized staining. (B) A representative of an ELT-1 over-expressing embryo displaying a moderate localized stain. (C) An ELT-2 over-expressing strain displaying a strong staining in at least half of the embryo. (D) An ELT-2 over-expressing embryo displaying a generalized strong staining. In an experiment where ELT-1 was over-expressed under heat-shock conditions ideal for ELT-1 over-expression, 4/10 embryos had a moderate localized stain and 6/10 embryos had a weak generalized stain. In a second experiment where ELT-1 was over-expressed under heat-shock conditions ideal for ELT-2, 2/6 embryos had a moderate localized staining and 4/6 had a weak generalized staining. Over-expression of ELT-2 under conditions ideal for ELT-1 over-expression caused 7/10 embryos to have a strong generalized staining and 3/10 to have a strong staining covering at least half of the embryo. Over-expression of ELT-2 under conditions ideal for ELT-2 over-expression caused 5/9 embryos to have a strong localized stain and 4/9 embryos had a strong staining in greater than half of the embryos. The presence or absence of the *gob-1* transcriptional reporter did not influence the degree of *ges-1* staining. Scale bar is 10 μm.


A
B
C
D

expression of ELT-2 caused the majority of the cells in the embryo to express *ges-1*. Over-expression of ELT-1 resulted in either no *ges-1* staining or a small localized patch of *ges-1* staining. This is expected since ELT-1 is known to be a weaker inducer of differentiation than ELT-2 (Fukushige et al., 1998; Gilleard and McGhee, 2001). Further controls were conducted to ensure the specificity of the over-expression results. The heat-shocked ELT-1 (n = 20) and ELT-2 (n = 45) over-expressing strains without the transcriptional reporter had no detectable β-galactosidase activity. In a separate experiment, the *gob-1* transcriptional reporter strain alone was heat-shocked under similar conditions. Out of approximately 100 heat-shocked embryos, 8 of these arrested due to the heat shock conditions. Five out of these eight arrested embryos could have β-galactosidase stained nuclei counted and of these an average of 33 nuclei could be detected (range of 26 to 40). It is expected that the majority of the staining cells in this control are intestine cells. Most of these eight arrested embryos displayed partial morphogenesis of the hypodermis and pharynx that was not evident in either the ELT-2 and ELT-1 over-expressing strains, suggesting a later stage of arrest. Despite a later stage of arrest, fewer staining cells were detected in the control than in the heat-shocked ELT-2 over-expressing strain. This suggested that the large number of staining cells under the ELT-2 over-expressing conditions were not due to a late arrest and subsequent expression of the reporter in non-intestinal cells, such as was detected in later embryos of the original reporter strain (see above). Overall it was concluded that ELT-2 driven ectopic expression of *gob-1* was a true over-expression phenomena and that

ELT-2 was sufficient to drive *gob-1* expression. The third experiment further tested the genetic relationship between *gob-1* and *elt-2*. RNAi knockdown experiments were used to create *gob-1* and *elt-2* simultaneous loss of function mutants **(see Table 7)**. The *elt-2*(RNAi) mutant alone caused the majority of the F1 population to arrest at the early L1 larval stage. The *gob-1*(RNAi) mutant alone caused the majority of the F1 population to arrest at the L2 larval stage. When *elt-2* and *gob-1* functions were knocked down simultaneously, the majority of the animals arrested as L1 larvae. With the above results, it can be concluded that the *elt-2* loss of function phenotype is epistatic to the *gob-1* loss of function phenotype. The earlier arrest of the *elt-2* mutants is perhaps not surprising, since *elt-2* is likely required for the expression of many genes required for intestine function (Fukushige et al., 2003). The fact that the *elt-2* loss of function phenotype is not enhanced by *gob-1* loss of function (e.g., enhanced embryonic arrest), suggests that the two genes are not acting in parallel pathways during early intestine formation. In the double knockdown situation there was an increase in the proportion of L4 larvae in the arrested F1 population. The arrested L4 larvae present in the double RNAi experiment are likely to be RNAi escapers. The degree of RNAi mediated knockdown tends to be reduced when two genes are targeted simultaneously (Kamath et al., 2000). *gob-1* and *elt-2* knockdown seem to be especially sensitive to the concentration of dsRNA, as an increase in the concentration dsRNA during the knockdown of each gene increases the penetrance of their loss of function phenotype.

The *elt-2*(RNAi) mutant is epistatic to the *gob-1*(RNAi) mutant

RNAi	Total F1	Adults	L4	L3	L2	L1	Dead Eggs
elt-2 (n=11) 24 hr	988	-	-	47 (4.8)	86 (8.7)	818 (82.8)	37 (3.7)
elt-2 48 hr	467	27 (5.8)	3 (0.6)		22 (4.7) (L2/L3)	415 (88.9)	-
gob-1 (n= 8) 24 hr	707	-	-	74 (10.5)	597 (84.4)	6 (0.8)	30 (4.2)
gob-1 48 hr	504	49 (9.7)	13 (2.6)		427 (84.7) (L2/L3)	15 (3.0)	-
elt-2 + *gob-1* (n= 10) 24 hr	1015	-	-	129 (12.7)	260 (25.6)	589 (58.0)	37 (3.6)
elt-2 + *gob-1* 48 hr	529	64 (12.1)	136 (25.7)		45 (8.5) (L2/L3)	284 (53.7)	-

Table 7: Young adults were injected with 500 µg/mL of *elt-2* and *gob-1* dsRNA either individually or in combination. F1 progeny were scored for stage of arrest at 24 hours and 48 hours post-injection. The number of progeny and the percent of their brood class that they represent (in parentheses) is presented. *gob-1* RNAi does not enhance the lethality associated with the *elt-2* RNAi phenotype. This suggests that *gob-1* does not have an early function redundant with *elt-2*.

CHAPTER SIX - Discussion

Description of the *gob-1* Loss of Function Mutant Phenotype

The *gob-1*(*ca17*) homozygous null mutant and the *gob-1* RNAi loss of function mutant result in an early larval lethal phenotype and a penetrant intestine defect. Based on these loss of function defects, we can conclude that *gob-1* is required for survival and has a critical role in maintaining intestine integrity in the worm. Nomarski optics and the bead feeding assay are able to detect intestine lumen defects immediately after hatching. This suggests that the lumen phenotype observed in *gob-1* mutants is in fact due to a developmental defect and not due to subsequent degradation of the intestine after feeding has begun. The earliest intestine defects can be detected at the comma stage when there appears to be a mildly disorganized localization of IFB-2 close to the apical surface of the intestinal cells. Electron microscopy of cross-sections of freshly hatched *gob-1* mutants detect a mild defect in the electron dense structure (the endotube) located just below the intestine apical surface. The breakdown of this electron dense structure and components of the endotube could account for defects in lumen structure in the *gob-1* mutant.

The intestinal microvilli in the *gob-1* mutant are shorter and less dense than in wild-type controls. The lumenal microvilli rely on the integrity of the meshwork of actin and intermediate filaments located just below the apical surface in order to have a base from which to project into the lumen

(Montgomery et al., 1999). The fact that the microvilli are shorter and less dense in the *gob-1*(RNAi) mutant could be a result of *gob-1*-related defects in the endotube structure. Potentially, the *gob-1* microvillar defect could be a secondary cause of starvation. This scenario seems unlikely due to the fact that, the *gob-1* mutant microvillar phenotype is qualitatively distinct from the phenotype of larvae that had been starved for 12 hours (Fukushige et al., 1998). Furthermore, the animals studied in the present EM experiment had a feeding period that was too short to illict a strong starvation response. The *gob-1*(RNAi) mutant villi phenotype could also result from the absence of bacteria in the lumen. This scenario also seems unlikely. EM cross-sections were taken of *gob-1* mutants that had an open lumen and therefore the likely presence of bacteria in the lumen (data not shown). The microvilli in these few pictures were shorter than those in the wild-type larvae.

gob-1 mutants also have more general intestine defects. The birefringent crystal-like structures present in the intestine of the *gob-1* loss of function mutants are distinct from normal gut granules and from granules seen in starved larvae. Birefringent granules are often associated with digestive organelles and this suggests that there may be a defect in these organelles in the *gob-1* mutant (Kostich et al., 2000). *gob-1* mutant animals arrest even with an open lumen and bacteria obviously inside suggesting that embryos are not dying solely because of the inability to feed. In later embryos, there is often the presence of lipid-like droplets in the head region. This is often associated with starvation or aging in wild-type animals and suggests that *gob-1* mutant embryos are in a

starved/stressed state prior to hatching (Herndon et al., 2002). The intestine is the major metabolic, digestive and stress responding organ of the worm. All of the above phenotypes suggest that there may be a metabolic or digestive defect in the intestine of the *gob-1* mutants.

From the EM and from the IFB-2 and AJM-1 immunohistochemistry, at least some potential causes of the *gob-1* defect can be eliminated. First, despite defects in some apical structures, the general apical/basal polarity of the epithelium has been maintained. This is evident by a normal adherens junction structure and the proper deposition of intermediate filaments and other apical determinants at the apical surface of the *gob-1* mutant. A second conclusion is that the *gob-1* dependant defects are not as severe as those of the *elt-2* mutant. Electron microscopy cross-sections do not detect the severe lumen malformations detected in the *elt-2*(*ca15*) homozygous null mutants (Fukushige et al., 1998). This suggests that the gut obstructions detected in the *gob-1* mutants may be occurring transiently along the length of the lumen. Alternaltively, loss of lumen surface adhesion may not be entirely complete even though the lumen and apical structures appear to be correctly formed.

gob-1 is expressed at the right time and place to be involved in intestine development. The *gob-1* promoter-driven reporter is first expressed in the intestine at the 8 E cell stage. Strong expression remains in the intestine for the life of the worm. The GOB-1 protein (albeit resulting from transgene over-expression) is detected in an identical tissue and temporal pattern as the transcriptional reporter. The GOB-1 protein is consistently cytoplasmic with

possible enrichment at the cell cortex, suggesting an enzymatic function in the cytoplasm or at the cell membrane. The sporadic detection of GOB-1 in the nucleus suggests it may have a role in the nucleus and potentially in gene transcription. GOB-1 expression outside the intestine begins at the comma stage. GOB-1's early expression in the intestine and manifestation of a lethal phenotype in this organ could explain why obvious defects in other organs, which display a later expression pattern, are not observed. Perhaps the use of a temperature sensitive allele could bypass this early lethality and allow a later *gob-1* function to be studied in other organs.

A necessary maternal contribution of *gob-1* seems unlikely but cannot be ruled out. RNAi to *gob-1*, which would result in reduction of both maternal and zygotic *gob-1* function, does not result in a phenotype that is more severe than the *gob-1*(c*a17*) homozygous mutant, which would be defective in only zygotic *gob-1* function. Homozygous animals die before they can produce F1 progeny, making it difficult to study the function of maternally provided GOB-1. It is also currently unknown if GOB-1 is expressed in the early embryo. The transcriptional reporter and transgenic array-driven over-expression of the protein only allows for detection of zygotically expressed genes because of the repression of expression of such constructs in the germline. Given the importance of trehalose in the early embryo (described below), it may also be useful to study *gob-1* maternal function in order to understand the full function of this gene.

Relationship of *gob-1* to *elt-2*

The initial goal of this project was to identify genes that may be acting in parallel to or acting as downstream effectors of *elt-2*. ELT-2 appears to be sufficient but not necessary for *gob-1* expression. This scenario is perhaps not surprising in that *elt-2* has been found to be sufficient but not necessary for expression of several early intestine specific genes (i.e., IFB-2 and *ges-1*) (Fukushige et al., 1998). The fact that *elt-2* is not necessary for expression of *gob-1* means that other factors must be acting redundantly to control *gob-1* expression in the intestine. Despite this loss of *elt-2* function result, ELT-2 is able to ectopically drive *gob-1* expression and this suggests that *gob-1* may potentially be a downstream effector of *elt-2* during later intestine differentiation.

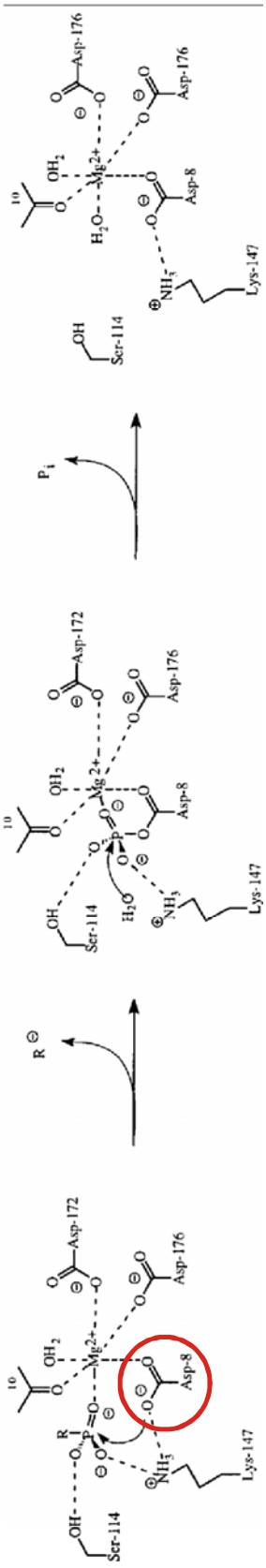
GOB-1 is a Member of the HAD-like Hydrolase Superfamily

GOB-1 is a member of the HAD-like hydrolase superfamily. To make this argument most convincing, it is necessary to review the characteristics that define this family of proteins and how these characteristics apply to GOB-1. Microbial dehalogenases, a group of enzymes that function to remove halogen groups from organic compounds during their detoxification, are the founding members of the HAD (Haloacid dehalogenase)-like hydrolase superfamily of proteins (Janssen et al., 2001). As crystal structures and catalytic mechanisms of numerous phosphatases were discovered, it became apparent that these

proteins employed a very similar catalytic mechanism in order to remove phosphate groups from their substrates (Cho et al., 2001; Collet et al., 1998; Collet et al., 1999; Ichiyama et al., 2000; Morais et al., 2000; Ridder and Dijkstra, 1999). The catalytic mechanism of dehalogenases and phosphatases is believed to involve the formation of an enzyme-substrate phospho-ester intermediate **(see Figure 40)**. This intermediate forms as a result of the nucleophilic attack of an absolutely conserved aspartate residue on the phospho-oxygen bond of the substrate. This intermediate is subsequently released by the hydrolysis of the phospho-ester bond by an activated water molecule, leading to the restoration of the enzyme, the release of a free phosphate molecule and the release of a phosphate free product.

Alignment of conserved amino acids of members of the HAD-like hydrolase superfamily of proteins has been used to define three highly conserved domains (Aravind et al., 1998; Collet et al., 1999; Koonin and Tatusov, 1994; Ridder and Dijkstra, 1999). Mutational analysis of conserved residues and establishment of conserved residue placement in the catalytic core of the enzyme-substrate complex allowed for a function to be assigned to each of the highly conserved residues in each domain. The domain nomenclature defined by Aravind et al. (1998) will be referred to for the remainder of this discussion. The Aravind et al. (1998) study is the most recent and has included the largest sequence sets of Domains I, II and III in their protein comparisons. The first and most highly conserved domain, domain I, is commonly referred to as the DXDX(T/V) domain (also known at the DKTGT domain in ATPases and the

Figure 40: One of the models for the catayltic mechanisim of HAD-like hydrolases. A phosphorylated molecule enters the active site of the hydrolase. In the first step of the reaction, there is a nucleophilic attack on the P-O bond of the phosphorous atom by the first aspartate residue in domain I (red circle). This results in release of the once phosphorylated molecule and a phospho-ester intermediate that is subsquently hydrolysed by attack of an activated water molecule. The bond that is targeted by water is the aspartate-carbon bond in L-2-haloacid dehalogenases and the phosphorous atom in ATPases. This last step results in the release of free phsphate and the restoration of the enzyme. In this particular model the predicted bonding of the serine/threonine residue in domain II, the lysine and two aspartate residues in domain III and the Mg^{2+} molecule is illustrated. Figure taken from Ridder and Dijkstra (1999).


Asp-172
Asp-176
Asp-8
Lys-147
Ser-114
Mg2+
OH2
H2O
NH3
R
Pi

DXYGT domain in dehalogenases). All family members possess an invariant first aspartate residue. In phosphotransferases, it is thought that this first aspartate residue is the nucleophile that acts on the phospho-oxygen bond to form a phospho-ester intermediate. The second aspartate is conserved in all members of the HAD-like super family except L-2-haloacid dehalogenase and epoxide hydrolase, and is thought to be required for the binding of a Mg^{2+} ion in the catalytic core. Also within this domain, there is a highly conserved threonine residue. This threonine residue is thought to stabilize the phospho-aspartate intermediate or orient the first aspartate for nucleophilic attack. The second domain, domain II, is defined by a conserved hydroxyl residue (serine or threonine) and is thought possibly to bind the phosphate group of the substrate. A highly conserved lysine residue followed by a pair of aspartates defines the third and final domain, domain III. The lysine is thought to interact with the phospho-aspartate intermediate. The first aspartate in this domain is thought to bind Mg^{2+}, while the second is thought to activate the water intermediate or stabilize the Mg^{2+}-enzyme complex. A Mg^{2+} ion is a necessary part of the catalytic core for all members of the HAD-like hydrolase superfamily except for the dehalogenases and epoxide hydrolases (Ridder and Dijkstra, 1999). The precise function of the Mg^{2+} ion in the catalytic core is still undefined but is expected to stabilize the catalytic core during substrate binding and the catalysis of the reaction.

Based on sequence alignments of highly conserved residues and their apparent substrates, at least five groupings have been identified for the HAD-like

hydrolase superfamily (Aravind et al., 1998; Koonin and Tatusov, 1994). One group is defined by the haloacid dehalogenases and the epoxide hydrolases. Another group includes members of the phosphotransferases, phosphoglycolate phosphatases, histidinol phosphate phosphatases and serine phosphate phosphatases. A third group is defined by nitrophenyl phosphatases. Members of the sugar phosphatases define the fourth group and the fifth, and most recently categorized group, are the P-type ATPases.

GOB-1 possesses all of the domains and conserved residues required to have the catalytic function of a HAD-like hydrolase **(see Figure 30)**. Outside these domains, GOB-1 is highly divergent from other family members. This is perhaps not surprising since most members vary widely with regards to the sequence content and spacing between the conserved domains. When just the conserved motifs defined by Aravind et al. (1998) are used for a sequence alignment, the GOB-1 protein fits nicely into the HAD-like hydrolase family and appears to be most closely related to the sugar phosphatases and the phosphoserine phosphatases **(see Figure 31)**. The most convincing evidence that GOB-1 is in fact a HAD-like hydrolase family member is its ability to dephosphorylate trehalose-6-phosphate (like its most closely related HAD family members) and its requirement for Mg^{2+} for this catalytic activity.

GOB-1 is a Trehalose-6-phosphate Phosphatase

Once GOB-1 was considered to be a potential member of the HAD-like hydrolases, the next step was to define its catalytic function. It was thought that a defined molecular function could potentially reveal GOB-1's physiological function during intestine development. Candidate approaches were used to find a GOB-1 hydrolase activity. The first series of experiments tested for a GOB-1 associated protein phosphatase activity. This seemed like a possible candidate function since GOB-1 was possibly related to serine phosphatases based on sequence alignment. Furthermore, the Eyes absent like proteins were the only known HAD domain containing proteins with a role in development and these HAD domains were found to have a protein phosphatase function (Rayapureddi et al., 2003; Li et al., 2003; Tootle et al., 2003).

The failure to find a significant GOB-1 associated phosphatase activity for phosphorylated proteins and the phosphorylated serine and tyrosine resides (phospho-L-tyrosine is the preferred substrate of the Eyes absent phosphatase as defined by Tootle et al. (2003) and Rayapureddi et al. (2003)) suggests that these are not the likely GOB-1 substrates. Furthermore, GOB-1 is not closely related to Eyes absent in sequence and a different *C. elegans* protein EYA-1 appears to be a more likely member of the Eyes absent family (reported in WormBase, Stein et al., 2001). With this said, it cannot be ruled out that under *in vivo* conditions, GOB-1 could act on a protein substrate. Other unrelated HAD

domain containing proteins have been found to act as protein phosphatases (Kobor et al., 1999; Selengut, 2001; Yeo et al., 2003).

The next series of experiments tested for GOB-1 function as a sugar phosphatase. This function also seemed plausible since GOB-1 had sequence similarity to proteins with known functions in trehalose synthesis. Indeed, trehalose-6-phosphate is the only substrate so far tested that has a phosphate rapidly removed by the GOB-1 protein. GOB-1 displays a robust T6P activity (a Km of 136 μM and a kcat of 7.77 s^{-1}) and other sugar substrates tested did not have their phosphates removed by GOB-1 at an appreciable rate. The GOB-1 specificity for T6P suggests that GOB-1 is not a general sugar phosphatase. Even glucose-6-phosphate, one of the building units of trehalose-6-phosphate (see below), was not hydrolyzed by GOB-1. The strong and specific phosphatase activity of GOB-1 for trehalose-6-phosphate suggests that it may have an *in vivo* role for synthesizing trehalose.

How Can the *gob-1* Loss of Function Defect be Explained by GOB-1 Phosphatase Function?

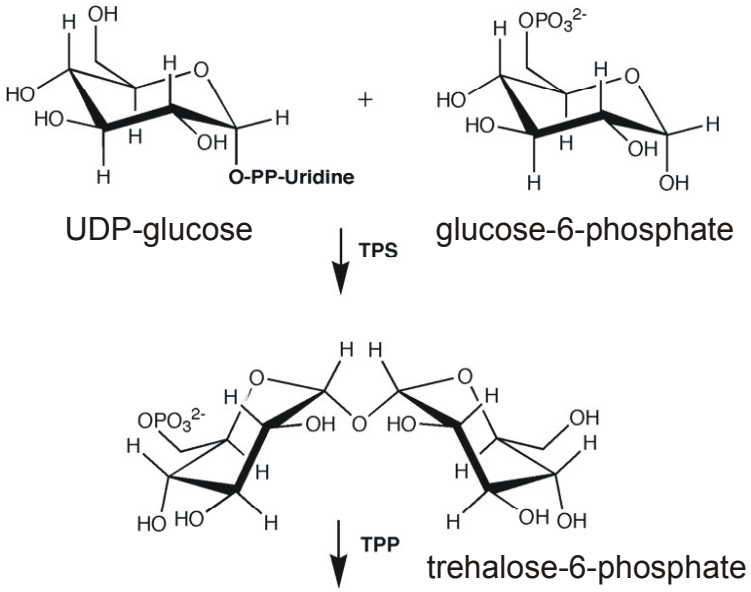
The finding that *in vitro* GOB-1 is a robust trehalose-6-phosphate phosphatase suggests that a defect in trehalose synthesis may be the reason for the *gob-1* loss of function associated defect. Below, the function of trehalose and the regulation of its metabolism will be reviewed. How these processes may or

may not apply to the phenotypes seen in the *gob-1* loss of function mutant will be discussed.

Trehalose is a disaccharide common to bacteria, plants, fungi, nematodes, and insects. Trehalose is most commonly synthesized in a two step pathway (Elbein et al., 2003) **(see Figure 41)**. The first step is the synthesis of trehalose-6-phosphate from UDP-glucose and glucose-6-phosphate. This step is carried out by the enzymatic activity of trehalose phosphate synthase (TPS). The domain responsible for this catalytic activity is defined as the glycosyltransferase 20 domain (Campbell et al., 1997). The second step of trehalose synthesis involves the removal of the phosphate from trehalose-6-phosphate. This reaction is catalyzed by a trehalose-6-phosphate phosphatase (TPP). The domain that is responsible for this phosphatase activity is a member of the HAD-like hydrolase superfamily and has been further classified as the IIB subfamily. The TPS/TPP synthesis pathway is the most common pathway discovered and is the only one found in eukaryotes with the exception of one fungus (Elbein et al., 2003; Tzvetkov et al., 2003). Alternative pathways exist in bacteria but these modes of synthesis appear to be peculiar to each species and are elaborated based on their dependence on trehalose.

In *E. coli*, the OtsA and OtsB proteins are required for the TPS and TPP activity respectively (Strom and Kaasen, 1993). In *Saccharomyces cerevisiae*, a complex of four proteins is required for trehalose synthesis (Bell et al., 1998; Reinders et al., 1997; Singer and Lindquist, 1998; Soto et al., 2002; Vandercammen et al., 1989). TPS1p possesses an active glycosyltransferase

Figure 41: Trehalose synthesis. The most common pathway for trehalose biosynthesis is a two step pathway. The first step involves the transfer of UDP-glucose to glucose-6-phosphate. This reaction is catalyzed by trehalose phosphate synthase (TPS) and results in the generation of trehalose-6-phosphtate (T6P). In the second step of the pathway, the phosphate molecule is removed from T6P. This reaction is catalyzed by trehalose-6-phosphate phosphatases (TPP) and results in the final product of trehalose. Figure taken from Elbein et al. (2003).


O-PP-Uridine
UDP-glucose
TPS
glucose-6-phosphate
TPP
trehalose-6-phosphate
trehalose

20 domain and this domain appears to be the only active TPS in *S. cerevisiae*. TPS2p possesses TPP activity and is the only active TPP in yeast. Interestingly, TPS2p has a domain with sequence similarity to the region containing the synthase domain in TPS1p but this domain is not functionally active (Bell et al., 1998). It is thought that this domain similarity, despite having no catalytic activity, may be required for complex formation. Inactive TPP-like domains have also been described in *Arabidopsis thaliana* and are thought to be involved in complex formation with other proteins involved in trehalose synthesis (Eastmond et al., 2003). In *S. cerevisiae*, there are also two other proteins TPS3p and TSL1p that are thought to have a largely regulatory function in the trehalose synthesis complex. In Drosophila, there is only one protein, Tps1, found to be involved in trehalose synthesis and this protein has both TPS and TPP domains, located in the N-terminus and C-terminus respectively. In nematodes, only trehalose phosphate synthase-like proteins have been found. In *C. elegans*, TPS-1 and TPS-2 are two similar proteins (54.2% identity across the proteins) with the glycosyltransferase 20 domain located in their N-termini (Pellerone et al., 2003). Sequence similarity searches have not identified a trehalose-6-phosphate phosphatase in any nematode (Behm, 1997; Pellerone et al., 2003). The specific activity of GOB-1 as a TPP suggests that GOB-1 and possibly the C-terminal domains of *C. elegans* TPS-1 and TPS-2 proteins are the missing trehalose-6-phosphate phosphatase in nematodes. The direct linkage of the trehalose phosphate synthase domain to the phosphatase domain in the same protein in *C. elegans* may have some functional relevance, since in yeast

and flies these domains are also kept in close proximity (reviewed by Singer and Lindquist, 1998; Vandercammen et al., 1989; Reinders et al., 1997; Bell et al., 1998; Soto et al., 2002).

Trehalose is proposed to have at least three important functions within the organisms that synthesize it. The first function involves storage of glucose (Elbein et al., 2003). Trehalose appears to be a major glucose store in yeast, plants, insects and nematodes and is the major circulating sugar in insects and nematodes. The trehalose disaccharide is unable to diffuse across membranes, so its synthesis provides a way for trapping imported glucose inside the cell. Trehalose is a non-reducing sugar and therefore can be stored at high levels in the cell without the adverse effect of interacting with cellular components.

A second function of trehalose has been largely investigated in yeast, where T6P is believed to serve as a signaling molecule to control metabolic pathways and growth (Elbein et al., 2003; Gancedo and Flores, 2004; Singer and Lindquist, 1998; Thevelein and Hohmann, 1995). In the presence of high levels of glucose, the yeast cell makes a metabolic switch from gluconeogenesis to glycolysis. Under these conditions, glucose is taken up into the cell and phosphorylated by several hexokinases (hexokinase PI, hexokinase PII and glucokinase). This results in the formation of glucose-6-phosphate (G6P), that is rapidly converted to fructose-6-phosphate (F6P) and then fructose-1,6-bisphosphate (F-1,6-P). When glycolysis is favored, F-1,6-P is shuttled into glycolysis and it appears that this step is highly regulated.

Trehalose synthesis was first hypothesized to have a role in glycolysis regulation by the observation that the *S. cerevisiae TPS1* mutant could not grow on a glucose supplemented medium (Reinders et al., 1997). It was later found that the *TPS1* mutant cells were blocked in growth due to a defect in the ability to shuttle phosphorylated glucose intermediates into the downstream glycolysis pathway (Blazquez and Gancedo, 1994). This unbalanced glycolytic flux resulted in high levels of hexose phosphates and a reduction in cellular levels of ATP due to the inability of enzymes downstream in the glycolysis pathway to sufficiently recover free phosphate from intermediate substrates (Bell et al., 1998). The nature of the cause of this imbalance in glycolysis is still in debate (Bonini et al., 2003; Ernandes et al., 1998; Hohmann et al., 1996). It appears that T6P in *in vitro* experiments can directly inhibit the *S. cerevisiae* hexokinase PII. This suggests that the level of control comes at the formation of glucose-6-phosphate and the entry of this molecule into glycolysis. Despite this finding, the up-regulation of hexokinase activity does not account for the full *TPS1* mutant phenotype. For instance, over-expression of hexokinase PII, a situation similar to the deregulation of hexokinase PII, causes a transient increase in G6P, F6P and F-1,6-P but does not cause the same block in glycolysis seen in the *TPS1* mutant (Ernandes et al., 1998). Other authors have suggested a T6P independent role for TSP1p protein in regulation of glycolysis. This is based on the observation that there is continued arrest of a *TPS1* (weak allele), *TPS2* double mutant. Despite the raised levels of T6P in this double mutant, the strain was still unable to grow on glucose (Hohmann et al., 1996).

In addition to the lethal effect of the *TPS1* deletion on glycolysis, high levels of T6P appear to lead to a decrease in thermotolerance in yeast (Gancedo and Flores, 2004). In *TPS2* mutants, there is an accumulation of T6P that results in thermosensitivity and osmotic sensitivity. A similar T6P related defect in chitin synthesis and cell wall integrity in *Aspergillus nidulans* has been reported (Borgia et al., 1996). The above results have lead to the conclusion that excess T6P can be detrimental to pathways involved in cell wall organization and maintenance of cell wall integrity.

The third function of trehalose, in most organisms that synthesize it, is the protection of proteins and lipid bilayers during stressful conditions (Elbein et al., 2003). Trehalose synthesis is dramatically up-regulated in response to desiccation, heat shock and cold stress. It is believed that trehalose is an excellent desiccation and cold shock stress mediator in that it is able to replace water and interact with the polar heads of lipids in the membrane to maintain the membrane in a liquid state. Trehalose's unique stereochemistry and its property of being less prone to crystallizing than other disaccharides such as sucrose, seem to make it ideal for this function. Trehalose also maintains proteins in a semi-folded state *via* associating with the hydroxyl groups and the polar residues of proteins.

In addition to the three major functions of trehalose mentioned above, other organisms use trehalose in other cellular functions. For instance, organisms such as the soil bacterium *Corynebacterium glutamicum* and the mycobacterium *Mycobacterium smegmatis* have trehalose-containing lipids that

are the major component of their lipid bilayer, and these lipids are required for proper membrane function (Klutts et al., 2003; Tzvetkov et al., 2003).

The role of trehalose in nematode physiology has been investigated previously (Behm, 1997; Pellerone et al., 2003). Trehalose is likely to be the major circulating sugar in nematodes. In the nematode *Ascaris suum*, trehalose is found in all tissues with the highest levels located in the reproductive system, muscles and hemolymph. In *C. elegans*, all tissues contain trehalose. However, eggs and dauer larvae contain the highest levels. Trehalose makes up to 2.3% of the dry weight of the *C. elegans* embryo and may be a stress protectant and/or aid in rehydration of the embryo at hatching. *tps-1* and *tps-2* transcripts are detected during all life stages of the worm (Pellerone et al., 2003). Despite the presence of trehalose during all life stages, knockdown of the *tps-1* and *tps-2* genes individually or in combination does not have an observable phenotype. Changes in viability, mobility, body shape, fertility, heat resistance and cold resistance were not obvious despite a drop in trehalose levels to 7% of wild-type. This suggests that the two major functions of trehalose as a glucose reserve and as a stress protectant may be dispensable for survival of the worm. This is in contrast to the phenotype found in flies where the level of Tps1 activity, and therefore trehalose synthesis, was related to anoxia resistance, and the *tps1* loss of function mutant resulted in larval lethality and abnormal organ morphology due to an increase in protein aggregation (Chen et al., 2002). Perhaps the incomplete knockdown of *tps-1* and *tps-2* function in *C. elegans* and the resulting low levels of trehalose are sufficient for necessary functions. Deletion of both

tps-1 and *tps-2* and therefore the effective elimination of trehalose synthesis may resolve the issue of whether trehalose has a necessary function in the worm.

The relationship of *gob-1* to trehalose metabolism in *C. elegans* is not clear for several reasons. First, if worms can potentially survive without trehalose, why would *gob-1* be required for survival? Second, if TPS-1, TPS-2 and GOB-1 all contain TPP domains, why can removal of *gob-1* function not be rescued by the TPP domains present in the C-terminal regions of TPS-1 and TPS-2. To explain the above scenarios, one could imagine that *gob-1* is absolutely required for trehalose synthesis. The partial knock down of *tps-1* and *tps-2* function with RNAi may allow for the presence of trehalose levels that are sufficient for survival. *gob-1* loss of function could potentially result in a stronger reduction of trehalose synthesis and lethality. Such a scenario could be addressed by measuring trehalose levels in *gob-1* mutants and determining if trehalose is reduced to even lower levels than those present the double *tps-1* and *tps-2* knockdown animals. Due to the larval lehality of *gob-1* loss of function and the failure of *gob-1* RNAi feeding to result in a penetrant defect, this has been technically difficult to test. An alternative way to test for the effect of trehalose levels is to generate the double mutant for null alleles of *tps-1* and *tps-2*. This null combination would elliminate production of trehalose. This situtation would test for the necessity for trehalose for survival.

In a different model, *gob-1* could be negatively regulating the TPS-1/TPS-2 associated TPS and TPP activity and in turn the excess production of trehalose

could cause a lethal phenotype. This particular lethality was seen with over-expression of Tps1 in Drosophila (Chen et al., 2002). To resolve this possibility, the *gob-1*, *tps-1* and *tps-2* triple mutant would create a scenario where trehalose production would be reduced, and the loss of *gob-1* as a negative regulator would not be required. The failure of the *tps-1* and *tps-2* double mutant to completely rescue the *gob*-1 mutant suggests that the deregulation of trehalose synthesis is not the reason for *gob-1* mutant lethality.

Another possibility is that GOB-1 is the only active T6P phosphatase in *C. elegans*. Despite the presence of all the amino acids required for phosphatase activity in TPS-1/TPS-2, these domains might not be catalytically active. The presence of catalytically inactive domains in proteins participating in the trehalose synthesis pathway have been detected in yeast and plants (Bell et al., 1998; Eastmond et al., 2003). In a scenario where GOB-1 is the only TPP in *C. elegans*, the *gob-1* mutant would result in a situation where the trehalose synthases are active but the trehalose-6-phosphate phosphatase is inactive, resulting in a cellular environment where T6P is in excess and potentially toxic to the cell. As noted earlier, T6P has been found to be a potent inhibitor of hexokinase PII in *S. cerevisiae* and high levels of T6P can interfere with cellular processes of the cell. Therefore, one could imagine the *gob-1* mutant lethality resulting from high concentrations of T6P and a deregulation of glycolysis or the breakdown of an important cellular pathway. The fact that the knockdown of the *tps1* and *tps2* in *C. elegans*, which in theory would reduce T6P formation in the

gob-1 mutant background, does not rescue the *gob-1* lethal phenotype strongly argues against the T6P over-production scenario.

gob-1 may have a T6P requiring function completely independent from the function of *tps-1* and *tps-2* in T6P synthesis. Indeed it has been speculated that the role of the multiple TPPs in some organisms may be to direct T6P into different pathways (Elbein et al., 2003). Despite the attractiveness of this model, this may not be the case as the *tps-1* and *tps-2* double knockdown and therefore a reduction in synthesis of T6P does not have an obvious phenotype. Complete reduction of *tps-1* and *tps-2* function and therefore elimination of T6P synthesis would strengthen this conclusion. Early in the studies of *gob-1*, it was thought that GOB-1 could be shuttling T6P into the assembly of carbohydrate-based complexes. The inability to detect a defect in carbohydrate deposition patterns using lectin staining suggests that GOB-1 is not directing T6P into at least some pathways of glycoprotein synthesis and supports the above conclusions.

Finally, *gob-1* could be required in a complex independent of its role as a T6P phosphatase. In this scenario, the GOB-1 protein alone could have a function in an unidentified pathway of metabolism or cellular function. A role in transcription or nuclear function is potentially plausible due to its occasional detection in the intestinal nuclei. Identification of proteins that directly interact with GOB-1 or genes that genetically interact with *gob-1* could shed light on a potentially new function.

Conclusions

In conclusion, *gob-1*(*ca17*) was identified in a genetic screen attempting to find genes that act either redundantly to *elt-2* or as a downstream effector of *elt-2*. *gob-1*(*ca17*) was positionally cloned and found to be a deletion of approximately 15 open reading frames of the X chromosome. RNAi to *gob-1* phenocopied the early larval arrest and gut obstructed phenotype of *gob-1*(*ca17*) and was thus concluded to be the gene responsible for this phenotype. The *gob-1* loss of function mutant causes an early defect in the intestine and the *gob-1* gene and GOB-1 protein are expressed specifically in the intestine during its later development. *gob-1* is therefore postulated to play a role in embryonic gut development, possibly by functioning in endotube formation and general metabolism of the intestine. GOB-1 is a member of the HAD-like hydrolase superfamily and the first identified trehalose-6-phosphate phosphatase in nematodes. GOB-1's function in trehalose metabolism is currently unclear and mounting evidence suggests that GOB-1 may have a function required for viability of the worm independent of trehalose synthesis. ELT-2 is sufficient but not necessary for expression of GOB-1 and this suggests that loss of *gob-1* function may be a component of the *elt-2* loss of function defect in the intestine.

Future Directions

The continued use of the Gob screen will be a valuable tool for identifying other genes involved in intestine development and maintenance. The less penetrant mutants isolated in this screen could be further mapped and characterized. Initial crosses of these mutants with *gob-1* and *elt-2* may assist in defining their relationship to these genes early in their characterization. Although useful in identifying genes required for intestine function, the Gob screen may not be the most efficient way to identify factors that are acting redundantly with *elt-2* during early gut specification. A more successful approach may include performing a genetic screen in an *elt-2* sensitized strain. For instance, a Gob-like screen could be performed in an *elt-2* heterozygous background and this genetic background may enhance the retrieval of genes that may not have a penetrant phenotype on their own but display a penetrant phenotype in a background of reduced *elt-2* function.

The identification of genes and/or proteins that interact with *gob-1*/GOB-1 will provide clues as to its function. The study of the genetic nulls of *tps-1* and *tps-2* will confirm initial prediction that a role in trehalose synthesis is not the primary reason for the defects detected in the *gob-1* loss of function mutant. The *tps-1* and *tps-2* deletion alleles have recently become available and I have generated a strain that is homozygous for the two deletion alleles. Although this strain needs to be studied in more detail, initial results suggest that the double mutant is viable. With a potential role in trehalose synthesis eliminated, other

functions of *gob-1* should be considered. Perhaps yeast two-hybrid screens could be used to find proteins that physically interact with GOB-1. The Gob screen and a modified Gob screen sensitized to *gob-1* may also identify genes that are genetically interacting with *gob-1* and therefore provide insight into the function of *gob-1*. *gob-1* is clearly important for intestine function and by understanding its function, we may gain insight into new mechanisms of intestine development and/or function.

REFERENCES

Adachi, R., and Kagawa, H. (2003). Genetic analysis of ryanodine receptor function in *Caenorhabditis elegans* based on *unc-68* revertants. Mol Genet Genomics *269*, 797-806.

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

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

An, J. H., and Blackwell, T. K. (2003). SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. Genes Dev *17*, 1882-1893.

Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998). The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. Trends Biochem Sci *23*, 127-129.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Ed.). (1994). Current Protocals in Molecular Biology. Volume 3. pp. 16.4.6, 16.4.7, 16.6.1 to 16.6.14 and 16.7.1 - 16.7. John Wiley and Sons, Inc.

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

Behm, C. A. (1997). The role of trehalose in the physiology of nematodes. Int J Parasitol *27*, 215-229.

Bei, Y., Hogan, J., Berkowitz, L. A., Soto, M., Rocheleau, C. E., Pang, K. M., Collins, J., and Mello, C. C. (2002). SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early *C. elegans* embryos. Dev Cell *3*, 113-125.

Bell, W., Sun, W., Hohmann, S., Wera, S., Reinders, A., De Virgilio, C., Wiemken, A., and Thevelein, J. M. (1998). Composition and functional analysis of the *Saccharomyces cerevisiae* trehalose synthase complex. J Biol Chem *273*, 33311-33319.

Bienz, M., and Clevers, H. (2000). Linking colorectal cancer to Wnt signaling. Cell *103*, 311-320.

Billin, A. N., Thirlwell, H., and Ayer, D. E. (2000). Beta-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator. Mol Cell Biol *20*, 6882-6890.

Blackwell, T. K., Bowerman, B., Priess, J. R., and Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. Science *266*, 621-628.

Blazquez, M. A., and Gancedo, C. (1994). Identification of extragenic suppressors of the *cif1* mutation in *Saccharomyces cerevisiae*. Curr Genet *25*, 89-94.

Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., Chiu, W. L., Duke, K., Kiraly, M., and Kim, S. K.

(2002). A global analysis of *Caenorhabditis elegans* operons. Nature *417*, 851-854.

Bonini, B. M., Van Dijck, P., and Thevelein, J. M. (2003). Uncoupling of the glucose growth defect and the deregulation of glycolysis in *Saccharomyces cerevisiae Tps1* mutants expressing trehalose-6-phosphate-insensitive hexokinase from *Schizosaccharomyces pombe*. Biochim Biophys Acta *1606*, 83-93.

Borgia, P. T., Miao, Y., and Dodge, C. L. (1996). The *orlA* gene from *Aspergillus nidulans* encodes a trehalose-6-phosphate phosphatase necessary for normal growth and chitin synthesis at elevated temperatures. Mol Microbiol *20*, 1287-1296.

Borgonie, G., van Driessche, E., Link, C. D., de Waele, D., and Coomans, A. (1994). Tissue treatment for whole mount internal lectin staining in the nematodes *Caenorhabditis elegans*, *Panagrolaimus superbus* and *Acrobeloides maximus*. Histochemistry *101*, 379-384.

Borgonie, G., Van Driessche, E., Link, C.D., Claeys, M., De Waele, D. and Coomans, A. (1997). Internal lectin binding patterns in the nematodes *Caenorhabditis elegans*, *Panagrolaimus superbus* and *Acrobeloides maximus*. Fundam. appl. Nematol. *20*, 173-186.

Bossard, P., and Zaret, K. S. (1998). GATA transcription factors as potentiators of gut endoderm differentiation. Development *125*, 4909-4917.

Bossinger, O., Fukushige, T., Claeys, M., Borgonie, G., and McGhee, J. D. (2004). The apical disposition of the *Caenorhabditis* elegans intestinal terminal web is maintained by LET-413. Dev Biol *268*, 448-456.

Bossinger, O., Klebes, A., Segbert, C., Theres, C., and Knust, E. (2001). Zonula adherens formation in *Caenorhabditis elegans* requires *dlg-1*, the homologue of the Drosophila gene *discs large*. Dev Biol *230*, 29-42.

Bowerman, B., Draper, B. W., Mello, C. C., and Priess, J. R. (1993). The maternal gene *skn-1* encodes a protein that is distributed unequally in early *C. elegans* embryos. Cell *74*, 443-452.

Bowerman, B., Eaton, B. A., and Priess, J. R. (1992). *skn-1*, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. Cell *68*, 1061-1075.

Brantjes, H., Roose, J., van De Wetering, M., and Clevers, H. (2001). All Tcf HMG box transcription factors interact with Groucho-related co-repressors. Nucleic Acids Res *29*, 1410-1419.

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

Brown, S., and Castelli-Gair Hombria, J. (2000). Drosophila *grain* encodes a GATA transcription factor required for cell rearrangement during morphogenesis. Development *127*, 4867-4876.

Buckley, M. S., Chau, J., Hoppe, P. E., and Coulter, D. E. (2004). *odd-skipped* homologs function during gut development in *C. elegans*. Dev Genes Evol *214*, 10-18.

Cadigan, K. M., and Nusse, R. (1997). Wnt signaling: a common theme in animal development. Genes Dev *11*, 3286-3305.

Campbell J.A. , Davies G.J. , Bulone V. , Henrissat B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. *326*, 929- 939.

Chen, Q., Ma, E., Behar, K. L., Xu, T., and Haddad, G. G. (2002). Role of trehalose phosphate synthase in anoxia tolerance and development in *Drosophila melanogaster*. J Biol Chem *277*, 3274-3279.

Cho, H., Wang, W., Kim, R., Yokota, H., Damo, S., Kim, S. H., Wemmer, D., Kustu, S., and Yan, D. (2001). BeF(3)(-) acts as a phosphate analog in proteins phosphorylated on aspartate: structure of a BeF(3)(-) complex with phosphoserine phosphatase. Proc Natl Acad Sci U S A *98*, 8525-8530.

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

Cogan, E. B., Birrell, G. B., and Griffith, O. H. (1999). A robotics-based automated assay for inorganic and organic phosphates. Anal Biochem *271*, 29-35.

Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998). A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. J Biol Chem *273*, 14107-14112.

Collet, J. F., Stroobant, V., and Van Schaftingen, E. (1999). Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases. J Biol Chem *274*, 33985-33990.

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

Consortium, I. H. G. M. (2001). A physical map of the human genome. Nature 409, 934-941.

Dixon, M., Webb, E., Thorne, C.J.R. and Tipton, K.F. (1979). Enzymes. 3rd Ed. pp. 55 – 138. Academic Press. New York and San Francisco.

Duncan, S. A., Manova, K., Chen, W. S., Hoodless, P., Weinstein, D. C., Bachvarova, R. F., and Darnell, J. E., Jr. (1994). Expression of transcription factor *HNF-4* in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: *HNF-4* is a marker for primary endoderm in the implanting blastocyst. Proc Natl Acad Sci U S A *91*, 7598-7602.

Eastmond, P. J., Li, Y., and Graham, I. A. (2003). Is trehalose-6-phosphate a regulator of sugar metabolism in plants? J Exp Bot *54*, 533-537.

Edgar, L. G., and McGhee, J. D. (1986). Embryonic expression of a gut-specific esterase in *Caenorhabditis elegans*. Dev Biol *114*, 109-118.

Elbein, A. D., Pan, Y. T., Pastuszak, I., and Carroll, D. (2003). New insights on trehalose: a multifunctional molecule. Glycobiology *13*, 17R-27R.

Ernandes, J. R., De Meirsman, C., Rolland, F., Winderickx, J., de Winde, J., Brandao, R. L., and Thevelein, J. M. (1998). During the initiation of

fermentation overexpression of hexokinase PII in yeast transiently causes a similar deregulation of glycolysis as deletion of *Tps1*. Yeast *14*, 255-269.

Firestein, B. L., and Rongo, C. (2001). DLG-1 is a MAGUK similar to SAP97 and is required for adherens junction formation. Mol Biol Cell *12*, 3465-3475.

Fujita, M., Hawkinson, D., King, K. V., Hall, D. H., Sakamoto, H., and Buechner, M. (2003). The role of the ELAV homologue EXC-7 in the development of the *Caenorhabditis elegans* excretory canals. Dev Biol *256*, 290-301.

Fukushige, T., Goszczynski, B., Tian, H., and McGhee, J. D. (2003). The evolutionary duplication and probable demise of an endodermal GATA factor in *Caenorhabditis elegans*. Genetics *165*, 575-588.

Fukushige, T., Hawkins, M. G., and McGhee, J. D. (1998). The GATA-factor *elt-2* is essential for formation of the *Caenorhabditis elegans* intestine. Developmental Biology *198*, 286-302.

Fukushige, T., Hendzel, M. J., Bazett-Jones, D. P., and McGhee, J. D. (1999). Direct visualization of the *elt-2* gut-specific GATA factor binding to a target promoter inside the living *Caenorhabditis elegans* embryo. PNAS *96*, 11883-11888.

Gancedo, C., and Flores, C. L. (2004). The importance of a functional trehalose biosynthetic pathway for the life of yeasts and fungi. FEMS Yeast Res *4*, 351-359.

Gaudet, J., and Mango, S. E. (2002). Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. Science *295*, 821-825.

Gebhard, A., and Gebert, A. (1999). Brush cells of the mouse intestine possess a specialized glycocalyx as revealed by quantitative lectin histochemistry. Further evidence for a sensory function. J Histochem Cytochem *47*, 799-808.

Gengyo-Ando, K., and Mitani, S. (2000). Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode *Caenorhabditis elegans*. Biochem Biophys Res Commun *269*, 64-69.

Gilleard, J. S., and McGhee, J. D. (2001). Activation of hypodermal differentiation in the *Caenorhabditis elegans* embryo by GATA transcription factors ELT-1 and ELT-3. Mol Cell Biol *21*, 2533-2544.

Gobel, V., Barrett, P. L., Hall, D. H., and Fleming, J. T. (2004). Lumen morphogenesis in C. *elegans* requires the membrane-cytoskeleton linker *erm-1*. Dev Cell *6*, 865-873.

Goldstein, B. (1992). Induction of gut in *Caenorhabditis elegans* embryos. Nature *357*, 255-257.

Goldstein, B. (1993). Establishment of gut fate in the E lineage of *C. elegans*: the roles of lineage-dependent mechanisms and cell interactions. Development *118*, 1267-1277.

Goldstein, B. (1995a). An analysis of the response to gut induction in the *C. elegans* embryo. Development *121*, 1227-1236.

Goldstein, B. (1995b). Cell contacts orient some cell division axes in the *Caenorhabditis elegans* embryo. J Cell Biol *129*, 1071-1080.

Hanna-Rose, W., and Han, M. (1999). COG-2, a sox domain protein necessary for establishing a functional vulval-uterine connection in *Caenorhabditis elegans*. Development *126*, 169-179.

Harlow, E. and Lane, D. (Ed.) (1998). Using Antibodies A laboratory Manual pp. 3- 20. Cold Spring Harbor Laboratory Press. New York.

Hecht, A., Vleminckx, K., Stemmler, M. P., van Roy, F., and Kemler, R. (2000). The p300/CBP acetyltransferases function as transcriptional coactivators of beta-catenin in vertebrates. EMBO J *19*, 1839-1850.

Hedgecock, E. M., and White, J. G. (1985). Polyploid tissues in the nematode *Caenorhabditis elegans*. Dev Biol *107*, 128-133.

Hermann, G. J., Leung, B., and Priess, J. R. (2000). Left-right asymmetry in *C. elegans* intestine organogenesis involves a LIN-12/Notch signaling pathway. Development *127*, 3429-3440.

Herndon, L. A., Schmeissner, P. J., Dudaronek, J. M., Brown, P. A., Listner, K. M., Sakano, Y., Paupard, M. C., Hall, D. H., and Driscoll, M. (2002). Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*. Nature *419*, 808-814.

Hobert, O., Mori, I., Yamashita, Y., Honda, H., Ohshima, Y., Liu, Y., and Ruvkun, G. (1997). Regulation of interneuron function in the *C. elegans* thermoregulatory pathway by the *ttx-3* LIM homeobox gene. Neuron *19*, 345-357.

Hodgkin, J., Horvitz, H. R., and Brenner, S. (1979). Nondisjunction mutants of the nematode *Caenorhabditis elegans*. Genetics *91*, 67-94.

Hohmann, S., Bell, W., Neves, M. J., Valckx, D., and Thevelein, J. M. (1996). Evidence for trehalose-6-phosphate-dependent and -independent mechanisms in the control of sugar influx into yeast glycolysis. Mol Microbiol *20*, 981-991.

Hopp, T. P., and Woods, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A *78*, 3824-3828.

Ichiyama, S., Kurihara, T., Li, Y. F., Kogure, Y., Tsunasawa, S., and Esaki, N. (2000). Novel catalytic mechanism of nucleophilic substitution by asparagine residue involving cyanoalanine intermediate revealed by mass spectrometric monitoring of an enzyme reaction. J Biol Chem *275*, 40804-40809.

Ishitani, T., Ninomiya-Tsuji, J., Nagai, S., Nishita, M., Meneghini, M., Barker, N., Waterman, M., Bowerman, B., Clevers, H., Shibuya, H., and Matsumoto, K. (1999). The TAK1-NLK-MAPK-related pathway antagonizes signalling between beta-catenin and transcription factor TCF. Nature *399*, 798-802.

Jacobsen, C. M., Narita, N., Bielinska, M., Syder, A. J., Gordon, J. I., and Wilson, D. B. (2002) Genetic mosaic analysis reveals that *GATA-4* is required for proper differentiation of mouse gastric epithelium. Dev Biol. 241, 34-46.

Janssen, D. B., Oppentocht, J. E., and Poelarends, G. J. (2001). Microbial dehalogenation. Curr Opin Biotechnol *12*, 254-258.

Kalb, J. M., Lau, K. K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P. G., and McGhee, J. D. (1998). *pha-4* is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the *C. elegans* pharynx. Development *125*, 2171-2180.

Kaletta, T., Schnabel, H., and Schnabel, R. (1997). Binary specification of the embryonic lineage in *Caenorhabditis elegans*. Nature *390*, 294-298.

Kamath, R. S., and Ahringer, J. (2003). Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods *30*, 313-321.

Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., and Ahringer, J. (2000). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. Genome Biology *2*, research0002.0001 - research0002.0010.

Katoh, M. (2002). Molecular cloning and characterization of *OSR1* on human chromosome 2p24. Int J Mol Med *10*, 221-225.

Klutts, S., Pastuszak, I., Edavana, V. K., Thampi, P., Pan, Y. T., Abraham, E. C., Carroll, J. D., and Elbein, A. D. (2003). Purification, cloning, expression, and properties of mycobacterial trehalose-phosphate phosphatase. J Biol Chem *278*, 2093-2100. Epub 2002 Nov 2001.

Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999). An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. Mol Cell *4*, 55-62.

Koonin, E. V., and Tatusov, R. L. (1994). Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. J Mol Biol *244*, 125-132.

Koppen, M., Simske, J. S., Sims, P. A., Firestein, B. L., Hall, D. H., Radice, A. D., Rongo, C., and Hardin, J. D. (2001). Cooperative regulation of AJM-1 controls junctional integrity in *Caenorhabditis elegans* epithelia. Nat Cell Biol *3*, 983-991.

Korswagen, H. C. (2002). Canonical and non-canonical Wnt signaling pathways in *Caenorhabditis elegans*: variations on a common signaling theme. Bioessays *24*, 801-810.

Korswagen, H. C., Herman, M. A., and Clevers, H. C. (2000). Distinct beta-catenins mediate adhesion and signalling functions in *C. elegans*. Nature *406*, 527-532.

Korswagen, H. C., Park, J. H., Ohshima, Y., and Plasterk, R. H. (1997). An activating mutation in a *Caenorhabditis elegans* Gs protein induces neural degeneration. Genes Dev *11*, 1493-1503.

Kostich, M., Fire, A., and Fambrough, D. (2000). Identification and molecular-genetic characterization of a LAMP/CD68-like protein from *Caenorhabditis elegans*. J Cell Sci *113*, 2595-2606.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Labouesse, M. (1997). Deficiency screen based on the monoclonal antibody MH27 to identify genetic loci required for morphogenesis of the *Caenorhabditis elegans* embryo. Dev Dyn *210*, 19-32.

Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B., and Evans, T. (1994). GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. J Biol Chem *269*, 23177-23184.

Lee, J., Ahnn, J., and Bae, S. C. (2004). Homologs of RUNX and CBF beta/PEBP2 beta in *C. elegans*. Oncogene *23*, 4346-4352.

Leung, B., Hermann, G. J., and Priess, J. R. (1999). Organogenesis of the *Caenorhabditis elegans* intestine. Dev Biol *216*, 114-134.

Levanon, D., Goldstein, R. E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., Paroush, Z., and Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. Proc Natl Acad Sci U S A *95*, 11590-11595.

Li, Q. L., Ito, K., Sakakura, C., Fukamachi, H., Inoue, K., Chi, X. Z., Lee, K. Y., Nomura, S., Lee, C. W., Han, S. B.*, et al.* (2002). Causal relationship between the loss of *RUNX3* expression and gastric cancer. Cell *109*, 113-124.

Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., Nigam, S. K., Aggarwal, A. K., Maas, R., Rose, D. W., and Rosenfeld, M. G. (2003). Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature *426*, 247-254.

Lin, R., Thompson, S., and Priess, J. R. (1995a). *pop-1* encodes an HMG box protein required for the specification of a mesoderm precursor in early *C. elegans* embryos. Cell *83*, 599-609.

Lin, W. H., Huang, L. H., Yeh, J. Y., Hoheisel, J., Lehrach, H., Sun, Y. H., and Tsai, S. F. (1995b). Expression of a Drosophila GATA transcription factor in multiple tissues in the developing embryos. Identification of homozygous lethal mutants with P-element insertion at the promoter region. J Biol Chem *270*, 25150-25158.

Lo, M. C., Gay, F., Odom, R., Shi, Y., and Lin, R. (2004). Phosphorylation by the beta-catenin/MAPK complex promotes 14-3-3-mediated nuclear export of TCF/POP-1 in signal-responsive cells in *C. elegans*. Cell *117*, 95-106.

Macara, I. G. (2004). Par proteins: partners in polarization. Curr Biol *14*, R160-162.

MacKintosh C. and Moorhead, G. (1999). Assay and purification of protein (serine/threonine) phosphatases, In Protein Phosphorylation: a Practical Approach. 2nd Ed. (ed. D.G. Hardie). pp. 153–181. IRL Press, Oxford.

Maduro, M. F., Lin, R., and Rothman, J. H. (2002). Dynamics of a developmental switch: recursive intracellular and intranuclear redistribution of *Caenorhabditis elegans* POP-1 parallels Wnt-inhibited transcriptional repression. Dev Biol *248*, 128-142.

Maduro, M. F., and Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the *Caenorhabditis elegans* endoderm. Dev Biol *246*, 68-85.

Mango, S. E., Lambie, E. J., and Kimble, J. (1994). The *pha-4* gene is required to generate the pharyngeal primordium of *Caenorhabditis elegans*. Development *120*, 3019-3031.

McComb, R.B., Bowers Jr., G.N., and Posen, S. (1979). Alkaline Phosphatase. pp. 313. Plenum Press. New York and London.

McKim, K. S., Howell, A. M., and Rose, A. M. (1988). The effects of translocations on recombination frequency in *Caenorhabditis elegans*. Genetics *120*, 987-1001.

McMahon, L., Legouis, R., Vonesch, J. L., and Labouesse, M. (2001a). Assembly of *C. elegans* apical junctions involves positioning and compaction by LET-413 and protein aggregation by the MAGUK protein DLG-1. J Cell Sci *114*, 2265-2277.

McMahon, M., Itoh, K., Yamamoto, M., Chanas, S. A., Henderson, C. J., McLellan, L. I., Wolf, C. R., Cavin, C., and Hayes, J. D. (2001b). The Cap'n'Collar basic leucine zipper transcription factor *Nrf2* (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. Cancer Res *61*, 3299-3307.

Mello, C.C. and Fire, A. (1995). DNA transformation. In *Caenorhabditis elegans*: Modern biological analysis of an organism (ed. H.F. Epstein and D.C. Shakes), pp. 452-482. Academic Press, San Diego.

Krause, M. (1995). Transcription and Translation. In *Caenorhabditis elegans*: Modern biological analysis of an organism (ed. H.F. Epstein and D.C. Shakes), pp. 483-512. Academic Press, San Diego.

Montgomery, R. K., Mulberg, A. E., and Grand, R. J. (1999). Development of the human gastrointestinal tract: twenty years of progress. Gastroenterology *116*, 702-731.

Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000). The crystal structure of bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. Biochemistry *39*, 10385-10396.

Nam, S., Jin, Y. H., Li, Q. L., Lee, K. Y., Jeong, G. B., Ito, Y., Lee, J., and Bae, S. C. (2002). Expression pattern, regulation, and biological role of runt domain transcription factor, run, in *Caenorhabditis elegans*. Mol Cell Biol *22*, 547-554.

Narita, N., Bielinska, M., and Wilson, D. B. (1997). Wild-type endoderm abrogates the ventral developmental defects associated with *GATA-4* deficiency in the mouse. Dev Biol *189*, 270-274.

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

Odenthal, J., and Nusslein-Volhard, C. (1998). fork head domain genes in zebrafish. Dev Genes Evol *208*, 245-258.

Ohno, S. (2001). Intercellular junctions and cellular polarity: the PAR-aPKC complex, a conserved core cassette playing fundamental roles in cell polarity. Curr Opin Cell Biol *13*, 641-648.

Otsuka, A. J., Jeyaprakash, A., Garcia-Anoveros, J., Tang, L. Z., Fisk, G., Hartshorne, T., Franco, R., and Born, T. (1991). The *C. elegans unc-104* gene encodes a putative kinesin heavy chain-like protein. Neuron *6*, 113-122.

Parviz, F., Matullo, C., Garrison, W. D., Savatski, L., Adamson, J. W., Ning, G., Kaestner, K. H., Rossi, J. M., Zaret, K. S., and Duncan, S. A. (2003). Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis. Nat Genet. 34, 292-6.

Pellerone, F. I., Archer, S. K., Behm, C. A., Grant, W. N., Lacey, M. J., and Somerville, A. C. (2003). Trehalose metabolism genes in *Caenorhabditis elegans* and filarial nematodes. Int J Parasitol *33*, 1195-1206.

Rayapureddi, J. P., Kattamuri, C., Steinmetz, B. D., Frankfort, B. J., Ostrin, E. J., Mardon, G., and Hegde, R. S. (2003). Eyes absent represents a class of protein tyrosine phosphatases. Nature *426*, 295-298.

Reboul, J., Vaglio, P., Tzellas, N., Thierry-Mieg, N., Moore, T., Jackson, C., Shin-i, T., Kohara, Y., Thierry-Mieg, D., Thierry-Mieg, J.*, et al.* (2001). Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. Nat Genet *27*, 332-336.

Reinders, A., Burckert, N., Hohmann, S., Thevelein, J. M., Boller, T., Wiemken, A., and De Virgilio, C. (1997). Structural analysis of the subunits of the trehalose-6-phosphate synthase/phosphatase complex in *Saccharomyces cerevisiae* and their function during heat shock. Mol Microbiol *24*, 687-695.

Reiter, J. F., Kikuchi, Y., and Stainier, D. Y. (2001). Multiple roles for Gata5 in zebrafish endoderm formation. Development *128*, 125-135.

Reuter, R. (1994). The gene *serpent* has homeotic properties and specifies endoderm versus ectoderm within the Drosophila gut. Development *120*, 1123-1135.

Ridder, I. S., and Dijkstra, B. W. (1999). Identification of the Mg2+-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY. Biochem J *339*, 223-226.

Riddle, D.L., Blumenthal, T., Meyer, B.J. and Priess, J.R. (Ed.) (1997). *C. elegans* II. Cold Spring Habor Laboratory Press. New York.

Roberts, D. J. (2000). Molecular mechanisms of development of the gastrointestinal tract. Dev Dyn *219*, 109-120.

Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y. H., Ali, M., Priess, J. R., and Mello, C. C. (1997). Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. Cell *90*, 707-716.

Roose, J., Molenaar, M., Peterson, J., Hurenkamp, J., Brantjes, H., Moerer, P., van de Wetering, M., Destree, O., and Clevers, H. (1998). The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. Nature *395*, 608-612.

Rose, A. M., and Baillie, D. L. (1980). Genetic Organization of the region around *unc-15* (I), a gene affecting paramyosin in *Caenorhabidtis elegans*. Genetics *96*, 639-648.

Schedin, P. Hunter, C.P., and Wood, W. B. (1991). Autonomy and nonautonomy of sex determination in triploid intersex mosaics of *C. elegans*. Development *112*,863-879.

Schierenberg, E. (1987). Reversal of cellular polarity and early cell-cell interaction in the embryos of *Caenorhabditis elegans*. Dev Biol *122*, 452-463.

Schlesinger, A., Shelton, C. A., Maloof, J. N., Meneghini, M., and Bowerman, B. (1999). Wnt pathway components orient a mitotic spindle in the early *Caenorhabditis elegans* embryo without requiring gene transcription in the responding cell. Genes Dev *13*, 2028-2038.

Segbert, C., Johnson, K., Theres, C., van Furden, D., and Bossinger, O. (2004). Molecular and functional analysis of apical junction formation in the gut epithelium of *Caenorhabditis elegans*. Dev Biol *266*, 17-26.

Selengut, J. D. (2001). MDP-1 is a new and distinct member of the haloacid dehalogenase family of aspartate-dependent phosphohydrolases. Biochemistry *40*, 12704-12711.

Shin, T. H., Yasuda, J., Rocheleau, C. E., Lin, R., Soto, M., Bei, Y., Davis, R. J., and Mello, C. C. (1999). MOM-4, a MAP kinase kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in *C. elegans*. Mol Cell *4*, 275-280.

Shivdasani, R. A. (2002). Molecular regulation of vertebrate early endoderm development. Dev Biol *249*, 191-203.

Shoichet, S. A., Malik, T. H., Rothman, J. H., and Shivdasani, R. A. (2000). Action of the *Caenorhabditis elegans* GATA factor END-1 in Xenopus

suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. Proc Natl Acad Sci U S A *97*, 4076-4081.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J., and Plasterk, R. H. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. Curr Biol *12*, 1317-1319.

Simon, T. C., and Gordon, J. I. (1995). Intestinal epithelial cell differentiation: new insights from mice, flies and nematodes. Curr Opin Genet Dev *5*, 577-586.

Simske, J. S., and Kim, S. K. (1995). Sequential signalling during *Caenorhabditis elegans* vulval induction. Nature *375*, 142-146.

Singer, M. A., and Lindquist, S. (1998). Thermotolerance in *Saccharomyces cerevisiae*: the Yin and Yang of trehalose. Trends Biotechnol *16*, 460-468.

Smit, L., Baas, A., Kuipers, J., Korswagen, H., van de Wetering, M., and Clevers, H. (2004). Wnt activates the Tak1/Nemo-like kinase pathway. J Biol Chem *279*, 17232-17240. Epub 12004 Feb 17211.

Soto, T., Franco, A., Padmanabhan, S., Vicente-Soler, J., Cansado, J., and Gacto, M. (2002). Molecular interaction of neutral trehalase with other enzymes of trehalose metabolism in the fission yeast *Schizosaccharomyces pombe*. Eur J Biochem *269*, 3847-3855.

Stainier, D. Y. (2002). A glimpse into the molecular entrails of endoderm formation. Genes Dev *16*, 893-907.

Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., and Spieth, J. (2001). WormBase: network access to the genome and biology of *Caenorhabditis elegans*. Nucleic Acids Res *29*, 82-86.

Strom, A. R., and Kaasen, I. (1993). Trehalose metabolism in *Escherichia coli*: stress protection and stress regulation of gene expression. Mol Microbiol *8*, 205-210.

Stubbs, M. D., Tran, H. T., Atwell, A. J., Smith, C. S., Olson, D., and Moorhead, G. B. (2001). Purification and properties of *Arabidopsis thaliana* type 1 protein phosphatase (PP1). Biochim Biophys Acta *1550*, 52-63.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev Biol *100*, 64-119.

Taraviras, S., Monaghan, A. P., Schutz, G., and Kelsey, G. (1994). Characterization of the mouse *HNF-4* gene and its expression during mouse embryogenesis. Mech Dev *48*, 67-79.

Thevelein, J. M., and Hohmann, S. (1995). Trehalose synthase: guard to the gate of glycolysis in yeast? Trends Biochem Sci *20*, 3-10.

Thomas, J. H. (1990). Genetic Analysis of Defecation in *Caenorhabditis elegans*. Genetics *124*, 855-872.

Thorpe, C. J., and Moon, R. T. (2004). nemo-like kinase is an essential co-activator of Wnt signaling during early zebrafish development. Development *131*, 2899-2909.

Thorpe, C. J., Schlesinger, A., and Bowerman, B. (2000). Wnt signalling in *Caenorhabditis elegans*: regulating repressors and polarizing the cytoskeleton. Trends Cell Biol *10*, 10-17.

Thorpe, C. J., Schlesinger, A., Carter, J. C., and Bowerman, B. (1997). Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. Cell *90*, 695-705.

Tootle, T. L., Silver, S. J., Davies, E. L., Newman, V., Latek, R. R., Mills, I. A., Selengut, J. D., Parlikar, B. E., and Rebay, I. (2003). The transcription factor Eyes absent is a protein tyrosine phosphatase. Nature *426*, 299-302.

Truett, G. E., Heeger, P., Mynatt, R. L., Truett, A. A., Walker, J. A., and Warman, M. L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). Biotechniques *29*, 52, 54.

Tzvetkov, M., Klopprogge, C., Zelder, O., and Liebl, W. (2003). Genetic dissection of trehalose biosynthesis in *Corynebacterium glutamicum*: inactivation of trehalose production leads to impaired growth and an altered cell wall lipid composition. Microbiology *149*, 1659-1673.

Vandercammen, A., Francois, J., and Hers, H. G. (1989). Characterization of trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase of *Saccharomyces cerevisiae*. Eur J Biochem *182*, 613-620.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Walker, A. K., See, R., Batchelder, C., Kophengnavong, T., Gronniger, J. T., Shi, Y., and Blackwell, T. K. (2000). A conserved transcription motif suggesting functional parallels between *Caenorhabditis elegans* SKN-1 and Cap'n'Collar-related basic leucine zipper proteins. J Biol Chem *275*, 22166-22171.

Walsh, S., Anderson, M., and Cartinhour, S. W. (1998). ACEDB: a database for genome information. Methods Biochem Anal *39*, 299-318.

Waterston, R. H., Thomson, J. N., and Brenner, S. (1980). Mutants with altered muscle structure of *Caenorhabditis elegans*. Dev Biol *77*, 271-302.

Weigel, D., Jurgens, G., Klingler, M., and Jackle, H. (1990). Two gap genes mediate maternal terminal pattern information in Drosophila. Science *248*, 495-498.

Weigel, D., Jurgens, G., Kuttner, F., Seifert, E., and Jackle, H. (1989). The homeotic gene *fork head* encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. Cell *57*, 645-658.

Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H., and Plasterk, R. H. (2001). Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. Nat Genet *28*, 160-164.

Wodarz, A. (2002). Establishing cell polarity in development. Nat Cell Biol *4*, E39-44.

Wood, W.B. and the Community of *C. elegans* Researchers. (1988). The Nematode *Caenorhabditis elegans*. Cold Spring Habor Laboratory Press. New York.

Yeo, M., Lin, P. S., Dahmus, M. E., and Gill, G. N. (2003). A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. J Biol Chem *278*, 26078-26085.

Zaret, K. (1999). Developmental competence of the gut endoderm: genetic potentiation by GATA and HNF3/fork head proteins. Dev Biol *209*, 1-10.

Zhu, J., Fukushige, T., McGhee, J. D., and Rothman, J. H. (1998). Reprogramming of early embryonic blastomeres into endodermal progenitors by a *Caenorhabditis elegans* GATA factor. Genes Dev *12*, 3809-3814.

Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997). *end-1* encodes an apparent GATA factor that specifies the endoderm precursor in *Caenorhabditis elegans* embryos. Genes Dev *11*, 2883-2896.