Study of Extracellular Matrix Synthesis in *C. elegans*

Carolina Ibáñez Ventoso

Wellcome Centre for Molecular Parasitology

University of Glasgow

Submitted for the degree of Doctor of Philosophy at the University of Glasgow

July 2003

Abstract

The extracellular matrix (ECM) is a complex network of secreted proteins and carbohydrates ubiquitous in metazoans. There are many different types of ECM, each composed of a characteristic set of components assembled into a structure of specialised function. In general, ECMs are essential for the organism's development and shape. The major components of ECMs are collagen proteins. The free-living nematode *Caenorhabditis elegans* is a broadly accepted model organism. *C. elegans* has two types of ECM: basement membranes and the cuticle. The cuticle is an exoskeleton that encloses the nematode, and is essential for maintenance and definition of post-embryonic body shape. *C. elegans* develops from embryo to adult through four larval stages. The epithelial monolayer of cells surrounding the animal, the hypodermis, synthesises five cuticles during the nematode life cycle. The first cuticle is formed within the egg, prior to hatching, and the remainder towards the end of each larval stage.

Because of the structural role of the cuticle, mutations in genes involved in assembly of this ECM can cause a spectrum of effects from lethality late in embryogenesis to alterations in the nematode shape. The severity of phenotype correlates with the severity of cuticle synthesis defects. Accordingly, two distinct mutant alleles that cause death after embryonic elongation, possibly due to failure in synthesising an intact cuticle, were characterised. One mutant, *ij15*, was isolated from a forward genetic screen previously performed (I. Johnstone, Glasgow University, Glasgow, UK). *ij15* defines mutationally the gene *stc-1*, which encodes a HSP70-like protein possibly localised in the secretory pathway. The other mutant, *h402*, defines mutationally the gene *let-607*. A second *let-607* allele, *h189*, which results in larval lethality at the L2 stage was also analysed in this study. *let-607* corresponds to the predicted gene F57B10.1, which encodes a putative bZIP transcription factor. Both *stc-1* and *let-607* are expressed in the hypodermis at all developmental stages. Furthermore, disruption of the function of either *stc-1* or *let-607* by mutation or RNAi affects cuticle synthesis in different ways. Thus, *stc-1* and *let-607* encode for a HSP70-like protein and a putative bZIP transcription factor required for synthesis of the cuticular ECM in *C. elegans*. In addition, this study defines *C. elegans* mutant phenotypes that can be used as indicators for gene products with controlling roles in the synthesis of this ECM.

List of Contents

Chapter 1

General Introduction

Chapter 2

Materials and Methods

Chapter 3

stc-1(ij15), *let-607(h402)* and *let-607(h189) C. elegans* mutants are defective in
cuticle synthesis

Chapter 4

stc-1 encodes a HSP70-like protein required for cuticle synthesis in *C.elegans*

Chapter 5

let-607 encodes a bZIP transcription factor required for cuticle synthesis in *C. elegans*

Chapter 6
General Discussion

References

List of Figures and Tables

Chapter 4

Chapter 5

Acknowledgements

Declaration

The work presented in this thesis was performed entirely by the author except where stated. This thesis contains unique work and will not be submitted for any other degree, diploma or qualification at any other University.
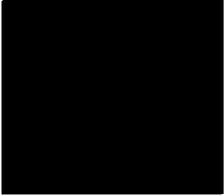
C. Ibáñez Ventoso
July, 2003

Abbreviations

Abs	absorbance
ACeDB	a *C. elegans* database
ADP	adenosine 5'-diphosphate
ATP	adenosine 5'-triphospate
BiP	immunoglobulin heavy chain binding protein
bZIP	basic leucine zipper
CDS	coding sequence
CGC	*Caenorhabditis* Genetic Centre
CIAP	calf intestinal alkaline phosphatase
DIC	differential interference contrast
DNA	deoxyribonucleic acid
cDNA	complementary deoxyribonucleic acid
gDNA	genomic deoxyribonucleic acid
EC	expression clone
ECM	extracellular matrix
EMS	ethyl methanesulfonate
ER	endoplasmic reticulum
EST	expressed sequenced tag
F_1	first filial generation
F_2	second filial generation
F_3	third filial generation
GFP	green fluorescent protein
HSP	heat shock protein
L1-4	larval 1-4 stages
LB	Luria-Bertani
MTP	microsomal triglyceride transfer protein
NGM	nematode growth media
Ni-NTA	nickel-nitrilotriacetic acid
OD	optical density
PAGE	polyacrilamide gel electrophoresis
PCR	polymerase chain reaction
P4H	prolyl 4-hydroxylase
PDI	protein disulphide isomerase

rtPCR	reverse transcriptase PCR
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
RNAi	RNA-mediated interference
dsRNA	double-stranded RNA
mRNA	messenger RNA
ssRNA	single-stranded RNA
SAGE	serial analysis of gene expression
SEM	scanning electron microscopy
SL	spliced leader
STCH	stress chaperone
SWLB	single worm lysis buffer
TBE	Tris borate buffer
TMP/UV	ultraviolet activated trimethylpsolaren
UGGT	UDP-Glc:glycoprotein glucosyltransferase
UPR	unfolded protein response
UV	ultraviolet
WLB	worm lysis buffer
WT	wild type
w/v	weight / volume
YAC	yeast artificial chromosome

Caenorhabditis elegans mutant phenotypes: blister (Bli), dumpy (Dpy), long (Lon), right roller (RRol), left roller (LRol), abnormal hermaphrodite tail (Tal), uncoordinated (Unc)

Measurements

bp	base pair
cm	centimetre
g	gram
h	hour
in	inch
Kb	kilobase
kDa	kilodalton
kV	kilovolt

L	litre
M	molar
ml	millilitre
mM	milimolar
m.u	map unit
ng	nanogram
nm	nanometre
rpm	revolutions per minute
sec	second
V	volt
μg	microgram
μM	micromolar
μl	microlitre
°C	degree centigrade
Ω	ohm
%	per cent

Chapter 1

General Introduction

1.1   BASICS IN *CAENORHABDITIS ELEGANS* BIOLOGY

Caenorhabditis elegans is a free-living soil nematode widely used as a model system in biological analyses due to many of its attributes, including small size (the adult is approximately 1 mm long), 97 Mb nuclear genome, rapid generation time and procreation by self-fertilisation and sexual reproduction. *C. elegans* is easily maintained on agar plates or liquid culture and strains can be stored frozen for years and possibly indefinitely. Other favourable features are its transparent body, anatomical simplicity and the constancy of cell number and cell position from individual to individual, which have made this nematode particularly advantageous for the study of development. *C. elegans* research has greatly benefited by three events: the description of the entire cell lineage and of the complete anatomy of the nervous system, and the completion of the genome sequence in 1998. From a technical point of view, this nematode is also an attractive system for different reasons, such as easy application of classical genetic approaches, DNA mediated transformation, reporter gene technology, DNA microarrays, gene knockouts and RNA-mediated interference (RNAi). All the above resources have resulted in an enormous expansion of the field with the obvious consequent generation of a vast amount of data. A full profit of this wealth of information would not have been possible without the existence of public and comprehensive databases that are always improving and keeping up with biological advances in *C. elegans*.

1.1.1   Life cycle

C. elegans progresses from egg to adult through four larval stages: L1, L2, L3 and L4 in 3.5 days at 20°C (Fig. 1.1). This cycle continues with the reproductive adult producing more than 300 progeny over a four-day period. In adverse conditions, such as high population density and insufficient food supply, L3 larvae adopt an alternative developmental fate called dauer. This is a non-growing stage specialised for survival and dispersal to favourable locations in their natural environment, endurable for several months. Dauer larvae recover to normal development when good conditions are encountered.

1.1.2   Anatomy

C. elegans anatomy is simple and well known (Fig. 1.2), mostly from detailed reconstruction of electron micrographs of serial sections. Despite its simplicity, the worm has many differentiated tissue types characteristic of more complex organisms, such as epidermis (or hypodermis),
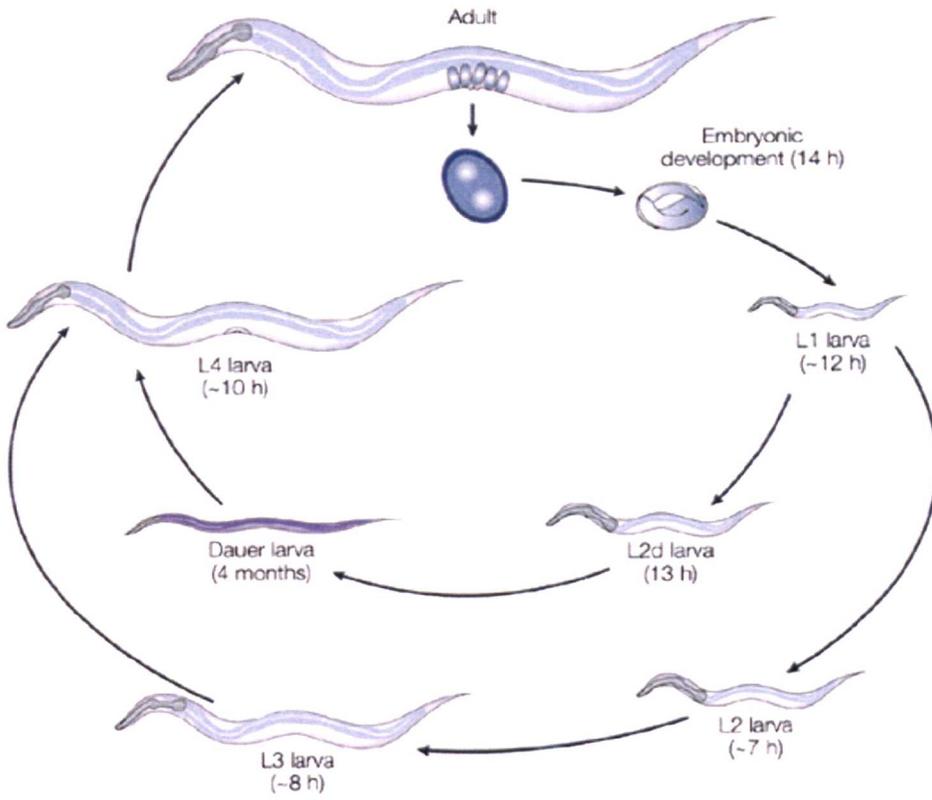
Figure 1.1: *C. elegans* life cycle

(adapted from Jorgensen and Mango, 2002)
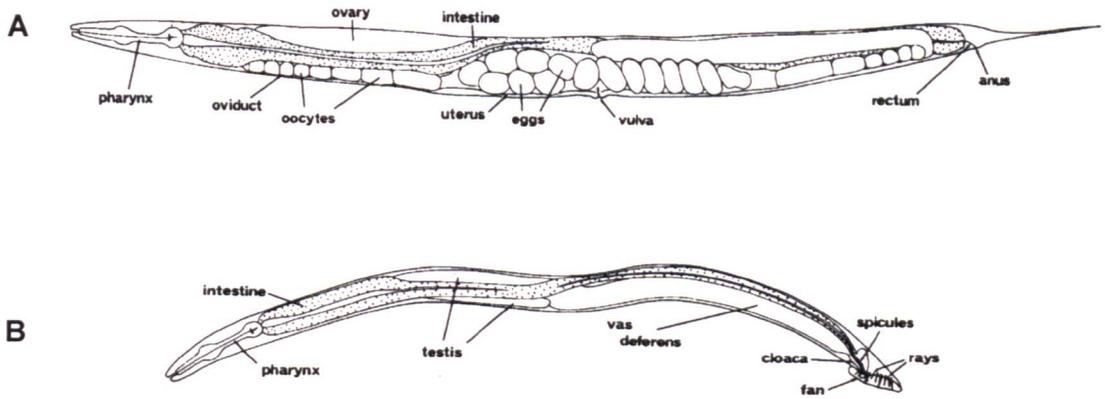
Figure 1.2: Drawings showing major anatomical features of *C. elegans* adult hermaphrodite (A) and adult male (B).

(adapted from Sulston and Horvitz, 1977)
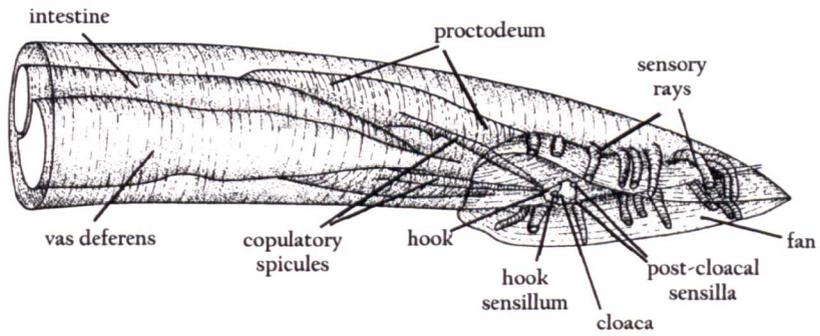
Figure 1.3: A cartoon of an adult male tail, left subventral view.

(from Sulston *et al.*, 1980)

muscle, nerve, intestine and gonad, which contains the germline. The basic body plan, most organs and structural differences between the two sexes are briefly discussed in this section.

All nematodes, including *C. elegans*, have the same broad body structure of two concentric tubes separated by a fluid filled space, the pseudocoelom. The outer tube comprises the cuticle, neurons, hypodermis and muscle. A thin basement membrane lines the pseudocoelom and separates muscle from hypodermis and nervous tissue. The pseudocoelomic cavity contains the inner tube, the intestine, as well as the gonads in adult worms. The pharynx, intestine and gonad are also covered by basement membrane.

The hypodermis and cuticle give form to the nematode body, which is then maintained by internal hydrostatic pressure. Characteristics of the hypodermis and cuticle are described in sections 3.2 and 3.3 respectively. The body-wall muscle consists of four longitudinal strips, two ventral and two dorsal, which are attached to the cuticle by fibrous elements that extend through the hypodermis. The ventral muscle strips and the dorsal muscle strips act antagonistically generating a sinusoidal movement under continuous phases of contraction and relaxation. Most of the neuron cells bodies are found around the pharynx, along the ventral side and in the tail. Processes projected by these neurons tend to organise into bundles that run mainly around the pharynx (the nerve ring) and the length of the animal (ventral and dorsal cords). The worm alimentary tract is formed by the pharynx, the intestine and the rectum. The bilobed pharynx is a composite of muscles, epithelial cells, and nerves. Food is ingested through its anterior section, processed in the terminal bulb by the grinder and subsequently pumped into the intestinal lumen that connects to the anus near the tail. The lumen is surrounded first by a ring of four cells followed by two rows of eight cells. A 180° twist occurs in the intestinal cells around the middle part of the animal body, coinciding with the localisation of the gonad primordium in the L1 larvae. Most intestinal nuclei become binucleate at the end of the first larval stage, excluding the anterior six nuclei that do not divide and the four posterior nuclei that may fail to undergo division (Sulston and Horvitz, 1977). At the end of the following larval stages, intestinal cells undergo endoreduplications of their DNA without nuclear divisions (Hedgecock and White, 1985). The intestine is a multifunctional organ. Its primary role is the synthesis and secretion of digestive enzymes and absorption of nutrients. It also nurtures germ cells by producing yolk proteins and acts as an energy store, as deduced from its cytoplasmic content of granules. Intestinal granules are refractile when viewed with polarised light and autofluorescent under 300-400 nm light. All the above organs are common in the two sexes. I shall now discuss some of the anatomical differences between adult hermaphrodites and males. The hermaphrodite

reproductive structures include a two-armed gonad, uterus and vulva. Each arm of the gonad is U-shaped and contains an ovary, oviduct and spermatheca connecting to a common uterus that opens to the exterior through the vulva. The male reproductive system is made up of a single-armed gonad and a specialised tail known as copulatory bursa. The gonad consists of a testis, seminal vesicle and a vas deferens that connects to the exterior through the cloaca. An adult male tail is shown in Figure 1.3. It is composed of an elongated bursa, cuticularised fan, sensory receptors and proctodeum. Nine pairs of sensilla extend over the fan, known as rays, and three others are located around the cloaca: one anterior associated to a hook (the hook sensillum) and two posterior. Two spicules and the gubernaculum are in the proctodeum. Each of the tail structures function specifically in the mating process.

1.2   C. ELEGANS GENETICS

Eukaryotic genomes include both nuclear and mitochondrial DNA. *C. elegans* has a small DNA content comprising a 13.794 Kb mitochondrial genome and a 97 Mb nuclear genome organised into six chromosomes, five autosomes (A) and a sex chromosome (X). Hermaphrodites are XX and males are X0, and arise spontaneously in hermaphrodite populations by X-chromosome non-disjunction at a frequency of about 0.1%. The ability of hermaphrodites to produce sperm and oocytes allows for self-fertilisation, resulting quickly in inbred lines. This mode of reproduction has other advantages. Firstly, mutants with severe defects that impede mating can be isolated and maintained as homozygous stocks. Secondly, maintenance of recessive lethal alleles is feasible in heterozygosis, since a heterozygous hermaphrodite for a lethal mutation will produce only one quarter dead progeny. Hermaphrodites, however, will preferentially cross-fertilise with males when these are present, thus allowing the exchange of genetic material between individuals. In addition to the above characteristics, the ability to handle large numbers of animals and the small genome make *C. elegans* a good organism for laboratory manipulation, the generation of mutants and their subsequent analysis.

Most of the mutations in *C. elegans* have been obtained by treatment with ethyl methanesulfonate (EMS), a potent and convenient mutagen that normally produces GC to AT transitions, and other types of point mutations and deletions at lower frequency. An effective means to induce deletions is by insertion and imprecise excision of transposons. This method requires the use of mutant strains with the machinery for transposition active in the germline. In addition to EMS and transposons, a variety of other mutagens have been successfully employed in *C. elegans* as can be found in Anderson (1995).

Mutations in genes associated with determined phenotypic defects allow information about the gene function to be obtained. This can be achieved by two different approaches. The classical strategy, namely forward genetics, starts with a mutant phenotype of interest followed by its characterisation, mapping of the mutation and positional cloning to identify the gene responsible. With the annotated sequence of virtually all *C. elegans* genes (see section 1.2.1), there has been an increase in the use of reverse genetics, going from gene sequence to mutant phenotype. Thus, researchers can concentrate on specific genes that have been implicated in a biological process of interest, or in human or animal diseases. Extensive functional information can now be obtained by performing homology searches, and from databases containing results of systematic genomic analyses (section 1.2.2.1). All the data compiled in this manner, *in silico*, can then be used as a starting point and therefore is an accelerator for in-depth studies either of *C. elegans* or of other organisms.

1.2.1 The genome project

The elucidation of the sequence of the *C. elegans* genome was the outcome of a large two-phase project carried out mainly by two laboratories. Firstly, an ordered physical map of the six chromosomes was constructed using a combination of cosmid, yeast artificial chromosome (YAC) and fosmid clones (Coulson *et al.*, 1988; Coulson *et al.*, 1991). The primary reason for cloning the entire genome was to facilitate the molecular cloning of a large number of loci that were identified by mutation. The free availability of all these clones since their generation has been of great value to the research community. Once the physical map was nearly completed, the systematic sequencing of the genome was a desirable next step. On top of this, the nematode was considered, together with other model organisms with a small genome, as a good system to test and develop sequencing techniques and an essential source of information for subsequent interpretation of the human sequence. Following sequencing of the genome, a variety of analyses were applied to predict the coding structures. The program Genefinder was used to identify possible genes that were then confirmed by comparison to end sequences of selected cDNAs and by homology BLAST searches against various databases, which often provided functional information as well. In addition, the Pfam protein database was employed to detect common protein domains in the predicted coding sequences. All data was collected and integrated under human supervision to produce the final gene annotation submitted to GenBank, EMBL and ACeDB databases. The mitochondrial genome was fully sequenced by 1992 (Okimoto *et al.*, 1992), whereas the sequence and gene annotation of the entire nuclear genome was essentially completed at the end of 1998 (*C. elegans* sequencing consortium, 1998). More recently, the genome sequence was subjected to a second prediction of coding regions using the

program Genie (http://www.wormbase.org). The current estimates suggest there are about 19,404 protein-coding genes, ~60% of which are significantly related to genes in yeast, *Drosophila*, and/or humans; and the remaining 40% appear to be nematode specific (Rubin *et al.*, 2000).

1.2.2 Post-genomic era

The ultimate motivation behind the sequencing of an organism's genome is to discern the meaning of that sequence, in particular, the function of its genes. A starting point to gain insight into a gene function is the characterisation of its mutant phenotype and expression profile.

There are two main methods to obtain mutant phenotypes in *C. elegans* using a reverse genetic approach: the polymerase chain reaction (PCR)-based knockout methodology and double-stranded (ds) RNAi. The first method consists of the identification of a deletion in the gene of interest from random pools of worms by means of PCR (reviewed by Jansen, 2002). Deletions can be generated using either transposons (normally Tc1) or chemical mutagens, such as EMS and UV activated trimethylpsoralen (TMP/UV), and result in stable gene inactivation. Transposon mutagenesis is a two-step procedure of isolating a transposon insertion in the gene of interest, followed by imprecise transposon excision causing a deletion (Zwaal *et al.*, 1993). Chemical mutagenesis is a more direct, one-step process for the detection of deletions in the gene of interest (Jansen *et al.*, 1997).

A complementary approach to the creation of genetic knockouts is the use of RNAi. This technique is widely utilised and provides a quick means to disrupt the function of a specific gene (reviewed by Maine, 2001). The effect is transient and it does not replace the need for the generation of mutants. There are various ways to induce RNAi: by injection of the dsRNA into the intestine (ideally), body cavity or gonad of a hermaphrodite young adult (Guo and Kemphues, 1995; Fire *et al.*, 1998); soaking L3-L4 larvae in concentrated dsRNA solutions (Tabara *et al.*, 1998); feeding L3-L4 worms with *E. coli* expressing dsRNA (Timmons and Fire, 1998; Timmons *et al.*, 2001); and *in vivo*, promoter driven expression of heritable inverted repeat genes in transgenic animals (Tavernarakis *et al.*, 2000).

The other means to reach an initial understanding of a gene function is the characterisation of its expression profile. The change in expression of a gene, that is when and where the gene is expressed, can be determined in *C. elegans* by traditional methods such as *in situ* hybridisation, immunohistochemistry, reporter gene fusion transgenes, and more recently, DNA chips, serial

analysis of gene expression (SAGE) and DNA microarrays (reviewed by Reinke, 2002). The above conventional techniques are dealt with detail in Methods in Cell Biology (1995); a later introduced method in *C. elegans*, microarray analysis, will be discussed here. DNA microarrays are glass slides containing discrete spots of DNA, each one corresponding to a coding region of an annotated gene. The gene fragments are amplified by PCR from either expressed sequence tags (EST) or genomic DNA. Hybridisation of two differently labelled cDNA samples to the microarray provides transcription expression profiles, by comparing the relative expression of each gene in both samples. Two groups have prepared DNA microarrays of most of the worm genes: the Kim laboratory at Stanford University, USA (see http://cmgm.stanford.edu/~kimlab for details), and the Kuwabara group at the Sanger Centre, UK (http://www.sanger.ac.uk/ Teams/Team20/microarrays/). Both microarray resources are accessible to the *C. elegans* community.

1.2.2.1 Functional genomics

Having the genome sequence in hand, the need to boost genetic analysis (only a small fraction of genes have been characterised thus far) and technical advances have prompted several groups to embark on diverse functional genomics projects as detailed in reviews by Sternberg (2001) and Kim (2001b). These projects have been designed to study specific biological aspects such as mutant phenotype, gene expression and protein-protein interactions of numerous genes in parallel. Some of the current studies are listed below.

A very informative approach to the role of a gene is the analysis of its loss of function phenotype. Four groups have reported the use of RNAi in order to document genes' knockdown effects on a large scale. Julie Ahringer's group targeted ~86% of the 19,427 predicted genes in *C. elegans* (Fraser *et al.*, 2000; Kamath *et al.*, 2003) and Tony Hyman's group ~96% of the annotated genes on chromosome III (Gonczy *et al.*, 2000). Ken Kemphues' group used cDNA clones from ovaries (Piano *et al.*, 2000), while Asaka Sigimoto's group used cDNAs from a mixed-stage library (Maeda *et al.*, 2001). Different methods of delivering dsRNA and annotation of the mutant phenotypes were applied in these studies. Both Hyman and Kemphues groups used the RNAi injection method and analysed the first few divisions of embryos by time-lapse differential interference contrast (DIC) microscopy. In contrast, Ahringer and Sagimoto groups characterised the RNAi phenotypes more broadly, the former applying the feeding strategy and the latter the soaking technique. The approach taken by J. Ahringer and co-workers provides the important advantage of generating a library of clones that contain the PCR fragments of all tested genes maintained in the form of bacterial stocks. This library is then a

permanent resource for more in-depth RNAi analyses (available from the UK HGMP Resource Centre, http://www.hgmp.mrc.ac.uk/geneservice/index.shtml). Detailed information of the above projects have been published recently (Hope, 2001; Kim, 2001a; Piano and Gunsalus, 2002). A second technology applied in a systematic fashion to inactivate genes is targeted deletion by chemical mutagenesis. The *C. elegans* Knockout Consortium is a group of four laboratories with the long-term goal of generating null mutants of all worm genes. At present, they are concentrating on deleting genes requested by researchers in the *C. elegans* community (http://elegans.bcgsc.bc.ca/knockout.shtml).

As previously mentioned, another way to elucidate gene function is the definition of the expression pattern. At the University of Leeds, England, Ian Hope has been collecting expression data for the annotated *C. elegans* genes using reporter gene fusion technology (Hope, 1991; Young and Hope, 1993; Lynch *et al.*, 1995; Hope *et al.*, 1998). At the National Genetics Institute (Mishima, Japan), Yuji Kohara aims to examine the expression pattern of a set of sequenced-tagged cDNAs at multiple stages of development by *in situ* hybridisation (http://nematode.lab.nig.ac.jp). Several genome-wide studies have been undertaken using DNA arrays, including the two analyses of the expression of most genes throughout *C. elegans* development carried out by Hill *et al.*, (2000b) using Affymetrix DNA chips and Jiang *et al.*, (2001) with DNA microarrays. The latter group also examined gene expression changes between males and hermaphrodites.

Finally, Marc Vidal's group at Harvard University, USA, is finding protein-protein interactions for all *C. elegans* proteins using the two-hybrid system. They first generated a library of clones containing a large number of protein-encoding open reading frames (Reboul *et al.*, 2001) and developed the two-hybrid technique for a large-scale approach. They have already described interactions between proteins involved in four specific biological processes and now aim to perform more general studies. A summary of technical basics plus the current status of the interaction mapping project is dealt by Rual *et al.*, (2002).

The data generated by the above projects will require new computational tools to be integrated with the current knowledge about *C. elegans*, other model organisms and humans. Thus, predicted functional information obtained by a single approach will consolidate within the newly synthesised context. Public access to this core of data will be essential and invaluable for individual labs to develop their research interests. Understanding of biological processes and networks is becoming feasible for the first time in a multicellular organism.

1.3 EXTRACELLULAR MATRIX

All metazoans produce extracellular matrices (ECMs), complex networks of secreted glycoproteins and carbohydrates. There are many types of ECM, differing between organisms, tissues and stages of development, and each is specialised for a particular function. ECMs play structural and instructional roles and are critical in cellular differentiation, development and determination of body form. In general, they provide a substrate for cell migration, act as a barrier or selective filter to soluble molecules, strengthen tissues, and interact with growth factors affecting their local concentration and biological activity. The most abundant component of ECMs is collagen. There are more than 20 collagen types in humans; a large number of mutations in approximately half of them have been identified and associated to a wide spectrum of diseases, manifested in several forms varying in phenotypic severity (Myllyharju and Kivirikko, 2001). Osteogenesis imperfecta, some cases of osteoporosis, osteoarthrosis, many chondrodysplasias, intervertebral disc disease, Alport syndrome, Bethlem myopathy, Knobloch syndrome, arterial aneurysms, and several subtypes of the Ehlers-Danlos syndrome and of epidermolysis bullosa are instances of these collagen related abnormalities. Some genetic alterations can result in excessive accumulation of collagen fibrils in certain tissues impairing their function. A variety of approaches are being carried out in order to alter collagen synthesis in fibrotic conditions at transcriptional, translational and post-translational levels (Prockop and Kivirikko, 1995). The approach followed to modify post-translational collagen synthesis consists of the use of drugs against collagen-specific processing enzymes.

1.3.1 Structural features of collagens

Collagen proteins are defined by their amino acid composition, high in glycine, proline and hydroxyproline (Hyp) and by their unique triple helical structure. The distinctive amino acid content is due to relatively large stretches of the repeating sequence Gly-X-Y, where X is often Pro and Y is often Hyp, in the polypeptide chains. Single polypeptide chains adopt a left-handed helix conformation and assemble into trimers to form a right-handed super-helix, a collagen molecule. Every third residue of each strand passes through the crowded centre of the triple helix, hence the requirement for a Gly at this position. Pro and Hyp are positioned on the surface of the molecule also for spatial and stereochemical reasons. The entire assembly is stabilised by the bulky and relatively inflexible Pro and Hyp residues, hydrogen bonds and disulfide bridges.

In humans, folding of the collagen molecule requires first proper association and in-register alignment of the three chains, which can then zip together in the C→N direction to form the three stranded helix (Engel and Prockop, 1991). The large globular C-terminal domains of fibrillar collagens, in particular their last 10 amino acids as demonstrated in Lim *et al.*, (1998), are critical in the initial steps of assembly. In contrast, FACIT collagens (see below) appear to contain the information necessary for trimer formation in both the distal collagen domain and the short C-terminal region (Mazzorana *et al.*, 1993). Homo and heterotrimers can be found in many different types of collagens. A large number of the disorders noted above are caused by heterozygous mutations with a dominant negative effect (Prockop and Kivirikko, 1995). Mutant collagen chains affect the function of co-synthesised normal chains interfering with their folding and/or posterior assembly. Because of the rigid structural requirements of collagen molecules, many of these mutations are single-base substitutions of glycine residues (Forlino and Marini, 2000; Olague-Marchan *et al.*, 2000).

1.3.2 Types of collagen in humans

Although sharing basic characteristics, collagens are a diverse group of proteins. They can be classified into different varieties and then into families on the basis of their structural features and the assemblies they form. The collagen classes are: fibril forming collagens (types I, II, III, V and XI), network forming collagens (types IV, VIII and X), FACIT collagens (fibril associated collagens with interrupted triple helices include types IX, XII, XIV, XVI and XIX), beaded filament forming collagen (type VI), collagen of anchoring fibrils (type VII), transmembrane collagens (type XIII, XVII) and the not fully characterised collagens types XV and XVIII (Prockop and Kivirikko, 1995).

FACIT collagens possess certain structural characteristics that make them a distinctive class in the collagen super-family (Shaw and Olsen, 1991). They are relatively small, are not synthesised as procollagens and comprise more than one collagen domain separated by non-collagen regions. Most of the collagen domains tend to include imperfections and a few cysteine residues. A significant number of cysteines are present in non-collagen domains, often located at the end of the previous collagen domains. Moreover, FACIT collagens are proteoglycans; they have glycosaminoglycan chains attached to non-collagen domains, and do not polymerise but bind to the surface of collagen fibrils.

1.3.3 ECM in *C. elegans*

There are two different types of ECMs in *C. elegans*: basement membranes and the cuticle. Basement membranes are closely associated with cell membranes separating most internal organs from the pseudocoelomic cavity and underlying the hypodermis. Type IV collagen is the major component of basement membranes and is exclusively found in this kind of ECM. The type IV collagen structure is highly homologous in different organisms such as *C. elegans*, parasitic nematodes and humans, indicating its conservation throughout evolution.

The cuticle is an exoskeleton that encloses the animal and lines mouth, pharynx and rectum. It is constituted of collagen proteins; non-collagenous proteins such as cuticlin, surface associated proteins and glycoproteins; and lipids. Chapter 3 deals in detail with this type of ECM and its primary components, the collagen proteins.

The favourable characteristics of *C. elegans* for experimental studies and the immense biological and genetic information available about this nematode, including its deciphered genome, make it an excellent organism to investigate processes common in multicellular animals. I have used *C. elegans* as a model organism to study collagen biosynthesis. For this, I characterised three distinct mutants failing to form an intact cuticle, identified the causative genes and defined them molecularly.

Chapter 2

Materials and Methods

2. MATERIALS

Unless otherwise stated all chemicals and other materials were purchased from the following commercial sources: AB gene (http://www.abgene.co.uk), Amersham Life Science (http://www.amersham.co.uk), Bio-Rad (http://www.biorad.co.uk), Difco (http://www.bd.com/industrial), Fisher Scientific (http://www.fisher.co.uk), GibcoBRL Life Technologies (http://www.lifetechnologies.com), Invitrogen (http://www.invitrogen.com), Molecular Probes (http://www.probes.com), New England Biolabs (NEB; http://www.neb.com), Novagen (http://www.novagen.com), Pierce (http://www.priercenet.com), Promega (http://www.promega.com/uk), Qiagen (http://www.qiagen.com), Roche (http://www.roche.com), Sigma (http://www.sigmaaldrich.com) and Stratagene (http://www.stratagene.com).

2.1 CHEMICAL ABBREVIATIONS

BCIP/NBT: 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium

BSA: bovine serum albumin

β-ME: β-mercaptoethanol

DABCO: 1,4-diazabicyclo-2,2,2,-octane

DMF: N,N'-dimethylformamide

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

dH$_2$O: distilled water

FITC: fluorescein isothiocyanate

IPTG: isopropyl-β-D-thiogalactoside

MOPS: 3-(N-morpholino)propanesulfonic acid

PBS: phosphate buffered saline

PEG: polyethylene glycol

SDS: sodium dodecyl sulphate or lauryl sulphate

Tris-HCl: 2-Amino-2-(hydroxymethyl)-1,3-propanediol-hydrochloride

Tween-20: polyoxyethylenesorbitan monolaurate

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

2.2 SOLUTIONS AND MEDIA

10 x gel loading buffer: 20 mM EDTA, 25% ficoll (type 400), 0.2% bromophenol blue.

Nematode growth medium (NGM) agar: 0.3% NaCl, 0.25% of Bacto-peptone (Difco), 1.7% agar (Difco), 0.0005% cholesterol (1 ml/L of 5 mg/ml stock in EtOH), in dH_2O. Autoclaved and, after the medium cooled, the following solutions were added using sterile technique: 1ml/L of 1 M $CaCl_2$, 1 ml/L of 1 M $MgSO_4$ and 25 ml/L of 1 M potassium phosphate pH 6, prior to pouring petri plates.

NGM agarose: as above but agar was substituted with agarose in the medium.

Potassium phosphate buffer (1 M, pH 6): obtained by mixing 13.2 ml of 1 M KH_2PO_4 and 86.8 ml of 1 M K_2HPO_4 to reach pH 6 in a 100 ml solution.

Alkaline hypochlorite solution: 2 volumes of 4 M NaOH:3 volumes of fresh 10-20% NaOCl.

Freezing solution: 5.85 g NaCl, 6.8 g KH_2PO_4, 300 g glycerol, 5.6 ml of 1 M NaOH plus dH_2O up to 1 L. Autoclaved and 3 ml of 0.1 M $MgSO_4$ added.

Luria-Bertani (LB) broth: 10 g bacto-tryptone, 5 g bacto-yeast extract, 5 g NaCl, dH_2O to 1 L. pH adjusted to 7.0 with NaOH. Sterilised by autoclaving.

LB agar: 10 g bacto-tryptone, 5 g bacto-yeast extract, 5 g NaCl and 15 g agar in 1 L of dH_2O. pH to 7.5 and autoclaved. If required, antibiotics were added to the cooled medium before pouring sterilely onto 9 cm plates. Unused plates were stored at 4°C for short periods.

SOB medium: 20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl, 10 ml of 250 mM KCl and dH_2O up to 1 L. pH was adjusted to 7.0 with NaOH and autoclaved. Before use, 5 ml of a sterile solution of 2 M $MgCl_2$ were added.

SOC medium: 20 ml of a filter-sterilised solution of 1 M glucose added to 1 L SOB medium.

M9 buffer: 3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl, 1 ml 1 M $MgSO_4$ in 1 L dH_2O. Mixed and autoclaved.

1 x worm lysis buffer (1 x WLB): 50 mM Tris-HCl pH 8, 100 mM NaCl, 50 mM EDTA, 1% SDS, 30 mM β-ME, 100 µg/ml proteinase K. Buffer was freshly made just before use.

TE buffer: 10 mM Tris and 1 mM EDTA in dH$_2$O. pH as required.

10 x Tris borate buffer (TBE): 465.2 g TRIS base and 205.2 g boric acid dissolved in 4 L of dH$_2$O by stirring. Then 14.88 g of EDTA added and mixed. The working solution was 1 x TBE.

Single worm lysis buffer (SWLB): 50 mM KCl, 10 mM Tris-HCl pH 8.3, 2.5 mM MgCl$_2$, 0.45% Tween-20, 0.01% (w/v) gelatin. Autoclaved and stored in aliquots at -20°C; proteinase K added to 60 μg/ml before use.

Chitinase solution: 20 mg/ml chitinase in 50 mM NaCl, 70 mM KCl, 2.5 mM MgCl$_2$. Stored in aliquots at -20°C.

Colony lysis buffer: 0.3 g ficoll, 0.1 g SDS, small amount of bromophenol blue dye and 1 x TBE buffer to 10 ml volume.

10 x injection buffer: 200 mM KPO$_4$ (pH 7.5), 30 mM K citrate (pH 7.5) and 20% PEG 6000. Stored at 4°C.

Staining solution for β-galactosidase activity: For 500 μl of staining solution the following was added in order: 145 μl of dH$_2$O, 40 μl 0.5 M NaH$_2$PO$_4$, 210 μl 0.5 M Na$_2$HPO$_4$, 70 μl 100 mM MgCl$_2$, 5 μl 4 mg/ml SDS, 10 μl 250 mM K$_4$FeCN$_6$, 10 μl 250 mM K$_3$FeCN$_6$, 1 μl 50 mg/ml kanamycin, 9 μl 2% X-gal in DMF (added last and mixed quickly to avoid precipitation). The solution could be stored for at least two days wrapped with foil at 4°C.

Sensitive staining solution for β-galactosidase activity: as above but heating the solution at 65°C before adding 12 μl of 20% X-gal. Mixed and retained at 65°C until used.

2% (w/v) X-gal: 0.2 g of X-gal dissolved in 10 ml of DMF, aliquoted and stored at -20°C. Aliquots were covered with aluminium foil to prevent damage by light.

PBST: 1 x PBS and 0.2% Tween-20.

Mounting solution: For 100 ml of solution, 50 μl glycerol, 10 μl DABCO and 40 μl 1 x PBS were mixed by vortex. Made fresh before use.

17

5 x SDS loading buffer: 0.225 M Tris-HCl pH 6.8, 50% glycerol, 5% SDS, 0.05% bromophenol blue, 0.25 M β-ME (or DTT).

Coomassie blue solution: 0.25 g of Coomassie brilliant blue R-250, 45% methanol, 10% glacial acetic acid. The solution was filtered through a Whatman no. 1 filter to remove particles.

Destain solution: 45% methanol, 10% glacial acetic acid.

20 x MOPS SDS running buffer: 1 M MOPS, 1 M Tris base, 69.3 mM SDS, 20.5 mM EDTA

2.3　OTHER MATERIALS

2.3.1　Markers

• **DNA markers:**

1 Kb DNA ladder: 75, 134, 154, 201, 220, 298, 344, 396, 506, 517, 1018, 1636, 2036, 3054, 4072, 5090, 6108, 7126, 8144, 9162, 10180, 11198, 12216 bp (Invitrogen). The ~1.6 Kb fragment served as reference band.

100 bp DNA ladder: 100, 200, 300, 400, 500, 517, 600, 700, 800, 900, 1000, 1200, 1517 bp (NEB). The 500 and 1000 bp fragments served as reference bands.

50 bp DNA ladder: 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800 bp (GibcoBRL Life Technologies).

Stock solutions: 1 μg/μl 1 Kb ladder, 0.5 μg/μl 100 bp ladder, 1 μg/μl 50 bp ladder.

Working solutions were prepared as follows:

500 μl of 50 bp ladder or 100 bp ladder: 25 μl DNA ladder, 50 μl 10 x gel loading buffer, 425 μl dH$_2$O.

500 μl of 1 Kb ladder: 12.5 μl DNA ladder, 50 μl 10 x gel loading buffer, 437.5 μl dH$_2$O.

10 μl of the working preparations were normally used in DNA analysis by electrophoresis.

• Protein markers:

Broad range 2-212 kDa: 2/3, 7, 14, 20, 27, 36, 43, 56, 66, 97, 116, 158, 212 (NEB).

SeeBlue Plus2 pre-stained standard: 3, 6, 14, 17, 28, 38, 49, 62, 98, 188 kDa (Invitrogen).

2.3.2 Plasmids and constructs

Table 2.1: Commercial vectors and constructs. A. Fire (Carnegie Institute of Washington, Baltimore, USA); J. Kramer (Northwestern University Medical School, Chicago, USA); I. Johnstone (Glasgow University, Glasgow, UK).

Name	Description	Use	Source /Company
pBluescript SK (+/-)	2958 bp; Amp^r; blue-white selection	standard cloning	Stratagene
pCR-Script Amp SK(+)	3000 bp; Amp^r; blue/white selection ligation efficiency increased by including a *Srf* I restriction enzyme	blunt-end cloning	Stratagene
pGEM-T	3000 bp; Amp^r; blue-white selection	TA cloning	Promega
pCR2.1-TOPO	3908 bp; Amp^r and Kan^r; blue-white screening	TA cloning	Invitrogen
pCRII-TOPO	3950 bp; Amp^r and Kan^r; blue-white screening; T7 and Sp6 promoter priming sites	TA cloning and *in vitro* RNA transcription	Invitrogen
pCR-XL-TOPO	3519 bp; Kan^r; *ccdB* eliminates background; blue-white screening	TA cloning for long PCR products	Invitrogen
pPD96.04	8119 bp; contains *GFP* and *lacZ* reporter genes	gene expression analysis	A. Fire
L4440 (pPD129.36)	2790 bp; RNAi feeding vector; contains two T7 promoters in opposite directions	RNAi feeding	A. Fire
pQE-30	3400 bp; expression vector; designed for expression of N-terminal 6 × His-tagged proteins	recombinant protein expression	Qiagen
pTAg	3816 bp; Amp^r Kan^r; blue/white selection	*C. elegans* DNA transformation	Novagen
pRF4	contains *rol-6(su1006)*; a dominant cuticle collagen mutation	*C. elegans* DNA transformation	J. Kramer
pMW0002	*dpy-7::GFP*; *dpy-7* cuticle collagen promoter in pPD95.67 (A. Fire)	*C. elegans* DNA transformation	I. Johnstone

Name	Description	Use	Source /Company
pMW0003	*col-12::GFP*; *col-12* cuticle collagen promoter in pPD95.67 (A. Fire)	*C. elegans* DNA transformation	I. Johnstone
pMW025	*vc5.3::GFP::lacZ*; *vc5.3* intestinal gene promoter in pPD96.04 (A. Fire)	*C. elegans* DNA transformation	I. Johnstone

2.3.3 Oligonucleotides

2.3.3.1 Primers used for stc-1

Table 2.2: Primers utilised for deficiency endpoint mapping.

Primer Name	Sequence (5'-3')	Deficiency characterised
M110-sense	CTGGGAGAGGGAAGCCGAATCAAG	*mnDf88/mnDf69*
M110-anti	TCCGGTCGCTGATATCCACCACTG	*mnDf88/mnDf69*
C26D10-sense	TTCAGAGACGCCGATAACGATCAAC	*mnDf88/mnDf69*
C26D10-anti	TCGGAAAGGCGATATTTTGGAAATG	*mnDf88/mnDf69*
F14E5-sense	ATATGATTTAACCAGCCGTGGACCG	*mnDf88/mnDf69*
F14E5-anti	AGGGCGGATGTGTGAGTTTCCTC	*mnDf88/mnDf69*
F22B5-sense	TCGTTGAGCAAAGCCCCACG	*mnDf88/mnDf69*
F22B5-anti	AAGTCTCTCCCAGTCCCATGGTGC	*mnDf88/mnDf69*
M05D6-sense	ACGTTCTCGCAGTATCTCCGAATCC	*mnDf88/mnDf69*
M05D6-anti	TGTGCAATTTGTTGGGGTGGGTC	*mnDf88/mnDf69*
W01C9-sense	CGGAGGTTGAGTGAAGAGAGCATCC	*mnDf88/mnDf69*
W01C9-anti	CAGTCGGCTGCTGTAGTCCCATAAG	*mnDf88/mnDf69*
D2085-sense	GCGACTTTGCATCTCGAGGACG	*mnDf88/mnDf69*
D2085-anti	CGATCACAAGCTTCGCCTGAAAAC	*mnDf88/mnDf69*
T13H5-sense	AAGGTACCCTTCTGAACTCTGC	*mnDf88/mnDf69*
T13H5-anti	AACGTTGGCTGGATTCTCAG	*mnDf88/mnDf69*
F14E5up-sense	GGGCACAGGTGTCACAGTGGATTTG	*mnDf88*
F14E5up-anti	TGTCGTGTCAAGACCGGGGTTG	*mnDf88*
F49E12up-sense	TTTCGATGGTGACCCACTGCTTATC	*mnDf88*
F49E12up-anti	TTGAAACTTACAAACTGGCGCCG	*mnDf88*
F49E12middle-sense	CTCTCCAGGCTTCAAGGTTCAATCC	*mnDf88*
F49E12middle-anti	AACGATTGCAATGAAACGACACTGC	*mnDf88*
F49E12down-sense	TATCTTCGGCTCTTCAAACGCAATG	*mnDf88*
F49E12down-anti	GGATACGAGTTCAGGAATGGCTTGC	*mnDf88*
M195up-sense	TCCTGGAGCAGTTCTTGGTGGTGG	*mnDf88*
M195up-anti	TGACTTCTGCTTCGACTGCCCAGC	*mnDf88*

Primer Name	Sequence (5'-3')	Deficiency characterised
M195down-sense	GGAATTCCCGCACCGAAGTTTTC	*mnDf88*
M195down-anti	CCTTGCAGCACTGTCCATCAGAGTG	*mnDf88*
T22C8up-sense	CCTTGCTAATTTAATGTCTCCGCCG	*mnDf69*
T22C8up-anti	GGGCGTGGGGAATGCTATTTACC	*mnDf69*
T22C8down-sense	TGCTCCCCATGCAATACCATATGAC	*mnDf69*
T22C8down-anti	CCGAATTGCGACAATGATTTCTGG	*mnDf69*
D2085down-sense	TCCGTGTGAAAATGGTTAATTGGGC	*mnDf69*
D2085down-anti	CCAATCCCATGTCACTTGTGACCAC	*mnDf69*
C34C6up-sense	CAGAAAAGATCCACAGCGAGGTGAG	*mnDf69*
C34C6up-anti	CGCCGACAAACAGTCCAAGAGTG	*mnDf69*
C34C6down-sense	AACCGAACATAAAAGTCCGTTTCCG	*mnDf69*
C34C6down-anti	TTTTGCGAATACAGTTGGACCGAAG	*mnDf69*
T01B7up-sense	TTTTCGGCGGATTTTATCCTGTCTG	*mnDf69*
T01B7up-anti	CGCGCCTTTAAAGTCGTAGAAGTGG	*mnDf69*
left_rol6-sense	TCTTTGTGGAAGAAACGAACACGGC	*mnDf69*
left_rol6-anti	AGCAGTTTGTCTCCTCAACGGGC	*mnDf69*
right_rol6-sense	GATTGGGTGTTTGCATTTGGCG	*mnDf69*
right_rol6-anti	TGGATGCTGCTGTTGTAGACCAAGG	*mnDf69*

Table 2.3: Primers used for cloning of *stc-1*.

Primer Name	Sequence (5'-3')	Use
senseF54C9.4(RNAi)	AATTAACCCTCACTAAAGGGAGAAGTG CTTCTTGGCTCTAGCTGGGCA	RNAi (also colony screening of F54C9 clones)
antiF54C9.4(RNAi)	TAATACGACTCACTATAGGGAGACGTT TTCCGTAACAAGCGTTCTGCT	RNAi (and colony screening of F54C9 clones)
senseF54C9.6(RNAi)	AATTAACCCTCACTAAAGGGAGATTAG AGCGTATCCATGGCTTCTCGA	RNAi (also colony screening of F54C9 clones)
antiF54C9.6(RNAi)	TAATACGACTCACTATAGGGAGACAAA ATGACTCGCTAATGCGGAAAT	RNAi (also colony screening of F54C9 clones)
senseF54C9.8(RNAi)	AATTAACCCTCACTAAAGGGAGATCAT CGTCCAACATTCCTCTGTTCA	RNAi (also colony screening of F54C9 clones)
antiF54C9.8(RNAi)	TAATACGACTCACTATAGGGAGACGTC AAGGGTTGCATGTTCAACA	RNAi (also colony screening of F54C9 clones)
senseF54C9.9(RNAi)	AATTAACCCTCACTAAAGGGAGACGTC GGAAATCGAAAAGGAACGC	RNAi (also colony screening of F54C9 clones)
antiF54C9.9(RNAi)	TAATACGACTCACTATAGGGAGAGACT TCATTTTCAATGGCTTCCGTC	RNAi (also colony screening of F54C9 clones)
senseF54C9.10(RNAi)	AATTAACCCTCACTAAAGGGAGATGTG TTCAGATCAGAAGCAATGGGA	RNAi (also colony screening of F54C9 clones)

Primer Name	Sequence (5'-3')	Use
antiF54C9.10(RNAi)	TAATACGACTCACTATAGGGAGAGATT ACGAAGAGCATCCAATCCCAA	RNAi (also colony screening of F54C9 clones)
senseF54C9.5	ACAACGTTTTCTCGTTTGGAAGCGA	cloning of F54C9.5
antiF54C9.5	ATCGTACTCGAACAAATTGT	cloning of F54C9.5
senseF54C9.10	TTTTCGCAACTCACCTACTTCCTCA	cloning of F54C9.10
antiF54C9.10	GACACACCTCGTGTTTGCAGCTAAA	cloning of F54C9.10
senseF54C9.3	GCCTGCCATCGCCTTCTTCT	cloning of F54C9.3
antiF54C9.3	CCCAGCTAGAGCCAAGAAGCACTAA	cloning of F54C9.3 and gene expression analysis
senseF54C9.2	CACCAAAGAAAGGAGCATCAGATGA	cloning of F54C9.2 and F54C9.3 (using also antiF54C9.3)
"2ND"senseF54C9.2	GAAAGGAGCATCAGATGAGCAACGA	cloning of F54C9.2
"2ND"antiF54C9.2	GCGGAAGACTGACTTATCGTTGCA	cloning of F54C9.2 and also spatial expression

Table 2.4: Other primers employed for *stc-1*.

Primer Name	Sequence (5'-3')	Use
senseF54C9.1	AACACGAGCACTTGATGATGGGTT	colony screening
antiF54C9.1	TTCCTTTCCCCTCAAACTACACGAA	colony screening
senseF54C9.2	CACCAAAGAAAGGAGCATCAGATGA	colony screening
antiF54C9.2	GCTTGGTATACGGTACCCAGCAACA	colony screening
senseF54C9.3	GCCTGCCATCGCCTTCTTCT	colony screening
antiF54C9.10	GACACACCTCGTGTTTGCAGCTAAA	colony screening
antiF54C9.11	GCTTTCAAAACGGAGCCGATAAAA	colony screening
rtpcr-anti-F54C9.2	GATGGAACATTCGTCAGCCTGATTT	temporal expression
rtpcr-sense-F54C9.2	GGAGAGTAATTTCCGAGCCGACA	temporal expression

Primer Name	Sequence (5'-3')	Use
hsp70-GFP sense-	GCGGGATCCATCGCCTTCTTCTGTC	spatial expression (also used 2ND-antiF54C9.2 and antiF54C9.3
antiRNAi-F54C9.2	ACCGAGTGTCGTTGCTTTCTTGTAA	RNAi
senseRNAi-F54C9.2	ATTTCGTCCACATCAGCAGTATCCA	RNAi
senseF54C9.2(E1-4)	GGCCTCCCAAACGATTATTTCCA	sequencing
antiF54C9.2(E1-4)	TCAGCGCCTGAAATTCGTCAA	sequencing
1st(E1-4;ANTI)	GGATTACCTCCACCACCTCCAAAAA	sequencing
2nd(E1-4;ANTI)	TGATGCAAAACGATTTATTGGACGA	sequencing
senseF54C9.2(E5-7)	TTCTAGCAATGCAGATTTGATCCCA	sequencing
antiF54C9.2(E5-7)	GTTGGCATACGGATTGCACAAAA	sequencing
1st(E5-7;ANTI)	GCTGACGAATGTTCCATCCACAAC	sequencing
2nd(E5-7;ANTI)	CCGTCAGAAAAATTGTGGGCAGA	sequencing
F54C9.2-EC1S	CGTAGGATCCATAGGTATCTACCAT GCTGTAACTGGTG	protein expression
F54C9.2-EC1A	CGTAAAGCTTGGAAATTACTCTCCG AACTTCCAT	protein expression
F54C9.2-EC2S	CGTAGGATCCCACAAAAAACAAGG AGTTGAAAATGTTG	protein expression
F54C9.2-EC2A	CGTAAAGCTTGTCCACATCAGCAGT ATCCAAATTTGC	protein expression

2.3.3.2 Primers used for let-607

Table 2.5: Primers employed for *let-607* sequencing and gene structure.

Primer Name	Sequence (5'-3')	Use
1st.fragment-sense	TCGGCATCCTACAGTACCCTCTGAA	sequencing
1st.fragment-anti	TGAATGAACAGAAAATAGCGCACCA	sequencing
2nd.fragment-sense	GCGCCCAGTCCCTACAAGGTTT	sequencing
2nd.fragment-anti	ACCACCTGGCACTGGGTCAATAT	sequencing

Primer Name	Sequence (5'-3')	Use
3rd.fragment-sense	GGGAAGAAGCGGTTTCTTAACGAAA	sequencing
3rd.fragment-anti	TGAATTACCGGGTCTTGAAGCGA	sequencing
1sense(1st)	CCATATCAAGCACTCGAAGACCCCT	sequencing; gene intron-exon structure
2sense(1st)	GCCTCTCTCCACCAGATTCAGGA	sequencing; gene intron-exon structure
3sense(1st)	AATGTGCAGCAGAGAAAGCAGCATA	sequencing
0anti(1st)	CAGATCCGACTTTCCGTCCATTTT	sequencing; 5' RACE; SL1+SL2 PCRs
1anti(1st)	CTGACCTTTGCCTAGAATGAAATGA	sequencing
2anti(1st)	GCAACCGAATTGGCATTTGTGTAG	sequencing; 5' RACE; gene intron-exon structure
1sense(3rd)	CCCAACTCAAAAAACTTCAAGCTCA	sequencing
2sense(3rd)	TGAATCGGAATGCTGTTCTCAATCA	sequencing; 3' RACE
3sense(3rd)	CGTACTTTGGGTGCATTCGAAGATC	sequencing; 3' RACE
4sense(3rd)	GGTGCCACAACAAAAGGTGCAATAT	sequencing
0anti(3rd)	TGCATTTTGTCCAAGTTGAGCTTGA	sequencing; 5' RACE; gene intron-exon structure
1anti(3rd)	GCATTCCGATTCATTTGGCGAT	sequencing
2anti(3rd)	CCCAAAGTACGCGAGGTACGGTAG	sequencing; gene intron-exon structure
3anti(3rd)	GACCAGGAGTTTGGAAGACGAATTG	sequencing; gene intron-exon structure
sense-E4-F57B10.1	AACCGAGAGTCCTCAATCCCGCA	sequencing

Table 2.6: Other primers utilised for *let-607*.

Primer Name	Sequence (5'-3')	Use
F57B10-1 a1	GGTTACTTCATGAGGATTGGTTAA	cloning of the gene
F57B10-1 up2	TGCGTGTTCAACTCTGGATATT	cloning of the gene; spatial expression
F57B10-1 GFPanti	CGGAGGATCCGAGCCTGGAGACCATATTCC	spatial expression
F57B10-1 ATG	ATGGACCAAGATTTTGACCT	checking gene structure
F57B10-1 rtpcr1	ATCGAGCAATTGGAAGATCG	temporal expression
F57B10-1 rtpcr2	TCGCAGAGTTCATGGAAGTT	temporal expression

Primer Name	Sequence (5'-3')	Use
ama-1a2	CAGTGGCTCATGTCGAGTTTCCAGA	temporal expression
ama-1b2	CGACCTTCTTTCCATCATTCATCGG	temporal expression
F57B10.1 T3sense	TATTAACCCTCACTAAAGGGAGAA TTTTCAGTTGTTTCCGATCCAG	RNAi
F57B10.1 T7anti	TAATACGACTCACTATAGGGAGAC CAACTGTGCAAACACC	RNAi
AUAP	GGCCACGCGTCGACTAGTAC	5' and 3' RACE
AAP	GGCCACGCGTCGACTAGTACGGG IIGGGIIGGGIIG	5' and 3' RACE
AP	GGCCACGCGTCGACTAGTACTT TTTTTTTTT TTTTTT	3' RACE
SL1	GGTTTAATTACCCAAGTTTGAG	5' end PCR
SL2	GGTTTTAACCCAGTTACTCAAG	5' end PCR
F57B10.1-ES	CGTAGGATCCATGGACCAAGAT TTTGACCTCGATG	protein expression
F57B10.1-EA	CGTACTGCAGTTAAAACATTTGG GTCTTTATTTTCTTAGC	protein expression

2.3.3.3 Vector primers

Table 2.7: Vector-specific primers.

Primer Name	Sequence (5'-3')	Use
M13(-20) Forward	GTAAAACGACGGCCAG	sequencing; colony screening
M13 Reverse	CAGGAAACAGCTATGAC	sequencing; colony screening
T3	AATTAACCCTCACTAAAGGG	sequencing; colony screening
T7	GTAATACGACTCACTATAGGGC	sequencing; colony screening
pPD96.04-anti	CCACTGACAGAAAATTTGTGCCCAT	colony screening
1224	CGCCAGGGTTTTCCCAGTCACGAC	colony screening
pEQ-30 Foward	TTTGCTTTGTGAGCGGATAAC	colony screening
pEQ-30 Reverse	AGCTAGCTTGGATTCTCACC	colony screening

25

2.3.4 Strains

2.3.4.1 Escherichia coli strains

Table 2.8: *E. coli* strains. CGC = *Caenorhabditis* Genetic Centre

Strain name and source	Genotype / comments
OP50 (CGC)	uracil auxotroph
Epicurian Coli XL10-Gold Ultracompetent Cells (Stratagene)	Tet^{r}Δ(mcrA)183Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac The [F' proAB lacI^{q}ZΔM15 Tn10 (Tet^r) Amy Cam^r] Used for standard transformations
Epicurian Coli XL10-Gold Kan Ultracompetent Cells (Stratagene)	Tet^{r}Δ(mcrA)183Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 Kan recA1 gyrA96 relA1 lac The [F' proAB lacI^{q}ZΔM15 Tn10 (Tet^r) Tn5 (Kan^r Amy] Recommended host strain for pPCR-Script Amp SK(+) vector
TOP10F' One Shot Chemically Competent Cells (Invitrogen)	F' [lacI^q Tn10(Tet^r)] mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL (Str^r) endA1 nupGΔ Used for transformations with pCR2.1-TOPO and PCRII-TOPO vectors
TOP10 One Shot Electrocompetent Cells (Invitrogen)	F' mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL (Str^r) endA1 nupGΔ Recommended strain for pCR-XL-TOPO vector
M15 Competent Cells (Qiagen)	Nal^S, Str^S, Rif^S, Thi$^-$, Lac$^-$, Ara$^+$, Gal$^+$, Mtl$^-$, F$^-$, RecA$^+$, Uvr$^+$, Lon$^+$ Host strain for expression of recombinant proteins
HT115 (DE3) (CGC)	F$^-$ mcrA mcrB IN(rrnD-rrnE)1 lambda$^-$ rnc14::Tn10(DE3 lysogen:lacUV5 promoter-T7 polymerase) (IPTG-inducible T7 polymerase) (RNAse III minus) Host strain for RNAi feeding experiments

2.3.4.2 C. elegans strains

Many worm strains were provided by the *Caenorhabditis* Genetic Centre (CGC; http://biosci.umn.edu/CGC/CGChomepage.htm). Other strains were obtained from I. Johnstone (University of Glasgow, Glasgow, UK), J. Rothman (UC Santa Barbara, Santa Barbara, USA), A. Rose (University of British Columbia, Vancouver, Canada) or T. Page (University of Glasgow, Glasgow, UK).

Table 2.9: *C. elegans* strains provided by the CGC or other researchers (I. Johnstone, J. Rothman, A. Rose and T. Page).

Strain name	Genotype	Source
CB450	*unc-13(e450) I*	CGC
CB458	*dpy-13(e458) IV*	CGC
CB61	*dpy-5(e61) I*	CGC
CB88	*dpy-7(e88) X*	CGC
DR103	*dpy-10(e128) unc-4(e120) II*	CGC
DR466	*him-5(e1490) V*	CGC
DR518	*rol-6(su1006) unc-4(e120) II*	CGC
IA019	*ijIs8 [col-12::GFP::lacZ]*	I. Johnstone
IA081	*stc-1(ij15)/+ II*	I. Johnstone
IA088	*unc-76(e911) V*; array pES[*dpy-7::GFP::lacZ*]	I. Johnstone
IA105	*ijIs12 [dpy-7::GFP::lacZ] I*	I. Johnstone
IA337	*dpy-10(sc48) II*	I. Johnstone
JR1838	*wIs84 [pJM66(elt-2::lacZ::GFP) + pRF4(rol-6(su1006dm))*]	J. Rothman
JR1986	*wIs118 [pDP#MM016B(unc-119+) + pMW025(vc5.3::GFP::lacZ)]*	J. Rothman
JR667	*unc-119(e2498::Tc1) III; wIs51(seam cell::GFP)*	CGC
KR503	*dpy-5(e61) let-607(h189) unc-13(e450) I; sDp2(I;f)*	A. Rose
KR727	*dpy-5(e61) let-607(h402) unc-13(e450) I; sDp2(I;f)*	CGC
ML581	*lin-26(mc15) unc-4(e120) / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
MT1679	*unc-105(n490) II; lon-2(e678) let-2(n821) X*	CGC
MT5104	*lin-31(n301) clr-1(e1745) dpy-10(e128) II*	CGC
NL2099	*rrf-3(pk1426) II*	CGC
N2	*C. elegans* wild type (var Bristol)	CGC
SP152	*him-1(e879) I; mnC1 dpy-10(e128) unc-52(e444) / zyg-11(mn40) unc-4(e120) II*	CGC
SP158	*spe-1(mn47) unc-4(e120) / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP211	*let-252(mn100) unc-4(e120) / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP378	*let-23(mn23) unc-4(e120) / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP379	*let-24(mn24) unc-4(e120) / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP638	*unc-4(e120) mnDf69 / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP754	*unc-4(e120) mnDf88 / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP757	*unc-4(e120) mnDf91 / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP781	*unc-4(e120) mnDf97 / mnC1 dpy-10(e128) unc-52(e444) II*	CGC
SP784	*unc-4(e120) mnDf100 / mnC1 dpy-10(e128) unc-52(e444) II*	CGC

Strain name	Genotype	Source
SP806	*unc-4(e120) mnDf108 / mnCl dpy-10(e128) unc-52(e444) II*	CGC
SP807	*unc-4(e120) mnDf109 / mnCl dpy-10(e128) unc-52(e444) II*	CGC
SU93	*jcls1* [pJS191(*ajm-1::GFP*) + pRF4(*rol-6(su1006)*)]	CGC
TJ415	*age-1(hx546) fer-15(b26) unc-4(e120) II*	CGC
TP12	*kals12 [col-19::GFP]*	T. Page

Table 2.10: *C. elegans* strains constructed in this work.

Strain Name	Genotype / Description
IA432	*stc-1(ij15) unc-4(e120) / mnCl dpy-10(e128) unc-52(e444) II*
IA445	N2 carrying *let-607::GFP::lacZ* in free array
IA446	*let-607(h189) / + I*
IA447	*let-607(h402) / + I*
IA448	*let-607(h402) / + I;* array pES[*dpy-7::GFP::lacZ*]
IA449	*let-607(h402) / + I; ijIs8 [col-12::GFP::lacZ]*
IA451	*stc-1(ij15) unc-4(e120) / mnCl dpy-10(e128) unc-52(e444) II* *ij15* rescued with [F54C5, F28C6, M111, D2085,C25G9] cosmids
IA452	*stc-1(ij15) unc-4(e120) / mnCl dpy-10(e128) unc-52(e444) II; ij15* rescued with F54C9 cosmid
IA460	*let-607(h402) / + I; wIs84* [pJM66(*elt-2::lacZ::GFP*) + pRF4(*rol-6(dm)*)]
IA461	*let-607(h402) / +I; wIs118* [pDP#MM016B(*unc-119+*) + pMW025]
IA462	*let-607(h189)/+ I; wIs118* [pDP#MM016B(*unc-119+*) + pMW025]
IA463	*let-607(h189)/+I; unc-119(e2498::Tc1) III; wIs51*(seam cell::*GFP*)
IA465	*dpy-5(e61) let-607(h402) unc-13(e450) / + I; let-607(h402)* rescued with F57B10.1 linear fragment
IA466	*let-607(h189) / + I; wIs84* [pJM66(*elt-2::lacZ::GFP*) + pRF4(*rol-6(su1006dm)*)]
IA469	*stc-1(ij15) unc-4(e120) / mnCl dpy-10(e128) unc-52(e444) II; ij15* rescued with F54C9.2+3 plasmid
IA471	*stc-1(ij15) unc-4(e120) / mnCl dpy-10(e128) unc-52(e444) II; ij15* rescued with F54C9.2 linear fragment
IA473	*let-607(h402)/+ I; unc-119(e2498::Tc1) III; wIs51*(seam cell::*GFP*)

Strain Name	Genotype / Description
IA478	*let-607(h402)/+ I; jcIs1[pJS191(ajm-1::GFP) + pRF4(rol-6(su1006))] IV*
IA479	*let-607(h189)/+ I; jcIs1[pJS191(ajm-1::GFP) + pRF4(rol-6(su1006))] IV*
IA497	N2 carrying 2006 bp 5' *stc-1::GFP::lacZ* in free array
IA498	N2 carrying 542 bp 5' *stc-1::GFP::lacZ* in free array

2. METHODS

2.4 TECHNIQUES FOR CULTURING *C. ELEGANS*

2.4.1 Preparation of basic culture

C. elegans strains were grown and maintained on NGM agar plates with a lawn of OP50 bacteria (CGC). The NGM medium was aseptically poured into petri plates, which then were stored in an airtight container at room temperature. Different sized petri plates were used depending on the purpose of the experiment: small (3.5 cm of diameter), medium (5.5 cm of diameter) and large (9 cm of diameter). Under sterile conditions, large plates were seeded in the centre with a few drops of OP50 bacteria, medium and small plates with two and one drops of OP50 culture respectively. Seeded and unseeded plates were placed at 4°C for longer storage if not used within a short time.

2.5 NEMATODE HANDLING

Individual animals were manipulated with a worm pick, which consists of a piece of platinum wire with a flattened, arrow-shaped end, sealed into a broken Pasteur pipette. The platinum wire cools quickly when flamed, allowing sterilisation of the pick on a Bunsen burner between worm transfers. For handling large amounts of animals, agar chunks were cut out from old NGM cultures and transfer onto fresh, seeded plates with the aid of a sterilised scalpel (of 11 in or 24 in blade).

N2 Bristol and most other strains were maintained in a 20°C incubator. Temperature sensitive mutants were also placed at 15°C or 25°C according to their characteristics and the aims of the study. For longer preservation of stocks, plates were wrapped with Parafilm[R] film and kept at 15°C.

2.5.1 Crossing

Crosses between hermaphrodites and males were carried out using standard methods (Hodgkin, 1999). In general, a few L4 hermaphrodites together with an excess of L4 males (e.g., 2-3 hermaphrodites and 6 males) were placed on a small seeded plate for mating. In some cases, however, single-pair mating (one hermaphrodite and one male) was necessary.

2.5.2 Scoring of progeny

Single hermaphrodites were picked onto small seeded plates, allowed to egg lay and were then removed the following day to facilitate counting of their broods. If the entire brood had to be scored, the mothers were transferred to fresh plates at daily intervals until no more eggs were produced.

2.5.3 Decontamination of *C. elegans* stocks

Contamination of cultures by fungi, yeast, or undesirable bacteria happened occasionally. This problem could be reduced by transferring some worms onto a fresh plate, where they were allowed to move before repeating the operation at least one more time. Yeast and bacteria could also be removed by treatment with alkaline hypochlorite solution, as eggs are resistant to bleach. Worms from the affected plate were placed into a drop of hyochlorite mix at one side of a seeded plate. The following day, newly hatched animals would have crawled away and spread over the bacterial lawn. The latter procedure was also carried out in order to obtain large amounts of L1 larvae.

2.5.4 Freezing and thawing nematode stocks

The wild-type and mutant strains can be stored indefinitely in liquid nitrogen (-196°C). The key factors for successful freezing and recovery are the correct developmental stage of the worms, the presence of glycerol in the freezing solution and a slow initial freezing. Two large plates were left close to starvation with the purpose of having populations mainly constituted by L1 and L2 larvae, since these survive freezing best. Each plate was then washed with 3 ml of M9 buffer to make three or four 0.5 ml aliquots to which an equal volume of freezing solution was added. Tubes were mixed and immediately placed in a polystyrene box at -70°C. After at least a day, one tube was thawed as a tester and the remainder of the tubes were passed to two different liquid nitrogen tanks. Three vials were kept as working stocks in one tank, and the other three were left as back-up tubes in the second tank. For thawing, a vial of worms was melted at room

temperature, its contents mixed and plated onto a large NGM Petri dish with bacteria. It was noticed that some mutant strains, specifically dumpy (Dpy) mutants, did not survive freezing as well as wild-type animals.

2.6 *E. COLI* CULTURES

E. coli strains were grown either on 9 cm LB agar plates or in LB broth, containing appropriate antibiotics (Table 2.11), if required. Plates were placed in a 37°C incubator, and liquid cultures in a 37°C shaker incubator for bacterial growth.

Table 2.11: Antibiotics employed for selective growth of *E. coli* strains. The concentrations of the stock solutions, their storage temperature and the working concentrations are specified for each drug.

	Stock solution		Working concentration
	concentration	storage	
Ampicillin	100 mg/ml in dH$_2$O	-20°C	100 μg/ml; 75 μg/ml (for cosmids)
Kanamycin	50 mg/ml in dH$_2$O	-20°C	25 μg/ml; 50 μg/ml (for cosmids)
Streptomycin	12.5 mg/ml in dH$_2$O	-20°C	12.5 μg/ml
Tetracycline	12.5 mg/ml in 50% ethanol	4°C	12.5 μg/ml

2.7 STANDARD MOLECULAR BIOLOGY PROCEDURES

2.7.1 DNA isolation

2.7.1.1 *Plasmid and cosmid DNA preparation*

Different kits were employed to isolate plasmids and cosmids depending on the amount and quality of DNA required. Large scale preparations (up to 100 μg or 500 μg of DNA) were carried out using the Qiagen plasmid midi or maxi kits. Purification of smaller amounts of DNA (up to 20 μg) was performed with the Qiagen plasmid mini kit or the Qiaprep spin miniprep kit. The latter was the system routinely used because it was fast and simple. It did not yield highly pure DNA, however, so for applications where that was important, such as *C. elegans* transformation, plasmid DNA was prepared by one of the other methods. Prior to DNA

extraction, 0.8 ml of the overnight bacterial culture was removed into 1 ml of 2% peptone and 40% glycerol to keep as stock. Aliquots were mixed and stored at -70°C.

2.7.1.1.1 Notes on cosmids

Cosmid clones were provided in soft agar stabs by Alan Coulson (The Sanger Centre, Cambridge, UK). As recommended, immediate handling proceeded upon receipt of the cosmids since these tend to be unstable. A glycerol stock was made by taking an inoculum straight from the stab into 1 ml of 2% peptone and 40% glycerol containing the required selective drug. This was mixed and frozen at -70°C for storage. Isolation of colonies was achieved by streaking from each stab onto selective LB agar plates. Large and very small colonies were avoided; medium, smallish colonies were picked to grow overnight cultures. The culture volumes for cosmids were as suggested in the kit manuals for low-copy plasmids. The volumes of resuspension, lysis and neutralisation buffers were adjusted accordingly.

2.7.1.2 C. elegans genomic DNA extraction

This was a two-day method adapted from Johnstone (1999) to isolate large molecular weight chromosomal DNA.

Notes: *NGM agarose medium was used because DNA prepared from worms cultured on NGM agar plates tends to be poorly digested with restriction enzymes.

*Use of wide tips and mixing by inversion was performed to prevent shearing of the DNA.

Several 9 cm plates of worms were grown on NGM agarose until the food source was nearly consumed. Worms were collected with M9 buffer into 15 ml Falcon tubes and allowed to settle on ice by gravity. Washes were repeated twice. To solubilise the tough exoskeleton and other worm proteins, six volumes of 1 x WLB were added to one volume of worms. The suspension was divided into screw cap eppendorfs that were placed on a dry block at 65°C for 4 h. The tubes were agitated periodically to aid in the digestion and then spun at 13000 rpm for 5 min in a benchtop centrifuge. The supernatant was removed into a new eppendorf and an equal volume of phenol/chloroform (pH 8) was added, mixed carefully by inversion and centrifuged for 10 min at maximum speed. The aqueous layer was transferred to an eppendorf, keeping the rest for back-extraction. Any DNA that remained associated with the interphase was back-extracted by adding 300 μl of dH$_2$O, mixing by inversion and spinning for 5 min at 13000 rpm. The aqueous layer was pooled with the previous, and the phenol/chloroform extraction process was repeated twice. The aqueous phase was then extracted once with chloroform. Two volumes of ice cold

100% ethanol were added, mixed and the aliquot was kept overnight at -20°C. The following morning, the nucleic acid was collected by centrifuging for 15 min at 13000 rpm to be subsequently washed in 70% ethanol. The pellet was resuspended in 1 ml of TE buffer (pH 8) for 1 h at room temperature on a roller platform. DNase-free RNase A (Sigma) was added to a final concentration of 100 µg/ml and incubated at 37°C for 1 h to allow removal of the RNA. At this stage, the solution was extracted once with phenol/chloroform and once with chloroform, and then 1/25 volume of 5 M NaCl plus 0.7 volumes of 100% of isopropanol were added to precipitate the DNA. The pellet was resuspended in 120 µl TE buffer and stored at 4°C.

2.7.1.3 Purification of DNA fragments

DNA fragments generated from diverse enzymatic reactions were purified using either the PCR purification or the gel extraction Qiagen kits. The PCR purification kit was used to rid fragments from contaminants such as primers, nucleotides, salts and enzymes. In cases where there were several DNA fragments present, gel extraction was employed to be certain that just the desired fragment was processed. Both methods are designed to purify up to 10 µg of DNA fragments with sizes between 70 bp or 100 bp and 10 kb. Similarly, pure DNA was obtained in a higher final concentration with the minielute kits. Only up to 5 µg of DNA fragments of sizes between 70 bp and 4 Kb can be extracted with the minielute system.

It should be noted that the gel extraction protocol was sometimes modified in order to recover the maximum amount of DNA possible. Elution with 50 µl of dH_2O, pre-heated to 50°C, was carried out three times. To ethanol precipitate the eluate, 1 µl of glycogen, 1/25 volume of 5 M NaCl and 2.5 volumes of 100% ethanol were added, then mixed and left overnight at -20°C. The following morning, DNA was collected by centrifuging for 5 min at 13000 rpm. The pellet was air dried, then resuspended in 10 µl of dH_2O.

2.7.2 Quantitation of nucleic acid

The following methods were used to estimate the concentration of DNA in samples. For approximate estimations, a percentage of the sample was electrophoresed in an agarose gel with size standards at known concentrations. A comparison of fluorescent intensity after ethidium bromide staining permitted an estimation of the DNA concentration of the sample.

For accurate estimation of DNA concentrations, spectrophotometric measurements were used. The formula below was applied to calculate the DNA concentration in a given sample:

$$ng/\mu l \; DNA = dilution\; factor \; x \; 50 \; x \; Abs_{260nm}$$

The amount of single-stranded (ss) RNA was determined only spectrophotometrically, knowing that an OD_{260nm} of 1 corresponds approximately to 40 ng/μl of RNA. The ratio between the 260 nm and 280 nm readings (OD_{260nm}/OD_{280nm}) was taken into account as an indicator of the purity of the sample and therefore of the accuracy of the estimation. The OD_{260nm}/OD_{280nm} should be 1.8 and 2.0 in pure DNA and RNA samples respectively.

2.7.3 Visualisation of nucleic acid

DNA and RNA samples were electrophoresed on agarose gels to estimate size and/or concentration, by comparison with appropriate standards. Gels ranged from 0.7% to 1.5% (w/v) of agarose in 1 x TBE. 1-1.5% gels were employed for electrophoresis of small fragments and for greater resolution of similar sized fragments; 0.7-0.8% gels were used otherwise. Gel loading buffer was added to all the samples prior to electrophoresis, which was normally carried out at 150 V/cm. For visualisation of the nucleic acids, gels were immersed in dH_2O containing ethidium bromide for 30-45 min and then exposed to UV light in a transilluminator connected to imaging systems (Bio-Rad).

2.7.4 PCR and PCR-based techniques

Standard PCR reactions were prepared on ice as a 25 μl reaction containing dH_2O, 1 x reaction buffer, 2.5 mM $MgCl_2$ if required, 200 μM dNTP, 100 ng of each primer, 100 ng DNA and 2.5 units of DNA polymerase. Different enzymes were employed depending on the target DNA, the desired accuracy and DNA yield: *Taq* DNA polymerase (AB gene); Vent DNA polymerase (NEB); and *TaqPlus* Precision PCR System, *TaqPlus Long* PCR System, *Taq* Extender PCR Additive, Cloned *Pfu* DNA polymerase and *PfuTurbo* DNA polymerase (Stratagene). The denaturation, annealing and extension parameters were modified in reactions according to the length of the target and the primers melting temperature. PCR reactions were performed using the Stratagene's RoboCycler Gradient 96 temperature cycler. The "hot start" technique whereby the enzyme was added after completion of the first denaturation cycle was conducted in all reactions using *Taq* polymerase.

2.7.4.1 Semi-quantitative reverse transcriptase PCR (rtPCR)

Semi-quantitative rtPCR was carried out to study gene expression throughout the worm's life cycle. Basically, this is a two-step procedure. Firstly, cDNA was generated by reverse transcription of mRNA samples prepared from synchronous cultures every 2 h from hatch to young adult (provided by I. Johnstone; Johnstone and Barry, 1996). Secondly, the resulting cDNA was used as a template for PCR of the test gene, either *let-607* or *stc-1*, and of a control, *ama-1*, which encodes the large subunit of RNA polymerase II (Bird and Riddle, 1989). rtPCR reactions contained 1 x *Taq* DNA polymerase buffer, 2.5 mM $MgCl_2$, 200 μM dNTP, 75 ng of each of the four primers, 100 ng/μl BSA, 1 μl cDNA and 2.5 units *Taq* polymerase, and were amplified under the following conditions: [94°C, 3 min] x1; [94°C, 30 sec; 57°C or 58°C, 30 sec; 72°C, 30 sec] x30-32; [94°C, 30 sec; 57°C or 58°C, 3 min; 72°C, 3 min] x1 (Johnstone and Barry, 1996). The fragments generated from *ama-1* and the test gene were relatively short and of different sizes. Oligonucleotides were designed such that they spanned an intron, so as to distinguish cDNA generated PCR products from those derived from contaminating gDNA.

2.7.4.2 Single-worm and single-embryo PCR

This procedure was modified from Williams (1995a). Single worms were transferred into a PCR tube after being washed in M9 buffer to minimise the presence of bacteria. Each transfer was checked under a stereomicroscope, 12.5 μl of SWLB were added and tubes were centrifuged briefly and frozen at -70°C for 10 min. Lysis of samples was for 1 h at 60°C followed by 15 min at 95°C for inactivation of proteinase K and proteases. Embryos were prepared in the same way with the exception that they were treated with chitinase before adding the SWLB. Single eggs were picked up using a pulled capillary filled with chitinase solution. Only a small volume of chitinase was transferred with each egg into the tube. 5 μl of the lysate was used for PCR analysis in which denaturation was conducted for 3 min at 94°C followed by 38 amplification cycles [94°C, 30 sec; 56-58°C, 1 min; 72°C, 1 min].

2.7.4.3 Spliced leader (SL) sequence PCR

PCR was performed on cDNA using a combination of a gene specific primer with an SL primer (either SL1 or SL2) to amplify the 5' end of the gene. This method indicates if the message of a gene is *trans*-spliced to an SL1 or SL2 sequence.

2.7.4.4 *Rapid amplification of cDNA ends (RACE) PCR*

RACE Gibco BRL systems were employed to amplify unknown 3' and 5' terminals of mRNAs (previously isolated by I. Johnstone). The manufacturer's protocols were followed, and an additional PCR with a "nested" gene specific primer to generate enough specific products was then performed. Following amplification, the 3' and 5' RACE products were visualised by gel electrophoresis and cloned into an appropriate vector for subsequent sequencing.

2.7.5 DNA digestions

Typically, DNA was cleaved with the appropriate restriction endonuclease in 20 μl reactions containing 1 x enzyme reaction buffer and, if required, 1 x 100 μg/ml BSA. Double digests were conducted simultaneously whenever possible. When the two restriction enzymes required different conditions for their activity, reactions were done sequentially with conventional phenol/chloroform extraction and ethanol precipitation steps between the two digestions. The incubation time was relative to the amount of DNA and enzyme added, thus it varied from 1.5 h to overnight. DNA digestions were performed at the temperature recommended by the manufacturer.

2.7.6 Cloning reactions

DNA fragments generated by PCR were inserted into various commercial cloning vectors, according to the manufacturers' recommendations. DNA inserts derived from digests were cloned into plasmid vectors with compatible ends that had been treated with calf intestinal alkaline phosphatase (CIAP; NEB) for 1 h at 37°C. Ligations were carried out in 20 μl reactions with 1 x T4 DNA ligase buffer and 0.4 units of T4 DNA ligase (AB gene) overnight at 16°C. The resulting recombinant plasmids were transformed into *E. coli* competent cells as described below.

2.7.7 Transformation of *E. coli*

Standard transformations were normally performed using Epicurian Coli XL10-Gold Ultracompetent Cells (Stratagene) and 2 μl of ligated DNA or approximately 5 ng of plasmid DNA. The manufacturer's protocol was followed with the exceptions that just 40 μl of cells and 1.6 μl of β-ME were utilised, and that SOC medium was employed instead of NZY broth. Positive clones were selected according to the antibiotic resistance acquired from the transformed plasmid and, when applicable, by β-galactosidase colour screening.

Transformations were carried out with other *E. coli* strains provided with specific cloning vectors. The above procedure was also valid with Epicurian Coli XL10- Gold Kan Ultracompetent Cells (Stratagene); the manufacturer's instructions were followed with TOP10F' One Shot Competent Cells (Invitrogen). For protein expression analyses, the strain M15 (Qiagen) was transformed modifying the recommended procedure as follows: use of 50 μl cells and 10 ng DNA construct, 30 sec heat shock and 0.9 ml SOC.

2.7.7.1 Transformation by electroporation

Electrotransformation was performed using a Bio-Rad electroporator to increase efficiency of transformations. The manufacturer's protocol was applied making use of 0.1 cm cuvettes and the following settings: gene pulser at 25 μF (capacitance), pulse controller 200 Ω (resistance), voltage at 1.8 kV.

2.7.8 Screening for positive transformants

Bacterial colonies grown on selective plates after transformation with the ligation reactions were screened either by PCR or colony lysis and examination on agarose gels. The PCR method relied on the visualisation of an amplified band on agarose gels when the colony contained the recombinant plasmid. Alternatively, a small amount of restreaked colony was dipped into 35 μl of colony lysis buffer and left for 15-20 min at room temperature. Samples were centrifuged at 13,000 rpm for 25 min in a bench microfuge and placed at 4°C for at least 35 min to facilitate loading. Supernatants were loaded onto a 1% agarose gel avoiding the loose pellet. Plasmids with inserts were identified by their slower electrophoretic mobility. All colonies were patched onto a selective plate prior to screening for future use.

2.7.9 Sequencing of mutant alleles

Series of single-worm and single-embyo PCRs were performed on *let-607(h402)* or *let-607(h189)* animals, and on *stc-1(ij15)* mutants to generate *let-607* and *stc-1* DNA respectively. Both genes were amplified in sections using either *Taq* polymerase or a combination of *Taq* and *Pfu* polymerises (*TaqPlus* Precision PCR System). PCR products were cloned into standard vectors, followed by insertion into *E. coli* cells and isolation of transformants. Clones were isolated, digested with specific restriction enzymes and visualised on agarose gels to confirm that the DNA inserts were intact. Sequencing of the gene sections was performed by the MBSU DNA Sequencing Service (Glasgow University) with the ABI stretch automated sequencer

(Applied Biosystems) or by Oswel (University of Southampton, Southampton, U.K) using the ABI 37796 lane fluorescent DNA sequencing technology (Applied Biosystems). Both strands of *let-607* inserts were sequenced; *stc-1* cloned sections were analysed in one direction. The resulting sequences were then compared to those predicted for the corresponding gene, *let-607* or *stc-1*, to identify the sequence alterations in the mutants. 5' and 3' RACE products were cloned and sequenced using the same procedure and services as above.

2.8 CLASSICAL GENETIC MAPPING

2.8.1 Two-factor crosses

A few *dpy-10(sc48) ij15/+* + hermaphrodites were self-crossed on individual plates to count their complete broods. Taking into account *ij15* is a recessive lethal, and therefore just one quarter of the recombinant progeny was detected, the following formula was applied to assess the distance between *dpy-10* and *ij15*:

Dpy recombinants / (Dpy + WT) = $(2p-p^2)/3$, where 'p' is the recombination distance

Some F_1 animals were tested for segregation of death to confirm that all classes of progeny were present.

2.8.2 Three- and multi-factor crosses

The mutation *ij15* was placed in *trans* to a series of sets of two or three visible markers (Table 2.12). The resulting heterozygous hermaphrodites were self-crossed on individual plates to examine their progeny for recombinants homozygous for one marker mutation. These recombinants were then isolated and tested for presence or absence of *ij15*. The proportion of recombinants carrying the lethal allele with respect to the total number of recombinants was used to order *ij15* in relation to a given set of markers. It should be noted that not all possible recombinants were obtained when a marker mutation causing death, sterility or any other unscorable phenotype was employed. In such cases, the data to map our allele derived only from one direction and therefore the estimation was less accurate. In order to counteract this inconvenience, a considerable number of three-factor crosses were carried out.

Table 2.12: Strains and genetic markers used in three- and multi-factor crosses. The localisation of the genes is given in map units (m.u).

CGC Strain	markers used for self-crosses	Position on chromosome II (m.u)
MT5104	*lin-31*	-4.85843
	clr-1	-1.36413
	dpy-10	-0.01626
DR103	*dpy-10*	-0.01626
	unc-4	1.73046
SP378	*let-23*	1.04694
	unc-4	1.73046
DR518	*rol-6*	0.816947
	unc-4	1.73046
TJ415	*fer-15*	0.783049
	unc-4	1.73046
SP158	*spe-1*	0.36787
	unc-4	1.73046
ML581	*lin-26*	0.485687
	unc-4	1.73046
SP152	*zyg-11*	0.628061
	unc-4	1.73046
SP379	*let-24*	0.698218
	unc-4	1.73046
MT1679	*unc-105*	0.679428

2.8.3 Mapping by deficiencies

ij15 was assigned to a specific segment of the map by means of complementation tests against a number of deficiencies (Table 2.13). Males heterozygous for a given deficiency were mated with *ij15/+* hermaphrodites. Then, 10 animals of the cross-progeny were picked and allowed to self in order to examine their broods. Complementation was indicated by the ability to isolate heterozygotes, identified by their segregation of 50% of dead progeny. In contrast to 25% if *ij15* was included in the deficiency.

Table 2.13: Strains and deficiencies employed to map *ij15*. The estimated deficiency extents are given in m.u.

CGC Strain	Name deficiency	Endpoints left	right (m.u)
SP754	*mnDf88*	0.245189	– 0.795166
SP781	*mnDf97*	0.306462	– 0.66859
SP806	*mnDf108*	0.582647	– 0.789388
SP807	*mnDf109*	0.613014	– 0.789388
SP784	*mnDf100*	0.526747	– 0.66859
SP757	*mnDf91*	0.659508	– 0.789388
SP638	*mnDf69*	0.800092	– 1.14889

2.8.4 Physical mapping of deficiencies using PCR

The end points of the deficiencies mnDf69 and mnDf88 were physically mapped in order to narrow down the region where *ij15* had been located. Homozygous embryos for one of the above deficiencies were employed for single-embryo PCRs with a series of primer pairs designed to amplify a small specific region of genome sequence. The absence or presence of amplified product indicated whether the test sequence was deleted by the deficiency or not. Tests of each primer pair were performed on four deficiency embryos and two wild-type embryos in parallel. In addition, oligonucleotides known to amplify a region within or outside the deleted segment were used as negative and positive controls respectively.

2.9 *C. ELEGANS* TRANSFORMATION

C. elegans transformation was performed to identify the genes responsible for *let-607(h402)* and *ij15* mutant phenotypes and to analyse *let-607* and *stc-1* expression patterns using reporter constructs.

2.9.1 Preparation of DNA samples

Plasmid and cosmid DNA was prepared using Qiagen kits and resuspended in TE buffer (pH 8). gDNA was purified by standard phenol/chloroform extraction. It should be noted that due to the

instability of cosmids, these were digested with restriction endonucleases and then run on agarose gels, to ensure that deletions or rearrangements (missing bands or bands of incorrect size) had not occurred and therefore the correct cosmid was being used.

Rescue experiments were carried out either with circular DNA or linear DNA. Injection mixes of circular DNA had an overall DNA concentration between 120 and 150 ng/μl of tester DNA plus the co-injection DNA in dH$_2$O. The co-injection DNA consisted of a marker plasmid, usually pMW0002 and pMW0003, and pTAg. pMW0002 and pMW0003 carry the *dpy-7* promoter::*GFP* and *col-12* promoter::*GFP* reporter gene fusions respectively; pTAg is a plasmid with ampicillin and kanamycin resistance genes that provides areas of homology to cosmids and plasmids with either of these cassettes. The above markers were normally used at a final concentration of 2.5 ng/μl each. The amount of pTAg varied in the mixtures to raise the DNA concentration to the desired total. Injection mixes of linear DNA contained 50-100 ng/μl of N2 gDNA plus the marker (either pMW0002 and pMW0003 or pMW025) and the test DNA at 0.5-2 ng/μl each. The DNA was linearised with restriction enzymes yielding blunt ends. Microinjections of gene reporter constructs were performed using 10 ng/μl of the recombinant plasmids and about 120 ng/μl of pTAg. Details of all the transformations conducted are given below.

2.9.1.1 Transformations carried out in the study of F57B10.1 – let-607

- Phenotype rescue -
Mix: 7543 bp F57B10.1 PCR product (0.5-2 ng/μl), *Fsp* I cut pMW0002 and pMW0003 (0.5-2 ng/μl) and *Pvu* II cut gDNA (~100 ng/μl).
Injected strain: *dpy-5(e61) let-607(h402) unc-13(e450)/+ + + I.*

- Expression analysis –
Mix: 10 ng/μl reporter construct and 120 ng/μl of pTAG.
Injected strain: N2 Bristol.

2.9.1.2 Transformations performed in the analysis of ij15 – stc-1

- Phenotype rescue -
In all cases the injected strain was IA432 [*stc-1(ij15) unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444) II*].

Injections with groups of cosmids

group A = C26D10, C10A8, C54D8 and F14E5; group B = F22B5, M05D6, C07H5 and C04A6; and group C = F54C9, F28C6, M111, D2085 and C34C6.

Mix: 20 ng/μl of each cosmid, pMW0002 and pMW0003 at 10 ng/μl each, together with 30 ng/μl of pTAg. Second attempt as before but using 10 ng/μl of each cosmid, pMW0002 and pMW0003 at 10 ng/μl each, together with 70 ng/μl of pTAg.

Injections with single cosmids in group C

Mix: 5 ng/μl cosmid DNA, pMW0002 and pMW0003 at 2.5 ng/μl each, and 130 ng/μl of pTAg.

Injections with F54C9 sections and specific coding sequences

A 1390 bp *Hind* III-*Xba* I fragment from cosmid DNA including F54C9.10 was cloned into pBluescript, and a 1600 bp F54C9.5 PCR fragment was cloned into TOPO pCR2.1 vector. Both constructs were injected at 15 ng/μl, with 2.5 ng/μl of pMW0002 and pMW0003, and 130 ng/μl of pTAg.

Digestion of F54C9 with *Spe* I and *Xho* I; *Pst* I and *Sal* I; and *Sma* I generated a range of fragments. The following were then cloned into pBluescript: *Spe* I-*Xho* I 9204 bp, *Spe* I-*Xho* I 8220 bp, *Spe* I 6756 bp, *Xho* I 1766 bp, *Pst* I 5613 bp, and *Pst* I 3698 bp. 15 ng/μl of each construct was tested using 2.5 ng/μl of pMW0002 and pMW0003, and 125 ng/μl of pTAg. The *Spe* I-*Xho* I 9204 bp fragment was digested further resulting in *Not* I-*Xho* I 5754 bp, *Not* I-*Spe* I 3450 bp and *Sac* I-*Spe* I 6611 bp fragments that were cloned into pBluescript. 7 ng/μl of each construct was used together with 2.5 ng/μl of pMW0002 and pMW0003, and with 125 ng/μl of pTAg.

2978 bp, 2004 bp, and 4449 bp PCR products containing the F54C9.2, F54C9.3, and F54C9.2-F54C9.3 genes respectively were cloned into TOPO pCR2.1 vector. Constructs were injected at 7 ng/μl with pMW0002 and pMW0003 at 2.5 ng/μl each, plus 125 ng/μl of pTAg.

The recombinant plasmid harbouring the F54C9.2 gene was linearised using *Msc* I, and injected at ~2 ng/μl with ~2 ng/μl of *Fsp* I digested pMW025 and ~100 ng/μl of *Pvu* II digested gDNA.

-Expression analysis-

Mix: 10 ng/μl of reporter construct and 120 ng /μl of pTAg.

Injected strain: N2 Bristol.

2.9.2 Microinjection procedure

The microinjection procedure and materials were essentially as described in Mello *et al.*, (1991) and Mello and Fire (1995). Needles were made using glass capillaries with inner filament and a microelectrode needle puller. Injection solutions were centrifuged for 15 min at maximum speed in a microfuge to sediment particulate matter before being loaded into the needles with the aid of a mouth pipette. Under the stereomicroscope, clean, healthy and well fed young adult hermaphrodites were immobilised on dried 2% agarose pads with a drop of paraffin oil. The mounted worms were then immediately injected employing a Zeiss Axiovert-100 inverted microscope attached to a Narishige MO-202 Joy-stick micromanipulator, which was connected to a nitrogen cylinder. Pressure was controlled by means of a foot pedal attachment. DNA samples were introduced into the gonad of the hermaphrodites. A worm pick dipped in M9 buffer was used to transfer the injected worms onto fresh plates, containing also a drop of the buffer to rid the animals of the paraffin oil. Survivors were subsequently passed to new plates.

Transformants from rescue experiments were identified by detection of the expression of the green fluorescent protein (GFP) from the reporter gene markers under UV light. Co-injection markers were not normally used for expression analyses since production of GFP from the test construct was already indicative of transformation.

2.10 REPORTER GENE CONSTRUCTION AND ANALYSIS OF GENE EXPRESSION

A 2119 bp *Hind* III-*Bam*H I fragment containing the upstream sequence and the first 115 bp of the gene *let-607* was fused, in-frame, with the reporter genes *gfp* and *lacZ* in the pPD96.04 vector (provided by A. Fire). In the case of *stc-1*, two different sized *Bam*H I-*Sph* I fragments, 2006 bp and 542 bp, corresponding to the 5' region and the ATG codon were inserted into the pPD96.04 vector. These gene fragments were obtained by PCR amplification from gDNA, followed by insertion into standard cloning vectors, digestion and gel extraction. Subsequently, they were fused with the reporter genes by standard recombinant DNA techniques. The expression constructs were microinjected to obtain transgenic lines as described in section 2.9.1. The gene expression patterns were interpreted from individuals of at least two independent transformed lines.

GFP expression was analysed in living animals and anaesthetised worms under a Nomarski microscope using UV light and a fluorescein isothiocyanate (FITC) filter set. To study the

distribution of the *lacZ* product (the enzyme β-galactosidase) mixed staged worms were washed from plates using either M9 buffer or 0.1% X-Triton in dH$_2$O. The worms were harvested and left in 100 µl of the solution, discarding the rest of the supernatant. 100 µl of 2.5% glutaraldehyde was added and left at room temperature for 15 min, mixing occasionally. The glutaraldehyde was removed by three subsequent washes in 1 ml of M9 buffer or 0.1% X-Triton. 100 µl aliquots containing the worms from the base of the tube were pipetted onto a clean slide and dried in a speed vacuum for 5-10 min. Slides were placed in –20°C acetone for 5 min. The excess acetone was drained off and the slides were allowed to air dry at room temperature. A few drops of the staining solution for β-galactosidase activity were added to each slide and spread evenly to cover the worms by placing a coverslip on top. The edges of the coverslip were sealed with clear nail varnish and the slides were left in a darkened humid chamber either at 37°C or room temperature until staining was apparent (from 1 h to overnight). Stained worms were examined by Nomarski microscopy. A staining solution containing a higher percentage of X-gal was also used for a more sensitive detection of the enzyme activity.

2.11 RNAI

2.11.1 RNAi by injection

RNAi by injection (Fire *et al.*, 1998) was carried out for the following predicted genes: F54C9.4, F54C9.6, F54C9.8, F54C9.9 and F54C9.10. Exon sequences containing 508 bp of F54C9.4, 574 bp of F54C9.6, 534 bp of F54C9.8, 553 bp of F54C9.9 and 445 bp of F54C9.10 were amplified by PCR using gDNA and gene specific primers containing T3 and T7 promoter sequences. The PCR products were purified by gel extraction and used as templates to synthesise ssRNA with Promega's RiboMAX large scale RNA production system-T7 and -T3. Equal volumes of sense RNA and antisense RNA were mixed with 10 x injection buffer and incubated for 10 min at 68°C, and 30 min at 37°C to make dsRNA. Solutions containing the dsRNA were injected into young adult N2 hermaphrodites following the procedure described for DNA transformation (section 2.9.2) with the following exceptions. The preparations were introduced anywhere in the body cavity or gonad of the animals. The injected worms were transferred onto fresh, seeded plates at 24-hour intervals to facilitate detection of the affected progeny. RNA samples not used for microinjection were stored immediately at -70°C.

2.11.2 RNAi by feeding

A *let-607* fragment of 1020 bp (424 bp of exon sequence) and an *stc-1* fragment of 1122 bp (891 bp of exon sequence) were employed for RNAi by feeding (Timmons and Fire, 1998; Timmons *et al.*, 2001). The gene fragments were generated by PCR from gDNA, inserted into standard cloning plasmids and then subcloned into the L4440 feeding vector (pPD129.36) provided by A. Fire. Ligations were first transformed into standard *E. coli* competent cells to facilitate screening of transformants. The correct clones were then isolated and transformed into the *E. coli* strain H115(DE3), which had been made competent following the procedure given by Lisa Timmons (http://www.ciwemb.edu). Transformants were selected on ampicillin; both ampicillin and tetracycline were used for maintaining stocks. Bacterial induction and preparation of plates were essentially as in Kamath *et al.*, (2000). Briefly, single colonies of HT115 bacteria carrying the desired constructs were picked and grown in 10 ml of LB with 100 μg/ml ampicillin for 8-16 h. 1 mM IPTG plus 100 μg/ml ampicillin NGM plates were inoculated with the bacterial cultures and left overnight at room temperature for induction. The following day, L3-L4 hermaphrodites were placed onto the RNAi plates and left at 15°C until adulthood. The animals were then transferred individually to fresh RNAi plates at daily intervals to facilitate scoring of their progeny. The RNAi effects were studied at 15°C, 20°C and 25°C; HT115(DE3) transformants containing just the L4440 vector were used as negative controls. Embryos were allowed to hatch on RNAi plates to disrupt the genes function during post-embryogenesis. Larvae and embryos were washed in M9 buffer before placing them onto the RNAi plates in order to minimise contamination by OP50 bacteria.

A second construct (provided by J. Ahringer; University of Cambridge, Cambridge, U.K) was used to perform RNAi of *let-607*. It contained a 1194 bp fragment (1017 bp of exon sequence) corresponding to a different region of the gene from that used in the first construct.

2.12 MICROSCOPY

2.12.1 Routine microscopy

Worms on NGM plates were visualised with Zeiss stereomicroscopes. The handling of transgenic animals expressing GFP was carried out with a Stemi SV-6 Zeiss microscope with fluorescence attachment. For closer examination of worms, specimens were collected and placed on 2% agarose pads with a drop of M9 buffer (care was taken not to transfer much bacteria). Larvae and adults were anaesthetised on pads containing sodium azide (6.5 μl of 10% sodium azide in 1 ml of 2% agarose). A coverslip was placed gently on top of the specimens

and was sealed with white soft paraffin (British Pharmaceuticals). Samples were viewed with a Zeiss Axioplan microscope equipped with DIC optics (Nomarski) and attachments for fluorescence microscopy: a tungsten halogen lamp and Zeiss filter sets for FITC/GFP and rhodamine. Images were taken with a Hamamatsu Orca camera plus Improvision Openlab 2.0.2 software (http://www.improvision.com). The computer program Adobe Photoshop 7.0 was used to process most of the images presented in this work.

2.12.2 Scanning electron microscopy

Scanning electron microscopy (SEM) was performed to examine the external appearance of *let-607(h402)* hatchlings and *let-607* RNAi N2 or DR466 adult progeny. Untreated N2 adult hermaphrodites and DR466 adult males were used as controls. Worms were washed off plates in M9 buffer, let settle down on ice and rinsed twice in M9 buffer. According to the size of the worm pellet, an appropriate volume of 2.5% glutaraldehyde in PBS was added and left on ice for 1 h and 30 min, mixing the tubes occasionally. Glutaraldehyde was removed by three rinses in 2% sucrose in PBS, leaving the same volume of buffer as that of worms in the last wash. An equal volume of 2% osmium tetroxide aqueous solution was added, mixed and left for 1 h at room temperature. Samples were then rinsed three times in dH_2O (10 min each wash), mounted between 1-micron membrane filters and placed in metal collars. Dehydration was conducted at room temperature for 10 min each in 30%, 50%, 70%, 90%, absolute (twice), and dried absolute acetone. After drying by "critical point dryer", worms were transferred onto an adhesive tape stuck on a metal stub. Silver paint was applied on the metal stub at each side of the tape and, once dry, the samples were coated with gold for examination under a Philips 500 SEM.

2.13 IMMUNOLOCALISATION STUDIES

Immunodetection of LET-607, STC-1, DPY-7 and CDC-25 was performed using a modified method from Miller and Shakes (1995). Specimens were transferred into a drop of M9 buffer on polylysine coated slides and allowed to adhere. A coverslip was placed gently on top of the animals and the slides were put on a metal block at -70°C for a few minutes. With the aid of a scalpel blade, the coverslips were flicked off and the slides were immediately immersed in -20°C methanol for 10 min, and then in -20°C acetone for 10 min. Samples were allowed to air dry, rehydrated in PBST for a few minutes and blocked for 20 min in either 1% dried milk in PBST or 10% goat serum. After removing the blocking agent with three washes in PBST, the samples were incubated with the primary antibody for 2 h at room temperature, washed twice in PBST and then incubated with the secondary antibody (Molecular Probes or Sigma) for 2 h at 37°C in

the dark. Slides were washed twice in PBST before the coverslips were mounted with mounting solution and sealed with clear nail varnish. Details of the antibodies used are given below; antibodies were diluted in 1% dried milk in PBST.

LET-607: polyclonal anti-peptide, tested at 1/5, 1/50, 1/250 and 1/1000 dilutions; antisera from second and final bleed used at 1/200 dilution.
Secondary antibodies: alexa fluor 594 anti-rabbit at 1/200 dilution and alexa fluor 488 anti-rabbit at 1/200 dilution (Molecular Probes).

DPY-7 (I. Johnstone): monoclonal, 1/50 dilution.
Secondary antibody: goat-anti-mouse FITC, 1/100 dilution (Sigma).

CDC-25 (I. Johnstone): polyclonal anti-peptide, 1/50 dilution.
Secondary antibody: alexa fluor 594 anti-rabbit at 1/200 dilution.

2.14 PROTEIN TECHNIQUES

2.14.1 Recombinant fusion proteins

A recombinant system from Qiagen (QIAexpressionist) was employed to generate LET-607 and STC-1 proteins. The complete coding sequence of *let-607* (2087 bp) and two distinct segments of *stc-1* (509 bp and 452 bp) were amplified from N2 cDNA using a combination of *Taq* and *Pfu* polymerases. The PCR products were first inserted into standard plasmids to facilitate subsequent cloning into pEQ-30 expression vectors. The pEQ-30 plasmid contains six histidine residues (6 x His) 5' to the polylinker, which acts as an N-terminal tag of the gene fragment inserted in-frame. The pEQ-30 recombinant plasmids were introduced into *E. coli* XL10-Gold ultracompetent cells and the resulting transformants were screened by colony PCR and restriction analyses. M15 cells were transformed with the isolated clones as explained in section 2.7.7. Small cultures of transformants were handled following essentially the manufacturer's recommendations as a test of the expression and purification of the recombinant proteins. Expressed proteins produced in insoluble aggregates were treated with 8 M urea for 1 h at 37°C and then purified on nickel-nitrilotriacetic acid (Ni-NTA) spin columns (Qiagen) using buffers of different pH. The same procedure was then scaled up to obtain a large quantity of the STC-1 protein. In this case, the purification step was carried out by the MRC Co-operative Group Protein Purification Lab (Glasgow University) using a BioCAD 700E Workstation (Applied

Biosystems). The Coomassie Protein Assay Reagent Kit (Pierce) was employed to assess the concentration of the purified protein.

2.14.1.1 SDS-polyacrylamide gel electrophoresis of proteins (SDS-PAGE)

SDS-PAGE gels were prepared for the analysis of LET-607 and STC-1 recombinant proteins. Resolving gels contained 12% acrylamide and stacking gels 5% acrylamide in all cases. The gels were made basically as in Sambrook (1989) using Bio-Rad apparatus. Protein samples and markers were denatured in 1 x SDS loading buffer for 5 min at 95°C, loaded on the gels and resolved for approximately 45 min at 200 V. Gels were simultaneously fixed and stained with methanol/glacial acetic acid and Coomassie blue solution overnight at room temperature on a rocking platform. Excess of the stain was removed by soaking the gels in destain solution for several hours (three changes). Finally, gels were hydrated in dH_2O and dried for 1 h at 65°C in a Bio-Rad drier.

2.14.2 Western blotting

The NuPAGE Western blot system (Invitrogen) was employed to check the reactivity of the affinity purified anti-LET-607 antibody to the LET-607 protein. N2 *C. elegans* cellular proteins were prepared as in Johnstone (1999) using a protease inhibitor cocktail (complete mini, Roche) and a glass homogeniser. Protein samples in 1 x SDS loading buffer with 0.5 M β-ME, and the protein marker were heated for five min at 95°C. Extracts and marker were electrophoresed for 45 min on a NuPAGE 4-12% Bis-Tris gel with NuPAGE 1 x MOPS SDS running buffer and NuPAGE antioxidant. The separated proteins were transferred to a nitrocellulose membrane as indicated in the manufacturer's manual. The membrane was then blocked with 1% dried milk (Amersham Life Science) in PBST overnight on a rocking platform. anti-LET-607 antibody solutions were prepared in 1% dried milk and PBST and left overnight on a rolling platform at 4°C. The proteins on the membranes were probed with the primary antibody solutions for 2 h at room temperature on a rolling platform, followed by four washes with PBST of 5 min each. Samples were exposed to the secondary antibody, an anti-rabbit Ig G coupled to alkaline phosphatase (Sigma), at 1/15000 dilution with gentle agitation for 2 h at room temperature. Following four washes of 5 min with PBST, membranes were treated with the alkaline phosphatase substrate BCIP/NBT, washed with dH_2O and let dry.

2.15 BIOINFORMATIC ANALYSIS

Genomic sequence and predicted proteins were generally extracted from the following sources: WormBase (http://www.wormbase.org), GenBank (http://www.ncbi.nlm.nih.gov), EMBL (http://www.ebi.ac.uk), and Sanger Centre (http://www.sanger.ac.uk/Projects). Analysis of sequence similarity was performed on http://www.wormbase.org/db/searches/blast, http://www.ncbi.nlm.nih.gov/BLAST or http://www.sanger.ac.uk/Projects/C_elegans/blast_server.shtml (BLAST searches), and on EMBL (http://www.ebi.ac.uk/), BLOCKS (http://www.blocks.fhcrc.org/) or Pfam (http://www.sanger.ac.uk/Software/Pfam/) (protein similarity and protein family searches). WormBase was used to obtain much of the *C. elegans* gene and protein information and to access other linked databases; ACeDB was useful for comparison of *C. elegans* genetic and physical maps.

The Vector NTi 6 software (http://www.informaxinc.com) was extensively used for various purposes including design of PCR and sequencing primers, analysis with restriction enzymes, electrophoresis of digest products, examination of sequences (ContigExpress program) and alignments of multiple molecules (AlignX program).

Chapter 3

stc-1(ij15), *let-607(h402)* and *let-607(h189) C. elegans*
mutants are defective in cuticle synthesis

3.1 *C. ELEGANS* DEVELOPMENT

C. elegans develops from egg to adult through four larval stages, L1-L4 (Fig. 3.1). The development occurs following a rigidly determined program; an invariant temporal and spatial pattern of cell divisions result in a fixed number of cells that form specific parts of the animal. A fully developed embryo consists of five hundred and fifty-eight cells in the hermaphrodite and five hundred and sixty cells in the male; post-embryonic cell divisions increase the number of somatic cell nuclei in the hermaphrodite to nine hundred and fifty-nine and in the male to one thousand and thirty-one.

The stereotyped development of *C. elegans*, the body transparency at all stages, together with the relatively small number of cells and simple anatomy, allowed the determination of the entire cell lineage by observation of living animals under Nomarski optics (Sulston and Horvitz, 1977; Sulston *et al.*, 1983).

3.1.1 Embryogenesis

The *C. elegans* embryo develops from zygote to a vermiform shape in about 14 h at 22 °C. The process can be divided in two phases of roughly the same extent (Sulston *et al.*, 1983). The initial phase comprises the formation of the eggshell by the zygote, and most cell divisions and cell movements. The eggshell is a tough, impervious structure that surrounds the embryo conferring it protection outside the mother. It consists of an inner vitelline membrane, a middle chitinous layer and an outer layer composed of lipids and cross-linked collagenous proteins. Gastrulation takes place in the early embryo with the ingression of specific cells from the ventral surface in a precise and orderly manner. During the second phase, tissues differentiate and morphogenesis shapes the ellipsoidal embryo to a long, thin larva, from comma to three-fold stage (Fig. 3.2). Sulston and co-workers (1983) demonstrated, by cell ablation, that hypodermal cells are essential for elongation of the embryo. Subsequently, Priess and Hirsh (1986) showed that these cells are in fact responsible for the elongation process and described the mechanism by which this occurs. Just prior to and during elongation, circumferentially oriented bundles of microfilaments of actin organise at the apical surfaces of hypodermal cells. Their contraction constricts the apical hypodermal surfaces generating pressure on internal cells and forcing the embryo to decrease in circumference approximately threefold, with a concomitant increase in length of about fourfold. Circumferentially oriented microtubules together with the embryonic sheath (an extracellular structure that covers the elongating embryo) are also required for elongation, since they are involved in the even transmission of the
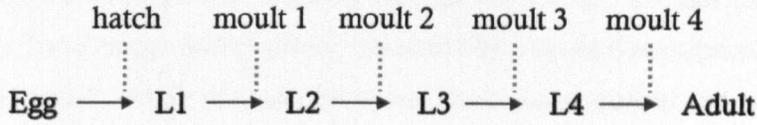
Figure 3.1: *C. elegans* development.

Embryogenesis ends with eclosion of the fully elongated animal; post-embryonic development comprises four larval stages (**L1-L4**) and one adult stage. The nematode moults four times during development, once at the end of each larval stage.
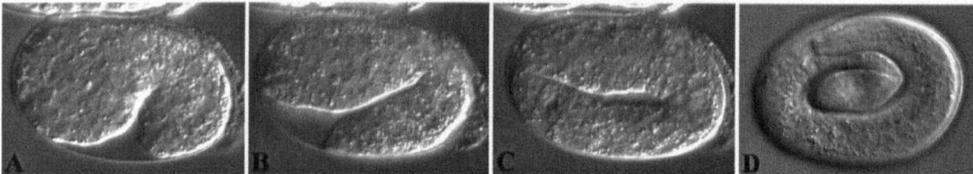
Figure 3.2: Major changes in body appearance of *C. elegans* embryo during elongation. A-D, DIC images of the same embryo at different elongation stages. A, comma (beginning of elongation); B-C, 1.5-fold and 2-fold respectively (middle stages of elongation); D, 3-fold (ending of elongation).

contractile forces along the animal. Once the embryo has fully elongated, the above cytoskeletal structures disorganise and the newly synthesised cuticle holds the worm in its final shape (Priess and Hirsh, 1986)

3.1.2 Post-embryonic development

Post-embryonic development proceeds through the L1, L2, L3 and L4 larval stages to adult (Fig. 3.1). These stages are temporally separated by a cyclical repeated phenomenon, termed the "moulting cycle", which involves the synthesis of a new cuticle and removal of the previous one. Somatic blast cells at the L1 stage divide during larval development contributing to the hypodermis, nervous system, muscle and somatic gonadal structures. The germline proliferates to fill the gonad as it is formed during larval development, and it continues its proliferation throughout adulthood. The external morphological differences between sexes becomes evident after the final moult, at the adult stage, with the presence of the vulva and tail structures specific of hermaphrodites and males respectively. The formation of new structures during post-embryonic development does not involve major changes in the overall architecture of the animal that was attained during embryogenesis. There is a continuous increase in body size throughout post-embryonic development.

3.2 THE HYPODERMIS

The hypodermis is the outer monolayer of epithelial cells that surrounds internal cells such as muscles, intestinal cells, and neurons. It does not enclose the animal where the pharynx, the anus, the excretory pore, sensilla and (in adult hermaphrodites) the vulva are connected to it. Specialised cells named interfacial cells mediate the connection between the hypodermis and the corresponding organs. In all the above openings, the interfacial cells present morphological and functional characteristics intermediate between the hypodermis and the other organs they interface.

As is common to other epithelial cells, two major features of the hypodermis are its apical-basal polarity and cell-to-cell contacts. The outer, apical surface is covered by an extracellular matrix known as the cuticle, which acts as an impervious barrier between the nematode and the environment. The inner, basal surface is closely apposed to body wall muscles by the other type of extracellular matrix in the nematode, the basement membrane. This typical polarisation is reflected in the distribution of the cytoplasmic components. Thus, it is most likely that the

apparatus involved in cuticle synthesis and secretion in hypodermal cells is proximal to the apical membrane.

Hypodermal cells are linked at their apical surface by tight intercellular junctions, the adherens junctions. In addition to mediating contacts between neighbouring cells, adherens junctions are implicated in maintenance of the apical-basal polarity as well as in transmission of contractile forces within individual cells across the entire epithelial layer. Such propagation of local forces is essential for any morphogenetic process (see section 3.1.1).

The hypodermis is a versatile multifunctional tissue that establishes the basic body form of the nematode, synthesises the cuticle, acts as a storage organ and is involved in the elimination of dead cells. Moreover, several classes of hypodermal cells are stem cells, generating a variety of hypodermal cells and neurons during *C. elegans* development.

3.2.1 Morphogenesis of the hypodermis

During embryogenesis, completion of gastrulation gives way to the formation of the hypodermis. Six parallel rows of epithelial cells: two dorsal, two lateral (left and right) and two ventral (left and right), are initially aligned along the dorsal area of the embryo leaving the ventral stem cells exposed (Podbilewicz and White, 1994). The two dorsal rows migrate and interdigitate to form a single row of cells. At the same time the left and right ventral rows spread around the embryo and meet at the ventral midline, where they form close contacts. These cell movements result in enclosure of an ellipsoidal embryo (comma), which corresponds to the preceding stage to elongation. In addition to the above rows of cells, there are small hypodermal cells located in the head and tail of the embryo. Some of the seventy-eight epithelial cells that constitute at first the hypodermis of the embryo fuse generating multinucleate cells (syncytia). These hypodermal fusions occur prior to and (mostly) in the course of elongation.

The hypodermis of the three-fold embryo and the newly hatched larva is comprised by six concentric rings hyp1-hyp6 in the head; hyp7, two lateral rows of ten cells (named seam cells) and two rows of six ventral cells (P cells) in the main body; and four cells hyp8-hyp11 in the tail (Sulston *et al.*, 1983). hyp1 through hyp6 contain three, two, two, three, two and six nuclei respectively. hyp7 has twenty-three nuclei, and hyp8 through hyp11 contain one, one, two and one nuclei respectively. The twenty-three nuclei hyp7 syncytium is cylindrical in its anterior part around the excretory pore and in its posterior part around the anus, and covers only the dorsal area of the main body at this stage. The lateral rows of ten seam cells (H0-H2, V1-V6,

T), that run along each side of the animal, and the six pairs of ventral P cells (P1/2-P11/12) are all embedded in hypodermal syncytia (hyp5, hyp6, hyp7). The six pairs of P cells are in close contact by their dorsal side with the left and right V1-V6 seam cells.

During larval development, most of the seam cells and all the P cells act as stem cells (Sulston and Horvitz, 1977). With the exception of the H0 left and right, H1-H2 and V1-V6 on each lateral side round up and divide at each larval stage generating syncytial hypodermal nuclei and seam cells. T cells only divide at the L1 and L2 stages. In general the anterior daughters of each division fuse to hyp7 while the posterior daughters are themselves seam cells. This cell division pattern is inverted for H1 during the first two divisions. At the same time that the anterior daughters are joining the syncytium, the stem daughters stretch longitudinally until they contact their anterior and posterior neighbours forming again a continuous lateral row of seam cells (Podbilewicz and White, 1994). H2, V5 and T differ from the other stem seam cells in that they also produce neuroblasts (Sulston and Horvitz, 1977). The seam cell divisions end about the middle of the L4 stage with the fusion of fifteen seam cells on each side to form a continuous lateral band. The two seam syncytia are embedded in the hyp7 syncytium but remain separated from it. In addition to a stem cell division, most seam cells undergo a proliferative cell division early in each larval stage. In the male, further divisions of V5, V6 and T cells produce the nine pairs of ray precursor cells that give rise to the eighteen sensory rays of the tail. Once the rays have differentiated, the posterior four ray precursor cells fuse with hyp7 while the anterior five ray precursors fuse together forming a distinct syncytium called the tail seam (set).

After the first division, about the mid-L1 stage, the anterior daughters of V2-V6 on each side send cytoplasmic processes towards the ventral midline breaking the contacts between P cells (Podbilewicz and White, 1994). The fusion of these seam cells to hyp7 results in the isolation of six pairs of cells (left and right) from the two initial rows of P cells. Subsequently, the isolated pairs of P cells migrate to the ventral cord where they form a single row of cells. P cells divide in the course of larval development producing neurons and hypodermal cells (Sulston and Horvitz, 1977). About the end of the L3 stage in the hermaphrodite, the six ventral P3-P8 cells divide generating twelve cells, six of which fuse to hyp7 while the other six become vulval precursor cells. Further divisions of the vulval precursor cells give rise to the twenty-two hypodermal cells constituent of the vulva (the hypodermal structure through which sperm enter and fertilised eggs are laid). The P3-P8 cells do not divide in the male; instead, the two posterior P10 and P11 cells divide to form part of the male tail. The hyp7 syncytium will cover the ventral side of the main body as a result of the incorporation of some of the P daughter cells. All

the hypodermal cells that join the hyp7 syncytium, either descendants from seam cells or P cell divisions, endoreduplicate their DNA and become tetraploid prior to fusion (Hedgecock and White, 1985). A total of one hundred and ten hypodermal cells fuse to hyp7, which contains one hundred and thirty-three nuclei in the adult (Sulston and Horvitz, 1977).

3.3   THE CUTICLE

The cuticle is the extracellular matrix that encloses the animal. It lines the apical surface of the hypodermis and therefore is pierced by the same openings as the epithelial tissue: the mouth, the anus, the excretory pore, sensilla and (in adult hermaphrodites) the vulva. The cuticle is a tough, flexible structure with several biological functions. It maintains post-embryonic body shape, acts as a hydroskeleton, allows movement, confers protection and permits interaction with the external environment.

3.3.1   Cuticle synthesis

In common with other nematodes, *C. elegans* synthesises five cuticles throughout its development: one at the end of embryogenesis and the remaining four at the end of each larval stage. With the exception of the first cuticle that is formed within the eggshell, each new exoskeleton is produced under an existing cuticle, which is then detached and shed during the moult (Singh and Sulston, 1978). Towards the end of each larval stage, the nematode enters a period of inactivity, referred to as lethargus, in which pharyngeal pumping and locomotion decrease gradually. The old cuticle begins to loosen at the head and mouth, then around the tail. Subsequently, the animal often spins and flips around its longitudinal axis, which further detaches the old cuticle from the hypodermis. The new cuticle is formed during this period of lethargus. Just prior to shedding of the old cuticle (ecdysis), the g1 pharyngeal gland secretes granules containing mainly proteases and the pharynx starts spasmodic contractions. The cuticular lining of the pharynx breaks, the old cuticle expands around the head and the animal pulls back repeatedly until the remainder of the cuticle in the pharynx is expelled. Once the pharynx lining is replaced, the nematode breaks the old cuticle pushing with its head and then crawls out of it.

As previously mentioned, the hypodermis is the tissue responsible for the production of the cuticular components. Large Golgi bodies are present in the cytoplasm of hypodermal cells between two and four hours before lethargus (Singh and Sulston, 1978). During the lethargus period, densely packed vesicles bud off from the Golgi bodies and approach the apical surface

where the cuticle is being laid down. These Golgi bodies and vesicles are indicative of synthetic activity in the secretory pathway and most likely carry proteins required for the formation of the cuticle.

At the beginning of the elongation of the embryo and during each larval moult, bundles of circumferentially oriented actin filaments organise and attach to the apical surface of the dorsal and ventral hypodermis of the animal. The arrangement of actin filaments into such bundles is temporal; they are present just prior to cuticle assembly and disorganise once the cuticle has been laid down. Costa *et al.*, (1997) have shown that the pattern of annuli in the L1 cuticle (see section 3.3.2) coincides with the pattern of the submembranous actin bundles in the hypodermis of the embryo. Contraction of actin bundles and concomitant elongation of the embryo modify the flat apical membrane of the hypodermal cells to a furrowed surface, which act as a template for the deposition of the first larval cuticle. The actin bundles present during post-embryonic moults may also have an important role in defining cuticle structure (Costa *et al.*, 1997). Moreover, contraction of these actin bundles may entail folding of the new cuticle that is being synthesised underneath the old cuticle. This would allow the formation of a new cuticle of larger dimensions than the old cuticle, permitting the rapid growth of the larvae after moult.

3.3.2 Cuticle structure

The *C. elegans* cuticle is a complex structure, presumably due to the diverse functions it performs. Each of the five cuticles produced during the nematode's life cycle has distinct biochemical and morphological properties (Cox *et al.*, 1981a; Cox *et al.*, 1981b). The adult cuticle has been thoroughly characterised by electron microscopy (Cox *et al.*, 1981b; Peixoto and de Souza, 1992). It is composed of six layers: epicuticle, external cortical, internal cortical, intermediate, fibrous and basal (Fig. 3.3A). The epicuticle is the outermost layer and presents a tri-laminate pattern resembling a membrane structure. The intermediate layer is probably fluid-filled and contains columns of filamentous material, termed struts, which connect the cortical and basal layers by fibres with a particular arrangement. The fibrous layer comprises two sublayers of fibres that spiral around the animal in opposite directions, each oriented at 60° angle relative to the animal's long axis. The basal layer has a loosely organised fibrillar appearance and is the innermost layer, adjacent to the hypodermis.

The L2, L3 and L4 cuticles are similar in structure and differ from the structurally distinct L1, dauer and adult cuticles (Cox *et al.*, 1981b). Each stage-specific cuticle has epicuticle, external cortical and internal cortical layers, which can differ from those of other stages. An intermediate

Figure 3.3: Structure of the adult cuticle in *C. elegans*.

A, scheme representing the different layers that constitute the cuticle: cortical (which includes the epicuticle, external and internal cortical), intermediate, fibrous, and basal. The fibrous layer is composed of two sublayers of fibres that meet each other at 60° angle and resemble a fish-bone pattern. Struts are predominantly organized into rows on either side of the annular furrows and are absent underneath the alae. **B,** traverse section through the animal illustrating hypodermal cells and structural features on the surface of the cuticle.

(Adapted from Johnstone, 1993).

layer separating the cortical from the basal layers is absent in juvenile cuticles. The L2-L4 cuticles have two fibrillar layers of different orientation, probably analogous to the fibrous layer in the adult cuticle. The L1 and dauer cuticles contain a striated layer with distinctive properties in place of the fibrous layer. In the L1 cuticle the striated layer apposes the hypodermis due to the absence of a separating basal layer. A surface coat or glycocalyx is loosely attached to the surface of the epicuticle of all stages.

The L1, dauer and adult cuticles display external specialisations, called alae, placed longitudinally on the right and left lateral sides of the animal, over the seam cells (Fig. 3.3B). As shown by cell ablation experiments, the seam cells appear to be the hypodermal cells responsible for the formation of the alae (Singh and Sulston, 1978). The alae are protruding ridges of the epicuticle and cortical layers and have stage-specific architecture (Cox _et al._, 1981b). Three ridges of equal size constitute the alae of adults; there are no struts within the intermediate layer beneath these alae. The dauer alae comprise five broad ridges; the outermost ridges on the left and right sides are more prominent than the three in the centre. The L1 alae are composed of three ridges, with the central ridge being much larger than the other two. The surfaces of all larval and adult cuticles have invariably circumferential, regularly spaced furrows or indentations termed annuli (Fig. 3.3). Annuli are absent on the lateral sides of the nematode's cuticle when the alae are present.

3.3.3 Cuticle components

The cuticle is acellular and is composed of structural and non-structural components. Structural components include collagens and non-collagen proteins such as cuticlins (Fujimoto and Kanaya, 1973). The basal, fibrous and internal cortical layers as well as the adult struts are predominantly made up of collagen proteins; the external cortical layer contains mainly non-collagen proteins (Fig. 3.3 Cox _et al._, 1981a). The cuticlin proteins CUT-1 and CUT-2 are present in low amounts in the cortical layer of the cuticle in all stages and more abundantly underneath the dauer alae (Ristoratore _et al._, 1994; Favre _et al._, 1995). Non-structural components constitute the epicuticle as well as the surface coat external to the cuticle (Maizels _et al._, 1993; Blaxter and Bird, 1997). The epicuticle is a lipid-rich layer while the surface coat contains surface-associated proteins and glycoproteins (carbohydrate-rich).

3.3.3.1 Cuticle collagens

Collagen proteins are the major constituents of the nematode cuticle, where they are extensively cross-linked by reducible disulphide bridges and non-reducible di- and isotrityrosine bonds (Cox, 1992; Page, 2001). They are encoded by a large gene family predicted to contain approximately one hundred and fifty-four members (*C. elegans* sequencing consortium, 1998; Johnstone, 2000). Cuticle collagen genes are generally dispersed throughout the genome (Cox *et al.*, 1985), and represent approximately 1% of its content. As inferred from a set of over thirty genes, they are small (<2 Kb), have from one to three short introns and encode polypeptide chains of 26-35 kDa (Kramer, 1997). A common structure defining cuticular collagen polypeptides predicted by the *C. elegans* genome has been proposed recently (Fig. 3.4; Johnstone, 2000). It consists of a non-collagen N-terminus, two main collagen domains and a non-collagen C-terminus. The N-terminal non-collagen region varies in length and comprises short conserved segments including a predicted signal peptide and a subtilisin-like cleavage site. Closely preceding the N-terminal collagen domain (of eight to ten Gly-X-Y repeats) there is a cluster of three cysteine residues (occasionally two). Before and just after the C-terminal collagen domain (of forty to forty-two Gly-X-Y repeats) there is respectively a cluster of two to three cysteine residues and a cluster of two cysteine residues. The C-terminal collagen domain normally has one or two small disruptions that vary in size and position between polypeptide chains. The structure of *C. elegans* cuticle collagens is most similar to that of vertebrate non-fibrillar FACIT collagens (section 1.3.2). Both comprise more than one collagen domain (which can be interrupted by short non-Gly-X-Y stretches) separated by non-collagen regions and have clusters of cysteine residues close to collagen domains (Fields, 1988). Cuticle collagen genes can be classified into six groups based on the pattern of the conserved cysteine residues (Johnstone, 2000). Consistent with this classification, collagens within the same group are more similar in their N- and C-terminal non-collagen regions than to collagens in other groups. The primary structure of the polypeptide chains most likely determines the association of these chains into collagen trimers as well as the interaction of different collagen molecules within the cuticle (Johnstone and Barry, 1996; McMahon *et al.*, 2003).

3.3.4 Cuticle collagen expression

Cuticle collagen genes are spatially and temporally regulated. Their synthesis is restricted to hypodermal cells (Singh and Sulston, 1978; Edwards and Wood, 1983), and is subject to a complex regulatory pattern throughout *C. elegans* development. This complex pattern stems from temporal regulation of these genes at several levels. (1) The expression of cuticle collagens
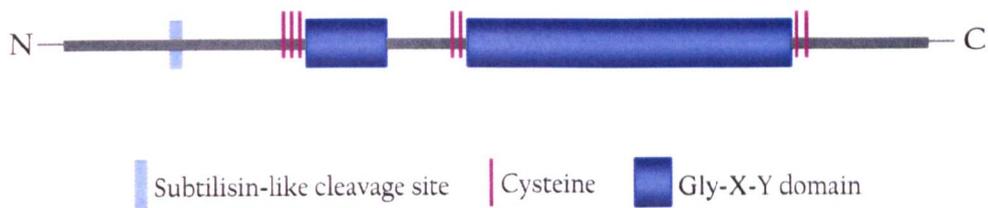
Figure 3.4: Schematic representation of the common structure of *C. elegans* cuticle collagen genes including the substilisin-like cleavage site, conserved cysteine residues and collagen (Gly-X-Y) domains. The N-terminal collagen domain contains between eight and ten Gly-X-Y repeats while the C-terminal collagen domain comprises between forty and forty-two Gly-X-Y repeats.

oscillates throughout post-embryonic development, peaking once during each larval stage (Johnstone and Barry, 1996). (2) Individual cuticle collagens are expressed at different times during the same moulting cycle and can be classified as early-expressed (four hours before cuticle secretion), middle-expressed (two hours prior to cuticle secretion) and late-expressed genes (coincident with cuticle secretion) (Johnstone and Barry, 1996). Such ordered expression possibly facilitates collagen trimerisation and interaction between collagens constituent of the same cuticle substructure (Johnstone and Barry, 1996; McMahon *et al.*, 2003). (3) Some cuticle collagens are expressed at all moulting cycles whereas other cuticle collagens are expressed only at specific cycles (Cox and Hirsh, 1985). (4) Levels of expression of individual cuticle collagens can vary at different moulting cycles (Cox and Hirsh, 1985). The type of collagens and their levels at each particular moulting cylce must contribute to the stage-specific characteristics of the cuticle. An additional level of complexity arises from expression of certain cuticle collagens only in a specific subset of hypodermal cells (Kramer, 1997).

3.3.5 Mutants with cuticular defects

Distinct mutations in genes with cuticle function in *C. elegans* can result in death late in embryogenesis, and/or gross anomalies in the animal's overall morphology (reviewed in Johnstone, 1994; Kramer, 1994a; Kramer, 1994b; Kramer, 1997). The morphological defects caused are diverse and have been defined as follows: blister (Bli), blisters on the adult cuticle; dumpy (Dpy), short and fat; long (Lon), long and thin; roller (Rol), helically twisted either to a left-handed screw direction (LRol) or to a right-handed screw direction (RRol); abnormal hermaphrodite tail (Tal) (Brenner, 1974; Kusch and Edgar, 1986). Certain mutations can alter body shape in more than one such way. Moreover, different single mutations within the same gene can produce distinct mutant phenotypes.

3.3.5.1 Cuticle collagen mutants

Thus far, fifteen cuticle collagen genes have been identified by mutagenesis to affect body morphology: *dpy-13* (von Mende *et al.*, 1988), *sqt-1* (Kramer *et al.*, 1988), *rol-6* (Kramer *et al.*, 1990), *dpy-7* (Johnstone *et al.*, 1992), *dpy-2* (Levy *et al.*, 1993), *dpy-10* (Levy *et al.*, 1993), *sqt-3* (Van der Keyl *et al.*, 1994), *dpy-3* (McMahon *et al.*, 2003), *dpy-8* (McMahon *et al.*, 2003), *lon-3* (Nyström *et al.*, 2002), *dpy-5* (C. Thacker and A.M. Rose, personal communication), *bli-1* and *bli-2* (J. Crew and J.M. Kramer, personal communication), *rol-1* and *rol-8* (http://www.wormbase.org). Several mutant alleles have been assigned to each of these genes. The phenotypes produced by mutations in these loci fall into the morphological classes

described above. In general, the isolated mutations alter the animal's morphology at certain or all post-embryonic stages. Recalling the structural role of collagen proteins within the cuticle, the defects reported in these fifteen collagen genes most likely affect cuticle assembly. This is supported by immunofluorescence and/or ultrastructural studies of the cuticles of some of these collagen mutants. As determined by several electron microscopy techniques, *rol-6(su1006)* and *sqt-1(sc13)* animals have a defective fibrous layer (Peixoto *et al.*, 1998; Peixoto *et al.*, 2000). By means of scanning electron microscopy and immunolocalisation analyses, mutations within *dpy-2*, *dpy-3*, *dpy-5*, *dpy-7*, *dpy-8*, *dpy-10* or *dpy-13* have been shown to affect the surface of the adult cuticle, specifically the width of the annuli or integrity of the annular furrows (McMahon *et al.*, 2003). The way in which the cuticle structure is disrupted depends on the particular function of the mutated collagen gene and on the nature of the mutation. Different collagen mutants with the most similar body abnormalities might display analogous cuticular defects, indicating that the corresponding collagen products may have similar roles and interact to form the same cuticle structure (Kramer and Johnson, 1993; McMahon *et al.*, 2003).

A temperature sensitive mutation within the locus *sqt-3*, namely *e2117*, has been reported to give embryonic death after elongation at restrictive temperature 25°C (Priess and Hirsh, 1986; Van der Keyl *et al.*, 1994). Priess and Hirsh (1986) have examined the cuticles of *sqt-3(e2117)* mutant embryos by transmission electron microscopy and shown that these lack the striated layer typical of the L1 stage. The severity of this defect is consistent with *sqt-3*, which codes for the cuticle collagen COL-1, being abundantly expressed in embryos (Kramer *et al.*, 1985; Van der Keyl *et al.*, 1994). Thus, major cuticular defects (such as absence of a whole cuticle layer) may result in lethality whereas minor defects (such as presence of an anomalous cuticle layer or absence of a specific cuticular feature) may impair cuticle function without affecting the viability of the animal.

Collagen polypeptides associate into trimers, which subsequently interact to form fibres that assemble into higher-order structures. This inherent characteristic of collagens makes loss of function phenotypes most important to interpret the role of collagens within the cuticle, as opposed to missense mutations that interfere with the function of normal collagens in the assembled cuticle. Null mutations in *dpy-7*, *dpy-13* or *dpy-10* cause strong recessive phenotypes (Dpy or Dpy LRol), indicating that these gene products might be essential for proper cuticle assembly (McMahon *et al.*, 2003; von Mende *et al.*, 1988; Levy *et al.*, 1993). In contrast, null mutations in *sqt-1* or *rol-6* result in weak recessive phenotypes (Tal or very weak Dpy), suggesting that these gene products might be dispensable for the cuticle structure (Kramer and

Johnson, 1993). To date, many independent mutagenic screens for mutants with morphological defects have been carried out. There are five cuticle collagen genes currently known to mutate to a severe loss of function phenotype: *dpy-5*, *dpy-7*, *dpy-10*, *dpy-13* and *lon-3*; the other ten cuticle collagens mutable to changes in body shape (see above) either have a weak loss of function phenotype or lack identified null alleles. Recently, six of these ten collagens (*bli-1*, *bli-2*, *dpy-2*, *dpy-3*, *dpy-8*, *sqt-3*) have been ascribed Bli or Dpy phenotypes by RNAi (Kamath *et al.*, 2003; McMahon *et al.*, 2003). Therefore, as inferred from null and RNAi phenotypes, only eleven of the predicted one hundred fifty-four cuticle collagen genes in the genome appear to have a gross loss of function phenotype. This indicates that a significant number of cuticle collagen genes might be functionally redundant, as would be expected for such a large multi-gene family.

The strict requirement of glycine as every third amino acid within the triple helix makes this residue especially vulnerable to mutation. Any other amino acid in lieu of glycine within the Gly-X-Y triplet may disrupt the proper formation of the triple helix, with the consequent degradation of the abnormal collagen molecule (Prockop and Kivirikko, 1995). Such glycine replacements in the nematode cuticle collagens are recessive and cause similar defects, although not always identical, to the null phenotype of the corresponding collagen. Different glycine substitutions in *sqt-1* result in a weak LRol and/or weak Lon in addition to the Tal phenotype seen in the null mutant (Kramer and Johnson, 1993). Moreover, two distinct, single glycine mutations in *dpy-10* produce a stronger Dpy phenotype than that of the null mutation (Levy *et al.*, 1993). For both *sqt-1* and *dpy-10*, the resemblance between the phenotype(s) caused by glycine substitutions and the null phenotype indicates that secretion of the glycine defective collagen might be considerably reduced. However, the fact that glycine substitutions produce slightly more severe defects than the null mutation in either *sqt-1* or *dpy-10* suggests that some of the glycine defective collagen may affect normal cuticle synthesis. In the case of *dpy-13* and *dpy-7* collagens, the defects caused by a glycine substitution do not differ noticeably from the null phenotype (von Mende *et al.*, 1988; McMahon *et al.*, 2003). McMahon *et al.*, (2003) have shown recently that only small amounts of glycine defective DPY-7 are secreted and assembled abnormally within the cuticle.

Single base substitutions of highly conserved residues within the N- or C-terminal non-collagen domains have been also found to give a detectable phenotype. Mutation of one of the two conserved cysteine residues in the C-terminal domain of *sqt-1* or *rol-6* causes a recessive LRol phenotype (Kramer and Johnson, 1993; Yang and Kramer, 1994). For some human collagens,

disulphide bonding between the C-terminal domains of the polypeptides is required for the formation of the triple-helical molecule (Engel and Prockop, 1991; Prockop and Kivirikko, 1995). This seems, however, not be the case of nematode cuticle collagens since these cysteine mutations cause distinct defects from the null phenotype, suggesting that the mutant collagen molecules assemble and interfere with cuticle function. Recently, Yang and Kramer (1999) have shown that loss of one of the two cysteines in *sqt-1* does not block assembly of the mutant collagen and its consequent secretion into the cuticle. The mutated cysteine appears to alter SQT-1 function by inhibiting the formation of normal cross-links between the collagen chains, with an ultimate effect on cuticle structure. Similar defects probably result from substitution for the analogous cysteine residues in the *rol-6* collagen.

Substitution of one of the two conserved arginine residues for a non-basic amino acid in the N-terminal domain of *dpy-10*, *sqt-1* or *rol-6* has been reported to have a dominant or semi-dominant effect (Levy *et al.*, 1993; Kramer and Johnson, 1993; Yang and Kramer, 1994). Specifically, such mutations give dominant LRol and recessive Dpy RRol phenotypes in *dpy-10*, dominant RRol and recessive Dpy phenotypes in *sqt-1*, and a semi-dominant RRol phenotype in *rol-6*. The two arginine residues are within a short segment conserved in cuticle collagens that is predicted to act as a cleavage site for a subtilisin-like protease (Yang and Kramer, 1994). Subtilisin-like proteases are involved in endoproteolytic processing of certain proproteins (Barr, 1991). Yang and Kramer (1999) have demonstrated recently that the arginine defective forms of SQT-1 are longer than the wild-type polypeptides, in a portion corresponding to the length of the region N-terminal to the cleavage site. Moreover, the region N-terminal to the cleavage site is present in the mutant forms but not in the wild-type polypeptides. These results confirm the presence of the protease cleavage site in the SQT-1 polypeptide, as well as indicating that mutation of one of the two arginine residues modifies this cleavage site and alters processing of the collagen. This processing might be an important requirement for collagen secretion since only small amounts of uncleaved SQT-1 are incorporated into the cuticle (Yang and Kramer, 1999). The reduced amount of the secreted mutant collagen appears to be sufficient to dominantly disrupt the cuticle structure. Assuming that processing by a subtilisin-like protease is common among cuticle collagens, the substitution for arginine residues detected in *dpy-10* and *rol-6* will affect the function of these collagens in a similar manner to *sqt-1*. This will also be consistent with the dominant and semi-dominant defects of these *dpy-10* and *rol-6* mutants.

3.3.5.2 Other cuticle defective mutants

Another class of genes that can be mutable to affect cuticle morphology encode collagen-processing enzymes. *bli-4* has been shown to encode a member of the kex2/subtilisin-like family of serine proteases (Thacker *et al.*, 1995). This class of serine proteases are involved in the maturation of a wide range of secretory protein precursors by endoproteolytic cleavage of their prodomain (reviewed in Thacker and Rose, 2000). The majority, if not all cuticle procollagens predicted from the *C. elegans* genome, contain a subtilisin-like cleavage site in the N-terminal domain. Analyses of fourteen different mutations in the *bli-4* gene have demonstrated that BLI-4 protease activity is required for normal cuticle synthesis, most likely by the processing of cuticle collagen genes (Peters *et al.*, 1991; Thacker *et al.*, 1995; Thacker *et al.*, 2000). Thirteen of these individual mutations result in late embryonic death, similar to that caused by the *sqt-3(e2117)* mutant allele; the other characterised mutation results in a viable blistering defect of the adult cuticle, analogous to the blister phenotype of some cuticle collagen mutants. *bli-4* is a complex locus that produces at least nine gene products in which a common N-terminus is alternately spliced to specific C-termini (Thacker *et al.*, 2000). The above lethal mutations have been found to reside in the common N-terminal region, and therefore are expected to affect the function of all the *bli-4* gene products. In contrast, the viable mutation deletes a region that is only included in five of the nine BLI-4 isoforms (Thacker *et al.*, 1999).

Prolyl 4-hydroxylase (P4H) is a multicomplex enzyme that catalyses hydroxylation of Pro residues in the Y position of the Gly-X-Y repeats of procollagen chains and acts as a molecular chaperone in the correct folding of collagen molecules. In *C. elegans*, P4H is generally a $\alpha_2\beta_2$ mixed tetramer, in which the α-subunits are the *phy-1* and *phy-2* gene products and the β-subunit is the *pdi-2* gene product (Myllyharju *et al.*, 2002; Myllyharju 2003). P4H is also found in lower amounts as $\alpha\beta$ dimers. In addition to its role as P4H β-subunit, protein disulphide isomerase (PDI) is involved in the catalysis of disulphide bonds as well as in the folding and assembly of diverse proteins, including collagens (see section 4.1.4). Recently, the enzyme P4H has been proved to be essential for nematode development and normal biogenesis of cuticle collagens (Friedman *et al.*, 2000; Hill *et al.*, 2000a; Winter and Page, 2000). Removal of P4H activity by the simultaneous disruption of *phy-1* and *phy-2*, or *pdi-2*, results in embryonic death, most likely due to impaired synthesis of the L1 cuticle (Friedman *et al.*, 2000; Winter and Page, 2000). Interestingly, a time course of the development of embryos deficient in *pdi-2* function has revealed that these animals elongate and subsequently retract back in a similar manner to *sqt-3(e2117)* mutant embryos (Winter and Page, 2000). Moreover, different null alleles of the *phy-1* gene result in a post-embryonic Dpy phenotype analogous to that of some cuticle collagen

mutants (Friedman *et al.*, 2000; Hill *et al.*, 2000a; Winter and Page, 2000). It should be noted that the *phy-1* gene is defined by the mutated *dpy-18* locus.

Lesions causing complete loss or reduction of an enzymatic function, such as processing of collagens, are expected to give mutant phenotypes (generally recessive) that can vary in degrees of severity depending on the reduction of the enzyme activity. This behaviour differs from that of genes encoding for structural components, such as cuticle collagens, in which distinct lesions of the same locus can result, some of them in a dominant manner, in a variety of morphological phenotypes.

3.3.6 Concluding remarks and project direction

In *C. elegans*, failure in synthesising a normal cuticle may result in lethality late in embryogenesis or in viable body shape changes. Collagens are the major components of the nematode cuticle, and therefore diverse quantitative and/or qualitative alterations of these can result in the above lethal and/or morphological defects. To date, late embryonic lethality has been reported as a result of disruption of SQT-3, BLI-4, P4H or PDI-2 function. SQT-3 is a cuticle collagen abundantly expressed in embryos with an important contributing role in the formation of a whole cuticle layer. Alteration of its function can result in the production of a severely malformed and dysfunctional cuticle with ultimate lethal effects. BLI-4, P4H and PDI-2 are enzymes involved in the processing of cuticle collagens. Alteration of the activity of any of these enzymes is therefore expected to result in severe cuticular defects by having a disruptive effect on all collagens synthesised to produce the cuticle. This is supported by the fact that the embryonic lethality resultant from lack of one of these three enzymatic activities is analogous to that consequent from a specific defect in SQT-3.

The main goal of this project has been the study of collagen biogenesis. Thus, I have focused on the analysis of *C. elegans* lethal mutants with phenotypes suggestive of major cuticular defects, as they possibly result from problems in the general process of cuticle production. Viable morphological defects are likely to be the result of impaired function of individual cuticle collagens and hence were avoided. Specifically, the mutants characterised in this project are *stc-1(ij15)*, *let-607(h402)* and *let-607(h189)*. *stc-1(ij15)* was isolated from a screen previously performed (C. Clucas and I. Johnstone, Glasgow University, Glasgow, UK) to detect mutations that cause death late in embryogenesis. *let-607(h402)* and *let-607(h189)* mutants were identified by A. Rose (University of British Columbia, Vancouver, Canada), and were of interest as

interference of *let-607* function resulted in dead eggs and Dpy animals, suggestive of cuticular defects.

3.4 RESULTS

3.4.1 Characterisation of *stc-1(ij15)*, *let-607(h402)* and *let-607(h189)* mutants

3.4.1.1 Definition of *stc-1(ij15)*, *let-607(h402)* and *let-607(h189)* mutant phenotypes

Phenotypic analysis of all mutants was performed using microscopy with Nomarski optics (Fig. 3.5) and, in the case of *let-607(h402)* larvae, by SEM (Fig. 3.6). A time course of *stc-1(ij15)* embryos development was performed. The mutant embryos are able to complete elongation but then disrupt at the ventral side along the middle-posterior part of their body (Fig. 3.5C). Eventually, these embryos lose most of their morphology as depicted in Fig. 3.5D.

The development of *let-607(h402)* homozygous embryos was also followed by time course. The mutants present a range of phenotypic defects and, according to these, have been classified in three categories. All animals elongate successfully to a three-fold embryo; it is at the last stage of embryogenesis that abnormalities first become evident. Class I mutants rupture from the ventral side at the mid-anterior part of their body (Fig. 3.5E). The rate of collapse and extrusion of cellular material is variable, ending in an amorphous mass of cells (Fig. 3.5F). A significant number of mutant embryos, however, do not disrupt but shorten in length and look wrinkled (Fig. 3.5G). These constitute the class II phenotype. Class III animals succeed in embryogenesis and hatch as small, fat larvae (Fig. 3.5J). They frequently present bulges and constrictions randomly about the body, and display gross cuticular defects as revealed by SEM (Fig. 3.6). Annuli and alae are severely irregular and tend to be absent over the regions of body swelling. The material torn on the surface of the larva shown in Fig. 3.6B could correspond to the outer layers pulling apart from the rest of the exoskeleton. This was not observed on wild-type worms manipulated and treated under the same conditions (see Fig. 5.9 A and C). The class III mutant larvae live for several days without obvious growth, lose motion progressively and then die. Both *let-607(h402)* embryos and hatchlings sometimes have abnormal tails, normally a "split tail". Of the total mutant progeny produced by *let-607(h402)/+* hermaphrodites, there is a higher incidence of embryonic lethality (~61% of class I and class II phenotypes) than larval death (~39% of class III phenotype).

Figure 3.5: Morphology of *stc-1(ij15)* and *let-607* mutants under Nomarski microscopy.
A-B, wild-type embryos at three-fold stage of development. **C-D,** *stc-1(ij15)* homozygotes. The mutant embryo bursts after elongation **(C)** and loses most of its structure **(D)**. **E-G,** *let-607(h402)* mutant embryos. **E** and **F** correspond to class I phenotype: the embryo oozes along the ventral midline at the three-fold stage **(E)** and eventually loses the shape **(F)**; **G** represents class II phenotype: the embryo does not disrupt after elongation but retract back. **H,** terminal phenotype of a *let-607(h189)* embryo. **I-K,** *let-607* hatchlings compared to an early stage wild-type larva. **I,** wild-type L1. **J,** *let-607(h402)* larva. **K,** *let-607(h189)* larva.

Facing pages 69-70
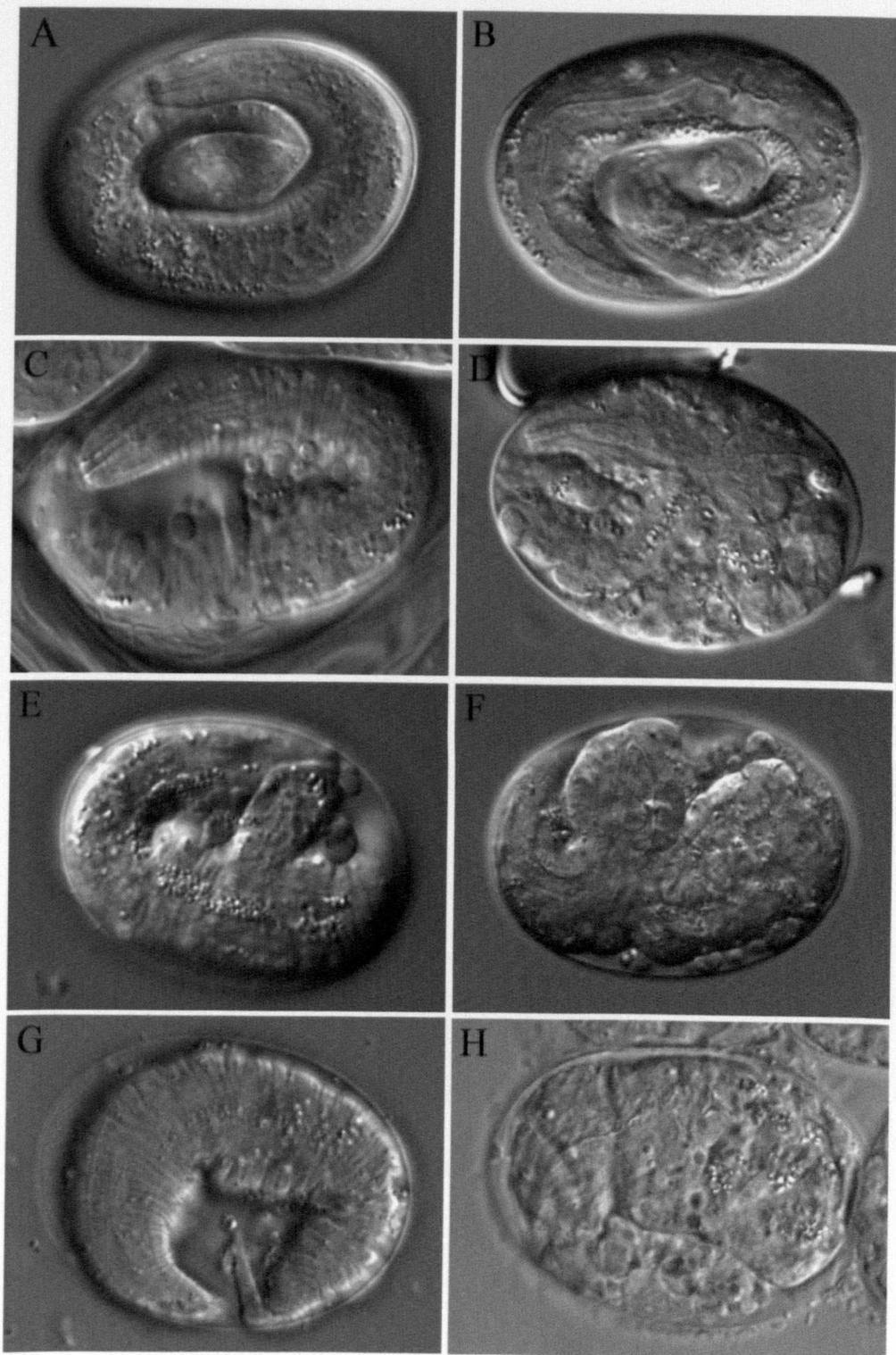
Figure 3.5: Morphology of *stc-1(ij15)* and *let-607* mutants under Nomarski mircroscopy. (Continued)

Figure 3.6: Scanning electron microscopy on *let-607(h402)* hatchlings.

Micrographs **A** and **B** depict diverse cuticular defects of class III *let-607(h402)* mutants. Annuli are indicated by arrows, the material detaching from the surface of the larva by a double arrow and the alae by arrowheads. The regions lacking the ala are denoted by an asterisk.
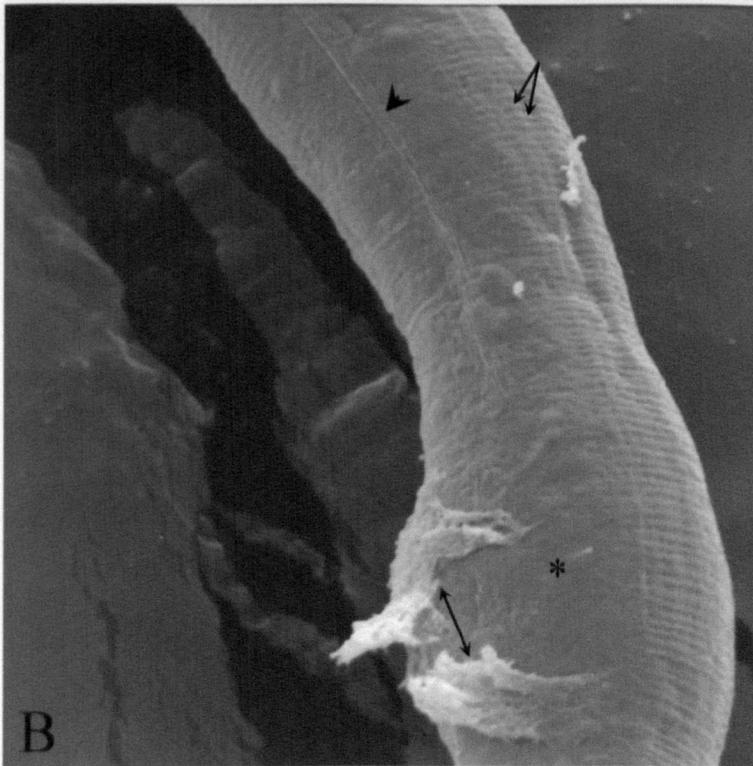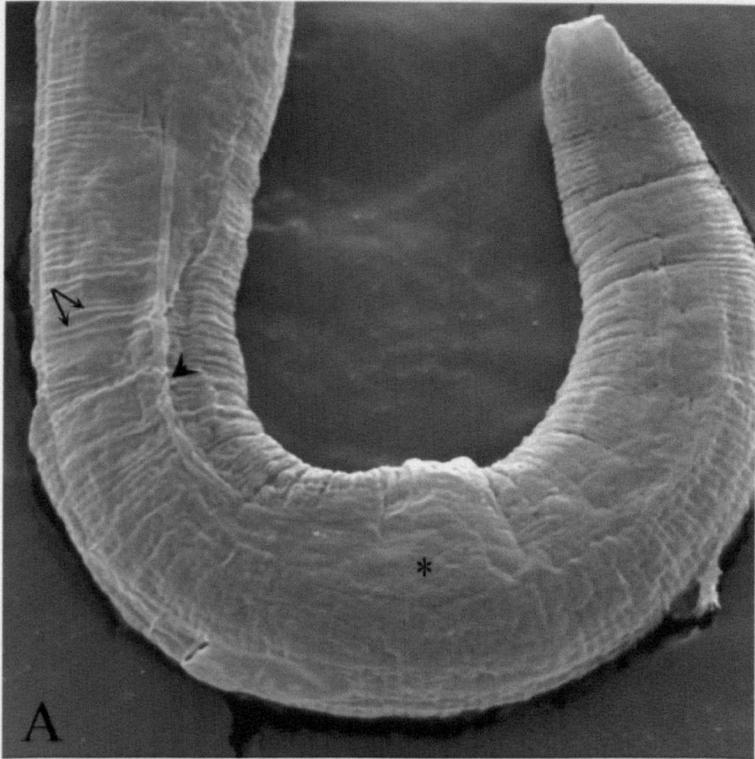
The allele *let-607(h189)* causes a distinct spectrum of phenotypes. About 1% of *let-607(h189)* homozygotes die at the terminal phase of embryogenesis (Fig. 3.5H); the remaining 99% hatch and do not present visible morphology anomalies (Fig. 3.5K). After a day or so on the plates, they become distinguishable from wild-type larvae in that they hardly move, are loopy at rest and their internal body contents adopt a clearer aspect. Interestingly, many of the larvae that are about to die shrink suddenly when touched with the worm pick.

3.4.1.2 *dpy-5 and unc-13 may influence the let-607(h402) phenotype*

The original strains containing the *let-607* alleles were KR727, genotype *let-607(h402) dpy-5(e61) unc-13(e450) I; sDp2(I;f)*, and KR503, genotype *let-607(h189) dpy-5(e61) unc-13(e450) I; sDp2(I;f)*. These strains were out-crossed twice with N2 males to ensure the loss of the duplicated chromosomal segment *sDp2(I;f)*. The procedure below was then carried out in order to isolate the *let-607* mutations for phenotypic analyses. *dpy-5 let-607 unc-13/+ + +* hermaphrodites were picked clonally, allowed to self and a few viable F_1 animals from plates containing *let-607* mutants were transferred individually onto new plates. This was repeated for several generations to increase the possibility of obtaining recombinant segregants carrying only the *let-607* mutant allele. To identify such recombinant animals, hermaphrodites heterozygous for *let-607* were mated with wild-type animals and the resulting F_1 heterozygous males with either *dpy-5(e61)/dpy-5(e61)* or *unc-13(e450)/unc-13(e450)* hermaphrodites. A 100% wild-type F_2 brood was indicative that *dpy-5(e61)* and *unc-13(e450)* mutations were not present in the male parents. *let-607(h189)/+ I* and *let-607(h402)/+ I* lines were named IA446 and IA447 respectively.

During the process of isolating *let-607* recombinants, an increase of class II and class III phenotypes together with a decrease of class I phenotype was observed in the total *let-607(h402)* progeny. I hypothesised that the reduced presence of mutants with the most severe *let-607(h402)* defect, class I phenotype, could result from a phenotypic enhancing effect of the *dpy-5(e61)* mutation and, in general, to mutations in *dpy* collagen genes. To study this further, the following Dpy lines heterozygous for either *let-607* allele were constructed: *dpy-5(e61) let-607(h402)/dpy-5(e61) +, dpy-7(e88) let-607(h402)/dpy-7(e88) +* and *dpy-13(e458) let-607(h402)/dpy-13(e458) +*; and *dpy-5(e61) let-607(h189)/dpy-5(e61) +, dpy-7(e88) let-607(h189)/dpy-7(e88) +* and *dpy-13(e458) let-607(h189)/dpy-13(e458) +*. The broods of single heterozygous mothers were scored for each line. No differences were detected between the ratios of embryonic and larval lethality obtained in a specific Dpy background and those from wild types heterozygous for a *let-607* allele. This suggests that both *dpy-5(e61)* and *unc-13(e450)* mutations, or other

unidentified linked mutations on the original chromosome, were the cause of higher *let-607* embryonic death.

3.4.1.3 Temperature sensitivity of let-607(h402) and let-607(h189) alleles

As the *let-607(h402)* mutant allele results in variable defects, *let-607(h402)* homozygotes were raised at the temperatures 15°C, 20°C and 25°C to determine if their phenotypes were temperature sensitive. Growth of *let-607(h189)* animals was also studied at the three temperatures, even though the *let-607(h189)* mutant allele generally causes larval lethality. Heterozygous hermaphrodites for either *let-607* mutant allele were cultured individually at each of the above temperatures, and their broods were scored to assess frequencies of embryonic and larval death. Numbers of total progeny, dead embryos and mutant larvae are indicated in Table 3.1. Comparison of the percentages of dead embryos and larvae obtained at the three temperatures indicate that both mutant alleles are not temperature sensitive.

A)

Temperature °C		15	20	25
Total scored progeny		2289	3108	2026
Dead embryos	Total number	311	461	226
	Percentage (%)	13.59	14.83	11.15
Dead larvae	Total number	207	295	225
	Percentage (%)	9.04	9.49	11.10

B)

Temperature °C		15	20	25
Total scored progeny		2687	3572	2268
Dead embryos	Total number	11	52	30
	Percentage (%)	0.41	1.45	1.32
Dead larvae	Total number	679	880	563
	Percentage (%)	25.27	24.64	24.82

Table 3.1: Representation of mutant embryos and larvae obtained at 15°C, 20°C and 25°C for *let-607(h402)* (A) and *let-607(h189)* (B) alleles. The percentages of dead embryos and dead larvae

homozygous for *let-607(h402)* or *let-607(h189)* did not vary notably between the three growth temperatures.

3.4.1.4 Age determination of let-607(h402) and let-607(h189) hatchlings

let-607 larval mutants were staged by different means. Certain *C. elegans* cells undergo stereotyped division programmes during post-embryonic development. The Ll larva has two lateral rows of 10 hypodermal cells, known as seam cells, and an intestine made of 20 cells. The seam cells, except the most anterior ones, divide at each larval stage; and most of the intestinal cells become binucleate at the beginning of the Ll lethargus. I examined intestinal nuclei and seam cells in *let-607* mutants by means of reporter construct expression and fluorescence microscopy. The strains JR1838, JR1986 and JR667 contain the genomic insertions *wIs84*, *wIs118* and *wIs51* respectively. *wIs84* and *wIs118* consist of distinct intestinal gene promoters fused to the *gfp* gene, causing expression of the GFP protein only in intestinal nuclei. *wIs51* comprises a seam cell-specific promoter::*GFP* fusion and results in restricted GFP expression to the seam cells. *let-607(h402)* and *let-607(h189)* heterozygotes were crossed with JR1838, JR1986 and JR667 hermaphrodites, and the F_1 progenies were screened for animals that expressed the reporter fusions and harboured one copy of either mutant allele. The isolated lines carrying the allele *let-607(h402)* plus the GFP insertion *wIs84*, *wIs118* or *wIs51* were named IA460, IA461 and IA473 respectively; and those containing the allele *let-607(h189)* plus *wIs84*, *wIs118* or *wIs51* IA466, IA462 and IA463 respectively. Observation of the mutant larvae at high magnification revealed that *let-607(h402)* hatchlings had mononucleated intestinal cells and about ten seam cells at each lateral side, and *let-607(h189)* mutants presented binucleated intestinal cells and more than ten seam cells per side.

Just hatched *let-607(h402)* animals were monitored under Nomarski microscopy to ascertain if they go through the first larval moult. The majority seemed not to attempt moulting and a few failed during the process. In addition, *let-607(h402)* dying animals had alae indicating the Ll stage. Observation of *let-607(h189)* larval development was more difficult as these are not easily distinguished from wild-type Lls. However, alae were not detected on *let-607(h189)* dying animals which suggests they passed the first moult.

Finally, the morphology of the gonad was observed in both types of mutants. *let-607(h402)* hatchlings present a four-cell gonad primordium typical of the Ll larvae; terminal *let-607(h189)* worms have a L2-stage gonad.

In summary, *let-607(h402)* hatchlings die predominantly as L1s from the following evidence: they tend to contain two rows of ten seam cells, mononucleated intestinal cells, alae, gonad of four cells and generally seem not to be able to moult. *let-607(h189)* worms seem to grow to L2 larvae and die at this stage as indicated by the presence of two rows of more than ten seam cells, binucleated intestinal cells, L2 gonad and the lack of alae.

3.4.1.5 Visualisation of hypodermal cell positions in let-607(h402) and let-607(h189) larvae

The arrangement of hypodermal cells was examined in *let-607(h402)* and *let-607(h189)* hatchlings using the *ajm-1::GFP* marker. *ajm-1* encodes for a protein component of adherens junctions of epithelia, the reporter construct of this gene shows the cell boundaries of individual hypodermal cells by GFP fluorescence (Köppen *et al.*, 2001). As detailed in section 3.2.1, the hypodermis of the newly hatched larva comprises six annular syncytia in the head; a large syncytium (mostly dorsal), two lateral rows of ten seam cells in close contact to two ventral rows of six P cells in the main body; and four cells in the tail. At the mid-L1 stage, the ventral P cells migrate to the ventral cord and form a single row of cells that is not apposed to the lateral seam cells. The boundaries of seam cells and P cells in *let-607(h402)* hatchlings, and of seam cells in *let-607(h189)* larvae are shown in Figure 3.7. The spatial location of these hypodermal cells and their contacts appear to be normal in both *let-607* mutants.

3.4.1.6 Cuticle synthesis in stc-1(ij15), let-607(h402) and let-607(h189) mutants

Cuticle synthesis in *stc-1(ij15)* and *let-607* mutants was examined by analysing expression and assembly of specific collagen components of the cuticle. Cuticle collagen gene expression was assessed with two reporter transgenes, *dpy-7::GFP* and *col-12::GFP*. As explained in section 3.3.4, cuticle collagen gene transcription occurs in discrete temporal patterns and is limited to hypodermal cells. *dpy-7* belongs to the first set of transcribed collagen genes and is expressed at comma stage during embryogenesis; *col-12* is expressed late at the three-fold embryo coinciding with secretion of the L1 cuticle. Transgene expression was observed in the disrupted body of *stc-1(ij15)* and *let-607(h402)* mutant embryos (Fig. 3.8). The study of *dpy-7::GFP* and *col-12::GFP* expression was performed by C. Clucas (Glasgow University, Glasgow, UK), as part of the genetic screen by which the mutant allele *stc-1(ij15)* was isolated.

Immunofluorescence experiments were performed with the DPY7-5a monoclonal antibody (McMahon *et al.*, 2003) to study cuticle secretion and the DPY-7 pattern in the above mutants

Figure 3.7: Hypodermal cell boundaries of wild-type and *let-607* animals.

A, lateroventral view of a wild-type L1 larva. B, lateral and lateroventral views of *let-607(h402)* hatchlings.

C, lateral view of a *let-607(h189)* larva. Arrowheads point to seam cells and arrows to P cells.

Facing page 76
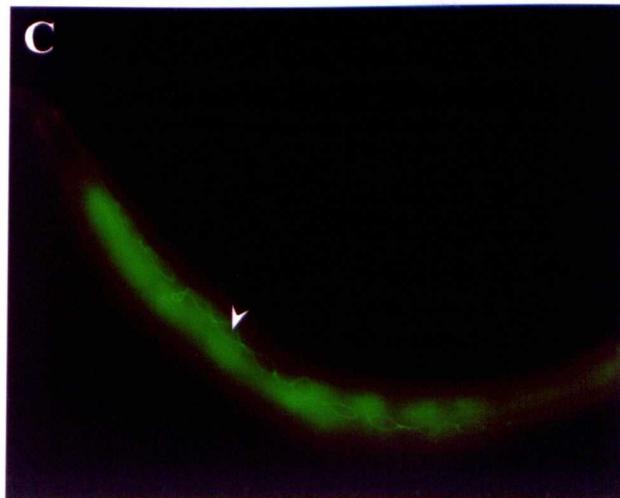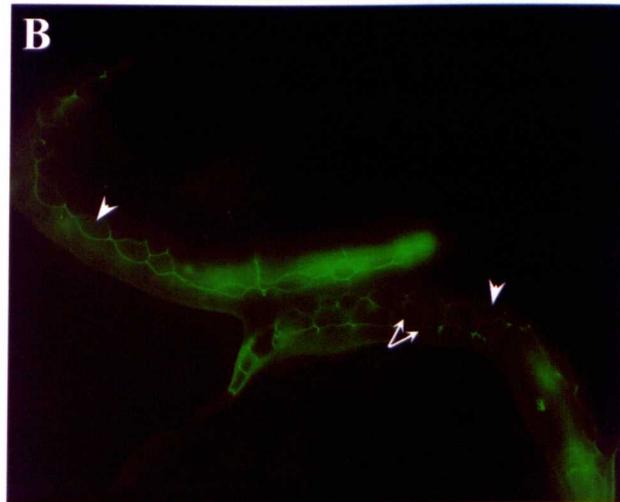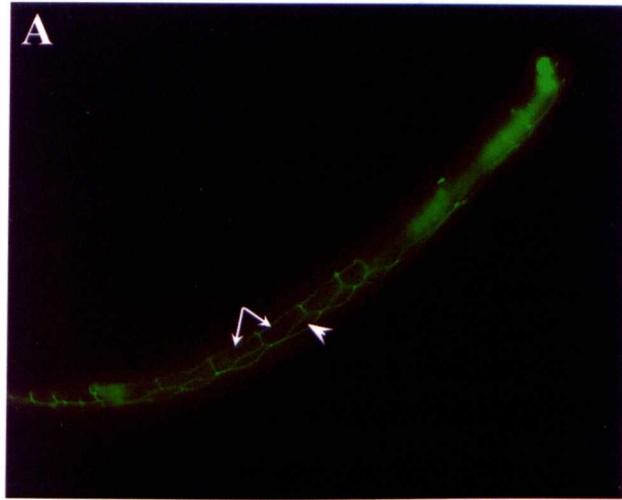
Figure 3.8: Collagen gene expression in *stc-1(ij15)* and *let-607(h402)* embryos.

A-B, wild-type embryo at three-fold stage expressing *dpy-7::GFP* or *col-12::GFP* transgene respectively.

C, *dpy-7::GFP* expression in a *stc-1(ij15)* mutant embryo. **D-E,** *let-607(h402)* mutant embryos expressing *dpy-7::GFP* or *col-12:GFP* respectively.

Facing page 77
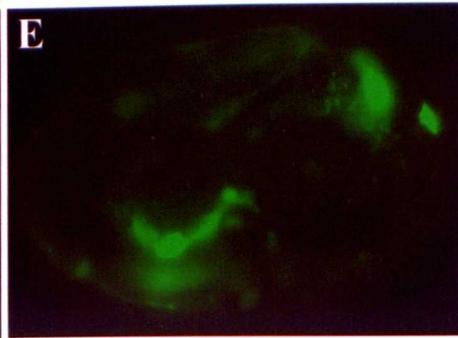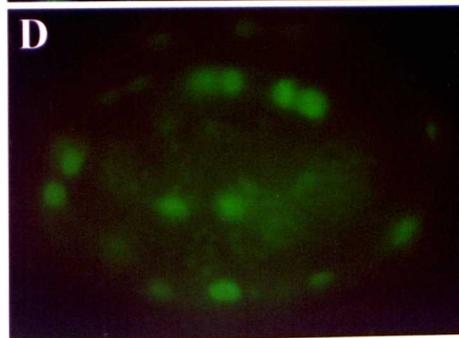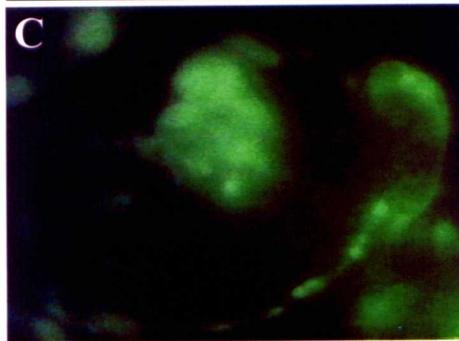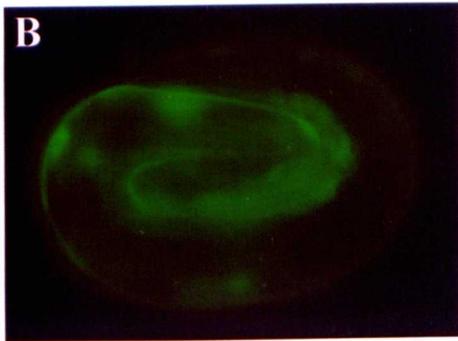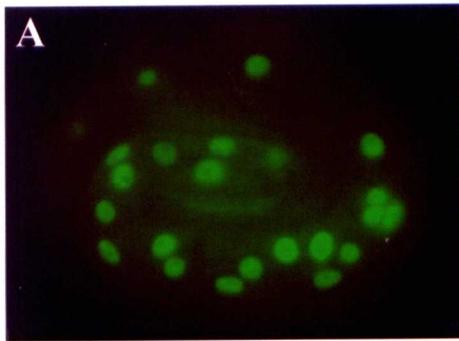
Figure 3.9: DPY-7 secretion in *stc-1(ij15)* and *let-607* mutants.

Left panels illustrate DPY-7 pattern on a three-fold wild-type embryo (**A**), *stc-1(ij15))* mutant (**C**), and *let-607(h402)* embryo (**E**). Right panels show DPY-7 localisation on an early stage wild-type larva (**B**), class III *let-607(h402)* mutant (**D**), and *let-607(h189)* larva (**F**).

Facing page 78
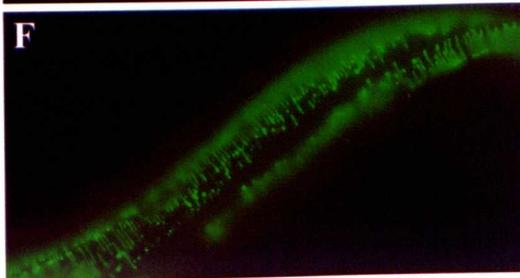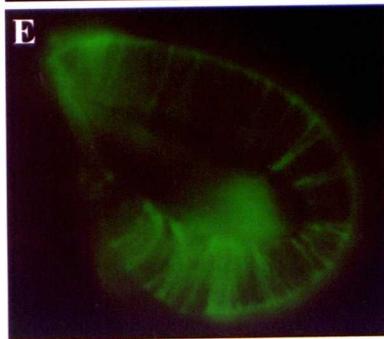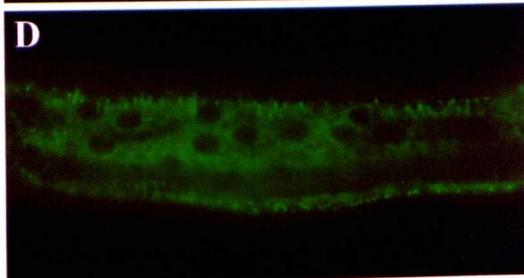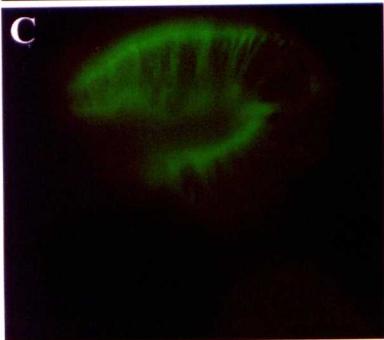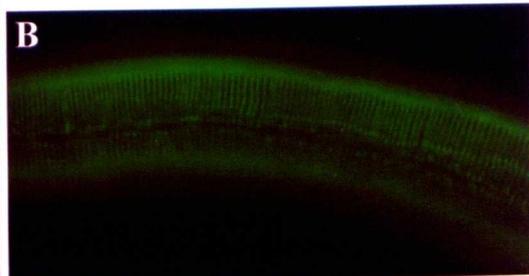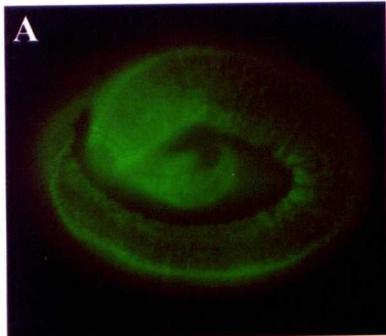
(Fig. 3.9). The DPY-7 protein is detected within the cuticle in circumferential bands or stripes (Fig. 3.9A and B), which are continuous except over the lateral sides of the animal when the alae are present. These stripes correspond to the evenly spaced indentations or furrows that define the annuli of the cuticle. As can be seen in Fig. 3.9C, *stc-1(ij15)* embryos secrete cuticle to some extent but their DPY-7 pattern is more compressed than that of the wild-type embryo (Fig. 3.9A). Similarly, DPY-7 secretion occurs in *let-607(h402)* class I and class II mutants and its distribution around the embryo is also more compact (Fig. 3.9E). *let-607(h402)* hatchlings (class III mutants) frequently show intracellular accumulation of DPY-7, presumably within the hypodermal cells (Fig. 3.9D). These animals also have absent or discontinuous and non-uniformly separated DPY-7 stripes. A disrupted pattern is also observed on *let-607(h189)* mutants (Fig. 3.9F). The characteristic DPY-7 bands tend to be normal at the anterior part of these mutant larvae but irregular or even absent from mid to end part of their body.

3.4.1.7 Cuticle strength analysis in let-607(h402) and let-607(h189) larvae

Two basic roles of the cuticle are to maintain the animal shape and to confer it protection from the environment. The latter function is achieved in part because of its high impermeability, which allows nematodes to be resistant to an ample range of solutions as well as facilitating osmoregulation. The cuticle integrity of *let-607(h402)* and *let-607(h189)* hatchlings was tested in three different solutions. *let-607(h402)* and *let-607(h189)* worms were separately placed in dH$_2$O, 1% β-ME and 1% SDS. More than half of *let-607(h402)* hatchlings burst and tended to shrink suddenly in dH$_2$O (~53.6%; n=20), all of them ruptured in 1% β-ME (100%; n=15) and some in 1% SDS (~21.4%; n=14). *let-607(h189)* mutants, however, did not burst in dH$_2$O and 1% SDS but significantly ruptured in 1% β-ME (70.8%; n=24). Wild-type L1s maintained their body shape when placed in any of the tested solutions. 1% SDS seemed to progressively dissolve the worms' body contents; this effect was notably faster in mutant larvae. Rupture of *let-607(h402)* hatchlings in the above three solutions and of *let-607(h189)* larvae in 1% β-ME indicated that these mutants are surrounded by weak cuticles.

3.4.2 Summary of results

I have characterised three independent phenotypes resultant from three distinct mutant alleles: *stc-1(ij15)*, *let-607(h402)* and *let-607(h189)*. *h402* and *h189* correspond to the same genetic locus, *let-607*. The main phenotypic traits of each mutant are recapitulated below. *stc-1(ij15)* mutants die at three-fold stage, late embryogenesis, concurrent with cuticle formation. They express specific cuticle collagen genes but fail in assembling a functional cuticle.

More than half of *let-607(h402)* homozygotes die at embryogenesis, about the time of cuticle secretion. They elongate normally but subsequently retract in length, or occasionally collapse and lose the worm shape. The rest of the mutants hatch and present severe morphological and cuticular defects. They remain arrested for several days and eventually die as L1s, few larvae seem to attempt moulting in vain. Despite transcribing specific cuticle collagen genes, *let-607(h402)* animals do not form an intact cuticle. Moreover, intracellular retention of collagen protein is observed in *let-607(h402)* hatchlings.

Homozygous animals for *let-607(h189)* rarely die at embryogenesis. Newly hatched *let-607(h189)* worms appear the same as wild types; they grow to L2 larvae but progressively deteriorate at this stage. Most of these mutants shrink significantly when prodded. Furthermore, the results presented reveal the presence of an aberrant cuticle. It should be noted that the *let-607(h402)* lesion has a more detrimental effect on the worm than *let-607(h189)*.

Chapter 4

stc-1 encodes a HSP70-like protein required for cuticle synthesis in *C. elegans*

4.1   INTRODUCTION

Organisms' protective response to a stress stimulus, such as a sudden increase of temperature, results in the inhibition of the expression of most proteins and synthesis of a small set of proteins, called heat shock proteins (HSPs). Some pre-existing cellular proteins may become partially denatured at elevated temperatures, exposing hydrophobic regions that tend to interact to form aggregates. HSPs recognise and bind to the exposed sequences enriched in hydrophobic amino acids normally through ATP expenditure. This results in stabilisation of the polypeptides in an extended conformation and prevents their aggregation. Folding of the partially denatured proteins is attempted when normal conditions are restored in cells. HSPs are widely distributed in prokaryotic and eukaryotic organisms and have been classified in families on the basis of their structural features. The majority of members of these families, however, are expressed constitutively and have essential roles under non-stress conditions. They are involved in *de novo* protein folding, translocation of proteins across membranes, degradation of malfolded polypeptides, assembly and disassembly of protein oligomers, as well as specialised functions in certain regulatory processes. Because of their task in assisting the non-covalent assembly/disassembly of protein-containing structures without being permanent components of these structures, both constitutively expressed and stress-induced HSPs have been referred to as molecular chaperones. The steric information required for correct folding is not provided by molecular chaperones but is contained in the primary sequence of the translated product. Moreover chaperones do not actively promote assembly, in contrast to a known distinct class of proteins that also intervene in folding of polypeptides, namely folding catalysts. This class of proteins are characterised by catalysing slow rate-limiting isomerisation steps resulting in the acceleration of folding of some proteins. The pathways by which polypeptides attain their native configurations in cells still remain poorly understood. Further studies in this topic are expected to shed light on folding processes as well as to reveal more examples of folding factors.

The endoplasmic reticulum (ER) is a specialised subcellular compartment in which modification and folding of secreted, plasma membrane and certain organelle proteins take place. In this compartment, there are at least ten general chaperones in addition to protein-specific chaperones that present restricted tissue expression (Stevens and Argon, 1999). The following sections deal with some of the resident ER molecular chaperones and enzymes together with their involvement in protein folding.

4.1.1 HSP70

HSP70s constitute a major class of molecular chaperones in prokaryotes and eukaryotes. In eukaryotic cells, HSP70s can be found in different compartments including the cytosol, ER, mitochondria and nucleus. Proteins belonging to this family have an average molecular weight of 70 kDa. They are composed of two domains: a highly conserved N-terminal ATPase domain of ~44 kDa and a less conserved C-terminal domain of ~25 kDa, which consists of a substrate binding region of 15 kDa and a more variable 10 kDa segment (Bukau and Horwich, 1998). Association of the C-terminal peptide domain with substrate is influenced by the ATPase activity of the N-terminal domain, since binding and hydrolysis of ATP causes protein conformational changes that alter the affinity of the C-domain for substrates (Bukau and Horwich, 1998). Therefore, HSP70s function in an ATP-dependent mode that is basically as follows. HSP70 binds substrate in an ATP-bound state, and this chaperone-substrate interaction is then stabilised by ATP hydrolysis. Subsequent release of ADP allows binding of a new ATP molecule, which triggers dissociation of substrate and re-establishes the starting point (Hartl, 1996; Bukau and Horwich, 1998). The ATPase activity, and hence, the chaperone functions of HSP70s is known to be regulated by the action of specific co-chaperone factors. Several studies have indicated that HSP70s have a preference for short, linear polypeptide sequences enriched in hydrophobic amino acids, which are probably present in all unfolded polypeptides (Frydman, 2001). These substrate binding properties permit HSP70s to fulfil various essential roles under both stress and normal conditions, including preventing aggregation and refolding of stressed, denatured proteins; assisting folding of newly translated proteins; aiding translocation of proteins across organellar membranes; disassembling oligomeric protein structures; and degradation of unstable, unfolded proteins (Gething and Sambrook, 1992; Hartl, 1996).

4.1.1.1 BiP

The immunoglobulin heavy chain protein (BiP) is the major ER representative of the HSP70 family, and was originally identified and described independently as the immunoglobulin heavy chain binding protein (Haas and Wabl, 1983) and as the glucose regulated protein, GRP78 (Pouyssegur *et al.*, 1977). It is roughly 78 kDa and is synthesised constitutively and abundantly under normal growth conditions. Its expression can be further induced upon accumulation of secretory precursors, mutant proteins that fail to assemble correctly, or unfolded polypeptides resultant from different stress conditions (Gething and Sambrook, 1992). In addition to the role attributed to all HSP70s in assisting protein folding, BiP is involved in translocation of secretory precursors across the ER membrane (Simons *et al.*, 1995) and in the transport back to the

cytosol (retro-translocation) of aberrant or misfolded polypeptides for degradation by the proteasome (Tsai *et al.*, 2002). BiP is a Ca^{2+} binding protein present as monomeric or oligomeric forms that can be subjected to post-translational modifications: phosphorylation and ADP ribosylation (reviewed by Gething, 1999). Modifications of BiP molecules appear to depend on the levels of unfolded polypeptides in the ER that influence the chaperone activity by affecting its state of aggregation. Increase in the levels of unfolded polypeptides results in a reduction of modification and of BiP oligomers; only unmodified, monomeric BiP molecules have been found associated with unfolded polypeptides. BiP chaperones the folding of a wide range of non-native polypeptides. This general function stems from its ability to interact with heptameric stretches of extensive diverse sequence, but usually enriched in hydrophobic residues (Flynn *et al.*, 1991). Thus, potential binding sites for BiP are expected to be common in nascent polypeptides.

4.1.2 HSP90

Proteins of the HSP90 family have a molecular weight of approximately 90 kDa and share a common domain structure. They are composed of well conserved N-terminal and C-terminal regions separated by a charged segment variable in length (reviewed in Buchner, 1999; Young *et al.*, 2001). The cytosolic representative of this family in eukaryotic cells, named HSP90, is highly abundant under normal growth conditions but its expression can be up-regulated upon stress. HSP90 functions as a homodimer in an ATP-dependent manner, and it has been shown to be essential in yeast and the fruitfly (Borkovich *et al.*, 1989; Cutforth and Rubin, 1994). The HSP90 dimerisation site resides in the last region of the C-domain, whereas its ATP-binding site is in the N-domain. Both N- and C-terminal domains are involved in binding of substrate polypeptides. The substrate binding ATPase cycle of HSP90 consists basically in association of a HSP90 dimer with its substrate polypeptide followed by interaction with ATP, hydrolysis of bound ATP releases the polypeptide by inducing a conformational change in the HSP90 dimer (Young *et al.*, 2001). Recently, a second ATP-binding site located on the C-terminus has been reported (Marcu *et al.*, 2000; Garnier *et al.*, 2002), however its involvement in the chaperone activity is unknown. HSP90 has been demonstrated to fulfil general and specialised chaperone roles (reviewed in Mayer and Bukau, 1999; Young *et al.*, 2001). The general activities of HSP90 have been shown by *in vitro* assays, where the chaperone suppresses the aggregation of and assists with the refolding of stressed, denatured proteins; there is as yet no direct evidence of a general implication of HSP90 in the folding of newly synthesised proteins *in vivo*. The specialised activities of HSP90 link this chaperone to various cellular processes including signal transduction, protein degradation and morphological evolution. There are several well

characterised examples of signalling pathways in which HSP90 is a key player (Buchner, 1999; Mayer and Bukau, 1999). In these, the chaperone in question has been found to participate in the maturation and conformational regulation of signalling molecules such as steroid hormone receptors or kinases. The implication of HSP90 in molecular evolution has been reported recently in animals and plants by Rutherford and Lindquist (1998) and Queitsch *et al.*, 2002. Interestingly, HSP90 appears to stabilise active wild-type-like conformations of mutant proteins that control morphogenesis, such as regulators of growth and development, as a consequence of its chaperone properties. Under conditions where the "buffering" capacity of HSP90 is not sufficient (e.g. stress), expression of the mutant products may occur resulting in a phenotypic diversity on which natural selection and ultimately evolution acts. Although the above mechanisms of action of HSP90 are known, the precise features of unfolded proteins recognised by this chaperone are still poorly understood and therefore a better understanding of how HSP90 aids in folding cannot yet be achieved.

4.1.2.1 GRP94

GRP94 (also known as endoplasmin, ERp99, HSP108 or gp96) is the ER member of the HSP90 family in vertebrates and in some invertebrates like *C. elegans* and plants. In fact, GRP94 orthologues constitute their own distinct subfamily within the HSP90 class. Although GRP94 is normally located in the ER lumen, secreted forms of this chaperone have also been detected (e.g. Peter *et al.*, 1992; Bruneau *et al.*, 1998). GRP94 is constitutively present in all cell types, but its expression can be further induced by certain stressful conditions (Kozutsumi *et al.*, 1988). It is a low-affinity, high-capacity Ca^{2+} binding protein, and association with Ca^{2+} appears to cause a change in the conformation of the chaperone. On the other hand, GRP94 is phosphorylated on Ser and Thr residues. Ca^{2+} binding and phosphorylation may regulate the activity of GRP94 (reviewed in Argon and Simen, 1999). Like HSP90, GRP94 is active as a homodimer, but the role of ATP in modulating its function is still unclear. Only a small set of proteins have been identified as targets of GRP94 thus far (Argon and Simen, 1999); the lack of common structural features among the known targets suggests that GRP94 is a specific chaperone. Not much is known about the physiological roles and the modes of action of GRP94 in the cell, except that it binds its substrates in a late stage of folding as shown for HSP90 (Melnick *et al.*, 1994).

4.1.3 Calnexin and calreticulin

Calnexin and calreticulin represent a new lectin subfamily with characteristic chaperone functions in the ER (Williams, 1995b; Michalak *et al.*, 1999). Lectins are defined as sugar binding proteins that generally have no catalytic activity towards carbohydrates. Calnexin is a 64 kDa type-I ER transmembrane protein whereas calreticulin is a 46 kDa soluble protein in the lumen of the ER, although it can also be found in other locations including the nucleus, cytoplasm and surface of certain cells. The luminal domain of calnexin shares extensive sequence homology with calreticulin. The three-dimensional structure of calnexin (Schrag *et al.*, 2001), and presumably that of calreticulin (Leach *et al.*, 2002), consists of two distinct domains: a globular domain and an extended arm composed of two proline-rich motifs repeated in tandem. Both lectins have been attributed multiple binding functions (Leach *et al.*, 2002). They bind to monoglucosylated oligosaccharides, interact with the thiol oxidoreductase ERp57 and, albeit still controversial, are capable of recognising and binding to non-native polypeptides. The primary recognition site of carbohydrates resides within the globular domain, and a much weaker secondary site in the arm domain. Two sites have been defined for the interaction with ERp57. One is a Zn^{2+}-independent site predominantly located in the repeated motifs of the arm domain; the globular domain appears to be involved in this interaction as well. The other is a Zn^{2+}-dependent site localised in the N-terminal region of the globular domain. The polypeptide-binding site resides primarily within the globular domain although the arm domain is also required. Moreover, calnexin and calreticulin associate with Ca^{2+} with high capacity and low affinity through their arm domains.

The majority of eukaryotic proteins that enter the secretory pathway are N-glycosylated during translocation through the ER membrane. N-glycosylation consists of the transfer of a pre-assembled triglucosylated oligosaccharide to an Asn residue within the sequence Asn-X-Ser/Thr, where X may be any amino acid except Pro. This sequence is required for glycosylation but is not always glycosylated. Immediately after the addition of the oligosaccharide to the protein, the first glucose is removed by glucosidase I, the second and third glucoses by glucosidase II. The third glucose can be reattached by UDP-glucose:glycoprotein glucosyltransferase (UGGT). Several lines of evidence indicate that both calnexin and calreticulin bind specifically to monoglucosylated oligosaccharides (reviewed by Parodi, 2000; Trombetta and Parodi, 2002). Removal of the glucose residue from the oligosaccharide by glucosidase II releases the glycoprotein from the lectins. Glycans can then be reglucosylated by UGGT, which recognises the folding status of glycoproteins as it can discriminate between their native and non-native conformations (Trombetta and Helenius,

1998). UGGT adds a single glucose to deglucosylated oligosaccharides only when linked to unfolded glycoproteins, which can be recognised again by the lectins. When bound to glycoproteins, calnexin and calreticulin appear to recruit ERp57 for disulfide bond formation in the folding substrates (Oliver *et al.*, 1999). Unfolded glycoproteins are predicted to undergo several cycles of binding to and release from these lectins until they acquire their final structure. In addition to recognising and binding monoglucosylated glycans, calnexin and calreticulin have also been suggested to act as classical molecular chaperones by interacting directly with unfolded proteins (Ware *et al.*, 1995; Ihara *et al.*, 1999). Although there are no apparent differences in their binding properties, both chaperones tend to associate with distinct protein species. This substrate preference may account for the membrane associated and soluble status of these lectins in the ER, interacting only with glycoproteins in their proximity. Williams and co-workers (Ihara *et al.*, 1999; Saito *et al.*, 1999) have reported that the presence of ATP enhances the function of both lectins. Expression of calnexin and calreticulin has been shown to increase in response to a variety of stresses (Williams, 1995b; Michalak *et al.*, 1999).

Besides assisting with the folding of glycoproteins, calreticulin is implicated in other cellular functions including regulation of Ca^{2+} homeostasis (reviewed by Michalak *et al.*, 1999). The Ca^{2+} binding and lectin properties of calreticulin permit this protein to modulate storage and transport of Ca^{2+} across the ER membrane. The latter has an effect on the levels of cytosolic Ca^{2+} and, consequently, on the regulation of diverse cell functions since Ca^{2+} is an important intracellular signalling molecule (Michalak *et al.*, 1999 and references cited therein). The *C. elegans* calreticulin gene, *crt-1*, has been functionally characterised recently (Park *et al.*, 2001). CRT-1 suppresses protein aggregation *in vitro* and may have important chaperone roles under stress conditions. In addition, it is required for proper reproduction in the worm, possibly influencing Ca^{2+} signalling in late oocyte and sperm development. The function of *C. elegans* calnexin, *cnx-1*, is less well understood. C.D. Johnson and co-workers (1999, personal communication) have determined that CNX-1 is required for larval development and suggest that it may be essential for nematode viability and in promoting cell death.

4.1.4 PDI

PDI is a multifunctional enzyme found in almost all tissues in mammals, and is involved in the modification, folding and assembly of a wide range of proteins of the secretory pathway. In addition to the ER, PDI can be present in other locations including the cytosol, endosomes and cell surface. Here I will focus on the characterised roles of this enzyme in the ER. As a protein-thiol oxidoreductase, PDI catalyses the formation and rearrangement of disulphide bridges on

nascent polypeptides accelerating their folding (Noiva, 1999). PDI has also been ascribed chaperone and anti-chaperone activities depending on its concentration (Gilbert, 1997). At high concentrations, it behaves as a *bona fide* chaperone since it inhibits the aggregation of both disulphide forming and non-disulphide forming proteins; when present at lower concentrations, it acts as an anti-chaperone facilitating the formation of aggregates. The purpose of these contrary properties is not known, however it has been suggested they may be useful in stressed cells to retain unfolded proteins in the ER and protect them from degradation until conditions revert to normal (Gilbert, 1997). In addition, PDI is the β-subunit of the enzymes P4H and microsomal triglyceride transfer protein (MTP). P4Hs catalyse the co- and post-translational hydroxylation of nascent procollagen chains, and have been characterised as $\alpha_2\beta_2$ tetramers in vertebrates and as mixed tetramers as well as $\alpha\beta$ dimers in *C. elegans* (Myllyharju *et al.*, 2002; Myllyharju, 2003). The P4H activity resides on the α-subunit, which also contains the peptide substrate-binding site (Myllyharju and Kivirikko, 1999). The β-subunit, or PDI, functions as a molecular chaperone preventing the aggregation and thus maintaining the active form of the α-subunits. PDI is also responsible for the location of P4H in the lumen of the ER since it possesses an ER-retention signal in its carboxyl terminus.

MTP is a heterodimer ($\alpha\beta$) enzyme essential for the assembly of specific proteins in hepatocytes and intestinal cells (Noiva, 1999). The function of PDI as a component of this enzyme also appears to be to keep the catalytic α-subunit in an active, non-aggregated form (Wetterau *et al.*, 1991).

The roles of PDI as a molecular chaperone and as a component of the aforementioned enzyme complexes are independent to its redox activity. PDI functions independently as a homodimer and, like other chaperones and foldases, its expression is increased upon stress conditions such as heat shock. It has been proposed that Ca^{2+} concentration may modulate the activity of PDI; this enzyme is an important Ca^{2+} storage protein in the ER (Noiva, 1999). Binding to Ca^{2+} by PDI appears to induce conformational changes which affect the enzyme interactions with substrates (Lucero and Kaminer, 1999). Moreover, the enzyme anti-chaperone properties are favoured by the physiological concentrations of Ca^{2+} present in the lumen of the ER (Primm *et al.*, 1996).

The involvement of PDI in collagen biosynthesis has been the subject of intense study *in vitro* and *in vivo* systems. It is now well known that this multifunctional protein intervenes in the folding of collagens at various stages, commented briefly hereafter. PDI catalyses the formation

of both intra- (Bulleid and Freeman, 1988) and inter-chain (Koivu and Myllylä, 1987) disulphide bonds. Intra-chain bonds are formed within the N- and C-propeptides of individual procollagen monomers; inter-chain bonds link the three chains composing a procollagen molecule. PDI interacts with the C-propeptides of procollagen monomers and thereby prevents the aggregation of such monomeric chains prior to trimerisation (Wilson *et al.*, 1998). Assembly of the triple helix begins with the association of the procollagen chains at their C-domains, followed by the formation of a nucleus of triple helix and completed by the propagation of the triple helical structure from the C-terminus to the N-terminus (Engel and Prockop, 1991). Bulleid and co-workers (Wilson *et al.*, 1998) suggest that PDI bound to monomeric C-propeptides could also function in coordinating the formation of procollagen trimers. Being the β-subunit of P4H, PDI also participates in the hydroxylation of procollagen chains and, as demonstrated by N. Bulleid's group (Wilson *et al.*, 1998; Walmsley *et al.*, 1999), it prevents premature association of procollagen polypeptides and retains misfolded trimers in the ER through binding to the triple helical domains. P4H disassociates from the triple helical domains upon complete assembly of the trimers. Moreover, PDI may preclude premature assembly of collagen X molecules in the ER as has been suggested by its interaction with the fully folded, triple helices of this collagen (McLaughlin and Bulleid, 1998). Collagen X molecules form higher order structures within the extracellular matrix without cleavage of their C- and N-propeptides. Thus, these molecules may have the ability to associate once they attain the triple helical form in the ER and PDI may be part of a mechanism to prevent this self-assembly within the cell. Dissociation of PDI from triple helical collagen X is pH-dependent and occurs at a pH about 6.0, likely in the early Golgi compartment.

There are two isoforms of PDI in the *C. elegans* genome, PDI-1 and PDI-2. PDI-1 has been proposed to intervene in the folding of cuticle proteins (Page, 1997), however its mode of function has not yet been determined. PDI-2 is the isoform found in P4H complexes. Recently, the activity of P4H has been shown to be essential for the production of an intact cuticle as well as for the development of the nematode (Friedman *et al.*, 2000; Hill *et al.*, 2000a; Winter and Page, 2000).

4.1.5 HSP47

HSP47 (also colligin or gp46) is a heat inducible, collagen binding glycoprotein with a molecular weight of 47 kDa (reviewed by Nagata, 1996; 1998) . This protein has been found only in vertebrates and belongs to the serpin (serine protease inhibitor) family, which includes α1-antitrypsin, α1-antichymotrysin, antithrombin III and ovalbumin. In contrast to most of the

serpin proteins, HSP47 resides in the ER and has no apparent protease inhibitor activity. The expression of HSP47 occurs exclusively in collagen secreting cells and parallels that of various collagens under non-stressed conditions. This correlated expression has also been detected in pathologies such as fibrosis (Masuda *et al.*, 1994), caused by an overproduction of collagens. Upon heat shock stress, however, collagen synthesis is normally decreased whereas HSP47 synthesis is dramatically enhanced (Hirayoshi *et al.*, 1991). The products of HSP47 present at normal or heat shock conditions contain distinct 5' regions as a result of alternative splicing (Takechi *et al.*, 1994). HSP47 has been shown to interact with types I to V collagens *in vitro* (Natsume *et al.*, 1994); its association with types I and IV has been confirmed by various *in vivo* analyses (e.g. Nagai *et al.*, 2000; Tasab *et al.*, 2000). Several lines of evidence have led to the suggestion that HSP47 interacts with procollagen at various stages during its maturation (Nagata, 1996). This interaction is pH sensitive, with dissociation of HSP47 from procollagens at a pH close to 6.3, probably at the *cis*-Golgi compartment. The precise role of HSP47 in the ER is unclear, therefore a range of possible functions are still associated with this protein-specific chaperone. As summarised in Nagata (1998), HSP47 could chaperone the folding and assembly of procollagen chains, prevent bundle formation of procollagen molecules, retain unfolded molecules in the ER, and/or falicitate transport of correctly folded molecules from the ER to Golgi. Bulleid and co-workers (Tasab *et al.*, 2000) have demonstrated recently that HSP47 binds with highest affinity to triple helical procollagen molecules. There has to be at least one Gly-X-Arg triplet in the collagen domain for this interaction to occur (Tasab *et al.*, 2002). They propose that the function of HSP47 in the ER might be to prevent lateral aggregation of procollagen molecules. Disruption of the *hsp47* gene in mice is lethal resulting in severe reduction of the mature processed form of collagen in several tissues (Nagai *et al.*, 2000). The phenotype of these knockout mice indicates that the role of HSP47 is critical in collagen biosynthesis and for the development of mice.

4.1.6 Integrated function of molecular chaperones in the ER

The ER resident molecular chaperones differ in their functional characteristics and assist folding of secretory proteins at different steps. Thus, a specific protein may associate with several distinct molecular chaperones during its transit through the ER until it attains its native conformation. The integrated function of the ER chaperones is illustrated by a number of studies. Melnick *et al.*, (1994) have reported the sequential interaction of BiP and GRP94 with immunoglobulin chains; Bass *et al.*, (1998) have determined that both calnexin and calreticulin together with BiP mediate the folding of insulin receptor monomers; Linnik and Herscovitz (1998) have shown that BiP, calreticulin and GRP94 intervene in the folding of apolipoprotein

B; finally, Ferreira *et al.,* (1994; 1996) have demonstrated that Hsp47, BiP and Grp94 associate with folding procollagens I and IV, in addition to other works (e.g. Wilson *et al.,* 1998) implicating PDI in the folding and assembly of fibrillar procollagens. Apart from this experimental data, the notion of chaperone co-operation in the ER is more than tempting because of its analogy to the better defined and understood, integrated action of chaperones in the cytosol of eukaryotic cells and in prokaryotes (reviewed in Scheibel and Buchner, 1998; Agashe and Hartl, 2000; Frydman, 2001).

This chapter contains mapping and cloning analyses of the *C. elegans* mutant allele *ij15. ij15* was found to define mutationally the gene *stc-1*, which is predicted to encode a novel HSP70-like protein. Possible functional implications of the previously uncharacterised *stc-1* product are discussed here.

4.2   RESULTS

4.2.1   Mapping of the mutant allele *ij15*

As with any other new mutation obtained by a forward genetic approach, *ij15* had to be first assigned to a chromosome and then to a chromosomal subregion. By means of sequence-tagged sites (STS)-mapping, the mutant allele was assigned to chromosome II, close to either STP36 (0.16 m.u.) or STP101 (-4.52 m.u.) polymorphic markers. Linkage to chromosome II was subsequently corroborated using the morphological marker *bli-2*. Having this preliminary positional data (obtained by C. Clucas, Glasgow University, Glasgow, UK), I carried out a series of crosses with visible genetic markers and deficiencies on this chromosome as described below.

4.2.1.1 Two-factor mapping

Two-factor self-crosses were performed using *ij15 dpy-10(sc48)/+ +* animals, permitting mapping of *ij15* relative to the *dpy-10* gene, which is the reference marker on chromosome II positioned at 0 m.u. The mutant allele *dpy-10(sc48)* causes a viable DpyRol phenotype. The above *cis* double heterozygotes were obtained following a series of crosses and isolation steps. Homozygous *dpy-10(sc48)/dpy-10(sc48)* hermaphrodites (strain IA337) were mated with wild-type males, and their F_1 males with *ij15/+* hermaphrodites. The outcrossed hermaphrodites were then picked clonally to identify *ij15/dpy-10(sc48)* animals by isolation of their DpyRol segregants. Broods of 100 DpyRol worms were examined for 25% embryonic death to obtain *ij15 dpy-10(sc48)/dpy-10(sc48)* hermaphrodites. Crossing of this line with wild-type males

yielded the required double heterozygous mutant for the mapping. The complete self-progeny of nine of these hermaphrodites were scored as represented in Table 4.1. Taking into account that one type of the resulting progeny was masked by the *ij15* lethal phenotype, the recombination distance between the mutant allele and *dpy-10* was calculated with the formula:

$$\frac{A}{A+W} = \frac{2p\text{-}p^2}{3} \quad ; \quad \begin{array}{l} \text{where } A \text{ is the total number of DpyRol progeny, } W \text{ that of} \\ \text{wild-type offspring and } p \text{ the recombination frequency} \end{array}$$

p = 0.00852, estimating *ij15* to be 0.852 m.u. from *dpy-10*.

Individual	F₁ phenotypes			Total brood
	embryonic death	DpyRol recombinants	wild-type	
1	68	3	205	275
2	76	2	219	297
3	81	0	247	328
4	70	0	214	284
5	75	0	223	298
6	76	2	251	329
7	89	2	258	348
8	90	2	277	369
9	76	1	216	293
TOTAL	701	12	2110	2823

Table 4.1: Progeny number for each of the three classes of F₁ phenotypes and total brood of nine self-crossed *ij15 dpy-10(sc48)/+ +* individuals. *ij15 dpy-10(sc48)/+ +* worms segregate wild types, dead embryos, and occasionally, DpyRol recombinants.

4.2.1.2 Multi-factor mapping

Multi-factor mapping of *ij15* entailed the construction and selfing of heterozygous animals carrying the mutant allele in *trans* to two or three genetic markers. Occasional F₁ recombinants were picked out clonally, and their broods were scored for the presence or absence of *ij15* (indicated by 25% dead embryos). I placed *ij15* relative to the markers used in each analysis according to the classes and proportions of recombinants obtained (Figure 4.1). The genotype of the *trans* heterozygotes constructed, the type and numbers of recombinants isolated together
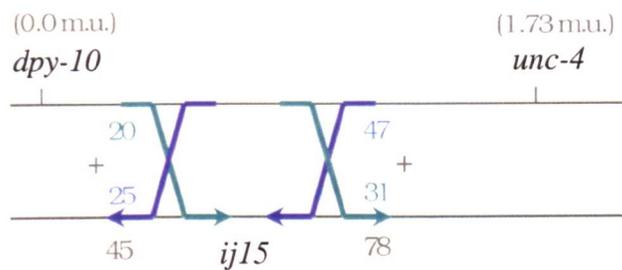
Figure 4.1: Scheme of chromosome II of heterozygous hermaphrodites used to order *ij15* relative to the *dpy-10* and *unc-4* gene markers, as an example of the multi-factor experiments performed. The self-progeny of these animals consisted mostly of wild types, Dpy Unc and *ij15* dead embryos. Dpy non-Unc and Unc non-Dpy animals resulted from rare recombinational events (depicted with coloured lines) between *dpy-10* and *unc-4*. Green lines represent DNA recombination that would generate Dpy non-Unc animals, while purple lines depict DNA recombination that would result in Unc non-Dpy worms. 51 (20 + 31) Dpy non-Unc recombinants and 72 (25 + 47) Unc non-Dpy recombinants were isolated, and their genotype determined by progeny testing. The *ij15* mutation was present in 20 of the 51 isolated Dpy non-Unc, and in 47 of the 72 detected Unc non-Dpy. Therefore, *ij15* is much closer to *dpy-10* than to *unc-4*. In addition to this, *ij15* was placed at approx. 0.63 m.u. on chromosome II as follows:

$$\frac{(45 \times 1.73)}{123}$$

where 45 is the number of recombinants not carrying *ij15*, 1.73 the distance in m.u. between *dpy-10* and *unc-4*, and 123 the total of recombinants (45 + 78)

with the assessed position of *ij15* in each self-cross are indicated in Table 4.2. A list of the genetic markers employed, their locations on the genetic map and the strains containing these markers can be found in Table 2.12.

The genotype of the recombinants was assessed from observations of their phenotype and the phenotypes of their broods. It should be noted that not all possible classes of recombinants could be detected in those three-factor crosses with marker mutations resulting either in lethality or a phenotype difficult to discern.

4.2.1.3   unc-105(n490)/ij15 unc-4(e120) cross

unc-105(n490) II; lon-2(e678) let-2(n821) X animals (strain MT1679) were outcrossed twice to isolate the dominant mutation *unc-105*, which lies at approximately 0.679 m.u. on chromosome II (Table 2.12). Hermaphrodites containing one copy of *unc-105* were mated with *ij15 unc-4/+ +* males to obtain the *unc-105/ij15 unc-4* heterozygotes required to order *ij15* relative to *unc-105*. These animals were allowed to self and their progeny inspected for wild-type or Unc-4 recombinants. The presence of a few F$_1$ wild types suggested that *ij15* was to the right of *unc-105*.

Mapping experiment	Genotype of self-cross heterozygotes	Recombinant			estimated *ij15* position
		Class	carrying *ij15* mutation	not carrying *ij15* mutation	
1	*lin-31(n301) clr-1(e1745) dpy-10(e128)/ij15*	Lin non-Dpy	17	0	on the right of *lin-31*
		Dpy non-Lin	difficult to determine		
		Clr	difficult to determine		
2	*dpy-10(e128) unc-4(e120) / ij15*	Dpy non-Unc	20	31	~0.63 m.u.
		Unc non-Dpy	47	25	
3	*lin-26(mc15) unc-4(e120) / ij15*	Lin non-Unc	NA	NA	~0.7 m.u.
		Unc non-Lin	54	11	
4	*spe-1(mn47) unc-4(e120) / ij15*	Spe non-Unc	NA	NA	~0.65 m.u.
		Unc non-Spe	35	9	
5	*fer-15(b26) unc-4(e120) / ij15*	Fer,non-Unc	ND	ND	close or to the left of *fer-15*
		Unc,non-Fer	24	0	

| Mapping experiment | Genotype of self-cross heterozygotes | Recombinant | | | estimated *lj15* position |
		Class	carrying *lj15* mutation	not carrying *lj15* mutation	
6	*rol-6(su1006) unc-4(e120) / ij15*	Rol non-Unc	0	11	near or to the left of *rol-6*
		Unc non-Rol	22	0	
7	*let-23(mn23) unc-4(e120) / ij15*	Let non-Unc	NA	NA	close or to the left of *let-23*
		Unc non-Let	55	0	
8	*zyg-11(mn40) unc-4(e120) / ij15*	Zyg non-Unc	NA	NA	~0.86 m.u.
		Unc non-Zyg	23	6	
9	*let-24(mn24) unc-4(e120) / ij15*	Let non-Unc	NA	NA	~0.8 m.u.
		Unc non-Let	82	9	

Table 4.2: Ordering of *lj15* on chromosome II assessed from the number and type of recombinants obtained in standard self-crosses. The map positions of the genetic markers in this table are specified in Table 2.12.

NA = Non-applicable

4.2.1.4 Mapping by means of deficiencies

A set of overlapping deficiencies, *mnDf88*, *mnDf97*, *mnDf108*, *mnDf109*, *mnDf100* and *mnDf91*, together with *mnDf69* were also employed as a means to position *ij15* on the genetic map (Table 2.13). Deficiencies *mnDf100*, *mnDf108*, *mnDf109* cause larval lethality in homozygosis; the remaining deletions result in embryonic death. Heterozygous males for each deficiency were mated with hermaphrodites carrying one copy of *ij15*. The resulting F$_2$ hermaphrodites were then cloned out and their broods scored for lethality: 25% was expected in a proportion of plates if *ij15* was covered by the tested deficiency; plates with 25% or 50% in the case where *ij15* was not contained in the deleted region. Some F$_2$ animals segregated 50% death in all tests. As a whole, the above deletions embrace two segments of chromosome II: one from approximately 0.245 to 0.795 m.u., the other from approximately 0.800 to 1.149 m.u. (Figure 4.2), which are separated by a small region or "gap". The complementation of *ij15* by all the deficiencies defining these segments placed *ij15* in the gap region.

The phenotypes of homozygous embryos for *mnDf69* and for *mnDf88* were examined using Nomarski optics. The progeny of *mnDf69/ij15* and *mnDf88/ij15* animals were observed, and it was determined in both cases that the deficiency embryos displayed defects distinct from those
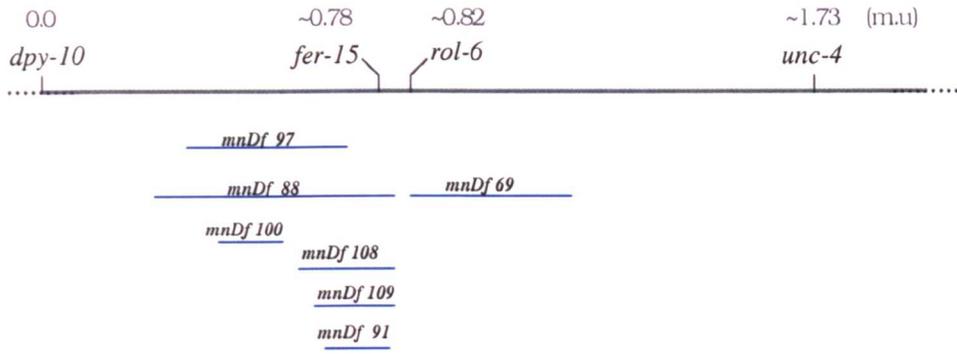
Figure 4.2: Scheme of the deficiencies employed to map *ij15* and their approximate coverage of chromosome II. As a whole, these deletions define two intervals of the map: ~0.24 - ~0.79 m.u. and ~0.8 - ~1.15 m.u., separated by a small gap. The solid grey line represents the 0.0 - 2 m.u. region of chromosome II, where the position of some gene markers is indicated.
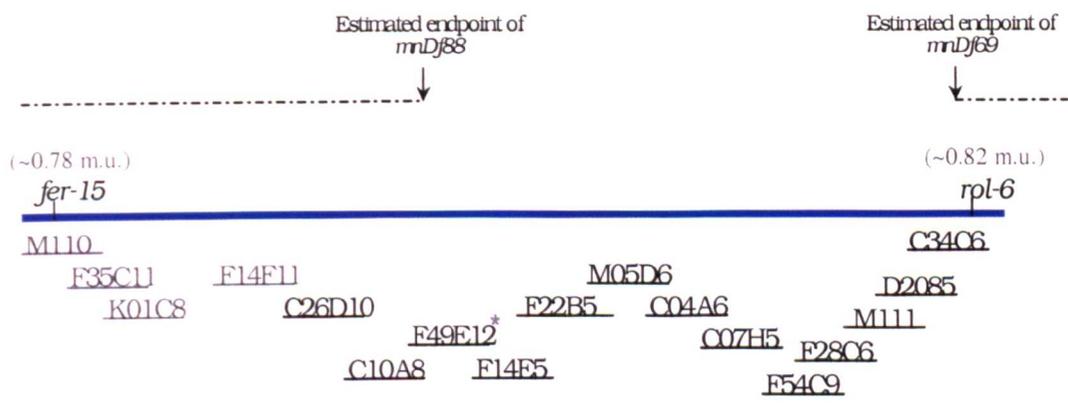
Figure 4.3: Diagram showing the estimated mnDf88 and mnDf69 extents, and the genetic and physical maps corresponding to the region expected to contain the *ij15* mutation. Seventeen overlapping cosmid clones comprise the genomic DNA segment between the genes *fer-15* and *rol-6*. The thirteen cosmids tested to rescue the *ij15* phenotype are coloured in black.

* C54D8 contains essentially the same genomic DNA region as the sequenced F49E12 cosmid, and it was the clone used for DNA transformation experiments.

of *ij15* embryos. This indicated that the parents were indeed *trans* heterozygotes and ruled out the possibility of having misscored percentages of lethality in the analyses.

4.2.1.5 Endpoint deficiency mapping by means of PCR

The endpoints of the above deficiencies defining the gap segment were known only approximately. To determine accurately the extents of these, a series of PCR reactions on mnDf88 and mnDf69 homozygous embryos were carried out with primers designed to amplify a short genomic sequence in the region concerned. Primers pairs were tested first on wild-type N2 gDNA for their compatibility and to find optimal conditions, then used on deficiency embryos and on wild-type N2 embryos as controls. The results indicate that the right end of mnDf88 is within the first ~11.5Kb of the genomic sequence contained in F49E12 cosmid, and that the left end of mnDf69 is between the last ~11Kb of cosmid C34C6 and the first ~12.5Kb of cosmid T01B7 (Fig. 4.3).

4.2.1.6 Estimated map position of ij15

Taken together, the mapping experiments performed gave slightly conflicting results. Data from multi-factor crosses 2, 3 and 4 placed *ij15* from ~0.63 to ~0.7 m.u. (Table 4.2), whereas the other multi-factor crosses and mapping analyses positioned *ij15* in a different interval of chromosome II. Multi-factor crosses 8 and 9 placed *ij15* at ~0.86 m.u. and ~0.8 m.u. respectively; and data from crosses 5, 6 and 7 suggested that *ij15* was close or to the left of *fer-15*, *rol-6* and *let-23* respectively (see Table 4.2). Moreover, *ij15* was mapped to the right of *unc-105* by means of crosses with the *unc-105* marker (section 4.2.1.3), at ~0.852 m.u. from *dpy-10* by two-factor crosses (section 4.2.1.1) and from ~0.795 to ~0.8 m.u. by deficiency mapping (section 4.2.1.4). The results from multi-factor crosses 5-9, crosses with *unc-105* and *dpy-10*, and deficiency mapping were consistent and therefore considered a better estimate of the position of *ij15* on the map as these data were obtained by different approaches. In addition, two-factor crosses and mapping by deficiencies are a more accurate means of mapping than three-factor crosses, which are employed to order mutations relative to known genes. The region from *fer-15* to *rol-6* (both cloned genes) was chosen initially as the most likely to contain *ij15*. This interval of the map is represented by seventeen overlapping cosmids (Fig. 4.3). The physical region was reduced to thirteen clones by molecular delimitation of the deficiencies endpoints, and was therefore the one focused on.

4.2.1.7 Test for allelism with let-252

Of all previous characterised genes available in databases, only *let-252* maps to the gap region between mnDf88 and mnDf69. The following genetic crosses were performed in order to test if *ij15* was allelic to *let-252*. *let-252 unc-4/+ +* heterozygous males (parental strain SP211) were mated with *ij15 unc-4/+ +* heterozygous hermaphrodites, and the generated broods were screened for the presence or absence of Unc-4 animals. The detection of Unc-4 progeny suggested *ij15* is not an allele of the gene *let-252*.

4.2.2 Molecular cloning of *ij15*

Once *ij15* was positioned within a reasonable sized interval, physical and genetic maps were compared to find out the number of clones spanning the segment. Isolation of cosmid clones was rather tedious since the bacterial clones carrying the cosmids tended not to grow easily or at all on the appropriate selective media. Moreover, cosmid clones frequently contained deletions, not bearing in such cases the entire expected genomic sequence. Thus, all cosmid preps were digested with specific restriction enzymes and visualised on agarose gels next to a DNA marker. The resulting patterns were then compared with those expected, predicted by the Vector NTi program, in order to ascertain the integrity of the cosmids DNA. Generally, many DNA preparations from different colonies had to be performed to isolate enough quantities of pure, intact cosmid for DNA transformation experiments.

4.2.2.1 Cloning of the gene identified by ij15

DNA transformation experiments were carried out with hermaphrodites of genotype *ij15 unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444) II*, named IA432. To construct this strain, *let-23(mn23) unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444) II* hermaphrodites (strain name SP378) were mated with N2 males and the F_1 *mnC1 dpy-10(e128) unc-52(e444)/+ +* males with *spe-1(mn47) unc-4(e120)/ij15 unc-4(e120) II* hermaphrodites. F_2 progeny were picked clonally, allowed to self and tested for segregation of 25% embryonic death, 50% wild types and 25% Dpy-10 Unc-52 animals to isolate the desired line. Genomic DNA clones were co-injected with pMW0002, pMW0003 and pTAg plasmids into the gonad of IA432 adult hermaphrodites as detailed in section 2.9. It is known that pools of DNA inserted by this means assemble into extrachromosomal arrays by homologous recombination and may become heritable depending on the array characteristics (Mello *et al.*, 1991). pMW0002 and pMW0003 contain the *dpy-7::GFP* and *col-12::GFP* reporter fusions respectively, and result in the production of GFP protein in the hypodermis. I relied on detection of GFP expression in the transformed animals

under the UV microscope as indicative of the presence of the test DNA on the arrays. The offspring derived from injected animals that contained the markers were transferred onto individual plates to examine their progenies. A different class of F_2 progeny, Unc-4 animals, was expected to arise in those cases that the wild-type copy of a gene in the array rescued the *ij15* phenotype. For each test DNA, several independent transgenic lines were studied to draw solid conclusions.

Initially, injections into the recipient strain of groups of four and five overlapping clones of the above thirteen cosmids were performed (see section 2.9.1.2). Several attempts were carried out using the DNA cosmids at different concentrations and rescue was eventually obtained with the group of five cosmids as shown by six transmitting lines. Clones within this group were individually tested and only the cosmid F54C9 rescued the mutant phentoype (two transgenic lines were obtained).

Taking into account the putative roles of the ten predicted coding sequences in F54C9 (Table 4.3) and the number of cDNAs isolated from each coding sequence, F54C9.4, F54C9.6, F54C9.8, F54C9.9 and F54C9.10 were considered the best candidates to result in the *ij15* phenotype if mutated. RNAi by injection was performed for these genes as described in section 2.11.1. None of them, however, appeared to phenocopy the *ij15* phenotype. At this stage, Julie Ahringer's group had determined the RNAi effects of most of the predicted genes on chromosome II. RNAi embryonic death had only been recorded for the F54C9.5 and F54C9.10 genes (A. Fraser, personal communication). I cloned these genes and tested them by DNA transformation. Both genes failed to rescue the *ij15* phenotype.

An alternative approach was then followed to identify the physical location of *ij15*. F54C9 was cut with specific restriction enzymes (*Spe* I and *Xho* I, *Pst* I and *Sal* I, and *Sma* I) to generate a spectrum of overlapping genomic fragments comprising different coding sequences. The fragments were gel isolated and cloned into appropriate vectors for DNA transformation. Figure 4.4 represents the constructs initially tested; rescue was obtained with the clone containing 9.2 Kb of F54C9 genomic sequence (three transgenic lines). This construct was then digested with *Not* I and *Xho* I, *Not* I and *Spe* I, and *Spe* I and *Sac* I and the resulting fragments tested (Fig. 4.5). The eight transgenic lines carrying the *Spe* I-*Sac* I construct segregated rescued Unc-4 animals. The *Spe* I-*Sac* I segment contains two predicted genes: F54C9.2 and F54C9.3. I did not expect F54C9.3 to be the gene identified by *ij15* since it was included in the *Not* I-*Xho* I fragment that failed rescuing the mutant phenotype. To assure however, that F54C9.2 was the
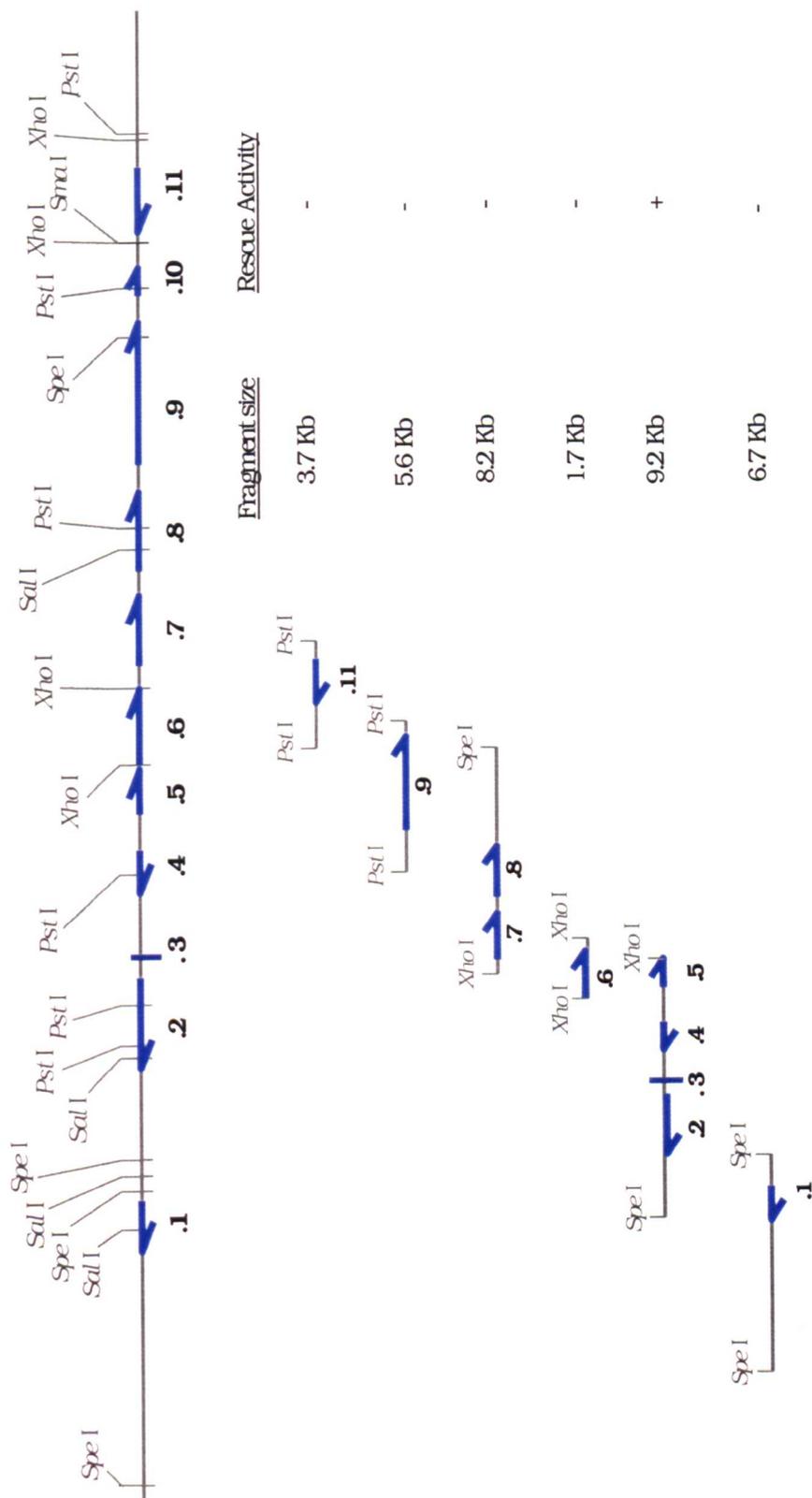
Figure 4.4: Fragments derived from digests of the F54C9 genomic sequence used in *ij15* rescue experiments.
Only the 9.2 Kb DNA region had rescue activity (+); lack of rescue is indicated with "-". The eleven predicted genes in F54C9 (.1 - .11) are represented as blue arrows or blue line. Restriction sites of the enzymes employed (*Sal* I, *Sma* I, *Spe* I, *Pst* I and *Xho* I) are shown in the F54C9 genomic sequence and consequent fragments.

Rescue Activity

-
-
-
-
+
-

Fragment size

3.7 Kb
5.6 Kb
8.2 Kb
1.7 Kb
9.2 Kb
6.7 Kb
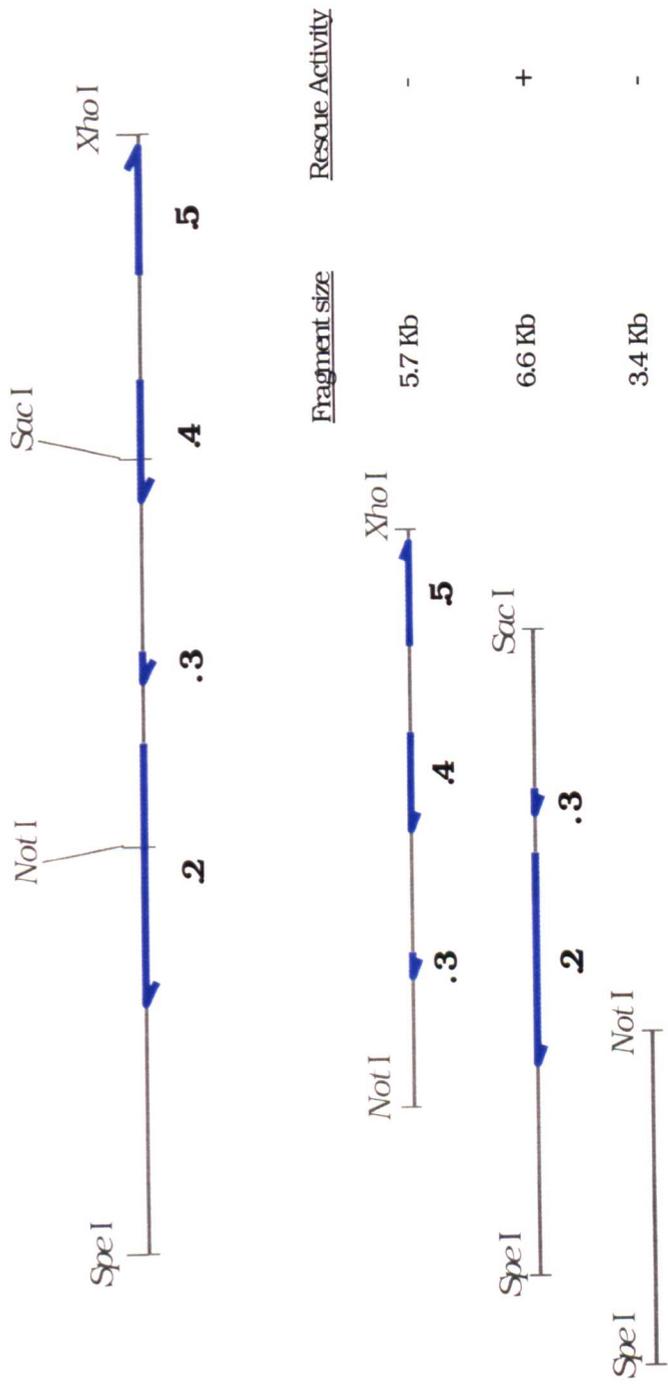
Figure 4.5: Fragments generated from digestions of the 9.2 Kb F54C9 region that rescued the *ij15* phenotype. The 6.6 Kb sequence containing predicted genes .2 and .3 was positive for rescue (+). Failed rescue by the transformed sequences is indicated by "-". Coding sequences (.2 - .5) are shown as blue arrows. Restriction sites of the enzymes used (*Not*I, *Sac*I, *Spe*I and *Xho*I) are represented in the 9.2 Kb sequence and fragments.

coding region responsible for rescue, individual fragments containing F54C9.2 alone, F54C9.3 alone and F54C9.2 plus F54C9.3 combined were cloned and tested by DNA transformation. F54C9.2 was sufficient for rescue of the mutant phenotype (two transgenic lines) and therefore is the gene identified by the *ij15* allele.

The phenotypic rescue observed in the above cases was incomplete for two reasons: not all *ij15 unc-4* homozygotes containing F54C9.2 on the array passed embryogenesis, and those that did so, grew to sterile Unc-4 adults. The development of Unc-4 animals to adulthood showed efficient rescue of the somatic *ij15* function; their sterility, however, might be indicative of failed rescue of the mutation defects in the germline. The multiple-copy content of transgenic free arrays has a negative effect for their expression in the germline (Kelly *et al.*, 1997). This problem can be overcome using the approach developed by Kelly *et al.*, (1997), which consists basically in the injection of linearised test DNA and markers at low concentration, together with short fragments of gDNA at high concentration. The linearised gDNA acts as a complex carrier, minimising the repetitive structure of the test DNA in the arrays and hence improving their function in the germline. By this means, I tested separately F54C9.2 and F54C9.3 using *vc5.3::GFP* transgene as selectable marker. The *vc5.3::GFP* reporter fusion results in expression of GFP protein in intestinal cells and is contained in the construct pMW025. Confirming previous results, the only coding region with rescue activity was F54C9.2. Partial rescue of germline function was obtained; some *ij15 unc-4* animals were sterile and other gave very low brood sizes. However, rescue was sufficient to maintain an *ij15 unc-4/ij15 unc-4* homozygous population (one transgenic line established).

F54C9.2 encodes a member of the stress 70 protein chaperone family, named STC-1, after its human and rat homologous STCH (stress chaperone) proteins (Otterson and Kaye, 1997).

Predicted CDS	Gene name	Transcriptional units
F54C9.1		Initiation factor 5A
F54C9.2		Heat shock 70-kDa protein (HSP70)
F54C9.3		unknown
F54C9.4	*col-38*	collagen
F54C9.5		60S ribosomal protein L5
F54C9.6		ATP-binding protein (CDC48/PAS1/SEC18 family)

Predicted CDS	Gene name	Transcriptional units
F54C9.7		unknown
F54C9.8	*puf-5*	Pumilio-family RNA binding domains
F54C9.9		unknown
F54C9.10	*arl-1*	GTP-binding ADP-ribosylation factor
F54C9.11		unknown

Table 4.3: Predicted genes and proteins in the cosmid F54C9.

CDS = coding sequence. (Data available in wormbase, http://www.wormbase.org).

4.2.3 Detection of the *stc-1(ij15)* lesion

I first sequenced the gene *stc-1* from *ij15* homozygotes in order to detect the genetic alteration causing the mutant phenotype. The predicted structure of *stc-1* consists of 7 exons separated by 6 intronic sequences that span 2418 bp. Two fragments comprising the whole gene were amplified from *ij15* embryonic gDNA by PCR: one from exon 1 to the start of exon 5, the other from the end of exon 4 to exon 7. Two copies of each fragment were obtained, cloned and sequenced. I detected three possible base changes comparing the mutant sequences with that predicted for wild-type *stc-1*: two within the exon1-exon4 section, the third in the exon5-exon7 fragment. Both altered bases within the exon1-exon4 region were in the same clone; the other cloned copy of the fragment was identical to the wild-type sequence. The base change within the exon5-exon7 section was clearly present in one clone. The two sequences of the other cloned copy of the fragment were not consistent: one contained the wild-type base, whereas the other had the altered base. Conclusions from this piece of data could not be drawn since these two sequences were of bad quality in the region containing the discrepancy.

A different approach was employed to test if any of these three base changes were real mutations. The three possible modifications implied the creation of new restriction enzyme sites in the *stc-1(ij15)* allele. I analysed the presence or absence of these restriction sites by digesting PCR amplified DNA derived from several different *ij15* embryos with each specific enzyme. Only the restriction enzyme polymorphism in the exon5-exon7 region (a *Hind* III polymorphism) appeared to be real, thus revealing the nature of the *stc-1(ij15)* lesion. The presence of this polymorphism, and thereby the mutation, was further verified as follows. Single IA432 hermaphrodites, genotype *stc-1(ij15) unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444) II,*

were allowed to self for the isolation of their F_1 progeny: Dpy Unc (*mnC1 dpy-10(e128) unc-52(e444)/mnC1 dpy-10(e128) unc-52(e444) II*), dead embryos (*stc-1(ij15)/stc-1(ij15) II*) and wild types (*stc-1(ij15) unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444) II*). The Dpy Unc animals do not contain the *stc-1(ij15)* allele (+/+; "wild types"), dead embryos have two copies of *stc-1(ij15)* (*ij15/ij15*; homozygotes), and wild types harbour one copy of the *stc-1(ij15)* allele (*ij15/+*; heterozygotes). The region comprising exon 5 to exon 7 of the gene *stc-1* was obtained from two +/+ individuals, four *ij15/ij15* embryos and two *ij15/+* animals by means of single-worm PCR. The DNA from each individual was then digested with the restriction enzyme *Hind* III and visualised by gel electrophoresis (Fig. 4.6). *Hind* III cuts only the exon5-exon7 *stc-1* fragment with the *ij15* lesion; therefore a distinct digestion pattern is seen for "wild-types", *ij15* homozygous and *ij15* heterozygous animals. The fact that just the base change in exon 7 is the true *ij15* mutation, suggests the other two base alterations detected in the exon1-exon4 segment by sequencing may result from either errors in the PCR polymerisation or sequencing steps. It is recalled that these two modifications were present only in one of the two fragment copies cloned.

The mutant allele *ij15* was isolated after mutagenesis with EMS, which normally causes GC to AT transitions (Hodgkin, 1999). Consistent with this, the mutation identified in *stc-1(ij15)* mutants is a GGA to GAA alteration, which is present at the 1980 bp within exon 7 and results in a glycine to glutamate substitution at amino acid 422 (G422E) in the protein sequence (Fig. 4.7). A protein alignment of the *C. elegans* STC-1, human and rat STCH, and human BiP is shown in Figure 4.8. Approximately 39% of the STC-1 amino acid sequence is identical to the mammalian STCH sequences and about 24% to that of the human BiP. As with other HSP70 members, the proteins here in question contain the highly conserved N-terminal ATPase domain distinctive of their family (see section 4.1.1). The glycine mutated in *ij15* animals appears to be a consensus residue in the ATPase domain.

4.2.4 *stc-1* expression

4.2.4.1 *Temporal expression*

RNA samples were taken from a synchronous culture of nematodes at two hour intervals during post-embryonic growth as described in Johnstone and Barry (1996). cDNA was then generated from these RNA samples by means of RT-PCR. To measure transcript abundance throughout the worm development, I employed the above cDNA samples and oligonucleotides specific for *stc-1* and *ama-1* genes. Primers were designed to amplify a short region of distinct size for each
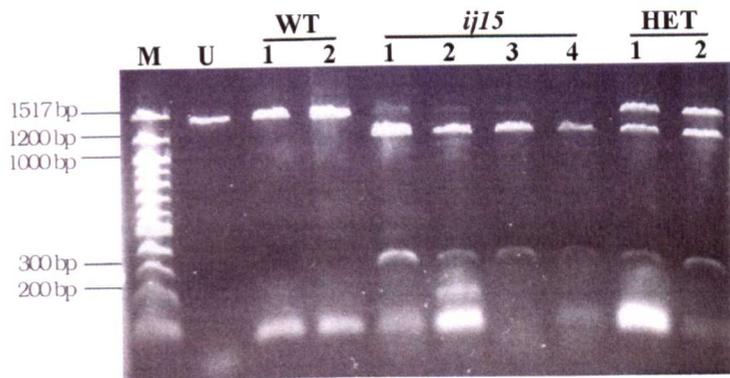
Fig. 4.6: *stc-1(ij15)* polymorphism in *+/+*, *ij15/ij15* and *ij15/+* animals.

A 1386 bp fragment comprising exons 5 to 7 of the gene *stc-1* was amplified separately by PCR from two +/+ individuals, four *ij15/ij15* embryos and two *ij15/+* animals, and then digested with the restriction enzyme *Hind* III. **M:** 100 bp DNA marker; **U:** *ij15/ij15* uncut DNA; **WT 1-2:** digested +/+ DNA; ***ij15* 1-4:** digested *ij15/ij15* DNA; and **HET 1-2:** digested *ij15/+* DNA. The exon5-exon7 region from a wild-type copy of the gene *stc-1* does not contain *Hind* III restriction sites and thus remains intact (1386 bp long) after treatment with *Hind* III (**WT 1** and **2** lanes). The *ij15* mutation is within exon 7 of *stc-1* and creates a *Hind* III site, as determined by the presence of two *Hind* III digested fragments of 1123 bp and 263 bp (***ij15* 1-4** lanes). Three bands are present after digesting exon5-exon7 *stc-1* DNA from *ij15/+* heterozygotes (**HET 1-2** lanes): the 1386 bp band corresponds to the wild-type fragment whereas the 1123 bp and 263 bp bands are the digestion products of the *ij15* fragment. The sizes of the relevant marker bands are indicated on the left of the gel. It should be noted that the PCR amplified DNA was not purified prior to digestion with *Hind* III so the ~100 bp band present in some lanes may correspond to PCR impurities.

Figure 4.7: Genomic organisation of *stc-1* and the point mutation identified in the *ij15* allele. The gene has seven exons (blue boxes) and six introns (grey lines) and is 2418 bp long. The GGA to GAA modification is at 1980 bp within exon 7 and results in a G422E change in the protein sequence.

```
         1                                                                                              100
Ce-STC-1   (1)  MAGATGSLLGNFQRMIRIKLSKKSYFIAAGGIFILLLAGWLGQKYGLPPPPPKIAGIDLGTTYSSIGIYHAVTGETIILPDDLGKKSVPSVVAFLPNGTV
rat STCH   (1)  --------------------MAGEMTILGSAVLTLLLAGYLAQQY-LPLPTPKVIGIDLGTTYCSVGVFFPGTGKVKVIPDENGHISIPSMVSFT-DGDV
human STCH (1)  --------------------MAREMTILGSAVLTLLLAGYLAQQY-LPLPTPKVIGIDLGTTYCSVGVFFPGTGKVKVIPDENGHISIPSMVSFT-DNDV
human BiP  (1)  ----------------------MKLSLVAAMLLLLSAARAEEEDKKEDVGTVVGIDLGTTYSCVGVF--KNGRVEIIANDQGNRITPSYVAFTPEGER
Consensus  (1)  MAKEMTILGSAVLTLLLAGYLAQQY LPLPTPKVIGIDLGTTYSSVGVFFPGTGKVKIIPDDNGHISIPSMVAFTPDGDV

         101                                                                                            200
Ce-STC-1  (101) LIGTRATEQQEHNPKRTIYDAKRFIGRNFEKNNKDFLSDQKRYPFKINLDSEGKAFFEIPLDSGTKNVYPEEIGSLIIGYLKSAAAKHLGVTLGQVISC
rat STCH   (79) YVGYESLEADSNPQNTIYDAKRFIGKIFTP--EELEAEIGRYPFKV-LHKNGMAEFSVTSNE-TIIVSPEYVGSRLLLKLKEMAEKYLGMPVANAVISV
human STCH (79) YVGYESVELADSNPQNTIYDAKRFIGKIFTA--EELEAEIGRYPFKV-LNKNGMVEFSVTSNE-TITVSPEYVGSRLLLKLKEMAEAYLGMPVANAVISV
human BiP  (75) LIGDAAKNQLTSNPENTVFDAKRLIGRTWND--PSVQQDIKFLPFKVVEKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTV
Consensus (101) YIGYEALEQADSNPQNTIYDAKRFIGKIFT    EELEADIKRYPFKV L KNGMAEFSVTSNE TKTVSPEYIGSRLLLKLKEMAEKYLGMPVANAVISV

         201                                                                                            300
Ce-STC-1  (201) PAEFNEKQRNFTAKAAEIAEMEVRRVISEPTAAALAYGLHKKQGVENVVVVDLGGGTLDVSVLWLQGGVFVTQAMAGNNRLGGQDFNDRVQKHLISKIAE
rat STCH  (175) PAEFDLQQRNSTIQAANLAGLKILRVINEPTAAAMAYGLHRVD-VFYVLVIDLGGGTLDVSLLNKQGGMFLTRAMSGNNKLGGQDFNQRLLQYLYKEIYQ
human STCH(175) PAEFDLKQRNSTIEAANLAGLKILRVINEPTAAAMAYGLHKAD-VFHVLVIDLGGGTLDVSLLNKQGGMFLTRAMSGNNKLGGQDFNQRLLQYLYKQIYQ
human BiP (173) PAYFNDAQRQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQRVMEHFIKLYKK
Consensus (201) PAEFNLKQRNSTIDAANIAGLKILRVINEPTAAAMAYGLHKKDGVFNVLVIDLGGGTLDVSLLNKQGGMFLTRAMSGNNKLGGQDFNQRLLQHLYK IYQ

         301                                                                                            400
Ce-STC-1  (301) KFGK-TIDNKEDIQQIRMEVEKGKIRLTNVPSTTISLNL--------------------------------------------KTVGKWNYELTRDEFET
rat STCH  (274) TYGF-LPSRKEEIHRLRQAVEMVKLNLTLHQSAQVSVLLTVEENDSQKPQNADSKLPEDQLTPGDGHVNRVFRPGLSDSTSAKSQVLFETEVSRKLFNT
human STCH(274) TYGF-VPSRKEEIHRLRQAVEMVKLNLTLHQSAQLSVLLTVEEQDRKEPHSSDTELPKDKLSSADDHRVNSGFGRGLSDKKSGESQVLFETEISRKLFDT
human BiP (273) KTGKDVRKDNRAVQKLRREVEKAKRALSSQHQARIEIESF--------------------------------------------YEGEDFSETLTRAKFEE
Consensus (301) TYGK VPSRKEEIQRLRQEVEMVKLNLTLHQSAQISVLLTVEEND   P  ADS LP D LS AD H VN  F  GLSD  SAKSQVLFETELSRKLFET

         401                                                   *                                        500
Ce-STC-1  (356) LNGDLLKAIELPITAALADANLDTADVDEIVLVGGSTQVPAVRKIVGRFF-KRSANYGVDPELAVVTGASVQAGVIGGGWP------LQVAAMELPAKRR
rat STCH  (373) LNEDLFQKILVPIQQVLKEGLLDKTEIDEVVLVGGSTRIPRIRQVIQEFF-GKDPNTSVDPDLAVVTGVAIQAGIDGCWP------LQVSALEIPNKHL
human STCH(373) LNEDLFQKILVPIQQVLKEGHLEKTEIDEVVLVGGSTRIPRIRQVIQEFF-GKDPNTSVDPDLAVVTGVAIQAGIDGGSWP------LQVSALEIPNKHL
human BiP (330) LAMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTGIETV
Consensus (401) LNEDLFQKILVPIQQVLKDG LDKTDIDEIVLVGGSTRIPRIRQVIQEFF GKDPNTSVDPDLAVVTGVAIQAGIDGG WP      LQVSALEIPNKHL

         501                                                                                            600
Ce-STC-1  (449) KRHFFTEEQKENEKDRTISEL-----------------------------------------------------------------------------
rat STCH  (466) QKTNFN--------------------------------------------------------------------------------------------
human STCH(466) QKTNFN--------------------------------------------------------------------------------------------
human BiP (430) GGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITND
Consensus (501) QKTNFN   K      S I

         601                                                                                            700
Ce-STC-1  (470) --------------------------------------------------------------------------------------------------
rat STCH  (472) --------------------------------------------------------------------------------------------------
human STCH(472) --------------------------------------------------------------------------------------------------
human BiP (530) QNRLTPEEIERMVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDFKAKKKELEEIVQP
Consensus (601)

         701              725
Ce-STC-1  (470) -----------------------
rat STCH  (472) -----------------------
human STCH(472) -----------------------
human BiP (630) IISKLYGSAGPPPTGEEDTAEKDEL
Consensus (701)
```

Figure 4.8: Protein alignment of the *C. elegans* (*Ce*-) STC-1, rat STCH, human STCH and human BiP. Amino acids in the alignment are coloured in black if they are not similar, green if they are weakly similar, black on grey if they are similar, blue if they are conserved (present at a given position in two or three of the aligned proteins), and red on grey if they are identical (present at a given position in all the aligned proteins). The amino acid coordinates for each protein are shown in brackets on the left. The glycine mutated in the *stc-1(ij5)* protein product (designated by an asterisk) is within the ATPase domain and is conserved in all the HSP70 proteins aligned above.

gene spanning an intron. The latter was important to distinguish amplified cDNA products from those derived from contaminating gDNA. *ama-1* encodes the largest subunit of RNA polymerase II (Bird and Riddle, 1989) and is expressed at constant levels in all post-embryonic stages. Comparison of *stc-1* and *ama-1* transcripts obtained for each sample showed that the gene *stc-1* is constitutively expressed throughout post-embryonic development.

4.2.4.2 Spatial expression pattern

Two different reporter constructs were made to study the spatial expression of the gene *stc-1*. 2006 bp and 542 bp fragments containing the *stc-1* upstream region and the ATG codon of the gene were introduced separately into the pPD96.04 plasmid vector, which carries both *lacZ* and *gfp* reporter genes. The resulting recombinant plasmids were microinjected into N2 young adults to generate transmitting transgenic lines (section 2.9). I analysed the distribution of the *lacZ* and *gfp* products in many transformed individuals of at least two independent lines for each construct as detailed in section 2.10. The earliest expression of the *stc-1::lacZ::GFP* fusion detected was at the proliferative phase (first half of embryogenesis), prior to gastrulation (see section 3.1.1). In comma-stage embryos, intestinal and hypodermal cells (Fig. 4.9A) were identified to express the reporter gene. The number of cell types showing expression increased during the worm development and these included hypodermal, intestinal and nerve cells, as well as the secretory cell and cells in the pharynx and the vulva (Fig. 4.9).

I did not detect differences in the distribution of the reporter proteins in transgenic animals for each construct, therefore concluded that the 5' 539-bp region may contain the regulatory elements sufficient to drive *stc-1* expression. In addition, expression of the reporter gene fusions suggested that the predicted start of the *stc-1* gene is correct.

4.2.5 STC-1 antibody production

Polyclonal anti-STC-1 antibodies were generated in rabbits using an STC-1 fragment representing about half of the protein. I constructed the protein fragment by a recombinant technique that consisted basically in the construction of an expression clone, expression of the fusion protein and purification by chromatography.

Two distinct expression clones (EC) were made by inserting a 509 bp and a 452 bp fragment amplified from *stc-1* cDNA (named EC1 and EC2, respectively) into the pEQ-30 expression vector. These constructs were introduced into an *E. coli* strain, followed by growing of

Figure 4.9: *stc-1::lacZ::GFP* expression.

Fluorescent GFP images (left panels) are compared with DIC photomicrographs (right panels). The two bottom DIC pictures correspond to worms stained for *lacZ* activity. Hypodermal cells expressing the reporter gene in an embryo at comma stage (**A-B**) and in a young larva (**C-D**). **E-L** are micrographs of adult worms. **E-H** images show expression in cells of the head including the nerve ring, cells within and surrounding the pharynx and the excretory cell (arrow). **I-J,** *stc-1::GFP* expression in intestinal cells (arrowhead) and possibly the nerve cord (arrows). **K-L,** *stc-1::lacZ* expression. Arrows indicate hypodermal cells and the arrowhead intestinal cells (**K**); asterisk marks cells in the vulva, arrows the nerve cord and the arrowhead intestinal cells (**L**).

Facing page 108
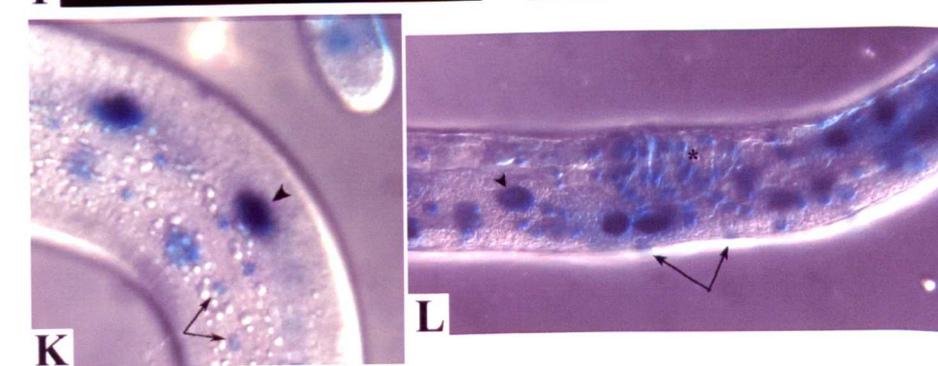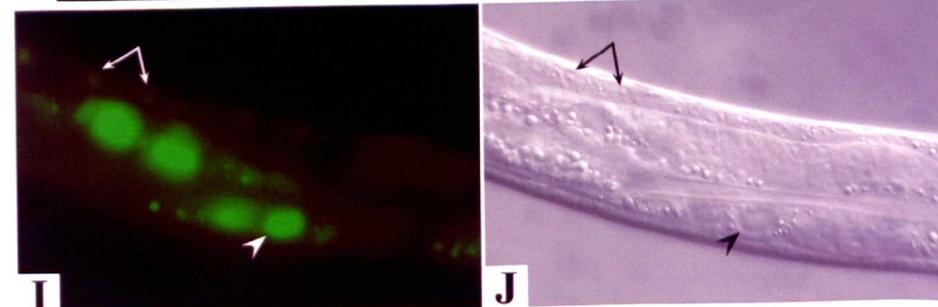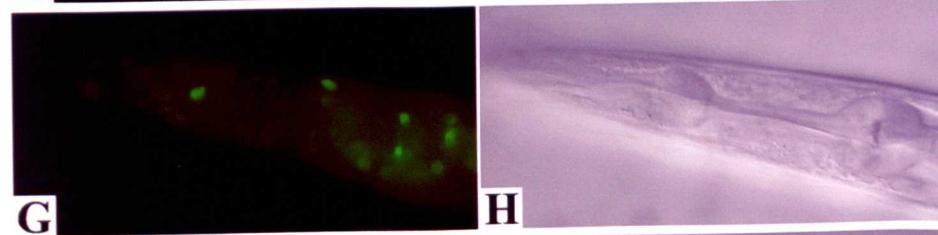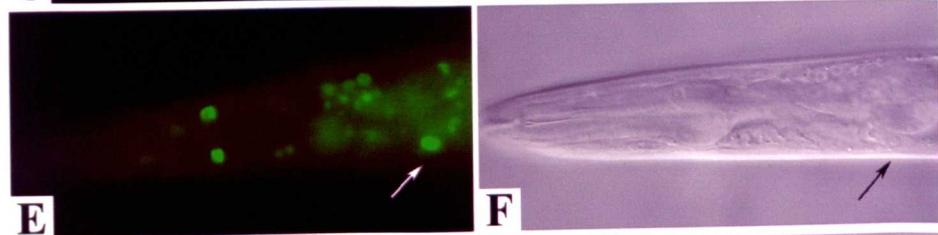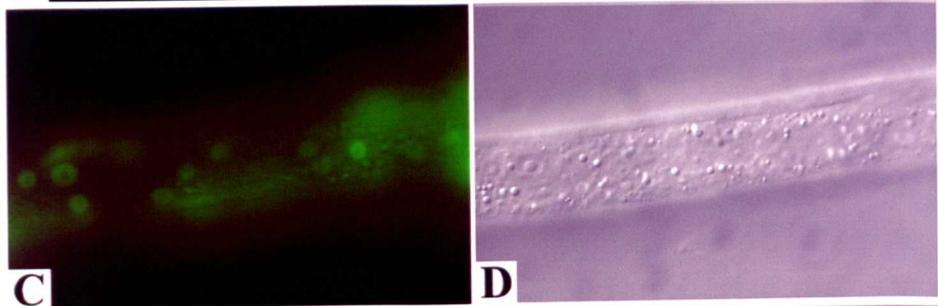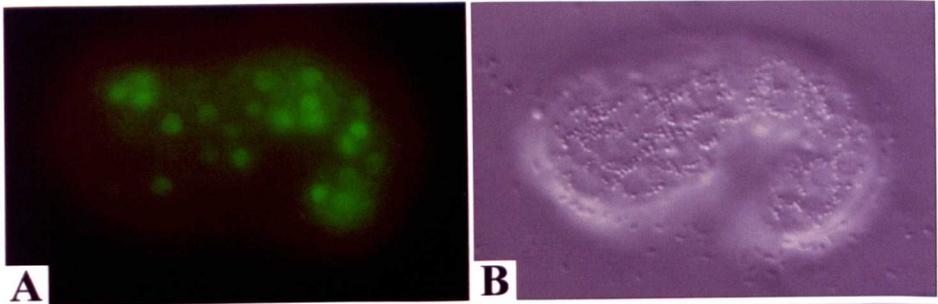
transformants and induction of expression of the clones. The expression vectors employed contained six repeats of histidine residues (6 x His) positioned such that the resulting expressed protein had a 6 x His tag at its N-terminus. I first carried out the complete procedure in order to check the functioning of both expression systems and to establish optimal conditions for subsequent large-scale purification of the proteins. Induction of expression of recombinant plasmids within the *E. coli* cells was performed with 1 mM IPTG at 37°C for three hours. These cells were lysed by chemical and mechanical means to release protein contents, and then centrifuged to separate soluble proteins (clear lysates) from cellular debris and other insoluble matter. The resulting clear lysates, as well as induced cultures, were analysed for the presence of the STC-1 proteins on an SDS-PAGE gel (Fig. 4.10). In many cases, recombinant proteins expressed in *E. coli* can associate producing insoluble aggregates (inclusion bodies). The expected 18.110 kDa STC-1 fragment from clone EC1 was detected in the induced samples (Fig. 4.10A) but not in the clear lysates (Fig. 4.10B), indicating that this protein was not in a soluble form. The 15.831 kDa protein fragment expected to result from clone EC2 was not present in any of the samples, suggesting that either the inductive conditions applied were not optimal for expression of this fusion protein or that the recombinant plasmid was not functional.

To solubilise the EC1 protein in inclusion bodies, the cellular debris and insoluble fraction from the lysis step was treated with 8 M urea for one hour at 37°C. Most of the protein came into solution under these denaturing conditions as can be seen in Figure 4.11. Despite absence of the recombinant protein in induced cells harbouring the EC2 construct, lysed insoluble samples were also treated with the denaturant. The results corroborated failure in expressing the EC2 protein.

Columns containing Ni-NTA matrices were employed to purify the 6 x His-tagged EC1 protein. This procedure is based on the binding capacity of histidine residues to nickel ions, which is lost if the pH of the solution is reduced. Pure STC-1 protein fragment was obtained by lowering the pH from 7.0 to 4.5 (Fig. 4.12). Once the inductive parameters, binding capacity and purification conditions to obtain the fusion protein were determined, the process was carried out on a large scale in order to isolate higher amounts of pure protein (Fig. 4.13). I calculated the concentration of protein in eluate 6 (1.725 mg/ml) and sent the required amount of sample to CovalAb, France (http://www.covalab.com) for antibody production.
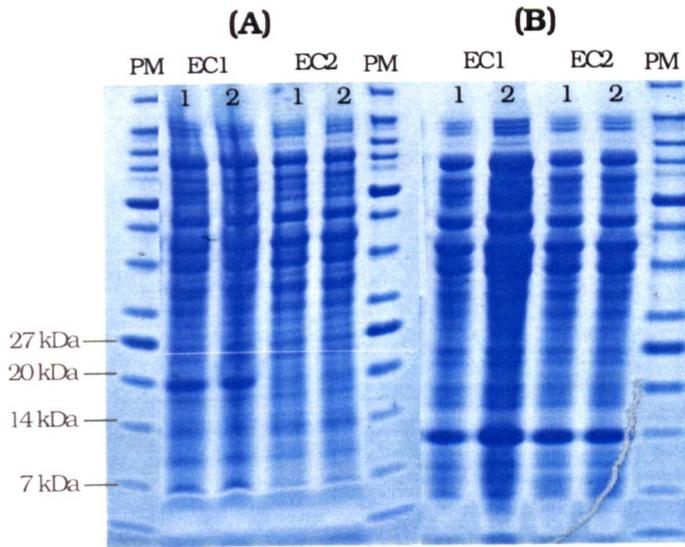
Figure 4.10: Protein content in solutions obtained following induction of protein expression (A) and in clear lysates (B). The two expression clones generated are indicated as EC1 and EC2. EC1 contained 509 bp of *stc-1* sequence and was expected to result in a 18.110 kDa protein fragment; a 452 bp gene region was present in EC2, which was expected to yield a 15.831 kDa STC-1 fragment. Size of relevant protein markers (**PM**) are shown on the left of the gel. Two transformant cells for each construct were analysed and have been numbered **1** and **2**.
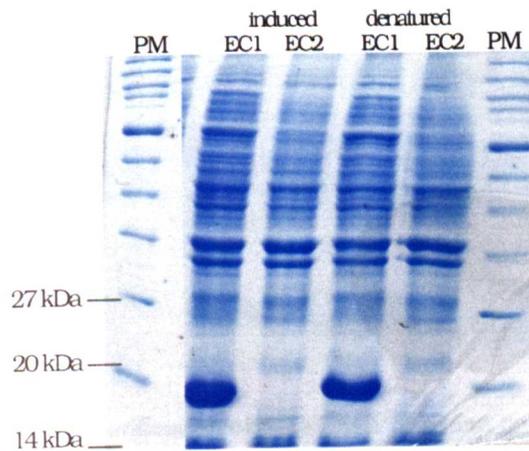
Figure 4.11: Protein content of induced cells and in lysed insoluble samples after denaturing treatment for expression clones 1 and 2 (EC1, EC2). PM: protein marker.
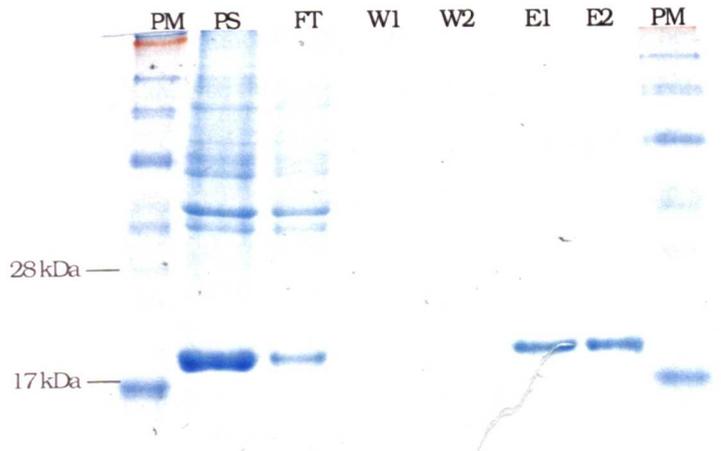
Figure 4.12: Purification of EC1 recombinant protein.

A Ni-NTA column was loaded with the 6 x His-tagged EC1 protein in an
8 M urea solution of pH 7.0. Washes were carried out at the same pH.
The protein was eluted at pH 4.5. **PM**: protein marker; **PS**: protein sample;
FT: flow-through; **W1-W2**: washes; **E1-E2**: eluates.
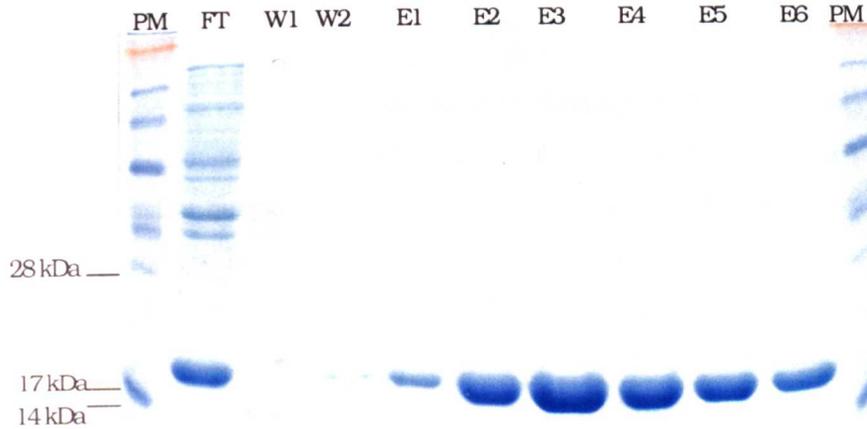
Figure 4.13: Purification of STC-1 EC1 fragment on large scale.

PM: protein marker; **FT**: flow-through; **W1**: 6.3 pH wash; **W2**: 5.9 pH wash;
E1-E6: 4.5 pH eluates. The protein was collected in twelve fractions, only
those with highest protein concentration are shown.

4.2.6 *stc-1* RNAi effects

The *stc-1* gene function was disrupted using the dsRNAi feeding technique. This method consists basically of feeding worms with an *E. coli* strain of expressing dsRNA of a plasmid-borne gene fragment upon inductive conditions. I employed 1122 bp of *stc-1* sequence (891 bp of coding sequence) to analyse the gene knockdown effects on the N2 and NL2099 strains at the temperatures 15°C, 20°C and 25°C. There was no apparent difference between the RNAi offspring obtained at the three temperatures, hence I only typed and scored that raised at 20°C. N2 and NL2099 worms fed with bacteria carrying the L4440 (pPD129.36) feeding vector without a gene insert, and thereby not expressing dsRNA, were utilised as negative controls. It should be noted that the NL2099 strain is homozygous for the mutant allele *rrf-3(pk1426)*, which renders these animals more sensitive to RNAi than wild-type N2s.

4.2.6.1 Inhibition of stc-1 function during embryogenesis

L3-L4 N2 and NL2099 larvae were placed on *stc-1* RNAi bacterial lawns to interfere with the function of this gene during embryogenesis of their F_1 progeny. These animals developed normally to adulthood but their broods were approximately half of those of control worms. RNAi-treated NL2099 worms had an average brood size of 53.9 ($n=15$), whereas non-treated NL2099 hermaphrodites produced a mean of 130.4 ($n=5$) F_1 animals. In the case of N2 hermaphrodites, the average brood size of RNAi-treated worms was 170.6 ($n=5$) compared to 306 ($n=3$) of non-treated animals.

stc-1 RNAi NL2099 hermaphrodites yielded approximately 16.9% dead embryos (Table 4.4). It should be noted, however, that NL2099 control worms laid dead embryos as well (the observed 'background' lethality ranged from 3.6% to 17.6%). The remaining F_1 progeny of the RNAi-treated hermaphrodites hatched and became severely uncoordinated (Unc) hardly moving forwards and backwards. These larvae developed more slowly than non-treated worms and appeared to arrest at the L3 stage as determined by examining the structure of the gonad with Nomarski microscopy. ~38.2% of these larvae (~31.7% of the total progeny) displayed obvious moulting defects: remains of previous cuticles attached to discrete parts of their body (Fig. 4.14A) and problems during cuticle synthesis (Fig. 4.14B). Most of the animals with the latter defects appeared to be unable to break through the old cuticle. I assume that the inability to eat is the cause of death in this case. Some larvae collapsed and died (~7.3%, ~6.1% of the total progeny), however a significant fraction of the progeny died shrinking noticeably or did so when poked with the worm pick (~39.6%, which corresponds to ~32.9% of the total progeny).

Figure 4.14: Moulting defects of *stc-1* RNAi NL2099 larvae.
Part of the old cuticle attached to the animal's mid-body (**A**); larva unable to break the cuticle remaining trapped during moulting (**B**).

Figure 4.15: Collagen DPY-7 distribution in the cuticle of a wild-type larva (A) and an *stc-1* RNAi NL2099 animal (B). The typical DPY-7 ribbons (**A**) tend to be discontinuous or are not present in some regions of the cuticle of *stc-1* RNAi NL2099 larvae (**B**). The disrupted DPY-7 bands sometimes have thickened endings.

Figure 4.14
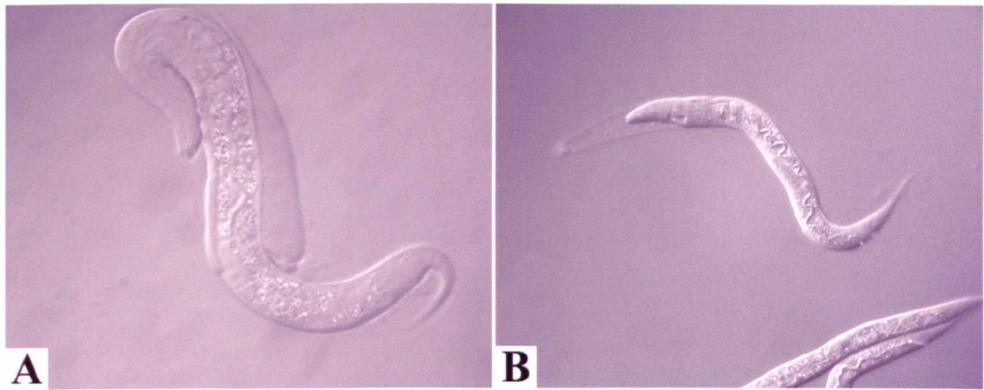
Figure 4.15
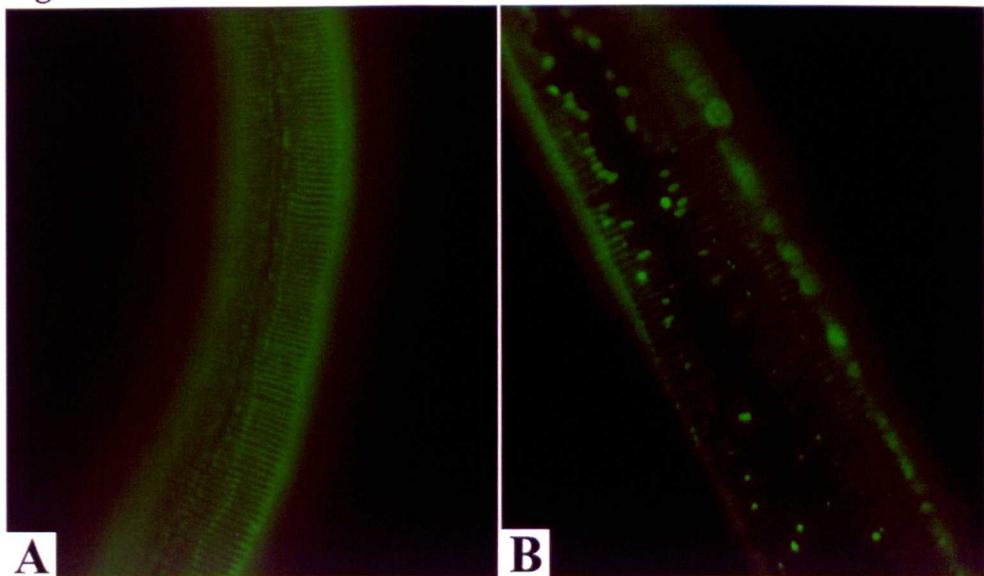
~14.9% of the hatched progeny (~12.4% of the total progeny) died for reasons difficult to determine under a stereomicroscope.

Most of the offspring of RNAi-treated N2 animals hatched and seemed to grow as healthy larvae (Table 4.5). These developed slower and were less active than the N2 control offspring. Some Unc animals were first observed at the L4 stage but they were not determined precisely since other defects affected movement of the worms at this stage. Approximately 28.8% of the progeny had difficulties in shedding the old cuticle, displaying severe constrictions and attached previous cuticles or more frequently, remaining trapped in the L4 cuticle. Occasionally, larvae shrunk and died or young adults burst through the vulva. The majority of the progeny, however, were not noticeably affected and developed as fertile adults.

(A) *stc-1* RNAi on NL2099 hermaphrodites

NL2099 animal	dead embryos	Larval defects					total progeny
		shrink	burst	old cuticle attached	encased in cuticle	other / undetermined	
1	11	16	2	2	8	7	46
2	13	19	3	4	3	5	47
3	12	13	1	2	4	8	40
4	8	12	3	3	9	8	43
5	12	22	9	6	18	4	71
6	10	24	3	7	26	11	81
7	5	9	0	4	2	3	23
8	2	4	1	3	4	2	16
9	5	35	2	11	20	9	82
10	8	24	5	0	24	6	67
11	3	16	3	3	12	3	40
12	20	22	10	7	28	9	96
13	4	13	1	2	16	10	46
14	9	14	1	1	8	8	41
15	15	23	5	2	17	7	69

NL2099 animal	dead embryos	Larval defects					total progeny
		shrink	burst	old cuticle attached	encased in cuticle	other / undetermined	
Total	137	266	49	57	199	100	808
Percentage (%)	16.9	32.9	6.1	7.1	24.6	12.4	100

(B) non-treated NL2099 hermaphrodites

NL2099 animal	total progeny	dead embryos	percentage of death (%)
1	119	14	11.8
2	170	19	11.2
3	142	25	17.6
4	111	5	4.5
5	110	4	3.6
Total	652	67	10.3

Table 4.4: F$_1$ dead embryos and total progeny of *stc-1* RNAi NL2099 hermaphrodites (A) and of NL2099 animals fed with bacteria harbouring the L4440 vector (B). Frequency in the progeny and types of larval defects and causes of larval lethality are specified in table A.

(A) *stc-1* RNAi on N2 hermaphrodites

N2 animal	dead embryos	Post-embryonic defects					total progeny
		shrink	burst	old cuticle attached	encased in cuticle	other / undetermined	
1	3	1	1	6	28	9	187
2	2	1	4	7	53	4	188
3	1	0	2	7	56	0	158
4	2	1	3	2	32	9	179
5	3	0	0	9	46	9	141

N2 animal	dead embryos	Post-embryonic defects					total progeny
		shrink	burst	old cuticle attached	encased in cuticle	other / undetermined	
Total	11	3	10	31	215	31	853
Percentage (%)	1.3	0.3	1.2	3.6	25.2	3.6	

(B) non-treated N2 hermaphrodites

N2 animal	total progeny	dead embryos	percentage of death (%)
1	372	6	1.6
2	238	0	0
3	308	1	0.3
Total	918	7	0.8

Table 4.5: Frequency of embryonic and post-embryonic death, and of defects detected in progenies of *stc-1* RNAi N2 hermaphrodites (A) compared with those of non-treated animals (B).

4.2.6.2 Cuticular defects of stc-1 RNAi NL2099 animals

The integrity of the cuticle of NL2099 larvae from *stc-1* RNAi hermaphrodites was analysed by means of immunofluorescence microscopy and treatment in a range of solutions. Immunofluorescence analyses were performed using the DPY7-5a antibody. As can be seen in Figure 4.15B, the collagen protein DPY-7 is significantly disorganised or even absent in the cuticle of the affected larvae.

The experiment mentioned in section 3.4.1.7 was performed to check if the cuticle of the RNAi NL2099 worms was weak. L2-L3 RNAi larvae were placed into dH_2O, 1% β-ME or 1% SDS. 64.3% of the animals in dH_2O disrupted through the anus or other points along their body ($n=14$), whereas 64.7% burst in 1% β-ME ($n=17$) and 5% in 1% SDS ($n=20$). NL2099 L2-L3 larvae not exposed to RNAi were treated in the above solutions as negative controls. 1% SDS disintegrated progressively the animals' body contents, leaving only their cuticles. This effect

was considerably faster in RNAi larvae than in non-treated worms. In contrast to the non-treated larvae, rupture of the RNAi animals in the tested solutions corroborates that an intact cuticle is not assembled in an *stc-1* knockdown background.

4.2.6.3 Inhibition of stc-1 function during post-embryogenesis

To study the function of *stc-1* during post-embryonic development, NL2099 and N2 embryos were placed onto *stc-1* RNAi bacterial lawns. RNAi-treated NL2099 animals developed to adulthood showing certain mild defects when compared to non-treated NL2099 worms: slightly slow growth rate, reduced mobility and brood size. Some adults became progressively sick and died; occasionally, a few worms burst through the vulva. No noticeable differences were detected between treated and non-treated N2 animals, with the exception of occasional cases of death due to rupture through the vulva.

4.3   DISCUSSION

The phenotypic characterisation of the *C. elegans* mutant allele *stc-1(ij15)* is dealt within chapter 3. *stc-1(ij15)* was isolated from a forward genetic screen performed previously (C. Clucas and I. Johnstone, Glasgow University, Glasgow, UK) to identify recessive alleles that cause death late in embryogenesis, specifically after the elongation process, when the first worm cuticle is produced. This chapter comprises detailed molecular analysis of the *stc-1* gene.

4.3.1   *stc-1(ij15)* embryos are defective in cuticle synthesis

stc-1(ij15) homozygotes elongate to a three-fold embryo and die at this stage as determined by Nomarski microscopy. They burst along the mid-posterior part of the ventral side of their body, with the consequent extrusion of internal cells and loss of their worm shape.

Johnstone and Barry (1996) have reported that the expression of cuticle collagen genes increases prior to cuticle synthesis in a temporally ordered manner. They detected three discrete peaks of expression, each corresponding to specific sets of cuticle collagen genes occurring at different times before cuticle assembly. Thus, cuticle collagen genes can fall into three distinct categories according to the period of their expression: early, middle and late. The collagen genes *dpy-7* and *col-12* are early and late expressed cuticle collagen genes respectively. In embryo development, *dpy-7* is first expressed at comma stage (approximately four hours before cuticle secretion) whereas the earliest *col-12* expression occurs at the three-fold stage (around

the time of cuticle secretion). The reporter constructs *dpy-7::GFP* and *col-12::GFP* were employed in the study of *stc-1(ij15)* mutants as markers of cuticle collagen gene transcription. *stc-1(ij15)* embryos transcribed both *dpy-7* and *col-12* reporter transgenes revealing that the function of the *stc-1* gene is not necessary for the transcription of these two cuticle collagen genes. Expression of the terminal marker *col-12::GFP* is normally associated with three-fold stage, and therefore is consistent with the post-elongation death of *stc-1(ij15)* mutants determined by Nomarski.

The localisation of DPY-7 in the embryo and within the cuticle has been well described by means of immunofluorescence experiments with an antibody, termed DPY7-5a, specific to this collagen protein (McMahon *et al.*, 2003). The secreted DPY-7 protein in wild-type animals is present in the furrows of the cuticular annuli, and is visualised as uniformly separated, circumferential bands when probed with the DPY7-5a antibody. The DPY-7 protein is not detected in the alae characteristic of L1 and adult stages, and hence it is not a component of these stage-specific cuticular structures. I employed the DPY7-5a antibody to investigate cuticle secretion in *stc-1(ij15)* mutants. DPY-7 bands corresponding to the secreted collagen protein were detected in *stc-1(ij15)* embryos, which indicates these mutants form the L1 cuticle to a certain degree. It is not known if the distribution of the secreted DPY-7 is altered in the *stc-1(ij15)* mutant background since the DPY-7 bands are squeezed together, most likely due to the collapse and loss of shape phenotype of the mutant embryos.

The fact that *stc-1(ij15)* embryos complete elongation normally, transcribe *dpy-7* and *col-12* collagen genes and secrete DPY-7 suggests the presence of an intact and functional hypodermis and indicates that the mutant embryos die at the terminal phase of embryogenesis, around the time of cuticle secretion. The elongation process of embryogenesis is driven by circumferential contractions of the cytoskeleton within the hypodermal cells (Priess and Hirsh, 1986). Once elongation is completed, the hypodermal cells commence to secrete the first larval cuticle and the cytoskeleton begins to disarrange. The secreted cuticle takes the role of maintaining the elongated animal form defined by the hypodermis from this point on. Supporting evidence to this important function of the cuticle after elongation is provided by the phenotypic defects of *sqt-3(e2117)* mutant embryos (Priess and Hirsh, 1986). *sqt-3* encodes the cuticle collagen COL-1. *sqt-3(e2117)* homozygotes are temperature sensitive and die during embryogenesis if raised at 25°C. At the restrictive temperature, the embryos elongate normally but subsequently shorten in length. Examination of the cuticles of these mutants by transmission electron microscopy reveals that they produce a defective cuticle (lack the striated layer), and hence fail in

maintaining the elongated shape. More recently, Winter and Page (2000) have observed similar phenotypes upon disruption of *phy-2* in a *dpy-18(e364)* background and of *pdi-2* in a wild-type N2 background by RNAi. The *dpy-18(e364)* allele is a presumed null mutant of the gene *phy-1*; *phy-1* and *phy-2* encode for the catalytic α-subunits of P4H and *pdi-2* for PDI, which functions as the β-subunit of P4H. The multicomplex enzyme P4H catalyses the hydroxylation of proline residues in procollagen chains, required for the stabilisation of the collagen triple helix by hydrogen bonds, and acts as a molecular chaperone in procollagen folding. PDI has also been ascribed other important roles in procollagen biosynthesis, including formation of disulphide bridges and chaperoning assembly. Considering the function of these enzymes, the retraction after elongation of *phy-2* RNAi *dpy-18* and *pdi-2* RNAi embryos may also be due to the synthesis of a malformed cuticle. It should be noted that Friedman *et al.*, (2000) have reported a different embryonic phenotype for double *phy-1* and *phy-2* knockouts (*dpy-18(ok162) III; phy-2(ok177) IV*): the embryos elongate usually to two-fold and then either arrest at this stage or often collapse into a disorganised mass of cells. *ok162* and *ok177* are deletion mutations, and most likely null alleles, of *phy-1* and *phy-2* respectively. The differences in the embryonic phenotypes described by the two groups could be explained if there was a reduction rather than complete disruption of *phy-2* function in the *dpy-18(e364)* background by RNAi and/or if *dpy-18(e364)* does not behave entirely as a null mutant (Hill *et al.*, 2000a). However, incomplete RNAi penetrance and/or residual *phy-1* function in *dpy-18(e364)* animals are probably not the explanation for such subtle differences in the embryonic phenotypes. The phenotypic defects described by Winter and Page (2000) are more likely to be correct as they were carefully examined by time courses of the development of the affected embryos and were consistent when comparing *phy-2* RNAi *dpy-18* and *pdi-2* RNAi animals; both *phy-2* RNAi *dpy-18* and *pdi-2* RNAi animals presumably lack the same enzymatic function, P4H. The phenotype of *phy-1 phy-2* double null mutants (*dpy-18(ok162) phy-2(ok177)*) was not determined accurately in Friedman *et al.*, (2000) since it was based on the observation of already dead mutant embryos. Clearly, a precise time course of the development of *phy-1 phy-2* double null mutants will clarify the incongruent reports.

stc-1(ij15) homozygous embryos die at the end of embryogenesis, in a similar yet distinct manner to the animals lacking *sqt-3, pdi-2*, or *phy-1 phy-2* combined functions discussed above. They do not retract in length after elongation but rupture and lose their structure, indicating that these embryos may die during rather than after cuticle secretion. This could be the case if there was insufficient secretion of the components required to assemble the cuticle by the time the hypodermal cytoskeleton disorganises. Lack of a supporting structure around the elongated

animal could result in the observed burst out of cells due to the internal hydrostatic pressure existing in the worm. Such hydrostatic pressure is important for locomotion of the hatched nematode.

4.3.2 *stc-1* encodes a HSP70-like protein

stc-1 encodes a 469 amino acid product which belongs to the HSP70 protein family. HSP70 proteins are molecular chaperones of approximately 70 kDa with a shared domain structure, consisting of a highly conserved N-terminal ATPase domain of ~44 kDa and a less conserved C-terminal substrate-binding region of ~25 kDa. STC-1 is an unusual member of this family in that it has a calculated molecular mass of about 51 kDa and lacks the C-terminal substrate-binding region. The predicted STC-1 sequence contains a signal peptide that may target the protein to the secretory pathway organelle, the ER. C-terminal sequences known to mark ER resident proteins for their retention in the compartment are not present in the STC-1 sequence. STC-1 is homologous to human and rat STCH proteins, the amino acid identity being approximately 39%. There is no STCH-like protein predicted by the Drosophila genome (Otterson *et al.*, 1994). However, the fact that the molecule sequence and structure is conserved in divergent species such as mammals and *C. elegans* suggests an important role for this HSP70-like protein in the cell. Human and rat *stch* are single copy, constitutively expressed genes with up-regulated expression upon incubation with the calcium ionophore A23187, but not by heat shock exposure (Otterson *et al.*, 1994; Otterson and Kaye, 1997). Moreover, the human *stch* gene appears to be transcribed in all cell types. Of note is the similar expression pattern observed with the ER resident HSPs GRP94 and BiP in mammals (Argon and Simen, 1999). The human STCH protein has been detected in the cellular cytoplasm in a distinct, but similar pattern to BiP, and by independent analyses within the lumen of a cellular organelle (Otterson *et al.*, 1994). Both human and rat STCH proteins contain an N-terminal signal peptide and several potential N-glycosylation sites (five in rat and four in human) throughout their sequence. The reported expression pattern and localisation results together with the signal peptide present in the protein product indicate a possible function for STCH in the secretory pathway. Finally, the ATPase activity of human STCH has also been measured and shown to be peptide independent in contrast to other well characterised HSP70s such as BiP and DnaK (Otterson *et al.*, 1994).

As an attempt to study the role of STCH in protein processing, Kaye *et al.*, (2000) have identified two STCH-binding proteins, named Chap1 and Chap2, by the yeast two-hybrid technique. Chap1 is a homologue of the *Saccharomyces cerevisiae* Dsk2 protein, which is

important for cell division; whereas Chap2 corresponds to the previously isolated human protein BAT3, and is homologous to the *Xenopus* scythe protein involved in apoptosis. The minimal amino acid region of Chap1/Dsk2 and Chap2/scythe proteins required for association with STCH has been determined and consists of two sequence stretches with strong similarity to the conserved domain of the chaperone binding proteins STI1 (also known as p60 or HOP) and HIP (Lassle *et al.*, 1997; Irmer and Hohfeld, 1997). The STCH binding site for Chap1/Dsk2 and Chap2/scythe locates within the ATPase domain and is highly conserved in HSP70 proteins. The fact that this region is not STCH unique suggests that Chap1/Dsk2 and Chap2/scythe may not be STCH specific but general HSP70 interacting proteins. Both Chap1/Dsk2 and Chap2/scythe are members of the large family of ubiquitin-like proteins characterised by the conserved UbL and UbA domains. The N-terminal UbL domain has been shown to be necessary for the interaction of some ubiquitin-like proteins with the 26S proteasome (Schauber *et al.*, 1998). Taken together, these observations suggest that the ubiquitin-like proteins Chap1/Dsk2 and Chap2/scythe may mediate HSP70 function in cell cycle and cell death processes.

4.3.3 *stc-1* is expressed in the cuticle-synthesising tissue at all developmental stages

Spatial expression of *stc-1* was examined by detection of the *stc-1* promoter-driven *gfp* and *lacZ* products. Two DNA fragments were employed containing variable lengths of the gene's upstream region plus the start codon ATG. Expression of the *gfp* and *lacZ* reporter products from the distinct 5' *stc-1* fragments was reproducible, indicating that the shorter upstream region (539 bp) contains the necessary regulatory elements for *stc-1* transcription. Moreover, the successful expression of the *stc-1::*reporter construct validated the predicted start of the gene. *stc-1* is expressed in the intestine and hypodermis of embryos, with earliest expression prior to gastrulation. At post-embryonic stages, nerve cells, the secretory cell and cells within the pharynx and the vulva express the *stc-1* gene, in addition to the hypodermis and the intestine.

The temporal expression of *stc-1* was determined at various points during larval and adult development by means of RT-PCR. Similar levels of transcript were detected at the distinct developmental time points measured, indicating *stc-1* may be constitutively expressed throughout post-embryonic development.

Cuticle collagens are produced exclusively in the hypodermal cells. The expression and synthesis of these collagens peak prior to the formation of each stage-specific cuticle (Johnstone and Barry, 1996). Although the expression of *stc-1* did not fluctuate as observed for cuticle

collagen genes, *stc-1* is transcribed in the cuticle synthesising tissue, the hypodermis, at all developmental stages.

4.3.4 *stc-1(ij15)* lesion may disrupt the structure conserved between HSP70s, actin and hexokinase ATPases

The *ij15* lesion results in a G422E substitution in the amino acid sequence that is within the ATPase domain common to HSP70 proteins. The three-dimensional structure of the ATPase fragment has been solved for a member of the HSP70 family, the bovine HSC70 (Flaherty *et al.*, 1990). It is composed basically of two domains separated by a cleft and an interdomain region. The interdomain region consists of two α-helices that traverse the base of the cleft in opposite directions. A single ATP molecule binds at the bottom of the cleft, just above the traversing α-helices. This structure is strikingly similar to those of actin and hexokinase, in spite of very low sequence identity between the HSC70 fragment and the other two proteins (Flaherty *et al.*, 1991). Like HSC70, actin and hexokinase are ATPases. It is known that the highly conserved, common structure between HSP70s, actin and hexokinases undergoes distinct conformational changes during the ATPase cycle, with ultimate effects on the protein function (Hurley, 1996; Schüler, 2001). Some of the conformational states and mechanisms by which these are produced differ between the three proteins. ATP hydrolysis and subsequent release of the inorganic phosphate leads to a scissors-type opening of the cleft of actin and hexokinase molecules that is mediated by shear movements of the crossover α-helices relative to each other and to the two domains (Hurley, 1996; Schüler, 2001). In contrast, binding of a nucleotide exchange factor to the ADP-bound state of HSC70 (and of its bacterial homologue, DnaK) induces rotation of the upper part of one of the domains, with the consequent opening of the cleft of the molecule (Harrison *et al.*, 1997; Sondermann *et al.*, 2001). A shearing motion in the interdomain region has not been reported to be involved in the cleft opening of this protein as yet.

Comparison of HSC70, actin and hexokinase structures and sequence alignments from members of each of the three protein families have revealed a pattern of conserved and equivalent amino acids within short segments that localise to the nucleotide binding site and interdomain region (Flaherty *et al.*, 1991; Bork *et al.*, 1992; Kabsch and Holmes, 1995; Hurley, 1996). In all three proteins, the two crossover α-helices make a close contact that involves three residues: a glycine on one α-helix and the first and last amino acids on the other α-helix (Bork *et al.*, 1992). The glycine appears to be a consensus residue in the three protein families as well as in other proteins structurally similar to actin and hexokinase; the other two amino acids are equivalent residues between the structural homologues. Moreover, the connection between the α-helices is

present in both the open and closed conformations of the common structure. It is conceivable that the high conservation of the glycine residue is imposed by a critical role of the helix-helix contact in the ATPase activity of the protein. The missense mutation identified here in the *stc-1(ij15)* protein (G422E) affects the invariant glycine residue (G372 in the bovine HSC70) involved in the interaction between the α-helices of the interdomain region. The properties of glycine and glutamate are significantly different. Glycine is non-polar and has the simplest amino acid structure, which confers it a special role in protein structure and folding. It allows much more structural flexibility than other amino acids, as it can fit where other residues would be too bulky. Glutamate is negatively charged and has the tendency to interact with H_2O molecules. Thus, it would be expected that the substitution of glycine 422 by glutamate would disrupt the connection between the two interdomain α-helices as well as its functional implications.

There are a vast number of studies reporting structural and functional effects of mutations within the ATPase domain of different actin, HSP70 and hexokinase proteins. In many of these reports, the mutated amino acids have been shown to be involved in the ATPase mechanism or in the coupling of the ATPase activity with the regulation of the protein function (e.g. Buchberger *et al.*, 1994; Flaherty *et al.*, 1994; Wilbanks *et al.*, 1994; Zeng *et al.*, 1998; Elefant and Palter, 1999; Johnson and Mckay, 1999; Barthel *et al.*, 2001). The G422 of the STC-1 ATPase domain is not expected to be involved directly in the ATPase mechanism but to have a structural role in the protein. Within the common three-dimensional structure, G422 would intervene in the connection of the two α-helices of the interdomain region, which mediates the transition between the distinct structure conformations adopted during the ATPase cycle. Thus, G422 may be a key residue required for either the proper folding of the common structure or for the switch between its conformations that regulates function. In the latter case, the common structure could bind to ATP and/or to other proteins but then incapable to respond to the interactions. The fact that the highly conserved G422 is essential for STC-1 *in vivo* function in *C. elegans* leads to the suggestion that this may be the case at least for other HSP70s and for proteins of the actin and hexokinase families.

4.3.5 Inhibition of *stc-1* function by RNAi

The function of the gene *stc-1* was targeted for disruption in N2 and NL2099 strains either during embryogenesis or post-embryogenesis by means of RNAi feeding. There was a more complete interference of the *stc-1* function in NL2099 animals than in N2s, consistent with the

fact that the NL2099 strain is more sensitive to the RNAi technique. I limit to comment and discuss in this section the RNAi results obtained with the NL2099 strain, as they are considered to be a closer resemblance of an *stc-1* null mutant phenotype. Moreover, the effects of interference of *stc-1* function were apparently identical at the three temperatures tested (15°C, 20°C and 25°C), and therefore are not temperature dependent.

4.3.5.1 *stc-1* may be required for fertility, cuticle synthesis and neuron function and is essential for development and viability

To favour efficient interference of the gene function during embryogenesis, hermaphrodites were fed with *stc-1* RNAi bacteria well before adulthood and production of the progeny, specifically from the L3-L4 stage. The brood size of the *stc-1* RNAi hermaphrodites was, on average, half that of non-treated worms, which suggests that the function of the *stc-1* gene is important in the fecundity of the nematode. Approximately 16.9% of the F_1 RNAi progeny died during embryogenesis. This percentage was not considered significant as similar frequencies of embryonic death were sometimes observed with non-treated NL2099 worms. The percentages of dead embryos segregated by non-treated hermaphrodites was variable and ranged from 3.6% to 17.6%. The embryonic development of the F_1 RNAi progeny could be monitored using Nomarski microscopy to determine if there is a distinct class of death from that associated with feeding the control bacterial strain, that is, if *stc-1* RNAi causes lethality during embryogenesis. If interference of *stc-1* function caused embryonic death, the phenotype of the RNAi lethal embryos should be characterised and compared with that of *stc-1(ij15)* mutant embryos. Such comparison would provide useful information regarding the functional consequences of the *stc-1(ij15)* lesion, whether it is a null mutation or not. The failure of *stc-1* RNAi to mimic the phenotype of *stc-1(ij15)* mutants could be explained if *ij15* was a dominant allele. STC-1(G422E) could have a dominant effect if, for instance, it retained the ability to associate with other chaperones and/or co-chaperones but was unable to disassociate from these, thus impairing chaperone function in the secretory pathway within the cell. However, such effect is unlikely as *ij15* is recessive over wild type.

All the hatched RNAi F_1 animals (~83.1%) became severely Unc, developed slowly and appeared to arrest at the L3 stage. Unc phenotypes are typically indicative of a defect in the neuromuscular system. *stc-1* is expressed in nerve cells of the head as well as in the nerve cord and, although not for certain, it seems not to be expressed in the body wall muscle. Taking into account these observations, it can be concluded that the *stc-1* protein product may be required for the optimal function of neurons. Moreover, *stc-1* may be essential for development and

viability as the described slow growth and larval arrest suggest. ~38.2% of the hatched progeny had difficulties during moulting, which indicates a possible function of *stc-1* in cuticle synthesis. Most of the worms with moult defects (~29.7%) were unable to break the old cuticle, dying trapped, possibly from starvation; the remainder (~8.5%) were noticeable because of their anomalous cuticular structures, formed due to incomplete shedding and retention of parts of old cuticles. Approximately 39.6% of the larvae either contracted considerably when prodded or died with a severe shrunken phenotype. The contraction of the body of these larvae was in length and therefore differs from the phenotype known as "shrinker" associated with some Unc mutants. Mutant animals with a shrinker phenotype contract dorsal and ventral body wall muscles simultaneously when prodded, and are defective in the biosynthesis or utilisation of the neurotransmitter GABA (e.g. *unc-25*, *unc-47* and *unc-30* Eastman *et al.*, 1999). A tendency to shrink and relax ("rubberband" phenotype) in response to a touch on the head has also been described for certain Unc mutants such as *unc-93* and *unc-110*, which have muscle activation defects. The cause of the observed contraction of some RNAi larvae is not known. There is, however, the possibility that such shrinking defects are due to failure in cuticle synthesis. If these animals were unable to produce a new cuticle, they would not maintain their shape and therefore retract in length upon loss of the connection between the old cuticle and muscle that occurs during moulting. Consistent with this is the fact that those RNAi larvae that shrink also die; in contrast to the above Unc mutants, whose shrinking defects do not cause death. Finally, a small percentage of the RNAi larvae burst and died (~7.3%) whereas ~14.9% died for undetermined reasons. The various RNAi phenotypes observed indicate that *stc-1* has more than one developmental role, agreeing with the expression of *stc-1* in several cell types.

4.3.5.2   *F₁ stc-1 RNAi larvae produce a defective cuticle*

Two different assays were employed to examine cuticle integrity in the F_1 larvae obtained from *stc-1* RNAi-treated NL2099 hermaphrodites. One assay was based on the immunodetection and distribution analysis of the cuticle collagen DPY-7. The DPY-7 bands of the affected larvae were notably disrupted. They tended to be fragmented or absent and sometimes had thickened endings. The absence or abnormal localisation of the secreted DPY-7 protein within the cuticle of the RNAi larvae reveals the inability of these animals to produce an intact cuticle.

The other assay consisted of placing L2-L3 RNAi larvae into dH_2O, 1% β-ME or 1% SDS. Non-treated worms at similar larval stages were employed as controls. In addition to its function in motility and maintenance of body shape, the cuticle is an impervious extracellular structure that acts as barrier between the nematode and the environment, conferring resistance to an

ample range of solutions as well as facilitating osmoregulation. The major components of the cuticle are collagen proteins that are extensively cross-linked primarily by disulphide bonds. In contrast to control worms, which maintained their body shape, the RNAi larvae disrupted at distinct points along their body in the above solutions. 64.3% burst in dH_2O, 64.7% in 1% β-ME and 5% in 1% SDS. Rupture of the RNAi larvae can be explained if these animals are surrounded by weak, non-functional cuticles that do not withstand the increase of the animal's hydrostatic pressure resultant from partial permeation of the solutions tested across the cuticle. In this analysis, dH_2O acts solely as a hypotonic medium whereas β-ME and SDS have an effect on the integrity of the animal. β-ME is a known sulfhydryl reducing agent (it breaks disulphide bonds) used as a 5% solution in the solubilisation of cuticle proteins (Cox *et al.*, 1981a). Here the effects of 1% β-ME could be sufficient to greatly weaken the cuticle of a significant proportion of RNAi animals causing them to burst. However, the percentage of RNAi larvae that burst in 1% β-ME is similar to that of RNAi worms that collapse in dH_2O. Taking into account β-ME was diluted to 1% in dH_2O, rupture of RNAi worms in 1% β-ME may be caused by the effects of dH_2O and not necessarily by those of β-ME. Therefore, 1% β-ME should be prepared in an osmotic medium to examine exclusively the effects of this reducing agent on the cuticles of RNAi larvae.

SDS is a denaturing detergent, and is employed for the isolation of cuticles from all cellular material (Cox *et al.*, 1981a). Sonication and treatment with 1% SDS leaves cuticles largely unaffected yet dissolves cellular membranes and denatures cellular proteins. The nematode's hydrostatic pressure may be maintained by an ion imbalance between the worm and its environment. H_2O molecules would enter into the nematode contributing to the internal hydrostatic pressure if there were higher ion levels in the animal relative to its surrounding medium. The effects of 1% SDS may disrupt such ion imbalance and entrance of H_2O molecules by osmosis, ultimately affecting the hydrostatic pressure. RNAi worms placed in 1% SDS adopt progressively a clear aspect due to internal disintegration of all cells, and occasionally burst. The lack of internal hydrostatic pressure may account for the small percentage of rupture of the RNAi larvae in 1% SDS. It should be noted that the internal body contents of the RNAi worms tended to dissolve faster than those of non-treated animals in this solution.

4.3.6 *stc-1* function may be dispensable for post-embryonic development

To disrupt the gene function post-embryonically, embryos from untreated mothers were allowed to hatch on plates seeded with *stc-1* RNAi bacteria. Thus, the animals were not exposed to the

inhibiting *stc-1* dsRNA during embryogenesis but from the L1 stage onwards in development. Treated worms reached adulthood and were not severely affected by *stc-1* RNAi. They developed slightly slower, had a smaller brood size and were less active than non-treated animals. Some of the affected adults died prematurely due to general sickness or, occasionally, rupture through the vulva. The range of mild defects observed could suggest that *stc-1* function is not essential for post-embryonic development. However, it should be noted that the *stc-1* RNAi could be partially penetrant. That is to say, an incomplete inhibition of the *stc-1* function could exist. This would be consistent with failure to induce embryonic lethality by *stc-1* RNAi during embryogenesis.

4.3.7 Concluding remarks

The *C. elegans* gene *stc-1* encodes a putative HSP70 chaperone homologous to the human and rat STCH proteins. STC-1, as well as its mammalian counterparts, contains an N-terminal signal sequence. Significantly, the human STCH protein is localised within the lumen of a membrane-based compartment and its expression is induced following incubation with the calcium ionophore A23187. This expression pattern is characteristic of mammalian ER chaperones such as GRP94 and BiP, whose functions are regulated by calcium ion fluxes. Taken together, these observations suggest that STC-1 may have a conserved function in the secretory pathway.

The cuticle defects caused by the genetic alteration *stc-1(ij15)* or by *stc-1* RNAi during embryogenesis indicate that the function of this gene may be required for cuticle synthesis. Consistent with this, *stc-1* is expressed at all developmental stages in the hypodermis, the tissue responsible for synthesis and secretion of the cuticle collagen proteins. These findings imply that STC-1 may be involved in the assembly of cuticle collagen proteins. If this were the case, STC-1 would be the first HSP-like molecule identified in *C. elegans* with a role in collagen biogenesis. Moreover, the functions of human and rat STCH proteins have not been reported as yet. By homology with STC-1, STCH could be required in the synthesis of human collagens.

4.3.8 Future work on *stc-1*

Immunofluorescence analysis with the anti-STC-1 sera should be carried out to determine the localisation of the STC-1 protein within the cell. The distribution of STC-1 should provide firm clues as to the function of this protein, specifically in view of its involvement in collagen biosynthesis suggested here.

As an approach to reveal the primary cause of the *stc-1(ij15)* mutant phenotype, the hypodermal function of STC-1 could be studied in *ij15* homozygotes by mutant rescue analysis. *dpy-7* is an early-transcribed cuticle collagen gene expressed exclusively in hypodermal cells. A *dpy-7::stc-1* gene fusion could be constructed to transform IA432 hermaphrodites, which are heterozygous for *stc-1(ij15)*. The progeny of stable transgenic lines should be then examined for rescue of the post-elongation death of *ij15* homozygous embryos. Phenotypic rescue would indicate that the *ij15* phenotype is the result of the impaired function of STC-1 in the hypodermis, whereas lack of rescue would imply that other tissue-specific functions of STC-1 cause the *ij15* lethality.

Requirement of the STCH protein in the assembly of human collagens is going to be studied in collaboration with a group in Manchester. For this, an anti-peptide antibody specific to human STCH has been raised to look for interactions between STCH and human collagens.

Finally, the human *stch* gene could be expressed in *stc-1(ij15)* homozygotes to see if it rescues at least the hypodermal function of the *C. elegans* STC-1. To favour efficient expression of the human gene in *C. elegans*, the *stc-1* promoter and 3' untranslated region should be fused to the stch cDNA coding sequence containing a synthetic intron, designed as a standard *C. elegans* intron. The strain IA432 could be transformed with the test construct and the progeny of stable transgenic lines examined for phenotypic rescue. Rescue of the *ij15* mutant phenotype by the human gene *stch* would indicate that the homologous sequence and similar structure of the *C. elegans* STC-1 and human STCH proteins fulfil a conserved function.

Chapter 5

let-607 encodes a bZIP transcription factor required for cuticle synthesis in _C. elegans_

5.1 INTRODUCTION

Gene transcription is modulated by protein factors, so-called transcription factors, which recognise and bind to short specific DNA sequences near the site where transcription begins. Transcription factors are defined by their structural motifs that can read the specific DNA sequences. A small number of protein structures constitute the DNA-binding domains of eukaryotic transcription factors, including helix-turn-helix, zinc finger, helix-loop-helix and basic leucine zipper.

The basic leucine zipper (bZIP) motif forms an α-helix conformation and is comprised of a basic DNA-binding region adjacent to a dimerisation domain (leucine zipper) (Landschulz *et al.*, 1988). Proteins that bind DNA through such motif are referred to as 'bZIP' transcription factors (reviewed in Hurst, 1996). Sequence similarity between bZIP transcription factors is normally restricted to the typifying motif (55-65 a.a). However, even within this motif a considerable degree of amino acid diversity exists. The position of the bZIP domain also varies. Although usually at the C-terminus, it can be located closer to the N-terminus of the protein. By sequence analysis, bZIP transcription factors have been found to be most similar in the DNA-binding domain, which frequently contains the sequence motif N_ _AA _ _ (C/S) R ten residues upstream the leucine zipper (Hurst, 1996). The most obvious conserved feature in the leucine zipper dimerisation domain is the periodic repetition of the hydrophobic amino acid leucine at every seventh position, with the number of repeats ranging from three up to six. In several proteins, one or more of these leucine residues are replaced by alternative amino acids, generally hydrophobic. In the α-helix (of seven amino acids per two helical turns) the evenly spaced leucine residues lie on the same side of the coil, giving a hydrophobic surface. Thus to be stable in solution, the coil bearing the leucines needs to interact with a matching surface which is provided by a second protein monomer. The structure and size of leucine residues appear to favour optimal packing at the dimer interface. As revealed by structural studies, the bZIP dimer is a pair of continuous α helices that form a parallel coiled coil over the leucine repeat regions but diverge outwards over the basic region helices, allowing grip of the DNA like a pair of molecular tweezers. As expected for proteins that bind DNA as dimers, bZIP factor binding sites show two-fold rotational symmetry (dyad symmetry) and thereby can be thought as two "half sites". Each half site of the binding site is contacted by one of the two basic helices in the dimer. In general the binding sites are either pseudopalindromic (where the two half sites overlap at a central base pair) or palindromic (where the two half sites adjoin one another).

Despite their limited sequence similarity, bZIP transcription factors can be classified into groups on the basis of most resemblance (Hurst, 1996). The majority of bZIP transcription factors form homodimers, however some can function as heterodimers. Heterodimer formation tends to occur between members of the same family group, but it can also happen across family groups. Since each monomer contacts one half site of the DNA-binding motif, heterodimers of proteins with different DNA-binding specificities will recognise a non-dyad symmetric site containing two distinct half sites. This potential to recognise hybrid DNA-binding sites greatly expands the number of gene promoters that can be regulated by bZIP transcription factors (Lamb and McKnight, 1991).

The majority of bZIP factors are activators of transcription (Hurst, 1996). The activity of probably all bZIP proteins is modulated by phosphorylation, via the action of diverse kinases. In many cases the phosphorylation sites are within protein domains involved in transcriptional activation, with phosphorylation enhancing the activity of these bZIP transcription factors. To be functional, all bZIP proteins must translocate to the nucleus upon synthesis. In general, sequences within the basic domain are sufficient to signal transport to the nucleus.

As promoter enhancer-binding factors, bZIP proteins are predicted to bind proteins associated with the RNA polymerase that constitute the transcription initiation complex. In some cases, bZIP factors require adaptor proteins to associate with the transcription initiation complex. Many other protein-protein contacts with bZIP factors involve other transcription factors, allowing cross-talk between different regulatory pathways. The interactions between bZIP factors and other proteins do not necessarily involve the bZIP domain.

5.1.1 *C. elegans* bZIP transcription factors

skn-1, *pha-1* and *xbp-1* products are some of the characterised transcription factors in *C. elegans* that belong to the bZIP family (reviewed in McGhee and Krause, 1997). SKN-1 and PHA-1, however, do not contain all the features that define the bZIP protein family. SKN-1 lacks the leucine zipper motif and binds to DNA as a monomer. This transcription factor has been reported to be required for the correct development of particular stem cells in the four-cell embryo (Bowerman *et al.*, 1992). In the absence of SKN-1, the MS and E daughters of the stem cell EMS undergo mistaken cell fate. Specifically, MS and E produce hypodermal cells instead of pharyngeal and intestinal cells respectively. PHA-1 differs from other bZIP transcription factors in that the basic DNA-binding region is not adjacent to the leucine zipper domain (Granato *et al.*, 1994). It is involved in the formation of the pharynx, possibly coordinating

pharyngeal morphogenesis and terminal differentiation (Schnabel and Schnabel, 1990; Granato *et al.*, 1994).

XBP-1 is homologous to the bZIP transcription factors HAC1 in *S. cerevisiae* and XBP-1 in mammals. As with its mammalian and yeast counterparts, the *C. elegans* XBP-1 acts as a transcriptional activator in the intracellular signalling pathway that is activated by the accumulation of unfolded proteins in the ER, termed the unfolded protein response (UPR) (Shen *et al.*, 2001; Calfon *et al.*, 2002). The UPR bZIP proteins enhance the activation of genes encoding diverse functions of the secretory pathway in order to augment the ER folding capacity, including ER-resident chaperones and folding catalysts (Fewell *et al.*, 2001; Ma and Hendershot, 2001).

As determined below, the mutationally defined locus *let-607* in *C. elegans* corresponds to the predicted gene F57B10.1. F57B10.1 appears to encode a novel transcription factor containing a bZIP domain. The characterisation of *let-607* as well as possible roles of the transcription factor protein are included in this chapter.

5.2 RESULTS

5.2.1 *let-607* is F57B10.1

During RNAi experiments irrelevant to the work presented here (C. Clucas, Glasgow University, Glasgow, UK), it was noted that disruption of the function of the predicted gene F57B10.1 resulted in dead embryos and Dpy progeny, phenotypes possibly consequent from cuticle synthesis defects (reasoned in section 3.3.5). From information available in databases, it was known that the *let-607(h402)* phenotype was rescued with the cosmid clones F57B10 and F48A9. The gDNA contained within these cosmids is contiguous except for a segment of the sequence, which is present at the end of F57B10 and at the start of F48A9. F57B10.1 lies on this overlapping region and therefore was a good candidate gene for *let-607*. In order to test this, I performed phenotypic rescue experiments of *let-607(h402)* with the F57B10.1 sequence. The standard procedure for rescue assays consists of injecting circular molecules of the test DNA and co-injection DNA, including markers. Cloning of F57B10.1 was attempted by different means. Firstly, a 7520 bp fragment containing the gene was amplified from gDNA by PCR. Amplification of this sequence failed frequently so the reaction was repeated several times to obtain the desired fragment. Despite using various cloning vectors and several attempts, insertion of the PCR fragment into a plasmid was not successful. As an alternative, subcloning

of a fragment comprising the gene from the cosmid clones was attempted. Even though more than one culture of bacteria containing the clones was manipulated, the cosmid DNA could not be obtained since the bacteria either did not grow on plates or appeared to have clones with deletions (i.e., the expected cosmid DNA sequence was not intact).

As shown by Kelly *et al.*, (1997), rescue of a mutant phenotype can also be achieved by injecting linear fragments of the test DNA and markers (at low concentrations) together with an excess of digested N2 gDNA. Due to the difficulties in cloning the F57B10.1 sequence into a vector, rescue analysis by this means was performed. The 7520 bp PCR fragment was co-injected with linearised pMW0002 and pMW0003 markers and digested N2 gDNA into *dpy-5(e61) let-607(h402) unc-13(e450)/+ + + I* hermaphrodites (section 2.9). The recipient line was constructed by out-crossing twice KR727 hermaphrodites [*let-607(h402) dpy-5(e61) unc-13(e450) I; sDp2(I;f)*] with N2 males. Presence or absence of Dpy Unc animals in the F_2 progeny indicated if F57B10.1 rescued the *let-607(h402)* phenotype. Heterozygotes for *let-607(h402)* had a wild-type phenotype and could not be distinguished from +/+ siblings at the time of injection and when handling animals containing the array. Injected and transgenic hermaphrodites were cloned individually to screen their offspring and determine their genotype. Moreover, the use of linear DNA for transformation experiments yields a low frequency of transmitting lines. Thus, a large number of animals had to be injected in order to obtain conclusive results. Two transgenic lines initially segregated some Dpy Unc animals. These lines, however, lost the array through generations and the DNA transformation experiments had to be repeated. This time, three transgenic lines with Dpy Unc progeny were established. These Dpy Unc animals were sick, did not produce many progeny, segregated dead embryos that were not transgenic, and could only be maintained in homozygosis for a few generations. These observations indicated that the Dpy Unc progeny resulted from partial rescue of the *let-607(h402)* phenotype by the F57B10.1 array and were not recombinant animals. In addition, Dpy Unc recombinants were expected at very low frequency since they could only be generated if two independent recombinant events occurred in the same chromosome: one between *dpy-5* and *let-607* genes, and the other between *let-607* and *unc-13* genes. These results showed that *let-607* corresponds to the Genefinder predicted gene F57B10.1. F57B10.1 encodes a putative transcription factor containing a bZIP domain.

5.2.2 *let-607* sequence in *h402* and *h189* mutants

The *let-607* predicted sequence consists of seven exons and six introns, and is 4223 bp long. The gene was amplified from the gDNA of *h402* or *h189* mutants by PCR in three different

sections: from exon 1 to exon 3, exon 4, and exon 5 to exon 7. The generated gene fragments were analysed to detect the *h402* and *h189* genetic defects as explained below.

5.2.2.1 Detection of the h402 mutation

Two copies of each *let-607* gene fragment generated from different *h402* mutants were cloned and sequenced. The resulting sequences were compared with that predicted for the wild-type *let-607* gene. A single base change was detected in the two clones carrying the exon1-exon3 fragment. Two other clones containing this region generated from distinct *h402* animals were then sequenced. The same base modification was present in these two clones confirming that the *h402* lesion is a CAA to TAA change at base pair 422 (exon 2) in the gene sequence, which would result in a glutamine to an ochre stop codon modification at amino acid 132 (Q132ochre) (Figure 5.1).

5.2.2.2 Definition of h189 lesion

Two copies of the exon1-exon3 region, three of the exon 4 fragment and two of exon5-exon7 were obtained from different *h189* worms, cloned and sequenced. Comparison of the mutant sequences with the predicted wild-type *let-607* sequence revealed the presence of two distinct base changes in the three cloned exon 4 fragments. *h189* mutants have been produced by EMS treatment (A. Rose, University of British Columbia, Vancouver, Canada). The EMS mutagen generates modifications in the worm genome randomly, and thereby the probability of having two mutations in the same gene is expected to be low. Both base alterations identified in *h189* animals were absent in *h402* mutants suggesting that the predicted wild-type sequence was correct. DNA containing *let-607* exon 4 was obtained from two different N2 larvae as well as from three other *h189* mutants by PCR, then cloned and sequenced. The two base changes were also present in these three *h189* clones but not in the N2 clones, ruling out the possibility of any of these modifications being a misprediction of the wild-type sequence. To sum up, two point mutations were identified in the *let-607* sequences of six different *h189* animals when compared with those of two N2 larvae and the predicted gene sequence. One mutation is a CCG to GCG change at 1809 bp (exon 4), which would result in a proline to alanine substitution at amino acid 200 in the protein sequence (P200A) (Fig. 5.1). The other mutation is at the start of intron 4 (2130 bp) and modifies the consensus 5' splicing signal from GT to AT.

The base modification at the start of intron 4 of the *let-607* gene disrupts the restriction site of the enzyme *Hph* I in *h189* mutants. The above exon 4 fragments analysed by sequencing were
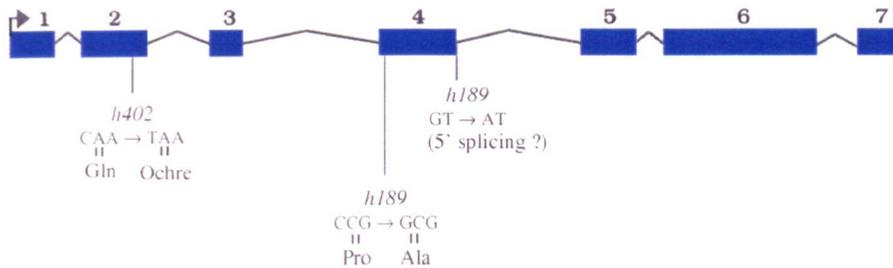
Figure 5.1: Diagram representing *let-607* and the point mutations identified in the *h402* and *h189* alleles. The gene is 4223 bp long and has seven exons (blue boxes) and six introns (grey lines). The *h402* lesion is a CAA to TAA change at 422 bp within exon 2 (Q132ochre). The two *h189* lesions are a CCG to GCG modification at 1809 bp within exon 4 (P200A) and a GT to AT change at 2130 bp, which is the first base of intron 4.
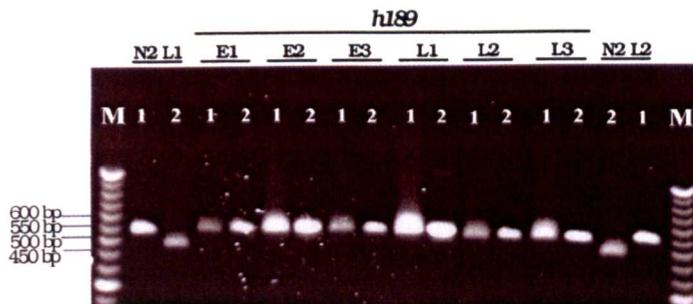
Figure 5.2: *let-607(h189)* polymorphism in N2 and *h189* larvae.

527 bp DNA fragments containing *let-607* exon 4 were obtained from two individual wild-type N2 larvae (named **L1** and **L2**) and six different *h189* animals (**E1**, **E2**, **E3**, **L1**, **L2** and **L3**) by PCR, and were subsequently analysed prior to digestion (**1**) and after digestion (**2**) with the restriction enzyme *Hph* I. The wild-type copy of *let-607* contains an *Hph* I site at the exon4-intron4 boundary. Digestion of the wild-type exon 4 fragments with *Hph* I results in two fragments of 466 bp and 61 bp (**N2L1-2** and **N2L2-2** lanes; only the 466 bp band is shown). The *let-607(h189)* lesion alters the first base of intron **4**, disrupting the *Hph* I site. As demonstrated in lanes **E1-2, E2-2, E3-2, L1-2, L2-2** and **L3-2**, *h189* exon 4 fragments remain intact (527 bp) after digestion with *Hph* I. **M**: 50 bp DNA marker; the size of the relevant DNA marker bands are indicated on the left side of the gel.

digested with *Hph* I and visualised on an agarose gel (Fig. 5.2). The *Hph* I site was present in the two N2 larvae but not in the six *h189* animals tested, thus verifying the nature and position of the mutation.

5.2.3 Analysis of the predicted *let-607* intron-exon structure

A series of PCRs on total cDNA, 3' RACE and 5' RACE assays were performed in order to corroborate the predicted structure of the *let-607* gene. Nine combinations of primers were designed to amplify different regions of the gene by means of PCR. As a whole, these regions covered most of the coding sequence. The primer pairs were tested first on N2 gDNA to study their compatibility and optimal conditions, and then were assayed on N2 cDNA. In all nine cases, the sizes of the PCR products derived from cDNA were as expected suggesting that the predicted exon-intron boundaries were correct. In addition, most of the predicted spliced gene structure is confirmed by the 3' and 5' sequences from cDNA clones (ESTs) available in databases.

Amplification of the 3' and 5' ends of *let-607* cDNA was carried out by PCR as described in section 2.7.4.4. The nested gene-specific primer used to generate the 3' end was designed to amplify the last 442 bp of the gene. Three independent PCR reactions were carried out with this primer. Two of these generated products of approximately 1.2 Kb; the other reaction resulted in the amplification of multiple fragments ranging from 500 bp to 600 bp. The PCR products were cloned and sequenced (two clones containing the 500-600bp fragments; one carrying the 1.2 Kb product). The cloned 3' ends were identical to the predicted terminal region of the gene. Therefore, the different sizes of the amplified fragments were due to a variety in length of the 3' non-coding region present before the poly-A tail.

A gene specific primer within exon 2 (235 bp from the predicted initiation codon) was utilised to generate the 5' end of the gene. The 5' RACE method was carried twice, but without success in amplifying the desired final product. A fragment of about 235 bp was obtained in a third attempt of the assay (Fig. 5.3). Two clones of this PCR product were sequenced and compared with the predicted N-terminal region of *let-607*. Only a small part of the amplified sequence corresponded to that predicted. BLAST searches were performed with the remaining part of the sequence distinct to *let-607* (about 129 bp) in the *C. elegans* genome and EST database. This region did not match any other *C. elegans* sequences, which would suggest that the isolated and sequenced PCR product was an artefact.

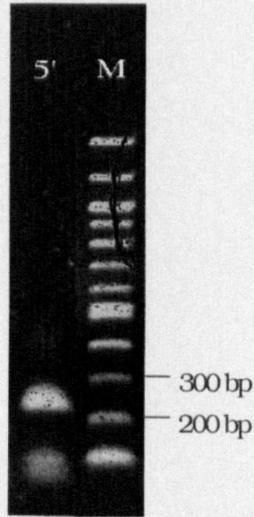**Figure 5.3:** *let-607* 5' RACE PCR product employed for sequencing.
Lane 5': 5' RACE PCR product of *let-607*. The band is approximately 235 bp
in size. **M:** 100 bp DNA marker. The relevant band sizes are indicated on the
right of the gel.

Many *C. elegans* transcripts are *trans* spliced at their 5' end to a short leader sequence, SL. There are two common SL sequences, SL1 and SL2; SL1 is most commonly found in mature mRNAs. Primers specific to *let-607*, SL1 and SL2 sequences were designed to amplify a short 5' region of the gene. PCRs were performed on cDNA using either of these SL primers and the *let-607* primer. Non-specific amplification was obtained in both cases, thus it was not possible to determine if SL1 or SL2 *trans* splicing occurs on the *let-607* transcript.

5.2.4 *let-607* expression pattern

The temporal and spatial expression of *let-607* was studied by RT-PCR analysis and reporter-gene fusion techniques respectively. Temporal expression studies were essentially carried out as explained in section 4.2.4.1 using oligonucleotides specific for the *let-607* and *ama-1* genes. In relation to the *ama-1* control products, *let-607* transcripts were present at a similar abundance in all larval and adult stages. The *let-607* gene is thus temporally constitutively expressed throughout the worm's post-embryonic development.

A *let-607*::reporter gene was constructed to identify which cell types express *let-607*. A 2119 bp fragment comprising the first 115 bp of *let-607* and the 5' region was fused in frame to the *gfp* gene within the pPD96.04 expression vector. The resulting construct was then introduced into N2 hermaphrodites by microinjection (see sections 2.9 and 2.10). *let-607* expression was first observed in hypodermal cells of comma-stage embryos, prior to elongation (see section 3.1.1), and continued to be on in later stage animals (Fig. 5.4 A and C). Expression in other cells was detected at post-embryonic stages, including cells within the pharynx and possibly nerve cells (Fig. 5.4D). GFP production from the above *let-607* construct corroborated the predicted start of the gene in study.

5.2.5 LET-607 localisation studies

An antibody reactive to LET-607 was obtained using the commercial antibody production facility Abcam Ltd., UK (http://www.abcam.com). Two distinct regions of the *let-607* gene encoding 14 amino acids were chosen such that they did not share homology to other *C. elegans* proteins. Both 14 aa segments were made synthetically and used simultaneously to immunise two rabbits. A series of blood samples were taken from each rabbit prior to isolating the antibody: one pre-immunisation and three after immunising the animals. Reactivity of the serum from each blood sample was tested against the peptides; the antibody was affinity-purified from the final serum showing the best response.

Figure 5.4: *let-607* **expression.**

Left panels are fluorescence images; right panels are DIC micrographs. **A-B,** comma-stage embryo. Expression of *let-607* is restricted to hypodermal cells (**A**). **C-F,** L1 larvae. **C** shows expression of *let-607::GFP* in hypodermal cells. A wide range of cells in the head express the *let-607* reporter construct, including pharyngeal cells and possibly nerve cells (**E**).

Facing page 139
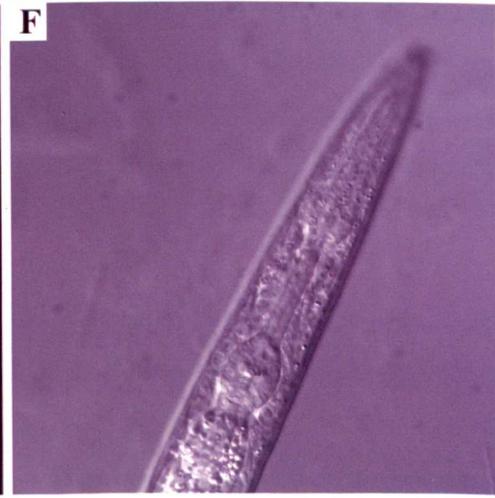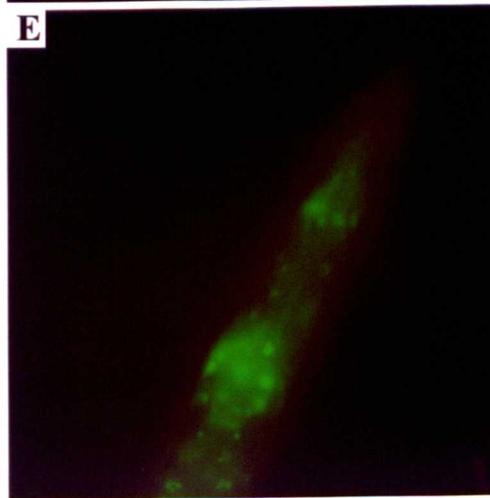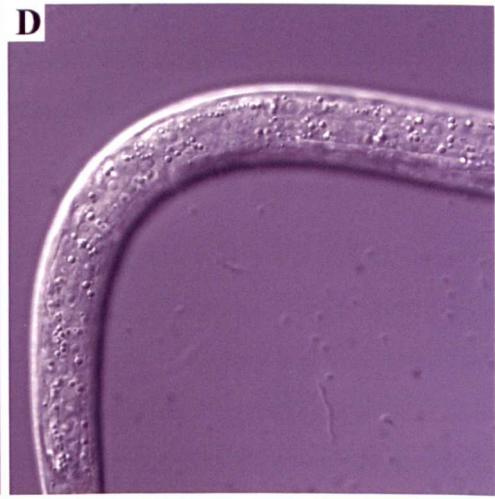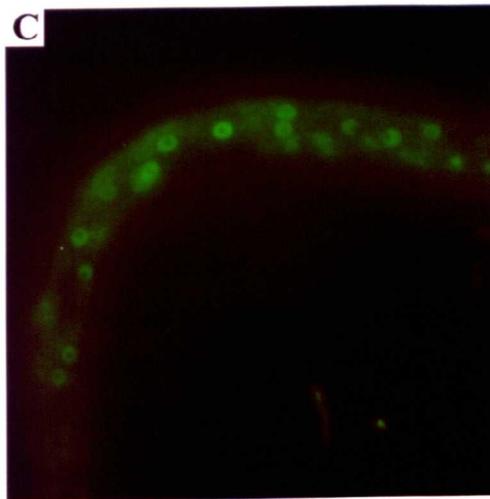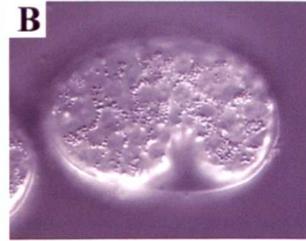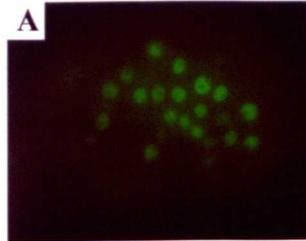
N2 embryos at different developmental stages were collected and prepared for immunodetection of the LET-607 protein as indicated in section 2.13. The affinity-purified anti-LET-607 antibody (primary antibody) was first tested at dilutions 1/50, 1/250 and 1/1000, using the alexa fluor 594 anti-rabbit secondary antibody at 1/200 dilution. As a control, the primary antibody was omitted in the procedure to determine any non-specific binding of the secondary antibody. Specific staining was not detected at any of the assayed concentrations of the primary antibody. The study was repeated using the primary and secondary antibodies at the above concentrations but the same negative results were obtained. The affinity-purified anti-LET-607 was finally employed at a 1/5 dilution with either the above alexa fluor 594 anti-rabbit or the alexa fluor 488 anti-rabbit secondary antibodies. In these cases, the incubation time, temperature and blocking method were also modified. The treatment with the primary antibody was performed overnight at 4°C instead of 2 h at room temperature. A combination of 10% goat serum and 1% dried milk were utilised as blocking agent; a solution containing only 1% dried milk had been used previously. Specific staining was not obtained with the affinity-purified anti-LET-607 at high concentration either.

Sera of the various blood samples from the immunised rabbits contain the affinity-purified anti-LET-607 antibodies as well as low-affinity anti-LET-607 antibodies, which were not isolated under the conditions employed in the purification process. The second bleed and the final bleed sera of the two rabbits were tested at 1/200 dilution using the secondary antibody alexa fluor 594. Some staining was noted on the samples but could not be defined due to strong background fluorescence. More analyses were carried out with both sera of each rabbit and either the alexa fluor 594 or the alexa fluor 488 secondary antibodies. The same modifications in the incubation time, temperature and blocking method mentioned above were also applied in these attempts. An improvement of the staining was not observed with any secondary antibody under the changed conditions.

As a positive control, staining with the previously characterised anti-CDC-25 antibody (Clucas *et al.*, 2002) was performed in parallel to my analyses with the affinity-purified antibody and the polyclonal sera. The expected fluorescent pattern corresponding to the CDC-25 protein was detected, validating the reagents and the permeabilisation step of the procedure as well as indicating that the secondary antibodies used were functional. Since the experiments were working optimally, it was assumed that specific staining was not obtained because the affinity-purified anti-LET-607 antibody and the anti-LET-607 polyclonal sera did not recognise the LET-607 protein under the employed conditions. Not all primary antibodies have the same

ability to react with fixed antigens. Although the method followed is generally appropriate for *C. elegans* immunolocalisation studies, other fixation procedures could result in a better preservation and/or exposure of the LET-607 antigenic epitopes to the antibody.

5.2.5.1 *Western blotting*

A western blot assay was carried out in order to check if the polyclonal antibody did react with the LET-607 protein. Four samples of *C. elegans* whole protein extract were loaded next to a protein marker on a polyacrylamide gel. Samples and protein markers were then transferred from the gel to a nitrocellulose membrane, which was left in blocking solution overnight at room temperature. Four different solutions containing the purified antibody, the antibody and either synthetic LET-607 peptide, or the antibody with both peptides, were prepared and placed at 4°C overnight. The antibody was at 1/500 dilution in all samples; the synthetic peptides were at 20 µg/µl final concentration in the corresponding solutions. The nitrocellulose membrane was cut into four strips, each containing an electrophoresed protein sample and a protein marker, which were incubated separately with the primary antibody solutions. An anti-rabbit secondary antibody conjugated with the enzyme alkaline phosphatase was then used at 1/15000 dilution. The membrane strips were finally treated with a solution containing the alkaline phosphatase substrate. The activity of this enzyme on its substrate produces a detectable colour change, allowing detection of any antibody-protein complexes present on the membrane. No bands were detected on the whole protein extract incubated with the anti-LET-607 antibody, suggesting that the antibody did not recognise the LET-607 protein. However, small quantities of LET-607 in relation to other worm proteins in the extract or the use of an insufficient amount of the antibody in the study could also result in failure to detect a protein species. In addition, an alternative secondary antibody could be used to ensure that the system of detection is sensitive enough. No further investigations were carried out in this respect and therefore answers to these other possibilities cannot be provided.

5.2.5.2 *Construction of a LET-607 expression vector*

A recombinant system was initiated in order to obtain the LET-607 protein to raise another antibody. This system implies the construction of an expression clone containing a region of the coding sequence of the gene, expression of the fusion protein and purification by chromatography. In order to make the construct, a 2087 bp PCR fragment comprising the whole *let-607* coding sequence was inserted 3' to the 6 x His tag in the pEQ-30 expression vector (see section 4.2.5 for more details). The resulting construct was introduced into *E. coli* cells,

followed by screening and isolation of transformants. Two different transformants were grown and induced with 1 mM IPTG for 3 h at 37°C for expression of the cloned recombinant protein. Induced cells were then lysed and harvested to obtain clear lysates containing cellular soluble proteins. Presence of the fusion protein in both induced cells and clear lysates was analysed on the SDS-PAGE gel shown in Figure 5.5. The expected 75.653 kDa LET-607 protein was not observed either in the clear lysates or in the content of induced cells. The latter indicated failure in expressing the recombinant protein, which could result from not having employed the optimal induction conditions or from an abnormal insert-vector fusion. Despite not detecting the recombinant LET-607 protein in samples from induced cells, the lysed cellular pellet of one of the two above transformants was treated with 8 M urea for one hour at 37°C (see section 4.2.5). LET-607 was also absent from the treated lysed pellet, corroborating that the recombinant protein had not been produced (Fig. 5.6).

5.2.6 *let-607* RNAi effects

The *let-607* gene function was inhibited by RNAi feeding in a range of strains: N2, NL2099, DR466, JR667 and TP12. The RNAi effects were first tested at the temperatures 15°C, 20°C and 25°C on N2 animals. No obvious differences were observed between the RNAi progenies raised at the three temperatures, so the following RNAi experiments were carried out at 20°C unless otherwise specified.

Interference of a gene function by RNAi feeding involves the following: construction of a vector expressing dsRNA of the gene of interest, transformation of a specific *E. coli* strain with the construct, and induction of the expression system prior to feeding the worms. In order to generate the expression construct, a 1020 bp fragment of the *let-607* gene was amplified from gDNA by means of PCR. The generated PCR fragment was then inserted into the L4440 (pPD129.36) vector to proceed as outlined above. The progenies of RNAi-treated animals appeared not to display any obvious defects. The RNAi experiments with this construct were carried out twice at the temperatures 15°C, 20°C and 25°C. The integrity of the clone was checked by PCR as well as by restriction analysis. The L4440 plasmid has two convergent T7 polymerase sequences at each side of the cloning site, which results in transcription of both strands of the DNA insert upon induction of T7 polymerase expression. PCRs carried out with the clone and T7 primers generated a fragment of the same size as the insert, suggesting that the T7 sequences and *let-607* were intact. A series of digestions of the clone were performed with a combination of restriction enzymes. In all cases, the digest products were as expected indicating
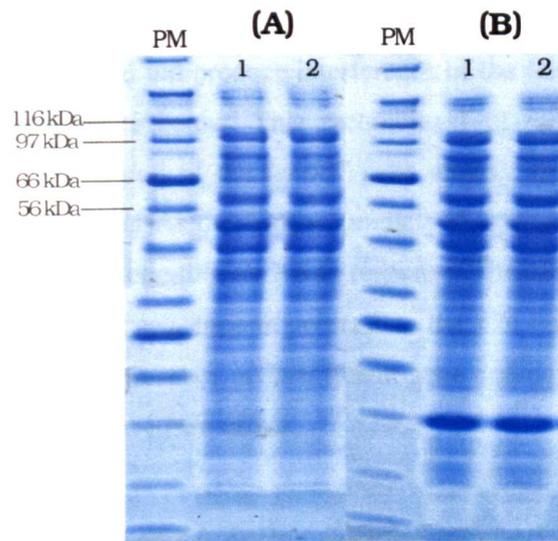
Figure 5.5: Protein content of induced cells (A) and in clear lysates (B).
Two different cells (numbered **1, 2**) harbouring the expression clone were tested
in parallel. The expected size of the recombinant LET-607 protein was 75.653 kDa.
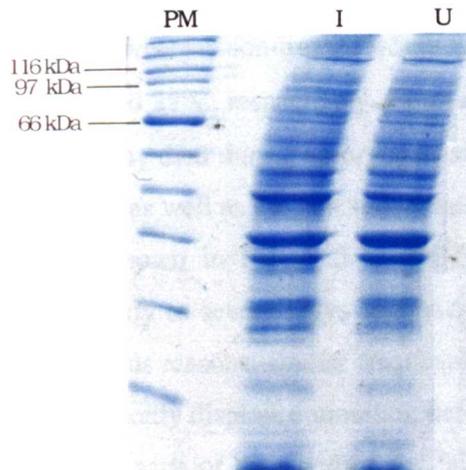Relevant protein markers (**PM**) are indicated on left side of the gel.

**Figure 5.6: Protein content of induced cells (I) and in lysed cell pellets
after treatment with 8 M urea (U). PM:** protein marker.

that the complete *let-607* insert was present in the plasmid. The results from these tests ruled out the possibility that the lack of an RNAi response was due to an aberrant construct.

A different construct assayed did produce interference of the *let-607* gene and it was the clone used in all the experiments below. This recombinant plasmid (provided by J. Ahringer) contained 1194 bp of a region of the *let-607* gene distinct from that used in my construct. Current knowledge of the RNAi mechanism predicts that long segments of coding sequence should be more effective in eliciting an RNAi response. The clone producing RNAi of *let-607* comprised 1017 bp of exon sequence in contrast to the 424 bp present in the first construct. Thus it is possible that the presence of insufficient gene coding sequence could have been the cause of failure in eliciting an RNAi effect.

5.2.6.1 Reduction of let-607 function during embryogenesis by RNAi

N2, DR466 and NL2099 animals at L3-L4 stage were fed with bacteria expressing *let-607* dsRNA in order to knockdown the gene during embryogenesis of their F_1 progeny. Bacteria carrying the L4440 vector without insert was used as control. I scored and analysed the broods of N2 treated hermaphrodites at 20 °C (Table 5.1A) and 25 °C (Tables 5.2A). Tables 5.1B and 5.2B show the progenies of non-treated animals (fed with the control bacteria) at 20°C and 25°C respectively. At both temperatures, the offspring of treated N2s was significantly less (about one third) than that of control worms. RNAi hermaphrodites had an average brood size of 77 ($n=$ 23) at 20°C and 88 ($n=24$) at 25°C, whereas non-treated worms produced on average 253 ($n=5$) and 260 ($n=4$) F_1 animals at 20°C and 25°C, respectively. Approximately 27% (at 20°C) and 15.1% (at 25°C) of the RNAi F_1 progeny died during embryogenesis; the remainder hatched as small, fat larvae. These dead embryos, as well as the hatched larvae, phenocopied the different classes of *let-607(h402)* mutants as shown in Figure 5.7. The RNAi larvae, in contrast to *h402* hatchlings, tended to grow slowly to severe Dpys with reduced motility. Most of these Dpys died before adult stage for various reasons; similar frequencies were observed at the two above temperatures tested. They commonly displayed moulting defects (Fig. 5.8 A and B): remains of old cuticle attached to different parts of their bodies and failure to moult (normally at the L4-adult moult). Some of them burst, through the vulva (Fig. 5.8C) or less frequently through the anus (Fig. 5.8D). The few Dpys that attained adulthood had obvious cuticular defects such as branched and broken alae plus discontinuous or absent annuli, as well as morphological defects of the vulva (Fig. 5.9B). In some cases the reasons of death were difficult to determine under a stereoscope.

Figure 5.7: *let-607* RNAi progenies phenocopy *let-607(h402)* mutants.

A-F, Nomarski micrographs. Right panels show the various types of *let-607(h402)* homozygotes: class I **(B)**, class II **(D)** and class III **(F)**. Left panels are *let-607* RNAi affected animals. **A**, **C**, and **E** correspond to **B**, **D** and **F**, respectively.

Facing page 145
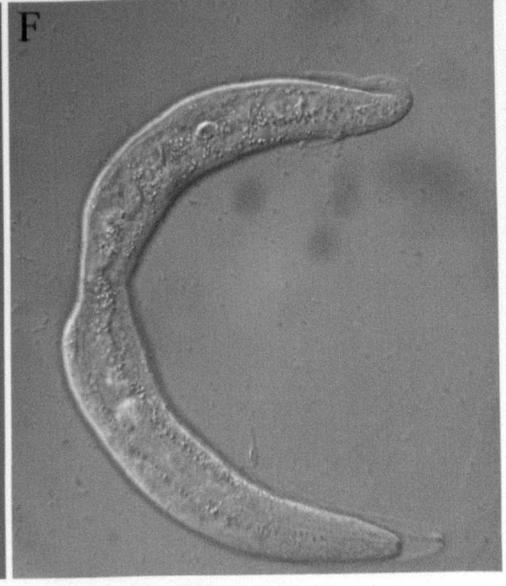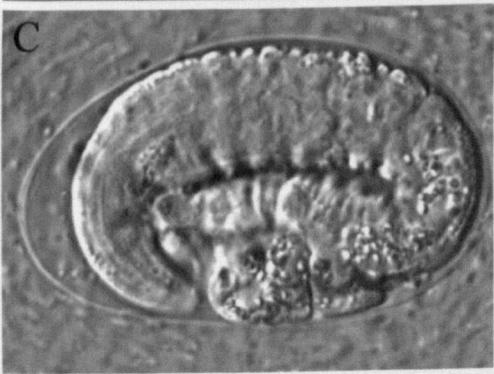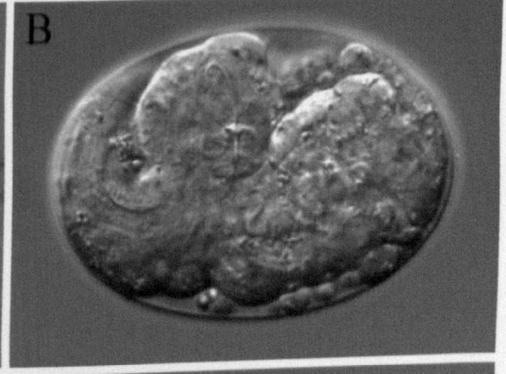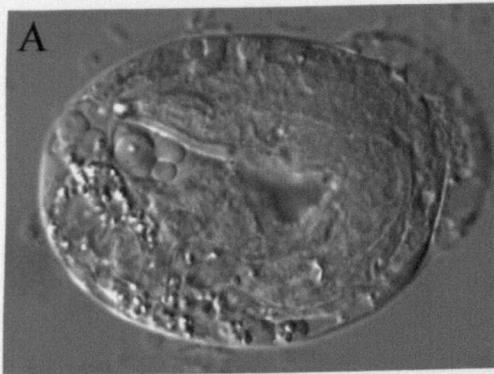
Figure 5.8: DIC micrographs of *let-607* RNAi Dpy worms.

Failed moult of an L4 **(A)** and of a younger larva **(B)**. Both larvae have constrictions at distinct points of their bodies caused by unshed previous cuticles. **C-D,** disruption through the vulva and the anus, respectively.
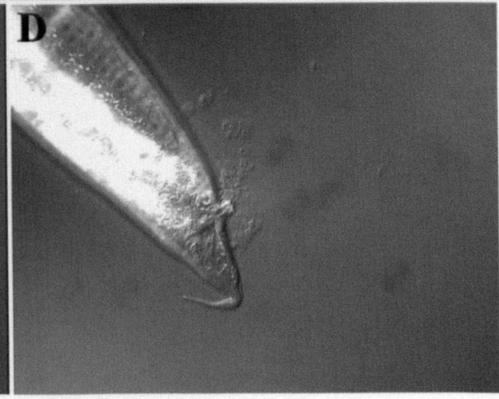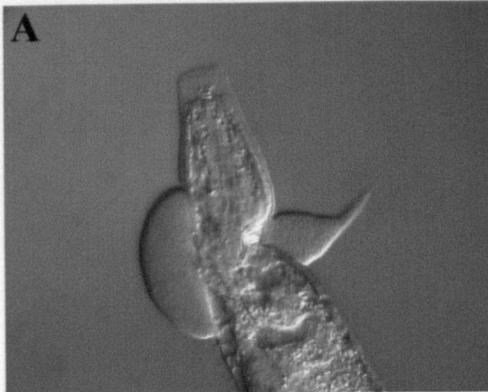
Figure 5.9: SEM images of non-treated and *let-607* RNAi adults.

A, non-treated N2 hermaphrodite. **B,** branched and discontinuous alae, disrupted annuli and abnormal vulva of an RNAi N2 Dpy hermaphrodite. **C,** non-treated DR466 male. **D,** RNAi DR466 Dpy male with severe cuticular defects on the body and tail.

Facing page 147
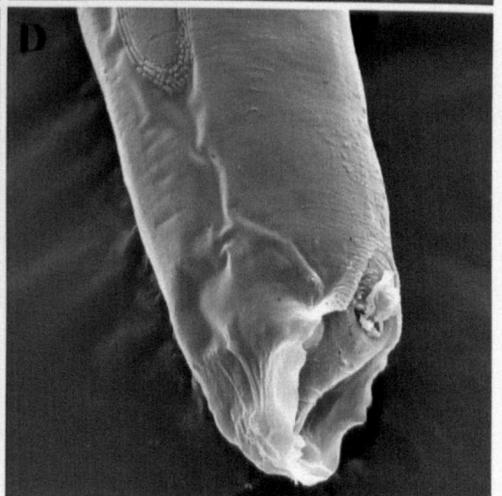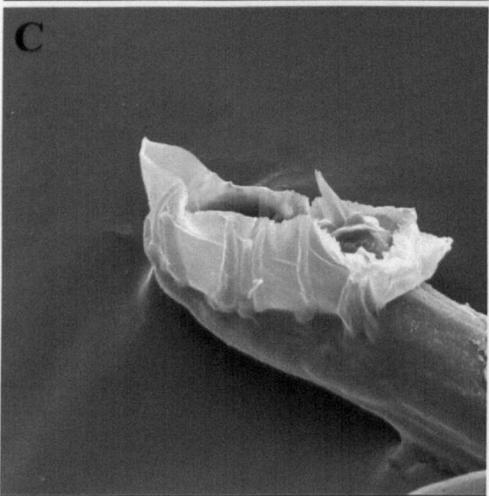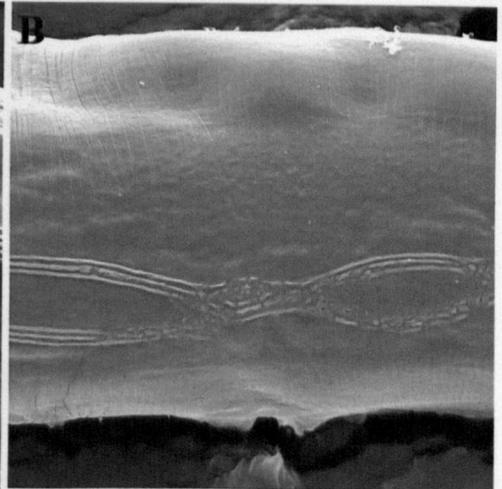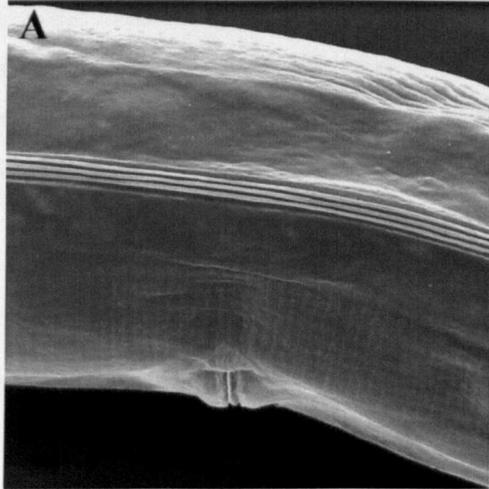
The DR466 strain is homozygous for the recessive mutant allele *him-5(e1490) V*, which results in segregation of approximately 33% males. The progeny of DR466 treated hermaphrodites was essentially as described above. The few Dpy males that developed to adulthood presented aberrant tails as depicted in Figure 5.9D.

let-607 knockdown effects were more severe with the RNAi sensitive NL2099 strain. RNAi-treated NL2099 hermaphrodites segregated some embryonic lethality and small, fat larvae normally with swellings and constrictions along the body. These affected larvae were similar to *let-607(h402)* class III mutants, except that they hardly moved on the plates and tended to curl. In contrast to the N2 RNAi larvae, most of the NL2099 treated progeny did not develop notably and died at an early larval stage. Those animals that survived early development outgrew the phenotype and arrested at late larval stages.

(A) *let-607* RNAi on N2 hermaphrodites at 20°C

N2 animal	dead embryos	Reasons for death of Dpy progeny			Dpy adults	slightly Dpy/WT adults	Non-Dpy larval deaths	total progeny
		moult defects	burst	other/ND				
1	9	41	10	6	5	0	0	71
2	41	35	6	6	5	0	5	98
3	15	29	11	17	5	5	3	85
4	11	42	8	22	4	0	0	87
5	0	34	8	2	3	0	0	47
6	11	33	4	16	1	1	5	71
7	32	36	0	12	4	0	1	85
8	28	25	5	18	9	0	5	90
9	61	31	8	9	3	0	5	117
10	19	29	2	8	3	2	0	63
11	21	32	7	10	15	1	4	90
12	34	67	16	19	7	2	7	152
13	7	7	2	4	0	2	1	23
14	27	39	9	6	6	0	3	90
15	5	22	4	1	2	0	1	35

N2 animal	dead embryos	Reasons for death of Dpy progeny			Dpy adults	slightly Dpy/WT adults	Non-Dpy larval deaths	total progeny
		moult defects	burst	other/ND				
16	48	51	11	12	6	0	3	131
17	12	53	9	5	6	0	5	90
18	70	48	4	7	6	0	6	141
19	11	34	7	9	8	1	0	70
20	3	21	9	5	4	3	2	47
21	2	4	0	0	0	3	0	9
22	2	26	4	5	5	0	1	43
23	10	19	2	7	0	1	1	40
Total	479	758	146	206	107	21	58	1775
Percentage (%)	27	42.7	8.2	11.6	6	1.2	3.3	100

(B) non-treated N2 hermaphrodites at 20°C

N2 animal	total progeny	dead embryos	percentage of death (%)
1	264	19	7.2
2	195	1	0.5
3	324	1	0.3
4	275	2	0.7
5	209	3	1.4
Total	1267	26	2

Table 5.1: Offspring of let-607 RNAi (A) and non-treated (B) N2 hermaphrodites raised at 20°C. The total broods and dead embryos are given in both cases. The types of progeny and causes of death of Dpy larvae as well as the corresponding frequencies are indicated in A. Of the reasons for death of Dpy larvae, moulting problems and burst were obvious. Other cases that were difficult to determine or occurred occasionally are stated as other/ND (non-determined).

(A) *let-607* RNAi on N2 hermaphrodites at 25°C

N2 animal	dead embryos	Reasons for death of Dpy progeny			Dpy adults	slightly Dpy/WT adults	Non-Dpy larval deaths	total progeny
		moult defects	burst	other/ND				
1	8	62	44	19	3	2	3	141
2	37	54	35	10	3	2	5	146
3	6	34	7	18	6	5	3	79
4	44	58	25	15	3	0	3	148
5	6	28	15	0	2	0	0	51
6	5	35	9	2	0	0	1	52
7	19	45	26	6	7	1	8	112
8	1	14	1	2	2	0	2	22
9	11	36	24	11	6	0	6	94
10	26	52	37	9	4	0	2	130
11	8	38	20	9	0	0	3	78
12	14	61	39	13	1	0	2	130
13	10	40	18	8	4	0	4	84
14	11	43	13	3	3	0	3	76
15	5	14	9	1	1	0	0	30
16	28	73	36	16	0	1	3	157
17	8	26	7	5	3	2	2	53
18	19	46	20	8	4	1	5	103
19	16	35	13	7	3	0	3	77
20	6	18	6	2	1	1	1	35
21	7	38	17	0	2	2	3	69
22	14	86	31	1	7	2	4	145
23	3	22	6	6	1	0	0	38
24	8	33	15	2	6	1	2	67
Total	320	991	473	173	72	20	68	2117
Percentage (%)	15.1	46.8	22.3	8.2	3.4	0.9	3.2	99.9

(B) non-treated N2 hermaphrodites at 25°C

N2 animal	total progeny	dead embryos	percentage of death (%)
1	323	2	0.6
2	277	0	0.0
3	190	0	0.0
4	251	0	0.0
Total	1041	2	0.2

Table 5.2: Progeny of *let-607* RNAi (A) and non-treated (B) N2 hermaphrodites raised at 25°C.
The other/ND (non-determined) class given as a reason for death of Dpy progeny corresponded to larval death that could not be defined or to defects that occurred at low frequency.

5.2.6.2 Analysis of the seam cells and cuticles of RNAi-treated animals

let-607 RNAi by feeding was performed on JR667 L4 hermaphrodites in order to analyse the seam cells of the RNAi Dpy progeny. JR667 contains the genomic insertion *wIs51*, which consists of a seam cell promoter gene::*GFP* marker fusion that results in restricted GFP expression to the seam cells. The position and number of seam cells in the RNAi Dpy animals was similar to that of the JR667 larvae. JR667 GFP fluorescence dims and disappears at early adult stage. Thus, it was not possible to examine the seam cells of RNAi Dpy adults with disrupted alae using this method.

The distribution of the collagen DPY-7 protein was examined in *let-607* RNAi N2 adults and *let-607* RNAi NL2099 larvae by immunofluorescence with the DPY7-5a antibody (McMahon *et al.*, 2003). As compared with the wild-type pattern (Fig. 5.10A), the DPY-7 protein tended to be absent or assembled abnormally into discontinuous bands in the cuticles of the RNAi N2 and NL2099 animals (Fig. 5.10 B and D, respectively). Interestingly, intracellular accumulation of the DPY-7 protein (presumably in the hypodermal cells) was observed in these animals (Fig. 5.10 C and E). It should be noted, however, that the localisation of the DPY-7 protein was variable in *let-607* RNAi NL2099 larvae. It varied from being accumulated in the cytoplasm of intracellular cells and absent or disorganised in the cuticle to resembling the wild-type pattern.

Figure 5.10: Immunofluorescence images of the DPY-7 distribution in untreated and *let-607* RNAi worms. A, mid-stage untreated N2 larva. The DPY-7 protein is typically secreted and assembled as regularly spaced bands. **B** and **C,** different focal planes of a *let-607* RNAi N2 adult. The DPY-7 bands are fragmented and discontinuous or missing in the cuticle surface of the worm **(B).** Cytoplasmic accumulation of the DPY-7 protein in internal cells **(C). D** and **E,** different focal planes of a *let-607* RNAi NL2099 larva. The DPY-7 bands are malformed or absent in the cuticle surface of the larva **(D).** Intracellular accumulation of the DPY-7 protein **(E).**
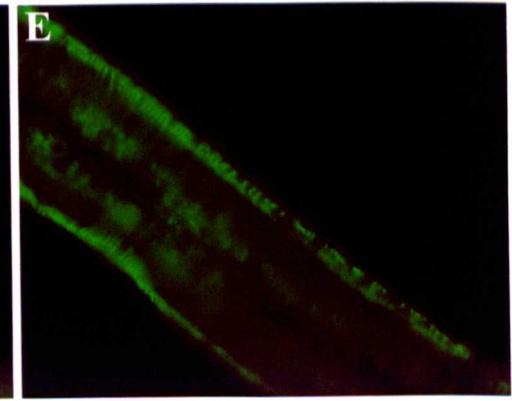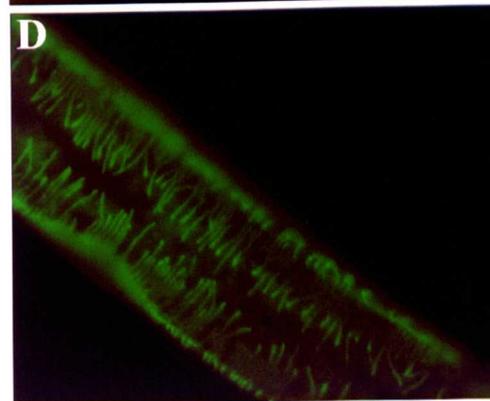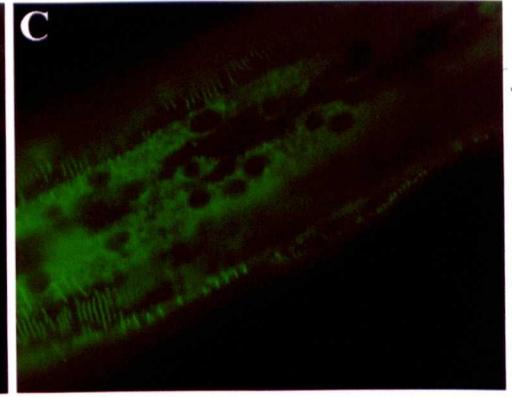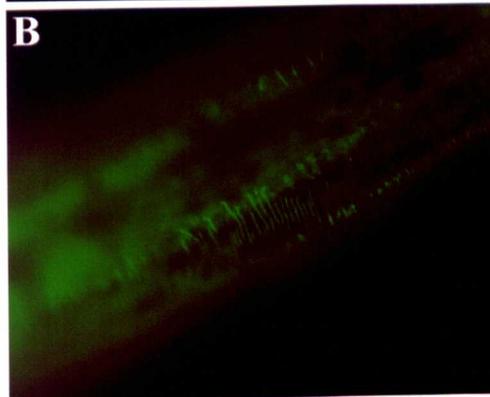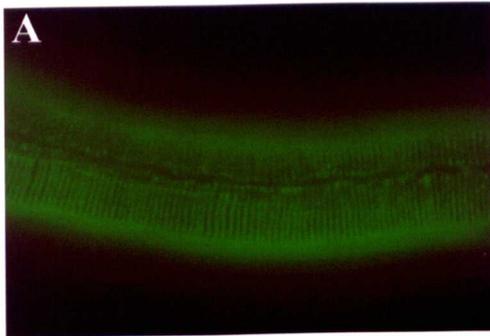
The strain TP12 contains a collagen gene (*col-19*)::*GFP* transgene inserted in the genome. The COL-19::GFP is expressed in the annuli and alae of adult worms as exemplified in Figure 5.11 A and B. TP12 L4 hermaphrodites were fed with bacteria expressing the *let-607* dsRNA to examine the distribution of COL-19 in a *let-607* knockdown background. As can be seen in Figure 5.11 C and D, the RNAi Dpy progeny that attained adulthood displayed a severely disrupted COL-19 pattern. The collagen tended to be in the form of fibres or fragments in both annuli and alae, which frequently bifurcated at various points (Fig. 5.11D).

The same assay to test the strength of the cuticle explained in section 3.4.1.7 was carried out with N2 RNAi Dpy larvae at L2-L3 and L4 stages. Approximately 64.7% (n=34) of L2-L3 larvae burst in dH$_2$O, ~91.7% (n=24) in 1% β-ME and ~96.3% (n=27) in 1% SDS. L4 larvae ruptured through the vulva at a frequency of ~57.9% (n=19) in dH$_2$O, ~68.7% (n=16) in 1% β-ME and 21% (n=19) in 1% SDS. In contrast to the RNAi Dpy animals, the control N2 worms used in the RNAi analyses did not burst in any of the above solutions. Worms lost progressively the characteristic colouration of their body contents when placed in SDS, apparently due to the dissolution of their body contents by this agent. The effect was visibly greater in RNAi larvae than in non-treated worms. The fact that tested larvae tended to burst in the three solutions when wild-type animals retained their body shape confirmed the inability of RNAi-treated animals to produce a functional cuticle when LET-607 was depleted during embryogenesis.

5.2.6.3 Inhibition of let-607 during post-embryogenesis by RNAi

N2 and NL2099 embryos were placed separately on *let-607* RNAi plates to inhibit the gene function after embryogenesis. Plates containing bacteria harbouring the L4440 vector were used as controls. N2 hatched larvae developed to adulthood slightly slower than control animals. The RNAi adults normally produced less progeny than controls, and a few sterile animals were detected as well. The treated adults were either wild type in length or slightly Dpy, but more frequently slightly small. They burst through the vulva or became progressively sick and died.

NL2099 hatched larvae developed more slowly and were obviously smaller than control animals. They became noticeably Dpy at the L4 stage and displayed various defects including moulting problems (Fig. 5.12), shrinking when poked, rupturing through the vulva and general sickness; some L4 larvae died for reasons difficult to determine. All the treated animals failed to develop as normal adults.

Figure 5.11: GFP images of the COL-19 pattern in wild-type TP12 and *let-607* RNAi Dpy adults.
A-B, typical COL-19::GFP localisation in the cuticle of a TP12 adult. The collagen protein is present in the annuli (arrow) and alae (double-headed arrow). C-D, altered cuticular arrangement of COL-19 in RNAi-treated Dpy TP12 adults. Arrows in D indicate the bifurcation of the alae.

Facing page 154
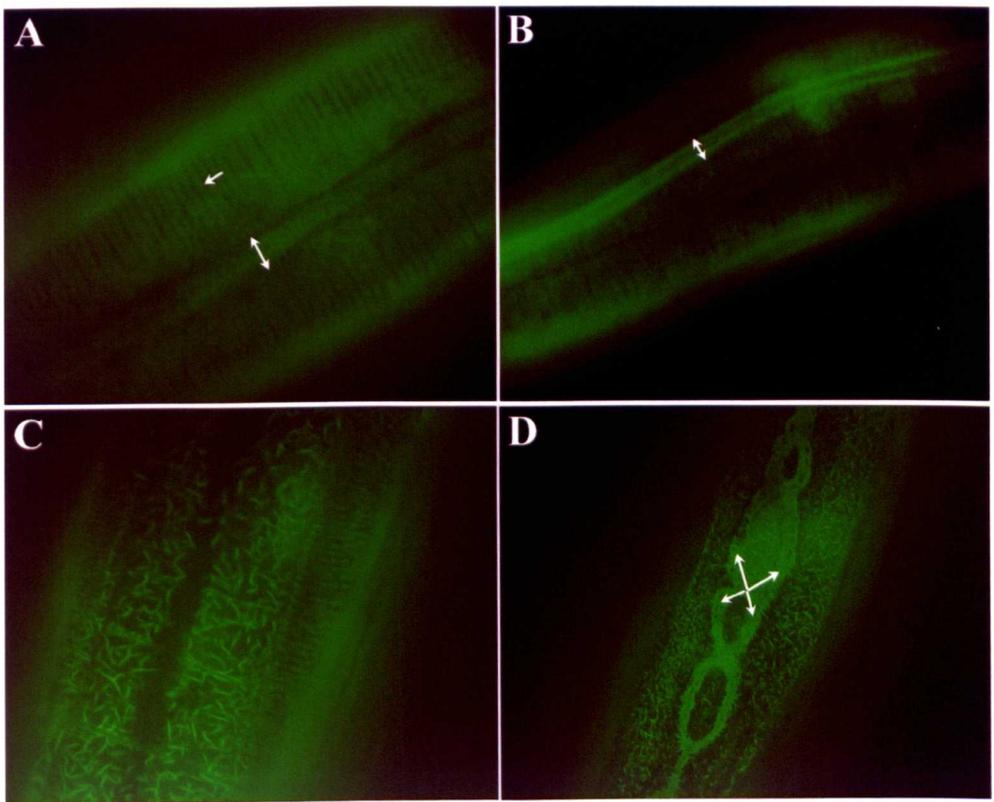
Figure 5.12: Defects of NL2099 worms upon interference of *let-607* post-embryonic function.

Left images (**A** and **C**) are *let-607* RNAi NL2099 L4-young adults treated post-embryonically; right image (**B**) represents a non-treated NL2099 young adult. All images were taken at the same magnification. *let-607* RNAi NL2099 animals (**A** and **C**) tend to be more fat and significantly shorter than non-treated NL2099 worms (**B**) at the L4-adult stages. Animals in **A** and **C** are attempting moult as inferred by partial detachment from the old cuticle (arrowheads). Body constrictions and remains of previous cuticles (arrow) can be seen in **C**. The body contents of the animal in **C** are grossly affected, probably due to general cell death.
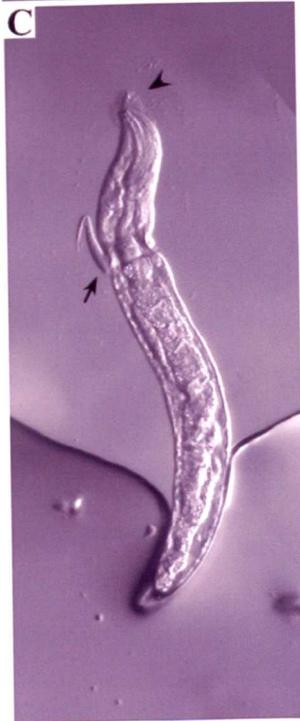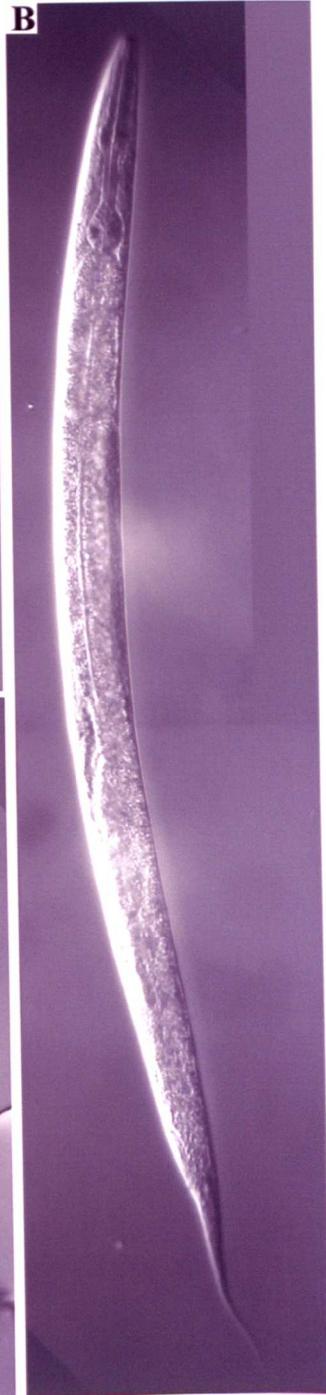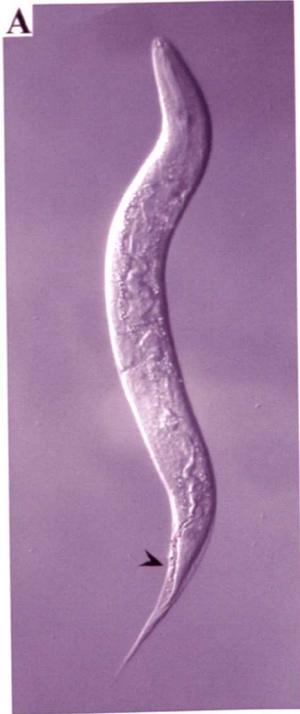
5.3 DISCUSSION

Chapter 3 deals in detail with the phenotypic analysis of two distinct mutant alleles of the *C. elegans* gene *let-607*: *h402* and *h189*. The identification of the *h402* and *h189* genetic alterations has been included in this chapter, where extensive molecular characterisation and further functional analysis of *let-607* can be found.

5.3.1 *let-607(h402)* and *let-607(h189)* may be null and hypomorphic mutants respectively

The *let-607* lesions *h402* and *h189* cause distinct phenotypic defects that differ in severity, as determined by Nomarski microscopy. *let-607(h402)* mutant traits are variable and occur after embryonic elongation. Approximately 61% of *let-607(h402)* embryos die at the three-fold stage retracting noticeably in length, or less frequently, rupturing along the ventral side of their body with the consequent loss of the elongated shape. The remainder ~39% hatch as small, fat larvae, which normally display bulges and constrictions randomly about the body. *let-607(h402)* hatchlings arrest at the L1 stage as inferred by the detection of a four-cell gonad primordium, mononucleated intestinal cells, about ten seam cells on each side of the larvae, alae and general failure to moult. They do not undergo perceptible growth and lose movement progressively. Additionally, *let-607(h402)* mutant embryos and hatchlings can have abnormally shaped tails, frequently "split tails".

The *let-607(h189)* mutation gives ~1% embryonic death and ~99% larval death. The frequency of *h189* embryonic death is practically negligible considering that N2 wild-type hermaphrodites produce occasional dead eggs. In contrast to *let-607(h402)* hatchlings, *let-607(h189)* newly hatched larvae do not display noticeable abnormalities in body shape and are not distinguishable from wild-type L1s. They appear to develop normally to L2 larvae and arrest at this stage as inferred by the presence of an L2-stage gonad, binucleated intestinal cells, more than ten seam cells on each side of the animal and lack of the cuticular alae. *let-607(h189)* arrested larvae lose motion progressively, rest curving their body as a loop and begin to lose the normal colouration of the internal body contents. Many of the larvae that are about to die shrink suddenly in response to a touch on the head.

The *h402* lesion identified would generate a stop codon at amino acid 132 in the protein sequence (Q132ochre) and is possibly a null mutation, as the predicted truncated product would lack most of the amino acid sequence including that of the putative bZIP domain. On the other hand, the *let-607(h189)* allele was found to contain two distinct nucleotide lesions. One would

result in a P200A alteration in the protein sequence, and the other is a modification of the consensus 5' splicing signal of intron 4. It is not known which of these two lesions causes the *h189* mutant phenotype. Although proline and alanine are both non-polar residues, and hence tend to localise to the interior of protein structures, they have distinctive roles in protein conformation. Proline residues reduce the structural flexibility of the protein and prevent hydrogen bonding as opposed to alanine residues, which are important in promoting the formation of hydrogen bonds. The P200A substitution is not within the predicted bZIP region. In contrast, the base change in the consensus 5' splicing signal interrupts the predicted bZIP domain. *let-23* and *dpy-10* RNA mutant products containing alterations at the end of introns can be spliced correctly, indicating that other elements besides the 3' splicing signal cooperate in the splicing of wild-type mRNAs in *C. elegans* (Aroian *et al.*, 1993). This has not been determined for the 5' splicing signal. *let-607(h189)* was produced by treatment with EMS (A. Rose, University of British Columbia, Vancouver, Canada). This mutagen generates modifications in the genome randomly, and therefore the probability of having two lesions within the same gene is expected to be low. However it could happen, and because a silent change would be tightly linked to a causative change, it would not be lost by recombination. Interestingly, Thacker *et al.*, (2000) have reported recently the identification of two alterations within a *bli-4* mutant allele. One causes a H127L missense mutation, which is silent; and the other a nonsense mutation, which is the lesion responsible for the mutant phenotype.

The fact that the *let-607(h402)* mutant allele gives more severe defects than *let-607(h189)* supports the notion of *h402* being a null or strong loss of function, and accordingly, *h189* a hypomorph. The variable *let-607(h402)* defects could be interpreted if the wild-type function of *let-607* was partly redundant.

5.3.2 *let-607(h402)* and *let-607(h189)* mutant phenotypes are not caused by a dysfunctional hypodermis

During the development of the *C. elegans* embryo, the hypodermis largely defines the form and size of the nematode. Thus, the final worm shape of the embryo reflects the morphogenetic process of the hypodermis (reviewed in Chin-Sang and Chisholm, 2000). The hypodermis begins as a set of cells along the dorsal part of the embryo that forms an epithelial sheet (Sulston *et al.*, 1983). The cells at the right and left lateral edges of the epithelial sheet migrate towards the ventral midline where they form close contacts, completing the enclosure of the embryo. At this stage the embryo is an ellipsoidal mass of cells that subsequently turns into a long, thin larva by contraction of the microfilaments of actin organised circumferentially at the apical

surface of hypodermal cells (Priess and Hirsh, 1986). An even transmission of these contractile forces, and thereby uniform elongation of the embryo, is mediated by the compounded action of an array of microtubules circumferentially oriented underneath the hypodermal surface and the embryonic sheath, which is an extracellular matrix layer surrounding the embryo.

A diverse range of gene functions have been determined mutationally to be required for the shaping of the embryonic hypodermis (reviewed in Chin-Sang and Chisholm, 2000). As seen for the disruption of the *let-607* function, the loss of function of the gene *mab-20* gives embryonic lethality and viable morphological defects. *mab-20* encodes a secreted member of the semaphorin family, named semaphorin-2a, that is involved in the correct migration of hypodermal cells during embryonic morphogenesis of the hypodermis (Roy *et al.*, 2000). In particular, it regulates the formation and stabilisation of contacts between hypodermal cells during enclosure of the embryo. Lack of semaphorin-2a MAB-20 results in the formation of ectopic hypodermal contacts with a consequent failure to properly enclose the ventral surface with hypodermis. Most mutants rupture from the defective ventral side during embryonic elongation, presumably due to the pressure generated on internal cells by the contraction of the hypodermal microfilaments of actin (Priess and Hirsh, 1986). The mutant embryos that escape lethality hatch, frequently presenting severe bulges and constrictions along the body. They are shorter than wild types and sometimes have a split tail. These larvae either outgrow the phenotype or fail to survive. The body bulges and constrictions of *mab-20* null mutants develop during embryonic elongation and appear to be caused by contraction of circumferentially misaligned hypodermal microfilaments. Misalignment of the microfilaments occurs within hypodermal cells that cluster as a result of ectopic cell contacts. The altered arrangement of the microfilaments is thought to affect the even distribution of the contractile forces across the hypodermis, causing the described body deformities that persist after elongation. Body bulges always coincide with clusters of hypodermal cells.

Clusters of hypodermal cells consequent from ectopic contacts were detected in *mab-20* mutant L1s using the *ajm-1::GFP* marker, which permits the visualisation of hypodermal cell boundaries (Roy *et al.*, 2000). In contrast to *mab-20* mutant hatchlings, the localisation, contacts and fusion of hypodermal cells appeared to be normal in both *let-607(h402)* and *let-607(h189)* mutant larvae, suggesting that *let-607* function is not required for hypodermal morphogenesis. Consistently, *let-607(h402)* embryos elongate as per wild type indicating the presence of an intact hypodermis. In addition, the normal elongation of these mutant embryos is suggestive of an optimal function of the hypodermal cytoskeleton.

5.3.3 *let-607* is necessary for cuticle synthesis at various developmental stages

Disruption of the *let-607* function by genetic alterations or RNAi affects *C. elegans* growth and viability, and also impairs cuticle synthesis at different stages of development. The cuticular defects of *let-607* mutants and RNAi animals were determined by means of Nomarski, SEM, immunodetection, GFP fluorescence and treatment in a range of solutions.

5.3.3.1 *let-607(h402)* mutants fail to produce an intact L1 cuticle

Just prior to elongation of the embryo, microfilaments of actin organise into bundles close to the apical surface of the dorsal and ventral hypodermis (Priess and Hirsh, 1986). These bundles are circumferentially oriented relative to the longitudinal axis of the animal. Contraction of the bundles of microfilaments squeezes the embryo circumferentially with the consequent elongation of this from an ellipsoidal mass of cells (comma stage) to a worm shape (three-fold stage). Once the elongation process is completed, the bundles of microfilaments disorganise and the newly secreted cuticle takes the role in maintaining the shape of the elongated animal. This structural role of the nematode cuticle is supported by the phenotypic defects of *sqt-3(e2117)* and *pdi-2* RNAi embryos, which shorten in length and die post-elongation as a consequence of producing an aberrant cuticle that fails to hold the worm shape (Priess and Hirsh, 1986; Winter and Page, 2000). *sqt-3* encodes the cuticle collagen COL-1 and is abundantly expressed in embryos (Kramer *et al.*, 1985; Van der Keyl *et al.*, 1994). The lesion *e2117* is temperature sensitive, and impedes the assembly of the L1-cuticle fibrous layer at the restrictive temperature of 25°C (Priess and Hirsh, 1986). *pdi-2* encodes an isoform of the cuticle-collagen modifying enzyme PDI. Thus, lack of the *pdi-2* function during embryogenesis may impair the processing of all collagen proteins, affecting ultimately their assembly into the cuticle (Winter and Page, 2000). By analogy with *sqt-3(e2117)* and *pdi-2* RNAi phenotypes, the *let-607(h402)* embryonic defects described here may result from failure to synthesise an intact L1 cuticle. The collapse and complete loss of the worm shape could be a consequence of a considerable reduction in the collagens being produced and secreted to form the cuticle. Disarrangement of the hypodermal bundles of microfilaments after elongation would dispossess the animal of a supportive structure for its elongated shape if the cuticle was not assembled by that time. Absence of the supportive structure would explain the extrusion of internal cells by the internal hydrostatic pressure in the worm. In this contention, those *let-607(h402)* mutant embryos that do not collapse but shorten in length would synthesise enough collagens to assemble a cuticular structure enclosing the animal but insufficient to retain its elongated form. Clearly, this

quantitative deficiency would not exclude the possibility of qualitative defects of the cuticle collagens.

There are two other mutant alleles of the cuticle collagen gene *sqt-3*, namely *sc63* and *e24*. Both *sqt-3(sc63)* and *sqt-3(e24)* mutants show temperature-sensitive morphological abnormalities post-embryonically; they have a Dpy body with bulges and constrictions at the L1 stage when grown at 25°C (Van der Keyl *et al.*, 1994). *sqt-3(e24)* newly hatched larvae are noticeably more Dpy than *sqt-3(sc63)* L1s at the restrictive temperature. *let-607(h402)* animals that survive embryogenesis hatch as Dpy larvae with bulges and constrictions similar to *sqt-3(sc63)* and *sqt-3(e24)* L1 mutants. SEM analysis revealed that the surface of the cuticle of *let-607(h402)* hatchlings is notably affected: annuli and alae tend to be aberrant and absent over the regions of body bulges. These observations support the notion of an abnormal cuticle being the primary cause of *let-607(h402)* morphological defects. Interestingly, alterations in some cuticle collagens (including *sqt-3*) result in a malformed hermaphrodite tail (Van der Keyl *et al.*, 1994; Kramer, 1997). *let-607(h402)* mutants can present abnormal tail morphology, however, to what extent these tail defects are analogous to those of cuticle collagen mutants is not known.

5.3.3.2 *let-607(h189)* shrinking defects are possibly due to impaired cuticle synthesis

Some Unc mutants such as *unc-25*, *unc-47* and *unc-30* contract both dorsal and ventral body wall muscles when touched with the worm pick. This defect is known as "shrinker" and has been associated with an altered biosynthesis or utilisation of the neurotransmitter GABA (Eastman *et al.*, 1999). A tendency to shrink and relax ("rubberband" phenotype) in response to touch has been ascribed to certain Unc animals such as *unc-93* and *unc-110*, which are defective in muscle activation. In contrast to the above shrinking defects of Unc mutants, arrested *let-607(h189)* larvae contract lengthwise and do not regain the worm shape when prodded, with the consequent death of the animal. Therefore, it seems unlikely that the shrinking defects of *let-607(h189)* mutants are caused by neuromuscular dysfunction. There is no obvious interpretation for this *h189* mutant trait. However, if *let-607(h189)* L2 larvae that are about to die were attempting the L2-L3 moult but had difficulties in synthesising the new cuticle, the severe contraction of muscles in shrinking could be explained by the absence of the new cuticle-muscle attachment upon loss of the connection between the old cuticle and the muscle, which normally occurs during moulting.

5.3.3.3 let-607 RNAi animals display a range of cuticular defects

The function of the *let-607* gene was targeted for inhibition during embryogenesis in the *C. elegans* strains N2, DR466, JR667, TP12 and NL2099 by means of RNAi feeding. *let-607* RNAi N2, DR466, JR667 and TP12 F$_1$ progeny presented similar defects, which were not temperature sensitive as determined by growth of N2-treated animals at the temperatures 15°C, 20°C and 25°C. *let-607* RNAi treated N2 hermaphrodites raised at 20°C produced a brood size of about one third that of non-treated animals, which indicates that *let-607* may have a role in the nematode's fecundity. Approximately 27% of the progeny died during embryogenesis and the remainder hatched as small fat larvae. Both *let-607* RNAi dead embryos and newly hatched larvae resembled *let-607(h402)* mutants, although *let-607* RNAi hatched progeny did not normally arrest at the L1 stage as found with *let-607(h402)* hatchlings but tended to develop slowly as severe Dpys with reduced motility. The majority of these Dpys died before adulthood mostly presenting moult defects, and less frequently rupturing through the vulva, or occasionally, through the anus. Moult defects and rupture through cuticle openings suggest that *let-607* RNAi animals are defective in cuticle synthesis. Consistent with this, SEM analysis of the few Dpys that developed as adults revealed gross defects on the cuticle surface, including bifurcated and fragmented alae, misaligned or absent annuli as well as malformation of the vulva, which is lined by cuticle.

The nematode male tail is composed of cuticular structures. Interference of the *let-607* function was performed on DR466 hermaphrodites, which segregate ~33% males, in order to determine if *let-607* is required for the formation of the male tail. As determined by SEM, the few Dpy males that attained adulthood had severely malformed tails, consistent with an impaired production of cuticle structures in a *let-607* RNAi background. Interestingly, male tail abnormalities have also been reported for *dpy-18* and *sqt-1* mutants (Baird and Emmons, 1990). *dpy-18* encodes the α-subunit of the collagen processing enzyme P4H, and *sqt-1* encodes a cuticle collagen.

NL2099 animals are more sensitive to RNAi than N2s. Indeed, inhibition of *let-607* in NL2099 hermaphrodites gave more severe defects than in N2 animals. NL2099-treated hermaphrodites segregated some dead eggs and small, fat larvae. The hatched progeny normally presented bulges and constrictions along the body, hardly moved and tended to arrest at an early larval stage. Occasionally, these worms developed outgrowing the phenotype and arrested at late larval stages. In comparison with *let-607* RNAi N2 worms, the affected NL2099 larvae more closely resembled *let-607(h402)* hatchlings. The fact that reduced function of *let-607* by RNAi

results in similar *let-607(h402)* phenotypes supports the contention that *h402* may be a null mutation.

5.3.3.4 Cuticle collagen assembly is disrupted in let-607 mutant and RNAi worms

The localisation of cuticle collagen proteins was analysed in *let-607* mutants and RNAi-treated animals by immunodetection with the DPY7-5a antibody (McMahon *et al.*, 2003) and reporter GFP fluorescence. As shown by McMahon *et al.*, (2003), DPY-7 is detected within the cuticle in individual circumferential bands that are absent over the lateral sides when the alae are present in the L1 and adult stages. The bands are distributed uniformly along the animal's body and most likely correspond to the furrows or indentations of the cuticular annuli. *let-607(h402)* mutant embryos secrete DPY-7 indicating certain production of the L1 cuticle. The distribution of the secreted protein, however, could not be determined since the DPY-7 bands were compressed as a result of the collapse or retraction of these mutants after embryonic elongation. Secretion of DPY-7 collagen supports the suggestion from time course analyses that *let-607(h402)* mutant embryos die post-elongation, about the time of the production of the first cuticle.

Interestingly, *let-607(h402)* hatchlings and *let-607* RNAi N2 adults frequently showed intracellular accumulation of DPY-7 protein, presumably within hypodermal cells. Consistent with this, the amount of DPY-7 within the cuticles of these animals was clearly reduced, since DPY-7 bands tended to be partially or completely absent. The secreted DPY-7 protein generally assembled into the cuticle anomalously in the form of discontinuous and unevenly separated bands, which would correspond with severely malformed annuli. This corroborates the findings from SEM analyses of *let-607(h402)* larvae and *let-607* RNAi adults. Despite the similar morphological defects between *let-607(h402)* and *let-607* RNAi NL2099 larvae, the DPY-7 pattern was variable between these RNAi individuals, ranging from some intracellular accumulation of the protein and absent or malformed DPY-7 bands to practically intact DPY-7 bands. These diverse effects could be interpreted if there was residual *let-607* activity in the RNAi NL2099 larvae and the collagen protein was progressively incorporated into the cuticle. Furthermore, the distribution of DPY-7 within the cuticles of *let-607(h189)* hatchlings was also abnormal, more frequently from mid to end part of the body where the DPY-7 bands tended to be absent or severely irregular.

TP12 is a *col-19::GFP* gene fusion integrated strain which has been reported as an efficient marker for the identification of factors involved in cuticle synthesis (Thein *et al.*, 2003). The COL-19::GFP product localises to the cuticular annuli and alae of adult worms. The *let-607* RNAi TP12 animals that developed to adults were generally severe Dpys that displayed an aberrant COL-19 pattern. The collagen tended to be fragmented into fibres in both annuli and alae, which were frequently branched at various points. The abnormal localisation of DPY-7 and COL-19 in the mutants and RNAi-treated animals indicate that the *let-607* function is required for cuticle collagen secretion and assembly.

The cuticular alae are synthesised by the underlying seam cells (Singh and Sulston, 1978). It is therefore reasonable to conjecture that the branching alae seen in *let-607* RNAi adult worms is a consequence of misposition of the secreting seam cells. This could not be proved using the seam cell-specific*::GFP* marker insertion *wIs51* (JR667 strain), as the seam cell GFP fluorescence was too low or absent at the adult stage. Of note is the correspondence of branching alae and mislocalisation of seam cells in *C. elegans* mutants with a weakened cuticle lacking the thioredoxin activity *pdi-3* (Eschenlauer and Page, 2003). This was shown using the strain JR667 and an antibody staining hypodermal cell boundaries, termed MH27. Another attempt should therefore be carried out in order to characterise the seam cells of *let-607* RNAi JR667 adults. Staining with the antibody MH27 would be advantageous for this analysis.

5.3.3.5 *let-607* mutant and RNAi larvae are surrounded by weak cuticles

Cuticle integrity of *let-607* mutant hatchlings and *let-607* RNAi N2 larvae at the L2-L3 or L4 stage was tested in dH$_2$O, 1% β-ME or 1% SDS. In contrast to non-treated N2 larvae, which maintained the body shape, the *let-607* mutant and RNAi animals tended to burst in the above solutions. The frequencies of rupture in dH$_2$O, 1% β-ME and 1% SDS were respectively ~53.6%, 100% and 21.4% for *let-607(h402)* hatchlings; ~64.7%, ~91.7% and ~96.3% for *let-607* RNAi L2-L3 larvae; and ~57.9%, ~68.7% and ~21% for *let-607* RNAi L4 larvae. *let-607(h189)* larvae only burst in 1% β-ME, at a frequency of ~70.8%. The *C. elegans* cuticle functions in motility, maintenance of the worm shape and acts as a barrier between the nematode and the environment, conferring resistance to a wide range of solutions and facilitating osmoregulation. The nematode's motion and shape is also controlled by internal hydrostatic pressure, which is regulated by an osmoregulatory system. The observed rupture of *let-607* mutants and RNAi larvae can be interpreted as a reduced physical resistance of their cuticles to withstand an increase in the internal hydrostatic pressure. In this analysis, dH$_2$O acts

as a hypotonic solution whereas β-ME and SDS affect the integrity of the animal as a sulfhydril reducing agent and a denaturing agent respectively. Entry of H_2O molecules into the animal is expected to result in the increase in hydrostatic pressure causing the collapse of more than half of the *let-607(h402)* hatchlings and RNAi-treated worms in this solution. Collagens within the nematode cuticle are extensively cross-linked primarily by disulphide bonding, and thereby can be solubilised using β-ME at 5% (Cox *et al.*, 1981a). Here 1% β-ME appears to affect considerably the strength of the cuticles of *let-607(h402)* and *let-607(h189)* mutants, and to a lesser extent those of the RNAi-treated larvae. It should be noted that β-ME was prepared to 1% in dH_2O, and therefore the frequency of rupture in dH_2O has to be taken into account to evaluate the effects of 1% β-ME on the cuticles of these mutants. Finally, the extensive cross-linking between cuticle proteins renders the nematode cuticle practically resistant to treatment with 1% SDS, which is used together with sonication to isolate cuticles from all cellular material (Cox *et al.*, 1981a). SDS may enter into the animal and dissolve all tissues by denaturing cellular proteins. This is consistent with the fact that worms placed in 1% SDS become progressively clear losing the typical colouration of their body contents; the effect on mutant and RNAi-treated animals was notably faster than on wild types. The nematode's hydrostatic pressure may be maintained by an ion imbalance between the worm and its environment. H_2O molecules would enter into the nematode contributing to the internal hydrostatic pressure if there were higher ion levels in the animal relative to its surrounding medium. 1% SDS may enter and dissolve cellular membranes, with the consequent disruption of the ion imbalance and entrance of H_2O molecules by osmosis. The lack of internal hydrostatic pressure may account for the small percentage of mutant and RNAi L4 larvae that burst in 1%SDS. An explanation for the high frequency of rupture of *let-607* RNAi L2-L3 larvae in 1% SDS has not been found. Burst of *let-607* mutant and RNAi larvae in the tested solutions indicates that these animals are surrounded by weak cuticles, as a consequence of impaired cuticle synthesis.

5.3.3.6 *let-607* RNAi animals treated post-embryonically may be defective in cuticle synthesis late in development

The post-embryonic function of the gene *let-607* was targeted for disruption by placing N2 or NL2099 embryos from untreated mothers on *let-607* RNAi plates. Thus, the animals were not exposed to *let-607* RNAi during embryogenesis but in the course of larval development. *let-607* RNAi NL2099 worms were noticeably more affected than N2s. RNAi-treated NL2099 larvae developed more slowly and were smaller than control worms. They became Dpy about the L4 stage; when they also displayed other defects including arrest at the L4-adult moult, shrinking when prodded, rupture through the vulva and general sickness. None of these animals developed

as healthy adults. These defects provide additional evidence on the requirement of the *let-607* function for post-embryonic cuticle synthesis and development.

5.3.4 *let-607* encodes a putative bZIP transcription factor

let-607 encodes a 688 amino acid product with a predicted bZIP region, characteristic of a diverse class of DNA-binding transcription factors. Interestingly, the *let-607* protein appears to be related to the transcription factors involved in UPR. The UPR transcription factors belong to the ATF/CREB family of bZIP proteins. UPR is the collective term given to several signalling pathways activated upon accumulation of unfolded proteins within the ER (reviewed in Welihinda *et al.*, 1999; Ma and Hendershot, 2001; Patil and Walter, 2001; Kaufman *et al.*, 2002). Diverse conditions can lead to such accumulation of unfolded proteins and thereby ER stress: elevated synthesis of secretory proteins, expression of mutant or misfolded proteins, decrease in calcium levels and perturbations in glycosylation. The UPR controls both cellular transcription and translation to limit further accumulation of unfolded proteins and re-establish ER homeostasis. At the transcriptional level, the UPR induces the expression of a large set of genes encoding many functions of the secretory pathway, including protein folding (protein catalysts and chaperones), degradation of unfolded proteins and lipid biosynthesis for ER membrane expansion (Fewell *et al.*, 2001). At the translational level, the UPR alters the patterns of protein synthesis by attenuating general translation as well as inducing transcription of UPR-activated genes (Kaufman *et al.*, 2002; Kaufman, 2002). In mammals, if these adaptive responses are not sufficient to counteract the ER stress, the prolonged UPR activation leads to cell death through apoptosis or necrosis. Thus, the degree and/or the duration of the ER stress conceivably determine(s) the appropriate response, either cell adaptation or cell death, via the UPR.

The UPR induced by ER stress is a mechanism conserved across evolution; it has been described in yeast, *C. elegans* and mammals. Recent evidence from studies in *C. elegans* and mice support the view that ER stress signalling has important physiological roles in tissues specialised in secretion, such as the intestine and pancreas (reviewed in Kaufman *et al.*, 2002; Kaufman, 2002). In *C. elegans*, double loss of function in the UPR transcriptional induction and translational regulation results in growth arrest at the L2 stage and necrotic intestinal degeneration (Shen *et al.*, 2001). This suggests that intestinal cells experience ER stress at the L2 stage, possibly due to an increase in synthesis of secretory proteins such as digestive enzymes and moulting hormones.

Five cuticles are produced by the underlying hypodermis during the *C. elegans* life cycle, each cuticle being synthesised towards the end of each developmental stage. The high production of collagens required for cuticle secretion must be accompanied by an increase in the ER membrane as well as in many functions of the secretory pathway, including protein processing enzymes, molecular chaperones and proteins acting in transport and secretion. It seems reasonable to think that a mechanism such as the UPR may be responsible for these diverse changes in the secretory pathway necessary to accommodate the load of cuticle collagens being expressed in the hypodermal cells. Thus I hypothesise that the transcription factor LET-607 could be required in an UPR-like mechanism during cuticle synthesis in the hypodermis. Of significance in this respect is the intracellular accumulation of DPY-7 collagen observed in *let-607(h402)* hatchlings and *let-607* RNAi animals. Intracellular accumulation of DPY-7 collagen has not been detected in several Dpy mutants that secrete small amounts of this collagen protein (McMahon *et al.*, 2003), suggesting that this phenotypic trait characteristic of *let-607(h402)* larvae and *let-607* RNAi worms could be a consequence of defects in more than one function in the secretory pathway.

If a form of UPR is involved in cuticle synthesis, this response could not include the attenuation of general protein translation and induction of cell death. Expression of cuticle collagens starts and is continued for four hours prior to cuticle deposition (Johnstone and Barry, 1996). This indicates that translation of cuticle collagens, most likely representing that of general proteins, is not attenuated in the hypodermis for this period of time. Furthermore, the level and the duration of cuticle collagen synthesis may require maintenance of a prolonged UPR, and clearly *C. elegans* does not induce cell death each time it undergoes cuticle synthesis. Supporting the notion of a specific UPR-like mechanism in the hypodermal cells is the fact that *C. elegans* null mutants lacking the functions of two key components of the UPR die at the L2 stage possibly due to intestinal dysfunction (Shen *et al.*, 2002), and not at the terminal phase of embryogenesis during production of the first nematode cuticle.

5.3.5 *let-607* may not be essential for the transcription of cuticle collagens, but it is expressed in the cuticle synthesising tissue at all developmental stages

Cuticle collagen gene expression was examined in *let-607(h402)* embryos (presumed null mutants) using the reporter transgenes *dpy-7::GFP* and *col-12::GFP*. Expression of cuticle collagen genes peaks once in each developmental stage, preceding cuticle synthesis. As determined by Johnstone and Barry (1996), the peaks of expression of individual cuticle

collagens are not coincident but temporally separated in three major discrete points. Accordingly, cuticle collagens can be classified as early-expressed (four hours prior to cuticle secretion), middle-expressed (two hours before cuticle secretion) and late-expressed (coinciding with cuticle secretion). *dpy-7* and *col-12* are early and late expressed cuticle collagen genes respectively. During embryonic development, expression of *dpy-7* begins just prior to elongation in the comma stage embryo, whereas that of *col-12* occurs after elongation at the three-fold stage and coincides with the formation of the L1 cuticle. *let-607(h402)* mutant embryos expressed both *dpy-7* and *col-12* reporter fusions indicating that the wild-type *let-607* function is not essential for the transcription of at least these two cuticle collagen genes.

The spatial and temporal expression of *let-607* was determined by promoter-driven GFP detection and RT-PCR respectively. Expression of *let-607* was first observed at comma stage in the hypodermis, the tissue synthesising the cuticle. During post-embryonic development, *let-607* was expressed in the hypodermis and in other cells including cells within the pharynx and possibly nerve cells. Interestingly, although the *let-607* function is not essential for the transcription of specific collagens, the earliest expression of this putative bZIP transcription factor coincides with that of cuticle collagens, and thus with the onset of cuticle synthesis.

As determined by Johnstone and Barry (1996), transcription of cuticle collagens peaks at each developmental stage prior to cuticle production. Similar levels of *let-607* transcripts were detected at the equal time intervals measured during larval and adult stages. The results presented here indicate that *let-607* expression is constitutive during post-embryogenesis, and therefore does not resemble the temporal reiterative waves of expression of cuticle collagens.

5.3.6 Concluding remarks

Here it has been determined that the *C. elegans* gene *let-607* encodes a putative transcription factor of the bZIP family. *let-607* is defined by two distinct mutations: *h402* and *h189*. *let-607(h402)* is predicted to be a null allele and *let-607(h189)* a missense allele. Accordingly, *let-607(h402)* mutants arrest at an earlier developmental stage and display more severe phenotypic defects than *let-607(h189)* animals. Either of the genetic lesions in the *let-607* gene or synthetic disruption of the *let-607* function by RNAi impairs cuticle production at different developmental stages, indicating that this putative transcription factor may be required for cuticle synthesis. Consistent with this possible function, *let-607* is expressed in the hypodermis (the cuticle synthesising tissue) at all developmental stages. Interestingly, the *let-607* product is

related by sequence to the bZIP transcription factors involved in the induction of the expression of a wide range of secretory pathway functions upon an increase of unfolded proteins in the ER.

5.3.7 Future work on *let-607*

In order to determine if *h402* is a null mutation, the *let-607(h402)* allele could be positioned in *trans* to a non-complementing deficiency such as hDf8. The *let-607(h402)* allele would be a true null if the phenotypes of *let-607(h402)/let-607(h402)* homozygous mutants were identical to those of *let-607(h402)/hDf8* animals. In contrast, *let-607(h402)* would not be a null if there was occasional ochre read-through resulting in residual *let-607* activity. This case would be indicated if *let-607(h402)/hDf8* heterozygotes presented more severe phenotypes than *let-607(h402)/let-607(h402)* mutants.

To detect which of the two lesions identified in the *let-607(h189)* allele is responsible for the mutant phenotype, plasmids containing either base change in the *let-607* gene could be constructed and injected singly into hermaphrodite adults heterozygous for the *let-607(h189)* lethal allele. Phenotypic rescue of the *let-607(h189)* homozygous progeny would imply that the incorporated base change is a silent mutation, as it does not affect the wild-type function of *let-607*. Lack of rescue would reveal the base change causing the *let-607(h189)* mutant phenotype. Moreover, the effects of the *h189* base change at the 5' end of intron 4 on the splicing of *let-607* mRNAs could be analysed by RT-PCR amplification of the *let-607* DNA from mRNA isolated from *h189* homozygous mutants. An increase in the length of the PCR amplified product corresponding to the size of intron 4 would indicate that this lesion impedes splicing, and possibly affects the wild-type function of the *let-607* product.

The intracellular accumulation of DPY-7 collagen and its effects on the cuticle structure of *let-607(h402)* hatchlings could be examined by transmission electron microscopy. This study was initiated in the course of the project, but it was abandoned prior to determining the best procedure due to diverse technical problems.

DNA microarray analysis would be convenient to identify possible target genes of the putative LET-607 transcription factor. This could be performed by comparison of the overall gene transcription of *let-607* RNAi NL2099 larvae with non-treated NL2099 worms at a similar developmental stage. LET-607 potential targets would be those genes that are transcribed in non-treated NL2099 animals (control population) but down-regulated in *let-607* RNAi NL2099 worms (test population).

General Discussion

The primary goal of my project was to identify genes required for the synthesis of the ECM using the free-living nematode *C. elegans*. Thus, I focused on the analysis of two previously uncharacterised *C. elegans* genes, *stc-1* and *let-607*, as possible candidates encoding factors required for cuticle synthesis. This was based on the observation that disruption of either *stc-1* or *let-607* function causes severe cuticular defects. The mutant allele *ij15* was isolated from a forward genetic screen previously performed (C. Clucas and I. Johnstone, Glasgow University, Glasgow, UK) to identify recessive mutants that die late in embryogenesis, about the time of cuticle secretion. *ij15* homozygous embryos elongate and subsequently burst losing the worm shape. These defects could be interpreted if insufficient material to assemble the L1 cuticle was secreted by the time the hypodermal cytoskeleton disorganises. The extrusion of cells could be driven by the internal hydrostatic pressure in the elongated embryo due to the absence of an enclosing supportive structure. Moreover, *ij15* mutants transcribe late expressed cuticle collagens and secrete cuticle to some extent, activities typically associated with a late-stage embryo. These reinforce the view that *ij15* embryos die during secretion of the L1 cuticle, as inferred by their development to a fully elongated worm. The lesion *ij15* defines the gene *stc-1*, which encodes a HSP70-like protein. *ij15* is a missense mutation that affects a consensus residue within the ATPase domain common to HSP70 proteins. Inhibition of *stc-1* function by RNAi feeding results in larval lethality about the L3 stage. The *stc-1* RNAi larvae display diverse phenotypic defects including abnormal cuticle synthesis. By sequence prediction, the *stc-1* product is homologous to the mammalian STCH proteins and possibly functions in the secretory pathway. *stc-1* is expressed in the hypodermis at all developmental stages. Taken together, these findings indicate that STC-1 may be a HSP70-like protein required in cuticle synthesis in *C. elegans*, in particular in the biogenesis of cuticle collagens.

Disruption of the function of the predicted gene F57B10.1 by RNAi was found to generate dead eggs and Dpy progeny. F57B10.1 encodes a protein related to the bZIP transcription factors involved in signalling pathways activated upon accumulation of unfolded proteins in the ER. F57B10.1 corresponds to the mutationally defined locus *let-607*. There are two recessive mutant alleles of *let-607*: *h402* and *h189*. The *let-607(h402)* mutant allele gives variable lethal phenotypes, all discernible after embryonic elongation. More than half of the *let-607(h402)* animals die late in embryogenesis, when the first nematode cuticle is secreted; whereas the remainder arrest at the L1 stage. *let-607(h402)* homozygous mutant embryos can burst after elongation losing the elongated shape, or more frequently, do not disrupt but shorten in length. The *let-607(h402)* animals that develop to the L1 stage present severe morphological defects; they are small and fat, and frequently have bulges and constrictions randomly about the body.

Like *stc-1(ij15)* embryos, rupture of *let-607(h402)* embryos could be caused by the internal hydrostatic pressure if a cuticular structure does not enclose the elongated animal when the cytoskeleton of hypodermal cells disarrange post-elongation. In this context, those *let-607(h402)* embryos that do not disrupt but retract back would produce a surrounding cuticular structure unable to hold the elongated shape but strong enough to contain the internal cells. Accordingly, *let-607(h402)* mutant embryos secrete to some extent the L1 cuticle after elongation. Those *let-607(h402)* animals that survive embryogenesis have body deformities possibly as a consequence of their cuticle defects.

The other *let-607* mutant allele, *h189*, generally results in larval arrest at the L2 stage. In comparison with *let-607(h402)* hatchlings, the *let-607(h189)* larvae do not present noticeable morphological defects but also produce an aberrant cuticle. The fact that *let-607(h402)* mutants display more severe phenotypic defects than *let-607(h189)* animals suggests that *h402* may be a null allele and *h189* is a hypomorphic allele. Consistent with this, *h402* is a nonsense mutation that would terminate LET-607 translation before the predicted bZIP domain. Furthermore, the defects of *let-607* RNAi animals partly phenocopy those of *let-607(h402)* mutants. Interference of *let-607* function by RNAi impairs cuticle synthesis and results in a range of cuticular defects. Despite wild-type LET-607 factor not being essential for the transcription of specific cuticle collagens, it is expressed in the hypodermis at all developmental stages. Thus, the putative transcription factor LET-607 may be important for the production of the cuticle in *C. elegans* possibly by inducing the expression of the cuticle synthetic machinery. Supporting this view, *let-607(h402)* hatchlings, *let-607* RNAi N2 adults and NL2099 larvae display intracellular accumulation of collagen, which could be interpreted if some part of the secretory pathway was not functioning optimally.

In conclusion, this work includes the characterisation of two novel protein factors required for the synthesis of the cuticle, a *C. elegans* ECM. Specifically, the HSP70-like protein STC-1 and the putative bZIP transcription factor LET-607. From evidence presented throughout this thesis, it could be hypothesised that STC-1 acts as a molecular chaperone in the biosynthesis of collagens components of this ECM. On the other hand, it can be speculated that LET-607 regulates the transcription of secretory pathway functions when there is high synthesis of secretory proteins, such as that of collagens required to form the cuticle. Therefore, these two factors could control distinct levels of the synthesis of the cuticular ECM in *C. elegans*. *stc-1* and *let-607* were identified by independent approaches on the basis of their mutant or RNAi

phenotypes. The roles determined here for the *stc-1* and *let-607* products validate the use of these *C. elegans* phenotypes as a means to detect factors involved in synthesis of the ECM.

References

Agashe, V. R. and Hartl, F. U. (2000). Roles of molecular chaperones in cytoplasmic protein folding. *Seminars in Cell and Developmental Biology* 11:15-25.

Anderson, P. (1995). In *Methods in Cell Biology*, Vol. 48 (Eds. H. F. Epstein and D. C. Shakes), pp. 31-58, Academic Press, San Diego, CA, USA.

Argon, Y. and Simen, B. B. (1999). GRP94, an ER chaperone with protein and peptide binding properties. *Seminars in Cell and Developmental Biology* 10:495-505.

Aroian, R. V., Levy, A. D., Koga, M., Ohshima, Y., Kramer, J. M., and Sternberg, P. W. (1993). Splicing in *Caenorhabditis elegans* does not require an AG at the 3' splice acceptor site. *Molecular and Cellular Biology* 13:626-637.

Baird, S. E. and Emmons, S. W. (1990). Properties of a class of genes required for ray morphogenesis in *Caenorhabditis elegans*. *Genetics* 126:335-344.

Barr, P. J. (1991). Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell* 66:1-3.

Barthel, T. K., Zhang, J., and Walker, G. C. (2001). ATPase-defective derivatives of *Escherichia coli* DnaK that behave differently with respect to ATP-induced conformational change and peptide release. *Journal of Bacteriology* 183:5482-5490.

Bass, J., Chiu, G., Argon, Y., and Steiner, D. F. (1998). Folding of insulin receptor monomers is facilitated by the molecular chaperones calnexin and calreticulin and impaired by rapid dimerization. *Journal of Cell Biology* 141:637-646.

Bird, D. M. and Riddle, D. L. (1989). Molecular cloning and sequencing of *ama-1*, the gene encoding the largest subunit of *Caenorhabditis elegans* RNA polymerase II. *Molecular and Cellular Biology* 9:4119-4130.

Blaxter, M. L. and Bird, D. M (1997). In *C. elegans II* (Eds. D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess), pp. 851-878, Cold Spring Harbor Laboratory Press, New York, USA.

Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proceedings of the National Academy of Sciences of USA* **89**:7290-7294.

Borkovich, K. A., Farrelly, F. W., Finkelstein, D. B., Taulien, J., and Lindquist, S. (1989). Hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temepratures. *Molecular and Cellular Biology* **9**:3919-3930.

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.

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

Bruneau, N., Lombardo, D., and Bendayan, M. (1998). Participation of GRP94-related protein in secretion of pancreatic bile salt-dependent lipase and in its internalization by the intestinal epithelium. *Journal of Cell Sciences* **111**:2665-2679.

Buchberger, A., Valencia, A., McMacken, R., Sander, C., and Bukau, B. (1994). The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171. *EMBO Journal* **13**:1687-1695.

Buchner, J. (1999). Hsp90 & Co. - a holding for folding. *Trends in Biochemical Sciences* **24**:136-141.

Bukau, B. and Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* **92**:351-366.

Bulleid, N. J. and Freeman, B. C. (1988). Defective co-translational formation of disulphide bonds in protein disulphide-isomerase-deficient microsomes. *Nature* **335**:649-651.

C.elegans sequencing consortium (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**:2012-2018.

Calfon, M., Zeng, H. Q., Urano, F., Till, J. H., Hubbard, S. R., Harding, H. P., Clark, S. G., and Ron, D. (2002). IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* **415**:92-96.

Chin-Sang, I. D. and Chisholm, A. D. (2000). Form of the worm: genetics of epidermal morphogenesis in *C. elegans*. *Trends in Genetics* **16**:544-551.

Clucas, C., Cabello, J., Bussing, I., Schnabel, R., and Johnstone, I. L. (2002). Oncogenic potential of a *C.elegans cdc25* gene is demonstrated by a gain-of-function allele. *EMBO Journal* **21**:665-674.

Costa, M., Draper, B. W., and Priess, J. R. (1997). The role of actin filaments in patterning the *Caenorhabditis elegans* cuticle. *Developmental Biology* **184**:373-384.

Coulson, A., Waterston, R., Kiff, J., Sulston, J., and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. *Nature* **335**:184-186.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J, and Waterston, R. (1991). YACs and the *C. elegans* genome. *Bioessays* **13**:413-417.

Cox, G. N., Kusch, M., and Edgar, R. S. (1981a). Cuticle of *Caenorhabditis elegans*: its isolation and partial characterisation. *Journal of Cell Biology* **90**:7-17.

Cox, G. N., Staprans, S., and Edgar, R. S. (1981b). The cuticle of *Caenorhabditis elegans*. II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. *Developmental Biology* **86**:456-470.

Cox, G. N., Carr, S., Kramer, J. M., and Hirsh, D. (1985). Genetic mapping of *Caenorhabditis elegans* collagen genes using DNA polymorphisms as phenotypic markers. *Genetics* **109**:513-528.

Cox, G. N. and Hirsh, D. (1985). Stage-specific patterns of collagen gene expression during development of *Caenorhabditis elegans*. *Molecular and Cellular Biology* **5**:363-372.

Cox, G. N. (1992). Molecular and biochemical aspects of nematode collagens. *Journal of Parasitology* **78**:1-15.

Cutforth, T. and Rubin, G. M. (1994). Mutations in HSP83 and CDC37 impair signaling by the sevenless receptor tyrosine kinase in *Droshophila*. *Cell* **77**:1027-1036.

Eastman, C., Horvitz, H. R., and Jin, Y. (1999). Coordinated transcriptional regulation of the *unc-25* glutamic acid decarboxylase and the *unc-47* GABA vesicular transporter by the *Caenorhabditis elegans* UNC-30 homeodomain protein. *Journal of Neuroscience* **19**:6225-6234.

Edwards, M. K. and Wood, W. B. (1983). Localisation of specific messenger RNAs in *C. elegans* by cytological hybridization. *Developmental Biology* **97**:375-390.

Elefant, F. and Palter, K. B. (1999). Tissue-specific expression of dominant negative mutant *Drosophila HSC70* causes developmental defects and lethality. *Molecular Biology of the Cell* **10**:2101-2117.

Engel, J. and Prockop, D. J. (1991). The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annual Review of Biophysics and Biophysical Chemistry* **20**:137-152.

Eschenlauer, S. C. P. and Page, A. P. (2003). The *Caenorhabditis elegans* ERp60 homolog protein disulfide isomerase-3 has disulfide isomerase and transglutaminase-like cross-linking activity and is involved in the maintenance of body morphology. *Journal of Biological Chemistry* **278**:4227-4237.

Favre, R., Hermann, R., Cermola, M., Hohenberg, H., Müller, M., and Bazzicalupo, P. (1995). Immuno-gold-labelling of CUT-1, CUT-2 and cuticlin epitopes in *Caenorhabditis elegans* and *Heterorhabditis sp.* processed by high pressure freezing and freeze-substitution. *Journal of Submicroscopic Cytology and Pathology* **27**:341-347.

Ferreira, L. R., Norris, K., Smith, T., Hebert, C., and Sauk, J. J. (1994). Association of Hsp47, Grp78, and Grp94 with procollagen supports the successive or coupled action of molecular chaperones. *Journal of Cellular Biochemistry* **56**:518-526.

Ferreira, L. R., Norris, K., Smith, T., Hebert, C., and Sauk, J. J. (1996). Hsp47 and other ER-resident molecular chaperones form heterocomplexes with each other and with collagen type IV chains. *Connective Tissue Research* **33**:265-273.

Fewell, S. W., Travers, K. J., Weissman, J. S., and Brodsky, J. L. (2001). The action of molecular chaperones in the early secretory pathway. *Annual Review of Genetics* 35:149-191.

Fields, C. (1988). Domain organization and intron positions in *Caenorhabditis elegans* collagen genes: the 54-bp module hypothesis revisited. *Journal of Molecular Evolution* 28:55-63.

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

Flaherty, K. M., DeLuca-Flaherty, C., and Mckay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346:623-628.

Flaherty, K. M., Mckay, D. B., Kabsch, W., and Holmes, K. C. (1991). Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein. *Proceedings of the National Academy of Sciences of USA* 88:5041-5045.

Flaherty, K. M., Wilbanks, S. M., DeLuca-Flaherty, C., and Mckay, D. B. (1994). Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. *Journal of Biological Chemistry* 269:12899-12907.

Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991). Peptide-binding specificity of the molecular chaperone BiP. *Nature* 353:726-730.

Forlino, A. and Marini, J. C. (2000). Osteogenesis imperfecta: prospects for molecular therapeutics. *Molecular Genetics and Metabolism* 71:225-232.

Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* 408:325-330.

Friedman, L., Higgin, J. J., Moulder, G., Barstead, R., Raines, R. T., and Kimble, J. (2000). Prolyl 4-hydroxylase is required for viability and morphogenesis in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of USA* 97:4736-4741.

Frydman, J. (2001). Folding of newly translated proteins *in vivo*: The role of molecular chaperones. *Annual Review of Biochemistry* 70:603-647.

Fujimoto, D. and Kanaya, S. (1973). Cuticlin: a noncollagen structural protein from *Ascaris* cuticle. *Archives of Biochemistry and Biophysics* 157:1-6.

Garnier, C., Lafitte, D., Tsvetkov, P. O., Barbier, P., Leclerc-Devin, J., Millot, J. M., Briand, C., Makarov, A. A., Catelli, M. G., and Peyrot, V. (2002). Binding of ATP to heat shock protein 90: evidence for an ATP-binding site in the C-terminal domain. *Journal of Biological Chemistry* 277:12208-12214.

Gething, M. J. and Sambrook, J. (1992). Protein folding in the cell. *Nature* 355:33-45.

Gething, M. J. (1999). Role and regulation of the ER chaperone BiP. *Seminars in Cell and Developmental Biology* 10:465-472.

Gilbert, H. F. (1997). Protein disulfide isomerase and assisted protein folding. *Journal of Biological Chemistry* 272:29399-29402.

Gonczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S. J. M., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S., Flohrs, K., Goessen, A., Leidel, S., Alleaume, A. M., Martin, C., Ozlu, N., Bork, P., and Hyman, A. A. (2000). Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* 408:331-336.

Granato, M., Schnabel, H., and Schnabel, R. (1994). Genesis of an organ: molecular analysis of the *pha-1* gene. *Development* 120:3005-3017.

Guo, S. and Kemphues, K. J. (1995). *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* 81:611-620.

Haas, I. G. and Wabl, M. (1983). Immunoglobulin heavy chain binding protein. *Nature* 306:387-389.

Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F. U., and Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* **276**:431-435.

Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. *Nature* **381**:571-580.

Hedgecock, E. M. and White, J. G. (1985). Polyploid tissues in the nematode *Caenorhabditis elegans*. *Developmental Biology* **107**:128-138.

Hill, K. L., Harfe, B. D., Dobbins, C. A., and L'Hernault, S. W. (2000a). *dpy-18* encodes an α-subunit of prolyl-4-hydroxylase in *Caenorhabditis elegans*. *Genetics* **155**:1139-1148.

Hill, A. A., Hunter, C. P., Tsung, B. T., Tucker-Kellogg, G., and Brown, E. L. (2000b). Genomic analysis of gene expression in *C. elegans*. *Science* **290**:809-812.

Hirayoshi, K., Kudo, H., Takechi, H., Nakai, A., Iwamatsu, A., Yamada, K. M., and Nagata, K. (1991). HSP47: a tissue-specific, transformation-sensitive, collagen-binding heat shock protein of chicken embryo fibroblasts. *Molecular and Cellular Biology* **11**:4036-4045.

Hodgkin, J. (1999). In *C. elegans: a practical approach* (Ed. I. A. Hope), pp. 245-269, Oxford University Press, New York, USA.

Hope, I. A. (1991). Promoter trapping in *Caenorhabditis elegans*. *Development* **113**:399-408.

Hope, I. A., Arnold, J. M., McCarroll, D., Jun, G., Krupa, A. P., and Herbert, R. (1998). Promoter trapping identifies real genes in *C. elegans*. *Molecular and General Genetics* **260**:300-308.

Hope, I. A. (2001). Broadcast interference - functional genomics. *Trends in Genetics* **17**:297-299.

Hurley, J. H. (1996). The sugar kinase/heat shoch protein 70/actin superfamily: implications of conserved structure for mechanism. *Annual Review of Biophysics and Biomolecular Structure* **25**:137-162.

Hurst, H. C. (1996). Leucine zippers transcription factors. In *Protein Profile*, 3rd Edn (Ed. P. Sheterline), Academic Press, London, UK.

Ihara, Y., Cohen-Doyle, M. F., Saito, Y., and Williams, D. B. (1999). Calnexin discriminates between protein conformational states and functions as a molecular chaperone *in vitro*. *Molecular Cell* 4:331-341.

Irmer, H. and Hohfeld, J. (1997). Characterization of functional domains of the eukaryotic co-chaperone hip. *Journal of Biological Chemistry* 272:2230-2235.

Jansen, G., Hazendonk, E., Thijssen, K. L., and Plasterk, R. H. A. (1997). Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nature Genetics* 17:119-121.

Jansen, G. (2002). Gene inactivation in *Caenorhabditis elegans*. *Current Genomics* 3:59-67.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., and Kim, S. K. (2001). Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of USA* 98:218-223.

Johnson, E. R. and Mckay, D. B. (1999). Mapping the role of active site residues for transducing an ATP-induced conformational change in the bovine 70-kDa heat shock cognate protein. *Biochemistry* 38:10823-10830.

Johnstone, I. L., Shafi, Y., and Barry, J. D. (1992). Molecular analysis of mutations in the *Caenorhabditis elegans* collagen gene *dpy-7*. *EMBO Journal* 11:3857-3863.

Johnstone, I. L. (1994). The cuticle of the nematode *Caenorhabditis elegans*: a complex collagen structure. *Bioessays* 16:171-178.

Johnstone, I. L. and Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development. *EMBO Journal* 15:3633-3639.

Johnstone, I. L. (1999). In *C. elegans: a practical approach* (Ed. I. A. Hope), pp. 201-225, Oxford University Press, New York, USA.

Johnstone, I. L. (2000). Cuticle collagen genes. Expression in *Caenorhabditis elegans*. *Trends in Genetics* 16:21-27.

Jorgensen, E. M. and Mango, S. E. (2002). The art and design of genetic screens: *Caenorhabditis elegans*. *Nature Reviews Genetics* 3:356-369.

Kabsch, W. and Holmes, K. C. (1995). Protein motifs. II The actin fold. *FASEB Journal* 9:167-174.

Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., and Ahringer, J. (2000). Effectiveness of specific RNA-mediated itnerference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biology* 2:0002.1-0002.10.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). Sistematic fucntional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421:231-237.

Kaufman, R. J. (2002). Orchestrating the unfolded protein response in health and disease. *Journal of Clinical Investigation* 110:1389-1398.

Kaufman, R. J., Scheuner, D., Schroder, M., Shen, X. H., Lee, K., Liu, C. Y., and Arnold, S. M. (2002). The unfolded protein response in nutrient sensing and differentiation. *Nature Reviews Molecular Cell Biology* 3:411-421.

Kaye, F. J., Modi, S., Ivanovska, I., Koonin, E. V., Thress, K., Kubo, A., Kornbluth, S., and Rose, M. D. (2000). A family of ubiquitin-like proteins binds the ATPase domain of Hsp-70-like Stch. *Febs Letters* 467:348-352.

Kelly, W. G., Xu, S., Montgomery, M. K., and Fire, A. (1997). Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* 146:227-238.

Kim, S. K. (2001a). Functional genomics: the worm scores a knockout. *Current Biology* 11:R85-R87.

Kim, S. K. (2001b). HTTP://C. elegans: mining the functional genomic landscape. *Nature Reviews Genetics* 2:681-689.

Koivu, J. and Myllylä, R. (1987). Interchain disulfide bond formation in types I and II procollagen. Evidence for a protein disulfide isomerase catalyzing bond formation. *Journal of Biological Chemistry* 262:6159-6164.

Kozutsumi, Y., Segal, M., Normington, K., Gething, M. J., and Sambrook, J. (1988). The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose regulated proteins. *Nature* 332:462-464.

Köppen, 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. *Nature Cell Biology* 3:983-991.

Kramer, J. M., Cox, G. N., and Hirsh, D. (1985). Expression of the *Caenorhabditis elegans* collagen genes *col-1* and *col-2* is developmentally regulated. *Journal of Biological Chemistry* 260:1945-1951.

Kramer, J. M., Johnson, J. J., Edgar, R. S., Basch, C., and Roberts, S. (1988). The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. *Cell* 55:555-565.

Kramer, J. M., French, R. P., Park, E. C., and Johnson, J. J. (1990). The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. *Molecular and Cellular Biology* 10:2081-2089.

Kramer, J. M. and Johnson, J. J. (1993). Analysis of mutations in the *sqt-1* and *rol-6* collagen genes of *Caenorhabditis elegans*. *Genetics* 135:1035-1045.

Kramer, J. M. (1994a). Genetic analysis of extracellular matrix in *C. elegans*. *Annual Review of Genetics* 28:95-116.

Kramer, J. M. (1994b). Structures and functions of collagens in *Caenorhabditis elegans*. *FASEB Journal* 8:329-336.

Kramer, J. M. (1997). In *C. elegans II* (Eds. D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess), pp. 471-500, Cold Spring Harbor Laboratory Press, New York, USA.

Kusch, M. and Edgar, R. S. (1986). Genetic studies of unusual loci that affect body shape of the nematode *Caenorhabditis elegans* and may code for cuticle structural proteins. *Genetics* 113:621-639.

Lamb, P. and McKnight, S. L. (1991). Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. *Trends in Biochemical Sciences* 16:417-422.

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240:1759-1764.

Lassle, M., Blatch, G. L., Kundra, V., Takatori, T., and Zetter, B. R. (1997). Stress-inducible, murine protein mSTI1 - Characterization of binding domains for heat shock proteins and *in vitro* phosphorylation by different kinases. *Journal of Biological Chemistry* 272:1876-1884.

Leach, M. R., Cohen-Doyle, M. F., Thomas, D. Y., and Williams, D. B. (2002). Localization of the lectin, ERp57 binding, and polypeptide binding sites of calnexin and calreticulin. *Journal of Biological Chemistry* 277:29686-29697.

Levy, A. D., Yang, J., and Kramer, J. M. (1993). Molecular and genetic analyses of the *Caenorhabditis elegans dpy-2* and *dpy-10* collagen genes: a variety of molecular alterations affect organismal morphology. *Molecular Biology of the Cell* 4:803-817.

Lim, A. L., Doyle, S. A., Balian, G., and Smith, B. D. (1998). Role of the pro-alpha2(I) COOH-terminal region in assembly of type I collagen: truncation of the last 10 amino acid residues of pro-alpha2(I) chain prevents assembly of type I collagen heterotrimer. *Journal of Cellular Biochemistry* 71:216-232.

Linnik, K. M. and Herscovitz, H. (1998). Multiple molecular chaperones interact with apolipoprotein B during its maturation. *Journal of Biological Chemistry* 273:21368-21373.

Lucero, H. A. and Kaminer, B. (1999). The role of calcium on the activity of ER calcistorin/protein-disulfide isomerase and the significance of the C-terminal and its calcium

binding. A comparison with mammalian protein-disulfide isomerase. *Journal of Biological Chemistry* 274:3243-3251.

Lynch, A. S., Briggs, D., and Hope, I. A. (1995). Developmental expression pattern screen for genes predicted in the *C. elegans* genome sequencing project. *Nature Genetics* 11:309-313.

Ma, Y. J. and Hendershot, L. M. (2001). The unfolding tale of the unfolded protein response. *Cell* 107:827-830.

Maeda, I., Kohara, Y., Yamamoto, M., and Sugimoto, A. (2001). Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Current Biology* 11:171-176.

Maine, E. M. (2001). RNAi as a tool for understanding germline development in *Caenorhabditis elegans*: uses and cautions. *Developmental Biology* 239:177-189.

Maizels, R. M., Blaxter, M. L., and Selkirk, M. E. (1993). Forms and functions of nematode surfaces. *Experimental Parasitology* 77:380-384.

Marcu, M. G., Chadli, A., Bouhouche, I., Catelli, M. G., and Neckers, L. M. (2000). The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the carboxyl terminus of the chaperone. *Journal of Biological Chemistry* 275:37181-37186.

Masuda, H., Hukumoto, M., Hirayoshi, K., and Nagata, K. (1994). Co-expression of the collagen-binding stress protein HSP47 gene and the $\alpha 1(I)$ and $\alpha 1(III)$ collagen genes in carbon tetrachloride-induced rat liver fibrosis. *Journal of Clinical Investigation* 94:2481-2488.

Mayer, M. P. and Bukau, B. (1999). Molecular chaperones: the busy life of Hsp90. *Current Biology* 9 :R322-R325.

Mazzorana, M., Gruffat, H., Sergeant, A., and van der Rest, M. (1993). Mechanisms of collagen trimer formation. Construction and expression of a recombinant minigene in HeLa cells reveals a direct effect of prolyl hydroxylation on chain assembly of type XII collagen. *Journal of Biological Chemistry* 268:3029-3032.

McGhee, J. D. and Krause, M. W. (1997). In *C. elegans II* (Eds. D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess), pp. 147-184, Cold Spring Harbor Laboratory Press, New York, USA.

McLaughlin, S. H. and Bulleid, N. J. (1998). Thiol-independent interaction of protein disulphide isomerase with type X collagen during intra-cellular folding and assembly. *Biochemical Journal* 331:793-800.

McMahon, L., Muriel, J. M., Roberts, B., Quinn, M., and Johnstone, I. L. (2003). Two sets of interacting collagens form functionally distinct substructures within a *Caenorhabditis elegans* extracellular matrix. *Molecular Biology of the Cell* 14:1366.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO Journal* 10:3959-3970.

Mello, C. and Fire, A. (1995). In *Methods in Cell Biology*, Vol. 48 (Eds. H. F. Epstein and D. C. Shakes), pp. 451-482, Academic Press, San Diego, CA, USA.

Melnick, J., Dul, J. L., and Argon, Y. (1994). Sequential interaction of the chaperones BiP and GRP94 with immunoglobulin chains in the endoplasmic reticulum. *Nature* 370:373-375.

Methods in Cell Biology (1995), Vol. 48 (Eds. H. F. Epstein and D. C. Shakes), Academic Press, San Diego, CA, USA.

Michalak, M., Corbett, E. F., Mesaeli, N., Nakamura, K., and Opas, M. (1999). Calreticulin: one protein, one gene, many functions. *Biochemical Journal* 344:281-292.

Miller, D. M. and Shakes, D. C. (1995). In *Methods in Cell Biology*, Vol. 48 (Eds. H. F. Epstein and D. C. Shakes), pp. 365-394, Academic Press, San Diego, CA, USA.

Myllyharju, J. and Kivirikko, K. I. (1999). Identification of a novel proline-rich peptide-binding domain in prolyl 4-hydroxylase. *EMBO Journal* 18:306-312.

Myllyharju, J. and Kivirikko, K. I. (2001). Collagens and collagen-related diseases. *Annals of Medicine* 33:7-21.

Myllyharju, J., Kukkola, L., Winter, A. D., and Page, A. P. (2002). The exoskeleton collagens in *Caenorhabditis elegans* are modified by prolyl 4-hydroxylases with unique combinations of subunits. *Journal of Biological Chemistry* 277:29187-29196.

Myllyharju, J. (2003). Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biology* 22:15-24.

Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hokawa, N., and Nagata, K. (2000). Embryonic lethality of molecular chaperone Hsp47 knockout mice is associated with defects in collagen biosynthesis. *Journal of Cell Biology* 150:1499-1505.

Nagata, K. (1996). Hsp47: A collagen-specific molecular chaperone. *Trends in Biochemical Sciences* 21:23-26.

Nagata, K. (1998). Expression and function of heat shock protein 47: A collagen-specific molecular chaperone in the endoplasmic reticulum. *Matrix Biology* 16:379-386.

Natsume, T., Koide, S., Yokota, K., Hirayoshi, K., and Nagata, K. (1994). Interactions between collagen-binding stress protein HSP47 and collagen: analysis of kinetic parameters by surface plasmon resonance biosensor. *Journal of Biological Chemistry* 269:31224-31228.

Noiva, R. (1999). Protein disulfide isomerase: The multifunctional redox chaperone of the endoplasmic reticulum. *Seminars in Cell and Developmental Biology* 10:481-493.

Nyström, J., Shen, Z. Z., Aili, M., Flemming, A. J., Leroi, A., and Tuck, S. (2002). Increased or decreased levels of *Caenorhabditis elegans lon-3*, a gene encoding a collagen, cause reciprocal changes in body length. *Genetics* 161:83.

Okimoto, R., Macfarlane, J. L., Clary, D. O., and Wolstenhome, D. R. (1992). The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics* 130:471-498.

Olague-Marchan, M., Twining, S. S., Hacker, M. K., McGrath, J. A., Diaz, L. A., and Giudice, G. J. (2000). A disease-associated glycine substitution in BP180 (type XVII collagen) leads to a local destabilization of the major collagen triple helix. *Matrix Biology* 19:223-233.

Oliver, J. D., Llewelyn Roderick, H., Lewelyn, D. H., and High, S. (1999). ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin. *Molecular Biology of the Cell* 10:2573-2582.

Otterson, G. A., Flynn, G. C., Kratzke, R. A., Coxon, A., Johnston, P. G., and Kaye, F. J. (1994). *stch* encodes the 'ATPase core' of a microsomal stress70 protein. *EMBO Journal* 13:1216-1225.

Otterson, G. A. and Kaye, F. J. (1997). A 'core ATPase', Hsp70-like structure is conserved in human, rat, and *C. elegans* STCH proteins. *Gene* 199:287-292.

Page, A. P. (1997). Cyclophilin and protein disulfide isomerase genes are co-transcribed in a functionally related manner in *Caenorhabditis elegans*. *DNA and Cell Biology* 16:1335-1343.

Page, A. P. (2001). In *Parasitic Nematodes* (Eds. M. W. Kennedy and W. Harnett), pp. 167-193, CABI Press, Oxon, U.K.

Park, B. J., Lee, D. G., Yu, J. R., Jung, S. K., Choi, K., Lee, J., Lee, J., Kim, Y. S., Il Lee, J., Kwon, J. Y., Lee, J., Singson, A., Song, W. K., Eom, S. H., Park, C. S., Kim, D. H., Bandyopadhyay, J., and Ahnn, J. (2001). Calreticulin, a calcium-binding molecular chaperone, is required for stress response and fertility in *Caenorhabditis elegans*. *Molecular Biology of the Cell* 12:2835-2845.

Parodi, A. J. (2000). Protein glucosylation and its role in protein folding. *Annual Review of Biochemistry* 69:69-93.

Patil, C. and Walter, P. (2001). Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Current Opinion in Cell Biology* 13:349-356.

Peixoto, C. A. and de Souza, W. (1992). Cytochemical characterization of the cuticle of *Caenorhabditis elegans*. *Journal of Submicroscopic Cytology and Pathology* 24:425-435.

Peixoto, C. A., de Melo, J. V., Kramer, J. M., and de Souza, W. (1998). Ultrastructural analyses of the *Caenorhabditis elegans rol-6 (su1006)* mutant, which produces abnormal cuticle collagen. *Journal of Parasitology* 84:45-49.

Peixoto, C. A., Alves, L. C., de Melo, J. V., and de Souza, W. (2000). Ultrastructural analyses of the *Caenorhabditis elegans sqt-1(sc13)* left roller mutant. *Journal of Parasitology* 86:269-274.

Peter, F., Nguyen, V. P., and Soling, H. D. (1992). Different sorting of Lys-Asp-Glu-Leu proteins in rat liver. *Journal of Biological Chemistry* 267:10631-10637.

Peters, K., McDowall, J., and Rose, A. M. (1991). Mutations in the *bli-4* (I) locus of *Caenorhabditis elegans* disrupt both adult cuticle and early larval development. *Genetics* 129:95-102.

Piano, F., Schetter, A. J., Mangone, M., Stein, L., and Kemphues, K. J. (2000). RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans. Current Biology* 10:1619-1622.

Piano, F. and Gunsalus, K. (2002). RNAi-based functional genomics in *Caenorhabditis elegans. Current Genomics* 3:69-81.

Podbilewicz, B. and White, J. G. (1994). Cell fusions in the developing epithelia of *C. elegans. Developmental Biology* 161:408-424.

Pouyssegur, J., Shiu, R. P. C., and Pastan, I. (1977). Induction of two transformation-sensitive membrane polypeptides in normal fibroblasts by a block in glycoprotein synthesis or glucose deprivation. *Cell* 11:941-947.

Priess, J. R. and Hirsh, D. I. (1986). *Caenorhabditis elegans* morphogenesis: the role of the cytoskeleton in elongation of the embryo. *Developmental Biology* 117:156-173.

Primm, T. P., Walker, K. W., and Gilbert, H. F. (1996). Facilitated protein aggregation. Effects of calcium on the chaperone and anti-chaperone activity of protein disufide-isomerase. *Journal of Biological Chemistry* 271:33664-33669.

Prockop, D. J. and Kivirikko, K. I. (1995). Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry* 64:403-434.

Queitsch, C., Sangster, T. A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature* 417:618-624.

Reboul, J., Vaglio, P., Tzellas, N., Thierry-Mieg, N., Moore, T., Jackson, C., Shin-i, T., Kohara, Y., Thierry-Mieg, D., Thierry-Mieg, J., Lee, H., Hitti, J., Doucette-Stamm, L., Hartley, J. L., Temple, G. F., Brasch, M. A., Vandenhaute, J., Lamesch, P. E., Hill, D. E., and Vidal, M. (2001). Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans. Nature Genetics* 27:332-336.

Reinke, V. (2002). Defining development through gene expression profiling. *Current Genomics* 3:95-109.

Ristoratore, F., Cermola, M., Nola, M., Bazzicalupo, P., and Favre, R. (1994). Ultrastructural imuno-localisation of CUT-1 and CUT-2 antigenic sites in the cuticles of the nematode *Caenorhabditis elegans. Journal of SubmicroscopicCcytology and Pathology* 26:437-443.

Roy, P. J., Zheng, H., Warren, C. E., and Culotti, J. G. (2000). *mab-20* encodes Semaphorin-2a and is required to prevent ectopic cell contacts during epidermal morphogenesis in *Caenorhabditis elegans. Development* 127:755-767.

Rual, J. F., Lamesch, P., Vandenhaute, J., and Vidal, M. (2002). The *Caenorhabditis elegans* interactome mapping project. *Current Genomics* 3:83-93.

Rubin, G. M., Yandell, M. D., Wortman, J. R., Miklos, G. L. G., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S. B., Gong, F. C., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J. Y., Li, Z. Y., Hynes, R. O., Jones, S. J. M., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. Q. H., Zhong, F., Zhong, W. Y., Gibbs, R., Venter, J. C., Adams, M. D., and Lewis, S. (2000). Comparative genomics of the eukaryotes. *Science* 287:2204-2215.

Rutherford, S. L. and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* 396:336-342.

Saito, Y., Ihara, Y., Leach, M. R., Cohen-Doyle, M. F., and Williams, D. B. (1999). Calreticulin functions *in vitro* as a molecular chaperone for both glycosylated and nonglycosylated proteins. *EMBO Journal* 18:6718-6729.

Sambrook, J, Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual* (Eds. N. Ford, C. Nolan, M. Fergurson), Cold Spring Harbor Laboratory Press, New York, USA.

Schauber, C., Chen, L., Tongaonkar, P., Vega, I., Lambertson, D., Potts, W., and Madura, K. (1998). Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature* 391:715-718.

Scheibel, T. and Buchner, J. (1998). The Hsp90 complex - A super-chaperone machine as a novel drug target. *Biochemical Pharmacology* 56:675-682.

Schnabel, H. and Schnabel, R. (1990). Early determination in the *C. elegans* embryo: a gene, *cib-1*, required to specify a set of stem-cell-like blastomeres. *Development* 108:107-119.

Schrag, J. D., Bergeron, J. J. M., Li, Y. G., Borisova, S., Hahn, M., Thomas, D. Y., and Cygler, M. (2001). The structure of calnexin, an ER chaperone involved in quality control of protein folding. *Molecular Cell* 8:633-644.

Schüler, H. (2001). ATPase activity and conformational changes in the regulation of actin. *Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology* 1549:137-147.

Shaw, L. M. and Olsen, B. R. (1991). Facit collagens: diverse molecular bridges in extracellular matrices. *Trends in Biochemical Sciences* 16:191-194.

Shen, X. H., Ellis, R. E., Lee, K., Liu, C. Y., Yang, K., Solomon, A., Yoshida, H., Morimoto, R., Kurnit, D. M., Mori, K., and Kaufman, R. J. (2001). Complementary signaling pathways regulate the unfolded protein response and are required for *C. elegans* development. *Cell* 107:893-903.

Simons, J. F., Ferro-Novick, S., Rose, M. D., and Helenius, A. (1995). BiP/Kar2p serves as a molecular chaperone during carboxypepidase Y folding in yeast. *Journal of Cellular Biology* 130:41-49.

Singh, R. N. and Sulston, J. E. (1978). Some observations on moulting in *Caenorhabditis elegans. Nematologica* 24:63-71.

Sondermann, H., Scheufler, C., Schneider, C., Höhfeld, J., Hartl, F. U., and Moarefi, I. (2001). Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. *Science* 291:1553-1557.

Sternberg, P. W. (2001). Working in the post-genomic *C. elegans* world. *Cell* 105:173-176.

Stevens, F. J. and Argon, Y. (1999). Protein folding in the ER. *Seminars in Cell and Developmental Biology* 10:443-454.

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

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

Tabara, H., Grishok, A., and Mello, C. C. (1998). RNAi in *C. elegans*: soaking in the genome sequence. *Science* 282:430-431.

Takechi, H., Hosokawa, N., Hirayoshi, K., and Nagata, K. (1994). Alternative 5' splice site selection induced by heat shock. *Molecular and Cellular Biology* 14:567-575.

Tasab, M., Batten, M. R., and Bulleid, N. J. (2000). Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. *EMBO Journal* 19:2204-2211.

Tasab, M., Jenkinson, L., and Bulleid, N. J. (2002). Sequence-specific recognition of collagen triple helices by the collagen-specific molecular chaperone HSP47. *Journal of Biological Chemistry* 277:35007-35012.

Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A., and Driscoll, M. (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nature Genetics* 24:180-183.

Thacker, C., Peters, K., Srayko, M., and Rose, A. M. (1995). The *bli-4* locus of *Caenorhabditis elegans* encodes structurally distinct kex2/subtilisin-like endoprotesases essential for early development and adult morphology. *Genes and Development* 9:956-971.

Thacker, C., Marra, M. A., Jones, A., Baillie, D. L., and Rose, A. M. (1999). Functional genomics in *Caenorhabditis elelgans*: an approach involving comparisons of sequences from related nematodes. *Genome Research* 9:348-359.

Thacker, C. and Rose, A. M. (2000). A look at the *Caenorhabditis elegans* Kex2/subtilisin-like proprotein convertase family. *Bioessays* 22:545-553.

Thacker, C., Srayko, M., and Rose, A. M. (2000). Mutational analysis of *bli-4/kpc-4* reveals critical residues required for proprotein convertase function in *C. elegans*. *Gene* 252:15-25.

Thein, M. C., McCormack, G., Winter, A. D., Johnstone, I. L., Shoemaker, C. B., and Page, A. P. (2003). *Caenorhabditis elegans* exoskeleton collagen COL-19: an adult-specific marker for collagen modification and assembly, and the analysis of organismal morphology. *Developmental Dynamics* 226:523-539.

Timmons, L. and Fire, A. (1998). Specific interference by ingested dsRNA. *Nature* 395:854.

Timmons, L., Court, D., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263:103-112.

Trombetta, E. S. and Helenius, A. (1998). Lectins as chaperones in glycoprotein folding. *Current Opinion in Structural Biology* 8:587-592.

Trombetta, E. S. and Parodi, A. J. (2002). N-glycan processing and glycoprotein folding. *Advances in Protein Chemistry* 59:303-344.

Tsai, B., Ye, Y. H., and Rapoport, T. A. (2002). Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nature Reviews Molecular Cell Biology* 3:246-255.

Van der Keyl, H., Kim, H., Espey, R., Oke, C. V., and Edwards, M. K. (1994). *Caenorhabditis elegans sqt-3* mutants have mutations in the *col-1* collagen gene. *Developmental Dynamics* 201:86-94.

von Mende, N., Bird, D. M., Albert, P. S., and Riddle, D. L. (1988). *dpy-13:* a nematode collagen gene that affects body shape. *Cell* 55:567-576.

Walmsley, A. R., Batten, M. R., Lad, U., and Bulleid, N. J. (1999). Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. *Journal of Biological Chemistry* 274:14884-14892.

Ware, F. E., Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M. A., and Williams, D. B. (1995). The molecular chaperone calnexin binds $Glc_1Man_9GlcNAc_2$ oligosaccharide as a initial step in recognizing unfolded glycoproteins. *Journal of Biological Chemistry* 270:4697-4704.

Welihinda, A. A., Tirasophon, W., and Kaufman, R. J. (1999). The cellular response to protein misfolding in the endoplasmic reticulum. *Gene Expression* 7:293-300.

Wetterau, J., Combs, K. A., McLean, L. R., Spinner, S. N., and Aggerbeck, L. P. (1991). Protein disulfide isomerase appears necessary to maintain the catalytically active structure of the microsomal triglyceride transfer protein. *Biochemistry* 30:9728-9735.

Wilbanks, S. M., DeLuca-Flaherty, C., and Mckay, D. B. (1994). Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. I. Kinetic analyses of active site mutants. *Journal of Biological Chemistry* 17:12893-12898.

Williams, B. D. (1995a). In *Methods and Cell Biology*, Vol. 48 (Eds. H. F. Epstein and D. C. Shakes), pp. 81-96, Academic Press, San Diego, USA.

Williams, D. B. (1995b). Calnexin: A molecular chaperone with a taste for carbohydrate. *Biochemical Cell Biology* 73:123-132.

Wilson, R., Lees, J. F., and Bulleid, N. J. (1998). Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen. *Journal of Biological Chemistry* 273:9637-9643.

Winter, A. D. and Page, A. P. (2000). Prolyl 4-hydroxylase is an essential procollagen-modifying enzyme required for exoskeleton formation and the maintenance of body shape in the nematode *Caenorhabditis elegans*. *Molecular and Cellular Biology* 20:4084-4093.

Yang, J. and Kramer, J. M. (1994). In vitro mutagenesis of *Caenorhabditis elegans* cuticle collagens identifies a potential subtilisin-like protease cleavage site and demonstrates that carboxyl domain disulfide bonding is required for normal function but not assembly. *Molecular and Cellular Biology* 14:2722-2730.

Yang, J. and Kramer, J. M. (1999). Proteolytic processing of *Caenorhabditis elegans* SQT-1 cuticle collagen is inhibited in right roller mutants whereas cross-linking is inhibited in left roller mutants. *Journal of Biological Chemistry* 274:32744-32749.

Young, J. M. and Hope, I. A. (1993). Molecular markers of differentiation in *Caenorhabditis elegans* obtained by promoter trapping. *Developmental Dynamics* 196:124-132.

Young, J. C., Moarefi, I., and Hartl, F. U. (2001). Hsp90: a specialized but essential protein-folding tool. *Journal of Cell Biology* 154:267-273.

Zeng, C. B., Aleshin, A. E., Chen, G. J., Honzatko, R. B., and Fromm, H. J. (1998). The roles of glycine residues in the ATP binding site of human brain hexokinase. *Journal of Biological Chemistry* 273:700-704.

Zwaal, R. R., Broeks, A., Vanmeurs, J., Groenen, J. T. M., and Plasterk, R. H. A. (1993). Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. *Proceedings of the National Academy of Sciences of USA* 90:7431-7435.