Harnessing Caenorhabditis elegans secretion to produce recombinant secreted proteins from other nematodes

ZMARZLY, KATARZYNA,KAROLINA

Katarzyna Karolina Zmarzly

Harnessing *Caenorhabditis elegans* secretion to produce recombinant secreted proteins from other nematodes

Acanthocheilonema viteae excretory-secretory 62 product can modulate the immune system, and is a potential treatment for rheumatoid arthritis and other autoimmune diseases. It has been shown to inhibit arthritis in mouse and *in vitro* models. However, the only known way of producing functional ES-62 involves growing *A. viteae* worms in two hosts (ticks and gerbils), isolating adult worms from infected gerbils and allowing them to secrete ES-62 into liquid media. This method is time-consuming, expensive, difficult to scale up and does not yield sufficient amount of protein for extensive analysis. Standard recombinant expression systems are not sufficient for expressing recombinant ES- 62 in a correctly folded, active form, because they cannot generate the nematode-specific post-translational modifications required for the biological activity of ES-62. The nematode *Caenorhabditis elegans* provides an alternative expression system. *C. elegans* is a model organism, with well-understood genetics, that can be easily manipulated, has a fast reproduction rate and can be grown in large cultures. We hypothesised that understanding the biology and sequence of endogenous ES-62, as well as the biology of *C. elegans* secretion, will help us design the best way to use *C. elegans* as an alternative expression system for secreted recombinant ES-62.

In this study, we characterised the endogenous ES-62 genomic fragment and found that it had a complex structure with a large second intron. We found that *C. elegans* secretes a simple mixture of proteins into the media, in a relatively short period of time. The most abundant of those proteins was identified by MALDI-TOF as an aspartyl protease ASP-6. We used control elements of *asp-6* to direct expression of ES-62 genomic fragment into the excretory/secretory pathway of *C. elegans*. Despite strong expression of the transgene, we found that the recombinant protein expression in transgenic *C. elegans* was weaker than anticipated.

Harnessing *Caenorhabditis elegans* secretion to produce recombinant secreted proteins from other nematodes

Katarzyna Karolina Zmarzly

Supervisor

Dr David Weinkove

School of Biological and Biomedical Sciences,

Durham University

Masters of Research 2016

List of Contents

List of figures **4**

Declaration **7**

Acknowledgements **9**

1. Introduction **10**

1.1 The hygiene hypothesis 10

1.2 Rheumatoid Arthritis 10

1.2.1 Current treatment 12

1.3 Filarial nematodes 13

1.3.1 Acanthocheilonema viteae 14

1.4 ES-62 anti-inflammatory mechanism 15

1.5 Production of parasite- derived ES-62 16

1.6 *Acanthocheilonema viteae* ES-62 17

1.6.1 ES-62 cDNA 17

1.6.2 ES-62 protein structure 18

1.6.3 ES-62 glycosylation 22

1.6.4 ES-62 PC modification 23

1.6.5 ES-62 homology to proteins from other parasites 24

1.7 Recombinant ES-62 27

1.8 *C. elegans* 29

1.9 Excretory- Secretory system of nematodes 30

1.10 *C. elegans* excretory-secretory protein 33

1.11 Challenges 35

1.12 *C. elegans* gene structure and regulation 36

1.13 Aims and hypothesis 38

2. Materials and methods **39**

2.1 Media and Buffer Preparation 39

2.2 Preparation of bacterial culture 40

2.3 Maintenance and synchronisation of *C. elegans* cultures 40

2.4 Preparation of *C. elegans* plate culture for protein secretion 41

2.4.1 Harvesting the worms and secreted media after protein secretion (for plate culture *C. elegans*) 41
2.5 Preparation of *C. elegans* liquid culture for protein secretion 42
2.5.1 Harvesting the worms and secreted media after protein secretion (for liquid culture *C. elegans*) 42
2.6 Extraction of genomic DNA from *C. elegans* 43
2.7 Extraction of genomic DNA from *A. viteae* 43
2.8 PCR 44
2.8.1 Nested PCR amplification of the genomic ES-62 fragment 44
2.9 Amplification of *asp-6::* genomic ES-62 DNA fusions 44
2.10 Sequencing PCR products 45
2.11 Microinjection of *C. elegans* 46
2.11.1 List of *C. elegans* transgenic strains used 47
2.12 Single worm lysis and diagnostic PCR 48
2.13 *C. elegans* pellets lysis after secretion 48
2.14 Concentration of *C. elegans* lysate 49
2.15 Concentration of *C. elegans* media 49
2.16 SDS-PAGE 50
2.17 Silver staining 50
2.18 Western blot analysis 51
2.18.1 Antibodies 51
2.19 Preparation of samples for silver stain (TCA precipitation followed by acetone precipitation) 52
2.20 Buffer exchange for media samples before MALDI-TOF MS analysis 52
2.21 Stage Tip fractionation of trypsin- digested samples 52
3. Results **54**
3.1 *C. elegans* secrete proteins into the media over a relatively short time period 54
3.2 Characterisation of *A. viteae* ES-62 genomic DNA fragment 58
3.2.1 Extraction of genomic DNA 58
3.2.2 ES-62 coding genomic DNA fragment 60
3.2.3 Comparison of the published cDNA sequence (Harnett) to *A. viteae* genome (sequenced by Blaxter lab) 60
3.2.4 Amplification of the ES-62 genomic fragment that aligns to cDNA 63

3.3 Making transgenic *C. elegans* 70
3.3.1 Preparation of DNA fusion products for microinjections 70
3.3.2 Screening injected worms for expression of the transgene 74
3.4 Expression of recombinant ES-62 protein in transgenic *C. elegans* 76
3.4.1 Identification of recombinant ES-62 in transgenic *C. elegans* lysates (Western Blot analysis of small scale plate cultures) 76
3.4.2 Identification of a recombinant ES-62 breakdown product in transgenic *C. elegans* media (large scale liquid culture) 79
3.4.3 Identification of the recombinant ES-62 in transgenic *C. elegans* lysates (Silver stain analysis of large scale liquid cultures) 85
3.4.4 MALDI-TOF MS analysis of transgenic *C. elegans* secreted media (large scale liquid culture) and parasite-derived ES-62 90
3.4.5 Identification of the recombinant ES-62 protein in secreted media from integrated 6B transgenic lines (small scale plate culture) 92
3.4.6 Comparing 6B and 10A genomic ES-62 transgenic *C. elegans* (large scale liquid culture) 95
4. Discussion **101**
4.1 *C. elegans* secrete proteins into liquid media 101
4.2 The structure of the ES-62 genomic DNA fragment 103
4.3 Expression of the recombinant protein 104
4.3.1 Expression of recombinant ES-62 in transgenic *C. elegans* lysates 105
4.3.2 Expression of recombinant ES-62 in transgenic *C. elegans* secreted media 108
4.3.3 Differences in recombinant protein expression in ES-62 genomic lines with identical transgenes 109
4.3.4 Regulation of folding and detection of recombinant protein 110
Conclusion and future work **113**
Appendix **115**
Bibliography **134**

List of figures

Figure 1.1. *A. viteae* lifecycle……..**14**

Figure 1.2. Diagram of the current method used for ES-62 production in the laboratory……....**17**

Figure 1.3. Diagram of the ES-62 functional residues.....**21**

Figure 1.4. Distribution of blast alignments to the *A. viteae* ES-62 cDNA sequence query…….....**25**

Figure 1.5. Graphical summary of conserved domains on ES-62 (protein BLAST)……**25**

Figure 1.6. Adult excretory- secretory system in *C. elegans*……..**32**

Figure 1.7. Expression of GFP labelled ASP-6 protein in adult transgenic *C. elegans*……**34**

Figure 3.1. Silver stain of *C. elegans* proteins secreted into the M9 medium over 4 hours……..**55**

Figure 3.2. Silver stain of *C. elegans* proteins secreted into the M9 medium over 4 hours……..**56**

Figure 3.3. Electrophoresis of dilutions of genomic DNA extracted from *A. viteae* and *C. elegans* worms……**59**

Figure 3.4. Schematic alignment of ES-62 cDNA sequence to *A. viteae* genome showing the distribution of the sequence between different scaffolds……**62**

Figure 3.5. Electrophoresis picture of A) outer primers PCR and B) inner/ nested primers PCR to amplify ES-62 fragment from scaffold 00478……**64**

Figure 3.6. Alignment of ES-62 cDNA fragments to the *A. viteae* genomic DNA……..**64**

Figure 3.7. Electrophoresis of the PCR result between olive green and grey fragments (scaffold 05826 and scaffold 00478)……**66**

Figure 3.8. Electrophoresis of the result of a PCR between light green and olive fragments……**66**

Figure 3.9. Summary of the PCR reactions performed to characterise the 5'end of the ES-62 genomic fragment……**68**

Figure 3.10. Schematic diagram of the full ES-62 genomic DNA……**69**

Figure 3.11. Schematic diagram of the *asp-6::*ES-62 fusion constructs……..**71**

Figure 3.12. Electrophoresis pictures of the fusion reaction between *asp-6* and genomic ES-62 fragment…….**73**

Figure 3.13. Electrophoresis of single-worm PCR to screen for transgenic lines……**75**

Figure 3.14. Western blot of concentrated worm lysates after isolation from plate culture……..**78**

Figure 3.15. Western blot of worm secreted media after isolation from liquid culture (rabbit α ES-62 antibody)……**80**

Figure 3.16. Western blot of worm lysates after isolation from liquid culture (rabbit α ES-62 antibody)……**81**

Figure 3.17. Western blot of worm secreted media and worm lysates after isolation from liquid culture (rabbit α de-glycosylated ES-62 antibody)…….**83**

Figure 3.18. Western blots of un-concentrated worm secreted media 45 days after isolation from liquid culture……..**84**

Figure 3.19. Silver stain of worm secreted media after isolation from liquid culture genomic ES-62 vs ES-62 cDNA transgenic lines……**87**

Figure 3.20. Silver stain of worm secreted media after isolation from liquid culture genomic ES-62 transgenic line……**88**

Figure 3.21. Silver stain of worm pellets, worm lysates and concentrated worm lysates after isolation from liquid culture……..**89**

Figure 3.23. *A. viteae* ES-62 sequence coverage by peptides identified by MS…….**91**

Figure 3.24. Western blot analysis of media from integrated 6B genomic ES-62 lines (small scale plate culture, rabbit α ES-62 antibody)……..**93**

Figure 3.25. Western blot analysis of media from integrated transgenic line 9* under non-denaturing conditions (small scale plate culture, rabbit α ES-62 antibody)……..**94**

Figure 3.26. Western blot of *C. elegans* secreted media immediately after isolation from liquid culture (rabbit α ES-62 antibody)……...**96**

Figure 3.27. Western blot of worm secreted media two weeks after isolation from liquid culture (rabbit αES-62 antibody)……**98**

Figure 3.28. Silver stain of worm secreted media after isolation from liquid culture…….**100**

Figure 4.1. Expression pattern of *asp-6::*ES-62cDNA::GFP fusion inside the intestinal cell of adult *C. elegans*…….**107**

List of figures in the Appendix

Figure 1. Alignment of *A. viteae es-62, Loa loa* leucyl aminopeptidase and *Brugia malayi* leucyl aminopeptidase protein coding DNA sequence……**115**

Figure 2. Alignment of *A. viteae es-62, Loa loa* leucyl aminopeptidase and *Brugia malayi* leucyl aminopeptidase protein sequence …….....**117**

Table 1. List of primers used for reactions mentioned in the Results (Section 3)…….....**118**

Table 2. List of peptides and scores after MALDI-TOF MS analysis of in- gel trypsin digestion of ES-62.…….....**119**

Table 3. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from **control** worms grown in liquid culture.……....**120**

Table 4. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from **6B ES-62 genomic DNA transgenic** worms grown in liquid culture.…….....**122**

Table 5. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from **IIL ES-62 cDNA transgenic** worms grown in liquid culture.…….....**124**

Table 6. MALDI-TOF MS results for 5 fractions from 6B asp-6 ES-62 genomic fragment transgenic line secreted media after liquid culture……....**126**

Declaration

I, Katarzyna Karolina Zmarzly, would like to declare that this thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

Acknowledgements

This work was partially funded by the NC3Rs organisation.

I would like to take this opportunity to thank my supervisor Dr David Weinkove, who supported me throughout my research with his knowledge, advice and encouragement.

I would like to offer my gratitude to Dr Sushmita Maitra for her unlimited support and endless knowledge, which she was kind enough to share with me. Thank you for your invaluable contribution to the 'ES-62 project', useful discussions, and above all, thank you for being a great friend.

I highly appreciate our collaborators- Prof William Harnett and his lab group, especially Dr Kara Bell for performing initial screening of transgenic lines and providing *A. viteae* worms, ES-62 protein and antibodies. I would like to thank Georgios Koutsovoulos from Blaxter lab for performing the re-assembly of *A. viteae* genome.

I thank Kyle Main for his contribution to this work during his undergraduate research project.

I would like to thank Miss Joanne Robson and Dr Adrian Brown for their help with analysing samples by Mass Spectrometry and Mrs Rebecca Manning for performing DNA sequencing.

I appreciate the wonderful help and advice from Claire Maynard, Dr Boo Virk, Ms Adelaide Raimundo, Razan Bakheet and the rest of the worm lab for their encouragement and support.

Finally, I would like to thank the worms, for being worms.

1. Introduction

1.1 The hygiene hypothesis

The hygiene hypothesis proposes that a lack of pathogens in areas with improved sanitation has led to an untrained immune system, and therefore contributed to an increase in autoimmune diseases (Strachan 1989). Epidemiological studies show that in developing countries, where the parasitic infections are endemic, the prevalence of autoimmunity and allergies remains significantly lower than in the industrialised world (Fleming & Cook 2006; Yazdanbakhsh et al., 2001; Rook 2012; Okada et al., 2010; Panda et al., 2013). This lead to a hypothesis that helminth-derived immunomodulatory excretory/secretory (ES) products can protect against the development of autoimmunity (Harnett M et al., 2008).

It has been observed in animal experiments that there is a negative correlation between helminth infection and autoimmune diseases and allergies (Weinstock and Elliot 2014). Parasitic helminths induce a strong Th2 (T helper Type 2) response in mammalian hosts. One consequence of this is the reciprocal attenuation of the Th1/Th17 branch of the immune response, thus having ameliorating effects on autoimmune diseases such as multiple sclerosis (MS), Crohn's disease, Type1 diabetes and rheumatoid arthritis which are mainly Th1/Th17 mediated (Okada et al., 2010).

1.2 Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic autoimmune disease affecting between 0.5% and 1% of the population in the developed world (Uhlig et al 2014; Vivar et al 2014). It is characterised by joint inflammation which leads to swelling and erosion of the cartilage and bone. If left untreated, RA eventually results in severe disability, pain, reduced health-related quality of life and premature mortality (Vivar et al 2014).

Rheumatoid arthritis involves a complex interplay among genotype, environmental triggers, and chance. The disease onsets with increased expression of adhesion molecules and chemokines in synovial microvessels. This attracts leukocytes, which accumulate and later proliferate in the joint. Synovial tissue biopsy experiments indicate that RA synovial membrane contains a variety of innate effector cells, including macrophages, mast cells, and natural killer cells. Macrophages act through the release of cytokines (e.g., tumor necrosis factor-α (TNF-α) and interleukins-1, 6, 12, 15, 18, and 23), reactive oxygen intermediates, nitrogen intermediates, the production of prostanoids and matrix-degrading enzymes, phagocytosis, and antigen presentation. All of the above make macrophages central effectors of the inflammation in the synovial membrane (McInnes, 2011; W & M Harnett, 2006; Tak & Bresnihan, 2000).

Pathogenesis of rheumatoid arthritis is mainly associated with TNF-α and interleukin-6 (Il-6), which has been confirmed by successful therapeutic blockade of the membrane and soluble TNF-α and the interleukin-6 receptors in RA patients. TNF-α acts through activation of cytokine and chemokine expression, expression of endothelial-cell adhesion molecules, protection of synovial fibroblasts, promotion of angiogenesis, suppression of regulatory T cells, and induction of pain. Interleukin-6 drives local leukocyte activation and autoantibody production but mediates systemic effects that promote acute-phase responses, anaemia, cognitive dysfunction, and lipid- metabolism deregulation (McInnes, 2011; Tak & Bresnihan, 2000). The suppression of differentiation and blocking the activity of regulatory T cells; shifts T-cell homeostasis toward inflammation (McInnes, 2011).

Prolonged, increased inflammation leads to a rapid bone erosion and affects 80% of patients within 1 year of diagnosis (McInnes, 2011). In the joint, macrophages stimulate osteoclast differentiation and invasion of the periosteal surface adjacent to articular cartilage. TNF-α and IL- 6 amplify osteoclast differentiation and activation. Osteoclasts are cells that degrade mineralized cartilage and bone with their acidic, enzymatic contents. Destruction of the bone and cartilage leads to deep resorption pits, which are filled by inflammatory tissue (Smolen et al., 2003, 2007). Therapeutic inhibition of TNF-

α and IL-6 has been used to slow down erosion in RA-affected joints (McInnes 2011; W & M Harnett, 2006; Tak & Bresnihan 2000).

1.2.1 Current treatment

The current treatment for RA involves initiation with the disease-modifying antirheumatic drugs (DMARDs) in almost all patients (Vivar et al 2014). DMARDs include cyclosporin, cyclophosphamide, gold injections, hydroxychloroquine, leflunomide, methotrexate, mycophenolate, and sulfasalazine. They act by dampening the immune system. Available DMARDs are slow-acting and can only temporarily retard the progression of the disease (Uhlig et al 2014; Vivar et al 2014). Many patients still experience premature work disability and co-morbidities. For societies, the economic burden of RA is high regarding direct and indirect costs, including modern drug treatment.

TNF-α blockers were the first biological agents approved for the treatment of rheumatoid arthritis; TNF-α blockade has become a central strategy of targeted anti-inflammatory therapy in the disease. The cytokine inhibitors that target TNF-α include adalimumab and golimumab (human monoclonal antibodies), certolizumab pegol (a pegylated humanized Fab' fragment of an anti- TNF-α monoclonal antibody), etanercept (a TNF-α receptor-Fc fusion), and infliximab (a chimeric monoclonal antibody) (Clements, 2011).

The second major advance in cytokine blockade in RA is tocilizumab, which targets the interleukin-6 receptor. It is a humanized monoclonal antibody, which has profound effects on systemic features, acute phase response, and synovitis (Clements, 2011).

Despite its role in RA, TNF is essential for the effective immune system; it also controls tumour growth. Thus, anti-TNF drugs, as well as most available DMARDs that interfere with the production of TNF-α, are associated with increased risk of acute viral and bacterial infections as well as dose-dependent increased risk of malignancies (Bongartz et al 2006). Patients treated with anti-TNF drugs

are, also, at higher risk of tuberculosis reactivation as well as herpes zoster (caused by *Varicella zoster* virus). Treatment with anti- TNF drugs can also increase the risk of melanoma in RA patients (Vivar et al 2014). There is a high demand for alternative therapies that will leave the immune system functional to fight infections. Filarial nematodes can modulate the immune system of the host leaving the defence against other infections intact (Harnett M et al 2008). The study of parasitic ES products provides new drug candidates for RA and other autoimmune diseases.

1.3 Filarial nematodes

Filarial nematodes are parasites of vertebrates, which require an arthropod vector host for the early life stages of the worm. Adult worms live in subcutaneous tissues (e.g. *Onchocerca volvulus*) or the lymphatic system (e.g. *Brugia malayi*, *Wuchereria bancrofti*) (Roberts & Janovy 2008). Most filarial nematodes can survive in the host for more than a decade without any symptoms for the infected individual (Subramanian et al., 2004). Some species, however, can cause severe medical conditions such as elephantitis, chronic skin lesions and eye damage that can lead to blindness (Harnett W et al., 2004). The parasites survival is promoted by their ability to modulate the host immune system. This immunomodulation consists of the reduction of the T helper (Th)1 associated interferon (INF)-γ and an increase in the Th2 associated anti-inflammatory interleukin (Il)-10 (Harnett W et al., 2004). ES products released by filarial nematodes have been shown to subvert the immune system of the host and have helped in maintaining infection and parasite survival (Harnett W & Parkhouse 1995). One of those products is ES-62, secreted by the rodent filarial nematode *Acanthocheilonema viteae* (Harnett W et al., 1989).

1.3.1 Acanthocheilonema viteae

Filarial parasites require two hosts to complete their lifecycle. Microfilariae of *A. viteae* are ingested by an arthropod vector (in the lab, tick *Ornithodorus moubata*), where they develop into infective larvae, which then migrate to the mouthparts of the tick. While feeding, the vector host infects the target host (in the lab gerbil *Meriones libycus*) with the larvae, which then develop into adult worms and produce progeny (Lucius & Textor 1995). Gerbil's blood, infected with a new generation of microfilariae, can be ingested again by the tick and the cycle repeats (Figure 1.1). Only L4 and adult worms live under the gerbil's skin; microfilariae are released into the bloodstream (Goodridge et al., 2005).

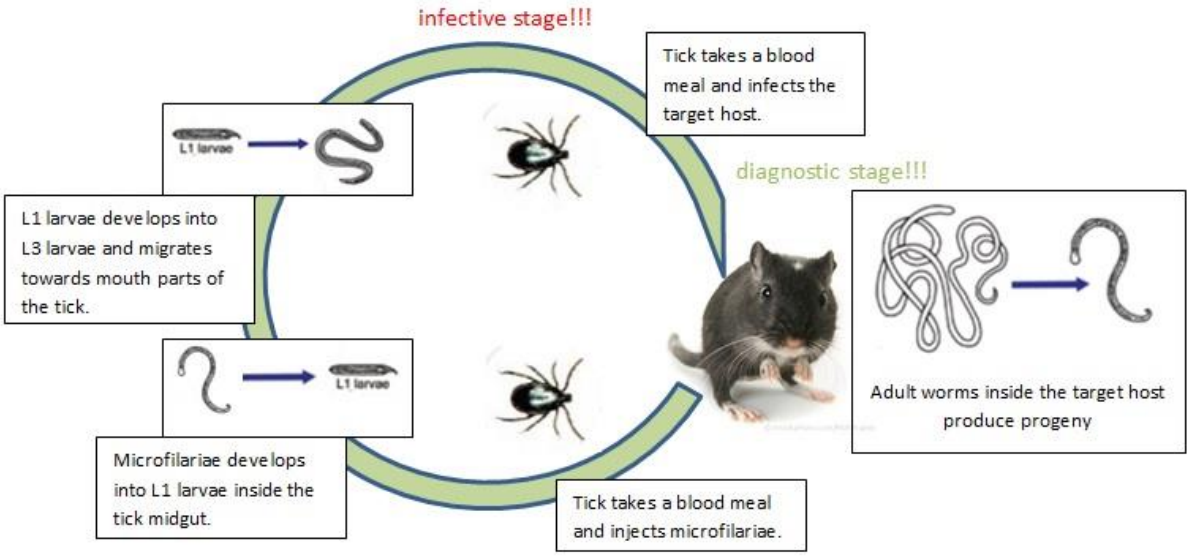


Figure 1.1. *A. viteae* lifecycle.

Most filarial infections result in a commensal relationship between the parasite and the host, where the parasite can survive for a long time, and the host is asymptomatic and appears immunologically tolerant to the parasite in the bloodstream (Goodridge et al., 2005). *A. viteae* survival inside the host has been attributed to the immunomodulatory activity of ES-62.

1.4 ES-62 anti-inflammatory mechanism

ES-62 was shown to inhibit the production of TNF-α and IL-6 cytokines in mice developing collagen-induced arthritis (CIA). ES-62 also suppressed the inflammation and the erosion of cartilage and bone, without inducing a compensatory increase in other types of immune response. ES-62 was able to suppress the CIA progression even when administered after the onset of the disease (McInnes et al 2003). The therapeutic potential of the protein was further investigated in human samples, through observations of primary cultures from RA joint- surrounding fluid and membrane, in the presence and absence of ES-62. In the presence of the protein, there was a significant suppression of lipopolysaccharide (LPS)- induced TNF-α and IL- 6 (Harnett W & M, 2006). This indicates that ES-62 can modify critical pro- inflammatory pathways in disease- relevant tissues *ex vivo*.

ES-62 has been shown to reduce Ag-driven B and T lymphocyte proliferation. It also modulates dendritic cell (DC) maturation to preferentially evoke Th2 responses. Moreover, ES-62 induces spleen cells and B1 B cells to produce IL-10 – an anti-inflammatory cytokine.

The mechanism by which ES-62 inhibits TNF-α production and the complete mechanism of action has yet to be fully defined (Rzepecka et al., 2014; Coltherd et al., 2016). The suppression of TNF-α, however, has been associated with inhibition of p38 mitogen-activated protein kinase (unpublished observations) that is required for such cytokine induction (Feng et al., 1999).

ES-62 has previously been shown to polarize the murine dendritic cells (DC) maturation in transgenic mice to induce subsequent Th2 responses. Therefore, ES-62 may modulate the cytokines through

altered DC function. (Whelan et al., 2000). Since DC in various maturation states are detected in RA synovia (Thomas et al., 1999; Page et al., 2002), this provides a potential antiarthritic action for ES-62 in altering local synovial DC maturation, autoantigen presentation, and Th1 functional polarization along with downstream effects on cytokine-releasing cells, including macrophages.

Due to the role ES-62 plays in the parasite's survival, it is an ideal target for new anthelmintics. Moreover, ES-62 opens a field for novel therapies for autoimmune diseases, especially rheumatoid arthritis. While the results are encouraging, a lot is unknown about the mechanism of action of ES-62, also, it is difficult to make large quantities of ES-62.

1.5 Production of parasite- derived ES-62

ES-62 is produced by L4 larvae and adult worms in the serum of an infected target host. So far, the only known way of producing functional ES-62 in the laboratory is to derive the protein from *A. viteae* adult females. The whole process starts with ticks that have to be infected with microfilaria. Infected ticks are then placed on gerbils to take blood meals. Infected gerbils are grown until parasites reach adulthood. *A. viteae* female worms live subcutaneously. Approximately ten adults can be isolated from one gerbil. Fifty female worms are cultured in 0.5L of glucose-rich liquid media, where they secrete ES-62. Liquid media with ES-62 is then separated from the parasites and subjected to purification and concentration steps. Approximately 6L of liquid media supernatant is needed to achieve 0.25 g of ES-62 (Figure 1.2) (Hewitson et al., 2009). The current process involves the use of a large number of animals and is time and labour intensive with not much yield at the end of the process (Harnett unpublished). To date, no recombinant active form exists.

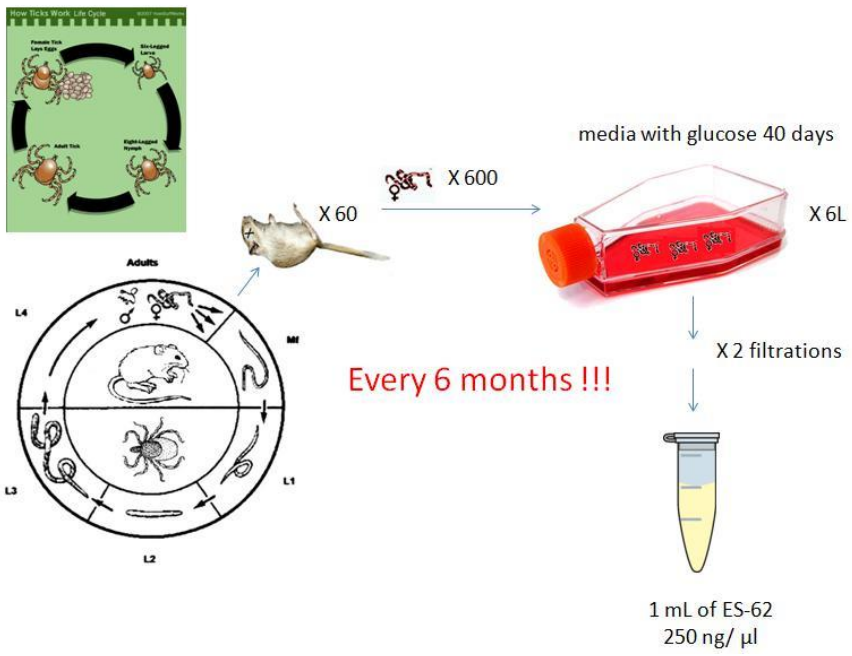


Figure 1.2. Diagram of the current method used for ES-62 production in the laboratory.

1.6 *Acanthocheilonema viteae* ES-62

ES-62 accounts for over 90% of the total protein secretions in adult-stage *A. viteae* worms (Harnett W et al., 1989), and can be detected in the serum of infected gerbils (Harnett W et al., 1989; Stepek et al., 2002). The ES-62 gene is transcribed throughout the life cycle of the *A. viteae* worm. The mRNA levels are, however, much higher in the adult worms than in the L3 larvae (~ 5% of the adult levels), or in the microfilariae (< 0.2% of the adult levels) (Stepek et al., 2004; Harnett W et al., 2005).

1.6.1 ES-62 cDNA

The ES-62 amino acid sequence was derived from the complementary DNA (cDNA). cDNA is a DNA copy synthesised from messenger RNA (mRNA) using reverse transcriptase. cDNA provides nucleotide sequence information of the coding region of the gene as well as any 5' or 3' sequences that are also transcribed and spliced, but not translated (Brown, 2002). An *A. viteae* cDNA expression

library was constructed as described by Adam et al (1996) and screened using a rabbit polyclonal serum raised against purified ES-62 (Harnett W et al., 1999). Three complete cDNAs of size ~ 1.5-1.6 kbp were isolated and sequenced. The sequences were compared to the N-terminal amino acid sequencing of purified ES-62 (Harnett W et al., 1999).

The 5' end of ES-62 mRNA was analysed by 5' rapid amplification of cDNA ends (RACE), which revealed that it codes for a signal sequence (Harnett W et al., 1999). Signal sequences are N-terminal extensions of newly synthesised secretory and membrane proteins. They direct the insertion of proteins into the membrane of the endoplasmic reticulum and are usually cleaved off by signal peptidase (Kapp et al., 2000). As ES-62 is a secreted protein (Harnett W et al., 1989), it is expected to lack the signal sequence. Protein sequencing will only reveal the N- terminal amino acids of the mature protein (after the signal sequence is cleaved off). RACE allows fast amplification of full-length cDNA when the sequence is only partially known, starting with mRNA, primers inside the known sequence and unspecific anchors (Yeku & Frochman 2011). The results suggested that ES-62 is synthesised as a preprotein and contains a signal sequence of 18 amino acids (Harnett et al., 1999).

1.6.2 ES-62 protein structure

The secreted protein consists of 472 amino acids with a predicted molecular mass of 52.8 kDa, and an isoelectric point of 5.96. (Harnett, W et al., 1999). On a denaturing SDS-PAGE gel, the protein runs at 62 kDa. The 9.2 kDa difference between the predicted and the experimentally determined molecular weight is likely to account for post-translational modifications.

The secondary structure of ES-62, predicted by the JPRED server based on the sequence information is 26% α-helix, 21% β-sheet and 53% other. The secondary structure of parasite- derived ES-62, determined experimentally by circular dichroism (CD) spectroscopy is 40% α-helix, 15% β-sheet,

and 15% turn. The difference in the prediction is again likely due to post-translational modifications and secondary structure elements not seen in the monomeric state (Ackerman et al., 2003).

The tertiary structure of the protein determined with DRAGON prediction software is presented in Figure 1.3., which was obtained from Goodridge et al., 2005. Sedimentation equilibrium experiment revealed that ES-62 is a tightly bound tetramer formed from dimers (Harnett W et al., 1993), which is a preferred structure of many aminopeptidases (Taylor 1993; Acosta et al., 1998).

The primary amino acid sequence of ES-62 contains homology to the M28 family of aminopeptidases as well as to glutamate carboxypeptidases (Ackerman et al., 2003). ES-62 was extracted from *A. viteae* culture medium by fast protein liquid chromatography. The purified product was able to hydrolyse Leu-AMC, confirming that the complex contained aminopeptidase activity (Harnett et al., 1989). Further confirmation was shown when the activity of ES-62 was inhibited by amastatin and EDTA. Amastatin is a competitive and reversible aminopeptidase inhibitor. EDTA is a known chelator and inhibitor of metallopeptidases. Bestatin is a transition-state analog of zinc, which makes it an inhibitor of zinc metalloenzymes. ES-62 activity was not inhibited by bestatin (Harnett et al., 1989). A low resolution structure analysis suggests that the active site cation is magnesium (Ackerman et al., 2003). This evidence suggests that ES-62 is an aminopeptidase with a strong Mg^{2+} coordination motif in its sequence. Aminopeptidase component of ES-62 is yet to be assigned a function. It is not required for the immunomodulatory function, as demonstrated by the ability of PC moiety to mimic the effect of ES-62 (as mentioned in section 1.6.4.). Since adult females produce ES-62 while residing under the host's skin, it is possible that ES-62 aminopeptidase plays a role in cutting through the extracellular matrix or aids the digestion of food.

No attempts were made to crystallise ES-62 due to restrictions on the amount of available wild-type protein, its stability, as well as the flexible heterogeneous carbohydrate present on ES-62. Instead, small angle X-ray scattering data was used to determine 19Å resolution dummy atom model (DAM), using DAMMIN program. HYDROPRO was used to calculate a sedimentation coefficient of 10.19

S for the DAM, which agreed with the experimentally determined sedimentation coefficient of 9.85 +/- 0.24 S (Harnett W et al., 2003; Ackerman et al., 2003). The results confirmed that ES-62 is an assembly of 2 dimers tilted by ~30° with respect to the long axis of the particle (Harnett W et al., 2003; Ackerman et al., 2003). ES-62 is a tetrameric glycoprotein (~240 kDa), comprising identical monomers of ~62 kDa (Ackerman et al., 2003).

Based on the analysis of the amino acid sequence as well as the homology with other proteins, ES-62 contains a putative metal coordination motif associated with aminopeptidase function (Figure 1.3). Atomic emission spectrum of ES-62 indeed contains a strong magnesium signal (Harnett et al., 2003). There are no high-resolution structures for ES-62 homologues (Ackerman et al., 2003).

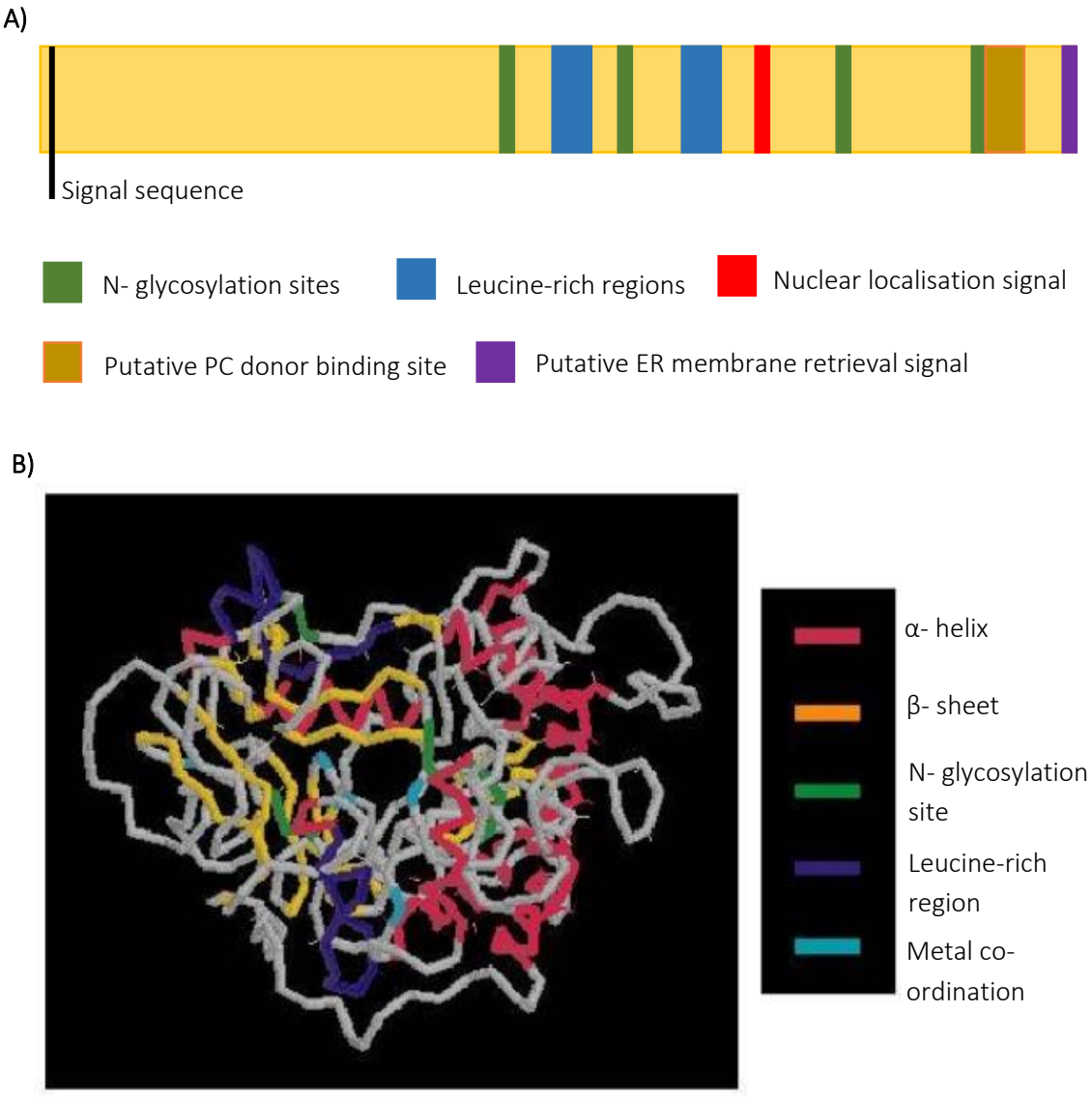


Figure 1.3. Diagram of the ES-62 functional residues. A) A schematic diagram showing the location of signal sequence, N-glycosylation sites, a possible PC donor interaction site, leucine- rich regions, nuclear localization signal site and possible ER membrane retrieval signal. **B)** A possible tertiary structure of the ES-62 monomer, predicted using dragon. The prediction includes positions of α-helices and β-strands, glycosylation sites, leucine-rich regions as well as metal co-ordination residues (The figure obtained from Goodridge et al., 2005).

1.6.3 ES-62 glycosylation

PROSITE is a database that allows determination of the function of uncharacterized proteins translated from genomic or cDNA sequences. It consists of descriptions of protein families, domains, functional sites and amino acid patterns. PRINTS is a compendium of protein fingerprints. PRINTS/PROSITE scanner was used to find conserved motifs for post- translational modifications of the ES-62 protein coding sequence. The ES-62 precursor contains four potential N-linked glycosylation sites at positions 4, 213, 254 and 400. The site at position 4 is in the signal sequence, so will be absent from the mature protein (Harnett, W et al., 1999). Schematic diagram of N-linked glycosylation sites is presented in Figure 1.3 (Goodridge et al., 2005).

To investigate the nature of the sugar modifications, ES-62 was labelled with [3H] glucosamine. The labelled ES-62 was exposed to enzymes: N-glycosidase F (that cleaves N-glycans from proteins) and O- glycosidase (that cleaves O-glycans from proteins). Only N-glycosidase F treatment resulted in the loss of radioactivity, which suggests that ES-62 may lack O-type glycans (Harnett W et al., 1993).

One type of previously uncharacterised N-glycan found on the molecule is trimmed to the tri-mannose core during oligosaccharide processing and then extended by N-acetylglucosamine (GlcNAc) residues (with or without core fucosylation) and finally decorated with the unusual post-translational modification of phosphorylcholine (PC) (Haslam et al., 1999; Harnett W et al., 1993; Ackerman et al., 2003).

ES-62 was originally detected by antibodies in serum samples from patients infected with parasitic nematodes. Such serum contained anti-PC antibodies and therefore ES-62 was suspected of containing PC (Harnett et al., 1989). Later on, ES-62 was found to interact with a monoclonal antibody as well as a myeloma protein reactive for PC (Harnett et al., 1990). Furthermore, it was discovered that ES-62 can be biosynthetically labelled with [3H]-choline (Harnett et al., 1995) and

the presence of PC on ES-62 was finally confirmed by Fast Atom Bombardment-MS (Haslam et al., 1997).

1.6.4 ES-62 PC modification

[3H] choline labels the PC modification. It was used to investigate the nature of the bond between PC and ES-62. Exposure to N-glycosidase F resulted in complete loss of radioactivity, suggesting that PC is attached via an N-type glycan. Further evidence for this was presented by culturing *A. viteae* worms with tunicamycin. Tunicamycin inhibits N-linked glycosylation. Worms cultured with tunicamycin produced PC- lacking ES-62 (Houston & Harnett W 1996).

Brefeldin A inhibits the transfer of proteins from the endoplasmic reticulum (ER) to the Golgi. When incubated with Brefeldin A, *A. viteae* worms did not secrete ES-62. The protein was detected in the worm extracts, but it lacked PC, which together with further experiments suggested that addition of PC is a post-ER, probably medial Golgi event (Houston et al., 1997).

ES-62 immunomodulatory activity depends on the PC moiety (Harnett W et al., 1999). Some protease activity has been attributed to the protein backbone (Pineda et al., 2014). The PC pattern is expressed in a range of pathogenic organisms: from bacteria and fungi to protozoa and gastrointestinal nematodes. PC modification on the surface of pathogens usually enables the detection of those pathogens by the host's antibodies or C-reactive protein (CRP) (Fallon, 2006). Only in nematodes PC can be present in glycoconjugates, specifically in phosphodiester linkage with N-type glycans in proteins (Houston & Harnett W 2004).

When compared to ES-62, administration of PC alone largely mimicked the results obtained with the whole protein. This suggests that the main active component of ES-62 is likely to be its PC modification. PC desensitizes B and T cells and dampens the proliferation of B and T lymphocytes *in vitro* and *in vivo.* This activity was compared to the loss of antigen receptors the phosphoinositide

3-kinase, protein kinase C and Ras mitogen-activating protein kinase pathways (Harnett and Harnett, 2001), however, the full mechanism of PC action in immunomodulation is unknown. It is possible that PC performs a dual immunomodulatory function by allowing the parasitic invasion, as well as being a target for innate and adaptive immune responses.

1.6.5 ES-62 homology to proteins from other parasites

At the beginning of this investigation, the only available sequence information for ES-62 was published cDNA sequence (Harnett et al., 1999). The cDNA sequence consists of 90 bp long 5'UTR, 1482 bp long protein coding region and 137 bp long 3'UTR followed by a polyadenylation chain. This sequence was used to establish if ES-62 is conserved in other nematode species including *C. elegans*. The nucleotide blast search revealed the high similarity between *A. viteae* ES-62 gene (coding sequence only) and that of human parasite proteins leucyl aminopeptidase from the 'eye worm' *Loa loa* (80% identity, 80% coverage) and *Brugia malayi* (78% identity, 47% coverage) (Figure 1, Appendix). However, other species contained highly similar short parts of the sequence that are likely to code for specific active domains (Figure 1.4). The protein sequence was analysed for similarities with other proteins. The analysis revealed highly conserved zinc- peptidase domains present in ES-62 (Figure 1.5).

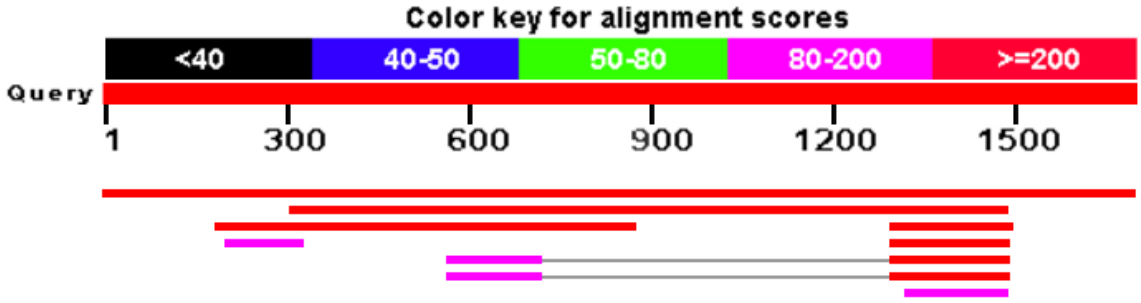


Figure 1.4. Distribution of blast alignments to the *A. viteae* ES-62 cDNA sequence query (including 5' and 3' untranslated regions). There appear to be two conserved parts between 300 and 900 nucleotide positions, especially a stretch around 600 nucleotides and at the end 3' end of the gene between 1300 and 1500 nucleotide position. These are likely to code for important, conserved domains (BLAST).

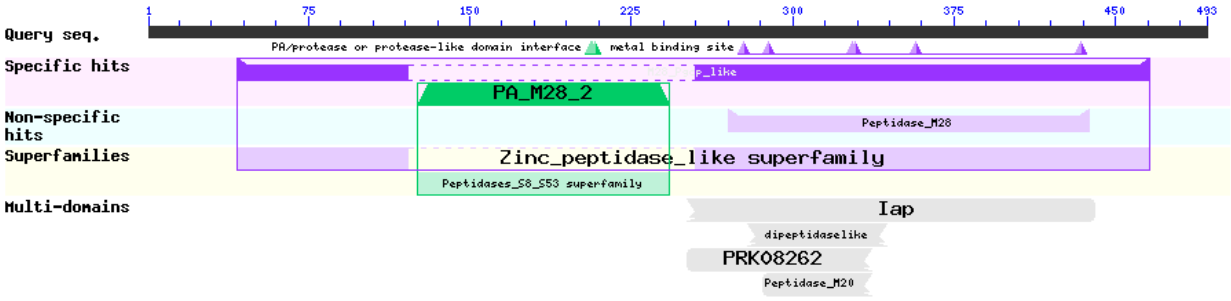


Figure 1.5. Graphical summary of conserved domains on ES-62 (protein BLAST). The BLAST alignment includes ES-62 in the Zinc-peptidase-like superfamily, due to the structural similarities, however, it was found that the active site of ES-62 contains a magnesium cation (Section 1.6.2.).

The blast search of the amino acid sequence again revealed a high degree of similarity amongst other species. The most similar being three leucyl aminopeptidases, two from *Loa loa* (71% identity, 90% coverage and 72% identity, 80% coverage) and one from *Brugia malayi* (67% identity, 80% coverage) (Figure 2, Appendix), as well as plasma glutamate carboxypeptidase from the pig roundworm *Ascaris suum* (57% identity, 87% coverage) and many other proteins from nematodes including *Pristionchus pacificus* (free-living nematode, satellite model organism to *C. elegans)*, *Wuchereria bancrofti* (human parasitic roundworm) and *Ancylostoma ceylanicum* (hookworm of humans and other mammals). The alignment to aminopeptidases was mainly due to a stretch of amino acids between ~250 and ~350 in ES-62 sequence. Aminopeptidases are usually metalloproteins (Taylor 1993), and examination of the sequence classified ES-62 to the zinc-peptidase-like superfamily (Figure 1.5). N-linked glycosylation sites predicted for ES-62 are not conserved in leucyl aminopeptidases.

Aminopeptidases are enzymes capable of degrading proteins by catalyzing the cleavage of amino acids from the amino terminus of protein or peptide substrates. They perform a wide variety of functions in organisms of animal and plant kingdoms. Some aminopeptidases contain a metal ion in their active site and, besides being enzymes, contribute to electron transfer reactions (cytochromes) or act as storage or transport proteins (Hoppert, 2011). Many aminopeptidases have been found in the secretory material of nematodes. Those aminopeptidases perform a wide variety of functions and can be involved in food metabolism, moulting, or fighting pathogens (Hoppert, 2011). Due to their enzymatic function, aminopeptidases have a potential to serve as effectors or inhibitors of biological processes. Filarial aminopeptidases have been found to play roles in modulating the host's immune response, embryogenesis, larval development, neuromodulation and digestion of peptides absorbed from the lumen of the intestine. They are potential drug targets and candidates for vaccines.

Human parasites produce ES-62 homologues with a high similarity at the amino acid level (*Brugia malayi* >70% and *Loa loa*- 80% homology) (BLAST). However, ES-62 homologue found in *B.*

malayi has been reported to lack PC raising the possibility of PC and protein moieties in *A. viteae* ES-62 having unconnected functions. Comparative studies on the N-glycans present in extracts of *A. viteae, O. volvulus* and *Onchocerca gibsoni* have revealed a high conservation of PC-containing N-glycans within filarial parasites (Pineda et al., 2014). As far as we are aware there is no ES-62 homologue found in free-living nematodes (Al-Riyami & Harnett W 2012).

Nevertheless, PC seems to be active regardless of the molecule that it is attached to. PC conjugated to ovalbumin (OVA) or BSA can protect mice against CIA in a similar fashion to that of ES-62 (Pineda et al., 2014). Confirming this, recombinant ES-62 produced in yeast (which lacks PC), failed to reduce the severity of CIA disease and the levels of pro-inflammatory cytokines in the joints. Interestingly, PC-OVA, unlike ES-62, did not alter the levels of anti-collagen IgG2a antibodies indicating that this aspect of ES-62 inhibition is PC-independent and might rely, at least partially, on the protein backbone or the attached glycans (Pineda et al., 2014).

1.7 Recombinant ES-62

Recombinant ES-62 lacking PC was expressed in yeast *Pichia pastoris* in order to investigate the function of the ES-62 protein backbone and confirm the role of the PC modification in immunomodulation (Egan et al., 2006). *Pichia pastoris* lacks the post-translational machinery necessary to attach PC to secreted proteins. The ES-62 gene was amplified and cloned into *E. coli*. The plasmid containing the ES-62 gene was sequenced, linearized and transformed into *P. pastoris.* Successful clones were allowed to express the recombinant protein. After 48 hours, the supernatant was separated by centrifugation and analysed by dot blotting with KK6 mouse monoclonal antibody that recognizes a conformational epitope on ES-62 (Egan et al., 2006). A positive clone was selected for a large-scale preparation of recombinant ES-62. This time, after the separation of the supernatant from the cells, it was filtered and concentrated by buffer exchange using a tangential flow system,

followed by further concentration step using Amicon centrifuge tubes with a 100 000 cut-off membrane (Egan et al., 2006).

The recombinant protein was recognised by a monoclonal antibody against a conformational epitope on the parasite- derived ES-62. However, the mouse polyclonal antibodies raised against the *A. viteae* ES-62 did not recognise the recombinant protein. Moreover, the mouse antisera generated against recombinant ES-62 failed to react with the parasite- derived ES-62 (Egan et al., 2006).

On a native-PAGE, under non-denaturing and non-reducing conditions, parasite-derived ES-62 runs as a tetramer. Under native conditions, recombinant ES-62 stained with Blue Colloidal-G stain ran as three, smear-like bands. The first smear ran slightly higher than the parasite- derived ES-62 and was identified as the tetramer of recombinant ES-62. Lower smeared bands were identified as dimers and monomers, which could lack secondary structure features stabilised by tetramerization (Egan et al., 2006). Under denaturing conditions parasite- derived ES-62 monomerizes (from a tetramer). The recombinant ES-62 failed to monomerize when examined by SDS-PAGE, under reducing conditions. The tetramer of the recombinant ES-62 was, however, recognised by the polyclonal rabbit α ES-62 serum (Egan et al., 2006).

Biophysical analysis of the recombinant ES-62 and comparison to *A. viteae* ES-62 revealed that the two varied in their secondary structure. The authors of the paper argued that the lack of immunological cross-reactivity would be due to the less folded structure of the recombinant material. The recombinant protein had most of its α-helix structures converted to β-sheets, which was likely to result in a shape change (Egan et al., 2006).

Finally, three single amino acid mutations were discovered in the recombinant protein. Each of those amino acids in native, *A. viteae* ES-62 is also conserved in *Brugia malayi* ES-62 and could take part in the correct protein folding. However, a mutation-free recombinant ES-62 was generated and accordingly to Egan it also failed to tetramerize completely. Similarly to the mutated recombinant

protein, the mouse antibodies raised against the mutation-free recombinant ES-62 failed to recognise the parasite-derived ES-62 (Egan et al., 2006).

The *P. pastoris* recombinant protein turned out to be much less effective and failed to inhibit arthritis in mice (Elliot & Weinstock 2012). The results suggest that the function of the ES-62 depends on the presence of the PC modification and a complete tetramerization. To date, the ability to post-translationally attach PC to an N-type glycan has only been attributed to nematodes.

1.8 *C. elegans*

Caenorhabditis elegans is a free-living nematode. In the lab, it is cultured on lab strains of *Escherichia coli* and grows exponentially at temperatures between 20-25°C. The full life cycle of *C. elegans* (egg to gravid adult) takes approximately three days. Gravid adult hermaphrodites produce up to 300 progeny each over the next three days leading to an incredible rate of growth. Most *C. elegans* are self-fertilizing hermaphrodites, they are homozygous and generate genetically identical progeny. Males arise infrequently, but male mating facilitates the isolation and maintenance of mutant strains as well as moving mutations between strains. *C. elegans* is usually grown on agar plates, but there are well-studied methods of harvesting the worms using liquid media (Brenner 1974).

Transgenic *C. elegans* can be made using a well-established microinjection technique (Evans 2006). A PCR fragment or a plasmid of the DNA coding for the desired protein can be introduced into the worm by injection into the gonad. Rescue selective markers, complementing the mutation, are included in the injection mix to enable selection of transgenic worms. The successful injection results in extrachromosomal arrays that are inherited by the progeny in a mosaic fashion. The resulting array has multiple copies of the transgene. Ionising irradiation can later be used to integrate the

extrachromosomal array into a chromosome and allow the stable transmission to all progeny and all cells.

The chosen DNA can be PCR- fused to *C. elegans* control elements (promoter, 5'UTR, 3'UTR), allowing *C. elegans* promoters to direct the desired expression of proteins from other species (Boulin et al., 2006).

C. elegans has previously been explored as an alternative expression system for genes from parasitic nematodes. A recombinant cystatin gene from *A. viteae* was detected in transgenic *C. elegans*, although it could not be purified by nickel chromatography (Pillai et al., 2005). More recently, an active cysteine protease of *Haemonchus contortus* has been expressed and purified using *C. elegans* (Murray et al., 2007). The purified protein resembled the native protein.

C. elegans has the potential to be a suitable expression system for parasitic nematode proteins as it provides machinery for correct folding and appropriate post-translational modifications (Pillai et al., 2005; Murray et al., 2007).

1.9 Excretory- Secretory system of nematodes

It is not entirely understood how *C. elegans* secretes proteins, what proteins are secreted and what are their functions. The secreted proteins could be a result of molting, uterine contents released with eggs or digestive enzymes. They could diffuse or leak from the soma, be released with pharyngeal pumping or be excreted through the cuticle (in *C. elegans* it is water permeable at all stages). The ES system of *C. elegans* was analysed by electron microscopy. It was found to consist of four cell types: excretory duct cell, excretory canal cell, excretory pore cell and excretory gland cells (Nelson et al., 1982). The morphology differs between the species, with some nematodes even lacking an excretory system (Davey & Kan 1968). However, many nematodes are permeated with excretory canals, part of a single H- shaped canal cell placed behind the pharynx (Figure 1.6).

The rate of excretion correlates to the osmotic strength of the media, which suggests that one of the functions of E-S system is osmoregulation (Weinstein 1952; Croll et al., 1972). The canal cell is often compared to a kidney responsible for maintaining the salt balance- excreting saline fluid via the duct and pore cells. The removal of metabolites has been hypothesised as another function of E-S system, and some nematodes concentrate injected dyes and use the excretory duct to expel them (Nelson et al., 1983). Gland cells are associated with excretions in many species; the secreted products

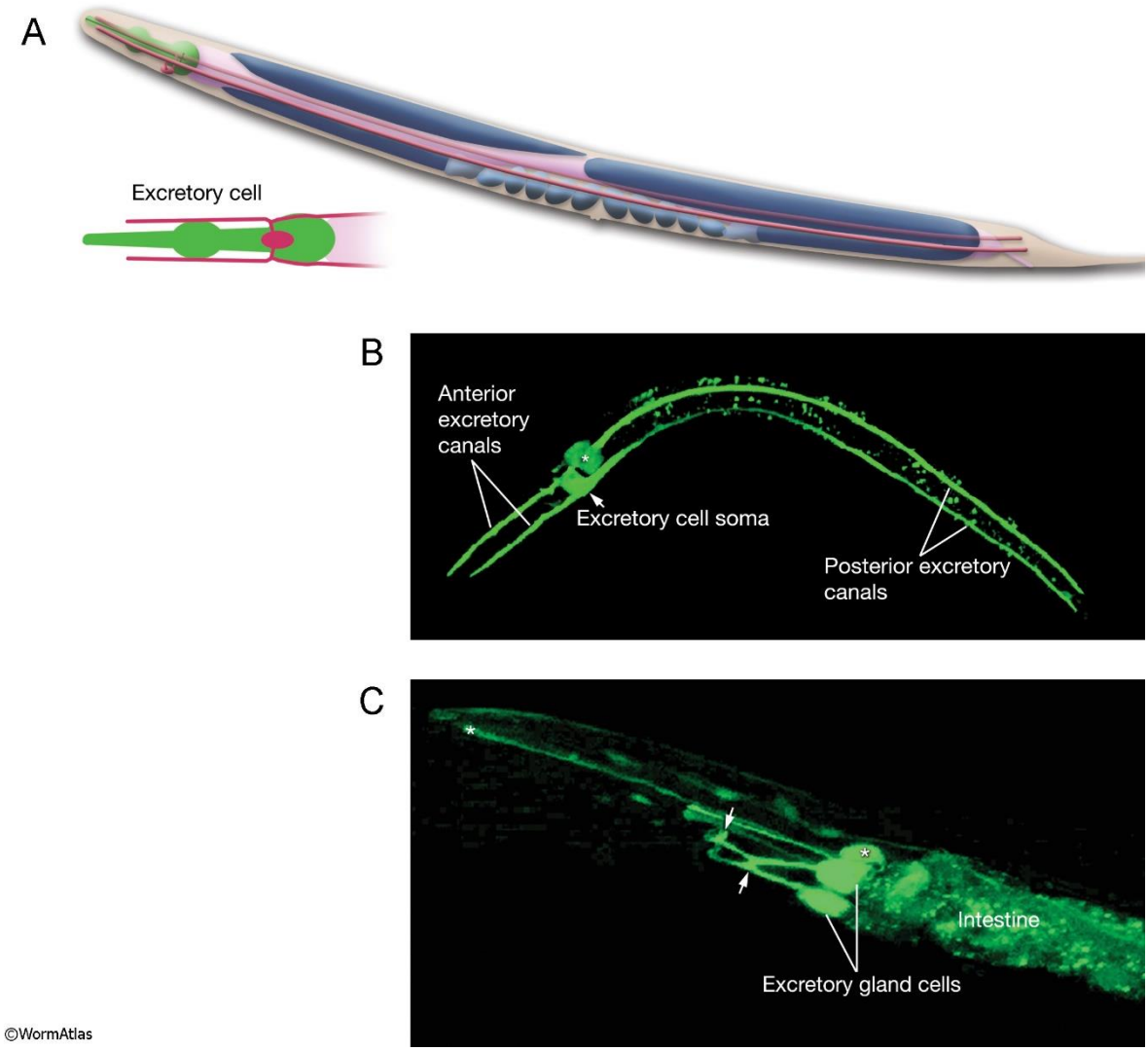


Figure 1.6. Adult excretory- secretory system in *C. elegans*.

A) Schematic showing the full length of the H-shaped excretory cell as seen from lateral side. In an adult animal, the excretory canals reach from the nose of the nematode to the tail region (WormAtlas).
B. Epifluorescent image of a transgenic adult animal (dorsal oblique view) expressing the *F22E10.1::GFP* reporter in the excretory cell. *Asterisk* marks the pharyngeal gland, which also expresses GFP (Image source: R. Newbury and D. Moerman, WormAtlas).
C. Epifluorescent micrograph of an animal showing expression of *B0403.4::GFP* in the excretory gland cell, ventral oblique view. The two small arrows point to regions where two cells fuse to make one syncytial cell. The anterior arrow marks the region where the excretory gland cell is suggested to receive synaptic input from the nerve ring. Pharyngeal gland cells also express this marker (*asterisks*) (Image source: R. Newbury and D. Moerman, WormAtlas).

could be released through the excretory duct. Gland cells concentrate secretory granules, their walls have membrane-like features and could be used for transporting products outside of the cell. The number of secretory granules is significantly higher in L4 and adult compared to younger *C. elegans* stages. Dauer larvae that undergo developmental arrest as well as starved larvae lose the secretory granules and the functional E-S system (Riddle et al., 1981). The pharynx was found to contain 5 cells that open into the lumen via short ducts (Albertson & Thomson 1976) and could be responsible for secreting digestive enzymes (Nelson et al., 1982) and release of molting peptidases (Singn & Sulston 1978).

The excretory-secretory molecules produced by parasitic nematodes have been postulated to originate from the excretory gland, connected to the duct and pore cells. The gland secretes materials from large membrane-bound vesicles (Lightowlers 1988). ES-62 was experimentally found in cells that underlie the gut and oesophagus of *A. viteae*. The protein was localised by the Immuno-Electron Microscopy with an antibody directed against the protein backbone. ES-62 was detected in secretory granules and the lumen of the worm gut (Harnett W 2003). It has been speculated that *A. viteae* ejects ES-62 protein through the pharynx, using pharyngeal pumping (Harnett W unpublished).

1.10 *C. elegans* excretory-secretory protein

ASP-6 is an aspartyl protease. It was reported to be excreted by *C. elegans* in a phosphorylcholine-substituted form, confirmed by the reactivity with the anti-PC antibody (TEPC-15) (Lochnit et al., 2006). The function of ASP-6, other than unspecific lysosomal protein degradation remains unclear. ASP-6::GFP fusion protein was used to localise the protein expression in adult worms. Diffuse fluorescence throughout the body cavity was observed, suggesting that the chimeric protein was secreted (Figure 1.7) (Lochnit et al., 2006), however, not necessarily externally. The translational fusion protein was seen in the pharynx. It is possible that the worm feeding organ is responsible for expelling the protein out of the body cavity.

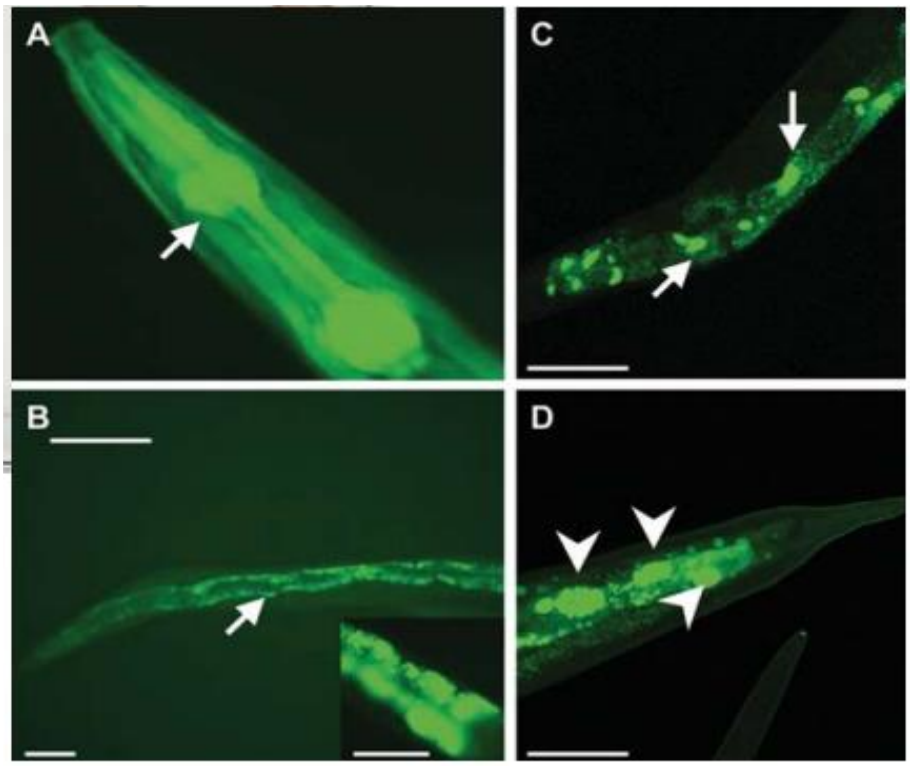


Figure 1.7. Expression of GFP labelled ASP-6 protein in adult transgenic *C. elegans*. A) The arrow indicates expression of asp-6 promoter+ASP-6::GFP in the pharynx (the feeding organ). **B)** The arrow indicates expression in the intestine. The diffuse expression can be observed in the body cavity. The detailed expression is shown in the bottom right corner. **C, D)** Transgenic animals expressing ASP-6 promoter::GFP in the intestinal cells (C- arrows) and in the epithelial cells (D- arrowheads) (Lochnit et al., 2006).

According to Lochnit et al ASP-6 protein is expressed in the intestine and the pharynx of gravid adult *C. elegans.* The pharynx is a feeding organ that works like a filter. Pharyngeal pumping consists of series of contractions and relaxations. During the contraction, pharynx takes in bacteria together with the liquid they are suspended in. During relaxation, the pharynx retains the bacteria and expels the fluid (Riddle, Blumenthal, Meyer et al., 1997). If ASP-6 protein is expressed in or reaches the pharynx from the intestine, it could be 'spat out' together with the bacterial fluid expelled by the pharynx during relaxation.

1.11 Challenges

Currently, ES-62 cannot be used as a drug. As a protein of ~240 kDa, it would be extremely costly to produce and purify it at industrial scale using current techniques for making parasite- derived protein. Moreover, ES-62 activity is attributed the post-translational attachment of the PC modification to an N-type glycan and, therefore, the production of the protein is highly nematode-dependent. The only functional ES-62 is the native, parasite- derived protein. In the lab, *A. viteae* ES-62 is made by growing parasites in gerbils, which leads to the death of all of the infected gerbils used for parasite recovery.

This dissertation aims to explore proteins secreted from *C. elegans* ES system as an alternative system for parasitic protein production.

The primary goal of the project was to find a method of engineering *C. elegans* to secrete foreign proteins into liquid media. If successful, the project would enable analysis of small parts of ES-62 to search for those suitable for therapeutic use. It would allow mutational analysis of ES-62 to define the functions of the amino acids of the protein. *C. elegans* as an expression system for recombinant foreign proteins would enable further investigation of proteins from human parasites. The

mechanism would also contribute to further understanding of the nematode excretory- secretory system.

C. elegans injected with ES-62 coding gene fused with the correct *C. elegans* promoter gene expressed in the excretory-secretory system, could produce *C. elegans* worms secreting large quantities of recombinant ES-62 while in liquid media.

1.12 *C. elegans* gene structure and regulation

Expression of genes, transcription into mRNA and translation into proteins are essential processes that are highly controlled in any organism. The number of regulators increases with the complexity of the organism (Vogel & Chothia 2006), but the core factors are highly conserved. Control elements for each gene are usually encoded for by the DNA regions located near that gene itself. The promoter is a DNA region directly upstream (5' end) from the transcription initiation site. It is a binding site for transcription factors, which determine if the gene will be active or silent (Levine & Tjian 2003). The promoter region can be used for directing the expression of genes in specific cells. In *C. elegans*, the control elements sufficient for correct expression of a gene are usually located in the DNA region up to 2000bp upstream from the translational start codon within the intergenic sequence (Reinke, Krause & Okkema 2005). Translational control relies on *cis*-regulatory elements including 5' untranslated region (UTR), which can regulate the length, secondary structure, upstream open-reading frames and specific sequences interacting with RNA binding proteins (Rhoads et al., 2006).

C. elegans introns are known to provide additional control of gene expression, especially larger introns may contain regulatory elements and need to be taken into consideration (Nam 2002; Okkema et al., 1993; Kostrouchova et al., 1998).

3'UTR plays a major role in mRNA expression. It can affect mRNA stability as well as the efficiency of translation. Moreover, 3'UTR and some introns contain micro RNAs (miRNAs), which function

as post-transcriptional regulators by targeting specific mRNAs for degradation or repression of translation (Wahid et al., 2010). unc-54 3'UTR is commonly used in *C. elegans* transformations to ensure transcript stability and efficiency (Boulin et al., 2006).

The regulation of transcription in *C. elegans* is mediated by RNA Polymerase II (Pol II) and is typical for eukaryotes. Protein coding genes contain a core promoter, where Pol II acts together with TATA Binding Protein (TBP) and TBP-Associated Factors (Dantonel et al., 2000; Kaltenbach et al., 2000; Lichtsteiner and Tjian, 1993; Walker et al., 2004). The regulation of transcription determines whether a gene will be active or not. The core promoter, upstream and downstream promoter regions, positive and negative enhancers act as the functional, regulatory parts. They contain cis-acting sequences, which bind trans-acting transcriptional factors (Levine and Tjian, 2003).

One unusual mechanism that *C. elegans* transcriptional regulation is capable of is trans-splicing. During trans-splicing the initial transcript 5' untranslated region (5'UTR) is replaced with a leader sequence (SL1), which is 22 nucleotides in size (Allen et al., 2011; Krause and Hirsh, 1987). The occurrence of trans- splicing accounts for the presence of polycistronic operons in *C. elegans* (Spieth et al., 1993). Polycistronic operons cluster genes and use a single set of transcriptional machinery to co-regulate them all. Trans-splicing is not, however, the only mean of transcriptional regulation. The promoters upstream differ from those within operons, which can account for independent regulation of some mRNAs (Allen et al., 2011; Huang et al., 2007; Yin et al., 2010; Morton and Blumenthal, 2011).

The core promoter sequence elements typically present in *C. elegans* have been investigated by Grishkevich et al. (2011). He observed that the five elements commonly present are a Specificity protein 1 (Sp1) like site (CNCCGCCC), T-blocks, which correlate with nucleosome eviction and gene expression levels (TTTT[N/T]), a TATA box (GTATA[TA][TA]AG), a trans-splicing site

(TTnCAG), and a Kozak site that includes the translation initiation codon ([CA]AA[CA]ATG) (Grishkevich et al., 2011 as mentioned by Reinke et al., 2005).

1.13 Aims and hypothesis

The primary aim of this thesis was to identify *Caenorhabditis elegans* excreted/ secreted proteins. The second aim was to characterise the genomic DNA structure of the gene coding for a secreted protein from *Acanthocheilonema viteae* - Excretory-Secretory 62 (ES-62) protein. The final aim was to use the ES-62 genetic information and the *C. elegans* excretory- secretory system information to make transgenic *C. elegans* that secreted functional recombinant ES-62 protein into the surrounding media. We hypothesised that understanding the biology and sequence of endogenous ES-62, as well as the biology of *C. elegans* secretion, will help us design the best way to use *C. elegans* as an alternative expression system for secreted recombinant ES-62.

2. Materials and methods

2.1 Media and Buffer Preparation

Nematode Growth Medium (NGM) was used for the maintenance of *C. elegans*. Worms were maintained on 6 cm or 9 cm plates. NGM (Brenner 1974) was prepared by dissolving 2.25 g (for 6 cm plates) or 4.5 g (for 9 cm plates) peptone, 18 g agar, and 2.75 g NaCl in 900 ml distilled water. After autoclaving for 20 minutes at 121°C, 0.9 ml of 5 mg/ ml cholesterol in ethanol, 0.9 ml 1 M $MgSO_4$, 0.9 ml 1 M $CaCl_2$ and 22.5 ml 1 M KH_2PO_4, buffer pH 6.0 were added in that order. 15 ml of media was poured into Petri dishes (6 cm). After agar solidification, the plates were stored at 4°C, for long- term storage.

Luria Broth medium was used for the culture of *E. coli*. This media consists of 10 g tryptone, 5 g yeast extract, and 10 g NaCl in 1 litre of distilled water, autoclaved for 20 minutes at 121°C.

M9 buffer was used as a media where worms were allowed to secrete proteins. To make the M9 buffer, 5.8 g $Na_2HPO_4 \cdot 2H_2O$, 3.9 g KH_2PO_4, and 5 g NaCl were dissolved in 1 L distilled H_2O, autoclaved for 20 minutes at 121°C.

S-media for liquid culture was prepared by adding autoclaved solutions of 10 ml 1M potassium citrate pH 6, 10 ml trace metals solution, 3 ml 1 M $Cacl_2$ and 3 ml 1 M $MgSO_4$ in 1 L of S-basal solution.

S Basal was prepared by dissolving 5.85 g NaCl, 1 g of anhydrous K_2HPO_4, 6 g of anhydrous KH_2PO_4, 1 ml cholesterol (5mg/ml in ethanol) in 1 L distilled H_2O, autoclaved for 20 minutes at 121°C.

Trace metals solution was prepared by dissolving 1.86 g disodium EDTA $\cdot 2H_2O$, 0.69 g $FeSO_4 \cdot 7H_2O$, 0.2 g $MnCl_2 \cdot 4H_2O$, 0.29 g $ZnSO_4 \cdot 7H_2O$, 0.025 g $CuSO_4 \cdot 5H_2O$ in 1 litre of distilled H_2O, autoclaved for 20 minutes at 121°C, and kept in the dark until needed.

Potassium citrate 1M pH 6.0 solution was prepared by dissolving 21.01 g of citric acid in 80 ml of distilled H_2O, adjusting pH to 6.0 with solid KOH and finally adjusting the volume to 100 ml with distilled H_2O, autoclaved for 20 minutes at 121°C.

2.2 Preparation of bacterial culture

E. coli is the food source for *C. elegans*. To prepare a culture, 30 ml of LB was pipetted into sterile 50 ml falcon tube. A disposable inoculating loop was used to inoculate the LB with OP50 *E. coli* bacteria from a frozen stock prepared previously. Once inoculated tubes were placed at 37°C on an orbital shaker set at 220 rpm and allowed to grow overnight. Cultures were then used for seeding NGM plates or stored at 4°C until needed.

E. coli pellets were prepared for liquid culture. The OP50 starter culture was prepared by inoculating bacteria in 1.5 ml of LB and shaking for 18 hours at 220 rpm at 37°C. On the next day 500ml of LB was inoculated with 1ml of OP50 starter culture in 2 L flasks, and was left shaking for 16 hours at 170 at 37°C. Bacteria were pelleted down at 5000 x g for 20 minutes, and the supernatant was discarded. The pellets were kept at 4°C or frozen at -80°C for long term storage.

2.3 Maintenance and synchronisation of *C. elegans* cultures

The protocol of *C. elegans* preparation was consistent with previously published guidelines (Stiernagle 2006) and used for all the experiments conducted during this study.

Eggs were prepared by treating adult hermaphrodites with alkaline bleach (7:8: NaOCl: 4 M NaOH) to remove the cuticle and any microbes attached to the worms. The eggs were washed with M9 buffer three times and then placed and pre-seeded on NGM plates with 100 µL OP50 bacteria (Weinkove et al., 2006). Animals were then left to grow at 20°C until starved, and then a 'chunk' was transferred

onto fresh NGM plates, pre-seeded with OP50 bacteria. Those animals were left to starve at 20°C. Newly starved plates were used for the preparation of plates for secretion experiments. A 'chunk' from the starved plate was transferred onto a fresh NGM plate, pre-seeded with OP50 bacteria. Worms were left at 20°C for 3- 4 days, until they were mostly young gravid adults (for liquid culture- just cleared of the bacteria). Those worms were later used for the secretion experiment or to establish a liquid culture.

2.4 Preparation of *C. elegans* plate culture for protein secretion

Nine cm plates were seeded with 200 µl of OP50 bacterial culture, and the culture was spread out to cover most of the plate. Plates were kept at room temperature for two days to allow the bacterial lawn to grow. A chunk from previously prepared starved worm plates was transferred onto the fresh plates with the bacteria. Worm plates were kept at 25°C for 3- 4 days until the cultures synchronized to mainly young gravid adults. The worms were washed off each plate with 3ml of sterile M9 and allowed to settle for 10 minutes. The supernatant was discarded, and the worms were washed with 30 ml of fresh M9 3 times. After the last wash, most of the M9 was removed, and the worm volume was estimated. An equal volume of fresh M9 was added to the worms. The worms were left to shake at 23°C, 177 rpm for 4 or 19 hours.

2.4.1 Harvesting the worms and secreted media after protein secretion (for plate culture *C. elegans*)

The mixture of worms and media was transferred into 2 ml Eppendorf tube and centrifuged at 828 x g for 1 min. The supernatant was transferred to a fresh tube and centrifuged at 9200 x g for 10 min to remove any remaining worm debris. The supernatant was transferred to a fresh tube and kept as the secreted worm media. The worm pellet was washed three times with 1 ml of fresh M9. It was

centrifuged at 828 x g for 1min after each wash, and the supernatant was removed. Protease inhibitors (cOmplete™ Protease Inhibitor Cocktail tablets, Roche) were added to both media and pellets. The media was stored at 4°C, and the pellets were frozen in aliquots at -20°C.

2.5 Preparation of *C. elegans* liquid culture for protein secretion

Bacterial pellets were thawed on ice (if taken from -80°C) and re- suspended in 10 ml of S medium. Two hundred and fifty ml of S Medium was added to a sterilized 1 L flask. The S Medium was inoculated with re- suspended *E. coli* OP50 pellet made from 2 L of an overnight culture. Four 9 cm plates of *C. elegans* (just cleared of bacteria) were washed with 5 ml of S media and added to 1 L flasks with 250 ml S Medium. The flask was kept on a shaker at 23°C at 220 rpm. Cultures were monitored every day under the microscope by pipetting 200 µl of the culture onto a clear NGM plate. As soon as the food supply was depleted (the solution was no longer visibly cloudy) and there were many adult animals in each drop, the culture was considered ready to be harvested. Usually on the 5^{th} or 6^{th} day.

2.5.1 Harvesting the worms and secreted media after protein secretion (for liquid culture *C. elegans*)

The liquid culture of worms was transferred from a 1L flask to a fresh 250 ml conical flask. The flask was kept tilted to allow the worms to settle and then the supernatant was removed. The worms were washed 5 times with 100 ml of fresh M9, each time adding the M9, swirling, allowing the worms to settle for 10 min, tilting the flask and allowing the worms to settle for another 5 min. Each time the supernatant was completely removed from the worms. After the final wash, the worms were removed into a fresh 50 ml conical tube to estimate the worm volume. The M9 volume of 1.5 x the volume of the worm pellet was used to inoculate the worm pellet inside a sterile 250ml flask. The worms were

left to shake at 200 rpm, at 23°C for 4 hours. After that, the worms and media mix was transferred to a 50 ml conical tube. The worms were pelleted by centrifuging at 3000 x g for 3 min. The supernatant was transferred to a fresh tube. Both worms and the media were kept. The media was centrifuged twice at 4000 x g for 15 minutes to pellet down any debris. The supernatant was removed and retained as the secreted worm media sample. The worm pellet was washed three times with 20 ml of fresh M9 by centrifuging at 3000 x g for 3 min, removing the supernatant and repeating. Protease inhibitors (cOmplete Protease Inhibitor Cocktail tablets, Roche) were added to both the media and the worm pellets. Media was kept at 4°C, and worm pellets were frozen at -20°C.

2.6 Extraction of genomic DNA from *C. elegans*

Fresh 1.5x Lysis Buffer was prepared on the day. To prepare the lysis buffer the following solutions were added: 0.6 ml 5 M NaCl, 1.5 ml 1 M Tris-HCl pH 8- 8.5, 1.5 ml 0.5 M EDTA, 0.75 ml of 10% SDS and 95.5 ml of distilled H_2O. Worms were washed off the plate with sterile M9, allowed to settle and the supernatant was removed. The volume of the pellet was estimated, and the lysis buffer was added at a volume of 1.5x the volume of the worm pellet. Finally 10 mg/ ml proteinase K (Sigma) solution was added to the lysis reaction in the ratio of 2:100. The lysis reaction was kept at 56°C for 2 hours, with shaking, at least, every 30 minutes. The worm lysis solution was used to extract the genomic DNA using an isolation column. (Sigma, bacterial genomic DNA extraction kit).

2.7 Extraction of genomic DNA from *A. viteae*

Fresh 1.5x Lysis Buffer was prepared on the day as described before. Frozen *A. viteae* worms were received from Harnett lab at Strathclyde University. 2- 3 worms were placed in the Eppendorf tube and sliced using a sterile scalpel. The volume of the pellet was estimated, and the lysis buffer was added at a volume of 1.5x the volume of the worm pellet. Finally, a volume of 2% of the lysis reaction

of 10 mg/ ml proteinase K solution was added to the lysis reaction. The lysis reaction was kept at 56°C for 3-4 hours (until there were no visible worm pieces left), with shaking, at least, every 30 minutes. The worm lysis solution was used to extract the genomic DNA using an isolation column. (Sigma, bacterial genomic DNA extraction kit).

2.8 PCR

All PCRs were carried out using Phusion high fidelity polymerase (New England Biolabs) in Techne TC-5000 Gradient Thermocycler. The reactions were set-up accordingly to the manufacturer's instructions. Specific primers used for each reaction are outlined in Table 1, Appendix.

2.8.1 Nested PCR amplification of the genomic ES-62 fragment

To amplify the genomic ES-62 fragment a nested PCR was carried out, that consisted of two PCR steps. Step 1 PCR used the pair of external primers for initial amplification of the ES-62 fragment from *A. viteae* genomic DNA. At this stage the product was not detectable by gel electrophoresis. Step 2 PCR used the step 1 PCR product as a template DNA and the pair of internal (or nested) primers to re- amplify the ES-62 fragment. Internal primers were designed so that, if successful, the second reaction would result in a slightly smaller fragment than that from the first reaction. Conditions for each reaction were set-up according to the manuacturer's instructions (Phusion, New England Biolabs).

2.9 Amplification of *asp-6::* genomic ES-62 DNA fusions

A rapid, PCR- based fusion was used to create transgenes (Hobert, 2002). The idea entails a simple fusion PCR, which joins two primary PCR products with a set of nested primers. *C. elegans asp-6* promoter fragment was amplified from *C. elegans* genomic DNA using a 5'primer (external) and a

3’ fusion primer with ES-62 genomic DNA overhang that will act as the 5’ primer to attach the ES-62 genomic fragment to result in the final 2-way fusion product for injections.

The primer exhausting PCR was performed to exhaust the left-over primers from the previous reactions (amplification of the promoter and amplification of the ES-62 genomic fragment) and stop them from interfering with the fusion PCR reaction. PCR products from the amplification of the promoter and amplification of the ES-62 genomic fragment were used as the template for the fusion PCR and were mixed in 1:1 ratio (based on the nucleotide content calculated as concentration x fragment size). No primers were used in this step. The product of the primer exhausting PCR is referred to as a ‘seed’.

The seed was used as a DNA template for the PCR fusion. The pair of internal primers was used to amplify the final product: internal 5’ primer inside *asp-6* promoter sequence and internal 3’ primer inside ES-62 sequence.

2.10 Sequencing PCR products

Fusion PCR reactions were mixed with 6X DNA Loading Dye (Thermo Scientific), which prepares the samples for loading on agarose gels. The loading dye contains bromophenol blue and xylene cyanol FF, which allows for visual tracking of DNA migration during electrophoresis. The samples were loaded onto 1% agarose gels, which were run at 100V. The bands of the size expected for the fusion PCR were excised from the gel under UV light. The excised bands were purified using a gel extraction kit (Qiagen), according to manufacturers’ instructions and eluted in Milli-Q® purified H_2O. If necessary, the samples were concentrated in a centrifugal dehydrator to a final concentration of 5 ng/ ml in 25 µl of the solution required for sequencing. The concentration was checked with the Nanodrop (Thermo Scientific).

2.11 Microinjection of *C. elegans*

Lin-15 C. elegans mutant worms were thawed and kept for at least two generations at 15°C. A 'chunk' from starved plates was cut out and transferred onto a seeded plate and maintained at 15°C for 4- 5 days until the worms were young gravid adults with a single row of eggs.

Preparation of injection pads: Injection pads were used to 'hold' the worm and keep it stationary for the microinjection. Pads were made with 2.5% agarose solution in water. The solution was mixed and boiled. Using a P200 pipette with its tip cut off, a drop around 100 µl of hot agarose was placed onto 50 × 22 mm glass coverslip. A clean coverslip was immediately placed on the top of the drop and lightly tapped. After 2 minutes when the agarose solidified, the coverslips were slid apart and left to dry overnight.

Preparation of needles: Injection needles were pulled, using the PC- 10E needle puller, which pulls the needle vertically from capillary using gravitational force of its weight. Settings used: 1 stage-pull, 68 °C, 2 heavy weights and 1 light weight.

Preparation of DNA for microinjection: Approximately 100 ng of fusion PCR product was mixed with 250 ng of pEKLI5 *lin-15* rescue plasmid. The injection mix was centrifuged at 15,550 x g for 10 minutes to pellet any debris. The supernatant was transferred to a fresh tube and stored at -20 °C. Shortly before microinjection, the DNA was thawed on ice. It was centrifuged at 15,550 x g for 5 minutes.

Microinjection set-up: The needle was back-loaded with 0.5 µl of DNA, placed in the collar of the instrument holder, connected to the injector and mounted on the manipulator to the optimal position under highest magnification (400x). The well maintained, well fed *C. elegans ∆lin-15* mutants (preferably young gravid adults with a single row of eggs) were chosen for microinjection. A small drop of oil was deposited on the side of the agarose on the injection pad. An eyelash attached to the glass capillary was used as a worm-pick to minimize the amount of OP50 bacteria on the injection

pad. The eyelash pick was firstly sterilized in 70% clean ethanol. A single young adult was picked by first aligning the eyelash pick with the worm, then angling it under the worm and letting it curl around the pick. The worm was picked from the area of the plate away from the bacterial lawn to avoid bacteria transferred to the injection pad. Working under the dissecting microscope, the worm was moved to the injection pad and placed on the side of the oil droplet. The worm was allowed to crawl off the eyelash and then was gently oriented and pressed down to 'stick' to the agarose. The injection pad with stationary worm was placed under the injection microscope. The highest magnification (400x) was used to position the worm in the middle, and the microscope was focused on the gonad. The loaded needle was brought into the picture and moved until it reached the same plane with the gonad. The needle was directed into the gonad, and the worm was injected. To recover the injected animals, the pad was transferred back under the dissection microscope. The worm was covered with a drop (2- 5 µl) of fresh M9 and transferred with an eyelash pick onto a clean NGM OP50 seeded plate. No more than 3 injected worms were placed on each fresh plate. Plates were kept at 20°C to allow the injected worms to recover and lay eggs. Worms were monitored daily for the presence of progeny. Approximately 7 days after injection the progeny was ready for initial screening. Each non-MUV worm was transferred onto a single, OP50 seeded plate. The worms were incubated at 25°C and allowed to lay eggs. After 3- 4 days the progeny was screened again. Each worm that produced non-MUV progeny was considered an independent transgenic line.

2.11.1 List of *C. elegans* transgenic strains used

6B – *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 genomic DNA fragment + 3'UTR

10A- *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 genomic DNA fragment + 3'UTR

IIL- *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 cDNA :: *asp-6* 3'UTR

9*- integrated 6B *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 genomic DNA fragment + 3'UTR

3-3 control- Δ*lin-15* mutant rescued with *lin-15* plasmid

2.12 Single worm lysis and diagnostic PCR

The protocol for single- worm PCR was adapted from the Plasterk lab protocols. Single worm lysis buffer was prepared by adding 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM Tris-HCl pH 8.3, 0.5% NP40, 0.5% Tween 20 and 0.01% gelatine. The buffer was kept at -20°C and thawed on ice before the lysis reaction. After thawing 0.1 mg/ml of proteinase K was added to the lysis buffer just before adding worms. Single worms were picked from plates and placed in a PCR tube in 5µl lysis buffer. The reaction was kept at 60°C for 60 minutes, followed by 15 minutes deactivation at 95°C. Then 5µl of MQ H_2O was added to the reaction. Diagnostic PCR was performed with Dream Taq Green PCR Master Mix 2x solution (Thermo Fisher Scientific) accordingly to the manufacturer's instructions.

2.13 *C. elegans* pellets lysis after secretion

To lyse the large worm pellets after liquid culture, the 1x Lysis buffer was used. The lysis buffer 1x was prepared by adding the following ingredients: 50 mM HEPES pH 7.4, 1 mM EGTA, 1 mM $MgCl_2$, 100 mM KCl, 10% glycerol, 0.05% NP-40 and 0.5 mM DTT.

The volume and the weight of the pellet were estimated at the beginning. The worms were freeze-thawed 3 times by keeping the tubes in liquid nitrogen for 30 seconds and then in a 45°C water bath until the worm pellet just thaws. The lysis buffer was chilled on ice. Approximately 1.7 g of the worm pellet was suspended in 5 ml of lysis buffer. More lysis buffer was added to the final 1:1 volume ratio of worm pellet to the lysis buffer. The worms were sonicated for 15 seconds (while on ice), using the 30% power and then left on ice for 45 seconds. The sonication and chilling cycle was repeated 3 times in total. After that, the worms were checked under the microscope. In order to break them open a second round of sonication was performed 6 times as follows: 20 seconds at 50% power,

40 seconds on ice. The sonicated mix was centrifuged at 16,000 x g for 30 minutes. After centrifugation the lysis separated into 3 layers, they were each kept separately. The concentration of each fraction was checked using the Nanodrop.

2.14 Concentration of *C. elegans* lysate

The lysate samples were concentrated using mini Amicon centrifuge filters. Lysates were centrifuged at 14,000 x g for 20 minutes until the volume decreased approximately 4 times. The concentrate was recovered by reversing the filter and centrifuging at 1000 x g for 2 minutes. The concentration was measured using a Nanodrop and with the 4 x volume reduction increased approximately 3 times.

2.15 Concentration of *C. elegans* media

Media was concentrated using Amicon centrifuge filters at 4°C in a centrifuge with a swing out rotor. It was centrifuged for 45 minutes at 4000 x g until the volume was reduced 10 times. If the volume of the media for concentration was larger than 15ml the same filter was reused. Flow through was kept to make sure there was no protein left. The concentrations of original media, concentrated media and flow- through were compared using Nanodrop measurements. With a 10 times volume reduction, the concentration increased approximately 10 times. The concentrated sample was further concentrated using mini Amicon centrifuge filters. It was centrifuged at 14,000 x g for 30 minutes until the volume decreased 5 times. The concentrate was recovered by reversing the filter and centrifuging at 1000 x g for 2 minutes. The concentration was measured using a Nanodrop and with the 5 x volume reduction increased approximately 3 times.

2.16 SDS-PAGE

Precast Run Blue SDS Protein gels and buffers (Expedeon) were used for SDS-PAGE according to manufacturer's description. The gels were either 10% or 12%. RunBlue LDS Sample Buffer 4X Concentrate from Expedeon was used to denature the samples and prepare them for running on the gel. The sample buffer consisted of a trace of bromophenol blue, 0.3125M Tris-HCl pH 6.8, 50% Glycerol, 10% SDS and 5% DTT (for denaturing gel). SDS is a detergent, which disrupts the tertiary structure of proteins. DTT is a reducing agent, which breaks down protein-protein disulphide bonds. SDS, DTT and high temperature bring folded proteins down to linear molecules. SDS has a second function of coating the protein with a negative charge, making the charge of the protein proportional to its molecular weight. The samples were mixed with the sample buffer and incubated for 10 minutes in 70°C water bath.

2.17 Silver staining

Plus one silver staining protocol devised by Joanne Robson (Durham University Proteomics) was followed to stain all SDS protein gels. Immediately after running, the gel was briefly washed in MQ water and then transferred to a fix solution (40% methanol, 10% acetic acid), the fixing step was performed twice for 30 minutes each time (or overnight with one change of the fixing solution) with shaking. After that the gel was placed in sensitizing solution (75ml methanol, 0.32g of anhydrous Na- thiosulphate and 17g Na-acetate trihydrate made to 250ml with MQ water) and kept for 30 minutes with shaking. Next the gel was washed 3 times for 10 minutes each time with MQ water. The gel was transferred to a silver solution (0.625g of silver nitrate made to 250ml with MQ water) and kept for 20 minutes with shaking. After that the gel was washed twice in MQ water, briefly and transferred to a developing solution (6.25g anhydrous Na-carbonate made to 250ml with MQ water, 100 µl formaldehyde added just before treatment) and shaken gently until the bands became visible. The reaction was stopped by discarding the developing solution and covering the gel with the stop

solution (3.65 g EDTA made to 250 ml with MQ water) for at least 10 minutes. After that, the gel was washed and kept in MQ water.

2.18 Western blot analysis

The transfer was prepared as follows: The PVDF membrane (Millipore) was dipped in methanol, washed with distilled water and then presoaked in transfer buffer for 10 minutes. The ready protein gel was taken out from the apparatus and presoaked in transfer buffer for 10 minutes. The sponges and filter papers were soaked in transfer buffer; the transfer was assembled. The transfer was carried out at 4°C at 25V overnight (approximately 23 hours). The blot was then washed with distilled water and blocked in 5% milk or BSA in TBS-Tween 0.1% for at least 2 hours at room temperature on a rotator. The blot was covered with primary antibody in 5% BSA or milk solution in TBS-Tween 0.1%, and the blot was incubated overnight in cold room 4°C (on a rotator). On the next day, the blot was washed with TBS-Tween 0.1%, 3X for 10 minutes. The blot was covered with secondary antibody in 5% milk or BSA solution in TBS-Tween 0.1% at room temperature for 30 minutes (on a rotator). The blot was washed with TBS-Tween 0.1%, 3X for 15 minutes and then with 1x TBS for 10 minutes. ECL solution (Chemiluminescent Substrate Reagent Kit, Pierce) was used accordingly to the manufacturer's instructions. The blot was placed in a film cassette and visualized in the dark room using film and developing machine.

2.18.1 Antibodies

Rabbit α ES-62 antibody and rabbit α de-glycosylated ES-62 antibody were used as primary antibodies as detection systems as described previously (Harnett et al., 1993).

2.19 Preparation of samples for silver stain (TCA precipitation followed by acetone precipitation)

The media with protease inhibitors was TCA precipitated. Fresh, cold 20% TCA was added to the media sample to a final concentration of 10%. The sample was mixed by inverting. It was kept on ice for 30min – 1 hr. After that, the solution was centrifuged at 15,550 x g for 15 min at 4°C. The supernatant was removed and then the sample was centrifuged again at 15,550 x g for 2 minutes at 4°C. The remaining supernatant was removed. Acetone was pre-chilled at -20°C overnight. 500 µl of cold acetone was added to the solution and vortexed. It was kept at -20°C for 10 minutes and then centrifuged at 15,550 x g for 10 min at 4°C. The supernatant was removed carefully without dislodging the pellet. The acetone wash was performed 3 times. The pellet was air-dried at RT for 5-10 min. The rehydration buffer was added. The solution was kept on ice for 1 hour and vortexed every 10 minutes.

2.20 Buffer exchange for media samples before MALDI-TOF MS analysis

Buffer exchange was performed for the samples in preparation for in- liquid trypsin digestion and MALDI-TOF MS. The buffer was changed from M9 to 50 mM Triethylammonium bicarbonate buffer. The buffer exchange was performed using Centrifugal Filter Units for concentration and purification of biological solutions Amicon Ultra- 15 Ultracel 100K accordingly to manufacturer's instructions.

2.21 StageTip fractionation of trypsin- digested samples

After trypsin- digestion, peptide samples were desalted and fractionated using ion exchange chromatography, performed with multi-StageTips. Step 1C from the protocol by Rappsilber et al (2007) was followed. Stagetips were used in order to prepare trypsin digested samples for MS

analysis. Stagetips are pipette tips containing small disks of beads with reversed phase and, in this case, anion- exchange surfaces. The anion-exchange disk allows for the separation of peptides in the mix based on their charge. The separation and elution of different fractions is allowed by passing increasing salt concentrations through the stagetip. The reversed phase disk is used for desalting of the sample. Fractionated samples are clean of salt and contain a simpler mixtures of peptides, which aids more accurate recognition (Rappsilber et al., 2007)

3. Results

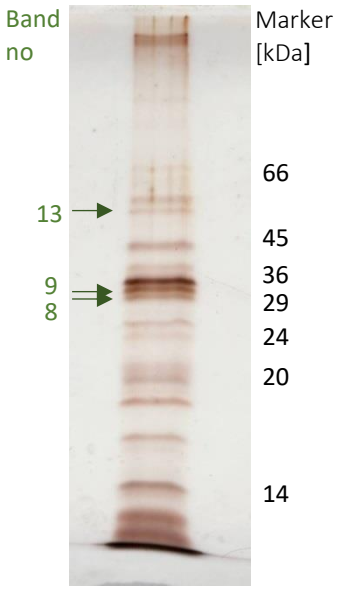
3.1 *C. elegans* secrete proteins into the media over a relatively short time period

In order to find out if *C. elegans* secrete proteins into the media, several experiments were performed where wild- type *C. elegans* worms grown on agar plates were washed off and separated from the *E. coli* bacteria. *C. elegans* were then allowed to swim in liquid media (method described in Section 2.4.). After 4 hours the worms and the media were separated by sedimentation and the media was analysed for the protein content using Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis (SDS-PAGE). SDS-PAGE is a method of separating proteins in their denatured form accordingly to their size. After the separation, proteins can be detected and visualised using staining methods. Silver staining was chosen for this experiment due to its sensitivity (in the very low nanogram range) and compatibility with downstream processing such as mass spectrometry. During silver staining the proteins separated by SDS-PAGE bind silver ions, which can later be reduced under appropriate conditions to visible silver metal (Chevallet et al., 2006).

The experiments revealed that *C. elegans* secrete a simple mixture of proteins into the media. Those proteins can be detected by silver staining (Figure 3.1 and 3.2).

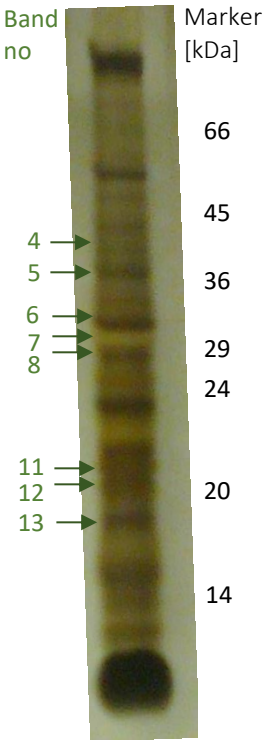
Matrix Assisted Laser Desorption/Ionization- Time of Flight Mass Spectrometry (MALDI-TOF MS) technique was used to identify proteins represented by individual bands on the gel. Protein identification by MALDI-TOF MS relies on the principle that digestion of each protein with a specific protease will result in a mixture of peptides with peptide masses that are unique to that protein (Webster & Oxley 2012).

The most intense bands were cut out from the gel and individually subjected to in-gel trypsin digestion by Joanne Robson and later Adrian Brown (Durham University Proteomics). MALDI-TOF MS was used to determine the mass of those peptides, which gave a characteristic dataset called a



Band no	Best Match	Protein score
13	hypothetical protein K12H4.7 [Caenorhabditis elegans]	60
9	Chain A, OmpF Porin Deletion (Mutant Delta 109-114)	285
	Aspartyl Protease family member (ASP-6) [Caenorhabditis elegans]	193
8	Chain A, OmpF Porin Deletion (Mutant Delta 109-114)	89
	Aspartyl Protease family member (ASP-5) [Caenorhabditis elegans]	219

Figure 3.1. Silver stain of *C. elegans* proteins secreted into the M9 medium over 4 hours (wild type N2 worms). Total loaded protein- 10 µg. Individual bands were cut out from the gel and analysed by MALDI- TOF MS. The significant results are presented in the table 3.1. *C. elegans* proteins are indicated in black, *E. coli* proteins are indicated in red. Only significant matches are presented. Probabilities were converted to scores for reporting purposes. A protein score higher than 57 has less than a 5% probability of being a random event and is considered an identity match.



Band no	Best Match	Protein score
4	GLCM4 Putative glucosylceramidase [Caenorhabditis elegans]	146
5	YM67 Putative serine protease K12H4.7 [Caenorhabditis elegans]	373
6	OmpF [Escherichia coli]	572
7	ASP6 Aspartic protease 6 [Caenorhabditis elegans] OmpF [Escherichia coli]	149 115
8	ASP6 Aspartic protease 6 [Caenorhabditis elegans] OmpF [Escherichia coli]	229 826
11	OmpF [Escherichia coli]	147
12	YUA6 uncharacterised serine carboxypeptidase F13S12.6 [Caenorhabditis elegans] OmpA [Escherichia coli]	94 425
13	YUA6 uncharacterised serine carboxypeptidase F13S12.6 [Caenorhabditis elegans] OmpX [Escherichia coli] DPS DNA protection during starvation protein [Escherichia coli]	206 230 262

Figure 3.2. Silver stain of *C. elegans* proteins secreted into the M9 medium over 4 hours (wild type N2 worms). Individual bands were cut out from the gel and analysed by MALDI- TOF MS. The significant results are presented in the table 3.2. *C. elegans* proteins are indicated in black, *E. coli* proteins are indicated in red. Probabilities were converted to scores for reporting purposes. A protein score higher than 52 has less than a 5% probability of being a random event and is considered an identity match.

Peptide Mass Fingerprint (PMF). The PMF for each band was later compared to the theoretical trypsin digest of proteins in the sequence database to find the best match. The protein score is a measure of the statistical significance of a protein mass fingerprinting (PMF) match. It is the sum of the highest ions score for each sequence. Ion score is based on a calculated probability (P) that the observed match between the experimental data and the database sequence is a random event. It depends on how many experimental peptide sequences match a single database sequence and how strong those matches are. It is also affected by the signal to noise ratio, rather than the concentration of each peptide (Perkins et al., 1999). On a 1D SDS gel, it is highly likely that each band will contain a mixture of proteins of very similar sizes. Those proteins, as well as other impurities, create noise, which interferes with the signal. For the investigation of the PMF data, it is possible to choose a specific organism database as well as the potential contaminants to reduce the unspecific identifications. In this case, we were interested in *C. elegans* proteins, but we also included the food source- *E. coli* protein database as a likely contaminant. A protein match with a significant score has less than 5% probability of being a random event and is considered an extensive homology or an identity match (Matrix science).

The MS analysis of individual bands from the *C. elegans* media gel identified the most abundant proteins, as judged by the intensity of the silver stain, as aspartyl proteases ASP-6 and ASP-5 (Figure 3.1 and 3.2). MS analysis consistently identified ASP-6 to be present in the medium where *C. elegans* were allowed to swim for 4 hours. Detection of ASP-6 inside the worm (Lochnit et al., 2006, Section 1.10.) and in the medium (Figure 3.1 and 3.2) suggests that *C. elegans* possess an intrinsic secretory pathway capable of secreting proteins externally. We hypothesized that *C. elegans* secretory pathway can be harnessed to produce recombinant protein. Using excretory- secretory system for protein expression will aid correct folding of the recombinant protein. It is advantageous when it comes to purification steps as there is no requirement for breaking down worm cuticle. Also, the secretome is much less complex than the proteome of the whole worm, which will also aid easier and faster purification.

3.2 Characterisation of *A. viteae* ES-62 genomic DNA fragment

3.2.1 Extraction of genomic DNA

The first strategy was to transform *C. elegans* with ES-62 genomic DNA fragment. In order to prepare the template DNA for amplification of the ES-62 fragment, genomic DNA from frozen *A. viteae* worms was prepared using a standard protocol (Section 2.7.). *C. elegans* genomic DNA for amplification of *C. elegans* control elements was prepared. Both DNA templates were analysed on a 0.8% agarose gel to check their quality and approximate concentration (Figure 3.3.). Due to past difficulties with the amplification of ES-62 gene from parasites, two methods were compared for extraction of *A. viteae* DNA: Edwards Prep and standard genomic DNA kit extraction. DNA extracted using standard genomic kit was used as the template for the PCR reaction. The kit extraction resulted in less concentrated DNA, with less RNA contamination, which should assist with the denaturation of the genomic DNA and aid amplification of ES-62.

A. viteae genomic DNA of sufficient quality and concentration can be extracted from frozen parasites using standard techniques.

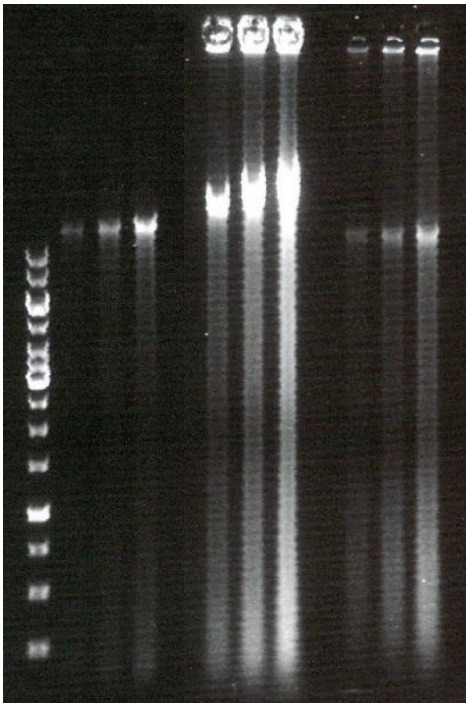


Figure 3.3. Electrophoresis of dilutions of genomic DNA extracted from *A. viteae* and *C. elegans* worms. One x is 1 µl of DNA in 5 µl of solution. Two x is 2 µl of DNA in 5µl of solution. Four x is 4 µl of DNA in 5 µl of solution. DNA samples were diluted with TE buffer. Ten µl of the diluted sample was mixed with 2µl 6x loading dye, mixed and spun down. The gel is 0.8% agarose in 1x TAE, 150 V, 45 minutes.

3.2.2 ES-62 coding genomic DNA fragment

A. viteae genomic DNA extracted from frozen parasites (Section 3.2.1.) was used as a template for amplification of the ES-62 gene. In order to amplify the whole ES-62 gene (introns and exons) from *A. viteae* genomic DNA, primers were designed based on the cDNA sequence data. The published untranslated regions included in the cDNA sequence were challenging for primer design as they were relatively short and very AT-rich with 5'UTR consisting of 90 base pairs (bp), 36% GC content and 3'UTR consisting of 137 bp, 23% GC content, followed by the polyadenylation chain. In order to amplify the ES-62 gene from the genomic DNA, primers were designed inside 5'UTR and 3' UTR. The PCR yielded a very faint band around 4000 bp long (undergraduate project, results not shown). This PCR product was then attached to a *C. elegans* promoter using the Hobert protocol for PCR fusion (Hobert 2002). The PCR- fusion yielded a strong band, around 4000 bp long. Sequencing of the band confirmed the presence of the ES-62 gene, but not that of the *C. elegans* promoter gene (undergraduate project, results not shown). Many reactions were performed in order to repeat the amplification of the 4000 bp fragment. However despite good DNA quality and concentration (Figure 3.3.), the use of different polymerase enzymes, primers and reaction conditions the result could not be repeated. As mentioned before during the undergraduate project, despite the lack of success of the fusion PCR, the re-amplification of the fragment led to increased yield, which later resulted in the idea to use a nested PCR to amplify ES-62.

A standard PCR is not sufficient for amplification of ES-62 gene from genomic DNA.

3.2.3 Comparison of the published cDNA sequence (Harnett) to *A. viteae* genome (sequenced by Blaxter lab)

The sequencing of *A. viteae* genome was performed by the Blaxter lab (University of Edinburgh), and the draft assembly was made available online in August 2012. Comparison of *A. viteae* published cDNA sequence (Harnett W et al., 1999) to *A. viteae* genome (sequenced by Blaxter lab) revealed

1240 nucleotides of the coding sequence that corresponded to 7 perfectly aligned exons and revealed the sequence of 6 introns, intron/ exon boundaries and 137 nucleotides of 3' UTR. Comparison of the cDNA to the genomic DNA confirmed the published sequence and revealed the rest of the 3'UTR structure. However, the 5'UTR and the N-terminal peptide from the cDNA did not align to the genomic DNA structure. Accordingly to the alignment, the 5'end of the cDNA (90 nucleotides of 5'UTR and 242 nucleotides of protein coding sequence) missing from the genomic sequence was divided into three smaller fragments (Figure 3.4). From the 5' end first fragment, 83 nucleotides of 5'UTR, aligned to the scaffold 0313. The second fragment, 107 nucleotides coding for the signal sequence and part of the mature end of the protein, did not align anywhere in the genomic sequence. The third fragment, 135 nucleotides aligned to scaffold 0582 in reverse orientation, with a short overlap in scaffold 0047. This is not unusual considering the genomic sequence data is from the draft assembly, and it is likely to be incomplete.

The large part of ES-62 cDNA aligned to the genomic fragment coding for ES-62 from *A. viteae* draft genome sequence. Three hundred and thirty three nucleotides of the 5' end of the cDNA sequence did not align to the genomic DNA in a continuous fashion, with parts missing from the genome. This could be due to the miss-alignment of the draft *A. viteae* genome or differences between genomic DNA and cDNA.

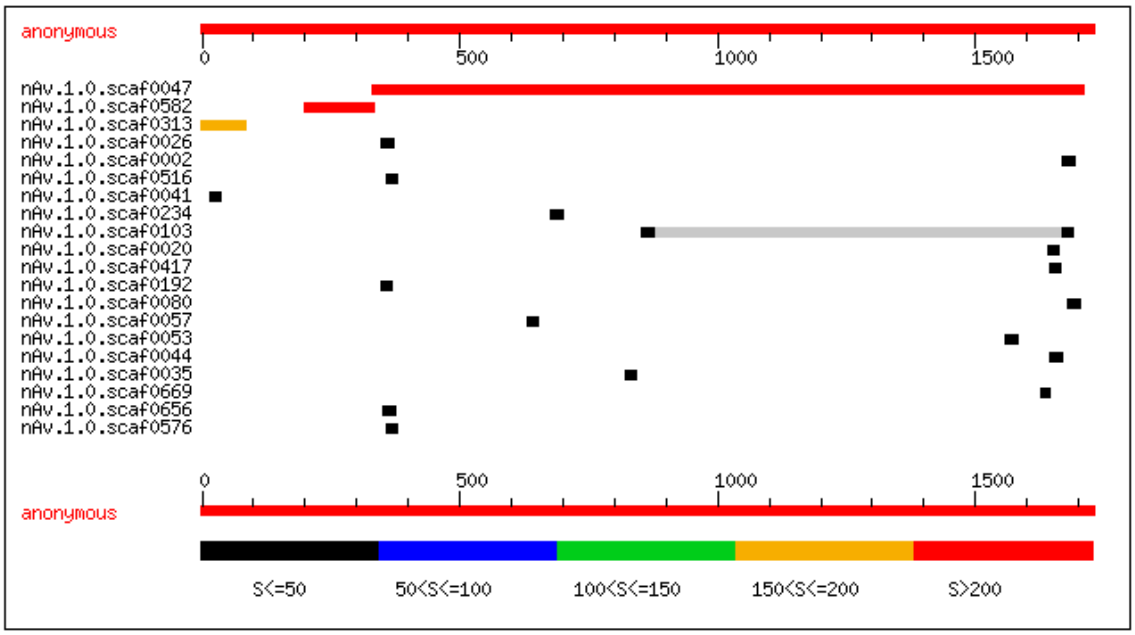


Figure 3.4. Schematic alignment of ES-62 cDNA sequence to *A. viteae* genome showing the distribution of the sequence between different scaffolds.

3.2.4 Amplification of the ES-62 genomic fragment that aligns to cDNA

Multiple PCRs were performed in order to amplify the full length genomic fragment using primers designed inside the 5'UTR (based on the cDNA sequence) and 3'UTR (based on both cDNA and genomic DNA sequences), which yielded no product (results not shown). Many attempts were made and a range of conditions used, with no effect. We hypothesised that ES-62 contains a large intron, missing from the genomic alignment, which makes the whole fragment difficult to amplify. Therefore, we decided to amplify the ES-62 gene in several parts, starting with the largest part that aligns to scaffold 00478.

First amplification of the genomic fragment that aligned to the cDNA (7 exons, 6 introns and 3'UTR) was unsuccessful (Figure 3.5 A). Genomic DNA of ES-62 is very AT-rich, with only 28% GC content. Parts of ES-62 DNA, especially introns, contain repetitive sequences and long strands of adenine and thymine nucleotides. If this is the case with the whole *A. viteae* genome, then a single set of primers may not be specific enough to amplify the desired gene. We attempted the re-amplification of the ES-62 fragment using a nested PCR method. The PCR product from the first amplification attempt was used as a template for the nested reaction. This meant that two pairs of PCR primers were used for the amplification of a single ES-62 fragment, which increased the specificity of the reaction. The nested PCR was successful and yielded a fragment approximately 3.7 kb long (Figure 3.5 B), which corresponded to the size of this fragment predicted by the genomic DNA sequence and the fragment obtained at the beginning of the investigation following the unsuccessful fusion reaction.

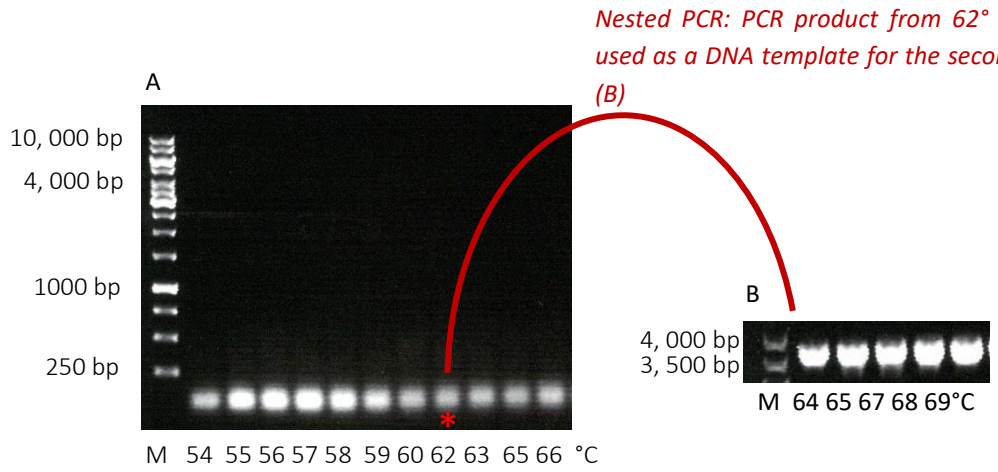


Figure 3.5. Electrophoresis picture of A) outer primers PCR and B) inner/ nested primers PCR to amplify ES-62 fragment from scaffold 00478. PCR reaction from A 62°C (indicated by red star) was used as a DNA template for the nested reaction (B). Temperature gradient was used to establish optimal annealing temperature for the PCR. Numbers under each lane indicate annealing temperature on the thermocycler. The gel is 1% agarose in 1x TAE, 100 V, 45 minutes

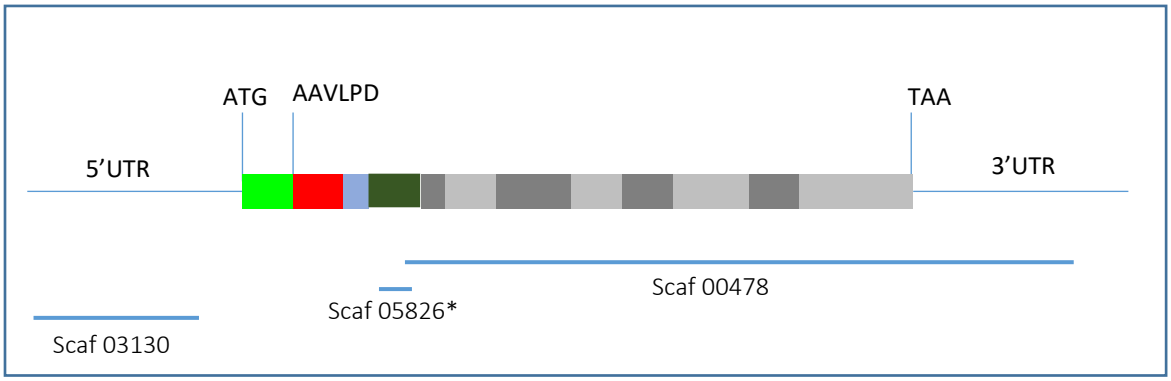


Figure 3.6. Alignment of ES-62 cDNA fragments to the *A. viteae* genomic DNA. ATG, the start of the mature protein (AAVL) and TAA are indicated. Different colours indicate parts that align to different scaffolds. Part of 5'UTR aligns to scaffold 03130, ATG, light green, red and light blue fragments do not align anywhere in the genomic sequence, the sequence following the signal peptide and before the start of first exon on scaffold 00478 (olive green fragment) aligns to scaffold 05826 in reverse orientation and different shades of grey indicate consecutive exons that align to scaffold 00478 of the genomic DNA.

A series of PCRs was performed attempting to bridge the gaps between sequences aligned to different scaffolds. Due to difficulties with amplifying the whole gene, it was hypothesised that exon 1 and intron 1 (5' end) are long and may be missing from the draft genome assembly. Figure 3.6 shows a schematic representation of cDNA divided into fragments accordingly to where they align in genomic DNA.

The sequence of the missing exon (olive green) and the first identified exon on scaffold 00478 (grey) have an overlap between scaffolds, however, they are differently orientated. The PCR with 5' primer designed inside the missing exon (olive green) and the 3' primer inside the first identified exon (grey) yielded one band significantly stronger than the rest (Figure 3.7). This band was purified, cloned and sequenced. The forward sequencing result aligned with the genomic DNA on scaffold 05826 (olive green region) in reverse orientation. After the alignment to scaffold 05826 had stopped, the sequencing result aligned with some sequence on the scaffold 00478 in the correct orientation, however, the result did not reach the sequence of the first identified exon (grey part, start of an exon alignment of cDNA to genomic DNA). The reverse sequencing result was aligned to the first exon on the 00478 scaffold (grey) and when it reached the 5' end of this exon, it covered 11bp on scaffold 05826. This sequencing result confirmed the information obtained by the alignment of cDNA and genomic DNA and connected the two scaffolds, but not different parts of the cDNA sequence.

The 5' cDNA fragments; ATG start of the first exon (light green), signal sequence (red), and the first exon (light blue) were not found in the genomic DNA. The PCR between the ATG start of the first intron (light green) and the 3' end of the first exon (olive green) yielded three distinct bands. They were all gel- purified and sequenced. The sequencing product was very short, and yielded sequence information only with the reverse primer for each band. It aligned with reverse 3' end of the first exon sequence (olive green) on scaffold 005826 and then with correct orientation to scaffold 005826.

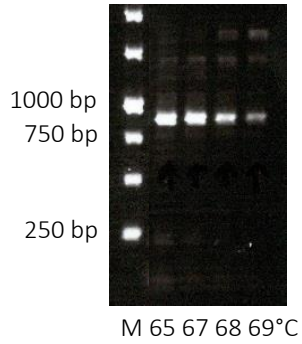


Figure 3.7. Electrophoresis of the PCR product between olive green and grey fragments (scaffold 05826 and scaffold 00478). A temperature gradient was used to establish optimal annealing temperature for the PCR. Numbers under each lane indicate annealing temperature on the thermocycler. The gel is 1% agarose in 1x TAE, 100 V, 45 minutes.

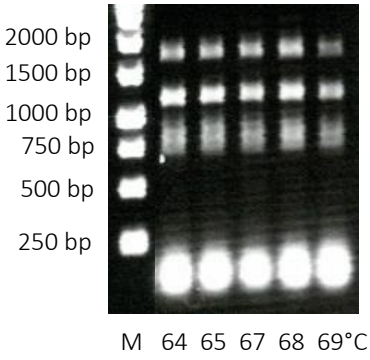


Figure 3.8. Electrophoresis of the result of a PCR between light green and olive fragments (within unidentified exon). A temperature gradient was used to establish optimal annealing temperature for the PCR. Numbers under each lane indicate annealing temperature on the thermocycler. The gel is 1% agarose in 1x TAE, 100 V, 45 minutes.

Reactions were also performed between 5' part of the unidentified exon (light green) and 3' part of the unidentified exon (olive) as well as the signal sequence within the unidentified exon (red) and the start of first identified exon on scaffold 00478 (grey). Both reactions yielded very short products, below 300 bp, and despite successful cloning did not yield any usable sequencing information. All of the reactions are summarized in Figure 3.9.

ES-62 cDNA sequence aligns to the *A. viteae* draft genome, but only partially. Attempts to amplify the whole genomic fragment failed. The part that aligns on scaffold 00478 can be amplified from the genomic DNA using a nested PCR and yields a 3.7 kb fragment with 7 exons, 6 introns and 3'UTR. The part that does not align continuously to the genomic DNA was divided into smaller fragments and some of them were amplified. Sequencing of those results failed to bridge the gaps between the fragments. This suggests that genomic ES-62 may contain a large intron and that the draft genome assembly may be incomplete.

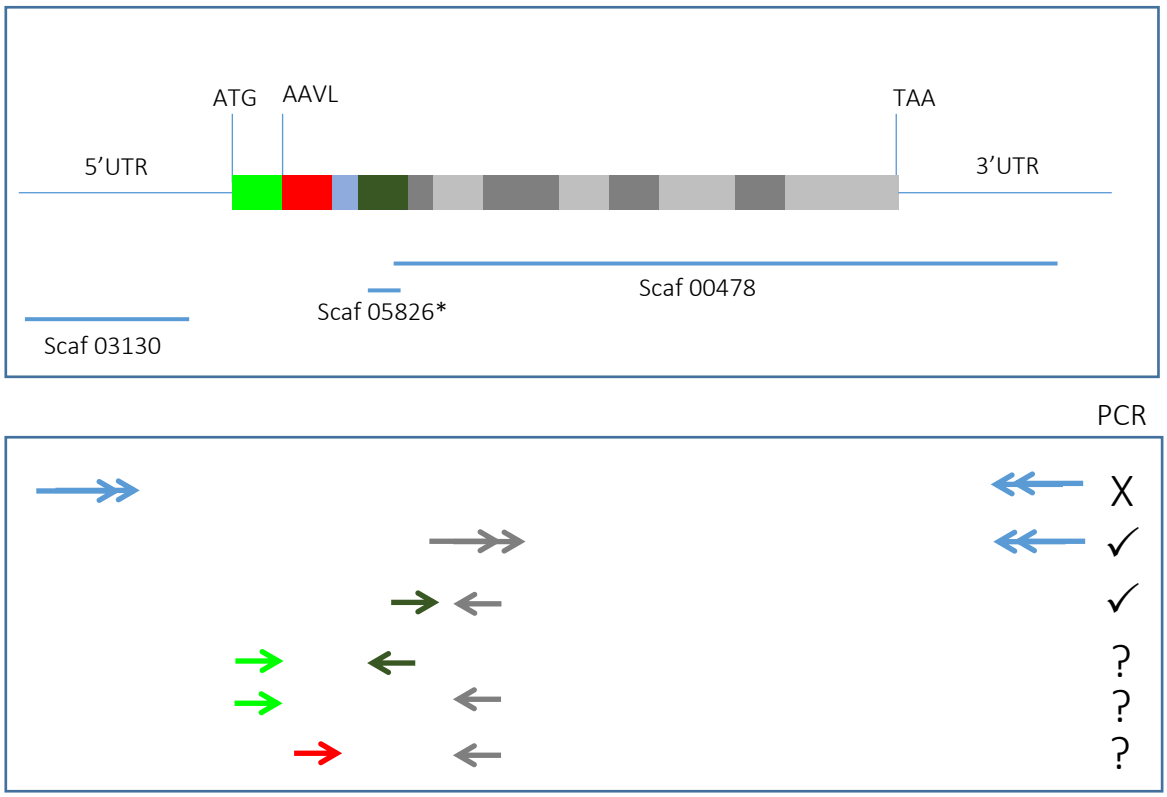


Figure 3.9. Summary of the PCR reactions performed to characterise the 5'end of the ES-62 genomic fragment. *Arrows indicate positions of primers for each reaction. Double arrows indicate nested PCR approach. X means there was no product.* '✓' means there was a product and it was sequenced. '?' means there was a product or multiple products but the sequencing was unsuccessful. The '*' indicates that the sequence is in reverse orientation.

The second assembly of ES-62 was performed in August 2015 and confirmed the presence of a large intron, missing from the draft genome assembly (per. Comm. Georgios Koutsovoulos, Blaxter lab). The ES-62 gene is approximately 10884 bp long; this size is approximate as there are still stretches of unknown nucleotides within introns. Based on the genomic sequence information and the sequencing of individual introns those unknown fragments are likely to be stretches of AT repeats. ES-62 consists of 10 exons and 9 introns (Figure 3.10.). The first exon is 111 nucleotides long and codes for the signal peptide and part of the mature protein. The first intron is 371 nucleotides long. The second exon is 131 nucleotides long. The second intron is approximately 6796 nucleotides long and is likely to be the reason of unsuccessful amplification attempts between ES-62 alignment on scaffold 00478 (grey) and 5' fragments (olive, red and green). The following exons are respectively 116, 101, 174, 159, 162, 180, 79 and 269 nucleotides long. Approximate sizes of the introns are respectively 520, 174, 233, 267, 290, 226 and 432 nucleotides long.

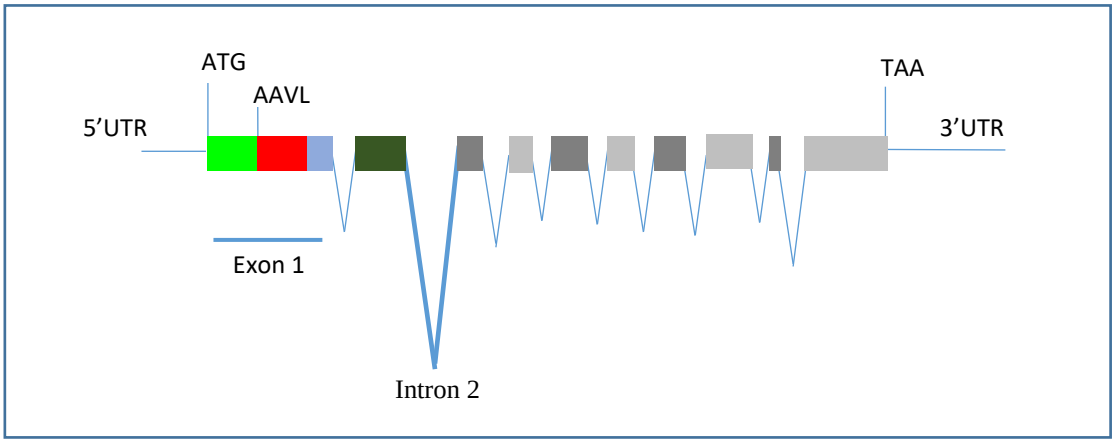


Figure 3.10. Schematic diagram of the full ES-62 genomic DNA. Blocks indicate positions of exons, lines indicate position of introns. Colors correspond to Figures 3.6 and 3.9 and indicate positions of ES-62 cDNA parts in the new alignment.

3.3 Making transgenic *C. elegans*

3.3.1 Preparation of DNA fusion products for microinjections

Control elements from the *C. elegans* gene coding for ASP-6 protein were used to direct expression of ES-62 in *C. elegans*. We hypothesised that under control of *asp-6* recombinant ES-62 protein will be secreted into the media.

The first strategy was to use ES-62 genomic DNA due to the gene- specific regulatory elements which may be present within introns. Furthermore, introns are important for stabilising the mRNAs and the pre-mRNA splicing promotes nuclear export and translation (Maniatis & Reed 2002; Nott et al., 2004). At the beginning of this study the 5'end of the genomic ES-62 fragment was unknown. Due to time constrains only the characterised part of the ES-62 genomic DNA was used in the fusion construct.

The second strategy was to use ES-62 cDNA, which codes for the whole mature protein. Sequencing of the only available cDNA clone (received from University of Nottingham) revealed that it contained 6 mutations, two of which were single base pair deletions and caused the sequence to stop prematurely. The cDNA required fixing of the mutations back to the original sequence before it could be used (cDNA repair was performed by Dr Sushmita Maitra).

Promoter, 5'UTR and signal sequence from *C. elegans asp-6* were fused to a part of ES-62 genomic DNA or cDNA using PCR. The 3'UTR was either taken from the *asp-6* gene (for ES-62 cDNA fragment) or from the ES-62 gene (for ES-62 genomic construct) (Figure 3.11).

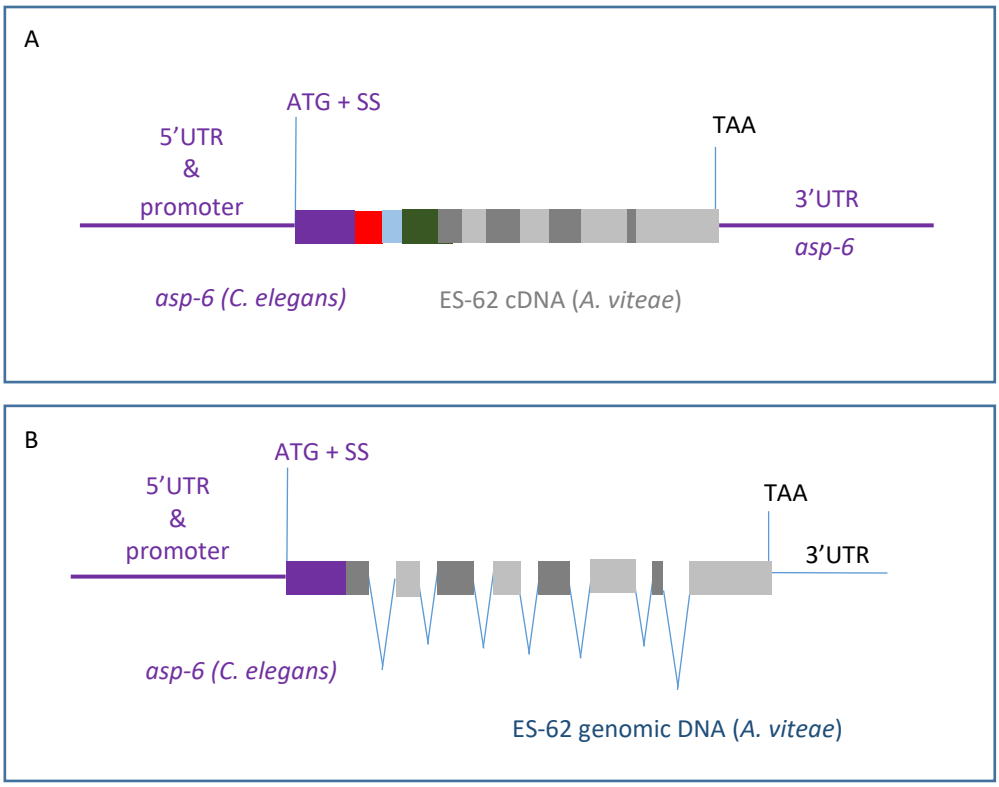


Figure 3.11. Schematic diagram of the *asp-6::*ES-62 fusion constructs. A) The *asp-6*::ES-62 cDNA construct consists of 5'UTR, promoter, ATG and signal sequence from *C. elegans* asp-6 gene fused to ES-62 cDNA from *A. viteae* and to 3'UTR from *C. elegans* asp-6 gene. B) The *asp-6*::genomic ES-62 construct consist of 5'UTR, promoter, ATG and signal sequence from *C. elegans* asp-6 gene fused to ES-62 genomic DNA fragment from *A. viteae* with ES-62 3'UTR. *asp-6*::genomic ES-62 construct contains a large part of genomic ES-62 fragment and not the whole gene.

In order to prepare DNA fusion products for microinjections, PCR fusions were performed using Phusion high-fidelity polymerase. Using high fidelity polymerase decreases the chances of mismatches occurring during the reaction. The promoter and ES-62 fragments were amplified, mixed and subjected to the 'primer exhausting' reaction in order to stop the first set of primers used to amplify individual fragments from interfering with the fusion reaction (Section 2.9.). This was done due to previous problems with performing fusion, which resulted in re-amplification of individual fragments rather than connecting them together. The 'primer exhausting' reaction was carried out before the fusion reaction and contained all ingredients needed for a PCR apart from the second set of primers (inner primers). The 'primer exhausting' reaction was then used as a DNA template to perform the fusion.

PCR to amplify the *asp-6* promoter fragment yielded a single band of around 1.8 kb size (Figure 3.12 A). The fusion fragment was contaminated with un-fused promoter fragments and short fragments below 500 bp (Figure 3.12 D). Those are known to aid the transformation of *C. elegans* and therefore the whole, un-purified fusion PCR product was used for injecting the worms. The correct fusion with no mismatches and the presence of both *C. elegans* promoter and ES-62 gene were confirmed by sequencing of purified fusion bands.

The *C. elegans* promoter DNA fragment can be attached to *A. viteae* ES-62 genomic DNA fragment to create an in-frame 2-way fusion product for *C. elegans* transformation. The fusion products created were: *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 (genomic DNA fragment + 3'UTR) and *asp-6* (5'UTR + promoter + signal sequence) :: ES-62 (cDNA coding for mature protein) :: *asp-6* 3'UTR (cDNA construct was made by Dr Sushmita Maitra).

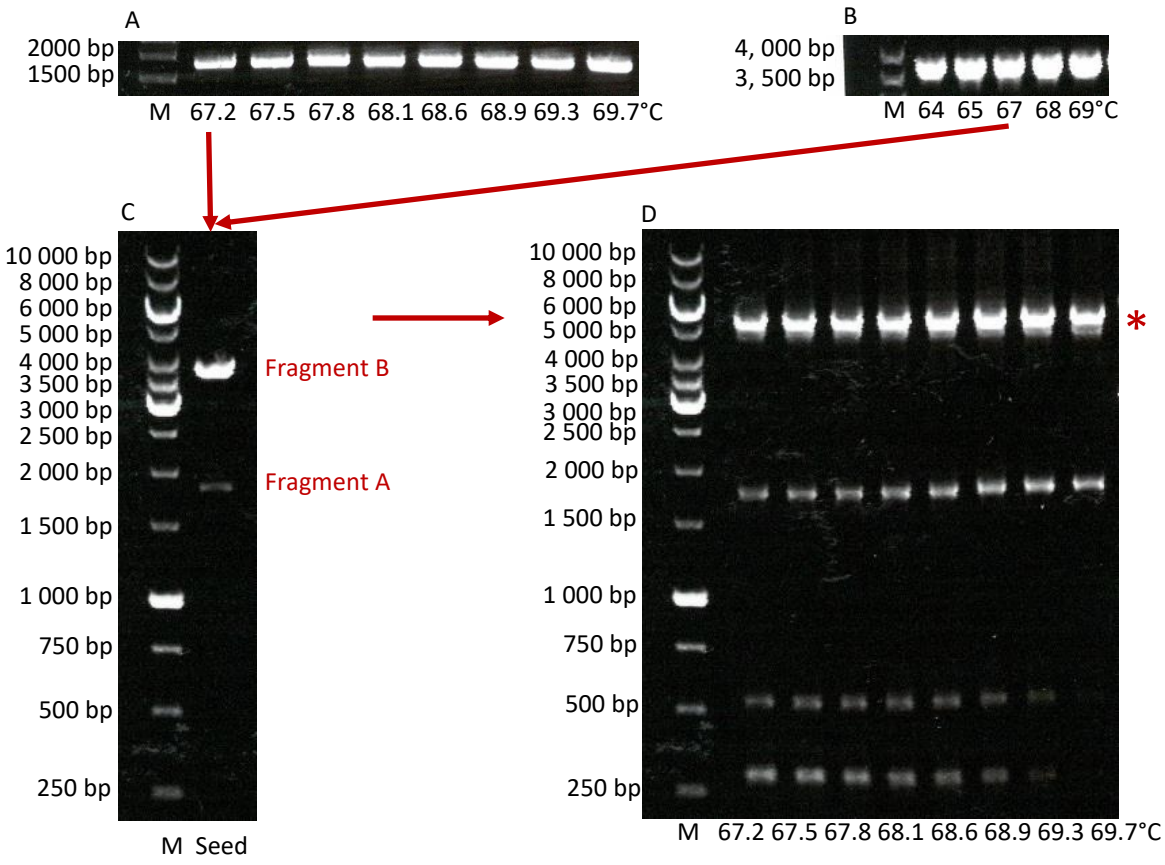


Figure 3.12. Electrophoresis of the fusion reaction between *asp-6* and genomic ES-62 fragment. A) *C. elegans asp-6* 5' UTR, promoter and signal sequence fragment B) nested PCR amplified *A. viteae* genomic ES-62 fragment from scaffold 00478. Fragments A and B were combined and subjected to a primer exhausting reaction. C) Seed- combined fragments A and B after the primer exhausting reaction. D) The seed (C) PCR was then used as a DNA template for the fusion reaction with 5' primer inside *asp-6* 5'UTR fragment (A) and 3' primer inside ES-62 3' UTR (B). A + B fusion product- *asp-6* 5'UTR, promoter, signal sequence + ES-62 genomic fragment, ES-62 3'UTR is indicated with a red star.

A temperature gradient was used to establish optimal annealing temperature for the PCR. Numbers under each lane indicate annealing temperature on the thermocycler. The gel is 1% agarose in 1x TAE, 100 V, 45 minutes.

3.3.2 Screening injected worms for expression of the transgene

Each DNA fusion product was injected into Δ*lin-15* mutant *C. elegans* together with the *lin-15* rescue plasmid. *Δlin-15* mutants are temperature sensitive and express multiple vulva phenotypes when kept at 20°C or higher. Worms that received the rescue plasmid expressed a wild-type phenotype when kept at 20°C or higher. We hypothesised that *C. elegans* rescued from *Δlin-15* mutation (wild type at 20 °C) also received the injected fusion product and expressed the ES-62 gene under control of the *C. elegans* promoter. However, after microinjection the injected genes form random extrachromosomal arrays, so in order to find transgenic lines the worms were subjected to three rounds of screening. For the initial screening injected mothers were kept at 20°C, allowed to reproduce and carefully monitored for the first set of progeny (F1 generation). When the F1 progeny reached the adult stage, worms which expressed a wild-type phenotype were separated and allowed to grow and reproduce. Those worms that produced wild- type progeny (passed the *lin-15* rescue plasmid on to the next generation) were considered a stable line. For the final screening, individual worms from each line were picked, lysed and analyzed by single-worm PCR in order to check for the presence of the injected product. The negative control was a WT *C. elegans* DNA, and a positive control was a fusion fragment used for the microinjection. PCR reactions were performed using Taq green polymerase and aimed to amplify the fragment of the injected gene across the fusion junction. All of the transgenic lines yielded the expected- sized product (Figure 3.13), which was later purified and confirmed by sequencing.

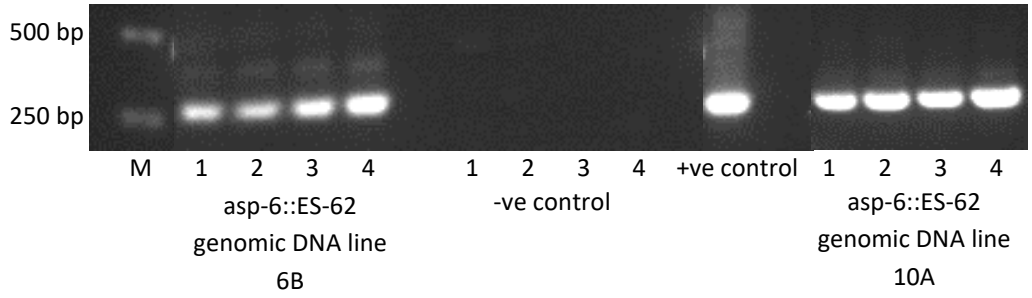


Figure 3.13. Electrophoresis of single-worm PCR to screen for transgenic lines. Numbers indicate repeats. DNA from Δ*lin-15* mutant worms rescued with *lin-15* plasmid was used as a template for the negative control. Fusion PCR product (asp-6::ES-62 genomic DNA) was used as a template for the positive control. The fragments were expected to be 308 bp long. M is a DNA hyperladder. The gel is 1% agarose in 1x TAE, 100 V, 45 minutes.

The progeny of injected worms was screened for the presence of the transgene. The results suggest that stable transgenic lines were created. The progeny (F1) of injected worms that show a wild-type phenotype at 20°C and produced wild-type progeny (F2) at 25°C was assumed to be a transgenic line. An additional third screening step was added to check for the presence of the transgene in the DNA of randomly selected F2 worms. Multiple PCRs were performed on DNA from individually picked worms, with primers spanning the fusion junction of injected transgene. All of the PCRs yielded the desired product suggesting that the stable lines indeed contained the transgene.

3.4 Expression of recombinant ES-62 protein in transgenic *C. elegans*

3.4.1 Identification of recombinant ES-62 in transgenic *C. elegans* lysates (Western Blot analysis of small scale plate cultures)

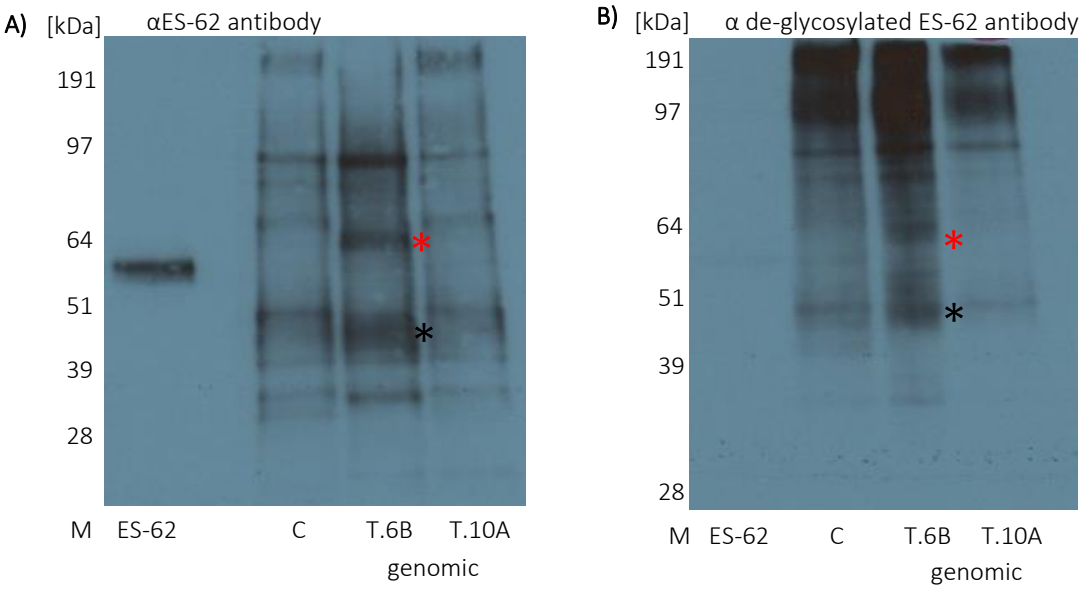
Three transgenic lines were used for the purpose of this study. One line made with (*asp-6 5'UTR+promoter+signal sequence*::ES-62 cDNA::*asp-6* 3'UTR)- IIL (cDNA construct and transgenic worms were supplied by Dr Sushmita Maitra) and two independent lines with (*asp-6 5'UTR+promoter+signal sequence*::ES-62 genomic DNA fragment+3'UTR)- 6B and 10A. Individual transgenic lines were maintained in plate cultures. Between four to fifty plates were prepared for each analysis. The preparation of samples is fully described in Materials and Methods (Sections 2.3. and 2.4.). Worms were allowed to swim in M9 phosphate buffer for 4 hours with shaking. Media were separated from worm pellets by sedimentation (Section 2.4.1.). Media for each transgenic worm were analysed by silver staining of SDS gels. There were, however, no detectable differences between the control and the transgenic worm media, which could be due to the very low concentration or lack of the secreted transgenic protein (results not shown). Moreover, the complexity of the proteins was higher than previously anticipated, which also dramatically affected the silver stain profile, making it challenging to distinguish between different bands. 1D gels separate proteins only accordingly to their size; this allows proteins of very similar size to travel together, and the less abundant proteins are covered by the more prominent ones.

Media and pellets were sent to our collaborators at Strathclyde University. Dr Kara Bell lysed worm pellets and concentrated both media and pellets using Amicon centrifugal filters. Concentrated media and concentrated worm lysates were analysed for the presence of recombinant ES-62 using western blotting with three antibodies used to detect the parasite- derived ES-62 (Dr Kara Bell, where indicated). The rabbit α ES-62 serum antibody (raised against the whole native protein) is a polyclonal antibody that recognizes the protein and possibly PC epitopes on ES-62. Rabbit α de-

glycosylated ES-62 is directed against the protein backbone of ES-62, as the *N*-glycans and PC – *N*-glycans were cleaved before immunization. The deglycosylation is carried out under denaturing conditions and, therefore, the antiserum may primarily recognize sequential rather than conformational epitopes: this may be particularly useful if recombinant ES-62 is folded differently in transgenic *C. elegans*. TEPC 15 is an IgA myeloma protein that specifically binds to phosphorylcholine (PC) (SIGMA).

Transgenic worm lysates from sample 6B (*asp-6*::genomic ES-62 fragment) reacted with all three of the antibodies, each time yielding a band relative to *A. viteae* ES-62 size (Figure 3.14 red stars, Dr Kara Bell). Two antibodies raised against parasite- derived ES-62 detected a 46 kDa band, also in 6B worm lysates. None of the secreted media samples showed any difference between the transgenic and the control (results not shown). Antibodies detected two protein bands, that could be the recombinant ES-62 in the concentrated worm lysates of one genomic DNA transgenic (6B). However, it was not detected in the second genomic DNA transgenic (10A), nor in the respective secreted media from 6B worms. Since the 10A transgenic line arose independently, a spontaneous mutation could have occurred, which resulted in miss-folding of the recombinant protein in one of the transgenic lines, but not the other.

In order to check for sequence differences in those two lines a series of sequencing reactions was performed (PCR reactions amplifying short, overlapping fragments were performed by Kyle Main-project student), however, no sequence differences were detected. We hypothesized that the difference between 6B and 10A worms arose due to different levels of protein expression in those independent lines, and the differences were further investigated.



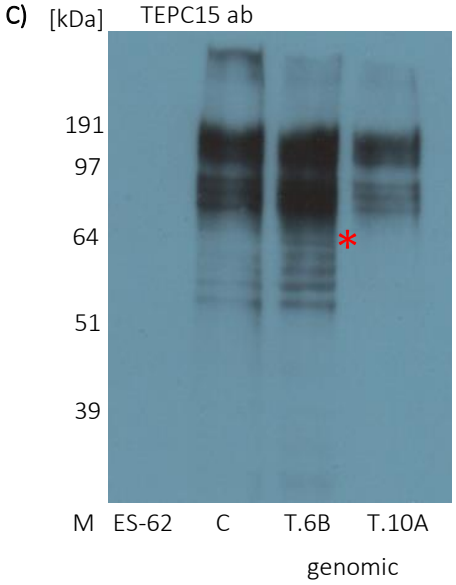


Figure 3.14. Western blot of concentrated worm lysates after isolation from plate culture. Samples: T.6B- 6B *asp-6* ES-62 genomic fragment transgenic line; T.10A- 10A *asp-6* ES-62 genomic fragment transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Red stars indicate a ~70 kDa band, black stars indicate a ~46 kDa band present in 6B genomic DNA transgenic line, but not in the 10A or control lines. Samples were probed with: A- rabbit α ES-62 serum antibody used at 1:500 in 5% BSA in TBS, 0.1%tween, B- rabbit α ES-62de-glycosylated antibody used at 1:200 in 5% milk in TBS, 0.1%tween, C- TEPC15 antibody used at 1:200 in 5% BSA in TBS, 0.1%tween (All of the samples were prepared by the author of this thesis, sample concentration and Western Blot analysis was performed by Dr Kara Bell).

3.4.2 Identification of a recombinant ES-62 breakdown product in transgenic *C. elegans* media (large scale liquid culture)

In order to increase the amount of material and the concentration, *C. elegans* were grown in liquid media. The first liquid culture experiment was performed with the [*asp-6* 5'UTR+promoter+signal sequence::ES-62 genomic DNA fragment+3'UTR] 6B transgenic worms, [*asp-6* 5'UTR+promoter::ES-62 cDNA::*asp-6* 3'UTR] IIL transgenic worms and the control [Δ*lin-15* rescued line]. Worms were grown in liquid culture (150 ml) for five to six days. Worms were separated from the growth media by sedimentation, washed and allowed to secrete into M9 buffer for 4 hours. Media and worms were separated as described in materials and methods (Section 2.5.1.). Media was concentrated using Amicon ultracentrifuge filters and analysed. Worm pellets were lysed, concentrated using Amicon ultracentrifuge filters and analysed alongside the media samples.

Experiments revealed the presence of a 27 kDa band on western blots. Rabbit α ES-62 serum antibody detected a 27 kDa polypeptide in the media (Figure 3.15) and worm lysates (Figure 3.16) from the 6B genomic ES-62 line, which was not present in the IIL ES-62 cDNA line or the control.

The 27 kDa polypeptide was detected by the rabbit α ES-62 serum antibody in 6B genomic ES-62 lines (media and lysates), which has not been detected in IIL cDNA ES-62 lines or the control. This may be a breakdown product of recombinant ES-62, which could be less stable than the parasite-derived protein. Rabbit α ES-62 serum antibody has a wide detection spectrum. It may recognise PC epitopes. The multiple polypeptides detected in the samples may reflect PC-containing proteins. The samples were tested with α de-glycosylated ES-62 antibody.

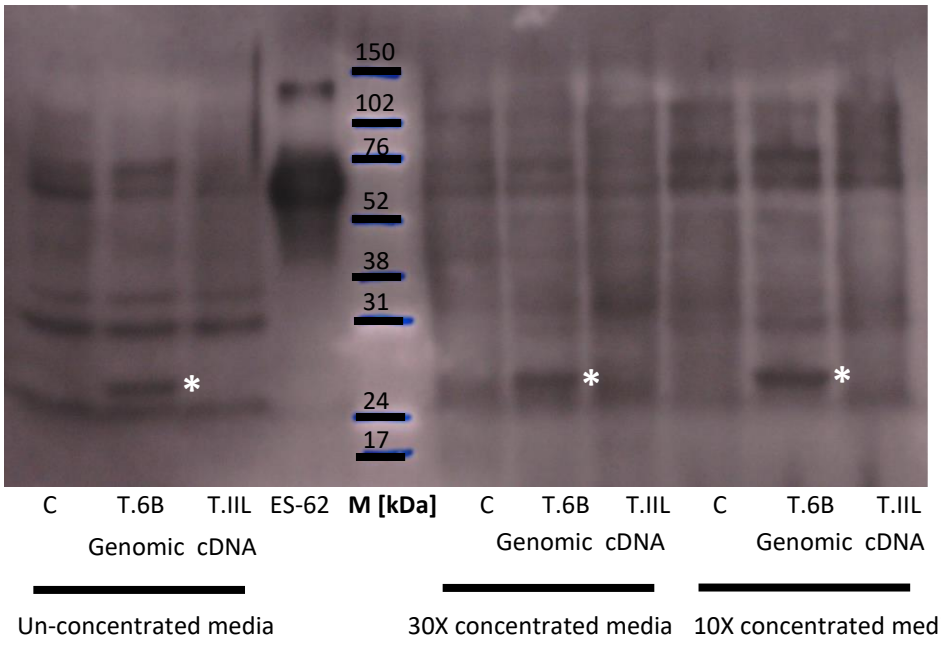


Figure 3.15. Western blot of worm secreted media after isolation from liquid culture (rabbit α ES-62 antibody). Samples: T.6B- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL- cDNA- IIL *asp-6* ES-62 cDNA transgenic line; control- *Δlin-15* mutants rescued with *lin-15* plasmid. Samples were concentrated using Amicon filters as described in materials and methods. The white star indicates a band of ~27kDa present in genomic DNA samples, but not in the cDNA or control lines. Samples were probed with rabbit α ES-62 serum antibody used at 1:500 in 5% BSA in TBS, 0.1%tween.

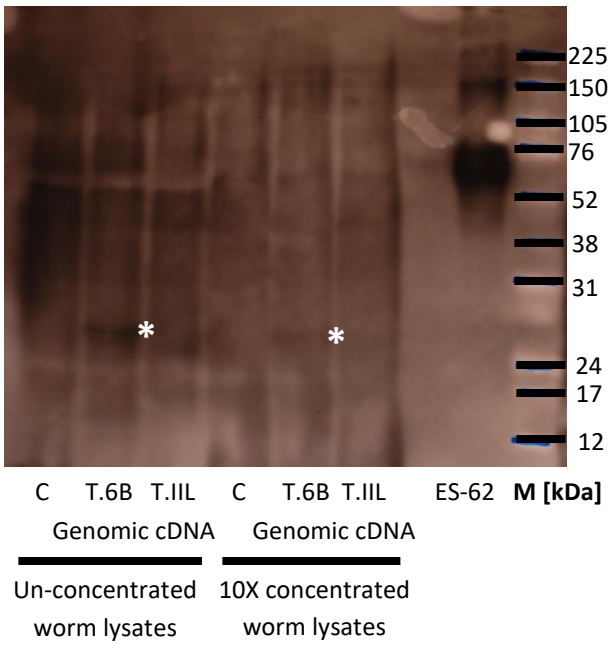


Figure 3.16. Western blot of worm lysates after isolation from liquid culture (rabbit α ES-62 antibody). Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL cDNA- IIL *asp-6* ES-62 cDNA transgenic line; C- control *Δlin-15* mutants rescued with *lin-15* plasmid. White star indicates a band ~27kDa present in genomic DNA samples, but not in the cDNA or control lines. Samples were probed with rabbit α ES-62 serum antibody used at 1:500 in 5% BSA in TBS, 0.1%tween.

In order to further investigate the nature of the 27 kDa polypeptide detected in genomic ES-62 transgenic line, a second antibody was used. Rabbit α ES-62 de-glycosylated antibody is directed against the protein backbone of ES-62 as the *N*-glycans and PC– *N*-glycans were cleaved before immunization. Also, deglycosylation is carried out under denaturing conditions, therefore, the antiserum may primarily recognize sequential rather than conformational epitopes: this may be particularly useful if ES-62 is folded differently by *C. elegans*.

The 27 kDa polypeptide was detected, however, this time in the secreted media from IIL ES-62 cDNA line, not in the ES-62 genomic line or the control (Figure 3.17 and 3.18).

The rabbit α ES-62 antibody recognized high MW bands (70 kDa) in the genomic and cDNA media samples (Figure 3.18 B red star). The bands appeared to be of the same size in both media samples from transgenic lines and were not seen in the control sample media. The 70 kDa band was similar to that detected previously in 6B genomic ES-62 transgenic worm pellets (Figure 3.14, Dr Kara Bell). However, the ES-62 cDNA and the genomic ES-62 transgenics were predicted to produce different recombinant ES-62 products. The genomic DNA line carries only a part of ES-62 (412 amino acids) and its predicted molecular weight is 46 kDa, if unmodified. The cDNA line carries the whole mature protein (474 amino acids), without the signal sequence, and its predicted molecular weight is 52.8 kDa without modifications. The 27 kDa polypeptide was recognised by the rabbit α ES-62 antibody in 6B genomic ES-62 transgenic media (Figures 3.15. and 3.18. B) and worm lysates (Figure 3.16) and could be a breakdown product of recombinant ES-62. The rabbit α de-glycosylated ES-62 antibody does not recognise the 27 kDa polypeptide in 6B genomic ES-62 samples. Interestingly, the last antibody recognises a 27 kDa polypeptide in the IIL ES-62 cDNA secreted media (Figures 3.17 and 3.18 A).

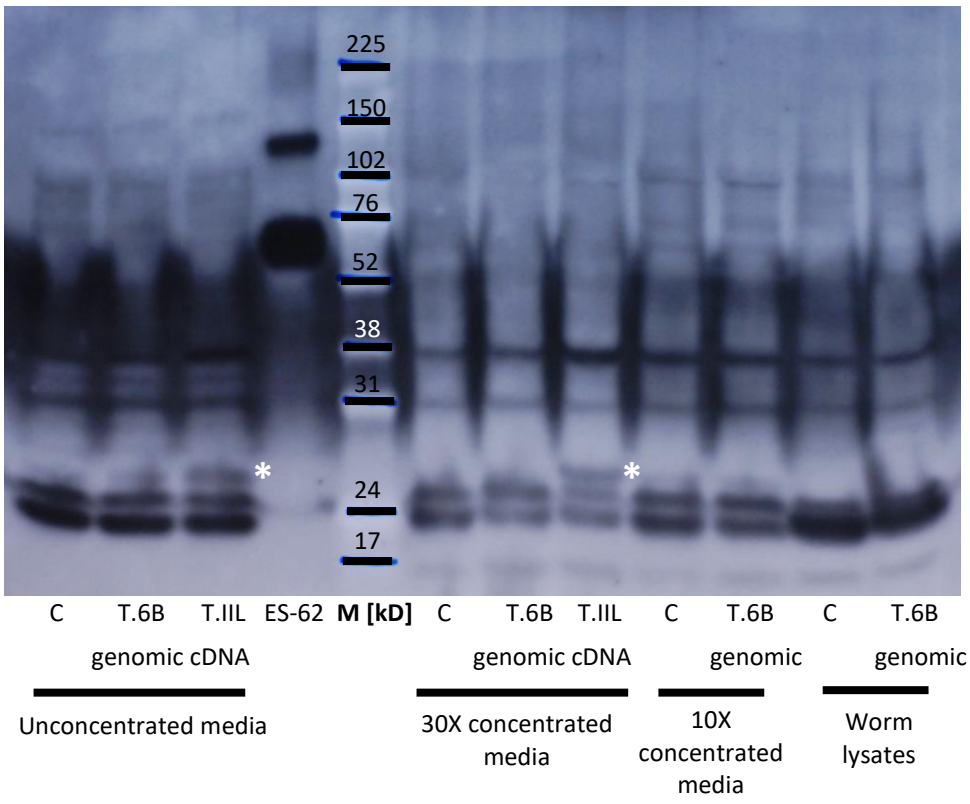


Figure 3.17. Western blot of worm secreted media and worm lysates after isolation from liquid culture (rabbit α de-glycosylated ES-62 antibody). Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL cDNA- IIL *asp-6* ES-62 cDNA transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Samples were concentrated using Amicon filters as described in materials and methods. White star indicates a band ~27kDa present in ES-62 cDNA transgenic samples, but not in the genomic ES-62 or control lines. Samples were probed with rabbit α ES-62 de-glycosylated antibody used at 1:200 in 5% milk in TBS, 0.1%tween.

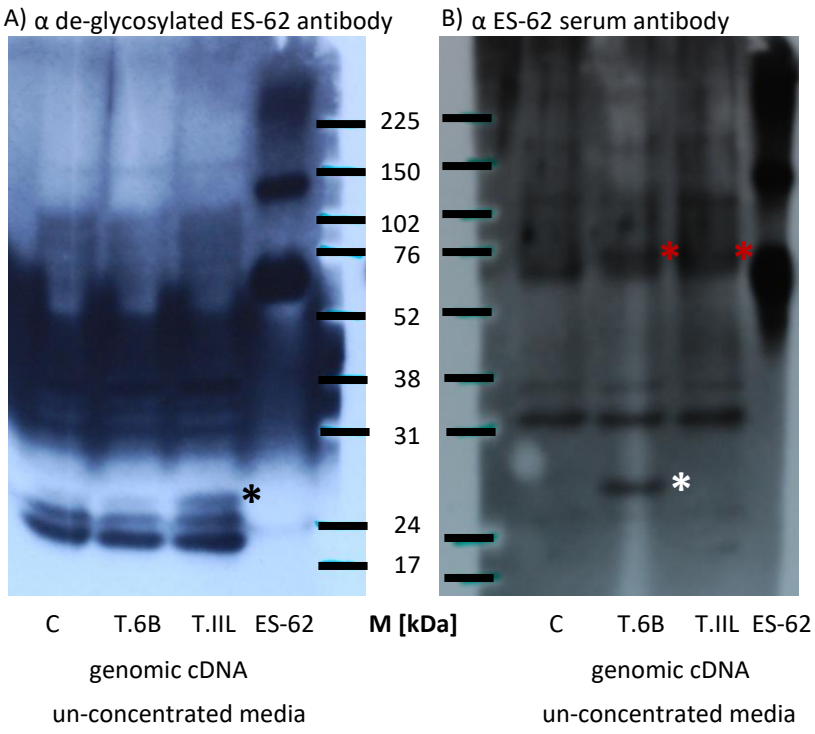


Figure 3.18. Western blots of un-concentrated worm secreted media 45 days after isolation from liquid culture. Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL cDNA- IIL *asp-6* ES-62 cDNA transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. A) Black star indicates a band ~27kDa present in secreted media from ES-62 cDNA transgenic lines, but not in the genomic ES-62 or control lines, detected by α ES-62 de-glycosylated antibody used at 1:1000 in 5% milk in TBS, 0.1%tween. B) White star indicates a band ~27kDa present in secreted media from genomic ES-62 transgenic lines, but not in the ES-62 cDNA or control lines, detected by α ES-62 serum antibody used at 1:1000 in 5% BSA in TBS, 0.1%tween. Red stars indicate a higher MW band ~70kDa recognised by α ES-62 serum antibody in transgenic media samples (genomic ES-62 and ES-62 cDNA) and not in the control.

3.4.3 Identification of the recombinant ES-62 in transgenic *C. elegans* lysates (Silver stain analysis of large scale liquid cultures)

Silver staining was chosen as an initial method of analysis of transgenic worm media. Results showed that *C. elegans* secrete a simple mixture of proteins that can be easily identified and distinguished on silver- stained SDS-PAGE gel (Figures 3.1. and 3.2.). First results of analysis of *C. elegans* secreted media confirmed that silver stain was sensitive enough to identify ASP-6 in the worm media. We expected to be able to see a difference in the SDS-PAGE profile of transgenic worms compared to the control. Using the silver staining method would allow us to excise this protein band and confirm its identity by MALDI- TOF mass spectrometry.

Originally transgenic worms were maintained in plate cultures. Between 4- 50 plates were prepared for each analysis of secreted proteins. Secreted worm media from each transgenic line were analysed by SDS-PAGE and silver staining. There were, however, no detectable differences between the control and the transgenic worm media, which could be due to the very low concentration or lack of the secreted transgenic protein (results not shown). Moreover, detected protein mixtures were more complex than previously anticipated. 1D gels separate proteins only accordingly to their size. There is a high chance that proteins of very similar size will run together, and the less abundant ones will be 'masked' by the more prominent bands.

Using large scale liquid culture for growing transgenic *C. elegans* increased the yield and allowed for the concentration of the product. We hypothesised that the silver stain of those concentrated proteins would allow better resolution and, therefore, better detection of transgenic proteins. Worm lysates were also prepared and concentrated in a similar manner to the secreted worm media. The concentration steps were performed using MW cut- off columns as described in Sections 2.15 and 2.14.

The liquid culture experiment was performed for the 6B asp-6::ES-62 genomic DNA and the IIL asp-6::ES-62 cDNA transgenic lines. The 6B transgenic worm lysates previously reacted with three antibodies were used to detect parasite- derived ES-62 (Figure 3.14, Dr Kara Bell).

The silver stain profile of *C. elegans* secreted media was complex. The MW cut-off and concentration did not show expected differences between concentrated and un- concentrated samples. Despite the use of a size- exclusion concentration method, lower MW polypeptides were detected. The silver stain of the secreted worm media did not show any difference between transgenic lines and the control (Figure 3.19).

The silver stain analysis was repeated for the 6B ES-62 transgenic *C. elegans* secreted media. The flow-through from the Amicon colums was analysed in order to investigate if the concentration step was working properly. A 46 kDa band was identified in the 6B ES-62 transgenic secreted media sample (Figure 3.20). A very faint high MW band was identified in the flow- through. No low molecular weight bands were retained by the column, which suggests that the size exclusion step was not working.

A 46 kDa band was also recognised in the 6B transgenic *C. elegans* lysates (Figure 3.21), this could be the truncated recombinant ES-62 that did not undergo any modifications. A 35 kDa band was identified in both genomic DNA and cDNA transgenic *C. elegans* lysates, which could be a breakdown product of recombinant ES-62.

The silver stain detected a 46 kDa band in the concentrated media (Figure 3.20) and concentrated lysates (Figure 3.21) from the 6B transgenic *C. elegans*. A similar band was previously identified by two antibodies raised against the parasite- derived ES-62, however only in 6B transgenic worm lysates and not in the media (Figure 3.14, Dr Kara Bell). 46 kDa is the predicted MW for truncated recombinant protein produced by genomic ES-62 transgenics. The results suggest that the

recombinant ES-62 is expressed in 6B transgenic *C. elegans*, but may not be secreted or is secreted in very low quantity.

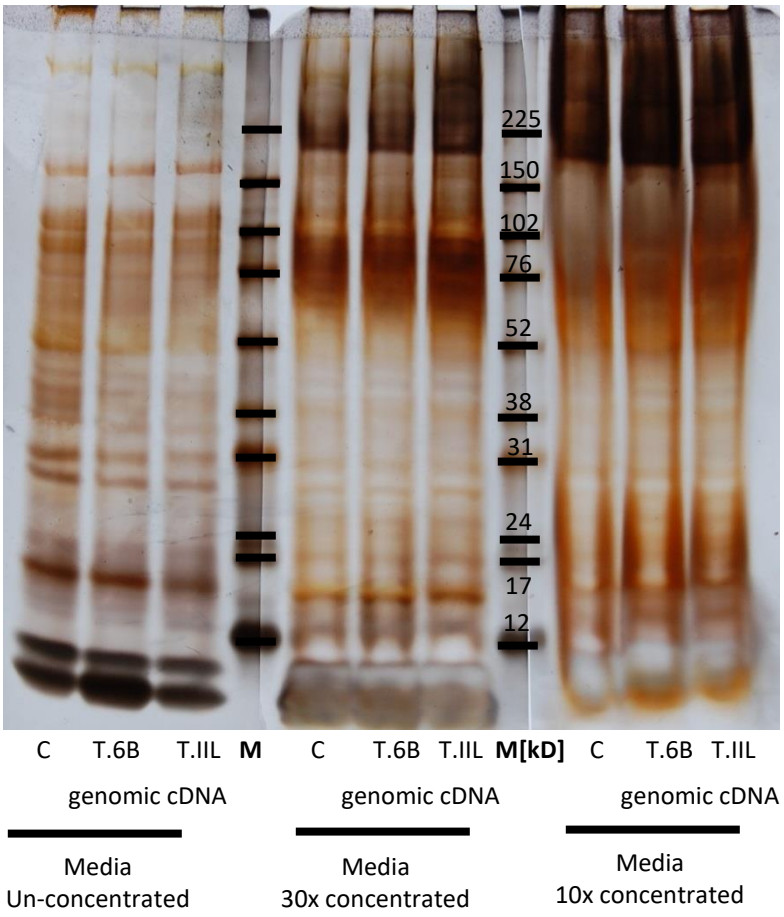


Figure 3.19. Silver stain of worm secreted media after isolation from liquid culture genomic ES-62 vs ES-62 cDNA transgenic lines. Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL cDNA- IIL *asp-6* ES-62 cDNA transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Samples were concentrated using Amicon filters (Section 2.15.)

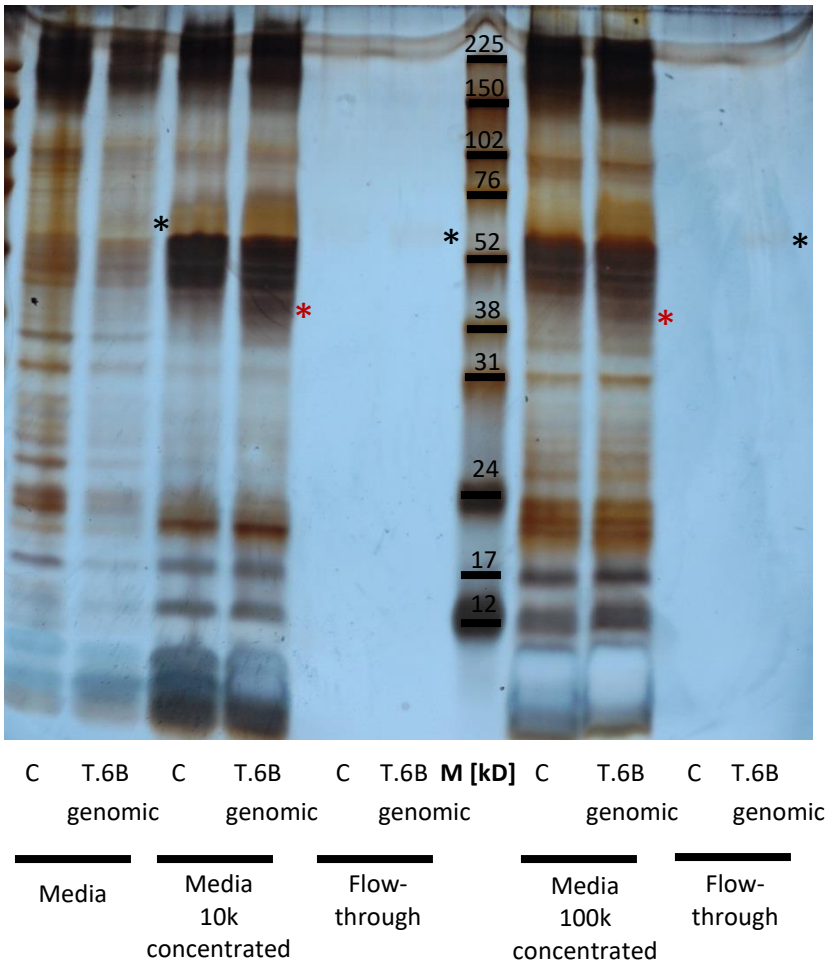


Figure 3.20. Silver stain of worm secreted media after isolation from liquid culture genomic ES-62 transgenic line. Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Samples were concentrated using Amicon filters (Section 2.15.). Red stars indicate a 46 kDa band detected in concentrated media from 6B transgenic worms. Black stars indicate faint 60 kDa bands present in the un-concentrated media from 6B transgenic worms and in the flow-through from the concentration step.

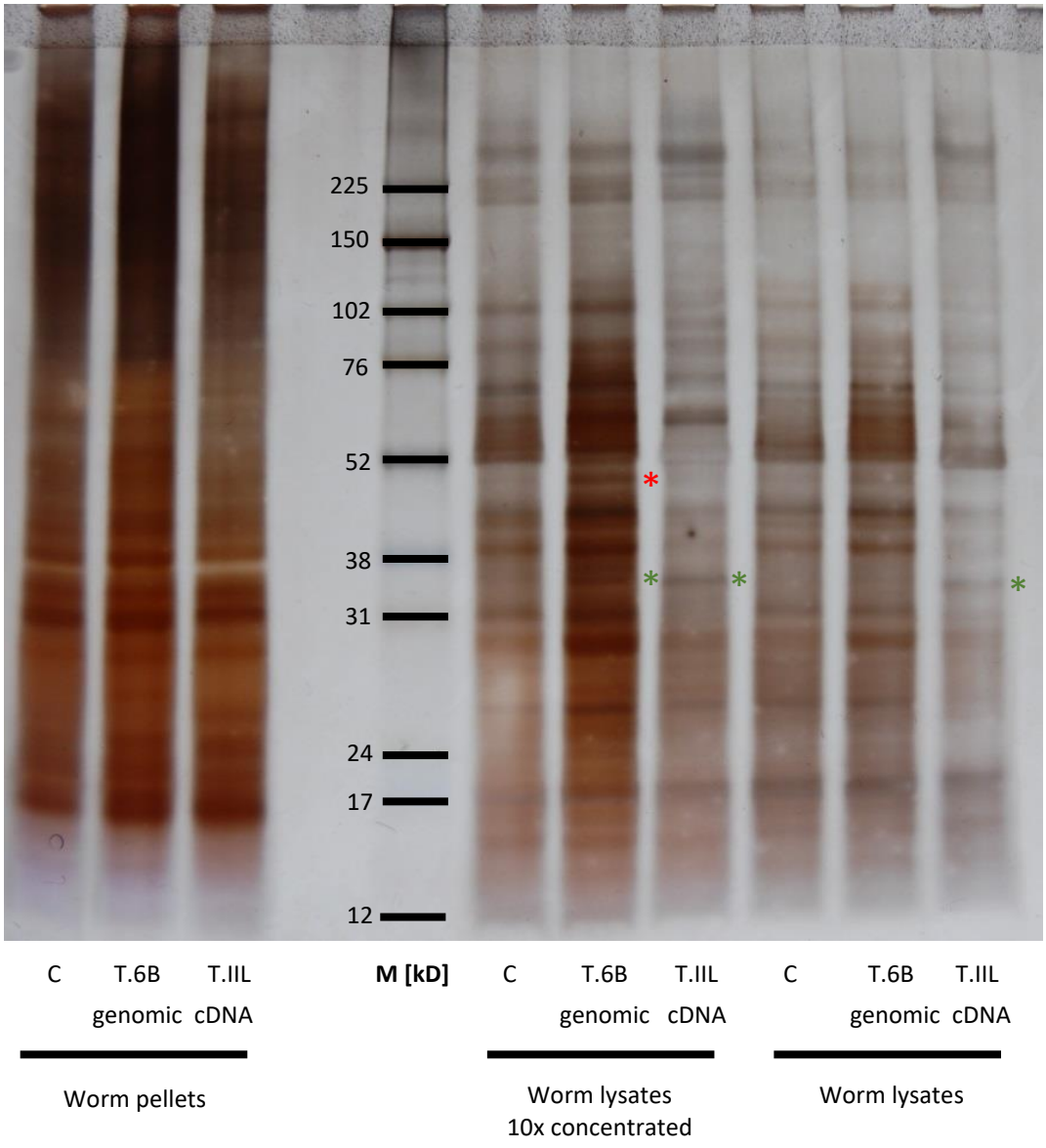


Figure 3.21. Silver stain of worm pellets, worm lysates and concentrated worm lysates after isolation from liquid culture. Samples: T.6B genomic- 6B *asp-6* ES-62 genomic fragment transgenic line; T.IIL cDNA- IIL *asp-6* ES-62 cDNA transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Red star indicates a 46 kDa band in T.6B. Green stars indicate a 35 kDa band in concentrated lysates of transgenic worms. Samples were concentrated using Amicon filters as described in materials and methods (Section 2.14.). Worm pellets were mixed with sample buffer and 0.1M DTT and kept at 70°C for 10 minutes. The samples were mixed, centrifuged and the supernatant was loaded on the gel.

3.4.4 MALDI-TOF MS analysis of transgenic *C. elegans* secreted media (large scale liquid culture) and parasite-derived ES-62

The samples were not concentrated enough in order to be excised from the gel and subjected to individual, in-gel trypsin digestion. Instead, the samples were subjected to buffer exchange as described in materials and methods (Section 2.20.) and then subjected to an in- liquid trypsin digest (Adrian Brown). This means that all of the proteins present in the media were digested by trypsin. Samples were analysed for the presence of recombinant ES-62. Results are shown in the Appendix (Tables 3, 4 and 5).

There was no ES-62 detected in any of the samples. This could be due to recombinant ES-62 not being secreted into the media. However, if the recombinant protein is secreted in low quantities, it is likely to be 'masked' by the abundance of other peptides. A large number of bacterial proteins were detected, not only *E. coli* proteins, but also proteins from pathogens.

The 6B [*asp-6* ES-62 genomic fragment] line secreted media was most likely to contain recombinant ES-62. In order to decrease the contamination by other peptides, 6B sample was fractionated using the StageTip (Section 2.21.). This method cleans, concentrates and separates peptides into fractions according to their isoelectric point. Each fraction contained a simpler mixture of peptides, which improved the number of identified proteins.

There was no ES-62 detected in the sample, even after fractionation. However, ASP-6 was consistently identified in the *C. elegans* secreted media (Table 6, Appendix).

In order to confirm the presence of *A. viteae* ES-62 in the available search- databases, the protein sample was analysed by MALDI-TOF MS. The ES-62 monomer was excised from the silver-stained SDS-PAGE gel and subjected to in-gel trypsin digestion. ES-62 was successfully identified by NCBI database, with the protein score of 698. Matched peptides covered 50% of the protein sequence (Figure 3.23). Individual peptides with their scores are presented in the Appendix (Table 2).

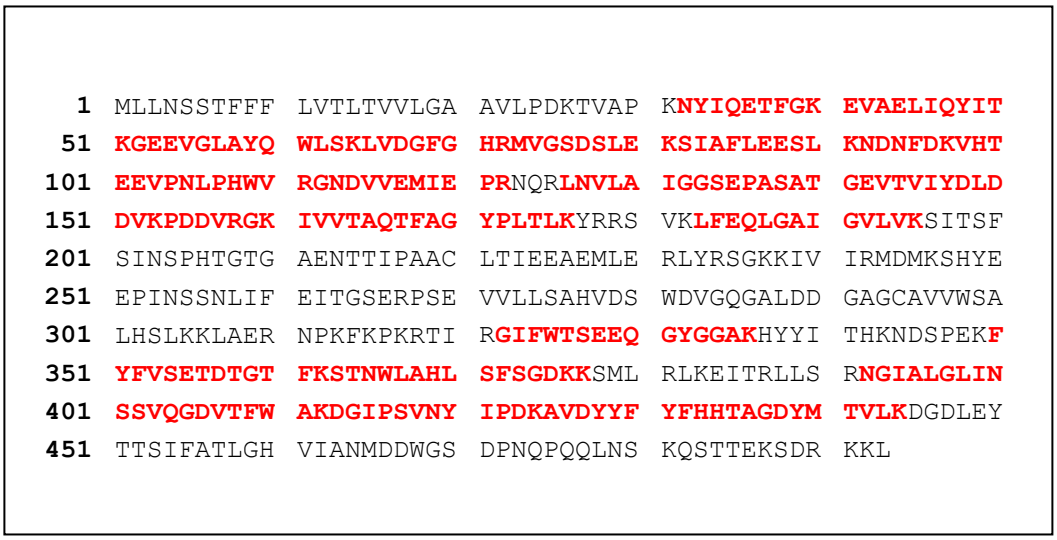


Figure 3.23. *A. viteae* ES-62 sequence coverage by peptides identified by MS. Matched peptides are shown in red.

3.4.5 Identification of the recombinant ES-62 protein in secreted media from integrated 6B transgenic lines (small scale plate culture)

The genomic lines, which presented promising results were irradiated in order to integrate the extrachromosomal arrays into a chromosome to help mitigate their genetic instability and variability. Irradiating transgenic strains induces chromosomal breaks and ligation of arrays to chromosomes during DNA repair. Because of this, mutations can arise in treated animals, so the recovered integrated strains were outcrossed by mating with wild type males (integration experiments and the initial screening of the integrated lines were performed by Dr Sushmita Maitra).

Integrated *C. elegans* transgenic lines were maintained on plates and later allowed to secrete proteins into liquid M9 media for 4 hours or overnight. Media samples were analysed by Western Blotting with the rabbit α ES-62 antibody in order to determine a transgenic line with the highest expression level of the recombinant ES-62 (Figure 3.24 supplied by Dr Sushmita Maitra). The media sample from 9* (*asp-6::*genomic ES-62 integrated) transgenic *C. elegans* reacted with the rabbit α ES-62 antibody and yielded a band pattern similar to that of parasite- derived ES-62.

The medium from the highest expressing line (9*) was tested under non-denaturing conditions (native- PAGE), in order to find out if the recombinant protein can form a tetramer. The native gel was probed with the rabbit α ES-62 antibody. A band of a similar MW to the parasite- derived ES-62 was identified in the transgenic *C. elegans* media (Figure 3.25). However, the identified protein was only present in the media, when *C. elegans* were allowed to secrete overnight.

The integrated 6B genomic ES-62 transgenic line (9*) secreted a protein which was recognised by the rabbit α ES-62 antibody in its denatured and native form. Under both denaturing and non-denaturing conditions, the media sample from 9* integrated transgenic yielded a similar band pattern to that of parasite- derived ES-62. Under non-denaturing conditions, the rabbit α ES-62 antibody recognised a band in the control media sample. The band in the control had a different profile to the

parasite- derived ES-62. However, the result suggests that *C. elegans* may possess an endogenous protein, which has similar modifications to parasite-derived ES-62.

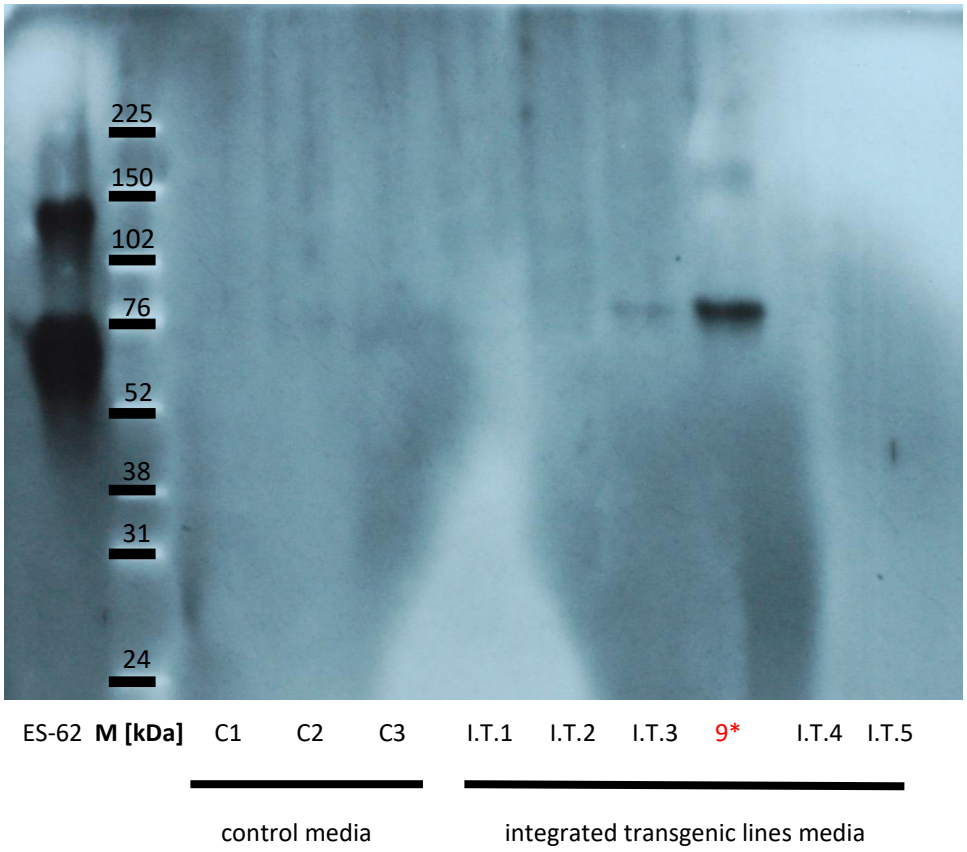


Figure 3.24: Western blot analysis of media from integrated 6B genomic ES-62 lines (small scale plate culture, rabbit α ES-62 antibody). Samples: C1- 3- integrated control lines (*Δlin-15* mutants rescued with *lin-15* plasmid). I.T.1- 5- integrated transgenic line (6B *asp-6* ES-62 genomic fragment). The rabbit α ES-62 antibody was used at 1:1000 in 5% BSA in TBS, 0.1%tween. (Integrated lines were prepared by Dr Sushmita Maitra, the figure was supplied by Dr Sushmita Maitra)

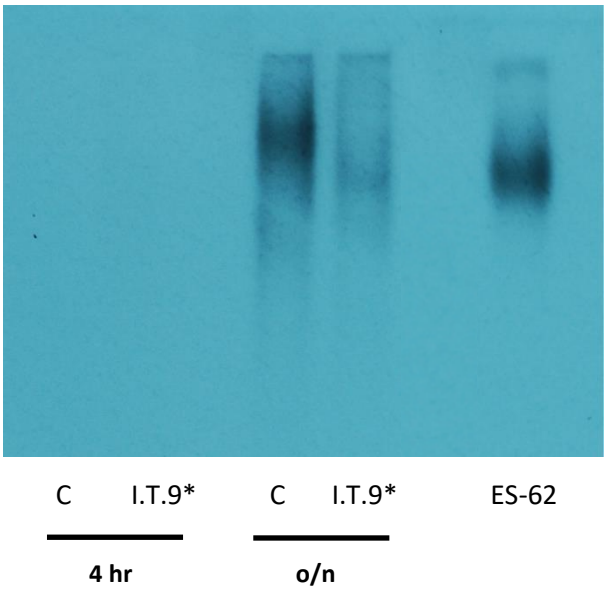


Figure 3.25. Western blot analysis of media from integrated transgenic line 9* under non-denaturing conditions (small scale plate culture, rabbit α ES-62 antibody). Samples: C- integrated control line (*Δlin-15* mutants rescued with *lin-15* plasmid). I.T.9*- integrated transgenic line (6B *asp-6* ES-62 genomic fragment). Worms were allowed to secrete into the M9 media for 4 hours or overnight (o/n). The rabbit α ES-62 antibody was used at 1:1000 in 5% BSA in TBS, 0.1%tween. The ES-62 protein was used as the reference, instead of the marker.

3.4.6 Comparing 6B and 10A genomic ES-62 transgenic *C. elegans* (large scale liquid culture)

Transgenic worms 6B and 10A are independent lines, however, they should be genetically identical. The transgenes were extracted from both lines and confirmed by sequencing (PCR reactions amplifying short, overlapping fragments were performed by Kyle Main- project student). There were no differences in the protein coding sequences of the two transgenics and no deviations from the native ES-62 gene sequence. Nevertheless, 6B and 10A transgenic lysates yielded different results when probed with the rabbit α ES-62 antibody (Figure 3.14, Dr Kara Bell).

In order to investigate the difference in recombinant protein expression in two genomic ES-62 transgenics, 6B and 10A, *C. elegans* were maintained in liquid culture. The secretion experiment was carried out as described before (Sections 2.5. and 2.5.1.). Secreted media samples from 6B and 10A genomic ES-62 transgenic *C. elegans* were probed with the rabbit α ES-62 antibody.

Figure 3.26 was supplied by Kyle Main, a project student in the laboratory. The liquid culture, sample preparation and Western Blot analysis was performed by Kyle Main under supervision of the author of this thesis. Figure 3.26 A 70 kDa band was detected in the 10A un-concentrated media (asp-6::ES-62 genomic DNA) (Figure 3.26, red star). It was not detected in the concentrated media, which suggest that the recombinant protein may be incompatible with the concentration columns. Another band only present in the un-concentrated 10A sample was 140 kDa (Figure 3.26, red star), which could represent a dimer form of the 70 kDa band. The profile of the 10A sample looked similar to the profile of parasite-derived ES-62 (a monomer at 62 kDa and a dimer at 124 kDa). However, the expected MW of correctly modified recombinant protein from genomic ES-62 transgenic *C. elegans* is 55 kDa.

Concentrated media from both 6B and 10A transgenics showed a 27 kDa band (Figure 3.26, white star), similar to the band previously detected in 6B genomic media (Figure 3.15 and 3.18 B) and 6B

worm lysates (Figure 3.16). There is also a 92 kDa band detected in all concentrated transgenic media (Figure 3.26, black star). 92 kDa corresponds to the predicted size for the recombinant ES-62 protein dimer.

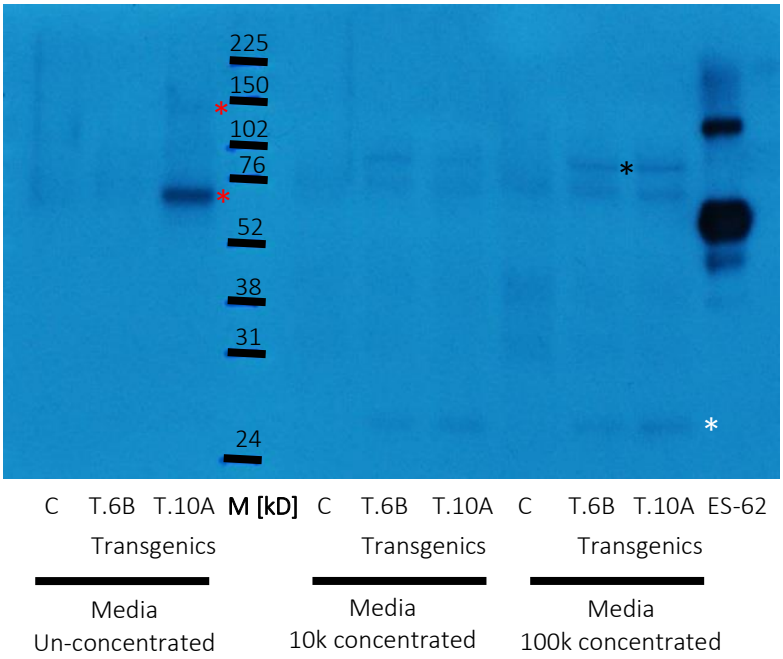


Figure 3.26. Western blot of *C. elegans* secreted media immediately after isolation from liquid culture (rabbit α ES-62 antibody). Samples: T.6B- 6B *asp-6* ES-62 genomic fragment transgenic line; T.10A- 10A *asp-6* ES-62 genomic fragment transgenic line; C- *Δlin-15* mutants rescued with *lin-15* plasmid. After isolation media was concentrated using Amicon centrifugal filters with 10 kDa MW cut-off and later with 100 kDa MW cut-off. Red stars indicate an ES-62- like band pattern in 10A transgenic media with a 70 kDa lower MW band and a 140 kDa higher band MW. Black star indicates 92kDa bands present in concentrated media samples from transgenic lines, but not in the control lines. White star indicate 27 kDa bands in concentrated transgenic media. Antibody: α ES-62 serum used at 1:1000 in 5% BSA in TBS, 0.1%tween (figure was supplied by Kyle Main a project student in the lab, the liquid culture, sample preparation and Western Blot analysis was performed by Kyle Main under supervision of the author of this thesis).

The media samples were kept at 4°C with protease inhibitors (cOmplete™ Protease Inhibitor Cocktail Tablets, Roche). In order to further investigate the nature of the 70 kDa and the 140 kDa bands as well as check for degradation products, the same samples were tested again after two weeks.

After two weeks the 70 kDa band could no longer be detected in 10A un-concentrated media (Figure 3.27). Instead, smaller MW bands, previously not seen were identified. A 35 kDa band was present in 10A transgenic media. A 27 kDa band was present in all of the transgenic samples and has been consistently been identified in genomic ES-62 transgenic worm lysates (Figure 3.16.) and media (Figures 3.15. and 3.18. B). This band appeared in all transgenic samples, which were concentrated or not analysed immediately after isolation, which suggests that it may be a product of degradation. Since the protein was only present in the transgenic samples, it may be a degradation product of recombinant ES-62.

The 92 kDa band was consistently identified in the concentrated transgenic media samples and not in the control, even after 2 weeks (Figure 3.27.). The MW prediction of recombinant genomic ES-62 is 46 kDa. If the protein is not modified, the ~92 kDa band could be a dimer of recombinant genomic ES-62, which is more stable than a monomer and does not degrade as easily. All of the bands would need to be investigated further in order to draw any conclusions about their nature.

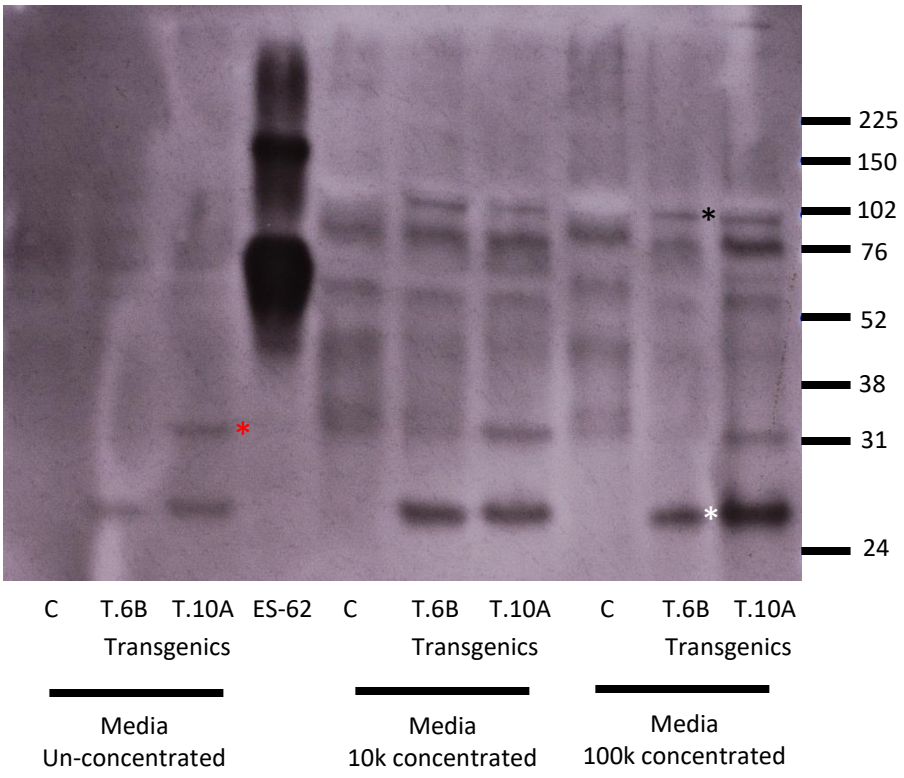


Figure 3.27. Western blot of worm secreted media two weeks after isolation from liquid culture (rabbit αES-62 antibody). Samples: 6B *asp-6* ES-62 genomic fragment transgenic line; 10A *asp-6* ES-62 genomic fragment transgenic line; Ct- *Δlin-15* mutants rescued with *lin-15* plasmid. The red star indicates a ~32 kDa band present in all samples from 10A transgenic lines. The black star indicates a band ~92kDa present in concentrated media samples from transgenic lines, but not in the control lines. The white star indicated a ~27 kDa band previously detected in genomic ES-62 samples, now present in all transgenic samples and not in the control. Antibody: α ES-62 serum used at 1:1000 in 5% BSA in TBS, 0.1%tween.

The media sample from the 10A genomic ES-62 line reacted with α ES-62 serum antibody and produced a band pattern similar to that of parasite- derived ES-62. In order to further investigate the nature of those bands, samples were analysed by SDS-PAGE and silver staining. Once detected on the gel, the band could be excised, subjected to a trypsin digest and analysed by MALDI- TOF MS.

The mixture of proteins secreted by the liquid culture worms identified by silver staining was complex (Figure 3.28). A 62 kDa band was identified in the un- concentrated 10A genomic ES-62 transgenic media (Figure 3.28 A). This band corresponds to the MW of parasite- derived ES-62. However, the predicted MW of recombinant genomic ES-62 protein is 55 kDa, if modified correctly. Moreover, a 62 kDa band was not identified in any previous experiments. The band could be an unspecific result arising from sample- to- sample variation.

The concentration of media using a MW cut-off columns yielded no difference in sample profiles. In fact, the proteins were best separated in un-concentrated samples.

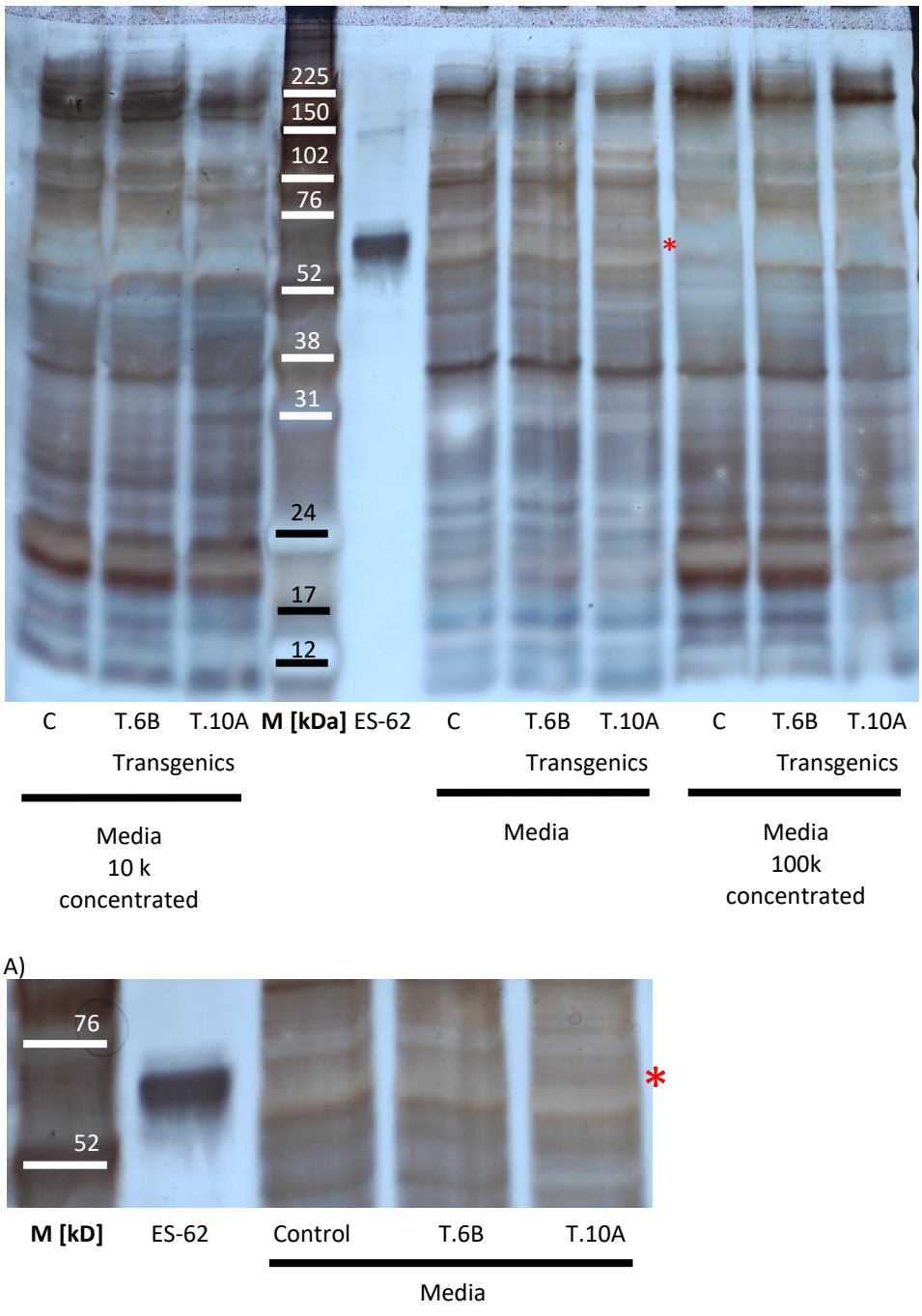


Figure 3.28 Silver stain of worm secreted media after isolation from liquid culture. Samples: T.6B- *asp-6* ES-62 genomic fragment transgenic line; T.10A- *asp-6* ES-62 genomic fragment transgenic line; C- control- *Δlin-15* mutants rescued with *lin-15* plasmid. Samples were concentrated using Amicon filters as described in materials and methods. The red star indicates a 64 kDa band detected in un-concentrated media from 10A transgenic worms. A) Close-up of the silver stain section, where the band was identified.

4. Discussion

4.1 *C. elegans* secrete proteins into liquid media

This study found that *C. elegans* secrete a simple mixture of proteins into the media in a relatively short period of time (4 hours). Results in figures 3.1., and 3.2. come from relatively small culture volume. The aim of the experiment was to detect the most abundant proteins, which can be seen on the silver stain. Moreover the worms were allowed to secrete the proteins over a relatively short period of time, to minimize overcrowding and worm lysis, which would result in contamination with *C. elegans* internal proteins. The fact that only simple mixture of proteins was detected could be due to the low concentration of the secreted media.

The most abundant proteins consistently identified in the externally secreted mix were proteases. They are likely to play part in *C. elegans* digestion by serving as enzymes breaking down bacterial cell wall. Wong et al (2006) observed that the expression of *asp-6* increases when *C. elegans* is infected with pathogens. Moreover, the *asp-6* knockout mutant expresses an altered response to a change in the environment (particularly different bacteria) compared to the control (unpublished comment on Wormbase by Herman, M.). This suggests that *C. elegans* secretes proteases as tools for interacting with the environment. Secreted proteases could play a role in defence against pathogens.

The media was fractionated by SDS-PAGE and proteins were detected by silver staining. MALDI-TOF MS analysis of bands excised from the silver stained gel consistently identified the most abundant proteins (as judged by the intensity of the silver stain) as aspartyl proteases (ASP-6 and ASP-5) (Figure 3.1 and 3.2). ASP-6 was also identified by direct MS analysis of *C. elegans* media subjected to in-solution trypsin digest (Table 3, Appendix). ASP-6 was previously described as a secreted protein by Lochnit et al (2006) (Section 1.10.). Considering the expression pattern of ASP-6 GFP reporter fusions in *C. elegans* intestine and pharynx (Lochnit et al., 2006) with our

identification of ASP-6 in the *C. elegans* media, we conclude that ASP-6 is secreted externally. This is why we used the *asp-6* promoter and UTRs to drive expression of ES-62 in *C. elegans.* We hypothesised that if we use control elements of an endogenous *C. elegans* gene coding for a secreted protein, to drive the expression of ES-62, then the recombinant protein would be likely to be secreted.

Many *E. coli* proteins were also identified in the *C. elegans* secreted media. This was expected as worms were fed on *E. coli* in the experiment. Despite extensive washing of the worms before the secretion step and thorough removal of the bacteria, *E. coli* proteins were abundantly present in the secreted worm media. The most often identified was an *E. coli* outer membrane protein OmpF (Inokuchi et al., 1982), which is likely to remain 'stuck' to the worms during the washing process and, therefore, is still detectable by MALDI- TOF MS analysis. Small parts of the broken down *E. coli* membrane could also remain in the *C. elegans* gut during the washing and only be released later from the worms.

The identification of the recombinant ES-62 relies on the simplicity of the media, especially if secreted at a low level. The abundance of bacterial products is likely to interfere with the MALDI-TOF MS readings and "mask" the much lest concentrated recombinant protein. *E. coli* proteins could also interfere with the purification. In the future, better methods could be employed to separate *E. coli* from *C. elegans* before collecting worm secretions. Alternatively, *C. elegans* could be grown in a nutrient rich media without bacteria as the source of food (Avery 1993). This would, however, be a dramatic change of conditions and its effect on *C. elegans* secreted products would need to be investigated.

4.2 The structure of the ES-62 genomic DNA fragment

This study aimed to characterise the full genomic ES-62 fragment. We compared the complete cDNA sequence of ES-62 (Harnett W et al., 1999) to the draft *A. viteae* genome sequence (Nematode Genomes, Blaxter lab, online release August 2012) and found that they did not align completely. Analysis of the *A. viteae* genome sequence revealed that the major part of the gene, from the third exon to the end, is represented in one scaffold (00478), but the first two exons are in different scaffolds. Several parts of the cDNA sequence did not align with any genomic sequence (Figure 3.4 and 3.6).

We performed a series of PCRs attempting to bridge the gaps between scaffolds (Figure 3.9). Based on our sequencing results and the cDNA sequence, the ES-62 genome fragment was reassembled (per. Comm. Georgios Koutsovoulos, Blaxter lab). We found that the genomic DNA fragment coding for ES-62 has a complex structure, with very large second intron (Figure 3.10). The ES-62 genomic sequence for the coding region is approximately 10884 bp long,consisting of 10 exons and 9 introns. The first exon is 111 nucleotides long and codes for the signal peptide and part of the mature protein. The first intron is 371 nucleotides long. The second exon is 131 nucleotides long. The second intron is approximately 6796 nucleotides long. A short part of the second intron is conserved in sequences from other parasitic nematodes (*Wuchereria bancrofti, Elaeophora elaphi, Brugia pahangi, Brugia malayi and Loa loa*) and is likely to code for important gene regulators. The following exons are respectively 116, 101, 174, 159, 162, 180, 79 and 269 nucleotides long. Approximate sizes of the introns are respectively 520, 174, 233, 267, 290, 226 and 432 nucleotides long.

The major region of genomic ES-62 from scaffold 00478 was amplified and was used as the transgenic fragment for the transformation of *C. elegans.* Due to the lack of 243 nucleotides from the 5' region, if produced, the recombinant protein would be truncated. The missing nucleotides code for 81 amino acids, including the signal sequence and 62 amino acids of the mature protein. Despite being truncated, the recombinant ES-62 contains all of the functional N-glycosylation sites.

Therefore, we hypothesised that the major fragment of the ES-62 gene could still form a modified recombinant protein.

We focused on using a genomic DNA construct for making transgenic animals rather than the cDNA due to the following reasons. Introns are important for stabilising the mRNAs and the pre-mRNA splicing promotes nuclear export and translation (Maniatis & Reed 2002; Nott et al., 2004). Furthermore, long introns are likely to contain gene- specific regulatory elements. The addition of synthetic introns was reported to substantially increase expression of reporter transgenes in *C. elegans* (Okkema et al., 1993). Use of cDNAs for protein expression in *C. elegans* is likely to be inefficient unless at least one intron is inserted. Furthermore, sequencing of the only available ES-62 cDNA clone (received from University of Nottingham) revealed that it contained 6 mutations, two of which were single base pair deletions and caused the sequence to stop prematurely. The cDNA required revision of the mutations back to the original sequence before it could have been used (cDNA repair was performed by Dr Sushmita Maitra).

Despite the presence of introns in the transgenic gene construct, the recombinant protein was not expressed correctly in transgenic *C. elegans*. Introns from the parasite have large stretches of As and Ts, which can interfere with the PCR amplification, and might cause problems with the *C. elegans* transgenes. In the future, specific introns, especially the second intron from the ES-62 genomic DNA could be incorporated into the cDNA and used as the transgene. Alternatively, synthetic ES-62 gene could be made with adapted codons designed by *C. elegans* codon adapter web tool for better expression of the recombinant protein in *C. elegans* (Redemann et al., 2011).

4.3 Expression of the recombinant protein

The final aim of this thesis was to produce transgenic *C. elegans*, which secrete recombinant ES-62 protein into the media. We made stable transgenic lines with the truncated ES-62 genomic fragment (missing 81 amino acids) under the control of *asp-6* 5'UTR, promoter and signal sequence. The

presence of the transgene was identified by single-worm PCR and confirmed by sequencing of the PCR product. Two independent genomic ES-62 transgenic lines (6B and 10A) had the entire ES-62 fragment sequenced and were found to have no mutations when compared to the *A. viteae* genome and ES-62 cDNA. Despite the presence of the transgene detected by PCR, we found the recombinant protein expression to be variable and weaker than anticipated.

Despite being truncated, the recombinant ES-62 contained all of the functional N-glycosylation sites. Therefore, we hypothesised that the major fragment of the ES-62 gene could still form a modified recombinant protein. The MW of *A. viteae* ES-62 under reducing and denaturing conditions is 62 kDa, with modifications accounting for 7 kDa,as the predicted MW of the protein, without any modifications, is ~55 kDa. The recombinant protein is 412 amino acids long, and its predicted MW is ~46 kDa. Despite being truncated, the recombinant ES-62 contains all of the functional N-glycosylation sites. Hence, if modified correctly, the recombinant ES-62 would be expected to be ~53 kDa.

4.3.1 Expression of recombinant ES-62 in transgenic *C. elegans* lysates

Western blot and silver stain analysis showed evidence of the recombinant ES-62 expressed inside transgenic *C. elegans*. Initially, transgenic *C. elegans* pellet samples and secreted media samples, from worms maintained on plates (small scale), were sent to Strathclyde University and analysed for the presence of recombinant ES-62. Western blot experiments detected a protein band of a size similar to the *A. viteae* ES-62 band in the 6B transgenic *C. elegans* lysed worm pellets. A polypeptide at ~70 kDa was recognised by two antibodies raised against parasite-derived ES-62 (whole native protein and de-glycosylated ES-62) and with the TEPC15 antibody directed against the PC modification (Figure 3.14., Dr Kara Bell). This band was not detected in the control, which suggests that it is specific to the 6B transgenic line. None of the antibodies reacted with lysates from the other

transgenic line of the same genotype- 10A, which suggests that there is a difference in the levels of expression between different transgenic lines.

6B sample also yielded a band at ~46 kDa, when stained with two antibodies raised against parasite-derived ES-62, but not the TEPC15. 46 kDa is an estimated size for the unmodified recombinant ES-62. A band of a similar size was also identified by the silver stain of 6B transgenic *C. elegans* lysate (Figure 3.21.).

It is possible that the ~46 kDa band accounts for the recombinant ES-62 that was not glycosylated or missing the PC modification. The attachment of the PC is likely to be a medial- Golgi event (Section 1.6.4.). It is possible that the recombinant protein cannot be transferred out of ER, is not modified and remains 'stuck' inside the *C. elegans* worm and, therefore, is not secreted.

Dr Sushmita Maitra transformed *C. elegans* worms with *asp-6*::ES-62cDNA labelled with GFP. She observed GFP fluorescence in intestinal cells of adult transgenic *C. elegans* (Figure 4.1.), which supports our theory that the recombinant protein is expressed inside the body of transgenic *C. elegans*.

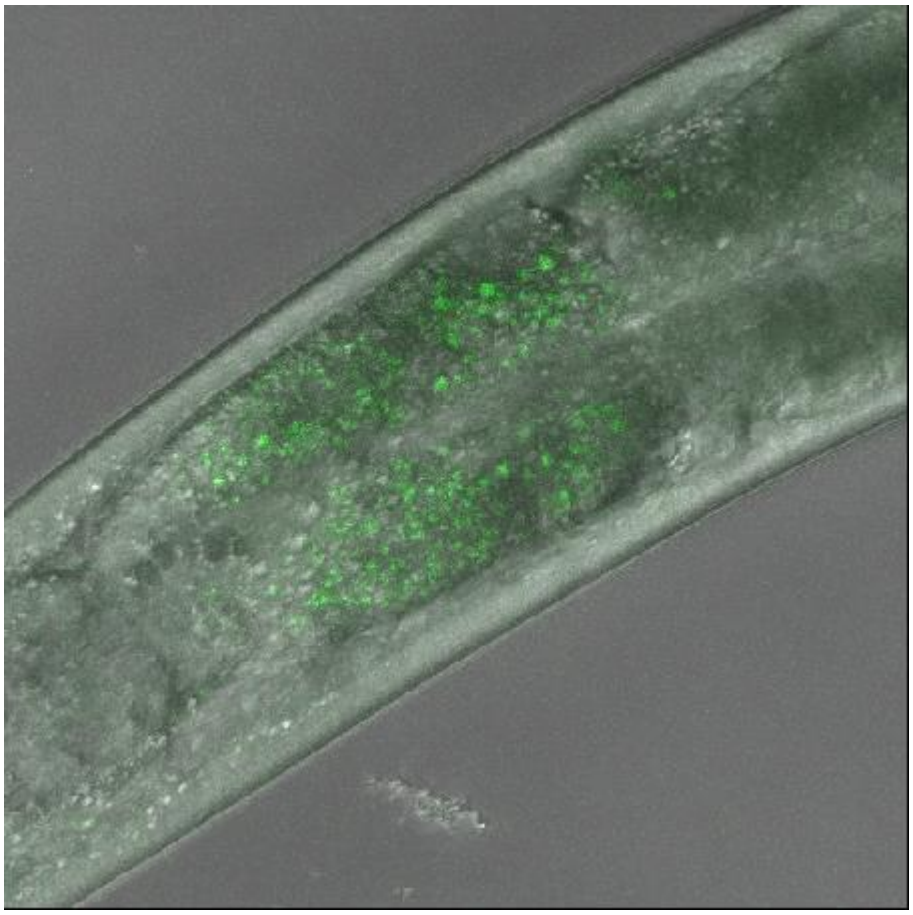


Figure 4.1. Expression pattern of *asp-6::*ES-62cDNA::GFP fusion inside the intestinal cell of adult *C. elegans* (figure supplied by Dr Sushmita Maitra).

4.3.2 Expression of recombinant ES-62 in transgenic *C. elegans* secreted media

We found limited evidence for the presence of the recombinant ES-62 in the transgenic *C. elegans* secreted media. Concentrated transgenic *C. elegans* media samples from plate cultures (small scale) were sent to the Strathclyde University, however, they showed a very weak signal and no difference between the control and the transgenic lines in any of the samples (results not shown).

Secreted *C. elegans* media, from large-scale liquid culture grown worms, were concentrated and analysed by western blotting, silver staining and MALDI-TOF MS.

Silver stain analysis detected a ~46 kDa band in the concentrated media from 6B genomic ES-62 transgenics (Figure 3.20). However, western blot analysis only recognised a 46 kDa product in the worm lysates and not in the media.

Secreted media from the integrated 6B transgenic line reacted with the rabbit α ES-62 antibody when the sample was run under non-denaturing conditions (native gel) (Figure 3.25.). Transgenics yielded a band of a similar size to parasite-derived ES-62, but only when the worms were allowed to secrete overnight. If this is recombinant ES-62 then it may be modified similarly to the parasite-derived protein. Since the band was not detected in 4-hour samples, it is possible that transgenic *C. elegans* require more time in the media for sufficient secretion.

Under native conditions, a band was also detected in the control media, however, it was larger than *A. viteae* ES-62 and the transgenic polypeptide. The result suggests that there may be an endogenous *C. elegans* proteins that react with the antibodies raised against a parasite- derived ES-62. Identification of recombinant ES-62 by MALDI-TOF is essential for the reliability of the result.

4.3.3 Differences in recombinant protein expression in ES-62 genomic lines with identical transgenes

Two independent lines, with identical transgenes with genomic ES-62 fragment, were compared (6B and 10A). Initially, only the 6B worm lysates reacted with the antibodies raised against the parasite-derived ES-62 (Figure 3.14. Dr Kara Bell), which suggested that there is a difference in expression between different lines.

Comparison of 6B and 10A transgenic *C. elegans* after growing them in liquid culture revealed a very prominent band ~70 kDa in un-concentrated media from 10A asp-6::ES-62 genomic DNA transgenic worms (Figure 3.26.). There was another band only present in the un-concentrated 10A sample ~140 kDa, which could be a dimer form of the ~70 kDa band. The profile of this sample looked similar to the profile of *A. viteae* ES-62, which forms a monomer at 62 kDa and a dimer at 124 kDa. This suggests that the correctly modified transgenic protein could be present in the un-concentrated media. This band was, however, not recognised by any other antibodies and is not a predicted MW for the recombinant protein. When tested after 2 weeks (sample kept at 4°C), the ~70 kDa and the ~140 kDa bands were not detected in the 10A transgenic un-concentrated media. Instead, two smaller MW bands were detected.

The ~27 kDa band was not detected in un-concentrated media immediately after isolation. It was, however, detected in concentrated transgenic samples (Figure 3.26.). This band appeared in all transgenic samples, which were processed (concentration) or not analysed immediately after isolation, which suggests that it can be associated with degradation. Recombinant ES-62 may be less stable than the parasite- derived protein. Since the band is only present in the transgenic samples, it may be a degradation product of recombinant ES-62.

The rabbit α ES-62 serum antibody has a wide detection spectrum, it may recognise PC epitopes. The multiple polypeptides detected by this antibody may reflect PC-containing proteins in *C. elegans*.

Both 10A and 6B transgenic media (10k and 100k concentrated) show a band that is not present in the control, at ~92 kDa, which did not degrade after two weeks. The MW prediction of recombinant genomic ES-62 is 46 kDa. If the protein is not modified, but still capable of folding, the ~92 kDa band could be a dimer of recombinant genomic ES-62, which is more stable than a monomer and does not degrade as easily. It is also possible, that the recombinant ES-62 was not denatured properly and hence the lack of the monomer. A 92 kDa band was not recognised by any other antibodies. A similar result was found when recombinant ES-62 was expressed in *Pichia pastoris* (Egan et al., 2006). The recombinant ES-62 from yeast reacted only with the rabbit α ES-62 antibody and failed to monomerize (Section 1.7.).

4.3.4 Regulation of folding and detection of recombinant protein

Many secreted proteases are synthesized in a preprotein form as precursors. The prepeptide or signal peptide (Von Heijne, 1986) transports the protein across the membrane of the endoplasmic reticulum, allowing the protein entry to the secretory pathway (Pfeffer and Rothman, 1987). The prepeptide is followed by the propeptide, a further N-terminal extension (30–250 amino acids in length). The propeptide has been found to be specific and essential in assisting the correct folding and the secretion of its associated protein (Eder and Fersht, 1995; Beggah, 2000). When the folding is complete, the propeptide is removed to generate the active enzyme. The lack of the N-terminal peptide specific for ES-62 in the transgenic construct may therefore impede the exit of the recombinant protein from the ER. It could also result in incorrect folding of the recombinant protein, which cannot be "externally" secreted, or, even if secreted, cannot be recognised by antibodies raised against native ES-62. This explains the difficulties in detection of recombinant protein in the secreted media. *Asp-6* was chosen to drive the expression of recombinant protein, however the presence of native *asp-6* could reduce the expression of the transgene. *Asp-6* deletion mutant is viable under standard conditions, therefore, in the future, the recombinant protein could be injected in the mutant background. Moreover, *asp-6*

may not be the best promoter to use for the expression of secreted proteins and other promoters could be tested.

Aminopeptidase H11 from *Haemonchus contortus* was previously expressed in *C. elegans* under control of *C. elegans* cathepsin L protease (promoter and 3'UTR), signal peptide of *H. contortus Hmcp-6* gene and synthetic *C. elegans* intron. Recombinant H11 was linked to a His tag and successfully purified by cobalt chelation chromatography (Roberts et al., 2013). This suggests that the protein may need a signal sequence from its native parasite gene in order to be properly expressed. Recombinant H11 was, however, not as effective as the native protein and failed to induce protective immunity in treated animals. This has been explained by the differences in preparation of native and recombinant proteins and suggests that harnessing the secretory pathway is a promising new way of expressing functional proteins.

Cystatin *Av17* gene from *A. viteae* was previously expressed in *C. elegans* under the control of *C. elegans hsp16/41* promoter and a synthetic intron. Protein expressed under the control of *hsp16/41* promoter on heat shock is targeted to the gut cells of transgenic worms (Pillai et al., 2005). Two constructs were prepared, one containing *Av17* cDNA and another containing *Av17* genomic DNA. The constructs were also attached to a His tag. The cDNA recombinant construct was recognised by anti-His antibodies and antibodies specific for *Av17*, however the protein could not be purified using Ni-chelate affinity chromatography. The expression of the recombinant protein resulted in the reduction in the number of eggs developing to adults in transgenic worms (Pillai et al., 2005). This has not been assessed in our study, however the toxicity of the recombinant protein could be a reason for low expression and therefore difficulties in detecting secreted recombinant protein in media. The expression of recombinant protease inhibitor could interfere with proteases involved in the differentiation or moulting of the worms.

In the same study, the expression pattern of Av17 was analyzed in transgenic lines of *C. elegans* obtained by microinjection of the Av17promoter::GFP construct (Pillai et al., 2005). The GFP expression was observed in the gland cells of the pharynx and the rectal gland cells of transgenic *C. elegans*, which was inconsistent with the immunostaining by indirect immunofluorescence using anti-Av17 antibodies and anti-*C. elegans* cystatin antibodies (Pillai et al., 2005). The immunofluorescence localized cystatins in *C. elegans* to the hypodermis and in developing stages (Pillai et al., unpublished results). It is possible that *C. elegans* transcription machinery is unable to process filarial regulatory sequences correctly, which could be another reason for difficulties in producing recombinant ES-62. It is also possible that the transgenic construct used was incomplete and did not contain all of the elements required to correctly express a filarial gene.

Conclusion and future work

This study has shown that *C. elegans* secrete a mixture of proteins into the surrounding media under laboratory conditions. One of those proteins was ASP-6, which has previously been reported to be PC- modified. After experimenting with several constructs and methods of production, we have shown that we can express recombinant ES-62 in the *C. elegans* worm, using a transgene in which a partial genomic fragment of ES-62 was coupled to the promoter, 5-UTR and the signal sequence of *asp-6*. However, the expression in the worm was not very strong. We found limited evidence for the presence of recombinant ES-62 in secreted transgenic *C. elegans* media.

The results suggest that *A. viteae* ES-62 is challenging to make in recombinant systems. The complex modifications required for the protein to be functional are reflected in the sequence of the ES-62 gene. The localisation of the very large intron in ES-62 genomic sequence contributed to our understanding of how ES-62 is controlled and provided a platform for future modifications.

The mechanism of *C. elegans* secretion needs to be investigated further, under different conditions in order to find out how it can be controlled and manipulated for optimal production of secreted proteins.

The main obstacle was the lack of reliable detection methods for the recombinant protein. In the future, this could be resolved by the use of different purification tags including a His tag or a GFP tag. Attachment of a His tag has previously been reported to aid detection of recombinant proteins in *C. elegans* (Roberts et al., 2013; Pillai et al., 2005) and there are commercially available anti-His antibodies. The use of His tag would also help with future purification of the recombinant protein.

The use of GFP tag is currently explored by Dr Sushmita Maitra. It has already enabled the detection of recombinant ES-62 inside the worm body cavity. The anti-GFP antibody recognised recombinant ES-62 in the transgenic worm lysates. The potential of GFP tag could be explored further in order to

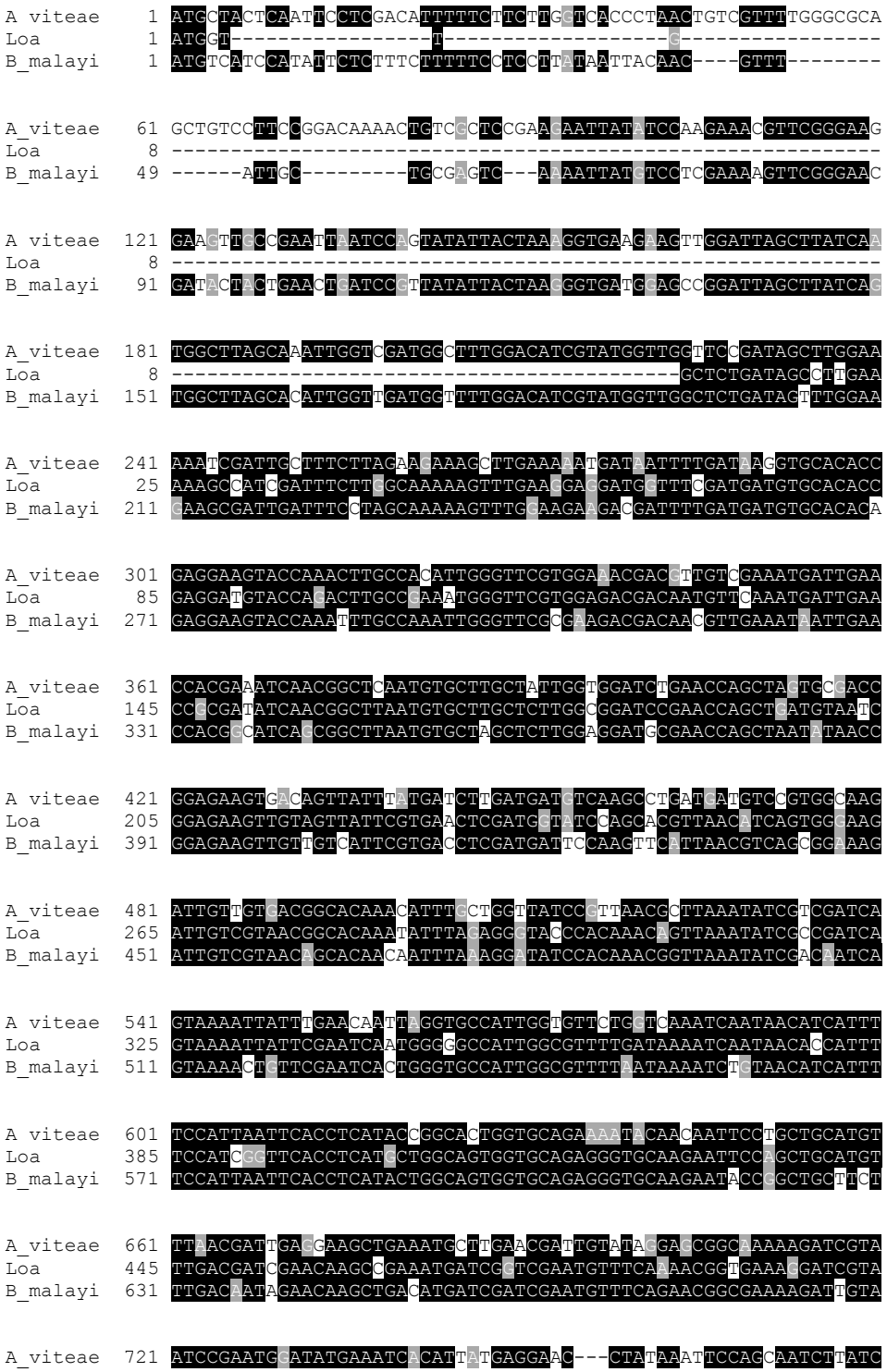
detect recombinant protein in the media. This has been attempted, however, no protein was detected, perhaps due to the small culture volume and low concentration of the protein.

Finally, the available ES-62 antibodies raised against the native filarial protein may not be optimised for the detection of recombinant ES-62 made in different organisms. It has been reported that mouse polyclonal antibodies raised against the *A. viteae* ES-62 did not recognise the recombinant protein in *P. pastoris* (Egan et al., 2006). Moreover, during this investigation, we found that antibodies raised against the native ES-62 had variable reactivity with transgenic worms and media samples. In order to improve the reliability of the detection of the recombinant protein, new antibodies could be raised against short, commercially synthesised peptides of ES-62. The use of optimised antibodies directed against ES-62 peptides would significantly improve the detection of the recombinant material, especially where the concertation is likely to be very low.

New transgenes could be designed in order to get a better expression of recombinant ES-62. The full-length genomic clone could be included to make transgenic lines. Alternatively, synthetic introns could be incorporated into full-length ES-62 cDNA to make a transgenic line, as the addition of introns and codon usage influences expression in *C. elegans*. The secretion of the recombinant protein could be optimised with the use of tissue-specific and inducible promoters. The expression of a transgene in an ASP-6 mutant background would alleviate the possibility of transgene suppression. Finally, to reduce variability in expression, the CRISPR/Cas9 system could be used to edit the genome and insert ES-62 coding sequence into *C. elegans* genes encoding secreted proteins.

Appendix

Figure 1. Alignment of *A. viteae* es-62, Loa *loa leucyl* aminopeptidase, and *Brugia malayi* leucyl aminopeptidase protein coding DNA sequence. Black colour indicates conserved residues, grey colour indicates conservative mutations, white colour indicates a divergence in the sequence.



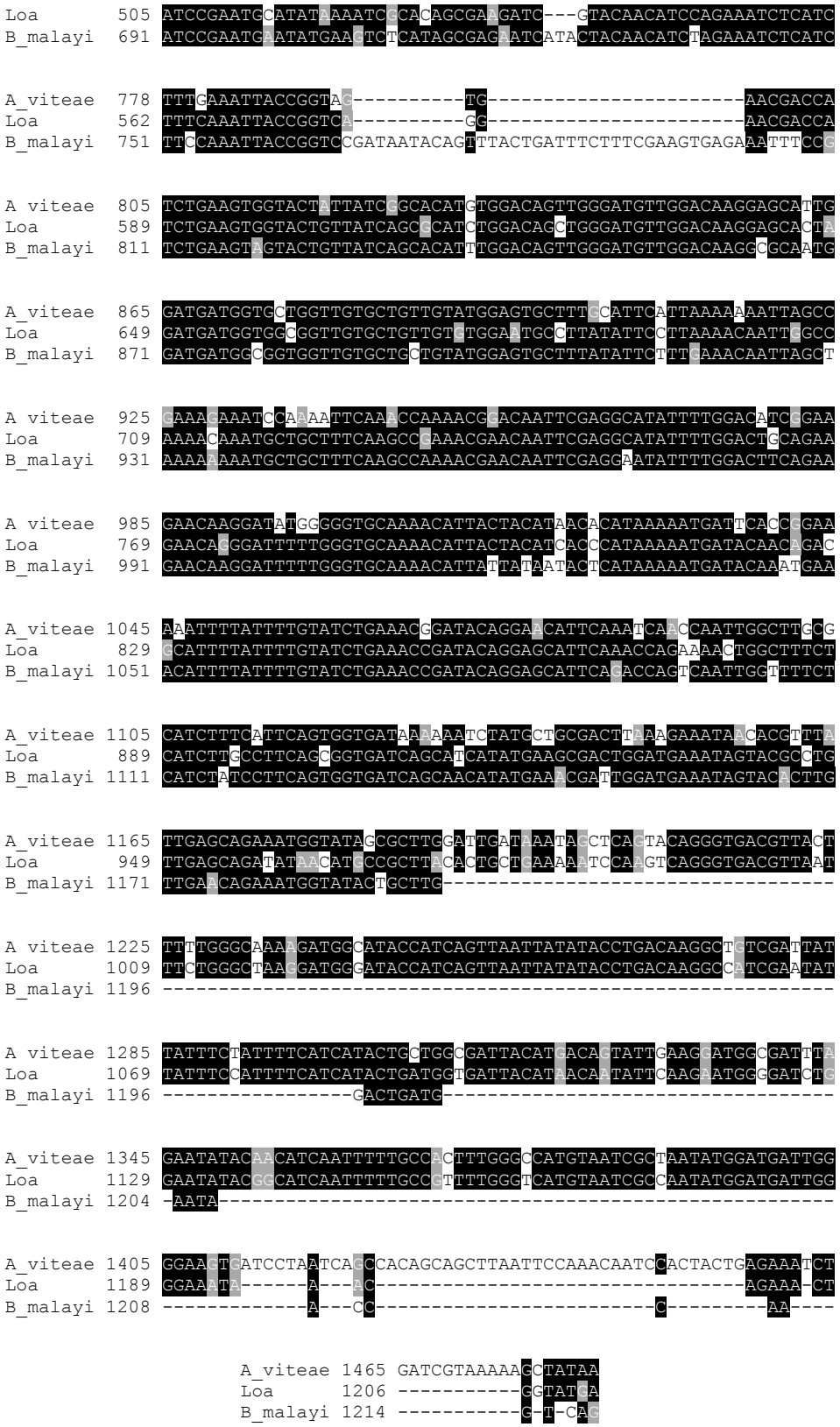

Loa 505 ATCCGAATGCATATAAAATCGCACAGCGAAGATC---GTACAACATCCAGAAATCTCATC
B_malayi 691 ATCCGAATGAATATGAAGTCTCATAGCGAGAATCATACTACAACATCTAGAAATCTCATC
A_viteae 778 TTTGAAATTACCGGTAG----------TG----------AACGACCA
Loa 562 TTTCAAATTACCGGTCA----------GG----------AACGACCA
B_malayi 751 TTCCAAATTACCGGTCCGATAATACAGTTTACTGATTTCTTTCGAAGTGAGAAATTTCCG
A viteae 805 TCTGAAGTGGTACTATTATCGGCACATGTGGACAGTTGGGATGTTGGACAAGGAGCATTG
Loa 589 TCTGAAGTGGTACTGTTATCAGCCCATCTGGACAGCTGGGATGTTGGACAAGGAGCACTA
B_malayi 811 TCTGAAGTAGTACTGTTATCAGCACATTTGGACAGTTGGGATGTTGGACAAGGCGCAATG
A_viteae 865 GATGATGGTGCTGGTTGTGCTGTTGTATGGAGTGCTTTGCATTCATTAAAAAAATTAGCC
Loa 649 GATGATGGTGGCGGTTGTGCTGTTGTGTGGAATGCCTTATATTCCTTAAAACAATTTGCC
B_malayi 871 GATGATGGCGGTGGTTGTGCTGCTGTATGGAGTGCTTTATATTCTTTCAAACAATTAGCT
A viteae 925 GAAAGAAATCCAAAATTCAAACCAAAACGGACAATTCGAGGCATATTTTGGACATCGGAA
Loa 709 AAAACAAATGCTGCTTTCAAGCCGAAACGAACAATTCGAGGCATATTTTGGACTGCAGAA
B_malayi 931 AAAAAAAATGCTGCTTTCAAGCCAAAACGAACAATTCGAGGAATATTTTGGACTTCAGAA
A viteae 985 GAACAAGGATATGGGGGTGCAAAACATTACTACATAACACATAAAAATGATTCACCGGAA
Loa 769 GAACAGGGATTTTTGGGTGCAAAACATTACTACATCACCCATAAAAATGATACAACAGAC
B_malayi 991 GAACAAGGATTTTTGGGTGCAAAACATTATTATAATACTCATAAAAATGATACAAATGAA
A_viteae 1045 AAATTTTATTTTGTATCTGAAACGGATACAGGAACATTCAAATCAACCAATTGGCTTGCG
Loa 829 GCATTTTATTTTGTATCTGAAACCGATACAGGAGCATTCAAACCAGAAAACTGGCTTTCT
B_malayi 1051 ACATTTTATTTTGTATCTGAAACCGATACAGGAGCATTCAGACCAGTCAATTGGTTTTCT
A_viteae 1105 CATCTTTCATTCAGTGGTGATAAAAAATCTATGCTGCGACTTAAAGAAATAACACGTTTA
Loa 889 CATCTTGCCTTCAGCGGTGATCAGCATCATATGAAGCGACTGGATGAAATAGTACGCCTG
B_malayi 1111 CATCTATCCTTCAGTGGTGATCAGCAACATATGAAACGATTGGATGAAATAGTACACTTG
A_viteae 1165 TTGAGCAGAAATGGTATAGCGCTTGGATTGATAAATACCTCAGTACAGGGTGACGTTACT
Loa 949 TTGAGCAGATATAACATGCCGCTTACACTGCTGAAAAATCCAAGTCAGGGTGACGTTAAT
B_malayi 1171 TTGAACAGAAATGGTATACTGCTTG----------
A viteae 1225 TTTTGGGCAAAAGATGGCATACCATCAGTTAATTATATACCTGACAAGGCTGTCGATTAT
Loa 1009 TTCTGGGCTAAGGATGGGATACCATCAGTTAATTATATACCTGACAAGGCCATCGAATAT
B_malayi 1196 ----------
A viteae 1285 TATTTCTATTTTCATCATACTGCTGGCGATTACATGACAGTATTGAAGGATGGCGATTTA
Loa 1069 TATTTCCATTTTCATCATACTGATGGTGATTACATAACAATATTCAAGAATGGGGATCTG
B_malayi 1196 ----------GACTGATG----------
A_viteae 1345 GAATATACAACATCAATTTTTGCCACTTTGGGCCATGTAATCGCTAATATGGATGATTGG
Loa 1129 GAATATACGGCATCAATTTTTGCCCTTTTGGGTCATGTAATCGCCAATATGGATGATTGG
B_malayi 1204 -AATA----------
A_viteae 1405 GGAAGTGATCCTAATCAGCCACAGCAGCTTAATTCCAAACAATCCACTACTGAGAAAATCT
Loa 1189 GGAAATA------A---AC----------AGAAA-CT
B_malayi 1208 ----------A---CC----------C----------AA-----
A_viteae 1465 GATCGTAAAAAGCTATAA
Loa 1206 ----------GGTATGA
B_malayi 1214 ----------G-T-CAG

Figure 2. Alignment of *A. viteae* ES-62, *Loa loa* leucyl aminopeptidase, and *Brugia malayi* leucyl aminopeptidase protein sequence. Black colour indicates conserved residues, grey colour indicates conservative mutations, white colour indicates a divergence in the sequence.

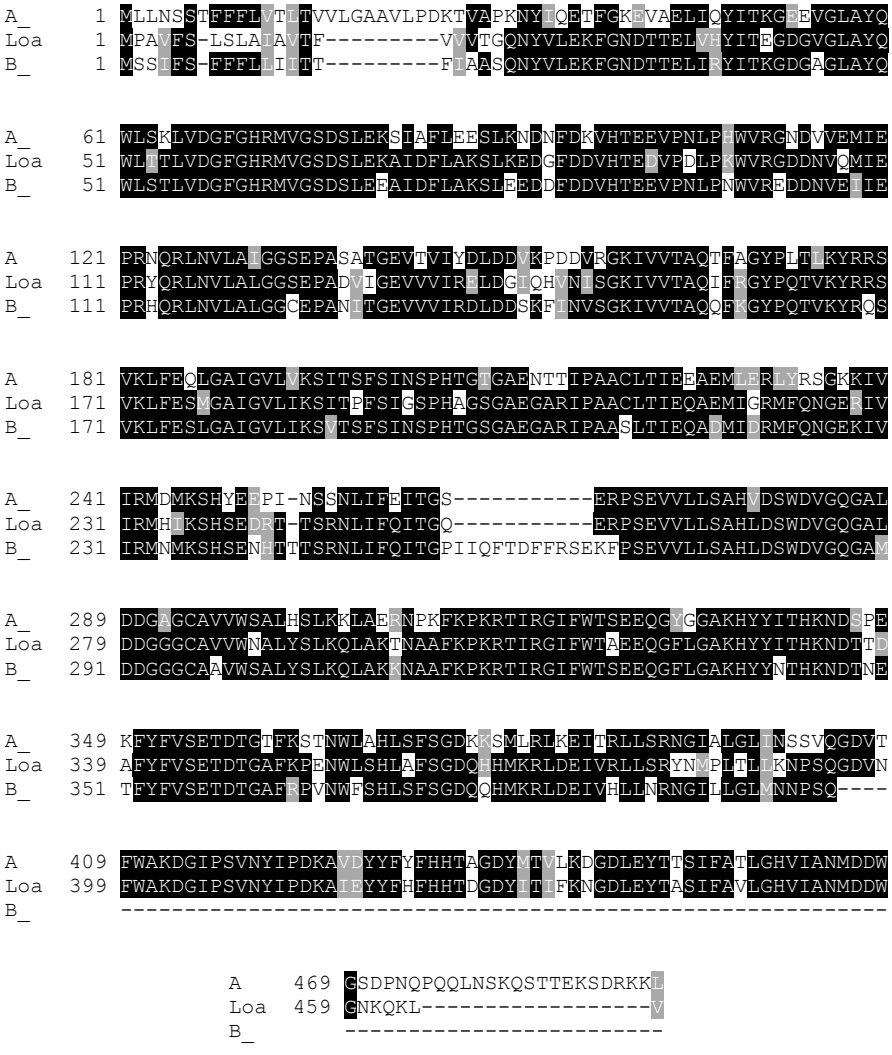


Table 1. List of primers used for PCR reactions mentioned in the results section figures: 3.5., 3.7., 3.8., 3.12., and 3.13.

Figure	PCR	5' primer	3' primer
3.5.	ES-62 (scaf 00478) outer	gtgtcacatagccacaaagctctc	gcttgctgtcttcttgctgact
3.5.	ES-62 (scaf 00478) inner	cctcctcaatgcacgctaacg	cgtcttacttcatgtttgcgaca
3.7.	olive- grey	GGAAGGAAGTTGCCGAATTAATCC	CACGAACCCAATGTGGCAAG
3.8.	green- olive	ACGACAGTTAGGGTGACCAAG	ACCATACGATGTCCAAAGCCA
3.12. A	*asp-6* promoter	tttgtgcgagactgtgtgaatgc	ggagcgacagttttgtccggaagGACTGAAGCCGATGCAAGGCCAAG
3.12. D	asp-6::ES-62 genomic fusion	gcaaactagaatgtccgtcgaaag	cgtcttacttcatgtttgcgaca
3.13.	fusion diagnostic	acgtcagtgcgcccatgag	CACGAACCCAATGTGGCAAG

Table 2. List of peptides and scores after MALDI-TOF MS analysis of in- gel trypsin digestion of ES-62. On average scores above 58 have less than 5% probability of being a random event. Significant matches are indicated in bold.

Peptide score	Peptide sequence
65	**R.MVGSDSLEK.S**
55	R.MVGSDSLEK.S + Oxidation (M)
40	K.NYIQETFGK.E
61	K.SIAFLEESLK.N
75	**R.GNDVVEMIEPR.N**
70	**R.GNDVVEMIEPR.N + Oxidation (M)**
61	**K.EVAELIQYITK.G**
43	**K.DGIPSVNYIPDK.A**
84	**K.LFEQLGAIGVLVK.S**
76	**K.GEEVGLAYQWLSK.L**
73	**K.FYFVSETDTGTFK.S**
96	**K.STNWLAHLSFSGDK.K**
61	**R.GIFWTSEEQGYGGAK.H**
51	**K.STNWLAHLSFSGDKK.S**
46	**K.VHTEEVPNLPHWVR.G**
79	**K.IVVTAQTFAGYPLTLK.Y**
73	**K.SIAFLEESLKNDNFDK.V**
117	**R.NGIALGLINSSVQGDVTFWAK.D**
73	**K.NYIQETFGKEVAELIQYITK.G**
42	**K.AVDYYFYFHHTAGDYMTVLK.D**
70	**K.NDNFDKVHTEEVPNLPHWVR.G**
90	**R.LNVLAIGGSEPASATGEVTVIYDLDDVKPDDVR.G**
102	**R.LNVLAIGGSEPASATGEVTVIYDLDDVKPDDVRGK.I**

Table 3. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from control worms grown in liquid culture. Sample: **Control** *Δlin-15* mutants rescued with *lin-15* plasmid. On average, individual ions scores > 57 indicate identity or extensive homology ($p<0.05$). Only significant matches and their scores are presented. *C. elegans* proteins are presented in blue. *E. coli* proteins are presented in red.

Score	Identified protein
163	membrane protein [Serratia liquefaciens]
150	MULTISPECIES: porin [Pseudomonas]
98	porin [Pseudomonas fluorescens]
94	protein F, partial [Pseudomonas fluorescens]
120	hypothetical protein MYCGRDRAFT_49141, partial [Zymoseptoria tritici IPO323]
112	transketolase [Serratia plymuthica A30]
111	Chain A, Conformational Variability In The Refined Structure Of The Chaperonin Groel At 2.8 Angstrom Resolution
111	ABC transporter [Pseudomonas fluorescens]
106	serine protease [uncultured organism]
101	Transport-associated [Yersinia bercovieri ATCC 43970]
90	hypothetical protein [Pseudomonas fluorescens]
88	Chain A, Structure Of Fructose-Bisphosphate Aldolase
87	Protein ASP-6 [Caenorhabditis elegans]
87	MULTISPECIES: peroxidase [Pseudomonas]
86	Protein CLEC-63 [Caenorhabditis elegans]
84	polymerase beta,RNA
82	aldehyde dehydrogenase [Acidovorax radicis]
82	manganese superoxide dismutase [Tatumella ptyseos]
77	RecName: Full=Azurin [Pseudomonas putida]
74	proton-translocating ATPase b subunit (uncF; gtg start codon) [Escherichia coli]
74	MULTISPECIES: ketol-acid reductoisomerase [Pseudomonas]
73	hypothetical protein AU67_04460 [Salmonella enterica subsp. enterica serovar Enteritidis str. SA20085285]
70	Protein F40F4.6 [Caenorhabditis elegans]
70	basement membrane proteoglycan [Caenorhabditis elegans]
69	acetyl-coenzyme A synthetase [Afipia clevelandensis]
68	MULTISPECIES: elongation factor Ts [Pseudomonas fluorescens group]
67	hypothetical protein SLIQ_07180 [Serratia liquefaciens FK01]
67	hypothetical protein NEMVEDRAFT_v1g221767 [Nematostella vectensis]
65	S-adenosyl-L-homocysteine hydrolase [Pseudomonas syringae pv. pisi str. 1704B]
65	Protein SPP-5 [Caenorhabditis elegans]

64	extracellular solute-binding protein family 3 [Pantoea sp. aB]
63	unnamed protein product [Escherichia coli]
62	MULTISPECIES: peptidoglycan-binding protein [Pseudomonas]
60	methyltransferase type 12 [Serratia liquefaciens]
60	antioxidant, AhpC/TSA family [Proteus mirabilis ATCC 29906]

Table 4. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from 6B ES-62 genomic DNA transgenic worms grown in liquid culture. Sample: **6B** *asp-6*::genomic ES-62 transgenic. On average, individual ions scores > 59 indicate identity or extensive homology ($p<0.05$). Only significant matches and their scores are presented. *C. elegans* proteins are presented in blue. *E. coli* proteins are presented in red.

Score	Identified protein
132	hypothetical protein [Pseudomonas fluorescens]
131	molecular chaperone GroEL [Serratia liquefaciens]
124	Chain A, Structure Of The Qb Replicase, An Rna-Dependent Rna Polymerase Consisting Of Viral And Host Proteins
94	MULTISPECIES: elongation factor Tu [Yersinia]
116	Transport-associated [Yersinia bercovieri ATCC 43970]
113	MULTISPECIES: azurin [Pseudomonas fluorescens group]
106	acetaldehyde dehydrogenase [Serratia liquefaciens]
102	proton-translocating ATPase b subunit (uncF; gtg start codon) [Escherichia coli]
99	serine protease [uncultured organism]
98	triosephosphate isomerase [Serratia liquefaciens]
95	molecular chaperone GroEL [Pseudomonas sp. Eur1 9.41]
91	MULTISPECIES: phosphoglycerate kinase [Rahnella]
90	Protein CLEC-63 [Caenorhabditis elegans]
88	Protein ENOL-1, isoform a [Caenorhabditis elegans]
87	ABC transporter [Pseudomonas fluorescens]
87	porin [Pseudomonas fluorescens]
87	unnamed protein product [Escherichia coli]
85	Chain A, Structure Of Fructose-Bisphosphate Aldolase
79	MULTISPECIES: ketol-acid reductoisomerase [Pseudomonas]
79	membrane protein [Serratia liquefaciens]
79	MULTISPECIES: porin [Pseudomonas]
79	polymerase beta,RNA
77	MULTISPECIES: peptidoglycan-binding protein [Pseudomonas]
77	metal ABC transporter substrate-binding protein [Serratia liquefaciens]
76	methyltransferase type 12 [Serratia liquefaciens]
75	MULTISPECIES: hypothetical protein [Pseudomonas putida group]
72	Protein ASP-6 [Caenorhabditis elegans]
71	catalase HP1 [Escherichia coli]
71	aldehyde dehydrogenase [Acidovorax radicis]
70	MULTISPECIES: glyceraldehyde-3-phosphate dehydrogenase [Serratia]

69	thioredoxin [Pseudomonas sp. Eur1 9.41]
69	von Willebrand factor A [Streptomyces seoulensis]
69	ribosomal protein S3 [cotton phyllosphere bacterium F]
67	unnamed protein product [Escherichia coli K-12]
66	MULTISPECIES: peroxidase [Pseudomonas]
65	glycerophosphoryl diester phosphodiesterase [Pseudomonas fluorescens]
65	MULTISPECIES: ferritin [Pseudomonas syringae group genomosp. 2]
64	Protein SPP-2 [Caenorhabditis elegans]
61	Protein LYS-7 [Caenorhabditis elegans]
60	unnamed protein product [Escherichia coli]

Table 5. MALDI-TOF MS results after in- solution trypsin digestion of 30x concentrated media from IIL ES-62 cDNA transgenic worms grown in liquid culture. Sample: **IIL** *asp-6*::ES-62 cDNA transgenic. On average, individual ions scores > 58 indicate identity or extensive homology ($p<0.05$). Only significant matches and their scores are presented. *C. elegans* proteins are presented in blue. *E. coli* proteins are

Score	Identified protein
140	MULTISPECIES: membrane protein [Pseudomonas]
137	hypothetical protein [Pseudomonas fluorescens]
129	Protein ASP-6 [Caenorhabditis elegans]
129	OsmY [Serratia liquefaciens]
128	MULTISPECIES: azurin [Pseudomonas fluorescens group]
112	50S ribosomal protein L22 [Pseudomonas syringae pv. pisi str. 1704B]
108	elongation factor Tu [Pseudomonas fluorescens]
102	serine protease [uncultured organism]
101	MULTISPECIES: porin [Pseudomonas]
100	membrane protein [Serratia liquefaciens]
99	peroxiredoxin, partial [Salmonella enterica]
94	MULTISPECIES: thioredoxin [Serratia]
93	PA1584, partial [synthetic construct]
90	MULTISPECIES: porin [Pseudomonas]
90	polynucleotide phosphorylase/polyadenylase [Pseudomonas syringae pv. oryzae str. 1_6]
88	aldehyde dehydrogenase [Pseudomonas fluorescens]
86	MULTISPECIES: phosphoglycerate kinase [Rahnella]
85	molecular chaperone GroEL [Serratia liquefaciens]
85	Protein CLEC-63 [Caenorhabditis elegans]
83	polymerase beta,RNA
82	hypothetical protein MYCGRDRAFT_49141, partial [Zymoseptoria tritici IPO323]
74	Chain A, Structure Of Fructose-Bisphosphate Aldolase
70	unnamed protein product [Escherichia coli]
70	Protein K06G5.1, isoform a [Caenorhabditis elegans]
70	proton-translocating ATPase b subunit (uncF; gtg start codon) [Escherichia coli]
64	amino acid ABC transporter substrate-binding protein [Pseudomonas fluorescens]
63	Protein F28B4.3 [Caenorhabditis elegans]
63	porin [Pseudomonas fluorescens]
62	elongation factor Tu [Synechococcus sp. WH 8103]

61	putative uncharacterized protein [Bacteroides sp. CAG:770]
60	catalase HP1 [Escherichia coli]
59	MULTISPECIES: peptidoglycan-binding protein [Pseudomonas]

Table 6. MALDI-TOF MS results for 5 fractions from 6B *asp-6* ES-62 genomic fragment transgenic line secreted media after liquid culture. Sample were concentrated using Amicon filters and subjected to buffer exchange (Section 2.20.), in-liquid trypsin digestion and StageTip fractionation as described in materials and methods (Section 2.21.). On average, individual ions scores > 57 indicate identity or extensive homology ($p<0.05$). Only significant matches and their scores are presented. *C. elegans* aspartyl proteases are indicated in blue. *E. coli* proteins are indicated in red. Fraction number is the concentration of Ammonium Acetate used in mM.

Fraction	Score	Identified protein
50	170	D1086.11, isoform b [Caenorhabditis elegans]
50	166	ILYS-5, isoform a [Caenorhabditis elegans]
50	165	C33G8.4 [Caenorhabditis elegans]
50	160	hypothetical protein C39D10.7 - Caenorhabditis elegans
50	151	SRP-7, isoform a [Caenorhabditis elegans]
50	150	SPP-5 [Caenorhabditis elegans]
50	148	C49G7.3 [Caenorhabditis elegans]
50	147	CBD-1 [Caenorhabditis elegans]
50	127	F15A4.6 [Caenorhabditis elegans]
50	118	unnamed protein product [Homo sapiens]
50	114	CPI-1 [Caenorhabditis elegans]
50	114	unnamed protein product [Homo sapiens]
50	113	protein T08A9.7 [imported] - Caenorhabditis elegans
50	113	K11D12.13 [Caenorhabditis elegans]
50	112	GRD-5 [Caenorhabditis elegans]
50	111	GRL-15 [Caenorhabditis elegans]
50	108	Hypothetical protein CBG06172 [Caenorhabditis briggsae]
50	107	T19B10.2 [Caenorhabditis elegans]
50	106	DPY-17 [Caenorhabditis elegans]
50	104	CLEC-223 [Caenorhabditis elegans]
50	103	FBN-1, isoform e [Caenorhabditis elegans]
50	96	E04F6.8 [Caenorhabditis elegans]
50	96	C08F11.12 [Caenorhabditis elegans]
50	95	Y57G11B.5 [Caenorhabditis elegans]
50	95	RecName: Full=Major outer membrane lipoprotein; AltName: Full=Murein-lipoprotein; Flags: Precursor [Serratia marcescens]
50	95	protein F32E10.3 [imported] - Caenorhabditis elegans
50	91	Y69H2.3, isoform c [Caenorhabditis elegans]
50	91	Chain A, Ompf Porin Mutant D74a
50	89	F30H5.3 [Caenorhabditis elegans]
50	86	Y42G9A.2 [Caenorhabditis elegans]

50	86	protein F10G7.10 [imported] - Caenorhabditis elegans
50	85	PHAT-5 [Caenorhabditis elegans]
50	85	Y69H2.3, isoform d [Caenorhabditis elegans]
50	84	unnamed protein product [Homo sapiens]
50	83	ZC373.2 [Caenorhabditis elegans]
50	81	M04G7.1 [Caenorhabditis elegans]
50	81	cytokeratin 9 [Homo sapiens]
50	80	F54D5.3 [Caenorhabditis elegans]
50	80	DAO-2 [Caenorhabditis elegans]
50	78	SPP-14 [Caenorhabditis elegans]
50	77	LYS-7 [Caenorhabditis elegans]
50	76	T12B5.15 [Caenorhabditis elegans]
50	75	GRD-10 [Caenorhabditis elegans]
50	74	THN-2 [Caenorhabditis elegans]
50	73	F20A1.1 [Caenorhabditis elegans]
50	72	D1086.6 [Caenorhabditis elegans]
50	71	Chain A, D-Ribose-Binding Protein From Escherichia Coli
50	70	TAG-293 [Caenorhabditis elegans]
50	69	C08F11.11 [Caenorhabditis elegans]
50	67	ASP-3 [Caenorhabditis elegans]
50	67	VHA-19 [Caenorhabditis elegans]
50	66	CCG-1 [Caenorhabditis elegans]
50	65	major sperm protein [Caenorhabditis elegans]
50	63	Y45F10C.4 [Caenorhabditis elegans]
50	63	TTR-45 [Caenorhabditis elegans]
50	60	F57F4.4 [Caenorhabditis elegans]
50	58	D1086.10, isoform a [Caenorhabditis elegans]
100	290	ILYS-5, isoform a [Caenorhabditis elegans]
100	222	SRP-7, isoform a [Caenorhabditis elegans]
100	213	SPP-5 [Caenorhabditis elegans]
100	173	C08F11.11 [Caenorhabditis elegans]
100	170	C49G7.3 [Caenorhabditis elegans]
100	169	C33G8.4 [Caenorhabditis elegans]
100	164	hypothetical protein C39D10.7 - Caenorhabditis elegans
100	161	CBD-1 [Caenorhabditis elegans]
100	158	T19B10.2 [Caenorhabditis elegans]
100	157	D1086.11, isoform b [Caenorhabditis elegans]
100	156	F20A1.1 [Caenorhabditis elegans]
100	155	ILYS-3 [Caenorhabditis elegans]

100	151	C06A8.3 [Caenorhabditis elegans]
100	138	THN-2 [Caenorhabditis elegans]
100	136	protein T08A9.7 [imported] - Caenorhabditis elegans
100	135	unnamed protein product [Homo sapiens]
100	98	Keratin, type II cytoskeletal 1 [Tupaia chinensis]
100	123	TTR-2 [Caenorhabditis elegans]
100	121	C14C6.5 [Caenorhabditis elegans]
100	73	C14C6.2 [Caenorhabditis elegans]
100	116	F15A4.6 [Caenorhabditis elegans]
100	112	ASP-3 [Caenorhabditis elegans]
100	111	CPI-1 [Caenorhabditis elegans]
100	110	IRG-3 [Caenorhabditis elegans]
100	108	Chain A, D-Ribose-Binding Protein From Escherichia Coli
100	106	GRD-10 [Caenorhabditis elegans]
100	105	SPP-14 [Caenorhabditis elegans]
100	102	Y57G11B.5 [Caenorhabditis elegans]
100	102	H23N18.5 [Caenorhabditis elegans]
100	102	K11D12.13 [Caenorhabditis elegans]
100	102	F30H5.3 [Caenorhabditis elegans]
100	99	cytokeratin 9 [Homo sapiens]
100	98	F42A10.7 [Caenorhabditis elegans]
100	97	LYS-1 [Caenorhabditis elegans]
100	95	NUC-1 [Caenorhabditis elegans]
100	94	C25E10.8 [Caenorhabditis elegans]
100	93	ZC373.2 [Caenorhabditis elegans]
100	92	GRL-15 [Caenorhabditis elegans]
100	92	M04G7.1 [Caenorhabditis elegans]
100	90	LYS-7 [Caenorhabditis elegans]
100	89	E04F6.8 [Caenorhabditis elegans]
100	88	F13G11.3 [Caenorhabditis elegans]
100	87	NPA-1, isoform c [Caenorhabditis elegans]
100	87	VHA-19 [Caenorhabditis elegans]
100	85	GRD-5 [Caenorhabditis elegans]
100	85	Hypothetical protein CBG06172 [Caenorhabditis briggsae]
100	83	T24B8.5 [Caenorhabditis elegans]
100	80	DAO-2 [Caenorhabditis elegans]
100	79	PHAT-5 [Caenorhabditis elegans]
100	78	D1086.6 [Caenorhabditis elegans]
100	78	DPY-17 [Caenorhabditis elegans]
100	78	Chain A, Glutamine Binding Protein Open Ligand-Free Structure
100	76	F54D5.3 [Caenorhabditis elegans]

100	76	PREDICTED: cationic trypsin-3-like [Myotis brandtii]
100	76	major sperm protein [Caenorhabditis elegans]
100	75	hypothetical protein CAEBREN_15245 [Caenorhabditis brenneri]
100	75	D1086.10, isoform a [Caenorhabditis elegans]
100	74	unnamed protein product [Homo sapiens]
100	74	putative polypeptide, similar to enteropeptidase, PIR Accession Number A53663, enterokinase encoded by GenBank Accession Number U09859, and bone morphogenetic protein encoded by GenBank Accession Number L24755, partial [Caenorhabditis elegans]
100	73	Chain A, Ompf Porin Mutant D74a
100	72	TTR-45 [Caenorhabditis elegans]
100	72	UBI 3 fusion protein (149 AA) [Neurospora crassa]
100	71	PREDICTED: uncharacterized protein LOC106097318 [Oreochromis niloticus]
100	70	SCL-3 [Caenorhabditis elegans]
100	70	CPR-5 [Caenorhabditis elegans]
100	69	F46F11.7 [Caenorhabditis elegans]
100	68	F53F4.13 [Caenorhabditis elegans]
100	68	CLEC-83 [Caenorhabditis elegans]
100	68	Y69H2.3, isoform e [Caenorhabditis elegans]
100	67	TAG-293 [Caenorhabditis elegans]
100	67	D1054.10 [Caenorhabditis elegans]
100	66	F48G7.8 [Caenorhabditis elegans]
100	65	ASP-6 [Caenorhabditis elegans]
100	65	F57F4.4 [Caenorhabditis elegans]
100	64	T21H3.1, isoform a [Caenorhabditis elegans]
100	63	CLEC-50 [Caenorhabditis elegans]
100	63	FAR-3 [Caenorhabditis elegans]
100	63	C08A9.10 [Caenorhabditis elegans]
100	63	C. briggsae CBR-FAR-1 protein [Caenorhabditis briggsae]
100	61	Y62H9A.5 [Caenorhabditis elegans]
100	60	K06G5.1, isoform a [Caenorhabditis elegans]
100	60	FIPR-28 [Caenorhabditis elegans]
200	202	SPP-5 [Caenorhabditis elegans]
200	188	C49G7.3 [Caenorhabditis elegans]
200	159	C08F11.11 [Caenorhabditis elegans]
200	148	TTR-2 [Caenorhabditis elegans]
200	147	ILYS-5, isoform a [Caenorhabditis elegans]
200	146	CBD-1 [Caenorhabditis elegans]
200	141	hypothetical protein C39D10.7 - Caenorhabditis elegans
200	130	protein T08A9.7 [imported] - Caenorhabditis elegans
200	118	SPP-14 [Caenorhabditis elegans]

200	118	DAO-2 [Caenorhabditis elegans]
200	112	ASP-13 [Caenorhabditis elegans]
200	109	unnamed protein product [Homo sapiens]
200	79	PREDICTED: keratin, type II cytoskeletal 6B-like [Macaca mulatta]
200	107	C06A8.3 [Caenorhabditis elegans]
200	105	F15A4.6 [Caenorhabditis elegans]
200	104	NPA-1, isoform a [Caenorhabditis elegans]
200	101	THN-2 [Caenorhabditis elegans]
200	100	T19B10.2 [Caenorhabditis elegans]
200	100	SRP-7, isoform a [Caenorhabditis elegans]
200	98	F42A10.7 [Caenorhabditis elegans]
200	94	ZC373.2 [Caenorhabditis elegans]
200	90	Chain A, Ompf Porin Mutant D74a
200	89	F13G11.3 [Caenorhabditis elegans]
200	87	FAR-3 [Caenorhabditis elegans]
200	86	PHAT-5 [Caenorhabditis elegans]
200	82	PREDICTED: keratin, type I cytoskeletal 10-like isoform 2 [Macaca mulatta]
200	81	Try10-like trypsinogen precursor [Mus musculus]
200	81	T24B8.5 [Caenorhabditis elegans]
200	79	D1086.11, isoform b [Caenorhabditis elegans]
200	79	CLEC-63 [Caenorhabditis elegans]
200	78	hypothetical protein [Streptomyces purpeofuscus]
200	78	VHA-19 [Caenorhabditis elegans]
200	78	LYS-7 [Caenorhabditis elegans]
200	78	Chain A, D-Ribose-Binding Protein From Escherichia Coli
200	78	Y47G6A.33 [Caenorhabditis elegans]
200	77	PREDICTED: LOW QUALITY PROTEIN: keratin, type II cytoskeletal 3 [Octodon degus]
200	75	ASP-3 [Caenorhabditis elegans]
200	73	H23N18.5 [Caenorhabditis elegans]
200	71	unnamed protein product [Mus musculus]
200	71	Y51F10.7 [Caenorhabditis elegans]
200	71	C14C6.5 [Caenorhabditis elegans]
200	71	F15B9.8 [Caenorhabditis elegans]
200	68	Y52B11A.8 [Caenorhabditis elegans]
200	68	F41G3.10 [Caenorhabditis elegans]
200	67	D1054.10 [Caenorhabditis elegans]
200	67	keratin, type I cytoskeletal 19 [Cricetulus griseus]
200	67	INS-31 [Caenorhabditis elegans]
200	66	F53F4.13 [Caenorhabditis elegans]
200	65	CPG-2 [Caenorhabditis elegans]
200	64	CPR-5 [Caenorhabditis elegans]

200	63	E04F6.9 [Caenorhabditis elegans]
200	62	hypothetical protein ZK6.10 - Caenorhabditis elegans
200	62	CPR-4 [Caenorhabditis elegans]
200	62	W03F11.1, isoform a [Caenorhabditis elegans]
200	60	F52E1.14 [Caenorhabditis elegans]
400	202	hypothetical protein C39D10.7 - Caenorhabditis elegans
400	184	C49G7.3 [Caenorhabditis elegans]
400	173	ILYS-5, isoform a [Caenorhabditis elegans]
400	145	PHAT-5 [Caenorhabditis elegans]
400	138	PREDICTED: keratin, type II cytoskeletal 6A-like [Equus caballus]
400	125	keratin 1 [Homo sapiens]
400	97	Keratin, type II cytoskeletal 1 [Pteropus alecto]
400	135	SPP-5 [Caenorhabditis elegans]
400	117	D1054.10 [Caenorhabditis elegans]
400	116	C08F11.11 [Caenorhabditis elegans]
400	115	F35A5.2 [Caenorhabditis elegans]
400	115	CBD-1 [Caenorhabditis elegans]
400	114	DAO-2 [Caenorhabditis elegans]
400	112	Y57G11B.5 [Caenorhabditis elegans]
400	112	K11D12.13 [Caenorhabditis elegans]
400	109	CLEC-83 [Caenorhabditis elegans]
400	107	ZC373.2 [Caenorhabditis elegans]
400	106	Y52B11A.8 [Caenorhabditis elegans]
400	106	thioredoxin [Escherichia coli]
400	103	LYS-8 [Caenorhabditis elegans]
400	102	GRL-15 [Caenorhabditis elegans]
400	101	SMO-1 [Caenorhabditis elegans]
400	99	F13G11.3 [Caenorhabditis elegans]
400	99	LYS-4 [Caenorhabditis elegans]
400	99	RecName: Full=Trypsin; Flags: Precursor [Sus scrofa]
400	97	TTR-2 [Caenorhabditis elegans]
400	97	C06A8.3 [Caenorhabditis elegans]
400	96	protein T08A9.7 [imported] - Caenorhabditis elegans
400	95	T19B10.2 [Caenorhabditis elegans]
400	95	Chain A, Ompf Porin Mutant D74a
400	94	SRP-7, isoform a [Caenorhabditis elegans]
400	94	M04G7.1 [Caenorhabditis elegans]
400	92	NSPC-7 [Caenorhabditis elegans]
400	90	K12H4.7, isoform a [Caenorhabditis elegans]
400	89	NSPC-20 [Caenorhabditis elegans]

400	88	cytokeratin 9 [Homo sapiens]
400	88	F08B12.4, isoform f [Caenorhabditis elegans]
400	88	Y47G6A.33 [Caenorhabditis elegans]
400	87	ASP-6 [Caenorhabditis elegans]
400	86	SPP-14 [Caenorhabditis elegans]
400	86	PHAT-6 [Caenorhabditis elegans]
400	82	CPG-3 [Caenorhabditis elegans]
400	81	T11B7.5 [Caenorhabditis elegans]
400	80	CLEC-63 [Caenorhabditis elegans]
400	78	F53F4.13 [Caenorhabditis elegans]
400	78	DPY-14 [Caenorhabditis elegans]
400	77	CPI-2, isoform a [Caenorhabditis elegans]
400	73	D1054.11 [Caenorhabditis elegans]
400	73	LYS-1 [Caenorhabditis elegans]
400	73	D1086.11, isoform b [Caenorhabditis elegans]
400	73	Hypothetical protein CBG01307 [Caenorhabditis briggsae]
400	73	unnamed protein product [Homo sapiens]
400	73	R09H10.3, isoform a [Caenorhabditis elegans]
400	72	Chain A, D-Ribose-Binding Protein From Escherichia Coli
400	71	CPR-4 [Caenorhabditis elegans]
400	70	F41G3.10 [Caenorhabditis elegans]
400	70	FAR-2 [Caenorhabditis elegans]
400	70	C08A9.10 [Caenorhabditis elegans]
400	69	F55B11.3 [Caenorhabditis elegans]
400	68	C25E10.8 [Caenorhabditis elegans]
400	67	F59A6.12 [Caenorhabditis elegans]
400	65	W03F11.1, isoform a [Caenorhabditis elegans]
400	65	F15B9.8 [Caenorhabditis elegans]
400	64	CPG-2 [Caenorhabditis elegans]
400	62	FAR-3 [Caenorhabditis elegans]
400	62	C. briggsae CBR-TAG-293 protein [Caenorhabditis briggsae]
400	62	CLEC-50 [Caenorhabditis elegans]
400	60	TTR-51 [Caenorhabditis elegans]
500	117	NSPC-7 [Caenorhabditis elegans]
500	95	C33G8.4 [Caenorhabditis elegans]
500	87	F20A1.1 [Caenorhabditis elegans]
500	83	F54D5.3 [Caenorhabditis elegans]
500	80	F41G3.10 [Caenorhabditis elegans]
500	79	LYS-1 [Caenorhabditis elegans]
500	75	C49G7.3 [Caenorhabditis elegans]

500	74	CBD-1 [Caenorhabditis elegans]
500	72	ASP-6 [Caenorhabditis elegans]
500	70	Chain A, Structure Of Caenopore-5 (81 Pro Cis Conformer)
500	71	hypothetical protein T18H9.2 - Caenorhabditis elegans
500	70	hypothetical protein C39D10.7 - Caenorhabditis elegans

Bibliography

Ackerman, C. J., Harnett, M. M., Harnett, W., Kelly, S. M., Svergun, D. I., & Byron, O. (2003). 19Å solution structure of the filarial nematode immunomodulatory protein, ES-62. *Biophysical journal*. *84*(1), 489-500.

Acosta, D., Goni, F. and Carmona, C. (1998). Characterization and partial purification of a leucine aminopeptidase from Fasciola hepatica. *Journal of Parasitology*, 84(1), 1-7.

Adam, R., Kaltmann, B., Rudin, W., Friedrich, T., Marti, T. and Lucius, R. (1996). Identification of Chitinase as the Immunodominant Filarial Antigen Recognized by Sera of Vaccinated Rodents. *Journal of Biological Chemistry*. 271(3), 1441-1447.

Allen, M. A., Hillier, L. W., Waterston, R. H., and Blumenthal, T. (2011). A global analysis of *C. elegans* trans-splicing. *Genome Research*. 21, 255-264.

Al-Riyami L., and Harnett W. (2012). Immunomodulatory properties of ES-62, a phosphorylcholine-containing glycoprotein secreted by Acanthocheilonema viteae. *Endocrine, Metabolic and Immune Disorders Drug Targets*. 12(1), 45-52.

Ambros, V. (2004). The functions of animal microRNAs. *Nature.* 431*, 350–355*.

Avery, L. (1993). The genetics of feeding in Caenorhabditis elegans. *Genetics. 133(4),* 897-917.

Bongartz, T., Sutton. A. J., Sweeting, M. J., Buchan. I., Matteson. E. L., Montori, V. (2006). Anti-TNF Antibody Therapy in Rheumatoid Arthritis and the Risk of Serious Infections and Malignancies: Systematic Review and Meta-analysis of Rare Harmful Effects in Randomized Controlled Trials. Journal *of the American Medical Association.* 295(19), 2275-2285.

Boulin, T., Etchberger, J. F., Hobert, O. Reporter gene fusions. (2006) In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA): WormBook; 2005-. Available from: http://www.ncbi.nlm.nih.gov/books/NBK19738/.

Brenner, S. (1974). The *genetics* of Caenorhabditis elegans. *Genetics.* 77(1), 71-94.

Brown, T. A. Genomes. (2002), 2nd edition. Oxford: Wiley-Liss; Chapter 7, Understanding a Genome Sequence.

Carusco, M-E, Jenna, S., Bouchecareilh, M., et al.,. (2008) GTPase-Mediated Regulation of the Unfolded Protein Response in *Caenorhabditis elegans* Is Dependent on the AAA^+ ATPase CDC-48. *Mollecular Cell Biology*. 28(13), 4261-4274.

Chalfie, M., Y. Tu, et al.,. (1994). Green fluorescent protein as a marker for gene expression. *Science. 263*, 802–805.

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nature Reviews. Genetics*. *8*, 93-103.

Chevallet, M., Luche, S., Rabilloud, T. (2006). Silver staining of proteins in polyacrylamide gels. *Nature Protocols*. 1(4), 1852-1858.

Cipollo, J. F., Costello, C. E., and Hirschberg, C. B. (2002). The fine structure of *Caenorhabditis elegans* N-glycans. *Journal of Biological Chemistry*. *277*, 49143–49157.

Clements, J.N. (2011). Treatment of rheumatoid arthritis: A review of recommendations and emerging therapy. *Formulary Journal, Clinical Pharmacology*.

Coltherd, J. C., Rodgers, D. T., Lawrie, R. E., Al- Riyami, L., Suckling, C. J., Harnett, W., and Harnett, M. M. (2016). The parasitic worm-derived immunomodulatory, ES-62 and its drug-like small molecule analogues exhibit therapeutic potential in a model of chronic asthma. Scientific Reports. 6, 19224.

Croll, N. A. (1972). Feeding and lipid synthesis of Ancylostoma tubaeforme preinfective larvae. *Parasitology*, *64*(03), 369-378.

Dantonel, J. C., Quintin, S., Lakatos, L., Labouesse, M., and Tora, L. (2000). TBP-like factor is required for embryonic RNA polymerase II transcription in *C. elegans*. *Molecular Cell*. *6*, 715-722.

Davey, K. G., and Kan, S. P. (1968). Molting in a parasitic nematode, Phocanema decipiens. IV. Ecdysis and its control. *Canadian journal of zoology*, *46*(5), 893-898.

Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the *C. elegans* germ line. *Genes & Development*. *14*, 1578–1583.

Egan, C. A., Houston, K. M., Alcocer, M. J., Solovyova, A., Tate, R., Lochnit, G., McInnes, I. B., Harnett, M. M., Geyer, R., Byron, O., Harnett, W. (2006) Lack of immunological cross-reactivity between parasite-derived and recombinant forms of ES-62, a secreted protein of Acanthocheilonema viteae. *Parasitology*. 132(Pt 2), 263-74.

Elliott, D. E., and Weinstock, J. V. (2012). Helminth-host immunological interactions: prevention and control of immune- mediated diseases. *Annals of the New York Academy of Science; The Year in Immunology*, 1-12.

Fallon, P. G., and Alcami, A. (2006). Pathogen- derived immunomodulatory molecules: future immunotherapeutics? *Trends in Immunology*. http://doi.org/10.1016/j.it.2006.08.002

Feng, G. J., Goodridge, H. S., Harnett, M. M., Wei, X. Q., Nikolaev, A. V., Higson, A. P., Liew, F. Y. (1999). Extracellular signal-related kinase (ERK) and p38 mitogen-activated protein (MAP) kinases differentially regulate the lipopolysaccharide-mediated induction of inducible nitric oxide synthase and IL-12 in macrophages:*Leishmania* phosphoglycans subvert macrophage IL-12 production by targeting ERK MAP kinase. *Journal of Immunology.* 163, 6403.

Fleming, J. O., Cook, T. D. (2006). Multiple sclerosis and the hygiene hypothesis. *Neurology* 67(11), 2085–6.

Goodridge, H. S., Stepek, G., Harnett, W., Harnett, M. M. (2005). Signalling mechanisms underlying subversion of the immune response by the filarial nematode secreted product ES- 62. Review article. *Immunology*. 115, 296-304.

Grishkevich, V., Hashimshony, T., and Yanai, I. (2011). Core promoter T-blocks correlate with gene expression levels in *C. elegans*. *Genome Research*. *21*, 707-717.

Harnett, M. M., Kean, D. E., Boitelle, A., McGuiness. S., Thalhamer, T., et al.,. (2008). The phosphorycholine moiety of the filarial nematode immunomodulator ES- 62 is responsible for its anti- inflammatory action in arthritis. Extended report. *Annalysis of the Rheumatic Diseases*. 67, 518-523.

Harnett, W. and Harnett, M. M. (1993). Inhibition of murine B cell proliferation and down-regulation of protein kinase C levels by a phosphorylcholinecontaining filarial excretory-secretory product. *Journal of Immunology.* 151, 4829–4837.

Harnett, W. and Harnett, M. M. (1999) Phosphorylcholine: friend or foe of the immune response? *Immunology Today.* 20, 125-129.

Harnett, W. and Parkhouse, R. M. E. (1995) In: Perspectives in Nematode Physiology and Biochemistry (M. L. Sood, Ed.), 207-242, M/S Narendra Publication House, New Delhi.

Harnett, W. and Parkhouse, R. M. E. Nature and function of parasitic nematode surface and excretory-secretory antigens. In: Sood MI, Kapur J, Eds. (1995) *Perspectives in Nematode Physiology and Biochemistry*. New Delhi: Narendra Publication House, pp. 207-242.

Harnett, W., Goodridge, H. S. and Harnett, M. M. (2005) Subversion of immune cell signal transduction pathways by the secreted filarial nematode product, ES-62. *Parasitol*ogy. 130, 63-68.

Harnett, W., Grainger, M., Kapil, A., Worms, M. J. and Parkhouse, R. M. E. (1989) Origin, kinetics of circulation and fate *in vivo* of the major excretory–secretory product of *Acanthocheilonema viteae*. *Parasitology*. *99*, 229-239.

Harnett, W., Harnett, M. M. (2006). Frontiers in research review: Cutting- edge molecular approaches to therapeutics. Filarial nematode secreated product ES- 62 is an anti-inflammatory agent: therapeutic potential of small molecule derivatives and ES- 62 peptide mimetics. *Clinical and Experimental Pharmacology and Physiology*. 33, 511-518.

Harnett, W., Houston, K. M., Amess, R., and Worms, M. J. (1993) *Acanthocheilonema viteae*: phosphorylcholine is attached to the major excretory-secretory product *via* an *N*-linked glycan. *Experimental Parasitol*ogy. 77, 498-502.

Harnett, W., Houston, K. M., Tate, R., Garate, T., Apfel, H., Adam, R., Haslam, S. M., Panico, M., Paxton, T., Dell, A., Morris, H., Brzeski, H. (1999). Molecular cloning and demonstration of an aminopeptidase activity in a filarial nematode glycoprotein. *Molecular and Biochemical Parasitology*. 104(1), 11-23.

Harnett, W., McInnes, I. B., and Harnett M. M. (2004) ES-62, a filarial nematode-derived immunomodulator with anti-inflammatory potential. *Immunology Letters*. 94, 27-33.

Harnett, W., Worms, M. J., Grainger, M., Pyke, S. D. M. and Parkhouse, R. M. E. (1990). Association between circulating antigen and parasite load in a model filarial system, Acanthocheilonema viteae in jirds. *Parasitology, 101*, 435-444

Harnett, W., Worms, M. J., Kapil, A., Grainger, M., Parkhouse, R. M. E. (1989). Origin, kinetics of circulation and fate *in vivo* of the major excretory–secretory product of *Acanthocheilonema viteae* Parasitology. 99, 229–239.

Harnett, W.; Harnett, M.M. and Byron, O. (2003) Structural/ functional aspects of ES-62-A secreted immunomodulatory phosphorylcholine-containing filarial nematode glycoprotein. *Current Protein & Peptide Science*. 4, 59-71.

Haslam, S. M., and Dell, A. (2003). Hallmarks of *Caenorhabditis elegans* N-glycosylation: complexity and controversy. *Biochimie. 85*, 25–32.

Haslam, S. M., Houston, K. M., Harnett, W., Reason, A. J., Morris, H. R., and Dell, A. (1999). Structural Studies of N-Glycan;s of Filarial Parasites. Conservation of phosphorylcholine-substituted glycans among species and discovery of novel chito-oligomers. *Journal of Biological Chemistry*. 274(30), 20953–20960.

Haslam, S. M., Khoo, K. H., Houston, K. M., Harnett, W., Morris, H. R. and Dell, A. (1997). Characterisation of the phosphorylcholine-containing N-linked oligosaccharides in the excretory-secretory 62 kDa glycoprotein of Acanthocheilonema viteae. *Molecular and Biochemical Parasitology*, *85*, 53-66.

Hewitson, J. P., Grainger, J. R, Maizels, R. M. (2009). Helminth immunoregulation: The role of parasite secreted proteins in modulating host immunity. *Molecular and Biochemical Parasitology*. 167(1- 9), 1- 11.

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

Hoppert, M. (2011). Metalloenzymes. In J. Reitner & V. Thiel (Eds.), *Encyclopedia of Geobiology.* (pp. 558-563). Inbook, Dordrecht: Springer Netherlands.

Houston, K. M., and Harnett W. (2004). Structure and synthesis of nematode phosphorylcholine-containing glycoconjugates. *Parasitology.* 129, 655–661.

Houston, K. M., and Harnett, W. (1996) Prevention of attachment of phosphorylcholine to a major excretory-secretory product of *Acanthocheilonema viteae* using tunicamycin. *Journal of Parasitology*. 82, 320-324.

Huang, P., Pleasance, E. D., Maydan, J. S., Hunt-Newbury, R., O'Neil, N. J., Mah, A., Baillie, D. L., Marra, M. A., Moerman, D. G., and Jones, S. J. (2007). Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. *Genome Research*. *17*, 1478-1485.

Inokuchi, K., Mutoh, N., Matsuyama, S., Mizushima, H. (1982). Primary structure of the ompF gene that codes for a major outer membrane protein of Escherichia coli K-12. *Nucleic Acids Research*. 10, 6957-6968.

Kaltenbach, L., Horner, M. A., Rothman, J. H., and Mango, S. E. (2000). The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during *C. elegans* embryogenesis. *Molecular Cell. 6*, 705-713.

Kapp, K., Schrempf, S., Lemberg, M. K., et al.,. (2000-2013). Post-Targeting Functions of Signal Peptides. In: Madame Curie Bioscience Database [Internet]. Austin (TX): Landes Bioscience. Available from: http://www.ncbi.nlm.nih.gov/books/NBK6322/

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

Kostrouchova, M., Krause, M., Kostrouch, Z., and Rall, J. E. (1998). CHR3: a *Caenorhabditis elegans* orphan nuclear hormone receptor required for proper epidermal development and molting. Development *125*, 1617-1626.

Krause, M., and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell 49*, 753-761.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.

Lichtsteiner, S., and Tjian, R. (1993). Cloning and properties of the *Caenorhabditis elegans* TATA-box-binding protein. *Proceedings of the National Academy of Science U. S. A. 90*, 9673-9677.

Lightowlers, M. W., Rickard, M. D. (1988). Excretory-secretory products of helminth parasites: effects on host immune responses. *Parasitology*. 96, S123-S166.

Lochnit, G., Grabitzki, J., Henkel, B., Tavernarakis, N., Geyer, R. (2006). First identification of a phosphorylcholine- substituted protein from *Caenorhabditis elegans:* isolation and characterizationof the aspartyl protease ASP-6. *The Journal of Biological Chemistry*. 387, 1487-1493.

Lucius, R., and Textor, G. (1995) Acanthocheilonema viteae: rational design of the life cycle to increase production of parasite material using less experimental animals. *Applied Parasitology*. 36(1), 22-33.

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature 416*, 499–506.

McInnes, I. B., and Georg S. (2011). The pathogenesis of rheumatoid arthritis. *New England Journal of Medicine* 365(23), 2205-2219.

McInnes, I. B., Leung, B. P., Harnett, M., Gracie, J. A., Liew, F. Y. and Harnett, W. (2003). A novel therapeutic approach targeting articular inflammation using the filarial nematode-derived phosphorylcholine-containing glycoprotein ES-62. *The Journal of Immunology*. 171, 2127-2133.

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

Mello, C., and Fire, A. (1995). DNA transformation. *Methods Cell Biology*. *48*, 451–482.

Morton, J. J., and Blumenthal, T. (2011). Identification of transcription start sites of trans-spliced genes: uncovering unusual operon arrangements. *RNA 17*, 327-337.

Murray, L., Geldhof, P., Clark, D., Knox, D. P., Britton, C. (2007). Expression and purification of an active cysteine protease of *Haemonchus contortus* using *Caenorhabditis elegans*. *International Journal of Parasitology*. 37, 1117-1125.

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

Natsuka, S., Adachi, J., Kawaguchi, M., Nakakita, S., Hase, S., Ichikawa, A., and Ikura, K. (2002). Structural analysis of N-linked glycans in *Caenorhabditis elegans*. *The Journal of Biochemistry*. (Tokyo) *131*, 807–813.

Nelson, F. K, Albert, P. S., Riddle, D. L. (1983). Fine structure of the Caenorhabditis elegans secretory-excretory system. *Journal of Ultrastructure Research*. 82(2), 156-71.

Nott, A., Le Hir, H., and Moore, M. J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes & Development*. *18*, 210–222.

Okada, H., Kuhn, C., Feillet, H., Bach, J. F. (2010). The 'hygiene hypothesis' for autoimmuneand allergic diseases: an update. *Clinical Experimental Immunology*. 160(1), 1–9.

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A., and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. *Genetics 135*, 385-404.

Page, G., Lebecque, S., Miossec, P. (2002). Anatomic localization of immature and mature dendritic cells in an ectopic lymphoid organ: correlation with selective chemokine expression in rheumatoid synovium. *Journal of Immunology.* 168, 5333.

Panda, A. K., Ravindran, B., Das, B. K. (2013). Rheumatoid arthritis patients are free of filarialinfection in an area where filariasis is endemic: comment on the article by Pineda et al.,. *Arthritis & Rheumatology* 65(5), 1402–3.

Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18), 3551–67.

Pillai, S., Kalinna, B., Liebau, E., Hartmann, S., Theuring, F., Lucius, R. (2005). Studies on Acanthocheilonema viteae cystatin: Genomic organization, promoter studies and expression in Caenorhabditis elegans. *Filaria Journal*.4, 9.

Pineda, M. A., McGrath, M. A., Smith, P. C., Al-Riyami, L., Rzepecka, J., Gracie, J. A., Harnett, W., and Harnett, M. M. (2012). The Parasitic Helminth Product ES-62 Suppresses Pathogenesis in Collagen-Induced Arthritis by Targeting the Interleukin-17–Producing Cellular Network at Multiple Sites. Arthritis & Rheumatism. 64(10), 3168–3178

Pineda, M. A., Rodgers, D. T., Al-Riyami, L., Harnett, W., Harnett, M. M. (2014) ES-62 protects against collagen-induced arthritis by resetting interleukin-22 towards resolution of inflammation in the joints. Arthritis and Rheumatism. 66(6), 1492–1503.

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols, 2(8), 1896-1906.

Redemann, S., Schloissnig, S., Ernst, S., Pozniakowsky, A., Ayloo, S., Hyman, A., A., & Bringmann, H. (2011). Codon adaptation–based control of protein expression in *C. elegans. Nature Methods* 8, 250–252.

Reinke, V, Krause, M, Okkema, P. (2005). Transcriptional regulation of gene expression in C. elegans In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA): WormBook.

Rhoads, R. E., Dinkova, T. D., Korneeva, N. L. (2006). Mechanism and regulation of translation in C. elegans. In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA). Available from: http://www.ncbi.nlm.nih.gov/books/NBK19664/

Riddle, D. L., Blumenthal, T., Meyer, B. J., et al., editors. (1997). C. elegans II. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. Section II, Feeding.

Riddle, D. L., Swanson, M. M., & Albert, P. S. (1981). Interacting genes in nematode dauer larva formation. *Nature*. *290*(5808), 668-671.

Roberts, L. S. and Janovy, J. (2008). *Foundations of Parasitology*, 8th edition, McGraw-Hill.

Rodgers, D. T., Pineda, M. A., McGrath, M. A., Al-Riyami, L., Harnett, W., & Harnett, M. M. (2014). Protection against collagen-induced arthritis in mice afforded by the parasitic worm product, ES-62, is associated with restoration of the levels of interleukin-10-producing B cells and reduced plasma cell infiltration of the joints. *Immunology*, *141*(3), 457–466.

Rook, G. A. (2012). Hygiene hypothesis and autoimmune diseases. *Clinical Reviews in Allergy and Immunology*. 42(1), 5–15.

Rzepecka, J., Coates, M. L., Saggar, M., Al-Riyami, L., Coltherd, J., Tay, H. K., … Harnett, W. (2014). Small molecule analogues of the immunomodulatory parasitic helminth product ES-62 have anti-allergy properties. *International Journal for Parasitology*, *44*(9), 669–674.

Smolen, J. S., Aletaha, D., Koeller, M., Weisman, M. H., Emery, P. (2007). New therapies for treatment of rheumatoid arthritis. *Lancet* 370, 1861-74.
Smolen, J. S., Steiner, G. (2003). Therapeutic strategies for rheumatoid arthritis. *Nature Reviews Drug Discovery*. 2, 473-88.

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

Stepek, G., Auchie, M., Tate, R., Watson, K., Russel, D. G., Devaney, E., and Harnett, W. (2002) Expression of the filarial nematode phosphorycholine-containing glycoprotein, ES-62, is stage specific. *Parasitology*. 125, 155-164.

Stepek, G., Houston, K. M., Goodridge, H. S., Devaney, E., and Harnett, W. (2004) Stage-specific and species-specific differences in the production of the mRNA and protein for the filarial nematode secreted product, ES-62. *Parasitology*, 128, 91-98.

Stevens, J., Spang, A. (2013) N-glycosylation is required for secretion and mitosis in *C. elegans*. PLoS One. 8(5), e63687

Stiernagle, T. (2006). Maintenance of *C. elegans*. In: WormBook: The Online Review of *C. elegans* Biology [Internet]. Pasadena (CA): Available from: http://www.ncbi.nlm.nih.gov/books/NBK19649/

Strachan, D. P. (1989). Hay fever, hygiene, and household size. *BMJ : British Medical Journal*, *299*(6710), 1259–1260.

Subramanian, S., Stolk, W. A., Ramaiah, K. D.; Plaisier, A. P., Krishnamoorthy, K., Van Oortmarssen, G. J., Dominic Amalraj, D., Habbema, J. D. and Das, P. K. (2004) The dynamics of *Wuchereria bancrofti* infection: a model-based analysis of longitudinal data from Pondicherry, India. *Parasitology*, 128, 467-482.

Tak, P. P., Bresnihan, B. (2000) The pathogenesis and prevention of joint damage in rheumatoid arthritis: advances from synovial biopsy and tissue analysis. *Arthritis & Rheumatology*. 43, 2619–2633.

Taylor, A. (1993). Aminopeptidases: towards a mechanism of action. *Trends in Biochemical Sciences*. 18, 167-72.

Thomas, R., MacDonald, K. P., Pettit, A. R., Cavanagh, L. L., Padmanabha, J., Zehntner, S. (1999). Dendritic cells and the pathogenesis of rheumatoid arthritis. *Journal of Leukocyte Biology*. 66, 286.

Uhlig, T., Moe, R. H., Kvien, T. K. (2014). The Burden of Disease in Rheumatoid Arthritis. *Pharmacoeconomics*. 32(9), 841-51.

Walker, A. K., Shi, Y., and Blackwell, T. K. (2004). An extensive requirement for transcription factor IID-specific TAF-1 in *Caenorhabditis elegans* embryonic transcription. *Journal of Biological Chemistry*. *279*, 15339-15347.

Webster, J., Oxley, D. (2012). Protein identification by MALDI-TOF mass spectrometry. *Methods in Molecular Biology*. 800, 227-40.

Weinkove, D., Halstead, J. R., Gems, D., & Divecha, N. (2006). Long-term starvation and ageing induce AGE-1/PI 3-kinase-dependent translocation of DAF-16/FOXO to the cytoplasm. *BMC biology*, *4*(1), 1.

Weinstein, P. P. (1952). Regulation of water balance as a function of the excretory system of the filariform larvae of *Nippostrongylus muris* and *Ancylostoma caninum*. *Experimental Parasitology*. 1, 363-376.

Weinstock, J. V., and Elliott, D. E. (2014). Helminth infections decrease host susceptibility to immune-mediated diseases. *The Journal of Immunology*. *193*(7), 3239-3247.

Whelan, M., Harnett, M. M., Houston, K. M., Patel, V., Harnett, W., Rigley, K. P. (2000). A filarial nematode-secreted product signals dendritic cells to acquire a phenotype that drives development of Th2 cells. *Journal of Immunology.* 164, 6453

Vivar, Nancy, and Ronald F. Van Vollenhoven. (2014). Advances in the treatment of rheumatoid arthritis. *F1000Prime Reports* 6, 31.

Vogel, C., and Chothia, C. (2006). Protein family expansions and biological complexity. *PLoS Computational Biology*. *2*, e48.

Yazdanbakhsh, M., van den Biggelaar, A., Maizels, R. M. (2001). Th2 responses withoutatopy: immunoregulation in chronic helminth infections and reduced allergicdisease. *Trends Immunology*. 22(7), 372–7.

Yeku, O., Frochman, M. A. (2011). Rapid amplification of cDNA ends (RACE), *Methods in Molecular Biology*. 703, 107-22.

Yin, J., Yu, L., and Savage-Dunn, C. (2010). Alternative trans-splicing of *Caenorhabditis elegans sma-9/schnurri* generates a short transcript that provides tissue-specific function in BMP signaling. *BMC Molecular Biology*. *11*, 46.

Zhao, Z., Fang, L., Chen, N., Johnsen, R. C., Stein, L., Baillie, D. L. (2005). Distinct regulatory elements mediate similar expression patterns in the excretory cell of *Caenorhabditis elegans*. *Journal of Biological Chemistry*. 46, 38787-38794.