INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

The *Caenorhabditis elegans* Clock gene *gro-1* encodes a metazoan N^6-(Δ^2) isopentenyl PPi: tRNA isopentenyl transferase.

Jason Lemieux

The Department of Biology

McGill University

Montreal, Canada

March 1999

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

Abstract

The *Caenorhabditis elegans* gene *gro-1* belongs to the Clock group of genes. The four known genes making up this class are believed to be involved in a general mechanism acting to coordinate the time-dependent processes in the organism. Mutation of these genes alter the timing of many disparate processes. This results in the mean lengthening of embryonic and post-embryonic development, as well as in a lengthening of the periods of a number of adult behaviours including pharyngeal pumping, swimming and the defecation cycle. These mutants also exhibit a significantly increased life span. The gene *gro-1* has been cloned and encodes a metazoan N^6-(Δ^2) isopentenyl PPi: tRNA isopentenyl transferase. In *S. cerevisiae* and bacteria this enzyme has been shown to modify tRNAs that code for codons begining with U. This modification consists of the isopentenylation of the adenine residue at position 37, adjacent to the anti-codon. Interestingly, *gro-1* is the fifth member of an operon. Preliminary expression studies with GFP reporter constructs suggest that as in yeast, GRO-1 is expressed in the cytoplasm and mitochondria, as well as in the nucleus.

Abrégé

Le gène *Caenorhabditis elegans gro-1* appartient au groupe Clock des gènes. On croit que les quatres gènes connus qui font partie de cette classification sont engagés dans un méchanisme général qui agit à coordiner les opérations qui sont sensible du temps dans l'organisme. La mutation de ces gènes change l'horaire de plusieurs opérations différentes. Il s'ensuit qu'il y a une prolongation du développement embryonique et post-embryonique. En même temps, il y a aussi une prolongation des périodes dans une grande partie du comportement des adultes y compris le pompage pharyngéal, la natation et le mouvement de défécation. Ces mutants démontrent aussi une vie énormement prolongée. La gène *gro-1* a été cloner et coder un metazoan N^6-(Δ^2) isopentenyl , PPi: tRNA ispoentenyl transferase. On a prouvé que cette enzyme, dans *s. cerevisiae* et bactéria, modifie tRNAs qui codent pour des codons commençants avec U. Cette modification consiste de l'isopentenylation du résidu de l'adenine à la position 37 près de l'anti-codon. C'est bien intéressant que le *gro-1* est le cinquième membre d'un operon. Des études préliminaires d'expression avec des mécaniques GFP rapporteur suggèrent que, comme dans la levure, GRO-1 s'exprime dans le cytoplasme et la mitochondria aussi bien que dans le nucléon.

Table of Contents

Abstract.....ii

Abrégé.....ii

Table of Contents.....iv

List of Figures.....vii

List of Tables.....viii

Acknowledgements.....viiii

Introduction.....1

Clocks.....1

Aging.....3

The Clock genes.....8

Cloning the Clock genes.....11

The modification of transfer RNA.....13

Isopentenyl adenosine in bacterial RNA.....15

Isopentenyl adenosine in the tRNA of S. cerevisiae.....18

Materials and Methods.....23

Publicly available data, strains and clones.....23

Microinjection and molecular techniques required for cloning.....23

DNA amplification by PCR.....25

Sequencing.....25

Search for the *e2400* mutation.....25

Confirmation of the predicted splicing patterns.....26

Determining membership in the *gro-1* operon..... 29

Identification and sequencing of the EST corresponding to the human homologue of *gro-1*..... .30

Simplifying the operon: pMQ8..... 30

The green fluorescent protein (GFP) reporter construct pMQ18.....33

Results..... ...35

Cosmid rescue..... 35

Subcloning.....35

The identification of the *e2400* lesion.....39

"Operonicity"..... 40

The conformation of predicted splicing patterns.....42

The human *gro-1* homologue..... 45

Expression pattern of the GRO-1::GFP reporter.....45

Discussion.....48

The Clock gene *gro-1* has been cloned..... 48

GRO-1 and its homologues..... 49

The *e2400* allele..... 50

gro-1 expression..... 51

The other genes in the *gro-1* operon.....52

How does *gro-1* fit into the "clock" hypothesis 54

Summary.....57

References.....59

Appendices.....69

Appendix A: Sequence of the *gro-1* operon 69

Appendix B: Human *gro-1* sequence.....79

Appendix C: pMQ8..... 80

Appendix D: Towards a GRO-1 antibody.....81

Appendix E: Towards GRO-1::GFP mislocalization.....83

Appendix F: Sequences of all the primers 86

List of Figures

Figure 1: Bipartite PCR and the construction of pMQ8.....31

Figure 2: The *gro-1::gfp* reporter construct..... 34

Figure 3: The *gro-1* genetic map..... 36

Figure 4: The *gro-1* physical map.....37

Figure 5: *gro-1* rescue..... 38

Figure 6: The *gro-1* operon..... 41

Figure 7: The gro-1 sequence..... 43

Figure 8: Alignment of *gro-1* and its homologues.....46

List of Tables

Table 1: Confirming the predicted splicing pattern of the *gro-1* operon..... 28

Acknowledgements

I would like to thank my supervisor Siegfried Hekimi for the opportunity to work with him on such an interesting and challenging project. I would also like to thank Bernard Lakowski, Tom Barnes, Jonathan Ewbank, Brent McCright and Claire Bénard for their support, guidence and instruction. Claire also deserves thanks for translating the abstract of this thesis.

I would like to take this opportunity to thank my parents, Yves and Janet, who have worked so hard to ensure my excellent education and upbringing, and who have always been there for support. It means more to me than you will ever know. Thank you.

Introduction

Clocks

Throughout history, man has endeavored to subdivide time into increasingly smaller units. Probably the earliest example of this is the week. In different parts of the world the week has varied from five to ten days, but third century Rome settled at seven. With the advent of the sundial, the water clock and the pendulum clock, man has managed to define the units we know today: the hour, the minute and the second. Without these invented measures of time, modern man would be lost in the rhythm of the universe that had for so long dictated human life. The earliest measures of time were an attempt to predict the cyclic nature of the universe; for example, the repeat of the seasons so crucial to agriculture. The Babylonians and the ancient Greeks invented complex calendars based on the lunar cycle, the most obvious of nature's cues. Their calendar had to be adjusted to the solar year by adding extra months. These adjustments undermined the attempts of the early calendar to coordinate people with their environment, as different towns and regions added months according to their own liking. The Egyptians were the first to devise a calendar based on the solar year. This 365¼ day calendar was adopted by Julius Caesar and its inaccuracy corrected by Pope Gregory XIII in 1582 to give us the Gregorian Calendar that we use today (the actual solar year of 365 days, 5 hours, 48 minutes and 46 seconds is 11 minutes and 14 seconds shorter than the original Egyptian year).

Independent of man's attempts to measure and organise time, evolution has developed its own clocks to keep organisms synchronised with their environments. Throughout biology there are behavioural and physiological examples of rhythmicity and timing. Circadian clocks allow organisms to stay in time with, for example, the day/night, lunar or seasonal cycles. When an organism is sequestered and cut off from environmental cues (such as light/dark changes), the timed behaviours continue. This "free-running" suggests that the circadian clocks are independent of environmental cues and an intrinsic part of the organism. Ultradian rhythms such as walking, heart rate, and sleep patterns are examples of other rhythmic biological processes.

The circadian clock has been well studied, both morphologically and genetically. Vertebrate circadian pacemakers have been localised to the pineal gland in the cases of birds, reptiles and fish, and to the hypothalamus in the case of mammals. These regions have been shown to secrete neuropeptides and hormones with a circadian rhythmicity (Klein *et al.*, 1991; Foulkes *et al.*, 1997). In a number of insects, circadian pacemakers are known to exist in the optic lobe (Fleissner, 1982; Loker, 1972; 1974; Page, 1984; and Sokolove and Loker, 1975).

Genetics has also contributed to the understanding of circadium clocks. *period* (*per*) and *timeless* (*tim*) are two genes known in *Drosophila melanogaster*. These genes are expressed in the CNS and are required for the maintenance of the flies

circadium rhythm. Different mutations have been known to shorten (per^S) and lengthen (per^l; Konopka and Benzer, 1971) as well as completely eliminate (per^0; Hardin *et al.*, 1990; Zerr *et al.*, 1990) the circadian rhythm. The *Neurospora crassa frequency* (*frq*) gene is involved in the regulation of the circadian clock responsible for timing asexual reproduction. The level of *frq* mRNA oscillates with the circadian rhythm, peaking at the time of spore production (Gardner and Feldman, 1980). Genes involved in mammalian circadian clocks have been identified. The *tau* gene of the golden hamster *Mesocricelius auratus* alters the normal 24 hour circadian rhythm associated with locomotion to 22 hours in heterozygous mutants and to 20 hours in homozygous mutants (Ralph and Menaker, 1988). In mice three homologues of the *Drosophila per* gene are known. *mper1*, *mper2* and *mper3* are all expressed in the CNS (Albrecht *et al.*, 1997; Shearman *et al.*, 1997; Sun *et al.*, 1997; and Tei *et al.*, 1997). Another gene, *clock*, has also been shown to effect circadian rhythm in the mouse (King *et al.*, 1997). In the nematode *C. elegans*, the Clock genes have been implicated in biological timing, based on the slowing of many developmental and behavioural processes in the mutants (Hekimi *et al.*, 1995; Wong *et al.*, 1995). The Clock genes are examined in detail below.

Aging

Of particular interest in any discussion of time is the life span of an organism. How long will it take for the cells or organs of an organism to fail and cause death? This is a problem that has been thoroughly debated and much scientific evidence has

been generated in the past couple of decades to support a number of the proposed theories.

Many of these theories support the notion of a degenerative senescence; that death is caused by the gradual failure of organs or critical processes. Among the more popular are the somatic mutation, the mitochondrial, the evolutionary, and the free-radical theories of aging. One of the earliest theories of aging is the rate of living hypothesis, which suggests that energy consumption is responsible for senescence (Pearl, 1928; Rubner, 1908; Sohal, 1976). This would mean that the higher the metabolic rate of an organism, the shorter its maximum life span (Rubner, 1908). More recent theories such as the mitochondrial and free radical theories (discussed below) have a basis in the rate of living hypothesis.

It has been proposed that the accumulation of mutations in somatic DNA is a factor promoting degenerative senescence (Bohr and Anson, 1995; Evans *et al.*, 1995; Miquel, 1992; Morley, 1995; Vilg and Gossen; 1993). The majority of these mutations would be corrected by the cell's repair mechanisms, but as time progresses the chance of hitting a very critical gene, such as a component of the repair mechanism itself, increases.

The mitochondrial and the mtDNA mutation theories of aging suggest that either through mtDNA mutation or simply mitochondrial tiring, dysfunction of the

mitochondria would lead to the death of the organism (Luft, 1994). This lethality would stem from the mitochondria's critical role as energy supplier to the cell.

The evolutionary theory of antagonistic pleiotropy suggests that natural selection acts in favor of factors promoting early reproduction rather than the preservation of non-germ-line cells (Kirkwood and Rose, 1991; Kirkwood, 1997; Rose and Finch, 1994). For a mutation to aid in post-reproductive preservation, it therefore must not interfere with processes involved in reproduction.

Probably the most popular mechanism used to explain aging is that of destructive free-radicals. This theory is supported by a large body of experimental evidence and also manages to include and support the above hypotheses. This mechanism involves the oxidative destruction of biological macromolecules such as lipid, nucleic acid and protein (Harman, 1956). A group of free radicals, collectively referred to as oxidants, can be produce *in vivo* from O_2 (reviewed in Jazwinski, 1996). Experiments have shown that mitochondria are a major source of endogenous superoxide anion ($O_2^{-}\bullet$), an oxidant, the result of an imperfect electron transport chain (Chance *et al.*, 1979). This would be a likely mechanism to explain the rate of living hypothesis; the more energy used by an organism the greater the level of oxidant production and the shorter the life span. The mutagenic nature of oxidants has been characterized (Feig *et al.*, 1994; Grollman and Moriya, 1993) and presents a mechanism for the accumulation of mutations. Furthermore, if these reactive by-products are produced by the electron transport chain, the mitochondria

would be the first organelle attacked. This is a likely mechanism to explain the mitochondrial and mtDNA theories of aging.

It has also been proposed that death may be programmed; the final stage of the developmental process. Such a notion is supported by the differential expression of genes in later adulthood, after reproduction is complete. With the adult body and reproduction complete, what could these genes be doing? (Kelly, 1989; Sameroff, 1983). Although, the idea of a programmed death and the theories of a degenerative senescence are by no means mutually exclusive. For example, the program could be a relief of the cell's oxidant defense and repair mechanisms in order to promote degenerative senescence.

As with the biological clock, genetics has also contributed to the understanding of aging. To suggest that genes have a role in senescence does not imply that a genetic program controls life span. Aging must be considered a stochastic process, where life span is affected by many environmental and biochemical determinants, and the genetic constitution of the organism also has its effects (Jazwinski, 1996). Mutations in many of the favorite model systems have been identified that affect, either positively or negatively, the maximum life span of the organism. Experiments in the budding yeast *Saccharomyces cerevisiae*, the nematode *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster* have all contributed to the genetic study of aging. Genetics has been used in mammalian systems, but has provided much less insight into the aging process.

In *Saccharomyces cerevisiae*, life span is measured by the number of progeny produced by a particular cell rather than by the time elapsed from birth or adulthood to senescence (Mortimer and Johnston, 1959). A mother cell dies after a finite number of cell divisions. A number of genes have been identified that seem to be involved in yeast ageing. *LAG1* has been shown to determine both mean and maximum life span (D'mello *et al.*, 1994), and *RAS1* and *RAS2* manipulation shortens and increases life span respectively (Sun *et al.*, 1990). *RAS2* acts as an interface between the cell and the environment, sensing the nutritional status of the cell and responding to stresses such as starvation and crowding (Broach and Deschenes, 1990). Jazwinski (1993) has reported an increase in metabolic capacity and efficiency in the *RAS2* long lived cells, and suggests that this is likely the cause of the long life.

A number of genes in *C. elegans* have been shown to affect aging. One group of genes that has been well studied are those involved in Dauer formation. The Dauer state is an alternative stage of larval development that L2 larvae enter under certain environmental stresses such as starvation and crowding (Riddle, 1987). The genes *daf-2* and *daf-23* are of the Daf-c (Dauer Formation constitutive) type. Some of these alleles are temperature sensitive and when the worms develop at the restrictive temperature they enter the Dauer pathway regardless of the growth conditions. But when these mutants are allowed to develop at the permissive temperature, normal development is observed and the Dauer pathway is only entered in the presence of

the proper cues. Once adults, if mutants are shifted to the restrictive temperature, an increase in longevity is observed (Kenyon *et al.*, 1993; Larsen *et al.*, 1995). A second set of genes involved in Dauer control is the Daf-d (defective) genes. *daf-16* is an example of a Daf-d gene and mutant animals are unable to enter the Dauer pathway (Riddle *et al.*, 1981). d*af-16* mutations also suppress the constitutive Dauer phenotype and increased life span resulting from *daf-2* and *daf-23* mutations (Kenyon *et al.*, 1993; Larsen *et al.*, 1995). The gene *age-1* (actually an allele of *daf-23*) exhibits long life and is suppressed by *daf-16* (Larsen *et al.*, 1995). The genetic interaction suggests that the Dauer genes may be involved in a common system affecting life span in *C. elegans* (Larsen *et al.*, 1995).

The Clock genes are another set of genes known to affect aging in *C. elegans* (Hekimi *et al.*, 1995; Lakowski and Hekimi, 1996). They are discussed in detail below.

The Clock Genes

The Clock genes are a class of genes that when mutated affect developmental and behavioral timing (Hekimi *et al.*, 1995; Wong *et al.*, 1995). There are four members to this class: *clk-1*, *clk-2*, *clk-3* and *gro-1*. The first three were uncovered by Hekimi *et al.* (1995) in a screen for maternal-effect viable mutations. Of them, only *clk-1* has been thoroughly studied: a comprehensive genetic analysis by Wong *et al.* (1995) and the molecular cloning by Ewbank *et al.* (1997). *gro-1* was uncovered as

a spontaneous mutant from the wild type strain PaC1 by Jonathan Hodgkin (Hodgkin and Doniach, 1997) and has been shown to exhibit the Clock mutant phenotype (Wong *et al.*, 1995; Lakowski and Hekimi, 1996).

The detailed phenotypic analysis by Wong *et al.* (1995) revealed that *clk-1* mutants exhibit a mean lengthening of embryonic and post-embryonic development, as well as a lengthening of the periods of a number of adult behaviors including pharyngeal pumping, swimming and the defecation cycle. Although not studied as thoroughly, the other *Clk* genes exhibit the same phenotype as *clk-1* (Hekimi *et al.*, 1995). Subsequent work by Lakowski and Hekimi (1996) has shown that mutations in the Clock genes also increase both mean and maximum life span.

Another interesting feature of the *Clk* genes is their maternal effect. The screen by Hekimi *et al.* (1995) that isolated *clk-1*, *clk-2* and *clk-3* was designed to only pick up genes with a maternal effect. *gro-1* was isolated originally based on its slow development and the maternal effect was only detected after closer examination (Wong *et al.*, 1995). In the case of maternal effect mutations, homozygous mutant progeny (*clk-1*/*clk-1*) from heterozygous hermaphrodites (*clk-1*/+) are phenotypically wild type. Due to the full zygotic rescue of the *Clk* genes (Hekimi *et al.*, 1995), only progeny homozygous for the mutant allele from homozygous mutant hermaphrodites will exhibit the mutant phenotype.

Using a genetic approach to study the *Clk* genes, Lakowski and Hekimi (1996) looked at the genetic interactions of the *Clk* genes by constructing double mutants. These double mutants took longer to develop than each of the individual mutations and most had a longer mean and maximum life span. Contradicting this general trend were double mutants made with *gro-1(e2400)*. While each of the mutations on their own increased mean and maximum life span, *gro-1* with *clk-2* and *clk-3* had a mean life span not significantly different than that of the wild type. However, this suppression of extended life span was not observed when *gro-1* was paired with *age-1*.

When the suppressing activity of *daf-16* was tested on the Clk genes, it was unable to suppress the long-life of *clk-1*, *clk-3* and *gro-1* (*clk-2* was not tested; Lakowski and Hekimi, 1996). This suggests that while the dauer genes may act by a common mechanism affecting life span, the Clk genes function in a distinct pathway (Lakowski and Hekimi, 1996). However, this evidence does not exclude the possibility that the Clk genes act downstream of the Dauer genes. But the fact that with the exception of the aging phenotype there is absolutely no overlap between the two phenotype and that these phenotypes are additive, lend strong support to the notion of distinct pathways.

An interesting aspect of the Clk genes that is absent from *daf-2* and *age-1* containing strains, is that development and mean adult life span are affected in a linear fashion (Lakowski and Hekimi, 1996). That is, when one of the Clk genes is

mutant, both the length of development and mean life span are increased. In the case of *daf-2* and *age-1*, when mean adult life span is increased, the length of development is decreased. This and the fact that *clk-1* has been shown to affect a diverse range of timed events has led to the proposal that the Clk genes are involved in some general physiological clock coordinating timed events in the development and life of the worm (Lakowski and Hekimi, 1996). Such a mechanism could impact on life span in a number of ways. For example, it seems that all of metabolism is slowed in the Clock mutants. This would result in the decreased production of the destructive metabolic by-products believed to be a cause of degenerative senescence (Harman, 1981; Orr and Sohal, 1994). The importance of the Clock genes is highlighted by the sterility seen at 25°C of *clk-2* and *gro-1* worms (Lakowski, 1998.). This sterility suggests an essential role. However the Clock genes effect life span, understanding the Clock mechanism would shed light on many important and fundamental biological questions.

Cloning the Clock genes

An essential step in the understanding of the Clock mechanism is determining the molecular identity of the genes involved. *clk-1* has been cloned by Ewbank *et al.* (1997) and shown to encode a protein similar to the *S. cerevisiae* gene *CAT5/COQ7*. Homologues have also been identified in rat (Jonassen *et al*, 1996) and human (Ewbank *et al.*, 1997).

The remaining three *Clk* genes have been genetically mapped to varying degrees of accuracy. *clk-2* has been assigned to the third linkage group (LGIII), along with *clk-1*, and mapped between *sma-4* and *mab-5* (Hekimi *et al.*, 1995). *clk-3* has been assigned to the second linkage group and mapped to the left of *eat-2*, near the end of the chromosome's right arm (Hekimi *et al.*, 1995).

Mapping by Hodgkin and Doniach (1997) had placed *gro-1* in the vicinity of *clk-1*. Examination of the *gro-1(e2400)* phenotype revealed that not only did it grow slowly as Hodgkin and Doniach (1997) reported, but it also had a reduced defecation rate, lived long and was maternally rescued (Wong *et al.*, 1995, Lakowski and Hekimi, 1996). Based on their similar phenotypes it was possible that the only *gro-1* mutant allele, *e2400*, could be another allele of *clk-1*. In order to facilitate the cloning of *gro-1*, Lakowski (1998) carried out further genetic mapping in an attempt to separate the two genes. Initially, *gro-1* was placed with *clk-1*, approximately 0.2 cM to the right of *dpy-17* on linkage group (LG) III. However, further mapping separated the two genes and indicated that *e2400* was 0.3 cM to the left of *e2519* (see Figure 3). With the cloning of *clk-1* by Ewbank *et al.* (1997) it was possible to transfer by mating into *gro-1* one of the extrachromosomal arrays that was able to rescue *clk-1(e2519)*. It was unable to rescue the *gro-1* mutant phenotype (Lakowski, 1998). Based on these results it was concluded that *gro-1* and *clk-1* were different genes. From the genetic mapping it was possible to estimated that *gro-1* should lie approximately 20 kb to the left of *clk-1* (Lakowski, 1998; see Figure 4).

With such precise mapping data, I was able to clone *gro-1* by injection rescue. *gro-1* encodes a protein homologous to the enzyme N^6-(Δ^2) isopentenyl PPi :tRNA isopentenyl transferase (IPP Transferase; Bartz *et al.*, 1970), which has been shown to modify an adenosine nucleotide of certain species of tRNA (reviewed by Björk, 1987.). This protein has homologues throughout evolution including yeast Mod5p (Dihanich *et al.*, 1987) and bacterial *miaA* (Connolly and Winkler, 1991). There exists a considerable body of work concerning these genes, their function and their regulation. The following is a review of this work.

The Modification of Transfer RNA

Transfer RNA from eubacteria to higher eukaryotes contain a wide variety of modified nucleosides. There have been more than 75 modifications identified, found in all three phylogenetic domains (Edmonds *et al.*, 1991). Many of the genes directly responsible for individual modifications have been identified. In some of these mutants, the absence of the particular modification has little or no effect on the cell's growth rate. For example, the *E. coli trmA5* and *S. cerevisiae trm2* mutants lack 5-methyluridine at position 54 (m^5U54) and their growth rates are altered very little as a result (Björk and Neidhardt, 1975; Hopper *et al.*, 1982). But the absence of other modifications cause a significant effect on growth rate. A reduction of 20-50% has been observed in the bacterial mutants *miaA*, *trmD* and *hisT* which lack the

modified nucleosides 2-methylthio-N^6-(*cis*-hydroxyisopentenyl)-adenosine at position 37 (ms^2i^6A37), 1-methylguanosine also at position 37 (m^1G37) and pseudouridine (ψ) at positions 38, 39 and 40, respectively (Ericson and Björk, 1986; Palmer *et al.*, 1983 and Björk *et al.*, 1989).

A number of roles have been assigned to these modifications. The primary and most obvious role is in translation. Modifications at position 34 of tRNA (the Wobble position of the anticodon) affect the tertiary structure of the base and its ability to form hydrogen bonds. This has the result of either extending or restricting the decoding capacity of the tRNA (reviewed by Persson, 1993). For tRNA specific for A-U rich codons, modifications at position 37 (3' end of the anticodon) act to stabilize and thus compensate for the weaker base pairing (reviewed by Grosjean and Chantrenne, 1980). Modifications at position 37 are also believed to aid in maintaining the translational frame (Jukes, 1973; Pieczenik, 1980). *trmD3* mutant bacteria lack the m^1G37 modified nucleoside which results in a +1 frameshift (Björk *et al.*, 1989; Hagervall *et al.*, 1992). The role of modified nucleosides in tRNA have also been linked, through their effects on the translation apparatus, to various regulatory processes in bacterial metabolism (Buck and Ames, 1984; reviewed by Persson, 1993). Based on the wide variety of phenotypes seen in the absence of individual modifications and on their involvement in many different metabolic and molecular processes, it is believed that modified nucleosides serve a crucial role in the fine tuning of the translational apparatus and that through this role may act in many regulatory processes.

The addition of an isopentenyl group is one of the modifications found on adenosine at position 37. Isopentenyl-adenosine 37 is found in tRNA specific for codons that begin with U. These include $tRNA^{Phe}$, $tRNA^{Tyr}$, $tRNA^{Cys}$, $tRNA^{Trp}$ and certain species of $tRNA^{Leu}$ and $tRNA^{Ser}$ (reviewed by Björk, 1987). While in eukaryotes A37 is simply isopentenylated, in bacteria the adenosine can be further modified, the primary derivative being 2-(methylthio)-N^6-(Δ^2-isopentenyl) adenosine (ms^2i^6A; Agris *et al.*, 1975; Geffer, 1969). For all the various derivatives, the first step is the addition of the isopentenyl group. This step is catalyzed by the enzyme N^6-(Δ^2) isopentenyl PPi : tRNA isopentenyl transferase (IPP Transferase; Bartz *et al.*, 1970). The gene encoding the transferase has been cloned in a number of organisms including bacteria (*miaA*; Caillet and Droogmans, 1988; Yanofsky and Söl, 1977) and yeast (*MOD5*; Dinanich *et al.*, 1987).

Isopentenyl Adenosine in Bacterial tRNA

In *miaA* mutants, the lack of the modified nucleosides at position 37 has effects on cellular growth, translation, and the regulation of gene expression. In *S. typhimurium*, *miaA*$^-$ cells exhibit an up to 50% decrease in growth rate. Significant changes in the sensitivity to a variety of amino acid analogs have also been reported (Ericson and Björk, 1986). At the level of translation, *miaA* mutants show a 30% decrease in translation chain elongation, a decrease in translational error, an increased sensitivity to codon context, and a decreased stability of codon-anticodon

interactions (Ericson and Björk, 1986; Bouadlown *et al.*, 1986; Vacher *et al.*, 1984). The expression of certain operons is also affected. There is a decrease in Leu operon expression and a relief of transcription termination at the Trp and Phe operon attenuators (Blum, 1988; Gourisharkar and Pittard, 1982; Yanofsky and Söll, 1977).

It has been proposed that modifications to tRNA, specifically at the position 3' of the anticodon, alter tRNA tertiary structure in a way that stabilizes the codon-anticodon pairing in order to reduce the chance of mispairing and the addition of incorrect amino acid (Jukes, 1973). Using purified $tRNA^{Phe}$ with and without the ms^2i^6A modification, Wilson and Roe (1989) have shown that in the absence of ms^2i^6A, the first position of the codon is allowed to "wobble", resulting in Phe incorporation at CUU codons (intended to code for Leu). This quite clearly demonstrates that the ms^2i^6A modified nucleoside is required to stabilize the pairing at the first position of the codon, and that in its absence "wobble" can occur. The role of ms^2i^6A in proper codon-anticodon pairing is further supported by the reduced efficiency of serine- and tyrosine-inserting UGA suppressor mutants in *miaA*$^-$ cells (Bouadlown *et al.*, 1986; Petrullo *et al.*, 1983). Misreading of the genetic code would, for obvious reasons, have disastrous consequences on the organism. But in *miaA*$^-$ cells, even though this A37 modification required for proper reading of the code is absent, a decrease in translation errors is observed (Bouadlown *et al.*, 1986). These seemingly contradictory results can be explained by the intrinsic proofreading of the ribosomal complex. Even though the absence of the A37 modification

destabilizes the codon-anticodon pairing, allowing the opportunity for the improper aminoacyl-tRNA to be loaded at the A-site of the ribosomal complex, when the ribosome attempts to shift the improperly paired aminoacyl-tRNA to the P-site during peptidyl transfer, the destabilized pairing fails and translation stalls. This dislodging of the improper aminoacyl-tRNA allows a second try at proper pairing and thus reduces the chance for the incorporation of erroneous amino acids. The stalling of the ribosomal complex due to an increase in the number of mispaired aminoacyl-tRNAs would explain the observed slow rate of translation. A global slow down of the rate of translation may explain the slow growth observed in *miaA* mutants.

miaA and A37 modification also seems to be involved in the expression of certain genes. In the absence of the modification, a decrease in Leu operon expression and an increase in Trp and Phe operon expression has been observed (Blum, 1988; Gourisharkar and Pittard, 1982; Zurwaski *et al.*, 1978; Yanofsky and Söll, 1977). There is an increased expression of the aromatic permeases (Buck and Griffiths, 1981) and it is also believed that there may be an increased expression of the tyrosine-sensitive first enzyme of common aromatic biosynthesis (McCray and Hermann, 1976). Interestingly, it has also been observed that the degree to which A37 is modified depends on the growth medium. Hydroxylation of ms^2i^6A to form ms^2io^6A occurs only when the cells are growing aerobically (Buck and Ames, 1984; Björk, 1980). The dependency of hydroxylation on oxygen does not exist for all tRNA modifications. The synthesis of the modified nucleosides cmo^5U and

mcmo^{5}U occurs during anaerobic as well as aerobic growth (Buck and Ames, 1984). Another example of the effect of growth conditions on A37 modification is the absence of methylthiolation (i^{6}A to ms^{2}i^{6}A) when the bacteria are starved for either iron or cystein (Buck and Ames, 1984). The involvement of *miaA* in gene expression and the dependency of the state of A37 modification on growth conditions suggests roles in addition to the fine tuning of the translational apparatus. It has been proposed that through the state of A37 modification, the translational apparatus is able to alter gene expression (possibly through an effect on attenuators) in an attempt to compensate for adverse growth conditions (Buck and Ames, 1984).

Isopentenyl Adenosine in the tRNA of *S. cerevisiae*

In comparison with the consequences seen in bacteria lacking isopentenylation at adenosine 37, in yeast the phenotype is very different. The gene encoding the isopentenyl transferase in yeast is *MOD5*. As in bacteria, when *MOD5* is mutant the adenosine at position 37 of tRNA specific for codons that begin with U is not isopentenylated, but there is no effect on amino acid acceptance nor on cell growth (Laten *et al.*, 1978).

In eukaryotes there are two populations of tRNA. The cytoplasmic tRNA is encoded by the nuclear genome and the mitochondrial tRNA is encoded by the mitochondrial genome. These two pools of tRNA do not mix. The mitochondrial genome encodes a limited number of proteins and many genes involved in

mitochondrial structure and function are therefore encoded by the nucleus. There are many instances, including tRNA modification, where analogous biochemical processes take place in the nucleus/cytoplasm and in the mitochondria (reviewed by Martin and Hopper, 1982). While in many cases the enzymes involved in these processes are encoded by distinct nuclear genes, there are examples where enzymes that act both in the nucleus/cytoplasm and in the mitochondria are encoded by single nuclear genes. For instance, *trm1* and *trm2* are nuclear genes involved in the post-transcriptional methylation of tRNA in the cytoplasm and the mitochondria (Hopper *et al.*, 1982; Philips and Kjellin-Straby, 1967). *MOD5* falls into this category of nuclear genes with both cytoplasmic and mitochondrial roles (Martin and Hopper, 1982).

Expression pattern studies have revealed that there are three different forms of Mod5p. The first (IPPT-I) is localized to the mitochondria, but is also found in small amounts in the cytoplasm (Najarian *et al.*, 1987; Gillman *et al.*, 1991). The second form (IPPT-II) was shown to be in the mitochondria (Najarian *et al.*, 1987; Gillman *et al.*, 1991). There are four mechanisms that could explain how *MOD5* encodes an enzyme that is targeted both to the mitochondria and cytoplasm. The mechanism could be based on differential initiation/termination of transcription, differential RNA processing, differential initiation of translation or differential protein processing. Analysis of the genomic *MOD5* sequence revealed that there was a second ATG ten codons downstream of the one which was predicted. Mutational analysis was used to eliminate each of the start codons in turn and it was

shown that IPPT-I was a product of initiation at the first start, and that IPPT-II was the product of initiation at the second (Gillman *et al.*, 1991). Using passenger proteins *in vivo* and *in vitro*, Boguta *et al.* (1994) were able to show that while the difference between IPPT-I and IPPT-II (the 11 aa N-terminal extension of IPPT-I) was essential for mitochondrial targeting, it was in fact the first 21 aa that were sufficient for targeting. The third form (IPPT-III) is the smallest and was found in all three compartments. Its size is consistent with initiation at the third ATG codon. The expression of IPPT-III alone is not sufficient to restore isopentenylation. How it is targeted to the mitochondria without the 21aa signal sequence and what its function may be are not known (Gillman *et al.*, 1991; Boguta *et al.*, 1994).

One surprising result was that IPPT-II, as well as being in the cytoplasm, is targeted to the nucleus (Boguta *et al.*, 1994). Based on a number of findings it is believed that isopentenylation of A37 in cytoplasmic tRNA occurs in the cytosol, after tRNA processing and export from the nucleus (Buguta *et al.*, 1994). The *rna1-1* mutation in yeast prevents intron excision which results in the nuclear accumulation of pre-tRNA. This accumulated pre-tRNA does not contain i^6A37, although other modified nucleosides are present (Hopper *et al.*, 1978; Knapp *et al.*, 1979). This however does not exclude the possibility that intron-containing pre-tRNA are not able to serve as a substrate for IPP transferase. It has also been observed that when yeast $tRNA^{Tyr}$ is injected into *Xenopus* oocytes, isopentenylated A37 is only found after intron excision in mature cytoplasmic tRNA (Nishikura and DeRobertis, 1981). The final supporting evidence is that when mature tRNA lacking the

isopentenylated adenosine nucleoside are injected into *Xenopus* oocytes, isopentenylation occurs in the cytosol (Boguta *et al.*, 1991). If tRNA isopentenylation does in fact occur in the cytosol, why then would the IPP transferase have to be targeted to the nucleus? Boguta *et al.* (1991) presented two possible explanations. It is possible that nuclear and cytosolic IPPT-II could modify different subsets of tRNA. For instance, some tRNA do not have introns. It is possible that nuclear IPPT-II modifies introns-less tRNA, while cytosolic IPPT-II might modify tRNA that do have introns. Alternatively, *MOD5* IPPT-II could also have functions in addition to tRNA modification that take place in the nucleus. There are examples where other tRNA modifying enzymes have been proposed to have additional roles. In bacteria missing the m^5U54 modification from the initiator $tRNA^{Met}$ as a result of environmental limitations, the $tRNA^{Met}$ is able to initiate without being formylated. These cells show no reduction in growth rate (Baumstark *et al.*, 1977; Samuel and Rabinowitz, 1974; Björk and Neidhart, 1975). However, when the gene *trmA* (responsible for synthesizing the m^5U54 modification in the initiator $tRNA^{Met}$) is mutated, the $tRNA^{Met}$ again does not require formylation to initiate, but growth rate is considerably reduced. In both cases the modification is absent. But in the case of the trmA mutant, growth rate is also reduced. This suggests that trmA has a function in addition to synthesizing the m^5U54 modification, and that this role somehow impinges on cellular growth (Persson, 1993). This may also be the case with *MOD5* in yeast.

These findings and the work done on the bacterial *miaA* propose a biochemical function for GRO-1 based on the similarity of the proteins. Different roles for this function in other systems have been proposed. Why the Clock phenotype is observed in *gro-1* mutants is not immediately clear but a number of possibilities are discussed here.

Materials and Methods

Publicly Available Data, Strains and Clones

C. elegans physical and genetic map data was obtained from the public repository at ncbi.nlm.nih.gov/repository.celegans/ace2 and viewed using the program AceDB, A *C. elegans* Database (Eeckman and Durbin, 1995). Database searches used the algorithm BLAST (Altschul *et al.*, 1990). The strains used include the wild type strains N2 (var. Bristol) and PaC1, *gro-1(e2400)* in the PaC1 strain (CB4512), and *gro-1(e2400)* out-crossed with N2 (MQ520) . Nematodes were cultured according to the methods developed by Sydney Brenner (1974). Cosmid clones were obtained from Dr. A. Coulson (Cambridge).

Microinjection and Molecular Techniques Required for Subcloning

Cosmid and plasmid DNA was prepared by the boiling miniprep technique (Holmes and Quigley, 1981). The concentrations were estimated after electrophoresis through a 0.4% agarose gel and visualized by ethidium bromide staining. The RNA was removed by precipitation with LiCl (Ausubel *et al.*, 1992). Cosmids were injected individually at the typical concentration of 50 μg/ml along with the coinjection marker (pRF4), also at a concentration of 50 μg/ml (Mello *et al.*, 1991). The methods for microinjection developed by Fire (1985) and Mello *et al.* (1991) were followed.

The constructs used to reduce the rescuing region were all based on the cosmid clone ZC395 (see Figure 5). The deletion construct pMQ2 was made by deleting the 29.9 kb SpeI fragment of ZC395, and pMQ3 by deleting the 31.4 kb NdeI fragment. The frame-shift construct pMQ4 was made by cutting pMQ2 at the unique ApaI site in the second exon of ZC395.7 and degrading the resulting 4 bp overhang with Mung Bean nuclease. pMQ5 was made by cutting pMQ2 at the unique NdeI site in the second exon of ZC395.6 and filling in the resulting 2 bp overhang with Klenow fragment of DNA polymerase.

The transgenic lines produced by microinjection were established by propagating rolling (Rol) lines. This was possible because of the use of the coinjection marker pRF4 which carries a dominant mutant allele of *rol-6*. Any worm expressing an extrachromosomal array in their epidermal cells with exhibit the Rol phenotype. The presence of pRF4 is used as a marker for animals that carry an extrachromosomal array and could be expressing the clone of interest. Lines were tested for rescue by comparing their developmental rates with that of N2 (wild-type) and *gro-1(e2400)*. Approximately 20 young adults (N2, *gro-1(e2400)* and the transgenic line separately) were allowed to lay eggs for one hour. The adults were then removed and the progeny allowed to develop. The time it took for the transgenic animals to reach adulthood was compared to each of the controls; if it was comparable to the wild-type then it was considered rescued for development

DNA Amplification by PCR

For PCRs using clean, abundant template, a single pair of primers was used. The notation A:B is used here, where A is the first primer and B the second. In cases where the template of interest is rare and non-specific amplification is likely, nested PCR was used. In this case, two separate PCR are run, the second with primers "nested" within those of the first, and the product of the first used as template. The notation A/C:D/B is used here, where A and B are the primers in the first reaction, and C and D are the primers in the second "nested" reaction.

Sequencing

DNA sequences were determined by the chain termination method (Craxton, 1993). The *fmol* DNA Sequencing System (Promega, http://www.promega.com) was used and the primers were end-labeled with γ-^{32}P. The products were separated by electrophoresis in a standard buffer gradient 6% polyacrylamide gel. After separation, the gels were exposed to x-ray film for visualization.

Search for the *e2400* mutation

Nematode genomic DNA was prepared using standard techniques from worms that had been prepared by sucrose flotation (Sulston and Hodgkin, 1988). The primers SHP93:SHP92 were used to amplify the *gro-1* region from N2, PaC1 and *gro-*

1(e2400) genomic DNA with the regimen (94°C for 20 sec, 55°C for 1 min, 72°C for 2 min) for 30 cycles. Taq Polymerase (Pharmacia, http://www.apbiotech.com) was used in the PCR and the bands were purified by agarose gel electrophoresis and Pharmacia BandPrep Kit. These products were sequenced in search of the *gro-1(e2400)* mutation. The primers used were: SHP93, 94, 95, 96, 97, 98, 99, 100, and 92. The sequence of these primers is given in Appendix F and their positions are illustrated in Appendix A.

Confirmation of predicted splicing patterns

The cDNA of each gene in the operon was sequenced to confirm the predicted splicing patterns. Mixed-stage nematode RNA was prepared using Trizol (GibcoBRL, http://www.lifetech.com) in a method based on that of Chomczynski and Sacchi (1987). This RNA was used to prepare reverse-transcribed libraries by the method of Frohman *et al.* (1988). The poly-T primer R_t was used for the amplification of the cDNA. Using this reverse-transcribed cDNA as template, the mRNAs of the five genes of the operon were amplified by PCR. Each of the genes was amplified in a number of pieces. The 5' end using a primer corresponding to the *trans*-spliced leader sequence SL2 (if the gene was *trans*-spliced, see "Operonicity" in the Results section) and two internal primers, and the 3' end using the primers R_i and R_o along with an internal primer. The primer R_t used to amplify the cDNA had two primer landing pads R_i and R_o built into its 5' end. This allowed priming for PCR at the end of each cDNA. *gop-1* was amplified in three parts, 5'

part was amplified with the nested primers SHP190:SHP174/SHP176, the middle part with SHP172/SHP173:SHP176, and the 3' end with SHP175:R_i/R_o. *gop-2* was amplified with the primers SL2:SHP143/SHP144 and SHP180:Ri/Ro. *gop-3* was amplified with the primers SL2:SHP184/SHP135 and SHP138:Ri/Ro. *hap-1* was amplified with the primers SL2:SHP130/SHP119 and SHP118:Ri/Ro. And finally *gro-1* was amplified with the primers SL2:SHP99/SHP100, SHP94:SHP99/SHP100, and SHP97:Ri/Ro. The sequences of these primers is given in Appendix F and their positions illustrated in Appendix A. The regimen used for each was: (94°C for 20 sec, 60°C for 1 min, 72°C for 2 min, for 20 cycles) for the first PCR, and (94°C for 20 sec, 60°C for 1 min, 72°C for 2 min, for 30 cycles) for the second, "nested" PCR. The products were purified by gel electrophoresis and sequenced.

Table 1 lists the primers and template amplified from cDNA used for confirming the predicted genomic organization of the *gro-1* operon. The primers were designed so that the splice sites of interest in the cDNA would easily fall within their sequencing reads. By looking at the positions of predicted splice sites in the sequence of the cDNA, unpredicted introns can be detected or the presence of predicted introns that do not really exist can be determined. Errors in the predictions will cause the position of correct splice sites to shift as a result of the added or subtracted nucleotides in the cDNA. The sequences of these primers is given in Appendix F and their positions illustrated in Appendix A.

Table 1: Confirming the predicted splicing pattern of the *gro-1* operon

GENE	PRIMER	STRUCTURE	TEMPLATE
gop-1	SHP141	splices 1 and 2	SHP190 : SHP174 / SHP176
	SHP172	splices 3 and 4	*ibid.*
	SHP173	splices 4, 5 and 6	*ibid.*
	SHP174	splices 5, 6 and 7	SHP172 / SHP173 : SHP176
	SHP185	splices 8, 9 and 10	*ibid.*
	SHP175	splices 10, 11 and 12	*ibid.*
	SHP176	splices 11, 12 and 13	SHP175 : R_i / R_o
	SHP177	splices 14 and 15	*ibid.*
	SHP178	poly-A	*ibid.*
gop-2	SHP179	5' end and splice 1	SL2 : SHP143 / SHP144
	SHP181	splices 3 and 4	SHP180 : R_i / R_o
	SHP144	splice 2	*ibid.*
	SHP182	poly-A	*ibid.*
gop-3	SHP183	5' end and splice 1	SL2 : SHP184 / SHP135
	SHP146	splice 2	*ibid.*
	SHP138	splices 3, 4 and 5	*ibid.*
	SHP184	splices 4 and 5	SHP138 : R_i / R_o
	SHP140	splices 6 and 7	*ibid.*
	SHP163	splices 7 and 8	*ibid.*
	SHP134	splice 8 and poly-A	*ibid.*
	SHP164	poly-A	*ibid.*
hap-1	SHP165	5' end and splice 1	SL2 : SHP130 / SHP119
	SHP118	splices 1 and 2	*ibid.*
	SHP119	splices 3 and 4	SHP118 : R_i / R_o
	SHP120	poly-A	*ibid.*
gro-1	SHP95	5' end and splice 1	SL2 : SHP95 / SHP99
	SHP159	splice 2	*ibid.*
	SHP96	splices 3 and 4	SHP94 : SHP99 / SHP100
	SHP99	splices 4 and 5	*ibid.*
	SHP98	splices 6, 7 and 8	SHP97 / SHP98: R_i / R_o
	SHP100	splices 6 and 7	*ibid.*
	SHP110	poly-A	*ibid.*

Determining membership in the *gro-1* operon

To determine which of the genes in the region were members of the *gro-1* operon, the 5' ends of each was amplified (see Figure 6 and "Operonicity" in the Results section). Two PCR were carried out for each gene, one using a primer corresponding to the first spliced leader sequence (SL1) and another using a primer corresponding to the second spliced leader sequence (SL2). These primers were used in conjunction with a pair of internal, gene-specific primers. The use of two internal primers allowed for nested PCR in order to increase the specificity of the reaction. For *gop-1* the internal primer pair used was SHP141/SHP142 and the expected product size (based on GeneFinder predictions) was ~570 bp. The primer pair for *gop-2* was SHP143/SHP144 to produce a product of ~510 bp. The primer pair SHP145/SHP146 was used to amplify the 440 bp *gop-3* 5' end. For *hap-1*, the primer pair SHP130/SHP119 was used in the amplification of the 465 bp 5' end. Finally, to amplify *gro-1*'s 390 bp 5' end, the primer pair SHP95/SHP99 was used. The sequence of these primers are given in Appendix F and their positions illustrated in Appendix A. The regimen used for each reaction was: (94°C for 20 sec, 60°C for 1 min, 72°C for 2 min, for 20 cycles) for the first PCR, and (94°C for 20 sec, 60°C for 1 min, 72°C for 2 min, for 30 cycles) for the second, "nested" PCR. The products were separated by gel electrophoresis and visualized by ethidium bromide staining and exposure to UV.

Identification and Sequencing of the EST corresponding to the human homologue of *gro-1*

A human Expressed Sequence Tag (EST) was identified using the BLAST algorithm (Altschul *et al.*, 1990) to compare the *gro-1* sequence with EST sequences in the public database. An entry (accession F07677) with a 5' read of the clone c-2ec05, very similar to *gro-1* was discovered. Upon investigation, a 3' read of the same clone was identified (accession Z40724), indicating that the clone contained the poly-A tail. Because the 5' read was very near to the 5' end of the *gro-1* sequence, of all the EST clones, c-2ec05 contained the most sequence. The clone was ordered from Genome Systems (http://genomesystems.com). The clone's sequence was determined by assembling the 5' and 3' reads of other ESTs found in the database, and sequencing the missing regions using primers based on the known sequence. The primers used were: dm2, SHP193, 212, 213, 226, 227, 239, 256, 263, 264, 265, and 266. The sequences of these primers is given in Appendix F.

Simplifying the Operon: pMQ8

Because *gro-1* is the fifth member of an operon (see "Operonicity" in the Results section) its promoter is more than 10 kb upstream (see Appendix A, pos. -9643 to -9351). In order to more easily assemble the various constructs, *gro-1* was placed immediately adjacent to its promoter, essentially deleting the first four genes of the

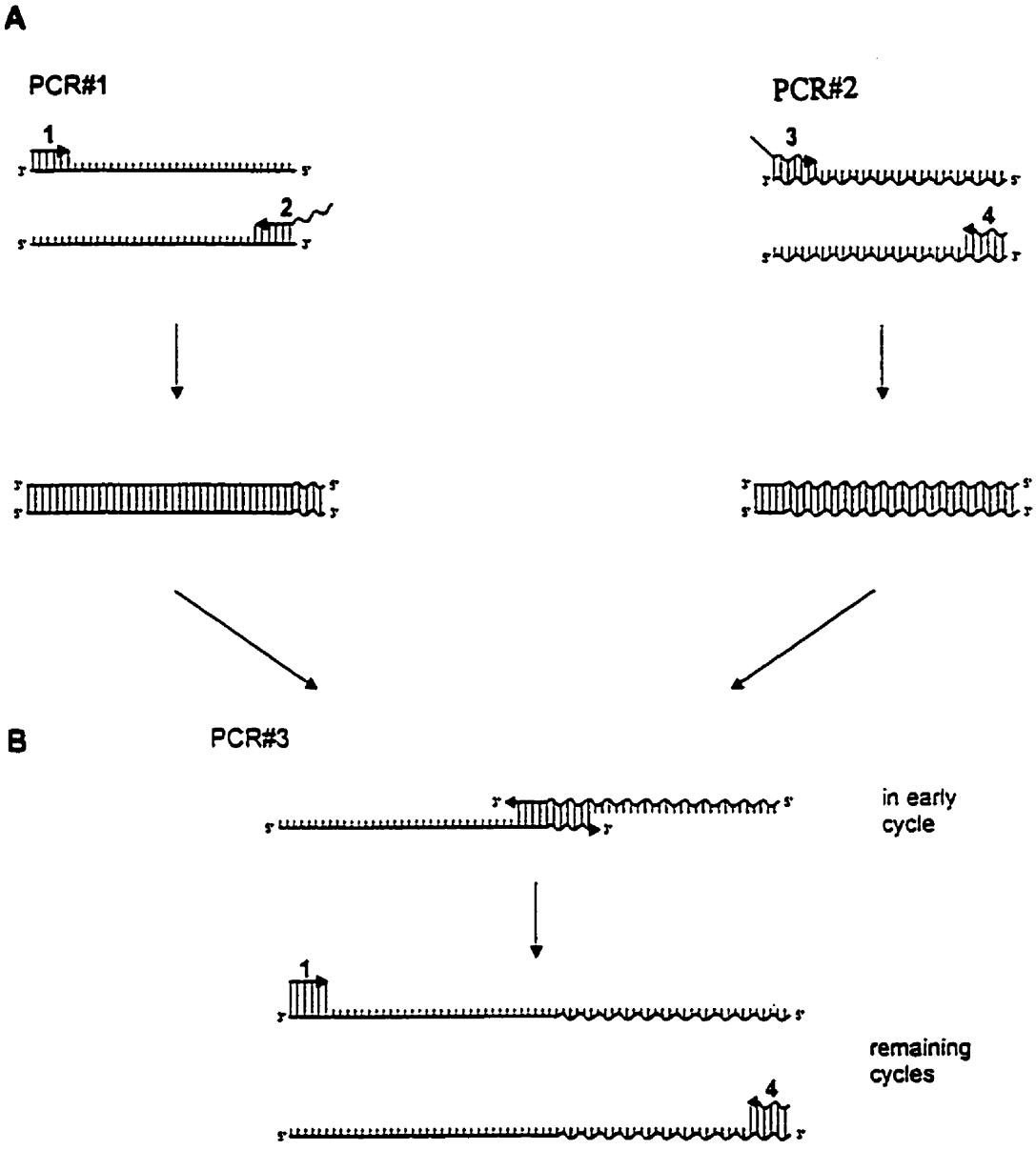


Figure 1: Bi-partite PCR. This technique involves three independent PCRs used to join two separate DNA pieces. A: PCR#1 and PCR#2 use primer pair 1:2 and 3:4 respectively to amplify the two pieces of DNA. The hybrid nature of primers 2 and 3 results in there being complementary portions in the products from each reaction. B.: After gel purification the products from PCR#1 and PCR#2 are used as templates for PCR#3. The complementary regions of the two templates anneal and present free 3' ends for extension, in early cycles, by the polymerase. This extension may be a rare event but when it occurs it produces the final, fused product which allows the flanking primers 1 and 2 to anneal and amplify the desired product during the remaining cycles.

operon. This may also eliminate any problems associated with the overexpression of the first four genes in transgenic animals. The absence of unique restriction endonuclease site that could be used to delete the region between *gro-1* and its promoter was not surprising considering the size of the region. Bi-partite PCR was therefore used to generate the desired product. The general concept of this technique is illustrated in Figure 1. The hybrid primers used were SHP159 and SHP160. The flanking primers used were SHP161 and SHP162, which had the restriction enzyme sites SacI and PstI, respective, built into their 5' ends in order to facilitate cloning. The sequence of these primers is given in Appendix F and their positions illustrated in Appendix C. PCR#1 used the primers SHP161 and SHP160, and N2 genomic DNA as template to amplify the promoter. PCR#2 used the primers SHP159 and SHP162, and N2 genomic DNA to amplify *gro-1*. PCR#3 used the flanking primers SHP161 and SHP162, and the gel purified products from PCR#1 and PCR#2 as template to fuse the two products. The high fidelity polymerase VENT (New England Biolabs, http://www.neb.com) was used because of its 3'-5' exonuclease activity. This feature reduces the chance of PCR error, but more importantly eliminates the additional nucleotides (usually A) added to the 3' end of Taq Polymerase products. The presence of these additional nucleotides would prevent the extension of the template's 3'end in the early rounds of PCR#3 (see Figure 1C). The product from PCR#3 was then cloned into the SacI/PstI site of vector pUC18 (Messing, 1983; Norrander *et al.*, 1983, and Yanisch-Perron *et al.*, 1983). The sequence of this construct is presented in Appendix C.

The Green Fluorescent Protein (GFP) reporter construct pMQ18

The *gro-1*::gfp reporter construct pMQ18 used the Fire GFP vector pPD95.77 (Fire, http://ciw2.ciwemb.edu/pub/FireLabVectors). The structure of the construct and the fusion protein produced are illustrated in Figure 2. Using pMQ8 (see above) as template, the primers SHP151 and SHP170 were used to amplify *gro-1* and its promoter with the regimen: (94°C for 20 sec, 55°C for 1 min, 72°C for 2 min 30 sec, for 30 cycles. The high fidelity polymerase VENT (New England Biolabs, http://www.neb.com) was used to reduce the chance of PCR error. The restriction enzyme sites SphI and XbaI, respectively, were built into the 5' ends of these primers. This allowed the PCR product to be cloned into the SphI/XbaI site of the vectors MCS, upstream of GFP. The primer SHP170 was designed to ensure that no frame-shift occurred between *gro-1* and GFP. The sequence of these primers is given in Appendix F and their positions illustrated in Appendix C.

This plasmid pMQ18 was microinjected, along with the marker plasmid pRF4, and transgenic lines established as described above. The GRO-1::GFP fusion protein was visualized in live transgenic worms using fluorescent microscopy.

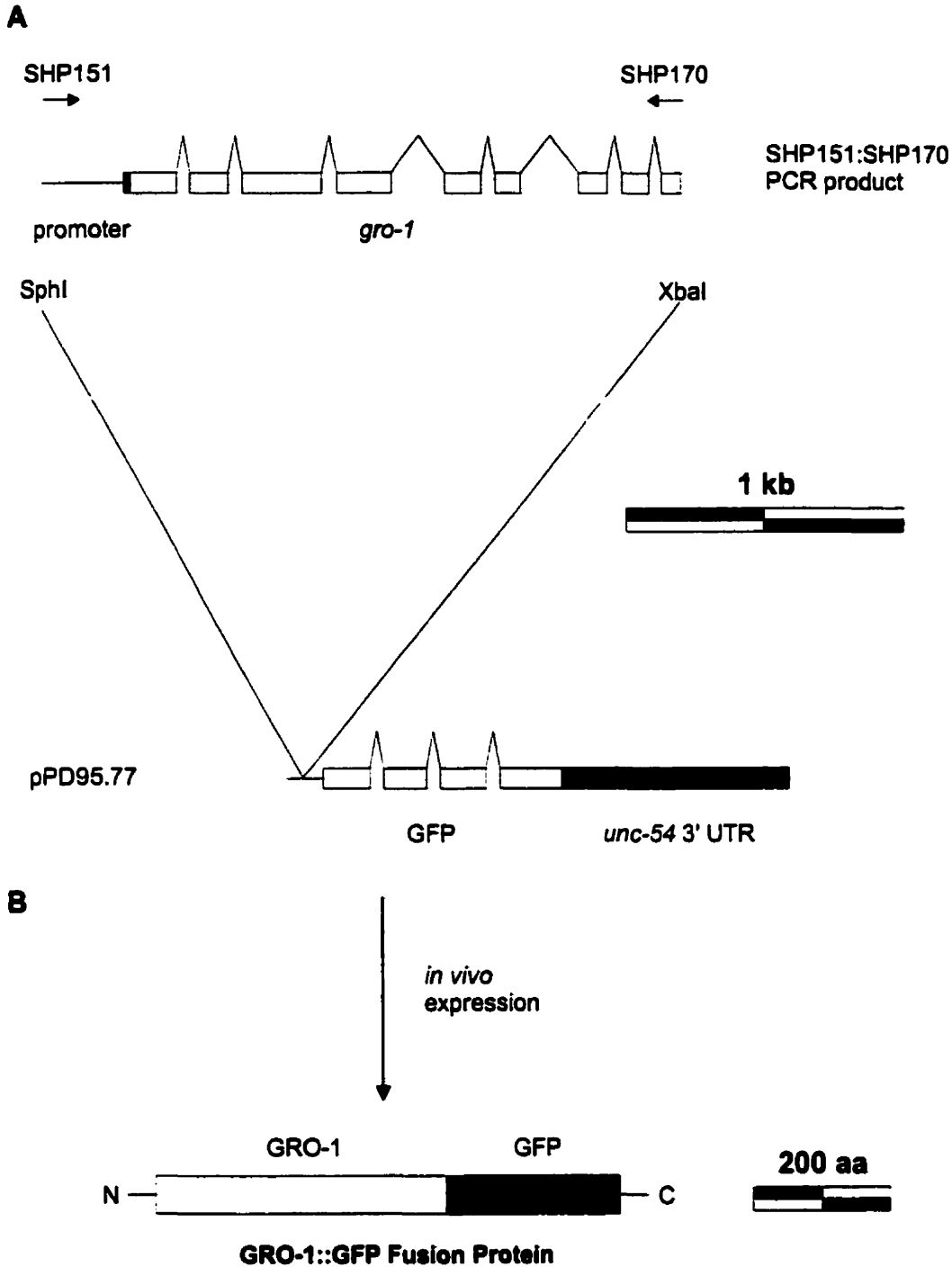


Figure 2: The *gro-1::gfp* reporter construct pMQ18. A: The structure of pMQ18. *Gro-1* and its promoter are amplified by PCR and cloned into the SphI/XbaI site of pPD95.77. B: This allows the expressionof a GRO-1::GFP fusion protein in transgenic worms.

Results

Cosmid Rescue

Based on the genetic mapping, *gro-1* was estimated to lie approximately 20 kb to the left of *clk-1* (see Figure 3, Lakowski, 1998). With *clk-1* already cloned (Ewbank *et al.*, 1997) and the estimated physical position based on the genetic data, it was possible to select a number of cosmid clones as candidates for injection rescue (see Figure 4). Of the eight cosmids selected, only B0498, C34E10 and ZC395 were able to rescue the mutant phenotype when microinjected (see Figure 5). Based on these results, attention was focused on the region common to the three overlapping clones. The 7 kb region was predicted by GeneFinder to encode four genes (see Figure 5).

Subcloning

Deletion constructs of the cosmid ZC395 were generated in an attempt to reduce the rescuing region and try to narrow down the candidate genes (see Figure 5). The first construct, pMQ2, deleted a large SpeI fragment of ZC395, leaving the 3.9 kb left end of the cosmid. This fragment contained the two predicted genes known to have homologues, ZC395.7 and ZC395.6. pMQ2 was able to rescue the mutant phenotype when microinjected. A second construct deleted a larger NdeI fragment of ZC395, leaving only ZC395.7. This construct, pMQ3, was unable to rescue *gro-1(e2400)*.

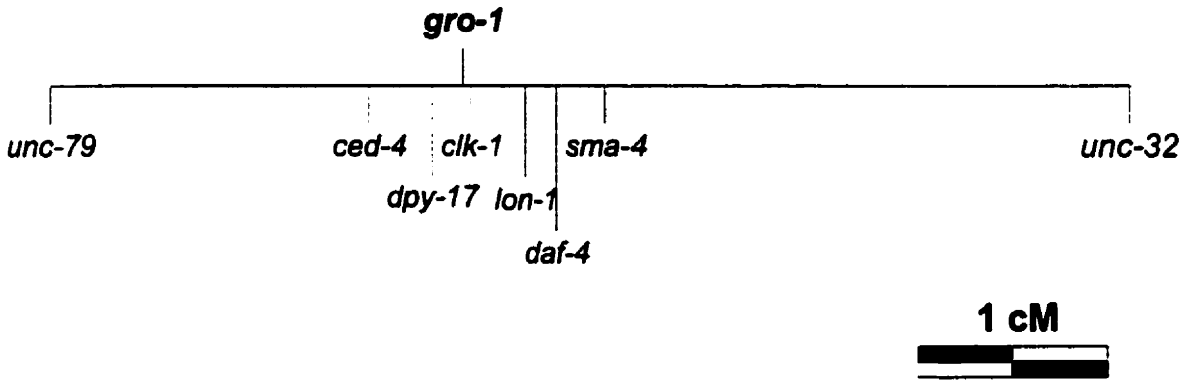


Figure 3: The genetic map of Linkage Group (LG) III in the region of *gro-1*. This data was generated by Lakowski (1998)

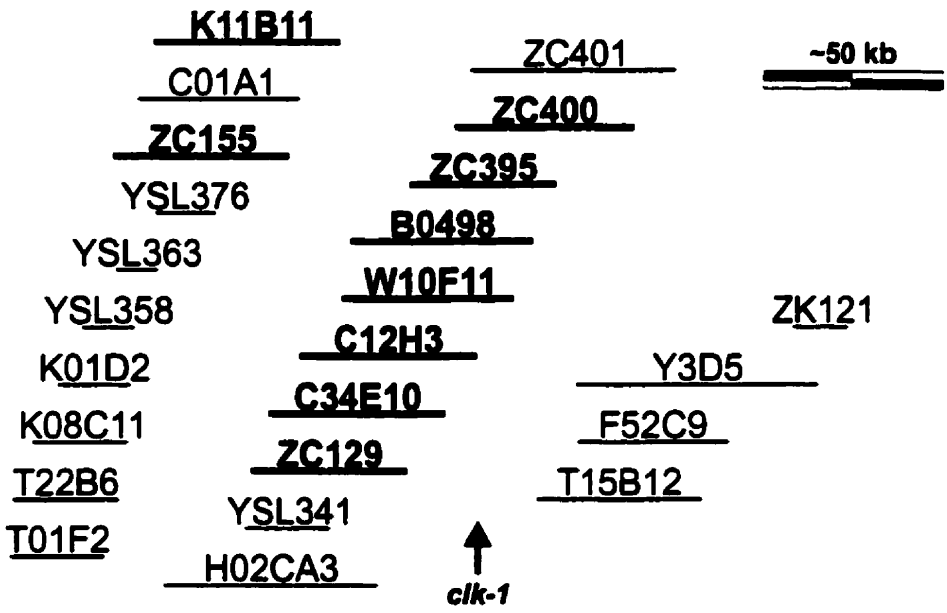


Figure 4: The physical map of the *gro-1* region. Illustrated are the cosmid clones of the region. While each clone is 40 kb in size, they are drawn to the number of bands that correspond to their fingerprint during map construction. The clones selected for microinjection are in bold.

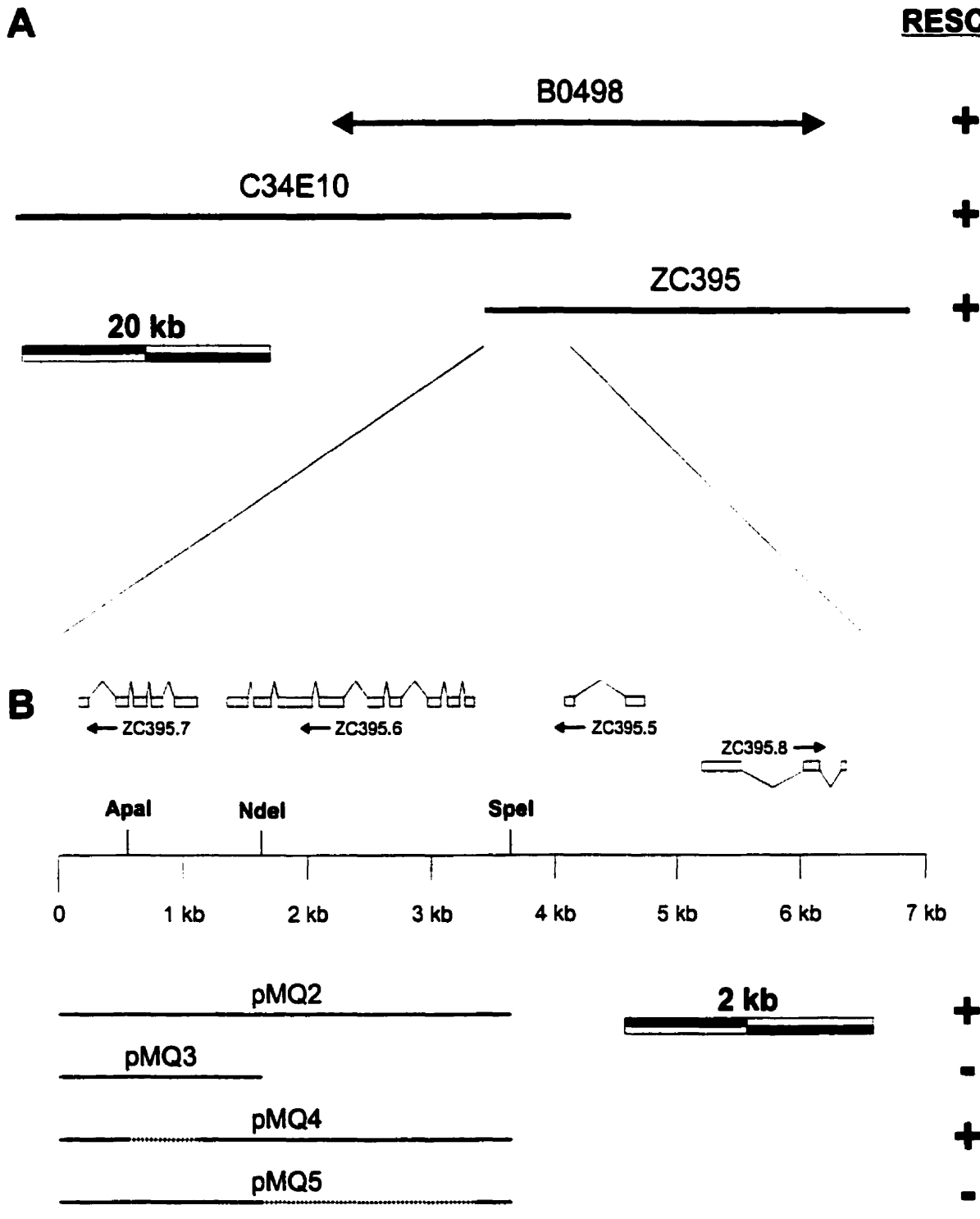


Figure 5: Clones used to identify the *gro-1* sequence. (+) to the right of the clone indicates its ability to rescue *gro-1(e2400)* when microinjected. (-) indicates its inability. A: The cosmid clones able to rescue *gro-1(e2400)* when microinjected. B: The plasmids constructed to subclone the 7 kb common rescuing region defined by cosmid clone rescue. Illustrated are the four predicted genes in this region. The endonuclease sites used in the subcloning are indicated. The (-----) regions illustrate the region that would be out-of-frame in the frame-shift constructs. The scale zeros at the left end of the cosmid clone ZC395.

Based on the negative result obtained by the microinjection of pMQ3, it was believed that ZC395.6 was in fact *gro-1*. In order to obtain a positive, and therefore more reliable, result, frameshift constructs were generated (see Figure 5). For the first of these, a unique ApaI site in the second exon of ZC395.7 was used to induce a frameshift by degrading the site's 4 bp overhang. The resulting construct, pMQ4, therefore encoded a non-functional copy of the gene. Another construct pMQ5, functionally knocked out ZC395.6 in the same fashion. When microinjected only pMQ4 was able to rescue the *gro-1* mutant phenotype. This positive result supported the results obtained with the deletion constructs and strongly suggested that ZC395.6 was in fact *gro-1*.

The constructs considered able to rescue caused the transgenic animals to develop at a wild type rate. These constructs also conferred a maternal effect: non-Rol worms also developed at an N2 rate.

Identification of the *e2400* Lesion

In order to confirm that *gro-1* is ZC395.6 the genomic copy of ZC395.6 was sequenced from mutant worms in an attempt to discover the *e2400* mutation. Sequencing with the primer SHP 99 (and SHP97) revealed that there was a 9 bp deletion and 2 bp insertion in the forth exon of ZC395.6 (actually the fifth exon of *gro-1*; see Figure 7). N2 and PAC1 genomic DNA were sequenced as wild-type controls. The sequences were consistent with those generated by the Genome Sequencing Consortium (Eeckman and Durbin, 1995). This mutation causes a

frameshift and a 33 aa out-of-frame extension in the mutant protein (see Figure 7C). With the agreement of a number of independent results, it was concluded that the gene ZC395.6 predicted by GeneFinder is in fact *gro-1*.

"Operonicity"

It has been shown that in *C. elegans* groups of genes can be organized into operons (Spieth *at al.*, 1993; Zorio *et al.*, 1994). For genes to be organized into such a transcriptional unit they must obviously share the same 5'-3' orientation and be closely positioned. The maximum intercistronic distance observed for two genes in an operon is 409 bp (Zorio *et al.*, 1989). Approximately 70% of *C. elegans* genes are *trans*-spliced at their 5' ends to a 22 bp sequence known as the first spliced leader (SL1) sequence. The downstream genes of operons are unique in their post-transcriptional modification; they are *trans*-spliced to a second spliced leader (SL2) sequence (Huang and Hirsh, 1989). Therefore, the presence of SL2 in an mRNA can be used as a marker indicating a gene's membership in an operon. There are four genes upstream of *gro-1* that are in close enough proximity to be in an operon. To examine this possibility, SL1- or SL2-specific primers were used in conjunction with a pair of gene-specific primers to amplify the 5' ends of these genes from Reverse Transcriptase-derived cDNAs (see Figure 6 and the Materials and Methods section). *gro-1* and the three genes upstream were all shown to be *trans*-spliced to SL2. Therefore, *gro-1* is the fifth member of a five-gene operon (see Appendix A for the full sequence and organization). The first three genes have no known homologues and have been named *gop* (gro operon gene) *1-3*, and the fourth gene is

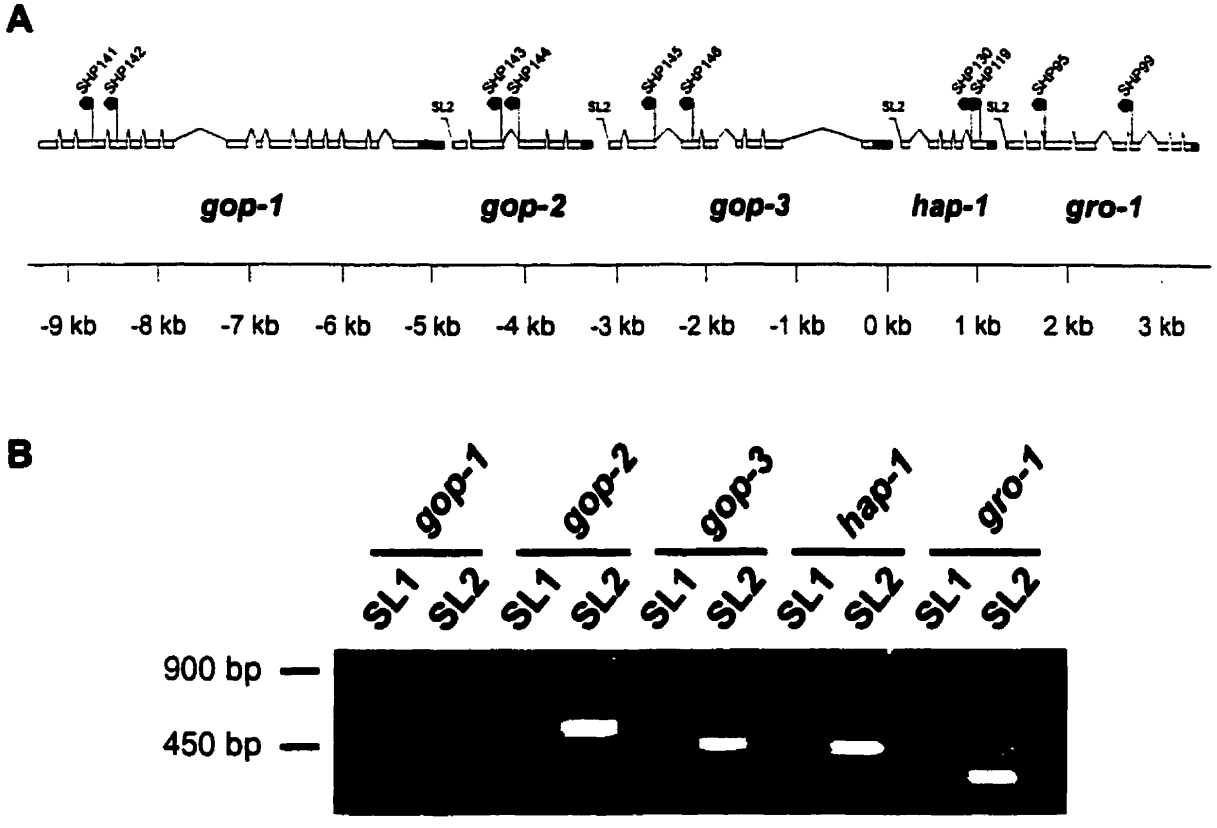


Figure 6: The *gro-1* operon. A: The five genes of the *gro-1* operon. Open sequence represents coding DNA, closed sequence represents non-coding DNA. The reverse primers used to determine the membership of each of the genes are illustrated. B: The PCR products generated using primers corresponding to SL1 or SL2 paired with two internal reverse primers for each of the genes in the *gro-1* operon.

homologous to the yeast gene HAM1 and has therefore been named *hap* (HAM1-like protein)-*1*.

Confirmation of Predicted Splicing Patterns

The splicing patterns predicted by GeneFinder for the five genes in the *gro-1* operon were confirmed by sequencing of their cDNAs. All the predictions were correct with three exceptions (see Appendix A for the full sequence of the operon and the splicing patterns). What was thought to be the last exon of *gop-3* actually splices over C34E10.9 (a small predicted gene on the other strand) to pick up two additional exons. This was originally detected when the *gop-3* 3' end was amplified and produced a product slightly larger than what was expected based on the prediction. The actual splicing pattern was then predicted manually and confirmed by sequencing. There was an additional intron in the predicted second exon of *gro-1*. This was detected when the positions of the first predicted splice and the 5' end occurred earlier in the sequence generated by SHP95. This suggested that there was fewer nucleotides between the primer SHP95 and the first predicted splice site. The presence of an additional exon, between the first and second predicted exons, was confirmed by sequencing with the primer SHP159. Finally, the true *gro-1* initiating AUG was actually 42 bp upstream of the predicted start. This was determined when sequencing the 5' end of *gro-1* with SHP95.

A

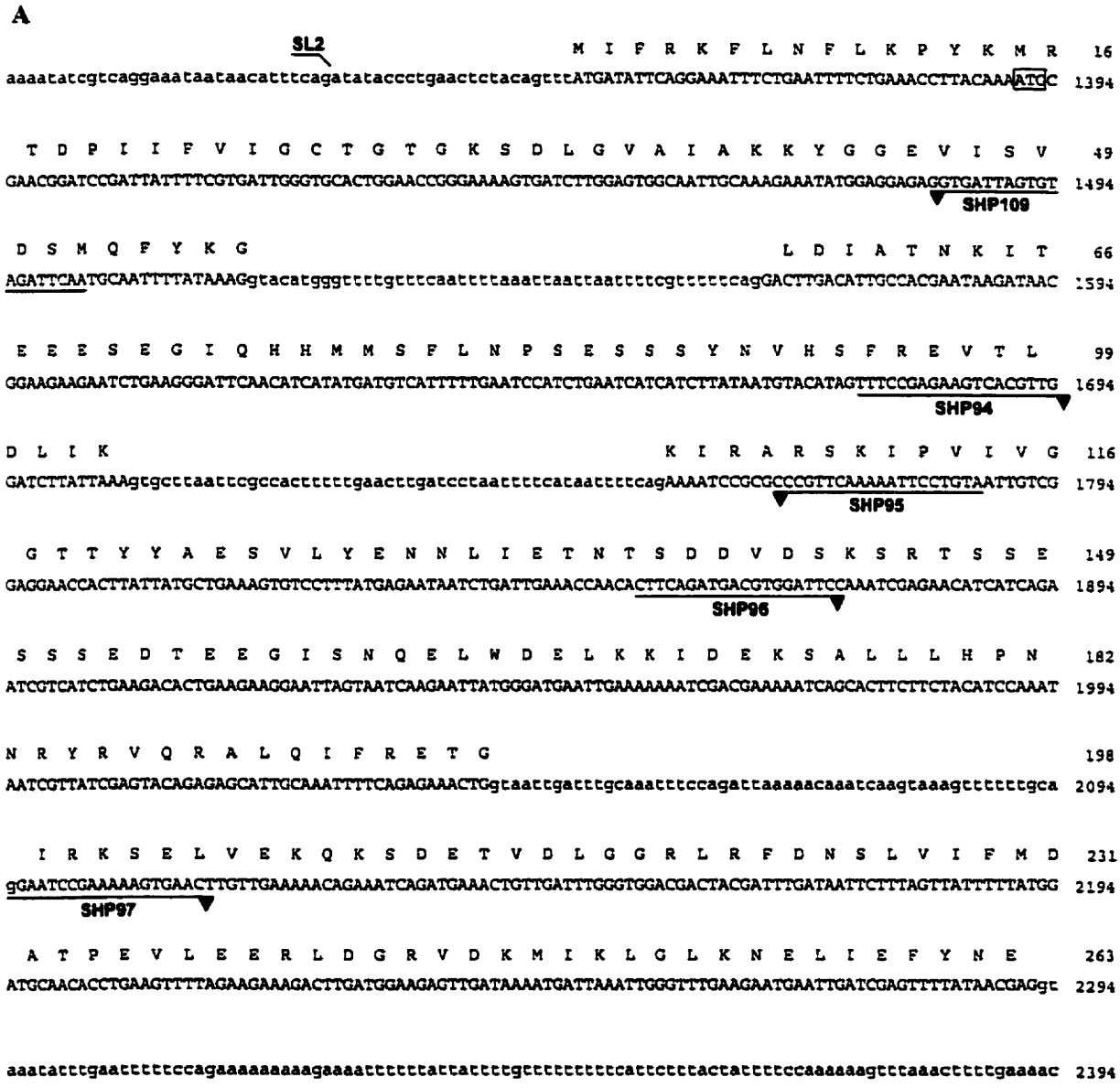
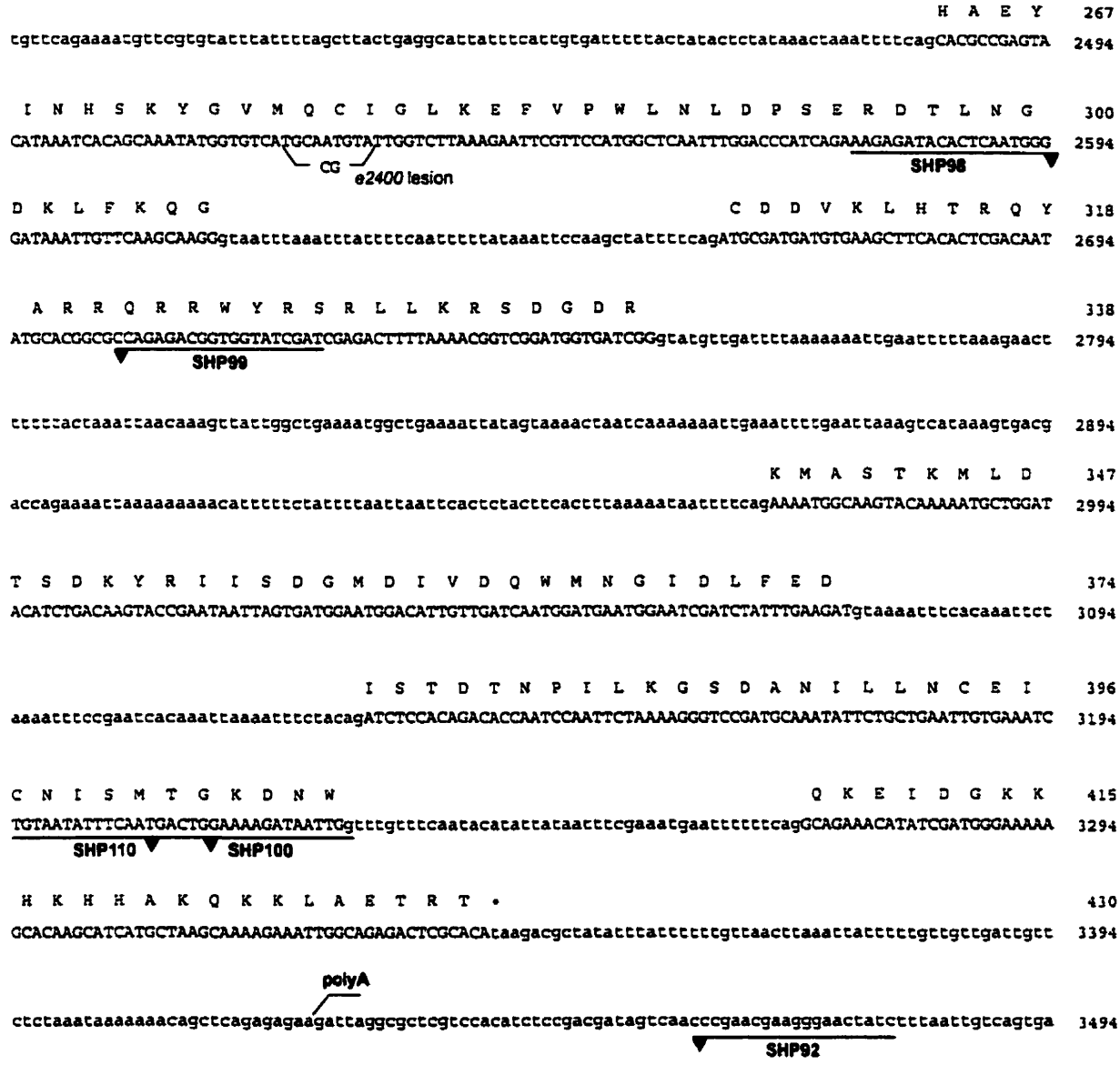
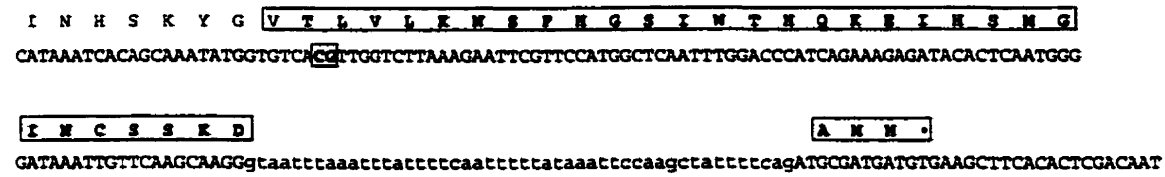


Figure 7: The *gro-1* sequence. A: The entire *gro-1* genomic sequence is shown. The coding sequence is in capitals. The protein sequence is given above the DNA sequence in single letter code. The *trans*-splice sites, the sites of poly-adenylation and the *e2400* mutation are shown. The proposed alternative initiator codon is boxed. B: The result of the *e2400* mutation. The 2 bp insertion and 33 aa out-of-frame extesion are boxed.

Figure 7 continued...



B



The human *gro-1* homologue

Using the BLAST algorithm (Altschul *et al.*, 1990) the clone c-2ec05, containing the Expressed Sequence Tag (EST) corresponding to the human *gro-1* homologue was identified. The sequence of the clone was determined and is presented in Appendix B. The protein sequence, based on the sequence of the cDNA, is very homologous to *gro-1* and the other known homologues (see Figure 8 for protein alignment).

The expression pattern of the GRO-1::GFP reporter

Two independent transgenic lines expressing the *gro-1::gfp* reporter construct pMQ18 were established. In general, the level of expression was low, making the signal weak. The first of the two lines (MQ698) expressed the reporter gene mainly in the gut, with some staining in a few other cell types in a small subset of the animals. The second line (MQ697) expressed the fusion protein in muscle cells, the pharynx, the celomacytes, the somatic gonad, the excretory canal and in the gut. In general, the fusion protein filled that cytoplasm and in most cases the nucleus. These preliminary results suggest that GRO-1 is localized to the cytoplasm and mitochondria, as well as to the nucleus.

A:

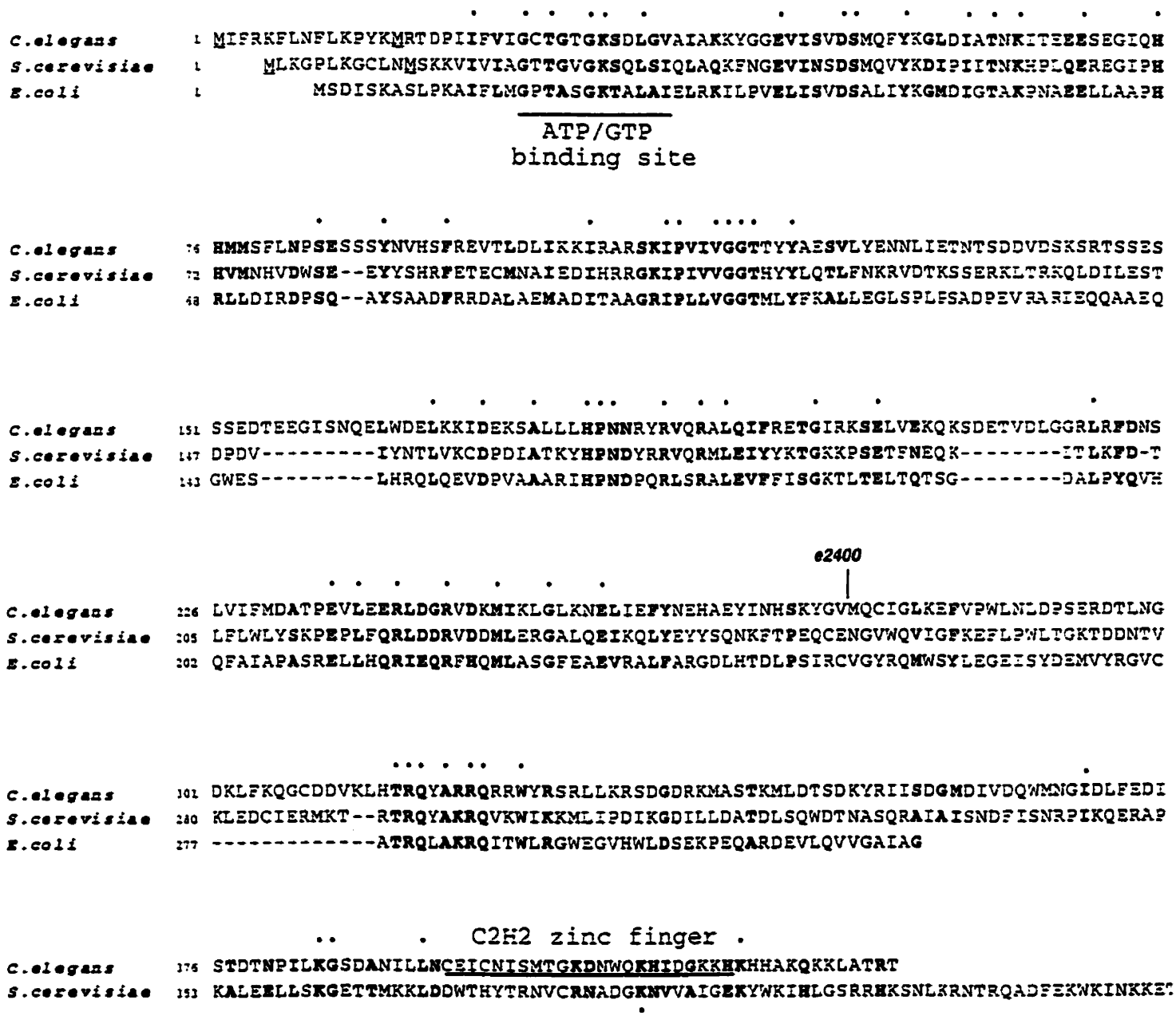


Figure 8: Alignment of GRO-1 and its homologues. A: Includes *C. elegans*, *S. cerevisiae* and *E. coli*. B: Includes *C. elegans* and human. Conserved residues are marked with a dot above, similar residues are in bold.

B:

```
                                  •• •      • •••••• •          •••  • •••• • •••• ••
hgro-1p   MASVAAARAVPVGSGLRGLQRTLPLVVILGATGTGKSTLALQLGQRLGGEIVSADSMQVYEGLDIITN
GRO-1             MIFRKFLNFLKPYKMRTDPIIFVIGCTGTGKSDLGVAIAKKYGGEVISVDSMQFYKGLDIATN

          •    •     ••• ••  •     •  •  ••    ••  • •• •••  •••• •• •• •
hgro-1p   KVSAQEQRICRHHMISFVDPL-VTNYTVVDFRNRATALIEDIFARDKIPIVVGGTNYYIESLLWKVLVN
GRO-1     KITEEESEGIQHHMMSFLNPSESSSYNVHSFREVTLDLIKKIRARSKIPVIVGGTTYYAESVLYENNLI

                         •   ••          •   •   •   •  •••    • • •• • ••••
hgro-1p   TKPQEMGTEKVIDRKVELEKEDGLV------LHKRLSQVDPEMAAKLHPHDKRKVARSLQVFEETGISH
GRO-1     ETNTSDDVDSKSRTSSESSSEDTEEGISNQELWDELKKIDEKSALLLHPNNRYRVQRALQIFRETGIRK

          ••    •    •     ••• • •          •    •• •••• ••• •    ••   ••      •
hgro-1p   SEFLHRQHTEEGGGPLGGPLKFSNPCILWLHADQAVLDERLDKRVDDMLAAGLLEELRDFHRRYNQKNV
GRO-1     SELVEKQKSDETVD-LGGRLRFDNSLVIFMDATPEVLEERLDGRVDKMIKLGLKNELIEF---YNEHAE

              •   •  • •• •••    •                   • •
hgro-1p   SENSQDYQHGIFQSIGFKEFHEYLITEGKCTLETSNQLLKKGPGPIVPPVYGLE--------------
GRO-1     YINHSKY--GVMQCIGLKEFVPWLNLDPSERDTLNGDKLFKQGCDDVKLHTRQYARRQRRWYRSRLLK

           ••     •          •                    •
hgro-1p   VSDVSKWEESVLEPALEIVQSFIQGHKPTATPIKMPYNEAENKRSYHL------------
GRO-1     RSDGDRKMASTKMLDTSDKYRIISDGMDIVDQWMNGIDLFEDISTDTNPILKGSDANILLN

          •  •    •   •  ••  ••   •
hgro-1p   CDLCDRIIIGDREWAAHIKSKSHLNQLKKRRRLDSDAVNTIESQSVSPDYNKEPKGKGSPGQNDQELKCSV
GRO-1     CEICNISMTGKDNWQKHIDGKKHKHHAKQKKLAETRT
```

C2H2 zinc finger

Discussion

The Clock gene *gro-1* has been cloned

The Clock genes have been implicated in a mechanism that controls and co-ordinates the timing of many metabolic, physiological and behavioural events in the development and life of the nematode *Caenorhabditis elegans*. Mutations in four Clock genes have been identified and result in a slowing of embryonic and post-embryonic development, the cell cycle and the periods of behavioural processes such as swimming and defecation (Hekimi *et al.*, 1995; Wong *et al.*, 1995). It is this effect on the temporal element of such a wide variety of seemingly unrelated biological events that has led to the proposal of a general "clock" that is responsible for coordinating the temporal aspects of the worm. Of the four Clock genes, only *clk-1* has been cloned (Ewbank *et al.*, 1997).

From the genetic mapping of *gro-1* (Lakowski, 1998) a number of cosmid clones to the left of *clk-1* were microinjected in an attempt to rescue the mutant phenotype. Three of these clones (B0498, ZC395 and ZC400) were able to rescue and the 7 kb common region was subcloned. Of the four genes predicted in this region, only ZC395.6 was able to rescue. After identification of the *e2400* lesion it was concluded that *gro-1* was indeed ZC395.6.

GRO-1 and its homologues

The *C. elegans* gene *gro-1* encodes a protein of 430 amino acids, homologous to the enzyme N^6-(Δ^2) isopentenyl PPi :tRNA isopentenyl transferase. The gene encoding this enzyme in bacteria (*MiaA*) and yeast (*MOD5*) are known. Using the *gro-1* sequence, a human Expressed Sequence Tag (EST) corresponding to the human homologue of *gro-1* was also identified. The protein in *H. sapiens*, *C. elegans* and *S. cerevisiae* are longer than that of *E. coli*. It is the N-terminal two thirds of the protein, common throughout evolution, where there is the most conservation at the amino acid level. There are about half a dozen blocks of amino acids that are highly conserved from bacteria to human. Based on the high degree of conservation it is likely that GRO-1 and the human GRO-1 are in fact the IPP-transferase in the worm and human respectively. This enzyme is responsible for the modification of tRNA that code for codons starting with U. The enzyme catalyzes the isopentenylation of the adenine residue at position 37, adjacent to the anti-codon. (see Figure 8 for the alignments of these proteins).

There are a number of recognizable motifs that are conserved in the IPP-transferase. Near the N-terminus there is an ATP/GTP binding site. This sequence is conserved in all four species (see Figure 8). Near the C-terminal ends of the human, worm and yeast protein is a C_2H_2 zinc finger motif. This motif is in the sequence that extends past the end of the bacterial protein. There are a number of other blocks of conserved sequence that do not correspond to any known motifs. One in particular may be of interest simply because it lies downstream of the *e2400* lesion. This

block and the zinc finger would be the two regions of conservation absent if the *gro-1* mutant protein were expressed.

The *e2400* allele

gro-1(e2400) was isolated as a spontaneous segregant from the wild type isolate PaC1 (Hodgkin and Doniach, 1997). It is the only mutant allele of *gro-1*. With the cloning of *gro-1*, the *e2400* lesion was sequenced and shown to be a 9 bp deletion and 2 bp insertion in the fifth intron (see Figure 7). The PaC1 strain is believed to contain an active transposon, the cause of the *e2400* mutation. While this lesion does not correspond exactly to the excision footprint of previously investigated transposons (Ruan and Emmons, 1987; Eide and Anderson, 1988; van Luenen and Plasterk, 1994), it is of a similar nature and could well have been caused by a transposon.

The mutation causes a frame shift that creates a 33 aa out-of-frame extension beginning at the 275th amino acid (see Figure 7). Because of the nature of the mutation, it is not clear at this point if *e2400* is a null mutation. If it were a massive deletion or if a stop codon was introduced in the first intron, it would be safe to conclude that no functional protein is expressed. But because the majority of the conserved sequence is present in the *e2400* protein, it is possible that if the out-of-frame extension does not cause the protein to be unstable or interfere with its

function, the protein could at least be partially functional. Further experiments are required.

gro-1 Expression

In yeast, Mod5p is found in the cytoplasm, nucleus and mitochondria. This expression pattern is brought about because different forms of the protein are generated by alternative translation. Initiation of translation at the first AUG codon produces a protein targeted to the mitochondria by an N-terminal targeting signal. The cytoplasmic and nuclear form are generated when translation is initiated at the second AUG codon, 14 codons downstream of the first. This shorter form lacks the N-terminal mitochondrial targeting sequence. By this mechanism, both the cytoplasmic and mitochondrial populations of tRNA are isopentenylated by the *MOD5*-encoded enzyme (Martin and Hopper, 1982). It is possible that GRO-1 is targeted by the same mechanism. As with *MOD5*, there are two ATG codons that could act as alternative sites for *gro-1* translation initiation (see Figure 7, the second ATG is boxed).

In order to determine the general expression pattern of *gro-1* in the worm and whether GRO-1 is expressed in the same sub-cellular compartments as Mod5p, a Green Fluorescent Protein (GFP) reporter construct was made. When introduced into *gro-1(e2400)* worms by microinjection, this GRO-1::GFP fusion protein was able to rescue the mutant phenotype. This indicates that whatever the expression

pattern of the fusion protein, it is active and targeted at least when and to where it is needed. The fusion protein was clearly expressed in muscle cells, the pharynx, the gut and the somatic. Expression in all cell types might be expected of a protein involved in a process such as tRNA modification. The reporter suggests that GRO-1 is localized to the nucleus and mitochondria, as well as to the cytoplasm. The same sub-cellular localization is seen in yeast but whether GRO-1 is targeted by a similar mechanism (see Introduction) can not be determined from such preliminary studies. This second ATG in *gro-1* could simply code for met and not a second start codon. This is a question that goes beyond what the molecular cloning can address and will have to be dealt with independently in the future (see Appendix E).

The other genes in the *gro-1* operon

The significance of *gro-1*'s membership in an operon may be greater than just the nature of its genomic organization and simple expression. In bacteria, genes that operate in a common system are grouped in operons in order to organize their expression. The most classical example of this if the *E. coli LacZ* operon. This operon contains three genes required for the metabolism of lactose. The organization of these genes in an operon allows for their coordinated expression and function (reviewed by Lodish *et al.*, 1996). It is unclear in *C. elegans* whether genes organized in operons are functionally involved. In most cases they are not, but some notable exceptions exist. For example, the *lin-15* locus has two

independently mutable activities and its cloning has revealed two non-overlapping transcripts organized in an operon (Huang *et al.*, 1994). The four other genes in the *gro-1* operon therefore deserve consideration.

BLAST (Altschul *et al.*, 1990) of the databases revealed homologues throughout evolution for each gene in the *gro-1* operon. *gop-1* has a human homologue of unknown function. *gop-2* also has a human homologue, as well as mouse and *S. cerevisae* homologues. None of the *gop-2* homologues have a known function, but sequence predictions indicate that it may be a GTP-binding protein. Homologues of *gop-3* of unknown function were found in human, mouse, rat, *S. pombe* and *S. cerevisae*. *hap-1*, as its name indicates, is homologous to the *S. cerevisae* gene *HAM1*. *hap-1* also has homologues in humans and mouse, in the bacteria *Bacillus subtilis*, *E. coli*, *H. influenza*, *Mycobacterium leprae* and *M. tuberculosis*, and in the primitives *Methanococcus jannaschi*, *Pyrococcus furious* and *Synechocystis sp.*

The presence of *gro-1* homologues and homologues of the other genes in the *gro-1* operon throughout evolution suggests that they may all be important and may operate in processes fundamental to the organism. This and their organization into an operon suggests that *gop-1*, *gop-2*, *gop-3* and *hap-1* may function in pathways along with *gro-1*. If this were the case, understanding these genes may give insight into the processes of the clock or aging, or tRNA modification and its role in cellular physiology.

A number of experiments can be imagined to address the possibility that *gro-1* and the other genes in the operon act in common pathways. The most direct approach would be to knock them out, either by a PCR based screen (see http://snmc01.omrf.uokhsc.edu/ revgen/RevGen.html) or by RNA interference (Fire *et al.*, 1998), to see if they produce phenotypes similar to *gro-1* (ie. a Clk phenotype). With mutations in the other genes, double mutants could be made to see if there is any genetic interaction between the genes. A less direct approach may involve the examination of their expression (either with antibodies, GFP reporters or Northern blots) in a *gro-1(e2400)* background, to see if it is altered with respect to the wild type.

How does *gro-1* fit with the "clock" hypothesis

The *gro-1(e2400)* phenotype is reminiscent of the bacterial *miaA* phenotype. The mutation affects a wide variety of physiological, metabolic and behavioral processes. Parallels can also be drawn between the regulatory role proposed for the Clock genes and for the isopentenylation of tRNA in bacteria. But how does the *gro-1* encoded IPP-transferase and the isopentenylation of tRNA fit into the concept of a "clock" mechanism responsible for temporal coordination in the worm? The results of Lakowski and Hekimi (1996) indicate that *gro-1* and the *clk* genes interact genetically, suggesting that they do in fact operate in the same mechanism effecting longevity. Based on what is known in bacteria and yeast and the mechanism by which the "clock" is believed to operate, at least two main hypothesis could be

imagined to marry the IPP-Transferase to the theory of the "clock". The first suggests that the transferase acts in the "clock" through the isopentenylation of tRNA that it catalyzes. Alternatively, there is evidence to support a role for IPP-Transferase other than tRNA modification.

A mechanism such as translation, a point of regulation in the expression of most genes, could be a powerful point of regulation, having a substantial impact on cellular physiology, metabolism and development. In bacteria it has been proposed that through the effect on translation, tRNA modification acts as an interface between the environment and the cells response to the absence of nutrients (Buck and Ames, 1984). This could be a point of regulation useful to a mechanism such as the "clock", trying to coordinate the temporal aspects of a wide variety of unrelated processes.

An alternative explanation for the involvement of IPP-Transferase in a "clock" mechanism is that the enzyme has a role in addition to the modification of tRNA, and that it is this role that is involved with the "clock". Isopentenylated adenosine has been shown to have many different roles in biology. The most common plant cytokines are adenine derivatives, such as io^6-adenine and ms^2io^6-adenine (Skoog and Armstrong, 1970). These cytokines affect respiration in plants (Skoog and Schnitz, 1979) and the induction of nitrate reductase is under cytokine control in plant tissues (Pandey and Sabharway, 1982; Kende *et al.*, 1971). In hamster cells it is believed that i^6-adenosine may be involved in triggering DNA replication during

the S phase of the cell cycle (Quesney-Huneeus *et al.*, 1982). With isopentenylated adenosine carrying out such a wide range of functions it does seem possible that the IPP-transferase could function in processes other than the modification of tRNA.

Subcellular localization experiments in yeast have provided results that support the theory of an additional IPP-transferase function. By a mechanism involving alternative translation, different forms of Mod5p are produced (see Introduction). The longer form is imported into the mitochondria, while the shorter form remains in the cytoplasm (Najarian *et al.*, 1987; Gillman *et al.*, 1991). Interestingly, this shorter form was also found to be targeted to the nucleus (Boguta *et al.*, 1994). This is surprising since there is a significant amount of evidence indicating that the cytoplasmic tRNA, while it is synthesized and receives certain modifications in the nucleus, is acted on by IPP-transferase in the cytoplasm (see Introduction). This would seem to suggest that the enzyme is targeted to the nucleus for no reason, unless it is to carry out this proposed additional function.

Another observation that hints at an additional role for IPP-Transferase is the increased level of spontaneous mutations in *miaA* mutants (Connolly and Winkler, 1989; 1991). Preliminary results indicate that there may also be an increased mutation rate in *C. elegans gro-1* mutants (Bernard Lakowski personal communication). These mutations are not simply phenocopies caused by isopentenyl deficient tRNA during translation, as they are heritable. In fact, it has been shown that in *miaA* mutants the translation error rate is actually reduced

(Bouadlown *et al.*, 1986). These mutations may be a result of the absence of DNA isopentenylation that are required, for example, for the proper replication of DNA. The fact the cell does not catch these errors suggest that perhaps the isopentenylation is required for the proper function of the cell's DNA repair mechanisms.

IPP-transferases involvement in DNA modification may also be able to explain the maternal effect of *gro-1*. It is difficult to imagine that an enzyme that modifies tRNA could result in a maternal effect. If maternal IPP-transferase or its mRNA were deposited in the oocyte, it seems unlikely that it would be capable of carrying out the isopentenylation of tRNA in all the cells and over the entire life of the worm. But if the maternal enzyme were to modify the oocyte or embryonic DNA, it could cause the observed maternal effect.

Summary

The *C. elegans* Clock gene *gro-1* has been cloned and it encodes a metazoan N^6-(Δ^2) isopentenyl PPi :tRNA isopentenyl transferase (IPP-transferase; Bartz *et al.*, 1970). There are homologues known in bacteria (*miaA*; Connoly and Winkler, 1991) and yeast (*MOD5*; Dihanich *et al.*, 1987). This enzyme is responsible for the isopentenylation of the adenosine residue at position 37 of certain tRNA (Bartz *et al.*, 1970). Preliminary expression studies suggest that GRO-1 is expressed in many cell types and that as in yeast, it is localized to the mitochondria, nucleus and

cytoplasm. A much more in depth study will have to be carried out in an attempt to determine the mechanism driving GRO-1 localization. An interesting result was that *gro-1* is the fifth member of an operon. This may identify four more genes of interest to the study of the "clock" or tRNA modification and its role in cellular physiology. *gro-1* has been shown to be involved with the *clk* genes, in a mechanism proposed to control and coordinate the temporal aspects of many processes in the development and life of the worm. While it is not possible to formulate a unifying theory of IPP-transferase involvement in the "clock" mechanism based on the results from bacteria, yeast and *C. elegans*, a number of interesting possibilities have been proposed.

Agris, P.F., Armstrong, D.J., Schafer, K.P., and Söll, D. 1975. Maturation of a hypermodified nucleoside in transfer RNA. *Nucl. Acids. Res* **2**: 691-698.

Albrecht, U., Z.S. Sun, G. Eichels, and C.C. Lee. 1997. A differential response of two putative mammalian circadian regulators, *mper1* and *mper2*, to light. *Cell* **91**: 1055-1064.

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

Ausubel, F.M., R. Brent, R.E. Kingston, D.M. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds. 1992. *Short Protocols in Molecular Biology*. John Wiley & Sons, New York.

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

Bartz, J.K., L.K. Kline, and D. Söll. 1970. N^6-(Δ^2-Isopentenyl)adenosine: biosynthesis *in vitro* in transfer RNA by an enzyme purified from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **40**: 1481-1487.

Baumstark, B.R., L.L. Spremulli, U.L RajBhandary, G.M Brown. 1977. Initiation of protein synthesis without formylation in a mutant of *Escherichia coli* that grows in the absence of tetrahydrofolate. *J Bacteriol.* **129**:457-471.

Boguta, M., L.A. Hunter, W.-C. Shen, E.C. Gillman, N.C. Martin and A.K. Hopper. 1994. Subcellular locations of *MOD5* proteins: mapping of sequences sufficient for targeting to mitochondria and demonstration that mitochondrial and nuclear isoforms commingle in the cytosol. *Mol. Cell. Biol.* **14**: 2298-2306.

Bohr, V. A., and R.M. Anson. 1995. DNA damage, mutation and fine structure DNA repair in aging. *Mutat. Res.* **338**: 25-34.

Blum, P.H. 1988. Reduced *Leu* operon expression in a *miaA* mutant of *Salmonella typhimurium*. *J. Bacteriol.* **170**: 5125-5133.

Bouadloun, F., T. Srichaiyo, L.A. Isaksson, and G.R. Björk. 1986. Influence of a modification next to the anticodon in tRNA on codon context sensitivity of translational suppression and accuracy. *J. Bacteriol.* **166**: 1022-1027.

Brenner, S. 1974. the Genetics of *Caenorhabditis elegans. Genetics* **77**: 71-94.

Broach, J.R., and R.J. Deschenes. 1990. The function of *ras* genes in *Saccharomyces cerevisiae*. *Adv. Cancer Res.* **54**: 79-139.

Buck, M., and B.N. Ames. 1984. A modified nucleoside in tRNA as a possible regulator of aerobiosis: synthesis of *cis*-2-methyl-thioribosylzeatin in the tRNA of *Salmonella*. *Cell* **36**: 523-531.

Buck, M., and E. Griffiths. 1981. Regulation of aromatic amino acid transport by tRNA: role of 2-methylthio-N^6-(Δ^2-isopentenyl)adenosine. *Nuc. Acids Res.* **9**: 401-414.

Björk, G.R. 1980. A novel link between the biosynthesis of aromatic amino acids and transfer RNA modification in *Escherichia coli*. *J. Mol. Biol.* **140**: 391-410.

Björk, G.R. 1987. Modification of stable RNA. p. 719-731. In F.C. Neidhardt, J.L. Ingrahm, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (ed.). *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology. Washington, D.C.

Björk. G.R., and F.C., Neidhardt. 1975. Physiological and biochemical studies on the function of 5-methyluridine in the transfer ribonucleic acid of *Escherichia coli*. *J. Bacteriol.* **124**: 99-111.

Björk, G.R., P.M. Wikstrom, and A.S. Bystrom. 1989. Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science* **244**: 986-989.

Caillet , J. and L. Droogmans. 1988. Molecular cloning of the *Escherichia coli miaA* gene involved in the formation of Δ^2-isopentenyl adenosine in tRNA. *J. Bacteriol.* **170**: 4147-4152.

Chance, B., H. Sies, and A. Boveris. 1979. Hydroperoxide metabolism in mammalian organs. *Physiol. Rev.* **59**: 527-605.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chlorophorm extraction. *Anal. Biochem* **162**: 156-159.

Cohen, A.H., S. Rossignol, and S. Grillner. 1988. *Neuronal Control of Rhythmic Movements in Vertebrates*. Wiley-Interscience, New-York.

Connolly, D.M., and M.E. Winkler. 1989. Genetic and physiological relationships among the *miaA* gene involved in the formation of 2-methylthio-N^6-(Δ^2-isopentenyl)adenosine tRNA modification, and spontaneous mutagenesis in *Escherichia coli* K-12. *J. Bacteriol.* **171**: 3233-3246.

Connolly, D.M., and M.E. Winkler. 1991. Structure of *Escherichia coli* K-12 *miaA* and characterization of the mutator phenotype caused by *miaA* insertion mutations. *J. Bacteriol.* **173**: 1711-1721.

Craxton, M. 1993. Cosmid sequencing. *Methods Mol. Biol.* **23**: 149-167.

Dinanich, M.E., D., Najarion, R. Clark, E.C. Gillman, N.C. Martin, and A.K. Hopper. 1987. Isolation and characterization of *MOD5*: a gene required for isopentenylation of cytoplasmic and mitochondrial tRNAs of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **7**: 177-184.

D'mello, N.P., A.M. Chilress, D.S, Franklin, S.P. Kale, C. Pinwasdi, and S.M. Jazwinski. 1994. Cloning and characterization of *LAG1*, a longevity-assurance gene in yeast. *J. Biol. Chem.* **269**: 15451-15459.

Eeckman, F.H., and R. Durbin. 1995. AceDB and macece. *Methods Cell Biol.* **48**: 583-605.

Edmonds, C.G., P.F. Crain, R. Gupta, T. Hashizume, C.H.Hocart, J.A. Kowalak, S.C. Pomerantz, K.O. Stetter, J.A. McClosky. 1991. Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). *J. Bacteriol.* **173**: 3138-3148.

Eide, D., and P. Anderson. 1988. Insertion and excision of *Caenorhabditis elegans* transposable elements Tc1. *Mol. Cel. Biol.* **8(2)**: 737-746.

Ericson, J.U., and G.R. Björk. 1986. Pleiotropic effects induced by modification deficiency next to the anticodon of tRNA from *S. typhimurium* LT2. *J. Bacteriol.* **166**: 1013-1021.

Evans, D.A., J.P. Burbach, and F.W. VanLeeuwen. 1995. Somatic mutations in the brain: relationship to aging? *Mutat. Res.* **338**: 173-182.

Ewbank, J.J, T.M. Barnes, B. Lakowski, M. Lussier, H. Bussey, and S. Hekimi. 1997. Structural and functional conservation of the *Caenorhabditis elegans* timing gene *clk-1*. *Science* **275**: 980-983.

Feig, D.I., T.M. Reid, and L.A. Loeb. 1994. Reactive oxygen species in tumorigenesis. *Cancer Res.* **54, Suppl.**: 1890s-1894s.

Feldman, J.F., and M.N. Hoyle. 1973. Isolation of circadian clock mutants of *Neurospora crassa*. *Genetics* **75**: 605-613.

Fire, A. 1995. Integrative transformation of *Caenorhabditis elegans*. *EMBO J.* **5**: 2673-2680.

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

Fleissner, G. 1982. Isolation of an insect circadian clock. *J. Comp. Physiol.* **149**: 311-316.

Foulkes, N.S., J. Boligin, S.H. Snyder, and P. Sassone-Corsi. 1997. Rhythmic transcription: the molecular basis of circadian melatonin synthesis. *Trends Neurosci.* **20**: 487-492.

Frohman, M.A., M.K. Dush, G.R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**: 8998-9002.

Gardner, G.F., and J.F. Feldman. 1980. The *frq* locus in *Neurospora crassa*: a key element in circadian clock organization. *Genetics* **96**: 877-886.

Geffer, M.L. 1969. The *in vitro* synthesis of 2'-o-methylguanosine and 2-methylthio-N^6(dimethylallyl)adenosine in transfer RNA of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **36**: 435-441.

Gillman, E.C., L.B. Slusher, N.C. Martin, and A.K. Hopper. 1991. *MOD5* translation initiation sites determine N^6-isopentenylation modification of mitochondrial and cytoplasmic tRNA. *Mol. Cell. Biol.* **11**: 2382-2390.

Gourishanker, J., and J. Pittard. 1982. Regulation of phenylalanine biosynthesis in *Escherichia coli* K-12: control of transcription of the *pheA* operon. *J. Bacteriol.* **150**: 1130-1137.

Grollman, A.P., and M. Moriya. 1993. Mutagenesis by 8-oxyguanine: an enemy within. *Trends Genet.* **9**: 246-249.

Grosjean, H., and H. Chantrenne. 1980. On codon-anticodon interactions. In molecular biology, biochemistry, and biophysics, chemical recognition in biology. F. Chapeville and L.-A. Haenni, ed. (Berlin: Springer-Verlag), Vol. 32, pp. 347-367.

Hagervall, T.G., T.M. Tuohy, J.F. Atkins and G.R. Björk. 1993. Deficiency of 1-methylguanosine in tRNA from Salmonella typhimurium induces frameshifting by quadruplet translocation. *J. Mol. Biol.* **232(3)**: 756-765.

Hardin, P.E., J.C. Hall, and M. Rosbash. 1990. Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature* **343**: 536-540.

Harman, D. 1956. Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* **2**: 298-30.

Harman, D. 1981. The aging process. *Proc. Natl. Acad. Sci.* **78(11)**: 7124-7128.

Hekimi, S., P. Boutis, and B. Lakowski. 1995. Viable maternal-effect mutations that affect the development of the nematode *Caenorhabditis elegans*. *Genetics* **146**:149-64.

Hodgkin J., and T. Doniach. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans. Genetics* **146**: 149-164.

Holmes, D.S. and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**: 193-197.

Hooper, A.K., F. Banks, and V. Evangelidis. 1978. A yeast mutant which accumulates precursor tRNAs. *Cell* **14**: 211-219.

Hopper, A.K., A.H. Furukawa, H.D. Pham, N.C. Martin. 1982. Defects in modification of cytoplasmic and mitochondrial transfer RNAs are cause by single nuclear mutations. *Cell* **28**: 543-550.

Huang, L.S., P. Tzou and P.W. Sternberg. 1994. The *lin-15* locus encodes two negative regulators of *Caenorhabditis elegans* vulval development. *Mol. Biol. Cell* **5(4)**: 395-411.

Huang, X.,-Y. and D. Hirsh. 1989. A second *trans*-spliced RNA leader sequence in the nematode *Caenorhabditis elegans. Proc. Natl. Acad. Sci.* **86**: 8640-8644.

Jazwinski, S.M. 1996. Longevity, Genes, and Aging. *Science* **273**: 54-59.

Jazwinski, S.M. 1993. In, *handbook of the Biology of Aging*, E.L. Schneider and J.W. Rowe, eds. (Academic Press, San Diego, CA, ed. 4, 1996), pp. 39-54; *Genetica* **91**: 35, 1993.

Jonnasen, T., B.N. Marbois, L. Kim, A. Chin, Y.R. Xia, A.J. Lusis, C.F. Clarke. 1996. Isolation and sequencing of the rat Cog7 gene and the mapping of mouse Cog7 to chromosome 7. *Arch. Biochem. Biophys.* **330(2)**: 285-289.

Jukes, T.H. 1973. Possibilities for the evolution of the genetic code from a preceding form. *Nature* **246**: 22-26.

Kende, H., H. Hahn, and S.E. Kays. 1971. Enhancement of nitrate reductase activity by benzyladenine in *Agrostemma gihago. Plant Physiol.* **48**:702-706.

Kenyon, C., J. Chang, E. Gensch, A. Rudner, and R. Tabtiang. 1993. A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366**: 461-464.

King, D.P., Y. Zhao, A.M. Sangoram, L.D.Wilsbacher, M. Tanaka, M.P. Antoch, T.D.L. Steeves, M.H. Vitaterna, J.M. Kornhauser, P.L. Lowery, F.W. Turek, and J.S. Takahashi. 1997. Positional cloning of the mouse circadian *clock* gene. *Cell* **89**: 641-653.

Kirkwood, T.B.L. 1997. Evolution and aging. *Nature* **270**: 310-304.

Kirkwood, T.B., and M.R. Rose. 1991. Evolution of senescence: late survival sacrificed for reproduction. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **332**: 15-24.

Klein, D.C., R.Y. Moore, and S.M. Reppert. 1991. Suprachiasmatic nucleus, In, The Mind's Clock. Oxford University Press, New York.

Knapp, G., R.C. Ogden, C.L. Peebles, and J. Abelson. 1979. Splicing of yeast tRNA precursors: structure of the reaction intermediates. *Cell* **18**: 37-45.

Konopka, R.J., and S. Benzer. 1971. Clock mutants of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **68**: 2112-2116.

Lakowski, B. 1998. Gentic factors affecting life-span in the nematode *Caenorhabditis elegans*. Ph.D. Thesis, McGill University.

Lakowski, B., and S. Hekimi. 1996. Determination of the life-span in *Caenorhabditis elegans* by four clock genes. *Science* **272**: 1010-1013.

Larsen, P.L., P.S. Albert, and D.L.Riddle. 1995. Genes that regulate both development and longevity in *Caenorhabditis elegans. Genetic* **139**: 1567-1583.

Laten, H., J. Gorman, and R.M. Bock. 1978. Isopentenyladenosine deficient tRNA from an antisuppressor mutant of *Saccharomyces cerevisiae*. *Nuc. Acids. Res.* **5**: 4329-4342.

Lodish, H., D. Baltimore, A. Berk, S.L. Zipurski, P. Matsudaira, J. Darnell. (1996) The Molecular Biology of the Cell 3rd ed. p. 406

Loher, W. 1972. Circadian control of stridulation in the cricket *Teleogryllus commodus*. *J. Comp. Physiol.* **79**: 173-190.

Loher, W. 1974. Circadian control of spermatophore formation in the cricket *Teleogryllus commodus. J. Insect Physiol.* **20**: 1155-1172.

Luft, R. 1994. The development of mitochondrial medicine. *Proc. Natl. Acad. Sci. USA* **91**: 8731-8738.

Martin, N.C., and A.K. Hopper. 1982. Isopentenylation of both cytoplasmic and mitochondrial tRNA is affected by a single nuclear mutation. *J. Biol. Chem.* **257**: 10562-10565.

McCray, J.W., and K.M. Hermann. 1976. Depression of certain aromatic amino acid biosynthetic enzymes of *Escherichia coli* K-12 by growth in Fe^{3+}-defecient medium. *J. Bacteriol.* **125**: 608-615.

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

Messing, J. 1983. New M13 vectors for cloning. *Methods enzymol.* **101**: 20.

Miquel, J. 1992. An update on the mitochondrial-DNA mutation hypothesis of cell aging. *Mutat. Res.* **275**: 209-216.

Morley, A.A. 1995. The somatic mutation theory of ageing. *Mutat. Res.* **338**: 19-23.

Mortimer, R.K., and J.R. Johnston. 1959. *Nature* **183**:

Najarian, D., M.E. Dihanich, N.C. Martin, and A.K. Hopper. 1987. DNA sequence and transcript mapping of *MOD5*: features of the 5' region which suggests two translational starts. *Mol. Cell. Biol.* **7**: 185-191.

Nishikura, K., and E.M. DeRobertis. 1981. RNA processing in microinjected *Xenopus* oocytes: sequential addition of base modifications in the spliced transfer RNA. *J. Mol. Biol.* **7**: 185-191.

Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**: 101.

Orr, W.C., and R.S. Sohal. 1994. Extension of life-span by overexpression of superoxid dismutase and catalase in *Drosophila melanogaster*. *Science* **263**: 1128-1130.

Page, T.L. 1984. Neuronal organization of a circadian clock in the cockroach *Leucophaea maderae*, In, Photoperiodic regulation of insect and molluscan hormones. R. Porter, G.M. Collins (eds.). Pitman, London. pp. 115-135.

Palmer, D.T., P.H. Blum, and S.W. Artz. 1983. Effects of the *hisT* mutation of *Salmonella typhimurium* on translation elongation rate. *J. Bacteriol.* **153**: 357-363.

Pandey, K.N., and P.S. Sabharwal. 1982. γ-irradiation activates biochemical systems: induction of nitrate reductase activity in plant callus. *Proc. Nat. Acad. Sci. USA* **79**:5460-5464.

Pearl, R. 1928. The Rate of Living. London: Univ. of London Press.

Persson, B.C. 1993. Modification of tRNA as a regulatory device. *Molec. Microb.* **8**: 1011-1016.

Petrullo, L.A., P.J. Gallagher, and D. Elseviers. 1983. The role of 2-methylthio-N^6-isopentenyladenoasine in readthrough and suppression of nonsense codons in *Escherichia coli. Mol. Gen. Genet.* **190**:289-294.

Phillips, J.H., and K. Kjellin-Straby. 1967. Studies on microbial ribonucleic acid. IV. Two mutants of *Saccharomyces cerevisiae* lacking *N*-2-dimethylguanine in soluble ribonucleic acid. *J. Mol. Biol.* **26**: 509-518.

Peiczenik, G. 1980. Predicting coding function from nucleotide sequence or survival of "fitness" of tRNA. *Proc. Natl. Acad. Sci. USA* **77**:3539-3543.

Quesney-Huneeus, V., H.A. Glaick, M.D. Siperstein, S.K. Ericson, T.A. Spencer, and J.A. Nelson. 1982. The dual role of mevalonate in the cell cycle. *J. Biol. Chem.* **258**: 378-385.

Ralph, P.E., and M. Menaker. 1988. A mutation of the circadian system in golden hamsters. *Science* **241**: 1225-1227.

Riddle, D.L. 1987. Post-embryonic development in *Caenorhabditis elegans. Int. J. Parasit.* **17**: 223-231.

Riddle, D.L., M.M. Swanson and P.S. Albert. 1981. Interacting genes in nematode dauer larva formation. *Nature* **290**: 668-671.

Rose, M.R., and C.E. Finch. 1994. Genetics and evolution of aging. In, *Contemporary Issues in Genetics and Evolution.* Dordrecht, Germany: Kluwer, vol. 3.

Ruan, K. and S.W. Emmons. 1984. Extrachromosomal copies of transposon Tc1 in the nematode *Caenorhabditis elegans. Proc. Natl. Acad. Sci.* **81**: 4018-4022.

Rubner, M. 1908. *Das problem der Lebensdauer und seine Beziehungen zu Wachstum und Ernehrung.* Munich: Oldenburg.

Samuel, C.E., and J.C. Rabinowitz. 1974. Initiation of protein synthesis by folate-sufficient and folate-deficient *Streptococcus faecalis* R. Biochemical and

biophysical properties of methionine transfer ribonucleic acid. *J. Biol. Chem.* **249**: 1198-1206.

Shearman, L.P., M.J. Zylka, D.R. Weaver, J.F. Kolakowski Jr., and S.M. Reppert. 1997. Two *period* homologues: circadian expression and photic regulation in the superchiasmatic nuclei. *Neuron* **19**: 1261-1269.

Skoog, F., and D.J. Armstrong. 1970. Cytokines. *Ann. Rev. Plant Physiol.* **21**: 359-384.

Skoog, F., and R.Y. Schmitz. 1979. Biochemistry and physiology of cytokines. In biochemical actions of hormones. G. Litwck, ed. (New York Academic Press). Vol. 6, pp. 335-413.

Sohal, R.S. 1976. Metabolic rate and life span. In: *Cellular Aging: Concepts and Mechanisms*, edited by R. Witler. Basel: Karger, p. 25-40.

Sokolove, P.G., and W. Loher. 1975. Role of eyes, optic lobes, and pars intercerebrallis in locomotory and tridulatory circadian rhythms of *Teleogryllus commodus. J. Insect Physiol.* **21**: 785-799.

Spieth, J., G. Brooke, S. Kuersten, K. Lea, and T. Blumenthal, 1993. operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. *Cell* **73**: 521-532.

Sulston, J., and J. Hodgkin, 1988. Methods, In, *The Nematode Caenorhabditis elegans* (Wood W.B. and the Community of *C. elegans* Researchers, eds) pp. 587-606. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Sun, Z.S., U. Albrecht, O. Zhuchenko, J. Bailey, G. Eichele, and C.C. Lee. 1997. RIGUI, a putative mammalian ortholog of the *Drosophila period* gene. *Cell* **90**: 1003-1011.

Tei, H.,H. Okamura, Y. Shigeyoshi, C. Fukuhara, R. Ozawa, M. Hirose, and Y. Sakaki. 1997. Circadian oscillation of a mammalian homologue of the *Drosophila period* gene. *Nature* **389**: 512-516.

Vacher, J., H. Grosjean, C. Houssier, and R.H. Buckingham, 1984. The effect of point mutations affecting *Escherichia coli* tryptophan tRNA on codon-anticodon interactions and on UGA suppression. *J. Mol. Biol.* **177**: 329-342.

VanLuenen, H.G.A.M. and R.H.A. Plasterk. 1994. Target site choice of the related transposable Tc1 and Tc3 of *Caenorhabditis elegans. Nucleic Acid Res.* **22**: 262-269.

VanVoorhies, W.A. 1992. Production in sperm reduces nematode lifespan. *Nature* **360**: 456-458.

Vilg, J., and J.A. Gossen. 1993. Somatic mutations ans cellular aging. *Comp. Biochem. Physiol B Biochem. Mol. Physiol.* **104**: 429-437.

Vitaterna, M.H., D.P. King, A.M. Chang, J.M. Kornhauser, P.L. Lowrey et al., 1994. Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian bahaviour. *Science* **264**: 719-725.

Wilson, C.R., and B.A. Roe. 1989. Presence of the hypermodified nucleoside N^6-(Δ^2-isopentenyl)-2-methylthioadenosine prevents codon misreading by *Escherichia coli* phenylalanyl-transfer RNA. *Proc. Natl. Acad. Sci. USA* **86**: 409-413.

Wong, A., P. Boutis, ans S. Hekimi. 1995. Mutations in the *clk-1* gene of *Caenorhabditis elegans* affect developmental and behavioral timing. *Genetics* **139**: 1247-1259.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strain: Nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene* **33**: 103.

Yanofsky, C., and L. Söll. 1977. Mutations affecting $tRNA^{Trp}$ and its charging and their effect on regulation of transcription termination at the attenuator of the tryptophan operon. *J. Mol. Biol.* **113**: 663-677.

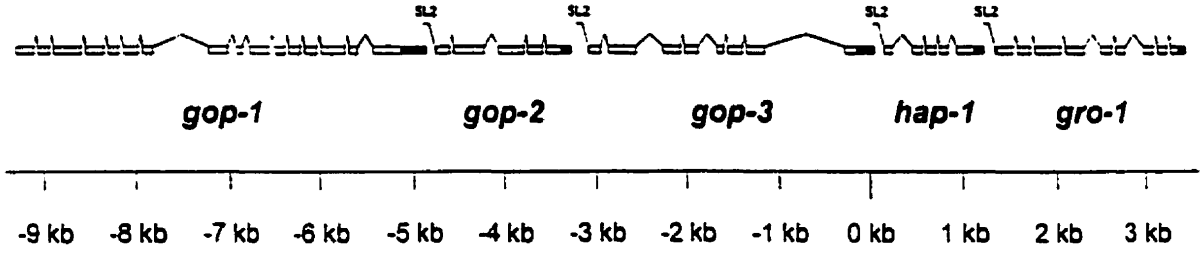
Zarawski, G., K. Brown, D. Killingly, C. Yanofsky. 1978. Nucleoside sequence of the leader region of the phenylalanine operon of *Escherichia coli. Proc. Nat. Acad. Sci. USA* **75**: 4271-4275.

Zerr, D.M., J.C. Hall, M. Rosbash, and K.K. Siwicki. 1990. Circadian fluctuations of *period* protein immunoreactivity in the CNS anf the visual system of *Drosophila. J. Neurosci.* **10**: 2749-2762.

Zorio, D.A., N.N. Cheng, T. Blumenthal, and J. Spieth. 1994. Operons as a common form of chromosomal organization in *C. elegans. Nature* **372**: 270-272.

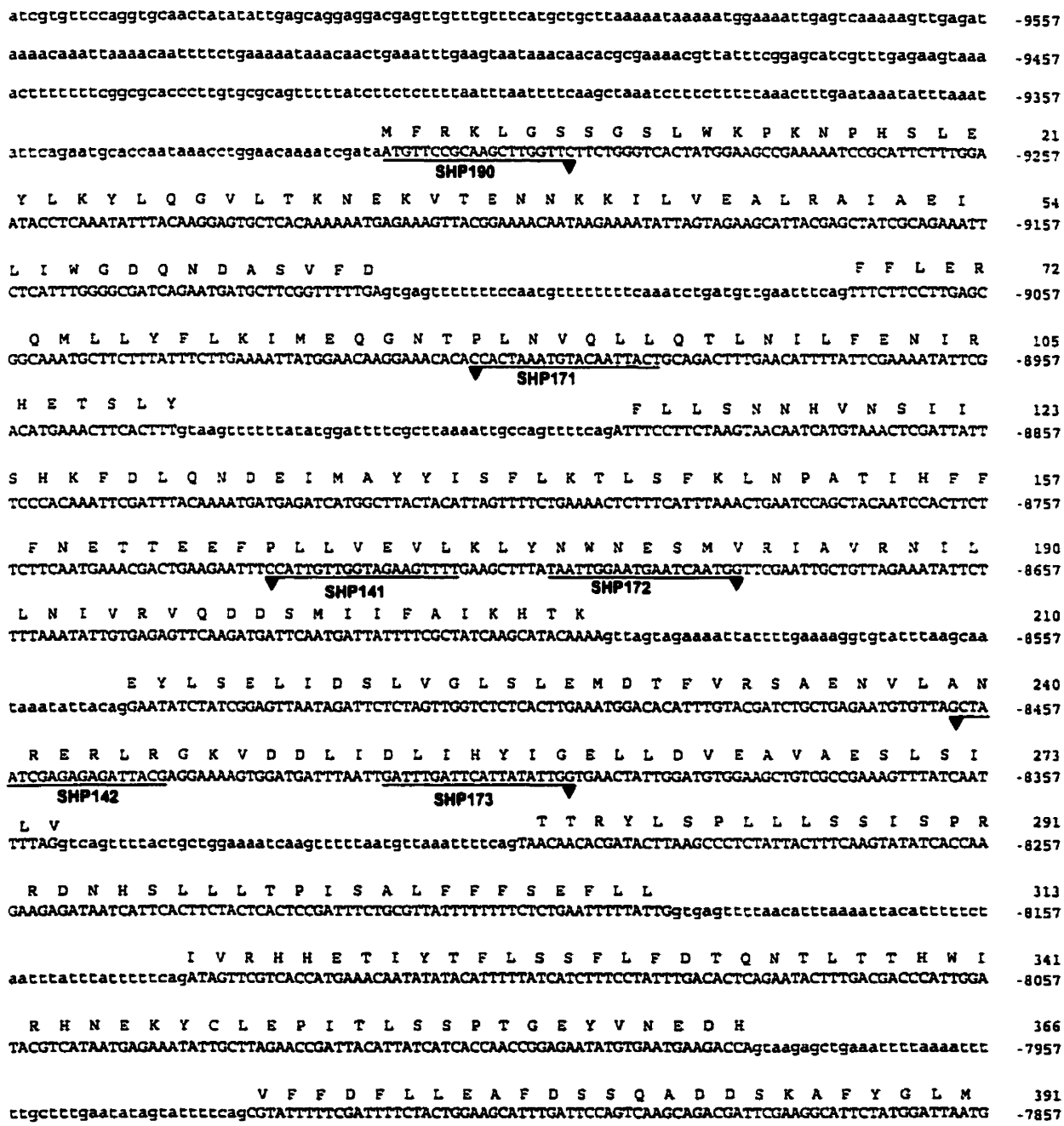
Appendix A: Sequence of the *gro-1* operon

Included in this appendix is the entire genomic sequence of the *gro-1* operon. The coding sequence is in capital letters. The protein sequence is given in single letter code above the DNA sequence. The *trans*-splicing sites, the sites of polyadenylation and the *e2400* mutation are marked. All the primers used to confirm the splicing and to identify the *e2400* mutation are illustrated. Below is a graphical representation of the genes in the *gro-1* operon. The nucleotide scale in the figure and in the sequence is zeroed at the left end of the cosmid clone ZC395.

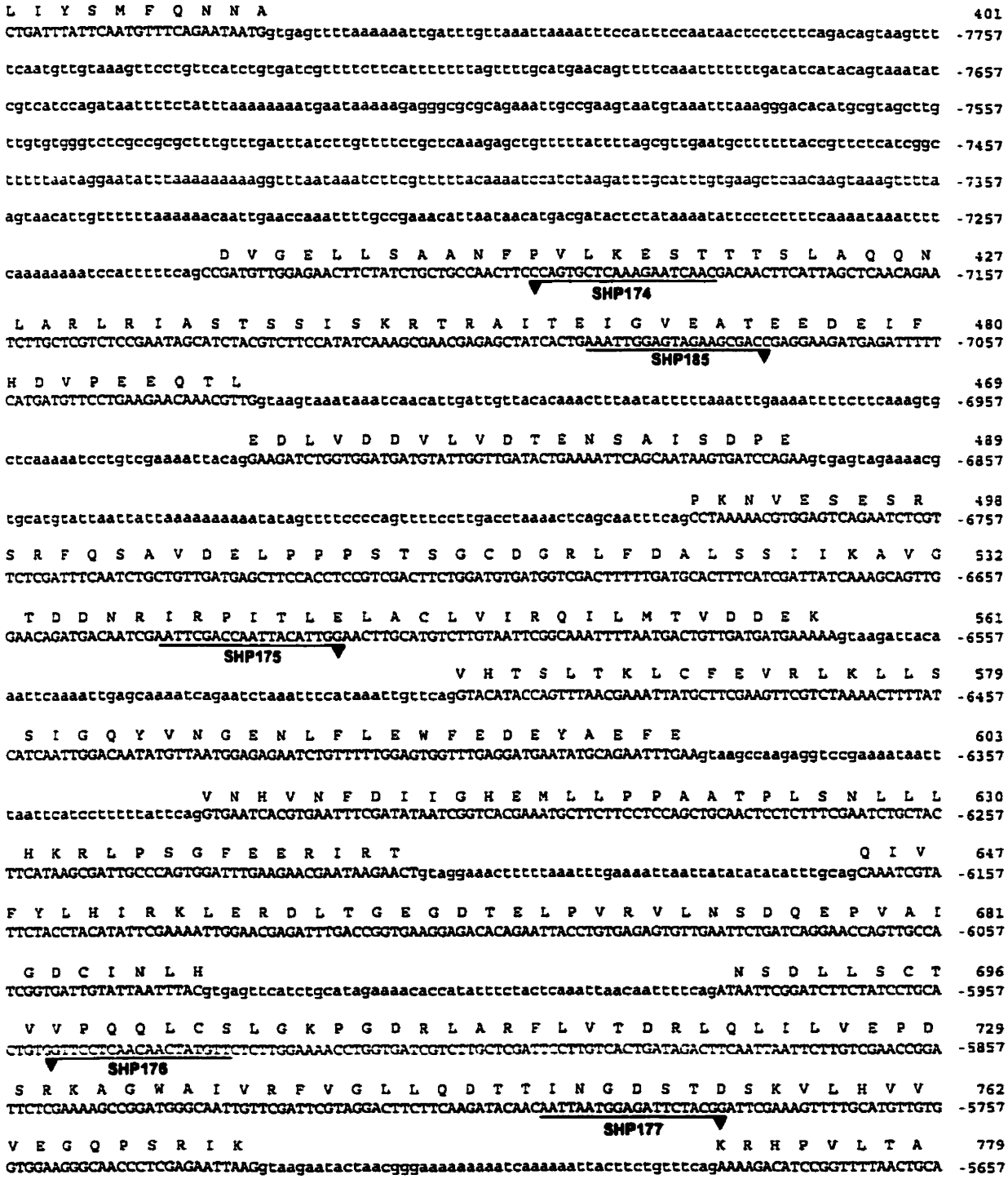


N.B. The coding sequence is represented by open boxes. 5' and 3' UTRs are represented by filled boxes. Trans-splice leader sequences are drawn as lines above the coding sequence.

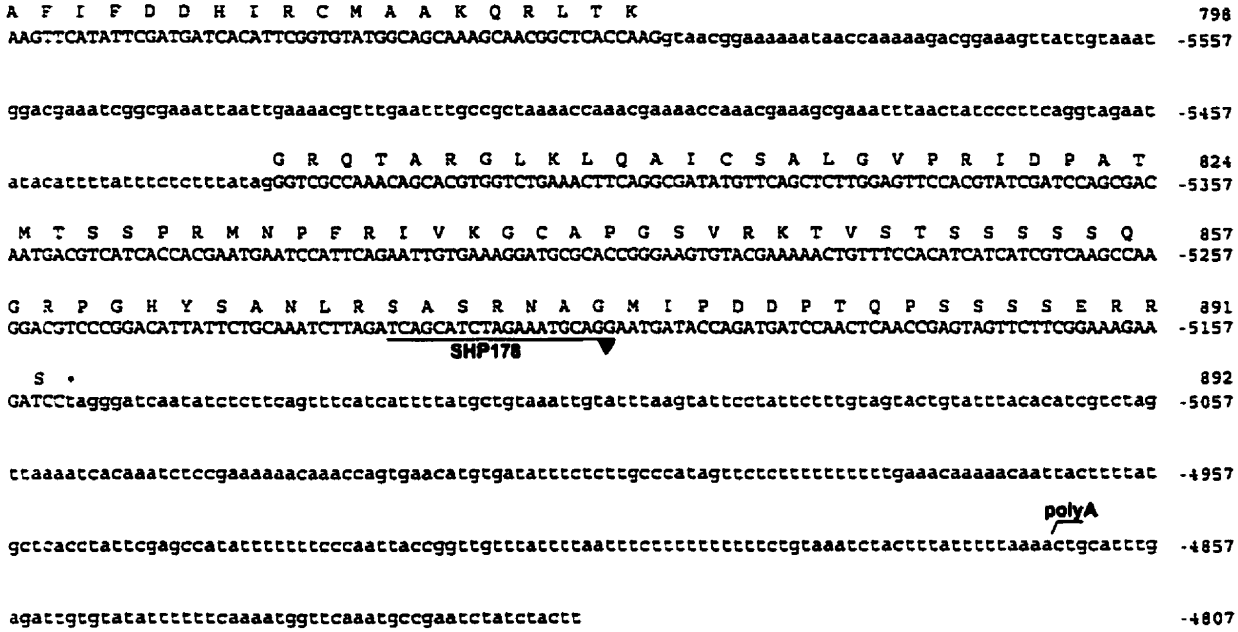
gop-1



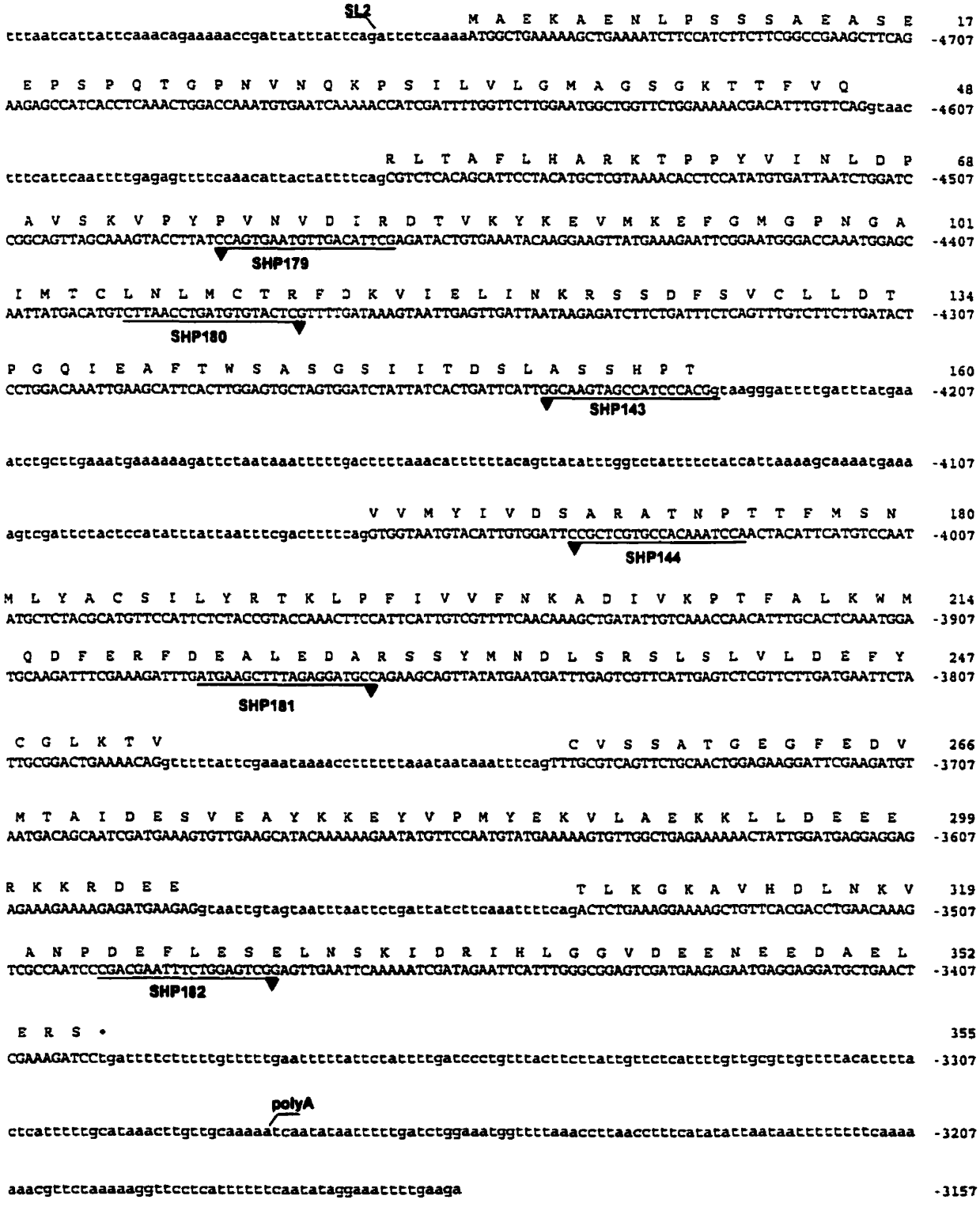
gop-1 continued...



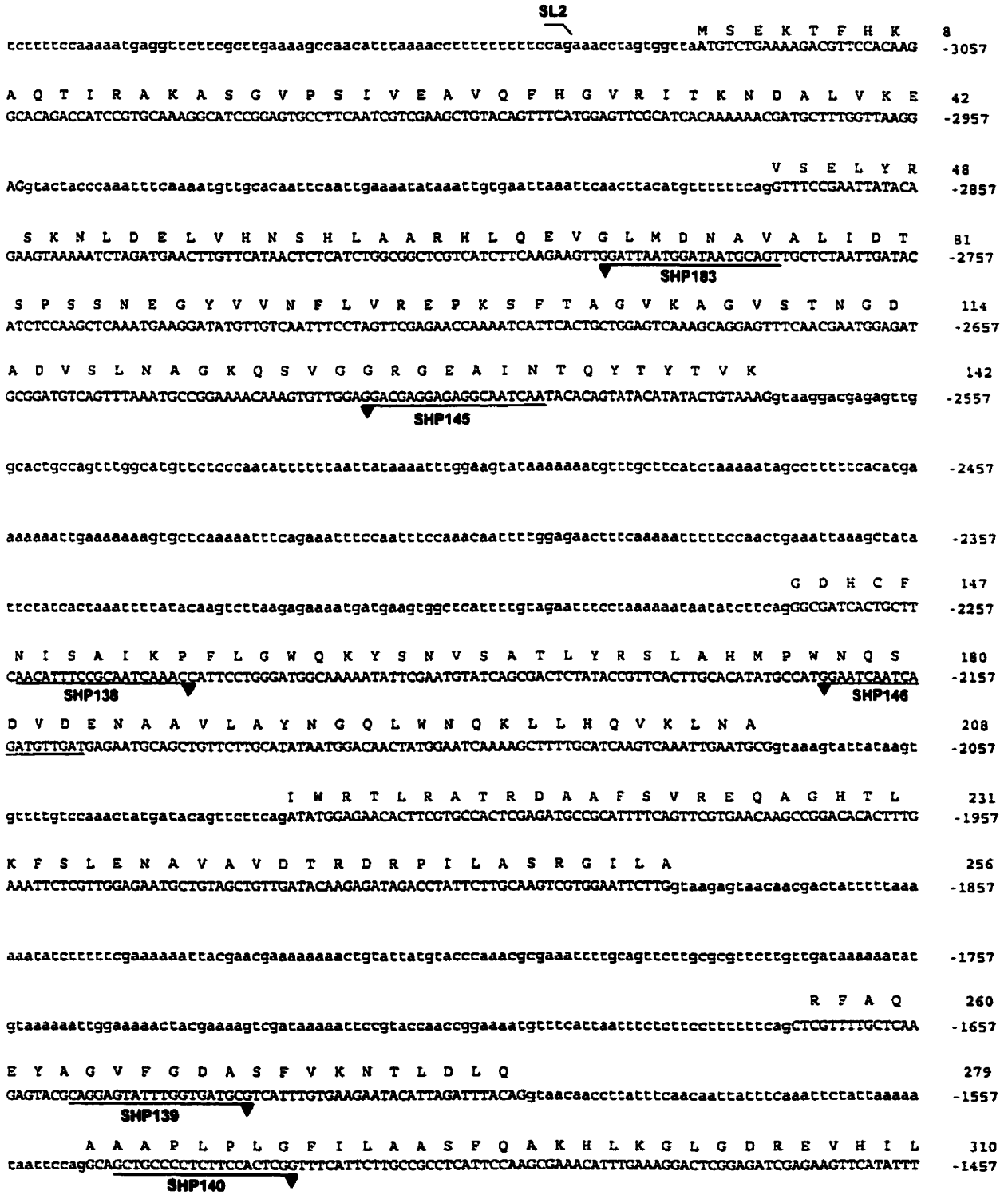
gop-1 continued...



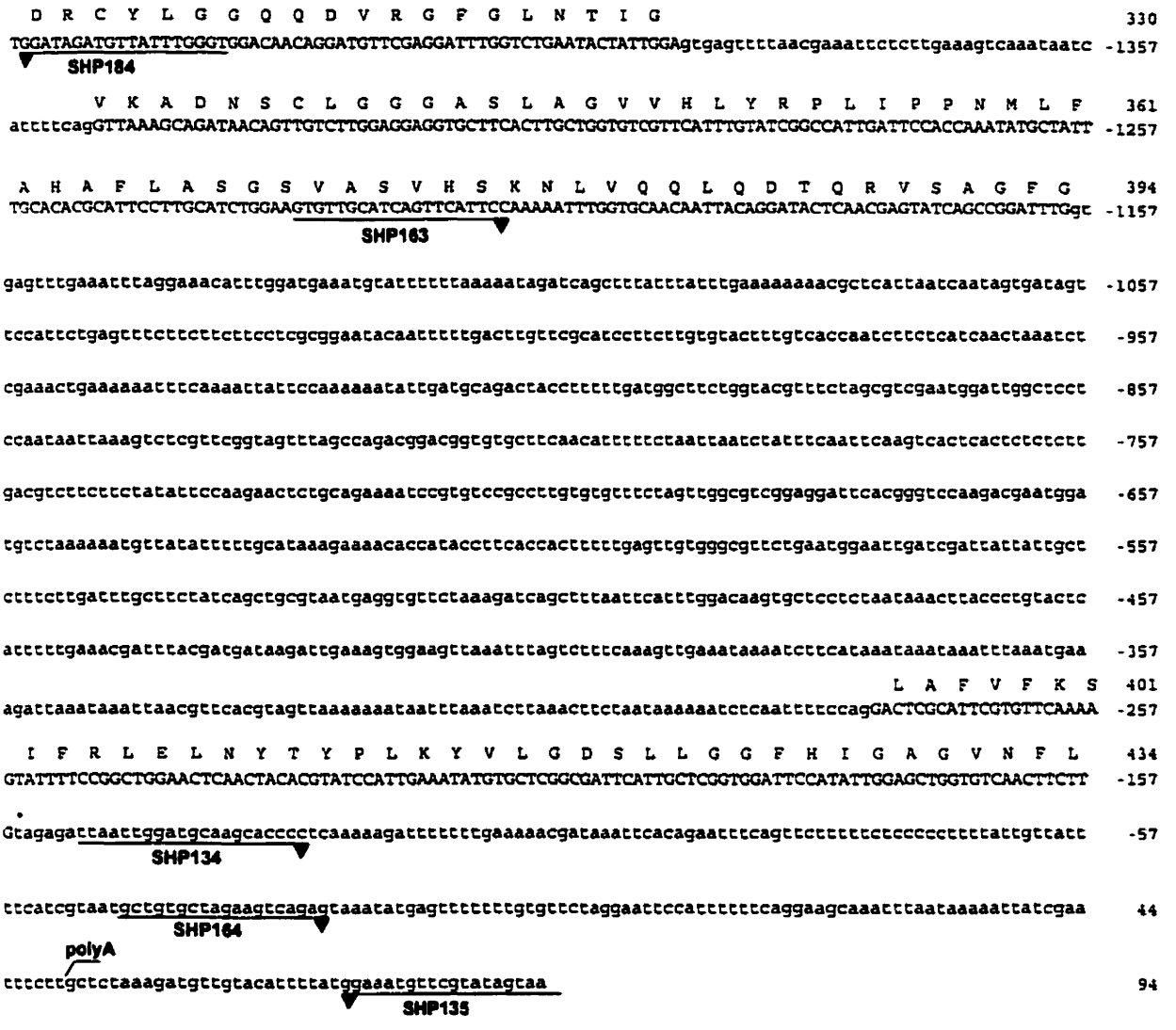
gop-2



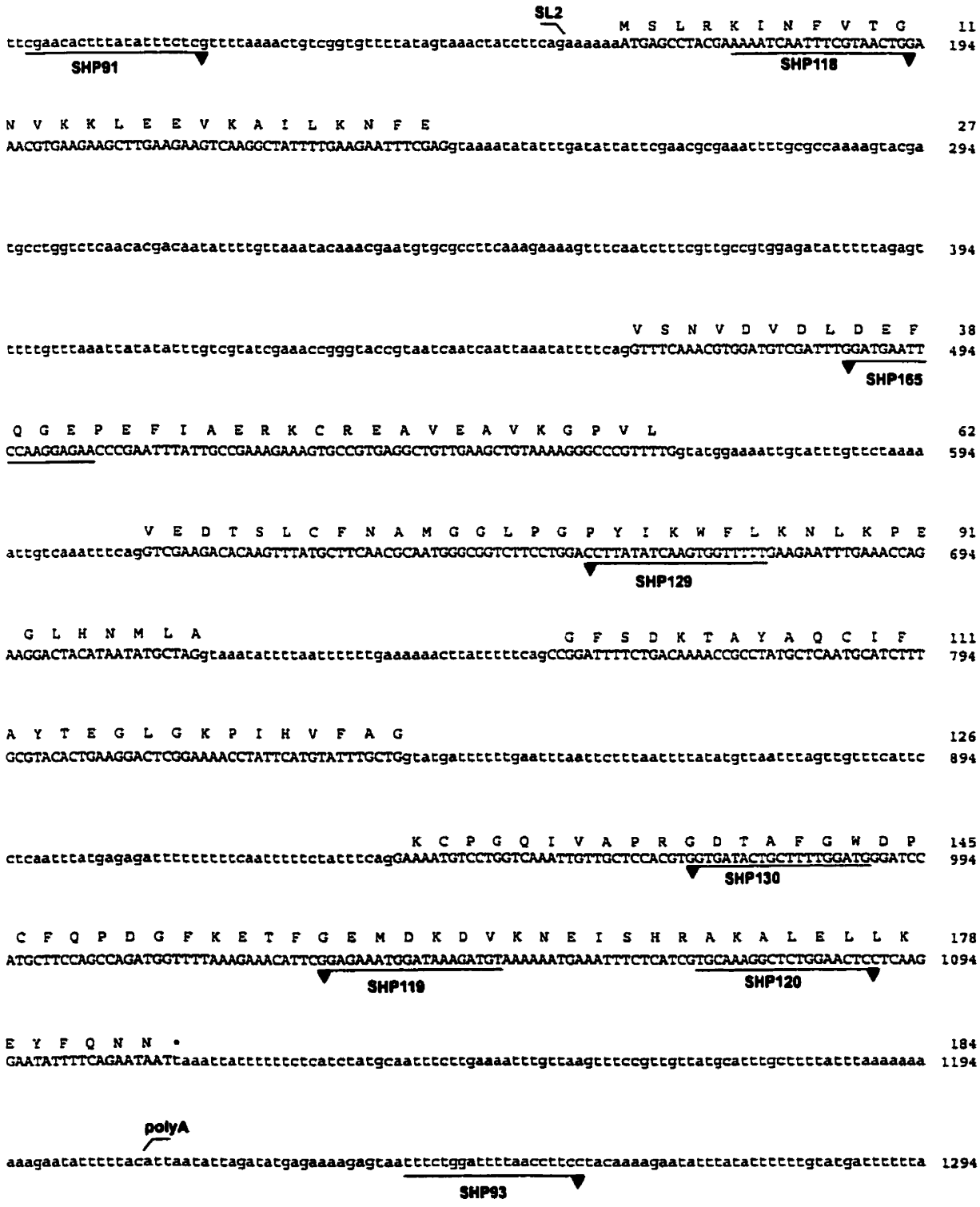
gop-3


SL2
M S E K T F H K 8
tcttttccaaaaatgaggttcttcgcttgaaaagccaacatttaaaaccttttttttccagaaacctagtggttaATGTCTGAAAAGACGTTCCACAAG -3057
A Q T I R A K A S G V P S I V E A V Q F H G V R I T K N D A L V K E 42
GCACAGACCATCCGTGCAAAGGCATCCGGAGTGCCTTCAATCGTCGAAGCTGTACAGTTTCATGGAGTTCGCATCACAAAAAACGATGCTTTGGTTAAGG -2957
V S E L Y R 48
AGgtactacccaaatttcaaaatgttgcacaattcaattgaaaatataaattgtgaattaaattcaacttacatgttttttcagGTTTCCGAATTATACA -2857
S K N L D E L V H N S H L A A R H L Q E V G L M D N A V A L I D T 81
GAAGTAAAAATCTAGATGAACTTGTTCATAACTCTCATCTGGCGGCTCGTCATCTTCAAGAAGTTGGATTAATGGATAATGCAGTTGCTCTAATTGATAC -2757
SHP183
S P S S N E G Y V V N F L V R E P K S F T A G V K A G V S T N G D 114
ATCTCCAAGCTCAAATGAAGGATATGTTGTCAATTTCCTAGTTCGAGAACCAAAATCATTCACTGCTGGAGTCAAAGCAGGAGTTTCAACGAATGGAGAT -2657
A D V S L N A G K Q S V G G R G E A I N T Q Y T Y T V K 142
GCGGATGTCAGTTTAAATGCCGGAAAACAAAGTGTTGGAGGACGAGGAGAGGCAATCAATACACAGTATACATATACTGTAAAGgtaaggacgagagttg -2557
SHP145
gcactgccagtttggcatgttctcccaatatttttaattataaaatttggaagtataaaaaaatgtttgcttcatctaaaaatagccttttcacatga -2457
aaaaaattgaaaaaaagtgctcaaaaatttcagaaatttccaatttccaaacaattttggagaactttcaaaaatttttccaactgaaattaaagctata -2357
G D H C F 147
ttctatcactaaatttttatacaagtcttaagagaaaatgatgaagtggctcattttgtagaatttcctaaaaaataatatcttcagGGCGATCACTGCTT -2257
N I S A I K P F L G W Q K Y S N V S A T L Y R S L A H M P W N Q S 180
CAACATTTCCGCAATCAAACCATTCCTGGGATGGCAAAAATATTCGAATGTATCAGCGACTCTATACCGTTCACTTGCACATATGCCATGGAATCAATCA -2157
SHP138
SHP146
D V D E N A A V L A Y N G Q L W N Q K L L H Q V K L N A 208
GATGTTGATGAGAATGCAGCTGTTCTTGCATATAATGGACAACTATGGAATCAAAAGCTTTTGCATCAAGTCAAATTGAATGCGgtaaagtattataagt -2057
I W R T L R A T R D A A F S V R E Q A G H T L 231
gttttgtccaaactatgatacagttcttcagATATGGAGAACACTTCGTGCCACTCGAGATGCCGCATTTTCAGTTCGTGAACAAGCCGGACACACTTTG -1957
K F S L E N A V A V D T R D R P I L A S R G I L A 256
AAATTCTCGTTGGAGAATGCTGTAGCTGTTGATACAAGAGATAGACCTATTCTTGCAAGTCGTGGAATTCTTGgtaagagtaacaacgactatttttaaa -1857
aaatatcttttttcgaaaaaattacgaacgaaaaaaaaactgtattatgtacccaaacgcgaaattttgcagttcttgcgcgctcttgttgataaaaaatat -1757
R F A Q 260
gtaaaaaattggaaaaactacgaaaagtcgataaaaattccgtaccaaccggaaaatgtttcattaatttctcttcctttttttcagCTCGTTTTGCTCAA -1657
E Y A G V F G D A S F V K N T L D L Q 279
GAGTACGCAGGAGTATTTGGTGATGCGTCATTTGTGAAGAATACATTAGATTTACAGgtaacaaccttatttcaacaattatttcaaatttctattaaaaa -1557
SHP139
A A A P L P L G F I L A A S F Q A K H L K G L G D R E V H I L 310
taattccagGCAGCTGCCCCTCTTCCACTCGGTTTCATTCTTGCCGCCTCATTCCAAGCGAAACATTTGAAAGGACTCGGAGATCGAGAAGTTCATATTT -1457
SHP140

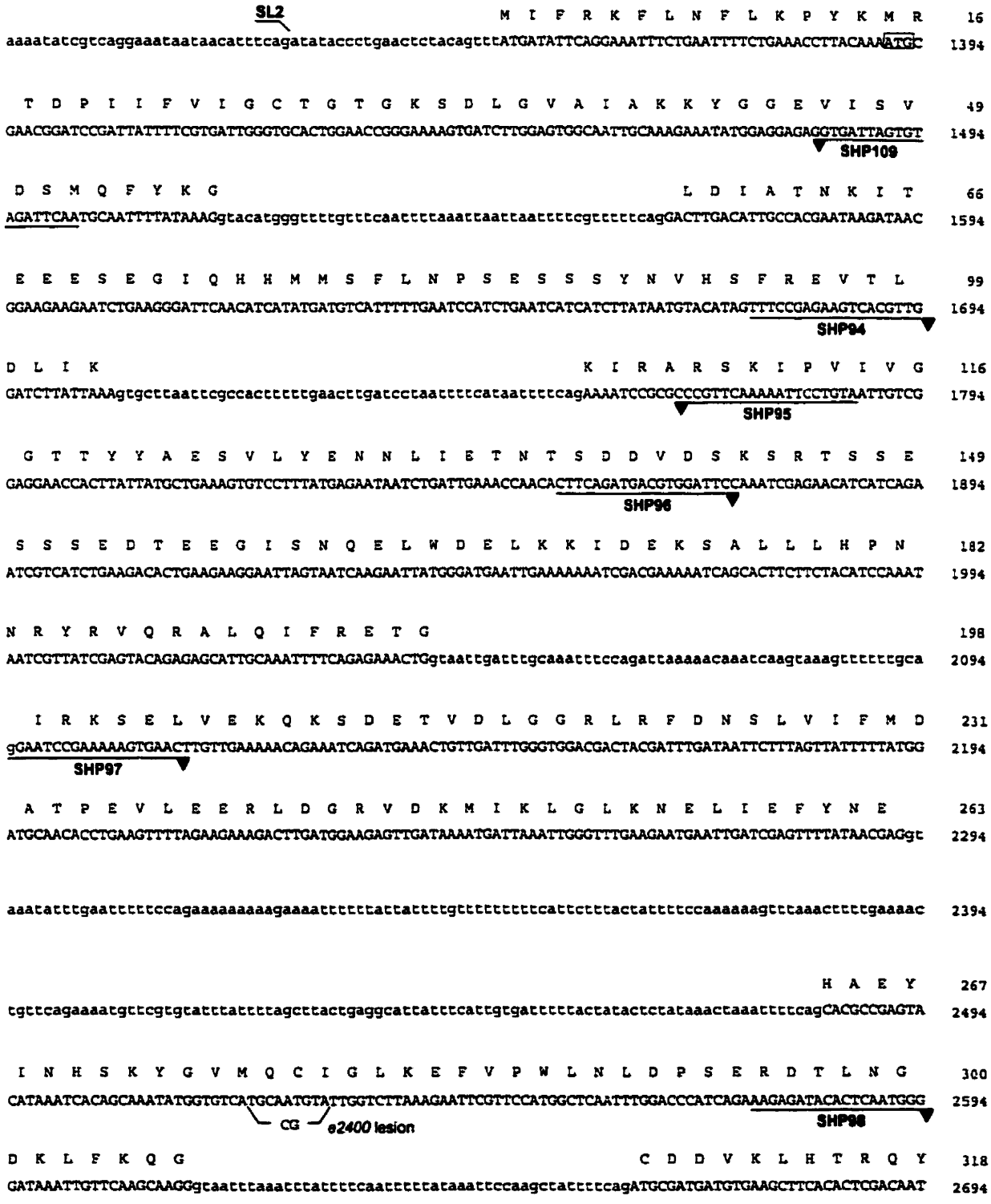
gop-3 continued...



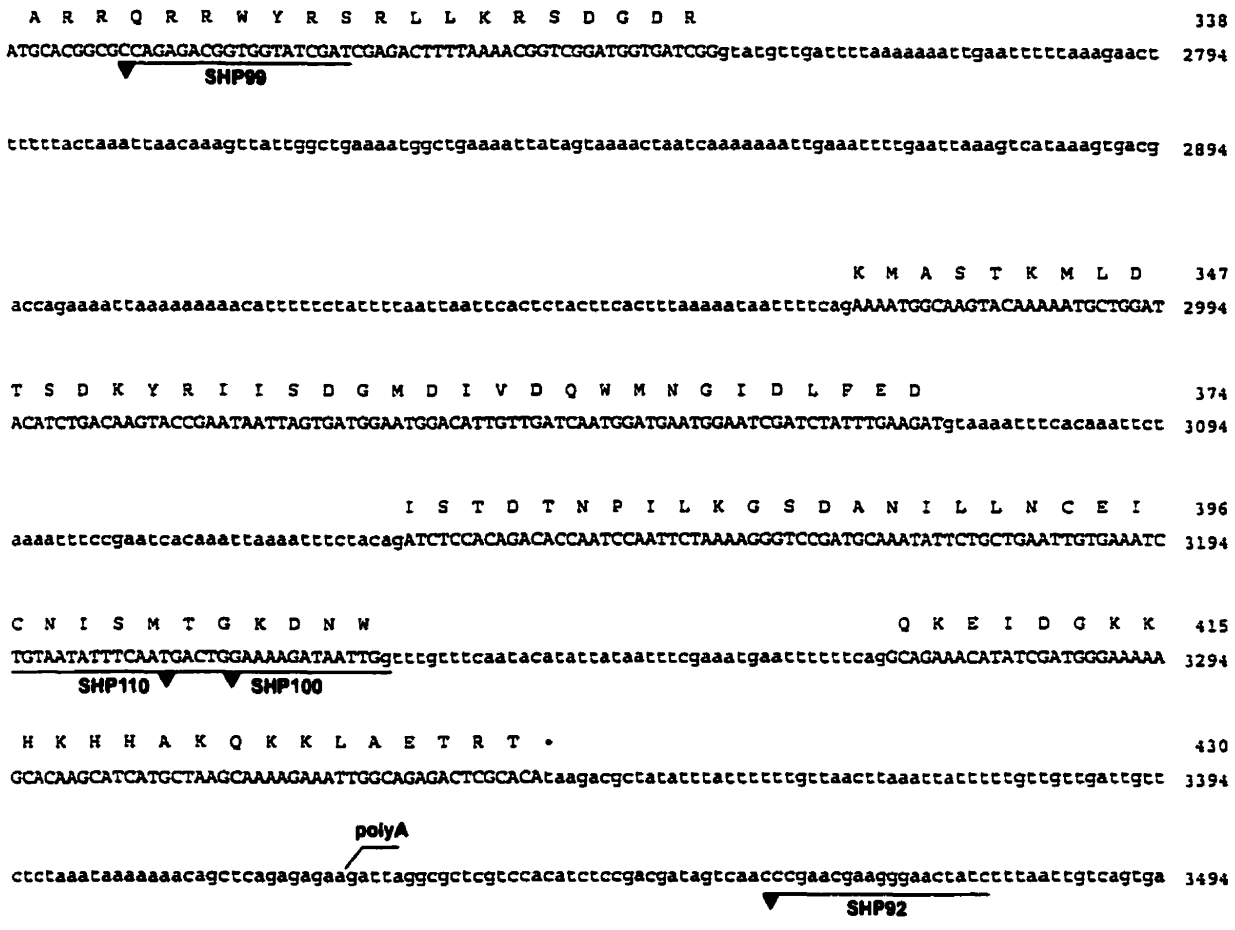
hap-1


SL2
SHP91
SHP118
SHP165
SHP129
SHP130
SHP119
SHP120
polyA
SHP93

gro-1


SL2
M I F R K F L N F L K P Y K M R 16
aaaatatcgtcaggaaataataacatttcagatataccctgaactctacagtttATGATATTCAGGAAATTTCTGAATTTTCTGAAACCTTACAAAATGC 1394
T D P I I F V I G C T G T G K S D L G V A I A K K Y G G E V I S V 49
GAACGGATCCGATTATTTTCGTGATTGGGTGCACTGGAACCGGGAAAAGTGATCTTGGAGTGGCAATTGCAAAGAAATATGGAGGAGAGGTGATTAGTGT 1494
SHP109
D S M Q F Y K G L D I A T N K I T 66
AGATTCAATGCAATTTTATAAAGgtacatgggttttgtttcaattttaaattaattaattttcgtttttcagGACTTGACATTGCCACGAATAAGATAAC 1594
E E E S E G I Q H H M M S F L N P S E S S S Y N V H S F R E V T L 99
GGAAGAAGAATCTGAAGGGATTCAACATCATATGATGTCATTTTTGAATCCATCTGAATCATCATCTTATAATGTACATAGTTTCCGAGAAGTCACGTTG 1694
SHP94
D L I K K I R A R S K I P V I V G 116
GATCTTATTAAAgtgcttaattcgccacttttgaacttgatcctaattttcataattttcagAAAATCCGCGCCCGTTCAAAAATTCCTGTAATTGTCG 1794
SHP95
G T T Y Y A E S V L Y E N N L I E T N T S D D V D S K S R T S S E 149
GAGGAACCACTTATTATGCTGAAAGTGTCCTTTATGAGAATAATCTGATTGAAACCAACACTTCAGATGACGTGGATTCCAAATCGAGAACATCATCAGA 1894
SHP96
S S S E D T E E G I S N Q E L W D E L K K I D E K S A L L L H P N 182
ATCGTCATCTGAAGACACTGAAGAAGGAATTAGTAATCAAGAATTATGGGATGAATTGAAAAAAATCGACGAAAAATCAGCACTTCTTCTACATCCAAAT 1994
N R Y R V Q R A L Q I F R E T G 198
AATCGTTATCGAGTACAGAGAGCATTGCAAATTTTCAGAGAAACTGgtaattgatttgcaaatttccagattaaaaacaaatcaagtaaagtttttgca 2094
I R K S E L V E K Q K S D E T V D L G G R L R F D N S L V I F M D 231
gGAATCCGAAAAAGTGAACTTGTTGAAAAACAGAAATCAGATGAAACTGTTGATTTGGGTGGACGACTACGATTTGATAATTCTTTAGTTATTTTTATGG 2194
SHP97
A T P E V L E E R L D G R V D K M I K L G L K N E L I E F Y N E 263
ATGCAACACCTGAAGTTTTAGAAGAAAGACTTGATGGAAGAGTTGATAAAATGATTAAATTGGGTTTGAAGAATGAATTGATCGAGTTTTATAACGAGgt 2294
aaatatttgaatttttccagaaaaaaaaagaaaatttttattattttgtttttttttcattctttactatttttccaaaaaagtttaaacttttgaaaac 2394
H A E Y 267
tgttcagaaaatgttcgtgtatttatttagcttactgaggcattatttcattgtgatttttactatactctataaactaaatttttcagCACGCCGAGTA 2494
I N H S K Y G V M Q C I G L K E F V P W L N L D P S E R D T L N G 300
CATAAATCACAGCAAATATGGTGTCATGCAATGTATTGGTCTTAAAGAATTCGTTCCATGGCTCAATTTGGACCCATCAGAAAGAGATACACTCAATGGG 2594
CG e2400 lesion
SHP98
D K L F K Q G C D D V K L H T R Q Y 318
GATAAATTGTTCAAGCAAGGgtaatttaaatttatttttcaatttttataaattccaagctatttttcagATGCGATGATGTGAAGCTTCACACTCGACAAT 2694

gro-1 continued...



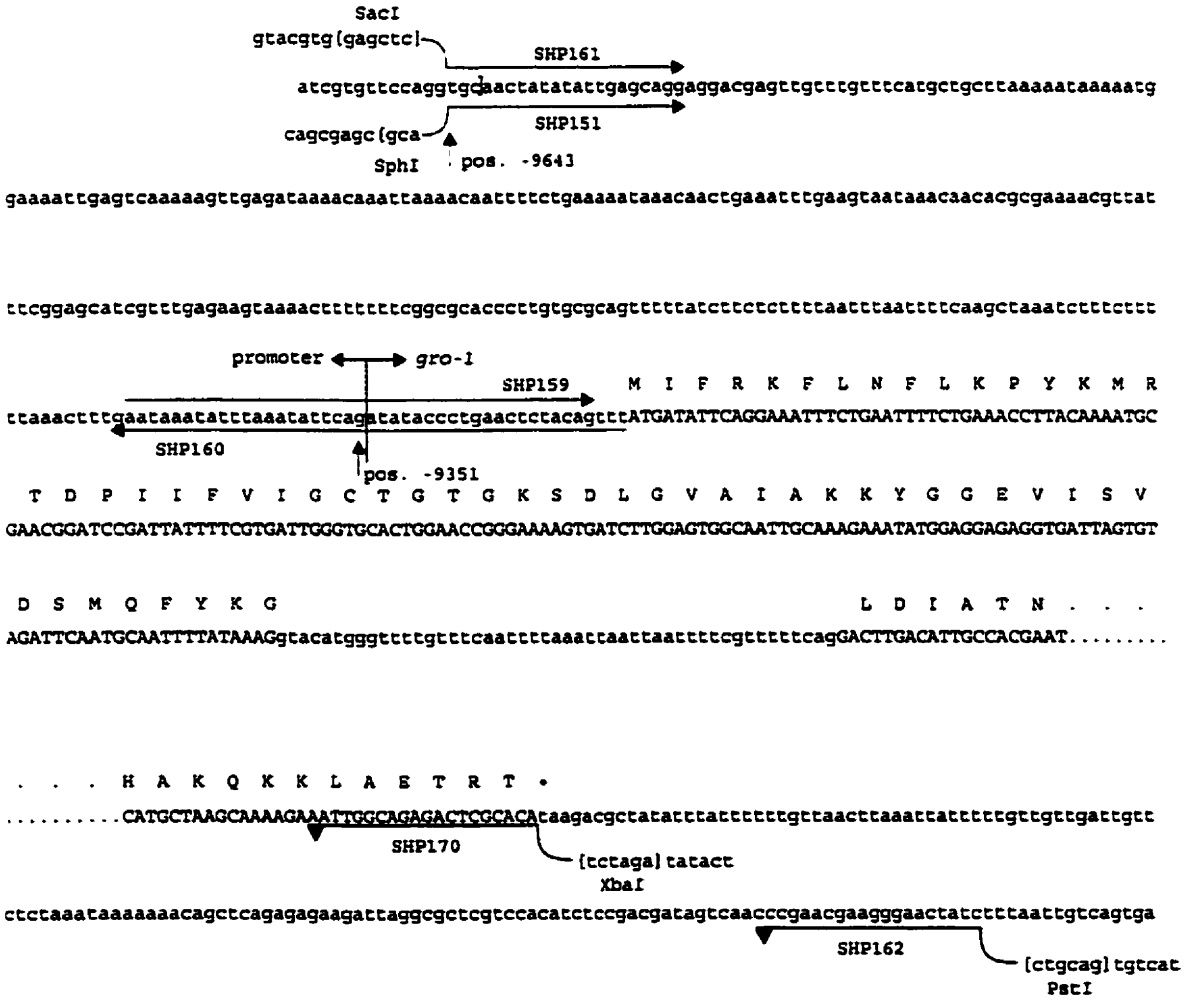
Appendix B: Human *gro-1* sequence

This is the sequence of the cDNA clone c-2ec05, the human *gro-1*. The clone does not contain the entire coding sequence, the 5' end is missing (see Figure 8 for protein alignment). Details of the isolation and sequencing of this clone can be found in the Materials and Methods section.

```
GCACGAGCAG TTCCTGTGGG CAGTGGGCTC AGGGGCCTGC AACGGACCCT    50
ACCTCTTGTA GTGATTCTCG GGGCCACGGG CACCGGCAAA TCCACGCTGG   100
CGTTGCAGCT AGGCCAGCGG CTCGGCGGTG AGATCGTCAG CGCTGACTCC   150
ATGCAGGTCT ATGAAGGCCT AGACATCATC ACCAACAAGG TTTCTGCCCA   200
AGAGCAGAGA ATCTGCCGGC ACCACATGAT CAGCTTTGTG GATCCTCTTG   250
TGACCAATTA CACAGTGGTG GACTTCAGAA ATAGAGCAAC TGCTCTGATT   300
GAAGATATAT TTGCCCGAGA CAAAATTCCT ATTGTTGTGG GAGGAACCAA   350
TTATTACATT GAATCTCTGC TCTGGAAAGT TCTTGTCAAT ACCAAGCCCC   400
AGGAGATGGG CACTGAGAAA GTGATTGACC GAAAAGTGGA GCTTGAAAAG   450
GAGGATGGTC TTGTACTTCA CAAACGCCTA AGCCAGGTGG ACCCAGAAAT   500
GGCTGCCAAG CTGCATCCAC ATGACAAACG CAAAGTGGCC AGGAGCTTGC   550
AAGTTTTTGA AGAAACAGGA ATCTCTCATA GTGAATTTCT CCATCGTCAA   600
CATACGGAAG AAGGTGGTGG TCCCCTTGGA GGTCCTCTGA AGTTCTCTAA   650
CCCTTGCATC CTTTGGCTTC ATGCTGACCA GGCAGTTCTA GATGAGCGCT   700
TGGATAAGAG GGTGGATGAC ATGCTTGCTG CTGGGCTCTT GGAGGAACTA   750
AGAGATTTTC ACAGACGCTA TAATCAGAAG AATGTTTCGG AAAATAGCCA   800
GGACTATCAA CATGGTATCT TCCAATCAAT TGGCTTCAAG GAATTTCACG   850
AGTACCTGAT CACTGAGGGA AAATGCACAC TGGAGACTAG TAACCAGCTT   900
CTAAAGAAAG GACCTGGTCC CATTGTCCCC CCTGTCTATG GCTTAGAGGT   950
ATCTGATGTC TCGAAGTGGG AGGAGTCTGT TCTTGAACCT GCTCTTGAAA  1000
TCGTGCAAAG TTTCATCCAG GGCCACAAGC CTACAGCCAC TCCAATAAAG  1050
ATGCCATACA ATGAAGCTGA GAACAAGAGA AGTTATCACC TGTGTGACCT  1100
CTGTGATCGA ATCATCATTG GGGATCGCGA ATGGGCAGCG CACATAAAAT  1150
CCAAATCCCA CTTGAACCAA CTGAAGAAAA GAAGAAGATT GGACTCAGAT  1200
GCTGTCAACA CCATAGAAAG TCAGAGTGTT TCCCCAGACT ATAACAAAGA  1250
ACCTAAAGAG AAGGGATCCC CAGGGCAGAA TGATCAAGAG CTGAAATGCA  1300
GCGTTTAA                                               1308
```

Appendix C: pMQ8

In order to simplify the *gro-1* operon, the construct pMQ8 was made (see Materials and Methods section). Illustrated here is the resulting sequence showing the fusion point where the hybrid primers (SHP159 ans SHP160) lie, and the ends where the flanking primers (SHP161 and SHP162) lie.



Appendix D: Towards a GRO-1 antibody

With *gro-1* cloned it became possible to raise an antibody specific to GRO-1. Antibodies are probably one of the most powerful tools in modern molecular biology, making possible many experiments designed to help understand the function(s) of the protein in question. The standard technique for raising antibodies involves the inoculation of a mammalian host (for example rabbit, mouse or donkey) with the protein of interest or a synthetic peptide corresponding to a short (usually around 12 aa) sequent of the protein of interest. The animal's immune system is then allowed to react to the foreign protein and produce antibodies in response. A number of bleeds are taken at regular intervals in order to test the progress of the immune reaction and boosts of the immunogen are administered if necessary. Once the protocol has run its course the animals are sacrificed and exsanguinated.

In the case of *gro-1*, the gene's coding sequence was cloned into the bacterial expression vector pET20b (Nonagon, http://www.novagen.com). This vector uses the T7 promoter to express the cloned gene in BL21 *E.Coli* cells when exposed to IPTG. The vector is also designed to tag the cloned gene with a hexa-His tag at its N-terminal. This His tag makes it possible to purify the expressed protein by passing it over a charged Ni^{2+} column. The bacterial expressed GRO-1 was isolated from inclusion bodies and dissolved in 6M urea and purified over the Ni column.

The protein was purified by SDS-PAGE and visualized by staining with comassie blue. Based of the coding sequence, GRO-1 was predicted to be 48 kDa, but the primary product of the bacterial expression was approximately 25 kDa. There were also a number of secondary products of varying lengths. Because these proteins were purified over the Ni column, they had to contain the His tag. pET20b places the tag at the N-terminal which rules out stop codon mutations in the cloned gene as an explanation for the truncated products. The only other explanations are that these are degradation products or that translation is initiating at an internal AUG. There are a number of internal ATG codons that could produce a protein of 25 kDa (see Appendix A, positions 2191, 2245 and 2520). In order to confirm this hypothesis, the N-terminus of the purified 25 kDa protein was sequenced by Sheldon Biotechnologies. The first 25 aa were sequenced and it was confirmed that this protein was produced from the internal initiation of translation at the ATG at position 2245. This protein therefore corresponded to 182 aa of GRO-1 (positions 248 to 430).

The urea was removed by dialysis and the precipitated protein sent to Commonwealth Biotechnologies (http://ww.cbi-biotech.com) where two rabbits were inoculated to raise the anti-GRO-1 antibody. The protocol was underway at the time of the writing of this thesis.

Appendix E: Towards GRO-1::GFP mislocalization

Mutational analysis by Gillman *et al.* (1991) was able to show that by differential initiation of translation, two different forms of Mod5p are produced: the longest contains an additional 11 aa at its N-terminus and is targeted to the mitochondria, while the shorter remains in the cytoplasm (see Introduction). Reporter construct experiments with the N-terminal of Mod5p have shown that it is the first 21 aa of Mod5p that are sufficient for mitochondrial targeting (Boguta *et al.*, 1994). The details of these experiments are given in the introduction.

It is very possible that GRO-1 is targeted by a similar mechanism. There is a second ATG 15 codons in from the actual start. In fact, this second ATG was predicted to be the start by GeneFinder, but sequencing of the 5' end of the cDNA revealed that the first ATG, at position 1349, was included in the message and therefore acted as the start codon (see Results section). GFP constructs where therefore designed which eliminated each of the start codons in turn in an attempt to manipulate the localization of GRO-1. For example, if the first start codon was eliminated, the longer form of the protein would not be produced, and based on the results in yeast, the IPP-transferase should be excluded from the mitochondria. Another construct eliminated the second start codon, this would in theory prevent the production of the shorter form and the IPP-transferase and therefore should only be localized to the mitochondria. These constructs have been designed and some of them constructed (see below). When these constructs are introduced into *gro-*

1(e2400), it is hoped that the nature of rescue is different than that produced by the normal protein (expressed by pMQ18, see Result section). This may enable the different *gro-1(e2400)* phenotypes to be linked to the specific localization of the IPP-transferase. For example, excluding it from the nucleus may help to address the question of an additional IPP-transferase function (see Discussion).

Site directed mutagenesis by bi-partite PCR was used to alter the *gro-1* sequence. Bi-partite PCR is explained in Figure 1. The difference is that instead of using the hybrid primers to join two separate pieces of DNA, two complementary "mutagenic" primers, containing the desired mutation, are used. The wild type gene is used as template and amplified in two parts, and because mutant primers are used the overlapping portions of the products from PCR#1 and PCR#2 contain the mutation. PCR#3 then re-establishes the intact gene, which now carries the desired mutation.

Three different constructs have been made that are designed to express only one of the two forms believed to be produced by alternative initiation. All three use the pMQ8 (see Materials and Methods section) as the template, the flanking primers SHP161 and SHP162, and a pair of specific mutagenic primers. The primers SHP161 and SHP162 have SacI and PstI, respectively, built into their 5' ends. This allows for easy cloning into pUC18. The sequences of the flanking and mutagenic primers are given in Appendix F. The positions of the flanking primers are illustrated in Appendix C.

The first attempts to express only the longer, mitochondrial form by changing the second ATG to GCG. The mutagenic primers used were SHP188 and SHP232. This construct has been designed but not built.

Two different approaches were used to try and express only the shorter of the two forms. The first construct, pMQ24, has the first ATG mutated to GCG. The mutagenic primers SHP187 and SHP231 were used along with SHP161 and SHP162. The vector backbone for this clone is pUC18 (Messing, 1983; Norrander *et al.*, 1983; Yanisch-Perron *et al.*, 1983). In order to make a mutant *gro-1::gfp* construct, the primers SHP151 and SHP170 were used to amplify the mutant gene from pMQ24. These primers have built into their 5' end the restriction enzyme sites SphI and XbaI, respectively, and allowed cloning into the Fire GFP vector pPD95.77 (Fire, http://ciw2.ciwemb.edu/pub/FireLabVectors). This mutant reporter construct is named pMQ25. The second approach was to place a stop codon before the second ATG so that any protein initiated at the first start would terminate. The mutagenic primers SHP358 and SHP359 were used along with the flanking primers SHP161 and SHP162 to make pMQ31. The vector backbone for this clone is pUC18 (Messing, 1983; Norrander *et al.*, 1983; Yanisch-Perron *et al.*, 1983). A GFP reporter was not made with this construct.

Appendix F: Sequences of all the primers.

Name	Orientation	Sequence (5'-3')
SHP92	reverse	GATAGTTCCCTTCGTTCGGG
SHP93	forward	TTTCTGGATTTTAACCTTCC
SHP94	forward	TTTCCGAGAAGTCACGTTGG
SHP95	reverse	TACAGGAATTTTTGAACGGG
SHP96	forward	CTTCAGATGACGTGGATTCC
SHP97	forward	GGAATCCGAAAAAGTGAACT
SHP98	forward	AAGAGATACACTCAATGGGG
SHP99	reverse	ATCGATACCACCGTCTCTGG
SHP100	reverse	CCAATTATCTTTTCCAGTCA
SHP110	forward	ACATTATAAAGTTACTGTCC
SHP118	forward	TTTTAGTTAAAGCATTGACC
SHP119	reverse	ACATCTTTATCCATTTCTCC
SHP120	forward	TGCAAAGGCTCTGGAACTCC
SHP130	reverse	CATCCAAAAGCAGTATCACC
SHP134	forward	TTAATTGGATGCAAGCACCCC
SHP135	reverse	ATTACTATACGAACATTTCC
SHP138	forward	TTGTAAAGGCGTTAGTTTGG
SHP140	forward	CGACGGGGAGAAGGTGACGG
SHP141	reverse	AAAACTTCTACCAACAATGG
SHP142	reverse	CGTAATCTCTCTCGATTAGC
SHP143	reverse	CCGTGGGATGGCTACTTGCC
SHP144	reverse	TGGATTTGTGGCACGAGCGG
SHP145	reverse	TTGATTGCCTCTCCTCGTCC
SHP146	reverse	ATCAACATCTGATTGATTCC
SHP151	forward	CAGCGAGCGCATGCAACTATATATTGAGCAGG
SHP159	forward	AATAAATATTTAAATATTCAGATATACCCTGAACTCTACAG
SHP160	reverse	AAACTGTAGAGTTCAGGGTATATCTGAATATTTAAATATTTATTC
SHP161	forward	GTACGTGGAGCTCTGCAACTATATATTGAGCAGG
SHP162	reverse	ATGACACTGCAGGATAGTTCCCTTCGTTCGGG
SHP163	forward	GTGTTGCATCAGTTCATTCC
SHP164	forward	GCTGTGCTAGAAGTCAGAGG

Appendix F continued...

Name	Orientation	Sequence (5'-3')
SHP165	reverse	GTTCTCCTTGGAATTCATCC
SHP170	reverse	AGTATATCTAGATGTGCGAGTCTCTGCCAATT
SHP171	reverse	AGTAATTGTACATTTAGTGG
SHP172	forward	ATTAACCTTACTTACTTACC
SHP173	forward	CTAAACTAAGTAATATAACC
SHP174	reverse	GTTGATTCTTTGAGCACTGG
SHP175	forward	AATTCGACCAATTACATTGG
SHP176	reverse	AACATAGTTGTTGAGGAAGG
SHP177	forward	AATTAATGGAGATTCTACGG
SHP178	forward	TCAGCATCTAGAAATGCAGG
SHP179	reverse	CGAATGTCAACATTCACTGG
SHP180		CTTAACCTGATGTGTACTCG
SHP181	forward	ATGAAGCTTTAGAGGATGCC
SHP182	forward	CGACGAATTTCTGGAGTCGG
SHP183	reverse	ACTGCATTATCCATTAATCC
SHP184	reverse	CACCCAAATAACATCTATCC
SHP185	forward	TTTAACCTCATCTTCGCTGG
SHP187		AATTTCCTGAATATCGCAAACTGTAGAGTTCA
SHP188		AATCGGATCCGTTCGCGCTTTGTAAGGTTTCAG
SHP190	forward	ATGTTCCGCAAGCTTGGTTC
SHP212	forward	GTG GACTTCAGAA ATAGAGC
SHP213	forward	CACTCAAATC TAGAACTCCC
SHP226	forward	GAGCTTGAAAAGGAGGATGG
SHP227	forward	CTGCTTTTAATAATAGAACC
SHP231		TGAACTCTACAGTTTGCGATATTCAGGAAATT
SHP232		CTGAAACCTTACAAAGCGCGAACGGATCCGATT
SHP239	forward	ATCCTTTGGCTTCATCGTGACCAGGC
SHP256	forward	GAGTCTGTTCTTGAACCTGC
SHP358		AATTTTCTGAAACCTTACTAAATGCGAACGGATCCG
SHP263	reverse	TCAATGTAAT AATTGGTTCC
SHP264	reverse	GCTCATCTAG AACTGCCTGG
SHP265	reverse	GGCCCTGGAT GAAACTTTGC

Appendix F continued...

Name	Orientation	Sequence (5'-3')
SHP266	reverse	ACTGGGAGAC AAACATACCC
SHP359		TCGGATCCGTTCGCATTTAG TAAGGTTTCAGAAAATTC
dm2 (SHP193)	forward	GACGGCCAGTGAATTCCCCT
R_t	NA	ATCGATGGTCGACGCATGCGGATCCAAAGCTTGAATTCGAGCTCTT TTTTTTTTTTTTTTTTT
R_o	reverse	ATCGATGGTCGACGCATGCGGATCC
R_l	reverse	GGATCCAAAGCTTGAATTCGAGCTC
SL1	forward	TTTAATTACCCAAGTTTGAG
SL2	forward	TTTTAACCCAGTTACTCAAG