



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>  
[research-enlighten@glasgow.ac.uk](mailto:research-enlighten@glasgow.ac.uk)

**Components of *Caenorhabditis elegans* cuticular  
biosynthesis and the molecular basis of  
morphological mutants**

**Melanie C. Thein**

**Wellcome Centre for Molecular Parasitology  
University of Glasgow**

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

**May 2005**

---

ProQuest Number: 10390775

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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



ProQuest 10390775

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

GLASGOW  
UNIVERSITY  
LIBRARY:

## **Dedication**

This thesis is dedicated to Mum, Dad and Steph  
and is in loving memory  
of Opa and Grandpa.

## **Declaration**

The work presented in this thesis was performed entirely by the author except where indicated. This thesis contains unique work and will not be submitted for any other degree, diploma or qualification at any other university.

Melanie Thein, May 2005

## Acknowledgements

First and foremost, I would like to thank my supervisor, Tony Page, who has been a constant source of encouragement, ideas and support, and who gave me the opportunity to do this project.

The following people are also thanked for their kind provision of reagents and materials: C. Shoemaker for the BA7.1 construct, Iain Johnstone for the anti-DPY-7 antibodies and staged cDNA samples, A. Fire for the pPD series of vectors, A. Frand for the pPAF vector, the CGC for nematode strains and J. Ahringer for the RNAi library. In addition, thank you to Gillian McCormack for *sqt-1*, *sqt-3* and *col-19* feeding strains, and to Alan Winter for the preparation of BV-expressed COL-19.

Thanks to my assessors, John Gilcard and Tansy Hammarton, for their helpful discussions and to Brett Roberts for proof-reading this thesis. I would like to acknowledge all the people, past and present, in the lab: Jill, Alan, Andrew, Brett, Caroline, Alex, Linda, Martyn, Sylvain, Collette, Iain, and Carol for their help. Many firm friendships have helped me through this time in Glasgow, so thank you to everyone on the 3<sup>rd</sup> floor of WCMP. I would especially like to say thanks to my friends Ali, Stalo and Claire. All of the above people and many other friends have made this a wonderful experience. Thank you also to Keith for bearing with me and for the encouragement.

This thesis is dedicated to my family and is in loving memory of Opa and Grandpa. Mum, Dad and Steph, you have given me everything and have been untiring in your support, thank you.

# Contents

Declaration	i
Acknowledgements	ii
Contents	iii
List of figures and tables	x
Abbreviations	xiii
Summary	xvi

<b>CHAPTER 1 GENERAL INTRODUCTION</b> .....	<b>1</b>
<b>1.1 <i>C. elegans</i> as a model organism</b> .....	<b>2</b>
<b>1.2 The <i>C. elegans</i> hypodermis</b> .....	<b>4</b>
<b>1.3 Actin is required for elongation during embryogenesis</b> .....	<b>5</b>
<b>1.4 The hypodermally secreted cuticle</b> .....	<b>5</b>
<b>1.5 Components of the <i>C. elegans</i> cuticle</b> .....	<b>9</b>
<b>1.6 Vertebrate collagens</b> .....	<b>10</b>
1.6.1 Vertebrate collagen processing.....	11
<b>1.7 Human collagen diseases</b> .....	<b>15</b>
<b>1.8 <i>C. elegans</i> cuticle collagens</b> .....	<b>17</b>
1.8.1 Structure of <i>C. elegans</i> cuticle collagens.....	18
<b>1.9 <i>C. elegans</i> morphological mutants</b> .....	<b>20</b>
<b>1.10 Body shape mutants</b> .....	<b>22</b>
1.10.1 Collagen gene mutants .....	22
1.10.1.1 Blistered collagen mutants .....	24
1.10.1.2 Roller collagen mutants .....	26
1.10.1.3 Squat collagen mutants.....	26
1.10.1.4 Dumpy collagen mutants.....	30
<b>1.11 Mutants in collagen biosynthesis</b> .....	<b>31</b>
1.11.1 PHY-1 (DPY-18) and its PDI subunit.....	31
1.11.2 A <i>cis-trans</i> isomerase (PPI) and additional roles of PDI-2 and PHYs.....	33
1.11.3 Proteases involved in collagen biosynthesis.....	34
1.11.3.1 An N-terminal collagen protease: BLI-4.....	34
1.11.3.2 Collagen C-terminal proteases: Astacins.....	36
1.11.4 Regulators of protease activity .....	38
1.11.4.1 BLI-5 is a serine protease inhibitor .....	38
1.11.5 Collagen cross-links and their catalytic enzymes.....	38
1.11.5.1 DPY-11 is a thioredoxin.....	39
1.11.5.2 Tyrosine-derived cross-links are catalysed by peroxidases.....	40

1.11.5.3	Duox: a dual oxidase that catalyses the formation of tyrosine cross- links 41	
1.11.5.4	Transglutaminases and PDIs .....	42
<b>1.12</b>	<b>Body length mutants.....</b>	<b>43</b>
1.12.1	The TGF- $\beta$ pathways.....	43
1.12.2	The DBL-1 pathway is branched and controls male tail patterning and body length. 44	
1.12.3	Interaction between components controlling body size and the cuticle .....	45
<b>1.13</b>	<b>Other factors involved in <i>C. elegans</i> morphogenesis and moulting .....</b>	<b>47</b>
1.13.1	Proteases .....	47
1.13.1.1	Moulting defects as a consequence of aberrant protease activity.....	48
<b>1.14</b>	<b>Regulators of moulting.....</b>	<b>50</b>
<b>1.15</b>	<b>Basement membrane collagens .....</b>	<b>50</b>
<b>1.16</b>	<b>RNA-mediated interference (RNAi) .....</b>	<b>51</b>
<b>1.17</b>	<b>Project Aims.....</b>	<b>55</b>
<b>2</b>	<b>CHAPTER 2 MATERIALS AND METHODS.....</b>	<b>56</b>
<b>2.1</b>	<b>Chemical abbreviations.....</b>	<b>57</b>
<b>2.2</b>	<b>Reagents.....</b>	<b>58</b>
<b>2.3</b>	<b>Standard media.....</b>	<b>58</b>
<b>2.4</b>	<b><i>E. coli</i> strains.....</b>	<b>61</b>
<b>2.5</b>	<b>Vectors and clones .....</b>	<b>62</b>
<b>2.6</b>	<b>Bacterial cultures.....</b>	<b>63</b>
2.6.1	Antibiotic concentrations for bacterial culture .....	63
2.6.2	Bacterial culture on solid media .....	63
2.6.3	Liquid cultures of bacteria.....	63
2.6.4	Making competent M15 cells .....	63
<b>2.7</b>	<b><i>C. elegans</i> culture.....</b>	<b>63</b>
2.7.1	<i>C. elegans</i> strains.....	64
2.7.2	Long term storage of nematodes .....	65
<b>2.8</b>	<b>Agarose gel .....</b>	<b>65</b>
<b>2.9</b>	<b>Purification and synthesis of genomic DNA and cDNA.....</b>	<b>65</b>
2.9.1	Phenol:chloroform extraction and ethanol precipitation of DNA and RNA ..	65
2.9.2	Genomic DNA isolation .....	66
2.9.3	Production of mixed stage cDNA.....	66
2.9.3.1	Total RNA isolation .....	67
2.9.3.2	Purification of mRNA .....	67

2.9.3.3	Synthesis of first strand cDNA.....	67
<b>2.10</b>	<b>Polymerase chain reaction (PCR)-based cloning procedures .....</b>	<b>67</b>
2.10.1	PCR conditions and polymerases .....	67
2.10.2	Purification of PCR products.....	68
<b>2.11</b>	<b>DNA transfer vectors.....</b>	<b>69</b>
<b>2.12</b>	<b>Transformation of <i>E. coli</i>.....</b>	<b>69</b>
2.12.1	Standard <i>E. coli</i> strains .....	69
2.12.2	HT115(DE3) strains .....	70
<b>2.13</b>	<b>Identification of bacterial transformants .....</b>	<b>70</b>
2.13.1	Bluc/white selection .....	70
2.13.2	PCR screening .....	70
<b>2.14</b>	<b>Plasmid DNA preparation .....</b>	<b>71</b>
<b>2.15</b>	<b>Restriction endonuclease digestion .....</b>	<b>71</b>
<b>2.16</b>	<b>5' end dephosphorylation of linearised vector DNA .....</b>	<b>72</b>
<b>2.17</b>	<b>Ligations into other vectors .....</b>	<b>72</b>
<b>2.18</b>	<b>Sequence analysis.....</b>	<b>73</b>
<b>2.19</b>	<b>Bioinformatic resources .....</b>	<b>73</b>
<b>2.20</b>	<b>Semi quantitative reverse transcriptase PCR (SQ RT-PCR).....</b>	<b>73</b>
<b>2.21</b>	<b>Southern blotting.....</b>	<b>75</b>
2.21.1	Probing the membrane.....	76
<b>2.22</b>	<b>Standard <i>C. elegans</i> techniques.....</b>	<b>76</b>
2.22.1	Generation of males.....	76
2.22.2	Crosses.....	77
2.22.3	Synchronising populations .....	77
2.22.4	Transformation of <i>C. elegans</i> .....	78
2.22.5	Microinjection .....	78
2.22.5.1	Behaviour of injected DNA.....	79
2.22.6	Maintenance of transgenic strains carrying a free array.....	79
2.22.7	TP12 integration of extrachromosomal arrays .....	79
<b>2.23</b>	<b>Single Worm PCR .....</b>	<b>80</b>
<b>2.24</b>	<b>Microscopy .....</b>	<b>80</b>
<b>2.25</b>	<b>Preparation of whole worms.....</b>	<b>80</b>
<b>2.26</b>	<b>Preparation of freeze-cracked specimens for immunolocalisation .....</b>	<b>81</b>
<b>2.27</b>	<b>Immunolocalisation procedure for whole worm staining.....</b>	<b>81</b>

<b>2.28</b>	<b>Optimisation of antibody staining using anti COL-19 antibodies .....</b>	<b>82</b>
2.28.1	'Heavy fixation' of freeze-cracked slides for immunolocalisation .....	83
2.28.2	Cuticle extraction.....	83
2.28.3	Pre-absorption of antisera.....	84
<b>2.29</b>	<b>Promoter analyses .....</b>	<b>84</b>
2.29.1	The <i>hpx-1</i> / <i>lacZ</i> promoter/reporter gene construct.....	84
2.29.2	Staining for $\beta$ -galactosidase activity .....	85
2.29.3	Cloning of the <i>hpx-1</i> promoter sequence into a vector with a destabilised EGFP .....	86
<b>2.30</b>	<b>RNA interference (RNAi) .....</b>	<b>87</b>
2.30.1	Construction of plasmids.....	87
	<i>dpy-7</i> .....	88
	<i>dpy-13</i> .....	88
	Table 2.6: Primers used for amplification of RNAi constructs. For each gene, the pair of primers utilised is given. In addition, the size of the amplification product, the transfer vector utilised and the restriction sites used for ligation into the L4440 plasmid, are given.....	88
2.30.2	RNAi by feeding.....	89
2.30.3	<i>In vitro</i> transcription of RNA .....	89
2.30.4	dsRNA microinjections .....	90
2.30.5	RNAi by soaking ( <i>col-19</i> ) .....	90
<b>2.31</b>	<b>Construction of a Ty-tagged COL-19 .....</b>	<b>90</b>
<b>2.32</b>	<b>Re-ligation of the Ty-tag into the <i>col-19</i> coding sequence .....</b>	<b>94</b>
2.32.1	Polyacrylamide gel purification of small DNA fragments.....	97
2.32.2	Ligation of the Ty-tag into the Gly-X-Y sequence of COL-19.....	98
<b>2.33</b>	<b>Promoter swap construction.....</b>	<b>99</b>
2.33.1	GFP fusion collagens under the <i>col-19</i> promoter.....	99
2.33.2	GFP fusion collagens under the <i>sqt-1</i> promoter .....	99
<b>2.34</b>	<b>Fusing <i>sqt-1</i> and its promoter to a GFP marker.....</b>	<b>102</b>
<b>2.35</b>	<b>Preparation of samples for scanning electron microscopy (SEM).....</b>	<b>105</b>
<b>2.36</b>	<b>Protein techniques .....</b>	<b>105</b>
2.36.1	SDS polyacrylamide gel electrophoresis (SDS PAGE) .....	105
2.36.2	Transfer of protein to PVDF membranes .....	106
2.36.3	Construct preparation for expression of <i>col-19</i> in a baculovirus system .....	106
2.36.4	Trypsin treatment of BV expression-derived recombinant COL-19 protein .....	107
2.36.5	Expression and purification of recombinant His-tagged HPX-1 protein .....	108
2.36.6	Purification of recombinant HPX-1 protein .....	108
2.36.7	Determination of protein concentration.....	110
2.36.8	Western Analysis of HPX-1 recombinant proteins .....	110
2.36.9	Peroxidase activity detection.....	110

<b>Table 2.7:</b>	<b>Primer list. Names and sequences of the primers utilised. ....</b>	<b>111</b>
-------------------	---	------------

<b>3</b>	<b>CHAPTER 3 THE EXOSKELETON COLLAGEN, COL-19, AS AN ADULT-SPECIFIC MARKER IN THE <i>C. ELEGANS</i> CUTICLE .....</b>	<b>115</b>
3.1	Introduction .....	116
3.2	Results.....	119
3.2.1	The wild type pattern of COL-19::GFP in the strain TP12 .....	119
3.2.2	<i>col-19</i> expression coincides with <i>col-12</i> .....	124
3.2.3	COL-19 is non essential but its RNAi-mediated depletion leads to adult-specific alae defects.....	127
3.2.4	Confirming the wild-type expression of COL-19.....	129
3.2.4.1	Immunofluorescence of COL-19 .....	129
3.2.4.2	Tagging the C-terminus of COL-19 with a Ty-tag.....	129
3.2.4.3	Inserting the Ty-tag to interfere with the Gly-X-Y repeat region .....	131
3.2.5	Trypsin digestion of recombinant COL-19 from a BV expression system ..	132
3.2.6	SQT-1 cannot form a functional GFP fusion .....	135
3.2.7	Investigating the role of collagen promoters .....	135
3.2.7.1	Collagen fusions under the influence of the influence of the <i>col-19</i> promoter	136
3.2.7.2	<i>col-19</i> under the influence of the <i>sqt-1</i> promoter .....	136
3.3	Chapter 3 Discussion .....	139
3.3.1	COL-19::GFP as a <i>C. elegans</i> cuticular marker .....	139
3.3.2	COL-19::GFP localisation represents that of wild type COL-19 .....	141
3.3.3	The role of COL-19 in the cuticle .....	142
3.3.4	The properties of COL-19::GFP that optimise it for being a useful marker	143
3.3.5	Investigation into the role of the promoter in collagen processing .....	144
3.4	Future work .....	146
<b>4</b>	<b>CHAPTER 4 ANALYSIS OF THE CUTICLES OF MORPHOLOGICAL MUTANTS AND AN RNAI SCREEN TO FIND GENES INVOLVED IN CUTICLE SYNTHESIS. ....</b>	<b>148</b>
4.1	Introduction .....	149
4.2	Results 1: Crosses between TP12 and <i>C. elegans</i> morphological mutants...	151
4.2.1	Five classes of COL-19::GFP disruption are observed .....	151
4.2.2	The COL-19::GFP expression pattern in Long and Sma mutants.....	157
4.2.3	The COL-19::GFP disruption in <i>rol-6</i> mutants .....	160
4.2.4	The COL-19::GFP disruption in Sqt mutants.....	161
4.2.5	The COL-19::GFP disruption in cuticle collagen Dpy mutants.....	166
4.2.5.1	Class III mutants are exemplified by <i>dpy-5</i> mutants .....	169
4.2.5.2	MH27 staining in <i>dpy-5(e61)</i> mutants.....	174
4.2.5.3	The effect of treating different stage larvae with <i>dpy-5</i> RNAi.....	174
4.2.5.4	Class II COL-19::GFP disruption is exemplified by <i>dpy-7</i> mutants .....	175
4.2.6	The COL-19::GFP disruption resulting from mutations in the collagen processing enzymes display type V COL-19::GFP disruption.....	178
4.2.6.1	COL-19::GFP disruption in <i>dpy-11</i> mutants .....	178
4.2.6.2	COL-19::GFP disruption in <i>dpy-18</i> mutants .....	181
4.2.7	The COL-19::GFP expression pattern in Bli mutants .....	182

4.2.7.1	BLI-4 subtilisin protease mutants exhibit a wild type COL-19::GFP expression pattern.....	182
4.2.7.2	The COL-19::GFP expression pattern in BLI-1 mutants .....	184
4.2.7.3	The COL-19::GFP expression pattern in <i>bli-5</i> mutants .....	184
<b>4.3</b>	<b>Results 2: An RNAi screen to identify components in cuticle synthesis.....</b>	<b>185</b>
4.3.1	Selection of gene targets.....	187
4.3.2	General trends from RNAi data.....	192
4.3.3	Correlation between gross body morphology and COL-19::GFP pattern....	193
4.3.4	Trends in the data .....	194
4.3.5	Tyrosinases and thioredoxins do not have severe RNAi effects .....	194
4.3.6	RNAi of metalloproteases .....	195
4.3.6.1	RNAi of astacins.....	195
4.3.7	RNAi of cuticle collagens .....	198
4.3.7.1	Identification of two novel collagen genes.....	199
4.3.8	RNAi of peroxidases .....	199
4.3.9	RNAi of other genes.....	199
4.3.10	Searching for trends within similarly expressed genes .....	205
<b>4.4</b>	<b>Discussion .....</b>	<b>205</b>
4.4.1	Examination of data from Rol mutants .....	206
4.4.2	Examination of data from Sqt mutants .....	207
4.4.2.1	<i>sqt-1</i> mutants .....	207
4.4.2.2	<i>sqt-3</i> mutants .....	209
4.4.3	Examination of Dpy mutants.....	210
4.4.3.1	Class III mutants.....	211
4.4.3.2	Class II mutants .....	212
4.4.4	Examination of <i>bli-1</i> mutants .....	213
4.4.5	Overview of the roles of collagens in the <i>C. elegans</i> cuticle .....	213
4.4.6	Examination of cuticle processing enzyme mutants.....	215
4.4.7	Overview of cuticular enzyme mutants .....	216
4.4.8	Examination of body size mutants.....	217
4.4.9	Analysis of data from the RNAi screen.....	217
4.4.9.1	Validity of the RNAi technique for this screen .....	218
4.4.9.2	Results from RNAi .....	219
4.4.9.3	Overview of the applications of the RNAi screen.....	221
<b>4.5</b>	<b>Future work .....</b>	<b>222</b>
<b>5</b>	<b>CHAPTER 5 ANALYSIS OF HPX-1 (ZK430.8), A NOVEL CROSS-LINKING ENZYME.....</b>	<b>223</b>
<b>5.1</b>	<b>Introduction .....</b>	<b>224</b>
5.1.1	Aim and introduction to animal haem peroxidases .....	224
5.1.2	Duox enzymes are H <sub>2</sub> O <sub>2</sub> generating enzymes with potential peroxidase activity .....	227
5.1.3	Peroxidase roles that do not involve cross-linking.....	229
5.1.4	Tyrosine cross-linking activity of peroxidases.....	231
5.1.5	<i>C. elegans</i> Duox enzymes .....	232

<b>5.2</b>	<b>Results</b> .....	<b>235</b>
5.2.1	Targeting of peroxidase containing genes by RNAi .....	235
5.2.2	Analysis of <i>bli-3</i> mutant alleles using TP12.....	237
5.2.3	<i>bli-3</i> RNAi using TP12.....	237
5.2.4	DPY-7 staining of <i>bli-3</i> RNAi-treated worms.....	241
5.2.5	Isolation of <i>bli-3</i> deletion mutant .....	242
5.2.6	Sequence analysis of <i>hpx-1</i> /HPX-1 .....	242
5.2.7	RNAi of <i>hpx-1</i> .....	247
5.2.8	DPY-7 staining of <i>hpx-1</i> RNAi-treated worms .....	253
5.2.9	SEM of <i>hpx-1</i> RNAi-treated animals .....	253
5.2.10	Isolation of <i>hpx-1</i> deletion mutant.....	256
5.2.11	Double targeting of <i>bli-3</i> and <i>hpx-1</i> .....	256
5.2.11.1	RNAi-targeting of <i>hpx-1</i> in the strain <i>bli-3(e767)</i> .....	257
5.2.11.2	RNAi targeting of <i>hpx-1</i> in <i>bli-3(n529)</i> .....	257
5.2.12	Spatial expression analysis of <i>hpx-1</i> .....	261
5.2.13	SQ RT-PCR of <i>hpx-1</i> .....	263
5.2.14	Temporal expression of <i>hpx-1</i> determined by a destabilised EGFP marker .....	263
5.2.15	Expression of recombinant HPX-1 .....	266
<b>5.3</b>	<b>Discussion</b> .....	<b>268</b>
5.3.1	Analysis of BLI-3, f53g12.3 and HPX-1 at the sequence level .....	269
5.3.2	<i>hpx-1</i> temporal and spatial expression.....	270
5.3.3	Assaying the peroxidase activity of recombinant HPX-1 .....	271
5.3.4	RNAi of <i>bli-3</i> and <i>hpx-1</i> .....	271
5.3.5	DPY-7 localisation in <i>hpx-1</i> and <i>bli-3</i> RNAi-treated animals.....	273
5.3.6	<i>bli-3</i> mutant strains .....	274
5.3.7	Double targeting of <i>bli-3</i> and <i>hpx-1</i> .....	275
5.3.8	Overview of <i>bli-3</i> and <i>hpx-1</i> .....	278
5.3.9	Function of HPX-1 .....	279
<b>5.4</b>	<b>Future work</b> .....	<b>281</b>
<b>6</b>	<b>CHAPTER 6 GENERAL DISCUSSION</b> .....	<b>283</b>
<b>6.1</b>	<b>Aim</b> .....	<b>284</b>
<b>6.2</b>	<b>Wild type expression of <i>col-19</i></b> .....	<b>284</b>
<b>6.3</b>	<b>Processing of collagens</b> .....	<b>284</b>
<b>6.4</b>	<b>The molecular bases of morphological mutants</b> .....	<b>284</b>
6.4.1	Lon mutants .....	284
6.4.2	Dpy mutants.....	285
6.4.3	Bli mutants.....	285
6.4.4	Sqt and Rol mutants.....	285
<b>6.5</b>	<b>Enzymes in cuticle synthesis and RNAi</b> .....	<b>286</b>
<b>6.6</b>	<b>HPX-1</b> .....	<b>286</b>

## Tables and figures

### Chapter 1:

Figure 1.1:	Diagrammatic representation of the <i>C. elegans</i> cuticle.	8
Figure 1.2:	An outline view of the key steps in collagen biosynthesis.	13
Figure 1.3:	The structure of a typical <i>C. elegans</i> cuticle collagen.	19
Figure 1.4:	Typical gross morphological phenotypes resulting from cuticle defects.	21
Table 1.1:	Morphological <i>C. elegans</i> collagen mutants and their phenotypes according to their genetic lesions.	23
Figure 1.5:	The body length-determining pathway.	46
Figure 1.6:	Mechanism of dsRNA-mediated mRNA silencing	53

### Chapter 2:

Table 2.1:	<i>E. coli</i> strains	61
Table 2.2:	Vectors and clones	62
Table 2.3:	<i>C. elegans</i> strains	64
Table 2.4:	Primers utilised for RT-PCR	74
Table 2.5:	Conditions of antibody staining	82
Table 2.6:	Primers for amplification of RNAi constructs	87
Figure 2.1:	Construction of a C-terminally Ty-tagged <i>col-19</i>	93
Figure 2.2:	Ligation of a Ty-tag into the <i>col-19</i> coding sequence	96
Figure 2.3:	Creation of a promoter swap construct: <i>sqt-1</i> coding sequence under the influence of the <i>col-19</i> promoter	101
Figure 2.4:	Creation of a promoter swap construct: <i>col-19</i> coding sequence under the influence of the <i>sqt-1</i> promoter.	104
Table 2.7:	Primers list	110

### Chapter 3:

Figure 3.1:	<i>col-19::gfp</i> fusion construct, BA7.1.	118
Figure 3.2:	Gene and protein structure of <i>col-19</i> /COL-19.	120
Figure 3.3:	The wild type TP12 cuticle: COL-19::GFP and DPY-7 localisation.	122
Figure 3.4:	The changing expression of COL-19::GFP in aging adult TP12 worms.	125
Figure 3.5:	Temporal expression of <i>col-19</i> via a SQ RT-PCR assay.	126

Figure 3.6:	Nomarski images of alae in untreated TP12 animals and animals targeted by <i>col-19</i> (RNAi).	128
Figure 3.7:	Expression pattern of Ty-tagged COL-19 in the cuticle of an adult worm	130
Figure 3.8:	Western analysis of BV-expressed recombinant COL-19 after trypsin treatment and under reducing and non-reducing conditions.	134
Figure 3.9:	Results from promoter swap experiments	138
<b>Chapter 4:</b>		
Table 4.1:	Table of morphological mutants.	155
Figure 4.1:	The 5 classes of COL-19::GFP disruption in morphological mutants.	153
Figure 4.2:	COL-19::GFP expression patterns of Lon, Sma and Rol mutants.	159
Figure 4.3:	<i>sqt-1</i> collagen mutants exhibit type IV COL-19::GFP disruption.	163
Figure 4.4:	<i>sqt-3</i> collagen mutants exhibit type IV COL-19::GFP disruption.	165
Figure 4.5:	Dpy collagen mutants exhibit one of two distinct types of COL-19::GFP disruption when crossed with TP12.	168
Figure 4.6:	<i>dpy-5(e61)</i> exemplifies type III COL-19::GFP disruption.	172
Figure 4.7:	Effects of treating different larval stages of TP12 animals with <i>dpy-5</i> RNAi feeding.	176
Figure 4.8:	<i>dpy-7</i> mutants crossed with TP12 exemplify class II COL-19::GFP disruption.	177
Figure 4.9:	<i>dpy-11</i> and <i>dpy-18</i> mutant phenotypes.	180
Figure 4.10:	Bli mutants display either wild type or type III COL-19::GFP patterns	183
Figure 4.11:	Venn diagram illustrating the choice of RNAi targets	189
Table 4.2:	Genes targeted by RNAi feeding and the resulting phenotypes	191
Figure 4.12:	Phenotypes resulting from RNAi of astcins	196
Figure 4.13:	Comparison of COL-19::GFP phenotypes in RNAi treated animals and uncharacterised morphological mutants	201
Figure 4.14:	Cuticle aberrations resulting from targeted RNAi of miscellaneous genes	203
<b>Chapter 5:</b>		
Figure 5.1:	Domain structure and topology of the Duox enzymes	233
Table 5.1:	List of different classes of haem peroxidases`	236
Figure 5.2:	Gross morphological and COL-19::GFP phenotypes of <i>bli-3</i> mutant alleles, ( <i>bli-3(e767)</i> and <i>bli-3(n529)</i> )	238

Figure 5.3: Gross morphological and COL-19::GFP phenotypes resulting from targeted RNAi of <i>bli-3</i>	240
Figure 5.4: Gene and protein structure of <i>hpx-1</i> /HPX-1	243
Figure 5.5: Clustal W alignment of the animal haem peroxidase domains of human, mouse, <i>C. elegans</i> , and sea urchin	246
Figure 5.6: Examples of gross morphological phenotypes observed in animals fed with <i>hpx-1</i> RNAi	250
Figure 5.7: Differing levels of cuticular disruption in animals treated with <i>hpx-1</i> RNAi	252
Figure 5.8: SEM images of <i>hpx-1</i> RNAi-treated worms	255
Figure 5.9: Gross morphological and COL-19::GFP phenotypes resulting from targeted RNAi of <i>hpx-1</i> in the strain <i>bli-1(e767)</i> (COL-19::GFP) compared to untreated animals	258
Figure 5.10: Gross morphological and COL-19::GFP phenotypes resulting from targeted RNAi of <i>hpx-1</i> in the strain <i>bli-1(n529)</i> (COL-19::GFP) compared to untreated animals	260
Figure 5.11: Tissue-specific expression of <i>lacZ</i> under the influence of the <i>lacZ</i> promoter sequence in L3 larvae	262
Figure 5.12: Results from SQ RT-PCR of <i>hpx-1</i>	264
Figure 5.13: Expression of a destabilised GFP marker under the control of the <i>hpx-1</i> promoter	265
Figure 5.14: Recombinant HPX-1	267
Figure 5.15: Investigating the activities of BLI-3 and HPX-1 in morphological mutants and in RNAi experiments	277

## Abbreviations

ADAM	adamalysin
Arg	arginine
Asn	asparagine
Bli	blister
BMP	bone morphogenic protein
Ca <sup>2+</sup>	calcium
CE	<i>Caenorhabditis elegans</i>
COL	collagen
CP	cysteine protease
Cys	cysteine
DIC	differential interference contrast
DNA	deoxyribonucleic acid
Dpy	dumpy
Duox	dual oxidase
ds	double stranded
ECM	extracellular matrix
EDS	Ehlers Danlos syndrome
EGF	epidermal growth factor
EGFP	enhanced GFP
EMS	ethyl methanesulphonate
Eml	embryonic lethal
EPO	eosinophil peroxidase
ER	endoplasmic reticulum
EST	expressed sequence tag
F1	first filial
FACIT	fibril-associated collagens with interrupted triple helices
Gly	glycine
GFP	green fluorescent protein
H <sub>2</sub> O <sub>2</sub>	hydrogen peroxide
His	histidine
HPX	animal haem peroxidase
HSP	heat shock protein
Lon	long
Lvl	larval lethal
Mab	monoclonal antibody
Mlt	moult defect
MMP	matrix metalloproteinase
mRNA	messenger ribonucleic acid
MPO	myeloperoxidase
NAS	astacin
NHR	nuclear hormone receptor
OD	optical density
OI	Osteogenesis Imperfecta
OVO	ovoperoxidase
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PCP	proprotein convertase protease
PDI	protein disulphide isomerase

PHY/P4H	prolyl 4 hydroxylase
PPI	peptidyl prolyl <i>cis-trans</i> isomerase
RISC	RNA interfering specificity complex
RNA	ribonucleic acid
RNAi	RNAi-mediated interference
Rol	Roller
RT-PCR	reverse transcription polymerase chain reaction
SEM	scanning electron microscopy
siRNA	small inhibitory RNA
Sma	small
SNP	single nucleotide polymorphism
Sqt	squat
ss	single stranded
Tet	tetracycline
TGase	transglutaminase
TGF- $\beta$	transforming growth factor- $\beta$
TPO	thyroid peroxidase
TSP	thrombospondin
Trp	tryptophan
ts	temperature sensitive
Tyr	tyrosine
Unc	uncoordinated
UV	ultraviolet
YAC	yeast artificial chromosome
Zn	zinc

## Measurements

bp	base pair
cm	centimetre
kb	kilobase
kDa	kilodalton
Mb	megabase
mg	milligram
$\mu$ g	microgram
ng	nanogram
mm	millimetre
$\mu$ m	micrometer
nm	nanometer
L	litre
ml	millilitre
$\mu$ l	microlitre
M	molar
mM	millimolar
$\mu$ M	micromolar
nM	nanomolar
$\mu$ mol	micromoles
nmol	nanomoles
pmol	picomoles
nm	nanometre

V	volts
rpm	revolutions per minute
g	acceleration due to gravity
Gy	greys

## Summary

The cuticle of *C. elegans* is an extracellular matrix primarily composed of highly processed collagens. It is the site of interaction with the environment, enables motility and controls morphology. Animals with aberrant cuticles have altered body shapes, including dumpiness and a twisted body shape or, exhibit blisters. During the life cycle, *C. elegans* progresses through 4 moults during which a new cuticle is synthesised and the old cuticle is shed. Synthesis of a new multi-layered cuticle occurs via the temporally-controlled secretion of collagens. The processes by which secreted *C. elegans* collagens are incorporated into the cuticle are believed to be similar to those of vertebrate collagens, which are co-translationally modified, folded into triple helices, and subsequently proteolytically cleaved at the C- and N-termini. In the *C. elegans* cuticle, such processed collagen trimers are highly cross-linked by an array of tyrosine-, glutamate, and cysteine-derived inter-chain cross-links in order to produce a cuticle with flexibility and high tensile strength. A *C. elegans* enzyme, BLI-3, originally termed Duox1 on the basis of the presence of a NADPH-oxidase and a peroxidase domain, has been identified as being an important catalyst of tyrosine cross-links in the cuticle. The H<sub>2</sub>O<sub>2</sub> required for the BLI-3 mediated cross-linking activity is supplied by the NADPH-oxidase domain.

A COL-19::GFP marker strain, TP12, has been created and has enabled the visualisation of the tagged collagen in regions and substructures of the cuticle including the annulae, seam cell-derived cuticle and the alae. By creating a number of crossed TP12 strains, this marker has been a useful tool in the analysis of the cuticle structure of many morphological mutants. These studies have enabled the roles of different collagens to be established and also have demonstrated the different mutabilities of discrete cuticle substructures such as the struts, fibrous layers, annular furrows and seam cell-derived cuticle.

TP12 has also functioned as a sensitive marker for cuticle disruptions resulting from RNAi-treatment of a number of genes and has shown these disruptions even in the absence of a gross morphological defect. This has identified at least one gene that has a role in cuticle synthesis that previous RNAi screens classified as being wild type.

The potential of IIPX-1, a putative peroxidase, to mediate cross-linking functions within the cuticle, has been investigated. The RNAi-induced blistered and body morphology defects, as

well as its expression being hypodermal and coincident with the moult, are consistent with a role in cuticle synthesis. The *Mlt* phenotype, which is especially apparent via SEM analysis, additionally implicates HPX-1 in being involved in the moulting process. The relationship between BLI-3 and HPX-1 has also been investigated because BLI-3 possibly represents the sole source of H<sub>2</sub>O<sub>2</sub> for HPX-1 mediated activity. HPX-1 is closely related to peroxidases in other organisms and it has been suggested, based on the roles of these homologues and the observations from these studies, that HPX-1 has synthetic and possibly degradative roles within the *C. elegans* cuticle.

**Chapter 1**  
**General Introduction**

## 1.1 *C. elegans* as a model organism

*Caenorhabditis elegans* is a free-living nematode belonging to the Rhabditidae family of the phylum Nematoda. It is typical of nematodes in having a vermiform body shape and life cycle that involves the progression through a number of larval stages. In conditions where there is an abundance of food, *C. elegans* develops through four larval stages (L1-L4) before maturation into an adult, a stage in which it remains for up to 15 days. However, there is also a specialised and resilient alternative L3 larval form, termed 'dauer', that is induced during conditions of low food and/or high population density. Dauer animals do not age until nourishment is re-established, when the life cycle can be re-entered. Transitions between larval stages are marked by moults, which are again characteristic features of all nematodes. Providing that conditions, such as incubation temperature, are constant, the timings of the moults are invariant, the first L1-L2 moult occurring 12 hours after hatching, and subsequent moults occurring every 6 hours thereafter.

As a metazoan animal, *C. elegans* shares a number of complex biological systems with higher eukaryotes, including a nervous system, defined developmental body plan and programmed cell deaths. However, these higher order processes are mediated by a limited number of cells that retain the organism's relative simplicity. This property of being the simplest known differentiated organism led Sydney Brenner, in 1963, to deem it useful as a model organism for the study of networks of development and their control mechanisms. *C. elegans* is ideal as a research tool both in biological and in practical terms. With regard to its useful biological characteristics, it has a limited and invariant number of cells (959 and 1031 somatic nuclei in hermaphrodites and males respectively), defined cell lineage (Sulston and Horvitz, 1977), and has many genes (~40%) that have high levels of homology to those of both parasitic nematodes and humans (Hashmi *et al.*, 2001; Kamath *et al.*, 2003). Practically, useful properties of the "wild type" Bristol lab strain (N2) include: its short 3-day life cycle; small size (1mm long); large population that is self fertilising (males are not required to maintain a population and a hermaphrodite lays ~300 eggs during a lifetime); transparent appearance; ease of maintenance, by culturing in the laboratory on solid or liquid media using *Escherichia coli* as food; ability to freeze stocks which can be thawed to retrieve viable animals; ease of generating mutants via treatment with chemicals or ionising radiation; ability to obtain transformed lines via microinjection (Mello *et al.*, 1991); and ease of genetic crossing (Brenner, 1974). Because of these properties, *C. elegans* became an ideal candidate for being the first multicellular organism

to have its genome completely sequenced and in 1998 the entire sequence was published (Consortium, 1998). The *C. elegans* genome sequence is approximately 100 million base pairs long and 80% of the sequence encodes for approximately 20,000 protein coding genes. This is in comparison with *Drosophila* that has a similar sized genome but only encodes for 12,000 genes. The post-genomic era has brought about a wealth of resources including the entire genome sequence being available in a library of cosmids and YACs; a dense map of single nucleotide polymorphisms (SNPs) (Swan *et al.*, 2002); an expressed sequence tag (EST) database; a bank of multiple alleles of mutant strains (the multiple alleles increasing the extent of analysis of gene function) (*C. elegans* genetics centre); two consortiums that are concerned with the production of knock-out strains (The *C. elegans* gene knockout consortium and the National Bioresource Project, Japan); and a freely accessible database (WormBase) that serves as a central resource to compile all gene and genomic data. These resources have facilitated the ability to perform functional genomic microarray experiments (Kim *et al.*, 2001), enabled members of families, e.g. collagens, to be identified via sequence homology, and have opened the way for comparative analyses with other organisms. Access to the entire genome sequence also hugely simplifies the design of primers for the cloning of genes.

An invaluable tool in *C. elegans* research has been the development of the RNA-interference (RNAi) technique that allows for specific targeted gene disruption (Fire *et al.*, 1998). High-throughput RNAi screens are possible due to the ease of administration via feeding, soaking or injection (the details of this technique are discussed in a later section (section 1.16)). A library of RNAi bacterial constructs, that covers approximately 86% of the total open reading frames, is now available (Kamath *et al.*, 2003) and has been used for a number of genome-wide screens, using N2 or an RNAi-sensitive strain (the *rrf-3* mutant strain) (Simmer *et al.*, 2002) (Simmer *et al.*, 2003). Screens have been non-specific, recording any of the resulting phenotypes (Fraser *et al.*, 2000) (Kamath *et al.*, 2003) (Simmer *et al.*, 2003) or have specifically sought for, for example, cell division defects (Gönczy *et al.*, 2000), fat storage defects (Ashrafi *et al.*, 2003), increased longevity (Lee *et al.*, 2003) or altered genome stability (Pothof *et al.*, 2003).

In terms of this study, a number of morphologically mutant phenotypes arise when certain genes, particularly those involved in the synthesis or composition of the cuticle, are targeted by RNAi or are mutated via mutagenic screens. This mutability of the genes that control the formation of the *C. elegans* cuticle to dramatically alter body shape into definable and readily identifiable

outcomes is a key advantage for using this organism in the analysis of cuticle composition and synthesis. This model organism is important because the cuticle of parasitic nematodes is a potential target for anti-helminth drugs.

## 1.2 The *C. elegans* hypodermis

The characteristic vermiform shape of the animal first develops during embryogenesis, being established by the rearrangement of a structure comprised of a thin layer of specialised epithelial cells. Within this 'hypodermis', the epithelial cells are arranged in a regular manner in which their apical surfaces face towards the exterior (White, 1988). Such an organisation results in a tissue that is only a single layer of cells thick but has a large area that encloses the animal. A properly formed hypodermis is an absolute requirement for embryogenesis because it is the main mediator of morphogenesis during the transformation from a spheroid- to a vermiform- shape (Sulston *et al.*, 1983). This process starts with contra-lateral migration and interdigitation of hypodermal cells occurs in opposite directions. The two migrating sets of cells meet at a ventral midline and are joined by adherin junctions, thereby establishing a cylindrical body shape (Priess and Hirsh, 1986; Simske and Hardin, 2001). The hypodermal cells then fuse to form a number of multinucleate syncytia, which are increased in size by subsequent embryonic and post-embryonic cell fusions (Sulston and Horvitz, 1977; Sulston *et al.*, 1983). 23 hypodermal cells fuse at this stage to form the largest syncytium, hyp7, which covers ~90% of the surface of the main body and comprises the ventral and dorsal hypodermis (Sulston and Horvitz, 1977; Sulston *et al.*, 1983). By the adult stage, contributions from continual cellular fusions increase the nuclear content of hyp7 from 23 to 139 nuclei and increase its size so that it covers the larger adult body and the openings of the excretory canal and anus (Sulston and Horvitz, 1977; Sulston *et al.*, 1983; White, 1988). There are 6 head- (hyp1- hyp6) and 4 tail- (5 in male) (hyp8-hyp12/13) hypodermal cells (White, 1988). All but three of these are syncytia but have no more than 6 nuclei, in contrast to hyp7 (White, 1988). Hyp6 fuses with the main body hyp7 during the mid L3 stage (Yochem *et al.*, 1998) while hyp8-hyp10 fuse in L4-adult moult in order to form the tail (Nguyen *et al.*, 1999). There are additionally two lateral/longitudinal rows of cells (H0-H2, V1-V6) which are termed the seam cells (Sulston and Horvitz, 1977) and which are distinct from other hypodermal cells. These divide to produce progeny that fuse with hyp7 (Sulston and Horvitz, 1977; Sulston *et al.*, 1983). However, via the controlling action of the GATA transcription factors, *elt-5* and *elt-6* (Koh and Rothman, 2001), they are themselves not fusegenic until the L4-adult moult, when they do fuse (Sulston and Horvitz, 1977) and lead to the formation of trilaminar structures on the cuticle surface called the alae (Sulston *et al.*, 1983)

(Cox *et al.*, 1981a). While most nuclei within syncytia are uniformly distributed and take up defined positions (White, 1988), some cells do move passively during development (Sulston and Horvitz, 1977).

### 1.3 Actin is required for elongation during embryogenesis

After the hypodermally-driven establishment of the basic cylindrical shape of the embryo, actin microfilaments form circumferentially at regular intervals at the apical membrane of hypodermal cells and, using their contractile forces, induce embryonic elongation. The control mechanisms dictating the position and orientation of the actin bundles is, as yet, unknown (Mei *et al.*, 2004). Another consequence of contraction is that the hypodermis is patterned by a regular arrangement of indentations that coincide with the position of the constricted actin bundles (Priess and Hirsh, 1986; Costa *et al.*, 1997).

Actin contraction also requires that the filaments be anchored to the adherin junctions. This anchorage is mediated by  $\alpha$ -catenins and cadherins (encoded by the *hmp* and *hmr* genes in *C. elegans*) and these most likely act to transmit the actin-derived force to the rest of the epidermis (Costa *et al.*, 1998; Mei *et al.*, 2004). Upon treatment of embryos with cytochalasin D, an inhibitor of actin contraction, the elongation step of embryogenesis is prevented (Priess and Hirsh, 1986). Seam cells also provide a significant force of elongation; during development they change shape and themselves become elongated. Failure of this non-muscle myosin-driven morphogenesis to occur prevents embryonic elongation. *mlc-4*, a non muscle II regulatory light chain has been implicated in this process on the basis that *mlc-4* mutants fail to elongate (Shelton *et al.*, 1999). In this way, myosins are implicated in the control of morphology.

The actin microfilaments required for embryonic elongation become disorganised after elongation has terminated (Priess and Hirsh, 1986). However, they re-position circumferentially prior to each larval stage transition (Costa *et al.*, 1997).

### 1.4 The hypodermally secreted cuticle

The hypodermis, during late embryogenesis, is necessary for maintaining the vermiform shape of the animal. However, following elongation and prior to hatching, the hypodermis produces and secretes an extracellular matrix (ECM), which takes over this role (Priess and Hirsh, 1986). ECMs are networks of macromolecules that are produced by all metazoans. In many organisms, the common roles of these lattice matrices are holding cells and tissues together,

tissue regeneration, macromolecule filtration, separating different tissue types, providing surfaces for cellular interactions and migrations and, as described above, maintaining morphology (Sibley *et al.*, 1993; Prockop, 1998; Norman and Moerman, 2000). However, more tissue-, developmental stage- and organismal- specific functions result from varied ECMs composition and structures (Cox *et al.*, 1981a; Kramer, 1994b).

*C. elegans* has two distinct ECMs (Kramer, 1994b): the aforementioned hypodermally-secreted ECM, termed the *cuticle*, and the *basement membrane*, which is secreted by muscles and organs (this is discussed in section 1.15). While the 20nm thick basement membrane lines the outside of the pseudocoelomic cavity that contains the musculature, intestines and gonads, the cuticle overlies the hypodermis and the nervous tissue and also lines the mouth, pharynx and rectum (White, 1988; Kramer, 1994b). Together, these two ECMs define the borders of the two concentric tubes that make up the basic vermiform body structure.

Since it is the most external surface of the animal, the role of the cuticle is multifarious. Its tough but flexible nature maintains body shape, provides a protective barrier against the environment and permits motility (via its attachments to muscle) (Kramer, 1994b; Johnstone, 2000).

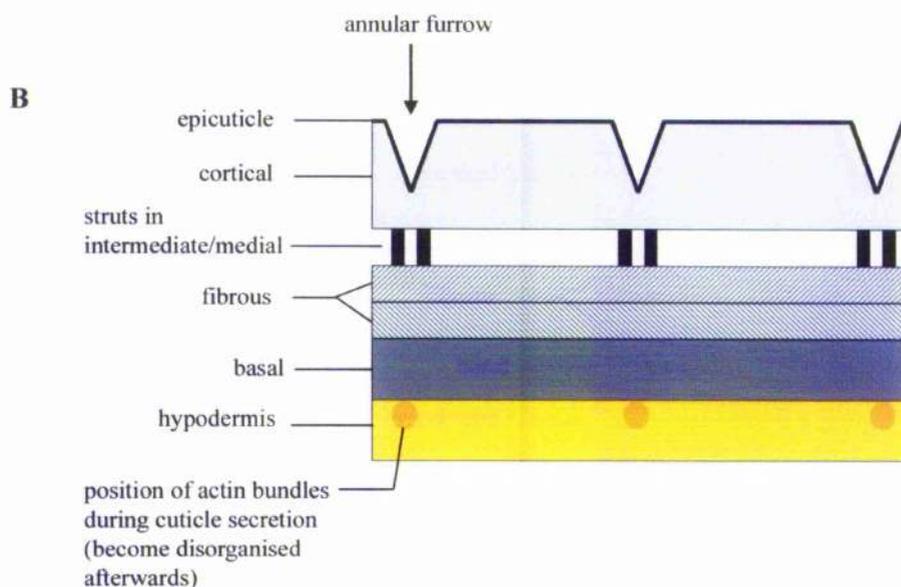
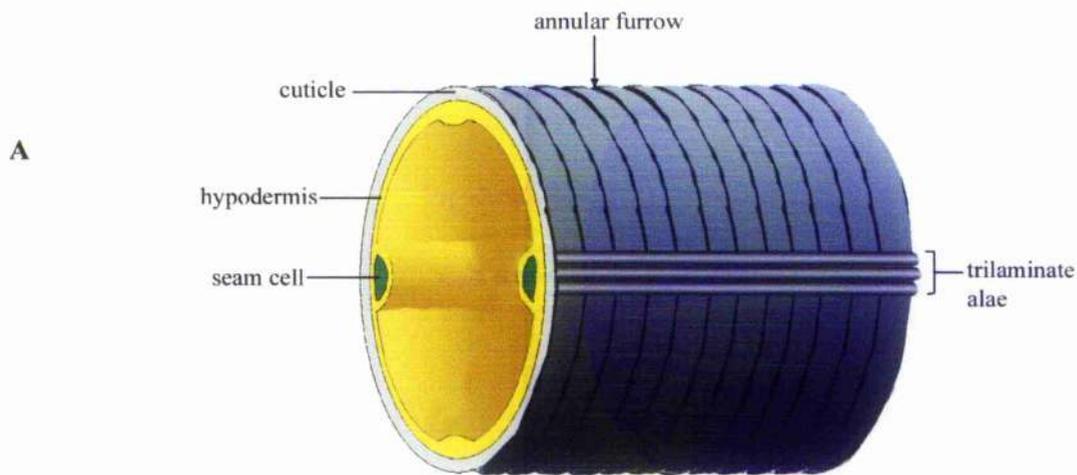
Secretion of cuticle components from the hypodermis in sequential steps facilitates the formation of a complex and versatile tissue that is made up of multiple layers, each of which has distinct components and levels of structural integrity (Johnstone, 2000; McMahon *et al.*, 2003). By altering the combination of layers within a cuticle, even further specialisations can be produced. This is key in the ability of *C. elegans* to adapt to changing environments, alter its interactions with external influences, or to cater for changing developing needs (Cox *et al.*, 1981a). While this is a relatively defunct mechanism for *C. elegans*, which is a free living organism, it is likely to be an evolutionary vestige from parasitic ancestors. Parasites require stage specific cuticles in order to adapt to differing host environments (Cox *et al.*, 1981a). The most important consequence for *C. elegans*, however, of producing new and progressively larger cuticles as the life cycle progresses, is that the growth during each larval stage can be accommodated for (Knight *et al.*, 2002).

At each larval stage, a new distinct and stage-specific cuticle is synthesised beneath the existing one and afterwards, the old defunct one is degraded and shed in a process called moulting. In

order to maintain a continual barrier between the worm and its environment, the new cuticle must be synthesised beneath the old one, a process that requires the prior proteolytic severing of connections between the old cuticle and the hypodermis. This process is termed apolysis or lethargus, the latter term being derived by the inactivity of the animal during this time due to the loss of connections between the musculature and the ECM (Singh and Sulston, 1978). The second event of the moult is characterised by an intense period of synthesis in which components of the new cuticle are deposited (Cox and Hirsh, 1985). The most external layers of the cuticle are laid down first, followed by the subsequent deposition of the underlying layers. The third and final event, termed ecdysis, requires the old cuticle to be shed, a process that is also proteolytically mediated. Moulting is observed in all nematode species and follows the same basic principles.

In terms of cuticle composition, adults have the most layers, numbering six in total. These are, from external to internal: the epicuticle, the cortical-, the medial-, the fibrous- (two of them) and the basal- layers (figure 1.1) (Cox *et al.*, 1981a). The resulting cuticle is approximately 0.5 $\mu$ m thick, which is relatively thick compared to other larval cuticles (Cox *et al.*, 1981a; Cox *et al.*, 1981b). This may be an adult-specific specialisation to enable it to withstand the increasing hydrostatic pressure of the adult body. As mentioned above, each of the cuticle layers has individual properties and is also ultra-structurally distinguishable. The larval stages possess different numbers and combinations of these layers, for example, L1 cuticles are composed of only the basal, cortical and epicuticle layers (Cox *et al.*, 1981a). Both the fibrous and medial layers are highly structured: the fibrous layer is composed of two rows of tightly organised fibres, each of which is positioned at approximately 60° in relation to the longitudinal axis of the animal. This results in the two layers spiralling around the animal in opposite directions. The medial layer, which is unique to the adult cuticle, is composed of fibrous columns, termed struts, that connect the cortical and fibrous layers (Cox *et al.*, 1981a). In contrast to these highly structured layers, the basal and cortical layers are generally amorphous and have less defined regularity and structure. However, stage-specific differences, including thickness, are observed in individual layers. For example, the basal layer in L1 larvae, instead of appearing predominantly amorphous, as it does in adults, exhibits L1-specific striations (Cox *et al.*, 1981a).

The epicuticle and the cortical layers exhibit patterns that are discernible on the surface of the worm as circumferential indentations (figure 1.1); these are termed annular furrows. The regions



**Figure 1.1: Diagrammatic representation of the *C. elegans* cuticle.** Figure A depicts a section of a transverse section of an adult cuticle (grey), overlying the hypodermis (yellow) from which it is secreted. The distinct hypodermal “seam cells” are shown in green. The L1- dauer- and adult-specific trilaminar alae, that run over most of the length of the worm, are synthesised above the seam cells. Circumferential indentations are also observed at regular intervals (these are termed the annular furrows). Figure B pictorially depicts the layers of the adult cuticle, again overlying the hypodermis (yellow). The position of the actin filaments in the hypodermis are indicated and it can be seen that the placement of the annular furrows and the struts coincide. However, these are only transiently positioned at the time of cuticle synthesis.

between these furrows are termed the annulae. The furrows are derived because the cuticle is secreted by the hypodermis which, at the time prior to the synthesis of a new cuticle, is transiently patterned, in an identical manner, by circumferential (repolymerised) actin filaments spaced at  $\sim 1 \mu\text{m}$  intervals (Costa *et al.*, 1997). As described in section 1.3, these are responsible for elongation during embryogenesis and, once this step has been completed become disorganised. However they reassemble prior to each larval moult. Because of the way that the cuticle is synthesised, in order that a cuticle is always present, the requirement for the actin filaments in these post-embryonic stages, is not to maintain body morphology, as it was during the embryo stage (Priess and Hirsh, 1986). Instead, the ability of the actin microtubules to bias the formation of what is essentially a creased cuticle, means that the surface area of the new cuticle is maximised, thereby accommodating for the growth of the animal (Costa *et al.*, 1997). A physical attachment between the cell membrane and the actin is likely to allow for the position of the actin bundles to be translated into the cuticle furrows (McMahon *et al.*, 2003).

The other discernible cuticle features are the alae, which, as stated previously, are seam cell-derived. These trilaminar ridges (figure 1.1) that run almost the entire length of the animal are only present in L1, dauer and adult stage cuticles (Cox *et al.*, 1981a).

### 1.5 Components of the *C. elegans* cuticle

The distinct property of each layer is dependent on its variable composition of collagens and cuticulins. For example, the struts, and the basal, fibrous and cortical layers are composed of cuticle collagens while the cortical layer also has cuticulin components (Cox *et al.*, 1981b). Further, dauer cuticles have high cuticulin content (e.g. the alae specific CUT-1 and CUT-6) (Sebastiano *et al.*, 1991; Muriel *et al.*, 2003). Because they are highly cross-linked, cuticulins are highly resistant to strong reducing agents and detergents (Lassandro *et al.*, 1994) and such a high content may facilitate the resilience of dauer larvae to adverse conditions.

*C. elegans* collagens are the primary cuticle components since they constitute  $>80\%$  of the soluble cuticle proteins (Cox *et al.*, 1981b). There is a large repertoire of collagens: 175 collagens are encoded by the genome of which most (between 154 and 173 (Myllyharju and Kivirikko, 2004)) function as components of the cuticle. The large number of different collagens increase the complexity and the variability of the cuticle (Johnstone *et al.*, 1992). This is in contrast to vertebrates and other nematode species since humans encode for only 27 collagen types (Myllyharju and Kivirikko, 2004) while *Ascaris suum* and *Haemonchus contortus* encode

~ 20 collagens (Kingston, 1991). Even more dramatically different is *Drosophila melanogaster*, whose genome contains only three collagen genes (Myllyharju and Kivirikko, 2004). *C. elegans* cuticle collagens are similar to those found in other nematodes but are smaller and somewhat distinct from vertebrate collagens. While some vertebrate collagens are over 150 kDa in size, *C. elegans* cuticle collagens are between 26-35 kDa (Cox, 1992). *C. elegans* cuticle collagen genes are generally dispersed throughout the genome (Kramer, 1994a). However, some cases of closely related genes have been identified as being only 1-3 kb apart (Kramer, 1994b). Collagen genes represent almost 1% of all *C. elegans* genes.

The processing of cuticle collagens into an extensively folded and cross-linked network is key in harvesting their structural properties and must be under precise control due to the unique folding properties that collagens exhibit. One example of such a constraint that must be overcome is the fact that single polypeptide chains of collagen cannot fold by themselves; they require the presence of two other collagen chains (Engel and Prockop, 1991). The step-wise events involved in collagen processing have been widely studied in vertebrate systems (Engel and Prockop, 1991) and serve as a model for the steps involved in *C. elegans* cuticle synthesis. There is some degree of type-specific processing of *C. elegans* collagens, which introduces yet another means of increasing the diversity of cuticle forms (Cox, 1992). The following sections will therefore describe how vertebrate collagens are folded and processed into functional higher order structures and the features of their sequences that facilitate this processing. The consequences of improper collagen processing are also discussed in the context of the resulting human disorders.

## 1.6 Vertebrate collagens

Vertebrate collagens form the fibrils that compose bones, skin and cartilage. In fact, collagen is the most abundant protein in mammals, constituting 30% of protein mass in humans (Myllyharju and Kivirikko, 2004). Different vertebrate collagens have distinct functions. While some collagens are fibril forming (e.g. type I and type II), other collagens form networks (e.g. IV, VIII and X) (Hulmes, 2002) or regulate fibril properties such as the fibril diameter (e.g. type V) (Kypreos *et al.*, 2000). Another class of collagens, the FACIT collagens (fibril-associated collagens with interrupted triple helices), modify the surface properties and stabilise fibrils (Reichenberger and Olsen, 1996).

### 1.6.1 Vertebrate collagen processing

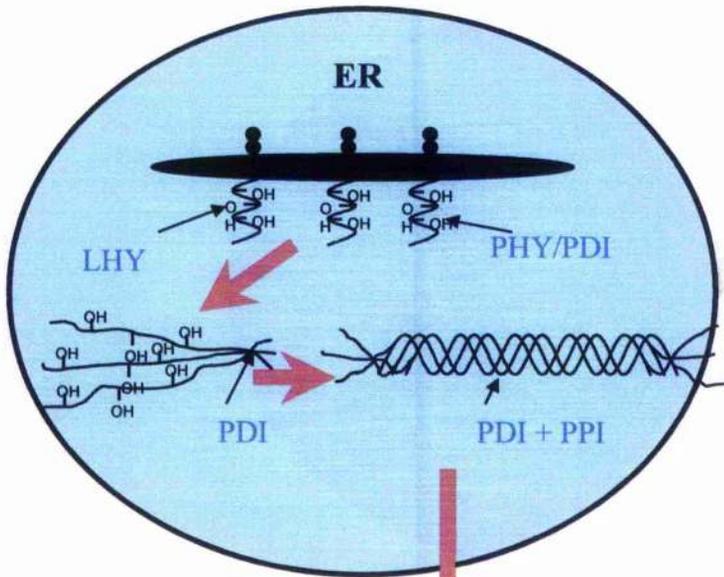
The initial processing steps of vertebrate collagen biosynthesis take place in the endoplasmic reticulum (ER) and, as reviewed by Page and Winter (2003), include: 1) the co-translational hydroxylation of proline and lysine residues (catalysed by prolyl 4-hydroxylase (PHY) and lysyl hydroxylase (LHY)) and the glycosylation of hydroxylated- or asparagine residues, 2) the association of three monomers at the C-terminal domain, 3) the *cis-trans* isomerisation of peptidyl proline bonds (which is rate limiting and catalysed by peptidyl prolyl *cis-trans* isomerase (PPI)) and 4) the propagation of a triple helix in a C- to N-terminal direction in order to form a left-handed helix (Kivirikko and Myllylä, 1982; Engel and Prockop, 1991; Prockop and Kivirikko, 1995). Subsequently, trimeric helical collagens are secreted but continue to be processed via proteolytic cleavage to remove the globular N- and C-terminal domains. The extensively processed trimers are cross-linked by an array of inter-chain bonds. The steps are depicted in figure 1.2.

The sequence of events that occurs prior to the propagation of the helix varies between collagen types. This is exemplified by the fact that the formation of C-terminal disulphide inter-chain bonds is the prerequisite for the initial association of type I collagens (Fessler and Fessler, 1974; McLaughlin and Bulleid, 1998b) but is not required in type III and FACIT collagens (Mazzorana *et al.*, 1993; Bulleid, 1996; Mazzorana *et al.*, 2001). Instead, in the latter collagens, the minimum requirement is a small section of spontaneously-folded triple helix (a minimum of two Gly-Pro-Hyp sequences) which, via lateral interactions, brings the C-termini in close proximity in order that disulphide bridges can subsequently form (McLaughlin and Bulleid, 1998a). The initial helix must be hydroxylated and therefore the presence of an active hydroxylating enzyme is essential for the trimerisation of type III and FACIT collagens (Shaw and Olsen, 1991) (Mazzorana *et al.*, 1993; Bulleid, 1996). Ablating the ability to form disulphide bonds (via mutation) in these collagens has no impact on helix formation (Bulleid, 1996). In contrast, type I collagens, do require disulphide bond formation but can trimerise (but not fold into helices) in the absence of hydroxylation (Fessler and Fessler, 1974; Doege and Fessler, 1986).

The processing steps listed above are facilitated by the protein sequence of collagens. Most vertebrate collagens encode N- and C-terminal proteolytic cleavage sites as well as a number of cysteines, which are involved in stabilising intra- and inter- disulphide bonds, the formation of which are catalysed by protein disulphide isomerase (PDI) (Bulleid and Freedman, 1988). However, not all collagens, for example the FACIT collagens, are proteolytically cleaved: this is

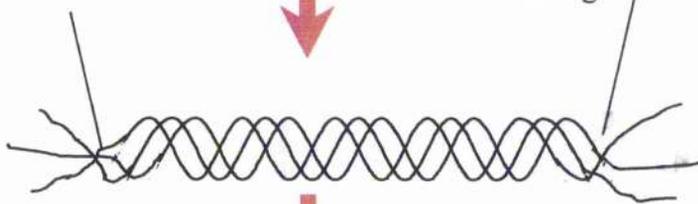
**Figure 1.2. An outline view of some of the key steps involved in collagen biosynthesis.** The initial steps are carried out in the endoplasmic reticulum (ER) and include the hydroxylation of proline and lysine residues, the nucleation of three collagen monomers and the propagation of the triple helix in a C- to N- terminal direction. Trimers are then secreted and their globular ends are proteolytically removed. The soluble trimers are extensively cross-linked by an array of bonds into a flexible and strong network. The enzymes of each of the steps in the ER are depicted: PHY = prolyl 4-hydroxylase, PDI= protein disulphide isomerase, LHY = lysyl hydroxylase, PPI = peptidyl prolyl *cis-trans* isomerase.

Facing page 13

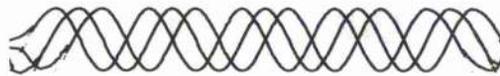


N-terminal protease  
recognition motif

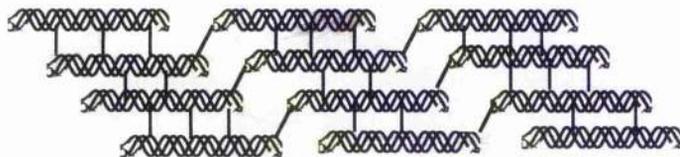
C-terminal protease  
recognition motif



Proteolytic processing



Cross-linked into a network



reflected by the absence of propeptides in these collagens (Mazzorana *et al.*, 1993; Mazzorana *et al.*, 2001). Another collagen-specific folding characteristic is that the triple helices of the transmembrane type XIII and XVII collagens may fold in the opposite direction from all others, i.e. in a N- to C- terminal direction (Myllyharju and Kivirikko, 2001).

The most distinguishing and invariantly encoded feature of collagens is the presence of one or more series of Gly-X-Y repeats which are essential for the formation and stability of the triple helical conformation. Helices, in turn, are vital for the structural integrity of collagens in the cuticle. Within this motif the most invariant constraint is the requirement for a Gly residue at every third position. Gly residues have the smallest side chain compared to other amino acids and thus impose the least sterical hindrance to the helix. Upon formation of a triple helix, the three chains are staggered by one residue in order that the Gly residues from each chain are buried in the central axis of the helix. Staggering of the three molecules also optimises the inter-chain backbone-to-backbone hydrogen bond formation (Melacini *et al.*, 2000; Xu *et al.*, 2002). Once a helix has formed, modifying enzymes, such as PHY (Kivirikko and Myllyharju, 1998), and proteases, such as trypsin (Cox *et al.*, 1981b), are not able to access the individual chains. This reflects the tightly folded nature of the helix.

Replacement of a Gly in the repeat region is extremely disruptive to the tightly folded tertiary structure since in non-Gly-X-Y regions, the helix unfolds (Jenkins and Raines, 2002). Renucleation of the helix is possible and is dependent on the sequence N-terminal to the Gly substitution (Xu *et al.*, 2002). In FACIT collagens there are numerous interruptions to the Gly-X-Y domain, which may be functional, imparting flexibility and “kinks” into the normally rod-like triple helix (Myers *et al.*, 2003) (Kassner *et al.*, 2004). These collagens are positioned on the surface of the fibrils and consequently, the terminal domains of FACIT collagens are able to project into the perifibrillar space (Kapyla *et al.*, 2004). In type XIX vertebrate collagens, molecules assume a “zig-zag” conformation and can even fold back on themselves (Myers *et al.*, 2003). The flexible nature of the collagens and the multiple conformations that can result have been attributed to facilitating interactions between collagen fibrils and other matrix components such as fibronectin and integrin, anchoring microfibrils to basement membranes and being a mechanism of altering the surface properties of individual fibrils in response to the microenvironment (Myers *et al.*, 2003) (Kapyla *et al.*, 2004). Type IX FACIT collagens, by interacting via covalent cross-links, form molecular bridges between cartilage collagen fibrils and other matrix components such as proteoglycans (Chou and Li, 2002).

Proline (Pro) and hydroxyproline (Hyp) frequently occur at the X and Y of the Gly-X-Y moiety respectively. In vertebrate type I collagens, of all Gly-X-Y triplets, 10% are Gly-Pro-Hyp, 22% are Gly-Pro-Y, 22% are Gly-X-Hyp, and 44% are Gly-X-Y (Brodsky and Ramshaw, 1997; Holmgren *et al.*, 1999). A Gly-Pro-Hyp sequence optimises the stabilisation of the triple helix (Sakakibara *et al.*, 1973) and this is illustrated by the increased melting temperature exhibited by collagens upon increased hydroxylation activity (Fessler and Fessler, 1974). In vertebrates, hydroxylation is required in order to raise the melting temperature of the helix above body temperature (Fessler and Fessler, 1974). The stabilising effect of Hyp and Pro is thought to be derived from their imide bonds which are more likely than amide bonds to adopt a *cis* conformation. This consequently reduces the number of conformations available and stabilises the helix (Vitagliano *et al.*, 2001; Xu *et al.*, 2002).

The processing of collagens into the cuticle requires chaperones such as the prolyl 4-hydroxylase complex (P4H), PDIs and Hsp47 (Walmsley *et al.*, 1999; Nagai *et al.*, 2000). Observations from proteolytic studies in vertebrate systems showed that the Hsp47 was specifically limited to triple helix formation (Nagai *et al.*, 2000). The necessary role of the chaperone is demonstrated by the lethality that Hsp47 null mutants cause in mice (Nagai *et al.*, 2000). However, no Hsp47 homologue has been identified in *C. elegans* so P4H may function as a collagen specific chaperone (Norman and Moerman, 2000).

## 1.7 Human collagen diseases

The absolute dependence on collagens to fold in the manner described above to attain proper function is demonstrated by the fact that a number of diseases are caused when these processes go awry. Because of the multifarious structures that are collagenous and the fact that collagen is the most abundant protein in vertebrates, collagen biosynthesis is one of the most vulnerable and mutation-sensitive systems in biology (Kuivaniemi *et al.*, 1991). Consequently, collagens are highly mutable to a number of human hereditary connective tissue disorders. While a handful of diseases have been attributed to mutations in catalysing enzymes for the above steps, the majority of them are the result of lesions in collagens, at sites necessary for their processing via the above steps or at sites required for folding. For most of the above steps, mutations (in collagen processing sites or in the processing enzyme) have been identified that are potent enough to lead to human disorders such as Osteogenesis Imperfecta (OI), Ehlers Danlos Syndrome (EDS) Stickler syndrome and Alport's syndrome. In addition, some forms of

osteoporosis and osteoarthritis are associated with collagen mutations. These diseases will be described in the context of the genetic lesions.

To date, over 1100 collagen gene lesions have been identified within a small subset of collagen genes. By far the most common lesions that lead to disorders are ones in type I collagens that result in Gly substitutions within the helical domain (Cabral *et al.*, 2001). These are particularly potent because the majority of the disease-associated lesions occur in loci encoding type I collagens which associate into heterotrimers consisting of two distinct collagen polypeptide chains in a 2:1 ratio (Prockop and Kivirikko, 1984). Consequently, mutation of a single gene will have wider implications than if all mutant collagens were sequestered to the same homotrimer and will alter chains containing wild type collagen polypeptides. Mutations in type I collagens have been found in OI, a disease associated with fragility of bones, and in severe cases prenatal lethality (Prockop and Kivirikko, 1984; Kuivaniemi *et al.*, 1991; Cabral *et al.*, 2001; Pace *et al.*, 2002) and in EDS, characterised by skin abnormalities and abnormal joints (Prockop and Kivirikko, 1984) (Kuivaniemi *et al.*, 1991). Similar mutations in type II and type IV collagens result in Sticker's syndrome, which is associated with myopia, vitreoretinal degeneration, retinal detachment or a lethal disorder of cartilage, and Alport's syndrome, associated with glomerular nephritis in which there is a progressive loss of kidney function, respectively (Kuivaniemi *et al.*, 1991). As discussed above in section 1.6.1, Gly substitutions are extremely potent due to the disruption to the tightly folded helix that they elicit. In addition, unfolded collagen monomers are the substrate of hydroxylating and glycosylating enzymes, and glycine substitution-induced prevention or delay of trimer formation results in collagens that are over modified and consequently destabilised (Prockop and Kivirikko, 1984).

Abnormal N- and C- terminal processing of collagens, whether as a consequence of lesions in enzymes or to the proteolytic site of collagens, have also been implicated in diseases such as EDS (Prockop and Kivirikko, 1984; Kuivaniemi *et al.*, 1991; Colige *et al.*, 1999) and Schmid Metaphyseal Chondrodysplasia (SMCD) (Wilson *et al.*, 2002). The latter disorder has also been attributed to improper signal peptide cleavage (Wilson *et al.*, 2002). A lethal form of OI has been found to be the result of collagens exhibiting mutations in the C-terminal propeptide, a motif that is required for correct chain recognition and assembly (Pace *et al.*, 2002). EDS is also caused by loss of activity of lysyl hydroxylases, which are required for the proper higher order cross-linking of trimers. Aberrant cross-linking via the loss of lysyl oxidases has also been

predicted to be the cause of Marvin's syndrome which is characterised by long, thin extremities, redundant ligaments and ruptured aorta (Prockop and Kivirikko, 1984).

Many null mutants for collagens are less disruptive than other lesions, consistent with the fact that many of the problems stem from integrating an aberrantly processed collagen into higher order structures (Prockop and Kivirikko, 1995). For example, the region of unfolded and flexible helix in Gly-substituted collagens may form kinks in helices that have implications on the stability of higher order structures, such as fibrils (Prockop and Kivirikko, 1995). However some lesions, such as a number of C-terminal processing lesions, are believed to prevent the incorporation of the mutant collagen into the cuticle. In this case, the resulting OI is a lethal variant (Pace *et al.*, 2002).

### 1.8 *C. elegans* cuticle collagens

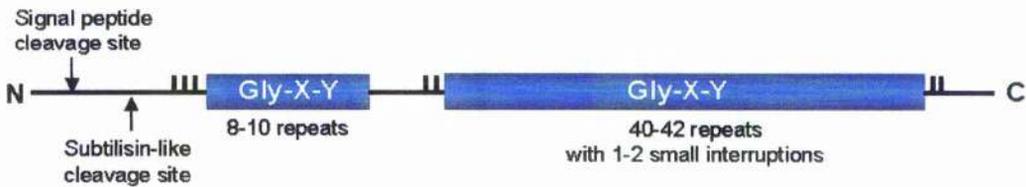
While vertebrate collagens are found in many structures, *C. elegans* collagens are restricted to the cuticle and to the basement membrane. As stated previously, a new cuticle is synthesised at the transition between each moult in order to accommodate for changing requirements and the necessity for growth (Cox *et al.*, 1981a). Consequently, secretion of the components from the hypodermis is only required during these periods. As a result, collagen expression is cyclical and coincident with the larval moults. The limited synthetic period of each moult is further subdivided into distinct peaks, lasting less than two hours, of collagen expression (e.g. 4 hours prior to, 2 hours prior to, and coincident with the moult) with each peak representing the expression of distinct sets of collagens (Johnstone *et al.*, 1992; Gilleard *et al.*, 1997). Collagens whose temporal expressions coincide are believed to interact with each other and in turn form distinct cuticle substructures (e.g. the annular furrows or the annulae) (McMahon *et al.*, 2003). However, not all collagens are expressed in every larval moult; for example, some are expressed predominantly in embryonic and early larval stages only (e.g. *sqt-3*) (Kramer *et al.*, 1985a) while others are adult-specific (e.g. *col-19*) (Liu *et al.*, 1995). Precise regulation of collagen temporal expression patterns facilitates the production of distinct stage-specific cuticles, cuticle substructures and layer compositions (Kingston, 1991). For example, it is hypothesised that the layers destined to become the most external are secreted first, and are therefore expressed first, while underlying layers are sequentially produced. Further cuticle specialisation is achieved by the limiting of expression of some cuticle components to subsets of hypodermal nuclei (Kramer, 1997).

Tight temporal control of hypodermal collagen expression is paramount to ensure that collagens are able to form correct substructures. This is carried out by the heterochronic pathway (Liu *et al.*, 1995), which also has a role in the control of stage specific events in post-embryonic cell division, cell cycle progression and developmental arrest (Ambros, 2000). Heterochronic genes are categorised as being early (hatching to beginning of L3) or late (L3-L4) and include *lin-29*, *lin-4*, *lin-14* and *lin-28* whose mutation lead to precocious or retarded collagen transcription (Ambros, 2000). LIN-29, a putative zinc finger transcription factor, is the downstream component of the pathway that targets 5' regulatory sequences (promoters) of collagens (Liu *et al.*, 1995; Rougvie and Ambros, 1995; Liu *et al.*, 1999). Cuticle collagen promoters are small; for example, ~125 bp of the 5' sequences of the cuticle collagen *dpy-7* is sufficient to drive tissue specific expression (Gilleard *et al.*, 1997) while 235 bp of *col-19* 5' sequence is sufficient for adult-specific expression (Liu *et al.*, 1995). However, this expression is weak and it is likely that further upstream 5' sequences may contain elements that affect levels of transcription (Liu *et al.*, 1995).

### 1.8.1 Structure of *C. elegans* cuticle collagens

Like vertebrate collagens, *C. elegans* cuticle collagens encode for N- and C-terminal domains and Gly-X-Y repeats. However, there are some *C. elegans*-specific features (Cox, 1992). It is therefore predicted that these hypodermally expressed collagens are processed in a similar, but not identical, manner as their vertebrate counterparts. It has not been conclusively established whether *C. elegans* cuticle collagens form heterotrimers or homotrimers despite the fact that genetic interactions between distinct molecules do occur.

The typical structure of *C. elegans* cuticle collagens is depicted in figure 1.3. In each of the ~175 cuticle collagens encoded by the *C. elegans* genome (Johnstone, 2000), there are two blocks of Gly-X-Y repeat sequences, the most N-terminal of which is comprised of 8-10 repeats. The second block of repeats is longer, being made up of 40-42 repeats, and has 1-2 small non-Gly-X-Y interruptions (Kingston, 1991). The presence of two distinguishable Gly-X-Y repeat regions makes the closest vertebrate relatives of these *C. elegans* cuticles the FACIT collagens (Kramer, 1994b), which are the only vertebrate collagens not to have a single, long Gly-X-Y region. However, the vertebrate FACIT collagens are over twice the size of their *C. elegans* counterparts. The interruptions exhibited in *C. elegans* collagens may facilitate more flexible helices and consequently will favour more complex matrix interactions.



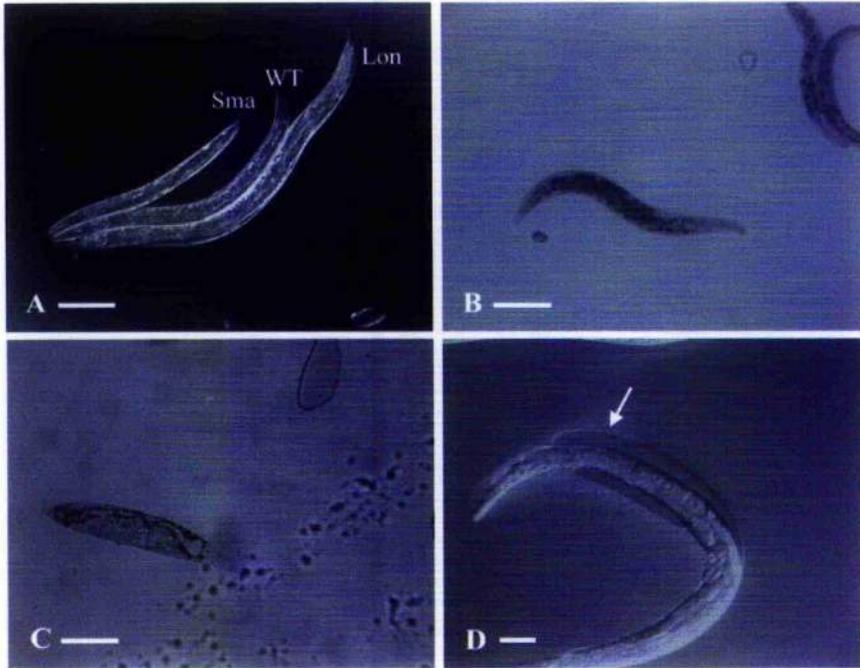
**Figure 1.3: The structure of a typical *C. elegans* cuticle collagen.** There are two stretches of Gly-X-Y repeats which are flanked by N- and C- termini (of varying length according to the collagen). The N-terminus has a number of conserved regions that are required for the proteolytic cleavage of the signal peptide (which functions to direct the collagen to the ER) and the subsequent removal of the collagen propeptide. The second repeat region is longer than the first and additionally has 1-2 short interruptions in the sequence. The vertical lines represent the regions of conserved cysteine residue clusters.

The N-terminal and C-terminal domains of *C. elegans* cuticle collagens are non-Gly-X-Y and are variable in size. The N-terminal contains a series of homology blocks (HB) (HBA-HBD) which, as well as including the signal peptide (HBD) that directs the proteins to the ER for secretion, are required for the proper proteolytic processing of the molecule at a subtilisin-like cleavage moiety (HBA): *C. elegans* collagens are cleaved at this site (Yang and Kramer, 1994). The C-terminal ranges from 14 – 400 residues in length in COL-19 and BLI-1 respectively (Page and Winter, 2003) (Crew and Kramer, personal communication). While it has been firmly established that vertebrate fibrillar collagens are cleaved within this domain, only recently has a proteolytic cleavage site been identified in a subset of collagens that is targeted by a *C. elegans* homologue of the BMP-1 astacin metalloprotease (DPY-31/NAS-35) (Novelli *et al.*, 2004). These N- and C-terminal proteolytic processing events of *C. elegans* cuticle collagens make them distinct from the FACIT collagens, which are not cleaved (Shaw and Olsen, 1991; Mazzorana *et al.*, 1993; Mazzorana *et al.*, 2001).

In a similar manner to vertebrate collagens, cysteine residues within the sequence of *C. elegans* collagens enable the formation of stabilising cross-links. Proof of these cross-links comes from the requirement for thiol reducing agents for solubilisation of the cuticle (Cox *et al.*, 1981a). Cross-links are formed between cysteines of the same single polypeptide chain, between chains of the same triple helix, or are formed in order to cross-link separate trimers. There are typically three clusters of highly conserved cysteine residues, which flank the Gly-X-Y regions. The number and position of cysteines are conserved in subsets of *C. elegans* collagens and have provided a means of classifying the members of the multigene family into one of six groups: group 1, 1A, 2, 3, *dpy-7* and *dpy-2* (Johnstone, 2000). These groupings can be viewed at <http://www.worms.gla.ac.uk/collagen/cecolgenes.htm>. The conservation among different nematode species argues for their functional importance (Kingston, 1991).

### 1.9 *C. elegans* morphological mutants

Early random mutagenic screens (using EMS) (Brenner, 1974) produced a variety of mutants including worms displaying gross morphological defects that deviated from wild type in terms of body shape and size. Body *shape* appears to be controlled by a precise system of collagen processing, collagen interactions and cuticle construction. Mutations within genes encoding for cuticle collagens and collagen processing machinery lead to dumpy (Dpy), blistered (Bli), roller (Rol) and moult defect (Mlt) phenotypes (figure 1.4). Body *size*, on the other hand, is under the influence of a TGF- $\beta$ -related signalling pathway within which lesions lead to long



**Figure 1.4: Typical gross morphological phenotypes resulting from cuticle defects.** A) Comparison of Sma, WT and Lon animals (*sma-2(e502)*, N2 and *lon-3(e2175)*). B) Dpy morphology (*dpy-17(e164)*). C) There can be differing levels of dumpiness, exemplified by this extreme Dpy (*dpy-14(e188)*). D) A Bli mutant. The fluid-filled blister is indicated with an arrow. Scale bars = 100 $\mu$ m. (Notice the different scale in figure D.)

(Lon) and small (Sma) worms (Padgett *et al.*, 1998). These two systems do not appear to be completely independent of each other as complex genetic interactions between components of both systems have been described.

### 1.10 Body shape mutants

45 loci have been identified in the *C. elegans* genome that affect body shape (Johnstone, 2000). Body shape mutants (Bli, Dpy, Sqt and Rol) are indicative of an aberrant cuticle. These mutants also often exhibit decreased mobility due to altered interactions between the ECM and the muscles (Eschenlauer and Page, 2003). In general, Dpy is epistatic to Rol, and Rol is epistatic to Bli (Higgins and Hirsh, 1977). Loci that have been attributed to body shape phenotypes include those of cuticle collagens (e.g. *dpy-5*, *bli-2* and *sqt-3*) and those that encode for enzymes with roles in cuticle biosynthesis (e.g. *dpy-11*, *dpy-18* which are a thioredoxin and a PHY respectively) (Page and Winter, 2003). A small proportion of Dpy genes (*dpy-21*, *dpy-26*, *dpy-27* and *dpy-28*) are X chromosome dosage compensation-related genes (Page, 2001). Collagen genes are described in the next section.

#### 1.10.1 Collagen gene mutants

*dpy-2*, *dpy-3*, *dpy-4*, *dpy-5*, *dpy-7*, *dpy-8*, *dpy-9*, *dpy-10*, *dpy-13*, *dpy-17*, *sqt-1*, *sqt-2*, *sqt-3*, *rol-6*, *bli-1* and *bli-2* are the only known mutation-sensitive cuticle collagen genes and these represent less than 10% of the 173 cuticle collagen genes in the *C. elegans* genome (refer to table 1.1). This implies that while these collagens are integral to the wild type cuticle morphology, there may be a level of redundancy between the others. Cuticle collagens are mutable to an array of phenotypes (Dpy, Bli, Rol, Lon and Sqt) all of which reflect aberrant cuticle structures (Kramer, 1994b; Johnstone, 2000). The collagens that are phenotypically mutable vary according to size and which collagen gene family they are in. This illustrates that there is no correlation between the molecular structure of a collagen and the resulting mutant phenotype (Johnstone, 2000). Some collagens display complex interactions (such as intergenic suppression or enhancement between genes) and this might reflect interactions between their gene products, e.g. between *rol-6* and *sqt-1* and between *dpy-2* and *sqt-1* (Levy *et al.*, 1993);(Kramer, 1994b). Table 1.1 lists the mutable collagens and describes any distinguishing features and notable phenotypes observed in different alleles.

Despite the complex genetic interactions that occur between alleles of multiple collagens, it has not yet been determined whether homotrimers or heterotrimers form during *C. elegans* cuticle

	Locus	Gene Identity	Distinguishing feature	Null/RNAi	Glycine Sub	Other	Ref
Bli Collagen mutants	<i>bli-1</i>	c09g5.6	long N- and C-terminal extensions	Bli			1
	<i>bli-2</i>	f59e12.12	long N- and C-terminal extensions	Bli			1
Sqt Collagen mutants	<i>sqt-1</i>	b0491.2	Interacts with <i>rol-6</i>	WT( <i>sc103</i> )	L.Rol/Lon ( <i>sc101</i> ) weak Lon ( <i>sc107, sc99</i> )	Dpy (N-terminal processing)( <i>sc7, e1350</i> ) Dpy (C-terminal crosslinking) ( <i>sc13, sc113</i> )	2,3,4
	<i>sqt-2</i>	c01b12.1		WT		Lrol (N-terminal processing)( <i>sc3, sc108</i> )	2
	<i>sqt-3</i>	f23h12.4	essential for embryogenesis	lethal ( <i>ts</i> ) ( <i>c2639</i> )	Dpy/Lethal ( <i>ts</i> ) ( <i>e2117, e24, sc63</i> )	L.Rol (signal peptide lesion)( <i>sc42</i> ) L.Rol (C-terminal processing) ( <i>sc2809, sc8, e2888</i> )	7, 17
Rol Collagen mutants	<i>rol-6</i>	t01b7.7	Interacts with <i>sqt-1</i>	WT ( <i>n1268, n1178</i> )		R.Rol (N-terminal processing)( <i>su1006</i> )	2,3,4
	<i>rol-8</i>	zk1290.3		WT			6
Dpy Collagen mutants	<i>dpy-2</i>	t14b4.6	interacts with <i>dpy-10</i> and <i>sqt-1</i>	DpyL.Rol	Dpy/DpyRol( <i>ts</i> ) ( <i>sc38, e8, e489</i> ) Dpy( <i>q292</i> )		5
	<i>dpy-3</i>	egap7.1	<i>S. cerevisiae</i> homologue is an actin interacting factor	Dpy Unc			11
	<i>dpy-4</i>	y41e3.2		Dpy		Dpy/Unc/Lvl ( <i>frameshift</i> )( <i>e1158, e1166</i> )	8
	<i>dpy-5</i>	f27c1.8		Dpy		Dpy (Gly-stop)( <i>e67</i> )	9
	<i>dpy-7</i>	f46c8.6		Dpy ( <i>qm63</i> )	Dpy (effective null) ( <i>e88</i> )		10,11
	<i>dpy-8</i>	c31h2.2					11
	<i>dpy-9</i>	t24d12.2		Dpy ( <i>e1164</i> )	Dpy ( <i>e12</i> )		8
	<i>dpy-10</i>	t14b4.7	interacts with <i>dpy-2</i> and <i>sqt-1</i>	DpyL.Rol ( <i>cg36</i> )	DpyRol ( <i>sc48, cg37</i> )	Dpy ( <i>splice</i> )( <i>e126</i> ) Dpy/DRol( <i>ts</i> ) (N-proteolytic processing)( <i>cn64</i> )	5
	<i>dpy-13</i>	f30b5.1		Dpy		Dpy (deletions/nulls?) ( <i>e184, e458</i> )	12
	<i>dpy-14</i>	h27m09.4	highest expression in embryos (reappears in L4s)	emb lethal Sma	weak Dpy/Dpy/Lethal ( <i>ts</i> )		14
	<i>dpy-17</i>		similar to <i>sqts</i>	Dpy ( <i>e164, c586</i> )	Dpy( <i>ts</i> )( <i>e1345</i> )		15
Lon Collagen mutants	<i>lon-3</i>	zk836.1	regulator of body length via TGF- $\beta$ -like pathway	Lon ( <i>sp6, sp23, cl417</i> )	Lon(semidominant) ( <i>cl418</i> )	Lon (deletion + transposon insertion) ( <i>e2175</i> )	16

Key: WT= wild type, Dpy = dumpy, Rol = roller, Lon = long, Unc = uncoordinated, Sqt = squat, Bli = blistered *ts* = temperature sensitive, emb = embryonic, Sma = small, References: 1= Crew and Kramer (personal communication); 2= Kramer and Johnson (1993); 3= Yang and Kramer (1994); 4 =Yang and Kramer (1990); 5= Levy *et al.* (1993); 6= Kusch and Edgar (1988); 7= Novelli *et al.* (2004); 8=Simmer *et al.* (2003); 9=Thacker and Rose (personal communication); 10=Johnstone *et al.* (1992); 11= McMahon, *et al.* (2003); 12= von Mende *et al.* (1988); 14 = Prasad and Rose (personal communication); 15= Smardon and Maine (personal communication); 16= Suzuki *et al.* (2002); 17 = van der Keyl *et al.* (1994).

Table 1.1: Collagen gene mutants and their phenotypes according to their genetic lesions.

The gene identities, any distinguishing features and allele names (in brackets and italics) are given. If no information is available, cells have been left blank. Refer to key for references and abbreviations.

biosynthesis. The genetic interactions may indicate that the collagens involved associate together in a trimer. However, it is also possible that they represent interactions at a higher level, for example, associations between triple helices or larger complexes. To date, only SQT-3 has been shown to be able to homotrimerise (A. Page, unpublished data).

Multiple alleles have been described for many of these genes, and for the majority of them, all the known morphological mutant alleles exhibit the same gross morphological phenotype. On the other hand, some collagen genes, such as *sqt-1*, have alleles with differing phenotypes (Kramer and Johnson, 1993). In both cases, the location of the genetic lesion dictates the *severity* of the resulting phenotype and in the latter cases, the lesion regulates the *nature* and severity of the overall phenotype. Some alleles elicit a temperature sensitivity, the timing of which coincides with the time of the moults (Higgins and Hirsh, 1977). This is further evidence that these alleles exhibit aberrant cuticle structures.

A study carried out by McMahon *et al.* (2003) analysed the cuticles of different collagen mutants using scanning electron microscopy (SEM) and immunolocalisation using monoclonal antibodies raised against the unique 40 residue C-terminus of the cuticle collagen, DPY-7 (McMahon *et al.*, 2003). These investigations revealed the cuticles of Dpy mutants can exhibit a wild type pattern of annular furrows or can be bald (i.e. devoid of discernible annulac). Mutant loci were grouped into two subsets accordingly (McMahon *et al.*, 2003). The collagens within each subset are expressed simultaneously in relation to each larval moult, interact with each other, and localise to distinct substructures. The collagens of one subset (DPY-7, DPY-2, DPY-3, DPY-8 and DPY-10) are integral components of the annular furrows. They must form complex interactions with each other since removal of any one of them prevents the incorporation of the rest into the cuticle with the consequential loss of the annular furrows. The collagens of the other subset (DPY-5/DPY-13) require the annular furrows in order to be properly incorporated into the cuticle in the annular cortex; in the absence of furrows, their localisation is disorganised. Consequently, these collagens are dependent on the proper localisation of the set of DPY-7 collagens. The relationship is not reciprocal however, since the DPY-7 collagens localise independently of DPY-5/DPY-13 (McMahon *et al.*, 2003).

#### 1.10.1.1 Blistered collagen mutants

The Bli phenotype is only manifested in adults and is characterised by: fluid filled swellings in the cuticle; a slightly reduced body length (700-800µm long as opposed to 1mm long wild type

animals); an expanded body circumference (500µm excluding blisters) wide versus 300µm wide wild type); slow movement; aberrant alae (multiple and discontinuous); shorter life span (3-4 days) (de Melo *et al.*, 2003) and possible slower development (*bli-4* animals take 30% longer to reach the adult moult) (Peters *et al.*, 1991). Of the six *bli* genes, two (*bli-1* and *bli-2*) are cuticle collagens and, consistent with the adult specific appearance of blisters, these are expressed in the L4-adult moult only (Crew and Kramer, personal communication). This is characteristic of all Bli phenotypes. Young adults have a normal cuticle with blisters appearing about 2 hours after this moult (Peters *et al.*, 1991). Although the severity of the Bli phenotype is allele-dependent, the general pattern is that the blistering becomes more severe as the worm ages. The less severe Bli phenotypes are usually characterised by relatively small blisters located on either side of the head of the worm. As most Bli worms age (or in severe alleles), the number and size of blisters increase and a corresponding decrease in movement and mating ability is observed (de Melo *et al.*, 2003);Crew and Kramer, personal communication). However, a phenomenon in which blisters can “heal” and the worm reverts to wild type has been noted in about 10% of blistered animals, although the percentage varies according to the allele.

Blisters result from the separation of the basal and cortical layers of the cuticle (Peters *et al.*, 1991). This was confirmed by the finding that struts are aberrant or absent in *bli-1(n361)* mutants after ultra-structural analyses, immunocytochemical, freeze fracture and transmission electron microscopy (TEM) (de Melo *et al.*, 2003). These also showed actual degradation of the fibrous and basal layers resulting in the degraded material accumulating inside the intermediate layer, which is normally empty space in wild type animals (de Melo *et al.*, 2003);Crew and Kramer, personal communication). An altered mechanism of osmoregulation is also thought to be involved (de Melo *et al.*, 2003).

The fact that *bli-1* mutants exhibit aberrant struts and that BLI-1::GFP localises to struts in wild type worms (Crew and Kramer, personal communication) illustrates the necessity of BLI-1 for the assembly of these structures. BLI-2 is also thought to compose them (Crew and Kramer, personal communication). Struts are only observed in the adult cuticle and crosses between *bli-1/bli-2* and different heterochronic mutants, which either fail to make the L4-adult switch or produce the adult cuticle prematurely, showed the presence of an adult cuticle is necessary for the manifestation of the Bli phenotype (Peters *et al.*, 1991). *bli-1* was identified using transformation rescue and encodes for an unusual cuticle collagen with greatly extended N- and

C- termini, each roughly 400 residues (Crew and Kramer, personal communication). Through interactions with other cuticle components, BLI-1 may use these long extensions to connect the layers flanking the medial layers. Some alleles displaying the Bli phenotype are results of lesions in the N- and C- terminal domains, illustrating their functional importance (Crew, 1999)). *bli-1* and *bli-2* have been shown to genetically interact (Crew and Kramer, personal communication).

### 1.10.1.2 Roller collagen mutants

Roller mutants are characterised by an abnormal crawling behaviour in which worms rotate with a left- or right- handed pitch (LRol and RRol) around their longitudinal axis (Cox *et al.*, 1980);(Higgins and Hirsh, 1977). The cuticle structure of these worms is twisted and exhibits aberrant alae. In wild type worms, alae extend straight down the body, parallel to the side of the worm, but in roller mutants they wrap around the body of the worm in a manner that reflects the pitch of the crawling (Cox *et al.*, 1980). The twisting is observed in the body musculature, the internal organs, the ectodermal cells and in the ventral nerve cord (Higgins and Hirsh, 1977). Even when the contents of the worm are removed through sonication, the cuticle remains twisted and illustrates that the phenotype is brought about by an aberrant cuticle structure rather than being a result of altered internal body mutations. In fact, sometimes the helical twisting is even increased, suggesting that the turgor pressure of the animal acts to partially unwind the twist (Cox *et al.*, 1980). Worms that exhibit LRol phenotypes normally exhibit a pitch of helix of a 1/2 turn along the length of the worm while those exhibiting RRol have a more severe helical twisting whereby a full turn is completed along the body length (Cox *et al.*, 1980).

*rol-6* is a mutation-sensitive cuticle collagen which exhibits a LRol or RRol according to the genetic lesion while null alleles exhibit a wild type body morphology (Kramer and Johnson, 1993). *rol-6* and *sqt-1*, additional cuticle collagens, have sequence similarity, and equivalent mutations in both genes result in the same phenotype (Kramer and Johnson, 1993). Therefore, the genetic lesions responsible for the LRol and RRol phenotypes will be discussed in the context of *sqt-1* in the next section.

### 1.10.1.3 Squat collagen mutants

Three *sqt* genes exist to date (*sqt-1*, *sqt-2* and *sqt-3*) and each encode cuticle collagens (Kramer *et al.*, 1988). *sqt-1* is by far the most characterised. Sqt mutants were historically defined by the fact that they are Rol as heterozygotes and Dpy/DpyRol as homozygotes. However, further

analysis has revealed that these are a complex class of genes and accordingly exhibit a range of allele-specific phenotypes that include Dpy, dominant and recessive LRol and RRol, wild type and one Lon allele (Kusch and Edgar, 1986). As stated, heterozygotes normally exhibit Rol phenotypes. Certain alleles even exhibit different phenotypes according to different larval stages (Cox *et al.*, 1980) e.g. *sqt-1(sc1)* are Rol in L3 and Dpy in L4/Adult (Kramer and Johnson, 1993; Park and Kramer, 1994). Some alleles, in particular *sqt-3* alleles, exhibit temperature sensitivity (van der Keyl *et al.*, 1994).

These varied phenotypes and the complex genetic interactions suggest that the overall morphology of the cuticle is dependent on the *sqt* gene products. In addition, it is interesting to note that the instance of a gene having a high frequency of dominant alleles (found for *sqt-1* and *sqt-3*) (Cox *et al.*, 1980) is a common attribute for *C. elegans* genes, such as actin, that encode for proteins that form multimeric structures (Park and Kramer, 1994). Certain mutations in these collagens can also result in abnormal alae which are beaded (i.e. they have minor interruptions) or display loops (Kramer, 1994b).

*sqt-1* was identified through Tc1 transposon insertion and the subsequent chromosome walking and deficiency mapping and it was the first gene to show the ability of mutant collagens to be sufficient to induce an altered phenotype (Kramer *et al.*, 1988). Its structure is that of a typical *C. elegans* cuticle collagen, and contains a 126 residue N-terminus, 150 Gly-X-Y repeats and a 25 residue C-terminus.

Worms that have a RRol phenotype *sqt-1(e1350)* and *sqt-1(sc1)* result when the conserved N-terminal arginine residues within the proposed ArgXXArg subtilisin-like cleavage site are replaced, thus preventing proteolytic processing at this site (Kramer and Johnson, 1993; Yang and Kramer, 1994; Yang and Kramer, 1999). These mutants show extremely aberrant cuticles that exhibit severe branching alae and annulae. The LRol phenotype of *sqt-1(sc13)* is caused by mutations that prevent cross-linking, such as substitutions of either, or both, of the two C-terminal cysteines (Kramer and Johnson, 1993; Yang and Kramer, 1994). Ablation of these residues allows trimer formation but prevents the formation of tyrosine cross-links since a prerequisite of tyrosine-associated cross-links between collagen monomers is the formation of disulphide bonds at a site adjacent to the tyrosine. Western blots probing for *sqt-1(sc13)* cuticle extracts show an accumulation of single collagen trimers in these mutants compared to wild type

in which SQT-1 trimer moieties are in dimer, tetramer and higher oligomer forms (Yang and Kramer, 1999).

Glycine substitutions in SQT-1 collagens lead to a mild phenotype in which body length is slightly reduced without an accompanying increase in body width (Kramer and Johnson, 1993). This is interesting since in many other collagens these lesions are associated with the most severe morphological defects (Kramer, 1994b; Page and Winter, 2003). Null alleles of all the SQT-1 collagens are the least potent lesions, resulting in almost wild type morphology, being only slightly Dpy with small abnormalities in the tail (Kramer and Johnson, 1993; Yang and Kramer, 1999). RNAi targeting of these genes also produces animals with only subtle morphological changes. The observations for the different allele lesions of *sqt-1* imply that the phenotype observed in non-null mutations is due, not to the removal of an essential collagen from the cuticle, but to a consequence of the insertion of an aberrant and disrupting "structural poison" (Kramer *et al.*, 1988). Based on its wild type null phenotype, the presence of SQT-1 in the cuticle is not essential.

Ultra-structural analysis of the LRol cross-linking mutant, *sqt-1(sc13)*, (Kramer and Johnson, 1993) reveals cuticle aberrations in the fibrous and intermediate layers of these mutants (Peixoto *et al.*, 2000). Compared to the wild type fibrous layers which has two layers of fibres arranged at 60° to one another to form a herring bone pattern (Cox *et al.*, 1981a), *sqt-1(sc13)* mutants exhibit only one fibrous layer that runs laterally. These mutants also display aberrant struts and an intermediate layer that is filled with electron-dense material, as opposed to being translucent in wild type animals (Peixoto *et al.*, 2000). This material completely fills the layer, extending between the cortical and basal regions, and is organised in a honeycomb fashion (Peixoto *et al.*, 2000). *sqt-1(sc103)* null worms have relatively wild type gross morphology but exhibit aberrations in the cuticle at the ultra-structural level since only one fibrous layer exists and struts are loosely organised (Peixoto *et al.*, 2000). However, unlike the Rol *sqt-1(sc13)* mutants, and more similar to wild type worms, the intermediate layer is empty. These observations suggest that a fibrous layer composed of just one parallel layer is not sufficient to induce a Rol morphology. It appears that these aberrations in the fibrous layer do not affect the overlying cuticle. It is likely that the distinguishing property between the LRol and RRol phenotypes is which of the two fibrous layers is lost since each one is in contact with different layers and non-fibre portions of the cuticle (Peixoto *et al.*, 2000). However, it may not be so simple

because *rol-6(su1006)* animals are observed to contain both fibrous layers yet still manifest the RRol phenotype (Peixoto *et al.*, 1998).

*sqt-1* and *rol-6* exhibit complex genetic interactions and this might illustrate the interaction of their gene products (Park and Kramer, 1994). This is further strengthened by the fact that, as mentioned above, equivalent mutants in *rol-6* and *sqt-1* have similar but not identical phenotypes. *sqt-1* phenotypes are generally stronger than the *rol-6* phenotypes (Park and Kramer, 1994).

SQT-3 is unique since it is the only collagen that is required for viability (Novelli *et al.*, 2004). The essential role of this collagen is restricted to embryogenesis. This is illustrated by a number of observations: firstly, it is expressed most abundantly in embryonic and L1 stages (Kramer *et al.*, 1985b) and secondly, many of its alleles are embryonic lethal when incubated at  $>15^{\circ}\text{C}$  (van der Keyl *et al.*, 1994). Thirdly, incubation at sub-optimal temperatures is only required for embryonic stages; once past L1, animals can be incubated at  $25^{\circ}\text{C}$  and maintain a wild type body morphology (van der Keyl *et al.*, 1994). Finally, some alleles such as *sqt-3(sc63)*, a glycine substitution, are Dpy at L1 but revert towards wild type as they progress towards the adult stage. On the other hand, some alleles (e.g. *sqt-3(e24)*) remain Dpy throughout all larval stages and become DpyRol in adults (van der Keyl *et al.*, 1994). This reflects that there is low-level expression of this collagen during post L1-larval and adult stages (van der Keyl *et al.*, 1994).

SQT-3 glycine substitutions exhibit larval lethal, Dpy or DpyRol phenotypes (e.g. in *sqt-3(e2117)*, *sqt-3(sc64)* L1s and *sqt-3(e24)* adults respectively) (van der Keyl *et al.*, 1994). The severity of phenotype is determined by the position of the substitution; the closer the substitution is to the N-terminal, the more severe the phenotype. This is because being closer to the N-terminus reduces the ability for the helix to renucleate (Xu *et al.*, 2002). Inability to renucleate reduces the thermal stability and/or affects the N-proteinase cleavage site (van der Keyl *et al.*, 1994). More recently, mutant alleles have been isolated that contain lesions in a C-terminal region that is the site of proteolytic cleavage by a BMP-1 homologue (DPY-31/NAS-35) (Novelli *et al.*, 2004). C-terminal processing site mutants exhibit LRol or weak Dpy phenotypes below  $15^{\circ}\text{C}$  (e.g. *sqt-3(e2809)* and *sqt-3(e2911)* respectively) but are lethal above the critical temperature. A temperature-sensitive null mutant has also been identified: *sqt-3(e2924)* (J. Novelli, personal communication).

#### 1.10.1.4 Dumpy collagen mutants

The Dpy phenotype, characterised by a short wide body, is a consequence of an aberrant cuticle that allows greater radial extensibility when subjected to high internal hydrostatic pressure (Kramer *et al.*, 1988). 10 of the 27 Dpy-causing genes *dpy-2* (Levy *et al.*, 1993), *dpy-3* (McMahon *et al.*, 2003), *dpy-4* (Simmer *et al.*, 2003), *dpy-5* (Thacker and Rose, personal communication), *dpy-7* (Johnstone *et al.*, 1992), *dpy-8* (McMahon *et al.*, 2003), *dpy-9* (Simmer *et al.*, 2003), *dpy-10* (Levy *et al.*, 1993), *dpy-13* (von Mende *et al.*, 1988) and *dpy-17* (A. Smardon, personal communication) are cuticle collagens. The genetic loci responsible for these collagen mutations were discovered through a variety of methods including low stringency hybridisation using a collagen, Tc insertion analysis, cosmid mapping, and rescues.

Unlike the Bli phenotype, which is only observed in adults (Peters *et al.*, 1991), the Dpy phenotype has been observed at all post-embryonic stages (e.g. all stages of *dpy-13(e184)* mutants have morphological defects). However, in some cases, there is gene- and allele-specificity in the extent of phenotype severity at different stages (e.g. *dpy-2(ed)* animals are non-Dpy during larval stages).

All the loci that are mutable to the Dpy morphology have null/RNAi phenotypes that are Dpy or DpyRol, indicating that, in contrast to *sqt-1*, whose exclusion from the cuticle has little effect on morphology, the presence of this subset of collagens in the cuticle is necessary for proper cuticle formation. DpyRol is assumed to be a more severe manifestation of Dpy and is associated with a significant reduction in the collagens within the cuticle, hence the many null mutants exhibiting this phenotype (Levy *et al.*, 1993). Glycine substitution mutants, identified to date, for the 'Dpy' loci are also Dpy or DpyRol but there is allele specificity in terms of the *severity* of the phenotype. Some loci (e.g. *dpy-7* and *dpy-5*) have null- and glycine substitution- alleles that exhibit equivalent degrees of dumpiness while other loci, such as *dpy-10* (Levy *et al.*, 1993), have null/RNAi phenotypes which are less severe than certain lesions. In the former case, equivalent phenotypes arise when the aberrant glycine-substituted collagens are excluded from the cuticle, making them effective nulls (McMahon *et al.*, 2003). In the latter case, the aberrant collagen is acting as a "structural poison" (Levy *et al.*, 1993) in a similar manner to the aberrantly processed SQT-1 collagens.

Some alleles of the Dpy collagen mutants are temperature-sensitive that exhibit either wild type→mutant or Dpy→DpyRol transitions when exposed to conditions over a critical temperature (e.g. *dpy-8(sc44)* and *dpy-2(sc28)* respectively). This temperature-sensitivity is presumably caused by temperature-induced destabilisation of the mutant collagens (Levy *et al.*, 1993).

In contrast to Bli-inducing collagens which are associated with the underlying struts, Dpy loci gene products localise to the annular cortex and furrows (McMahon *et al.*, 2003). This illustrates that the localisation of a mutant collagen within the cuticle dictates the resulting morphological phenotype.

### 1.11 Mutants in collagen biosynthesis

A number of morphological mutants have been identified as having lesions in genes encoding cuticle processing enzymes. These include the *bli-3*, *bli-4*, *bli-5*, *dpy-11*, *dpy-14*, *dpy-18* and *dpy-31* loci. These will be discussed in the context of their functions below. It is interesting to note that all the alleles for each locus result in the same phenotype. Therefore, whether Bli or Dpy phenotypes result as a consequence of lesions to the enzymes, may be indicative of the enzymes having different substrates. To recap, the processing of collagens requires hydroxylation of proline and lysine residues, *cis-trans* isomerisation, proteolytic processing of the pro-collagen N- and C- termini, and numerous cross-linking steps. Morphological mutants have been found for all the enzymes necessary for these steps apart from PPIs, which are thought to be important but not to lead to defects due to high levels of redundancy (Page and Winter, 2003).

#### 1.11.1 PHY-1 (DPY-18) and its PDI subunit

Hydroxylation of proline residues, catalysed by prolyl 4-hydroxylase (PHY), occurs co-translationally and results in the necessary hydroxylation of approximately half the Y position prolyl residues in the Gly-X-Y repeat region in both nematode and vertebrate collagens (Cox *et al.*, 1981b; Kivirikko *et al.*, 1992; Kivirikko and Myllyharju, 1998). This process is key for the thermal stability of the final triple helix structure (Sakakibara *et al.*, 1973). In both systems, PHY activity requires that it be in a complex with a protein disulphide isomerase (PDI) subunit (the  $\beta$  subunit) whose enzyme activity is not required for the hydroxylation step in collagen synthesis but which is vital in the complex for maintaining the PHY subunit (the  $\alpha$  subunit) in a non-aggregated, catalytically active conformation as well as sequestering it to the ER via a

KDEL (vertebrate) or H/K-E/T-C/E-L (nematode) motif region (Kivirikko and Pihlajaniemi, 1998; Eschenlauer and Page, 2003). The PDI subunit, which is effectively acting as a chaperone, cannot be substituted with another chaperone molecule, such as BiP, and therefore the ability to maintain PHY activity is PDI-specific (Kivirikko 1998). Collagen molecules that are monomeric or unfolded are the only substrates for PHY activity because once folded into a helix, the enzyme can no longer access the proline residues. Unfolded or monomeric collagens are actually necessary for the PHY complex since the subunits of the enzyme dissociate in their absence (Kivirikko, 1998). This necessity for monomeric collagens is the reason that hydroxylation of prolines occurs as one of the first steps in collagen processing.

The *C. elegans* genome encodes two PHY genes, *phy-1* and *phy-2*, that have cuticular roles. *phy-3* (Riihimaa *et al.*, 2002) and *phy-4* do not (Winter and Page, 2000). *phy-1* was identified as the locus responsible for *dpy-18*: microinjection of the gene is able to rescue the Dpy phenotype of these mutants while microinjection of a mutated form of the gene does not enhance the null phenotype mutants (Friedman *et al.*, 2000; Hill *et al.*, 2000; Winter and Page, 2000). In addition, RNAi of the *phy-1* gene phenocopies *dpy-18* alleles (Friedman *et al.*, 2000; Hill *et al.*, 2000; Winter and Page, 2000). Three *C. elegans* PDIs (PDI-1, PDI-2 and PDI-3) exist, however, both PDI-1 and PDI-3 have roles distinct from the PHY-PDI complex (refer to sections 1.11.2 and 1.11.5.4).

Active PHY/PDI complexes exist as either a DPY-18/PHY-2/(PDI-2)<sub>2</sub> mixed tetramer or as DPY-18/PDI-2 (or PHY-2/PDI-2) dimers. The mixed tetramer is by far the most abundant form (Winter and Page, 2000). Due to the presence of the PDI-2 subunits in all of these complexes, it is not surprising that RNAi of *pdi-2* produces a more severe effect than when either of the PHY subunits is eliminated alone. Interference of PDI-2 results in embryonic lethality while *dpy-18* RNAi (as described above) produces Dpy animals and RNAi of *phy-2* does not affect the gross morphology. However, combined interference of both PHY subunits (through RNAi of *phy-2* in *dpy-18* null background) is embryonic lethal. The viable RNAi phenotypes of the single *phy* interferences is explained by the fact that despite the elimination of the mixed tetramer, the dimer forms of the enzyme still exist and, are actually upregulated in the absence of one of the  $\alpha$  subunits (Myllyharju *et al.*, 2002). All RNAi-treated worms exhibit reduced hydroxyproline content (Winter and Page, 2000).

The human forms of PHY complexes differ from their *C. elegans* counterparts in that neither dimers (DPY-18/PDI-2 or PHY-2/PDI-2) nor mixed tetramers (i.e. containing two different PHY subunits (PHY-PDI)<sub>2</sub>) exist; only tetramers containing two molecules of the same PHY subunit are observed (Kivirikko and Myllyharju, 1998). The formation of either a dimer or tetramer depends on the PHY subunit since an active dimer is formed when *C. elegans phy-1* is expressed with human PDI while an active tetramer is formed between two *C. elegans* PDI-2 subunits and two human PHY subunits (Myllyharju *et al.*, 2002).

### 1.11.2 A *cis-trans* isomerase (PPI) and additional roles of PDI-2 and PHYs

PDI-1 lies on an operon with *cyp-9*, a peptidyl prolyl *cis-trans* isomerase (PPI) (Page, 1997) which catalyses the *cis-trans* isomerisation of Hyp and Pro residues, a step that is necessary for, and another rate limiting step of, collagen triple helix propagation (Bächinger, 1987). Expression of these enzymes is hypodermal and peaks coincident with the larval moults (Page, 1997). 26 members of the PPIs family are encoded in the *C. elegans* genome (Page *et al.*, 1996; Page and Winter, 2003). This includes the FK506-binding proteins and the cyclophilins of which *cyp-9* is one (WormBase). This level of redundancy may explain the fact that targeted RNAi of *cyp-9* does not produce observable morphological or cuticular defects (Page and Winter, 2003).

In addition to its role in maintaining PHY in an active conformation, PDI-2 acts as a chaperone, both with and without PHY, that retains non-helical and non-hydroxylated collagens in the ER (Walmsley *et al.*, 1999) and also catalyses the formation of inter- and intrachain- disulphide bonds through its thioredoxin-like activity (Eschenlauer and Page, 2003). PDI-1 and PDI-3 (the latter of which is discussed in section 1.11.5.4) share these additional functions. The thioredoxin-like activity of these enzymes is involved in the breaking, rearrangement and formation of disulphide bonds (Blasco 2003); this is a rate-limiting step in collagen biosynthesis (Page 1997).

Prolyl hydroxylating enzymes have important roles in addition to collagen-synthesis: EGL-9, a PHY-related enzyme, has already been demonstrated to hydroxylate the  $\alpha$  subunit of a hypoxia-inducible transcription factor (HIF) (Epstein *et al.*, 2001). On the basis of *egl-9* null mutants being eggless (Egl) (Trent *et al.*, 1983) and resistant to the cyanide in *Pseudomonas aeruginosa* toxin (Gallagher and Manoil, 2001), it appears that this enzyme may have broader roles (Epstein *et al.*, 2001; Freeman *et al.*, 2003). The HIF-pathway is conserved in nematodes

and higher vertebrates but the nematode system represents a less complex version, having only a single HIF-hydroxylating enzyme compared to the three in humans (Myllyharju, 2003). In these systems, the hydroxylating activity of the PHY-like enzymes at a Leu-X-X-Leu-Ala-Pro motif maintains HIF in a state whereby it is targeted for ubiquitin-mediated degradation (Epstein *et al.*, 2001; Freeman *et al.*, 2003; Myllyharju, 2003). Loss of hydroxylating activity prevents degradation and consequently, HIF is transported to the nucleus where it activates cell survival mechanisms e.g. via the transcription of genes such as glucose transporter-1, and insulin-like growth factor (Freeman *et al.*, 2003). The HIF-hydroxylating enzymes constitute a distinct family of PHY enzymes because *phy-2* and *dpy-18* mutants are not defective in the HIF system (Epstein *et al.*, 2001).

### 1.11.3 Proteases involved in collagen biosynthesis

Proteases are required for the cleavage of the N- and C- termini of vertebrate collagen precursors and this processing is an essential step for the proper folding of such collagens (Prockop, 1998; Page and Winter, 2003). A vast number of proteases exist, not all of which are involved in collagen processing, and are classified via a complex grouping system into families and subfamilies according to their active site metals, binding site sequences and functional domains. The metalloproteinase group is one category of proteases and these are implicated in the regulation, synthesis, remodelling and degradation of the ECM in a variety of organisms. Subdivisions of metalloproteinases include the M12 type proteases, adamalysins (ADAMs) and matrixins. Further subgroups of the M12 proteases include the reprotlysins and astacins while a subgroup of the matrixins are the matrix metalloproteinases (MMPs). As will be explained in section 1.13.1.1, astacins are required for morphogenesis and for collagen processing. The proteases described in this section are only those associated with collagen processing during cuticle synthesis. Here, the defective N- and C-terminal processing of collagens in *bli-4* and *dpy-31* mutants (respectively) is discussed in terms of the necessary roles that the encoded enzymes have in cuticle synthesis.

#### 1.11.3.1 An N-terminal collagen protease: BLI-4

BLI-4 was identified as the *C. elegans* homologue of a family of calcium-dependent subtilisin-like serine endoproteases (also termed proprotein convertase proteinase (PCPs)) (Thacker and Rose, 2000; Thacker *et al.*, 2000a). This family of proteases cleaves a wide range of proteins (such as pro- $\alpha$ -mating factor in yeast, prohormones, proneuropeptides, growth factors and adhesion molecules in humans, bacterial toxins and viral glycoproteins) (Thacker *et al.*,

2000b) in order to convert them from inactive precursors to active mature peptides. Although related to the bacterial subtilisins, eukaryote homologues are much more specific in their cleavage, recognising only C-terminal sites containing dibasic residues, Arg-X-X-Arg (Thacker and Rose, 2000).

Despite being most similar to vertebrate FACIT collagens, which are not subjected to N-terminal proteolytic cleavage (Mazzorana *et al.*, 1993), all of the characterised *C. elegans* collagens do contain the characteristic subtilisin-like (R/K)XX(R/K) cleavage recognition site at their N-terminus (HBA) (Kramer, 1994b). In experiments analysing ROL-6 and SQT-1 collagens (e.g. from the strains *rol-6(su1006)*, *sqt-1(sc1)* and *sqt-1(e1350)*), cleavage at this recognition site was found to be necessary for normal collagen function and for efficient incorporation of the collagens into the cuticle (Yang and Kramer, 1994). All *C. elegans* collagens contain this recognition site, which may be targeted by the proteolytic activity of BLI-4 (Thacker and Rose, 2000). Four subtilisin-like genes exist in the *C. elegans* genome (*kpc-1*, *kpc-2/egl-3*, *kpc-3/aex-5*, *kpc-4/bli-4*) and of these only the *bli-4* gene product has been implicated in cuticle processing due to its Bli phenotype, a phenotype associated with cuticle defects (Thacker and Rose, 2000).

All but one of the *kpc-4/bli-4* alleles tested to date are non-viable, being embryonic or larval lethal (Thacker *et al.*, 2000a). The viable *bli-4(e937)* mutants exhibit adult specific cuticular blistering (Peters *et al.*, 1991). The lethality of most *bli-4* alleles indicates that the encoded protease is essential in early development (including late embryogenesis) (Peters *et al.*, 1991) while the Bli phenotype illustrates its importance to the adult cuticle. Its role in early development could be in the maintenance of the larval cuticle or processing of another substrate not associated with the ECM (Peters *et al.*, 1991) (Thacker and Rose, 2000). Since *bli-4(e937)* worms appear wild type at larval stages, it seems that the larval and adult cuticle requirements are independently mutable (Peters *et al.*, 1991). As well as being hypodermally expressed, *bli-4* expression is also observed in neural tissue (Thacker and Rose, 2000), strengthening the idea that this enzyme has distinct roles outside of the cuticle.

The *bli-4* locus is complex and, through alternative splicing of 12 exons, encodes for at least nine different gene products (A-I) each of which differs in their substrate-determining C-terminal (which can range in length from nine to 286 residues) (Thacker and Rose, 2000). The complexity of the gene and the consequential varied substrate specificities of the resulting

isoforms have been proposed to facilitate the differing cuticle compositions at each stage of the life cycle (Thacker and Rose, 2000). While all of the isoforms are affected in lethal mutants, it is the isoforms A, E, F, G and H which are eliminated (via a 3325 bp deletion) in the viable *bli-4(e937)* mutant and which therefore must be specifically involved in the construction of the adult cuticle (Thacker and Rose, personal communication). The argument for adult-specific processing being affected in this mutant strain is strengthened by the fact that an adult cuticle is a prerequisite (but not sufficient) for the formation of blisters. This was shown by crosses with *bli-4* and heterochronic mutants, *lin-29* and *lin-14* (Peters *et al.*, 1991). The observations above, in addition to the fact that the Bli phenotype is a result of aberrant struts and that *bli-4(e937)* mutants fail to secrete BLI-1::GFP (Crew and Kramer, personal communication), suggest that the targets of the mutant isoform/s of the *bli-4(e937)* strain are the adult-specific and strut-localised BLI-1 and BLI-2 collagens (Crew and Kramer, personal communication). It is probable that collagens whose molecular lesions lead to Dpy and Sqt phenotypes are processed normally since these phenotypes are both epistatic to Bli (Higgins and Hirsh, 1977). Thus, if collagens that produced these phenotypes were mutant, the Bli phenotype would not be observed. This is exemplified by the fact that *sqt-3*, *dpy-10*, *dpy-13* and *dpy-5* morphological mutants suppress the *bli-4* Bli phenotype (Peters *et al.*, 1991). Typical for the subtilisin-like family, all isoforms of BLI-4 are themselves synthesized as inactive precursors that contain a signal peptide, a prodomain domain, a catalytic domain, a conserved middle (or P) domain (which directs folding and catalytic activity), and a variable C-terminal domain (Thacker and Rose, 2000).

A BLI-4 homologue (termed 'blisterase') has recently been identified in *Onchocerca volvulus*. Its cleavage site is RRKR. Seven subtilisin-like homologues exist in humans (Poole *et al.*, 2003).

### 1.11.3.2 Collagen C-terminal proteases: Astacins

While BLI-4 processes the N-terminal of collagens, a metalloprotease of the astacin family can target the C-terminal pro-collagen domain (Novelli *et al.*, 2004).

The first collagen C-terminal protease to be identified was in a vertebrate system. Protease activity was attributed to a bone morphogenic protease-1 (BMP-1), a member of the astacin family of zinc-dependent M12 metalloproteases (Li *et al.*, 1996; Prockop *et al.*, 1998). Pro-collagen C proteinase (PCP) activity is exhibited by both BMP-1 and its longer splice

variant TOLLOID (Tld) (Takahara *et al.*, 1994). The latter functions in the establishment of the dorsoventral axis of *Drosophila melanogaster* embryos (Hishida *et al.*, 1996; Li *et al.*, 1996; Wardle *et al.*, 1999). Vertebrate BMP-1/Tld astacin proteases have a wide range of substrates and consequently have many roles in both developmental and mature systems (Hartigan *et al.*, 2003). In a variety of organisms, roles include hatching (Hartigan *et al.*, 2003), morphogenesis and pattern formation (Wardle *et al.*, 1999). In terms of the aforementioned morphogenesis, in addition to cleavage of collagen, astacin activity targets other ECM molecules including prollysyl oxidase, laminin 5, probiglycan and chordin (Suzuki *et al.*, 1996; Wardle *et al.*, 1999; Rattenhol *et al.*, 2002).

Astacins are distinguished from other zinc-dependent metalloproteases (metzincins) by a typical HEXXHXXGFXHEXXRXDR protease domain sequence (HEXXH within this sequence is the zinc binding motif) (Bond and Beynon, 1995). In addition to this protease domain, they usually have 1) a signal sequence, 2) a prodomain that mediates furin-mediated activation, and 3) a substrate-determining C-terminal extension that can include different combinations of the following domains: SXC, epidermal growth factor- (EGF)-like, CUB (domains that are similar to- and named after their presence in complement components C1r/C1s, embryonic sea urchin proteins *Uegf* and in *BMP-1*) and thrombospondin-like (TSP) domains (Mohrlen *et al.*, 2003) (Bond and Beynon, 1995).

A publication by Novelli *et al.* (2004) has attributed a *C. elegans* BMP-1 homologue to the *dpy-31/nas-35* locus, and has implicated it in the C-terminal processing of at least a subset of *C. elegans* collagens (Novelli *et al.*, 2004). Loss-of-function mutants exhibit cuticle defects (e.g. *Dpy* phenotype) as well as larval- and embryonic- lethality, particularly at temperatures above 15°C. Embryonic and L1 stages are particularly susceptible to lesions within this locus, and this is consistent with the fact that the only proven substrate of this protease is SQT-3, a collagen that is expressed in embryo and L1 stages. SQT-3 encodes a YCALD motif that is believed to be the protease recognition site. All members of the SQT-3 collagen family invariantly encode this motif and it is therefore likely that these collagens are targets of C-terminal proteolysis by DPY-31/NAS-35 (Novelli *et al.*, 2004). DPY-31/NAS-35 shares 43% similarity to human BMP-1 and contains: a signal peptide, a prodomain, an astacin catalytic domain, an EGF motif, a CUB domain and a TSP-type-1 repeat (Novelli *et al.*, 2004).

There are more astacins encoded by the *C. elegans* genome than in any previously studied organisms; 40 active astacins have been identified in *C. elegans* compared to 2-3 in organisms with larger genomes (Mohrlen *et al.*, 2003). There appears to be a level of redundancy in these genes since very few of these have visible RNAi phenotypes. Each *C. elegans* astacin has been assigned into one of six subgroup according to their domain structures (Mohrlen *et al.*, 2003) DPY-31/NAS-35 falls into subgroup V. Other members of this family are discussed in a later section as their roles are not directly related to collagen processing.

#### 1.11.4 Regulators of protease activity

##### 1.11.4.1 BLI-5 is a serine protease inhibitor

The locus responsible for *bli-5* morphological mutants was recently confirmed to encode for a serine protease inhibitor (Simmer *et al.*, 2003). It is interesting to note the apparent contradiction in the fact that disruption of a serine protease such as BLI-4 (described above) results in a similar phenotype to a protease inhibitor such as BLI-5. How BLI-5 activity contributes to cuticle synthesis is as yet unknown. However, it is likely that its role is in the control of proteolysis. Again, due to the Bli phenotype, it appears that the *bli-5* lesion may only affect the cuticle associated with the struts. *bli-5(e518)* is a Ser-Lys mutation (Simmer *et al.*, 2003) that occurs outside of the defined trypsin inhibitor active site and it is not known how, or if, such a mutation affects the inhibitory activity of the enzyme.

##### 1.11.5 Collagen cross-links and their catalytic enzymes

During cuticle biogenesis, trimer processing is followed by the formation of higher order structures. In *C. elegans* and other nematodes, these are mediated through intermolecular disulphide-(Schofield *et al.*, 1974), tyrosine-derived-(Fujimoto, 1975; Fujimoto *et al.*, 1981) and glutamine-derived bonds (Mehta *et al.*, 1990). While disulphide bonds, which are formed by thioredoxins, are soluble in detergent or 2- $\beta$ -mercaptoethanol (2-ME)-containing buffers, (Cox *et al.*, 1981b; Cox, 1992), the latter two covalent bonds are non-reducible as 25% of the adult cuticle remains insoluble after  $\beta$ -ME treatment (Cox *et al.*, 1981b). Differing extents of solubility are observed in different larval stages, suggesting stage-specific cross-linking. Dauer cuticles, in particular, are highly cross-linked and this extent of cross-linking must contribute to the resilience of this stage to adverse conditions (Cox *et al.*, 1981b). While tyrosine-derived bonds are formed by the catalytic action of animal haem peroxidases, transglutaminases (TGase) catalyse the formation of glutamate-derived bonds (Cariello *et al.*, 1990; Greenberg *et al.*, 1991). As well as the non-reducible nature of these bonds, the latter of these bonds is additionally

resistant to proteolysis, only being destroyed upon the complete degradation of the peptide chains (Greenberg *et al.*, 1991). An additional cross-linking bond is observed in *C. elegans* cuticles and is hydroxylysine derived. However, it is restricted to dauer cuticles (Cox *et al.*, 1981a). This is interesting since these are the predominant intermolecular bonds in vertebrate collagens (Bornstein, 2003), with disulphide and glutamate-derived cross-links being less abundant, but necessary. Tyrosine-derived cross-linking is not observed in vertebrate collagens; however, they have been identified in the (cataract) vertebrate lens (Garcia Castineiras *et al.*, 1978; Fetterer and Rhoads, 1990). The different types of bonds observed in vertebrates and nematodes presumably alters the properties of the distinct collagens and this may reflect the differing structures that they form; the majority of vertebrate collagens form fibrils unlike the non-fibrillar, cuticular *C. elegans* collagens.

The necessity of cross-linking bonds in nematodes and other organisms has been demonstrated by the mutant/lethal phenotypes observed upon disruption of cross-linking enzyme activity; TGase deficiency in mice, for example, results in neonatal death (Matsuki *et al.*, 1998) while phenotypes that range from mild blister to lethal are observed in *C. elegans* when certain animal haem peroxidase enzymes are mutated or are ablated (Dyox) (Edens *et al.*, 2001). Dpy phenotypes are also associated with the absence of, or lesions in, thioredoxins (Ko and Chow, 2002). Further demonstration for the necessity of collagen cross-linking in *C. elegans* comes from collagen mutants with lesions in covalent bond-forming residues. For example, mutations that remove a conserved C-terminal domain cysteine residue in SQT-1 collagens prevent the formation of tyrosine-derived cross-links and result in a severely disrupted cuticle (Yang and Kramer, 1999).

#### 1.11.5.1 DPY-11 is a thioredoxin

Thioredoxins are found in invertebrates and vertebrates and are cellular redox cofactors that catalyse various thiodisulphide exchange reactions which result in disulphide bond formation e.g. during antioxidant defense, modulation of apoptosis and immune responses in vertebrates (Sadek *et al.*, 2003). The reducing activity of this important family of enzymes, that includes PDI along with other ER proteins in the *C. elegans* genome, is driven by one or more essential CXXC motif(s) within the redox active site (Martin, 1995). Thioredoxins have an essential role in development (Baird and Emmons, 1990); knock-out mice for a thioredoxin gene are embryonic lethal (Matsui *et al.*, 1996). They are also implicated in the control of *C. elegans* body morphology. The latter function, which includes regulation of the male tail, was

demonstrated by the identification of the locus responsible for the morphological mutant, DPY-11, as being a novel and active thioredoxin in the *C. elegans* genome (Ko and Chow, 2002). It represents one of 38 that are potentially encoded. Despite its substrate being as yet unidentified, its hypodermal expression and Dpy mutant phenotype implicate it in the formation of the cuticle (Ko and Chow, 2002). It has been proposed that its ECM target may be cuticular collagens that may be subjected to cross-linking or disulphide bridge formation (Ko 2002). However, another potential way that thioredoxin activity could mediate the cuticle morphology is through acting on signalling molecules (Ko and Chow, 2002). This was proposed when *dpy-11* was implicated as one of the *ram* (ray morphology abnormal) genes, whose products were shown to mediate interactions between the rays and the cuticle (Baird and Emmons, 1990). These signalling-molecule-targets were proposed to be deposited in the ECM to facilitate interactions between the ray and the cells (Baird and Emmons, 1990). Thus, modification of the cuticle by these thioredoxins may not be through their direct action on the collagen molecules themselves.

The DPY-11 protein comprises 256 amino acids and contains signal peptide-, thioredoxin-, transmembrane- and C-terminal- domains, all of which are essential. Membrane anchorage and the C-terminus KKTK retention signal are both important to DPY-11's biological function (Ko and Chow, 2002).

The partial loss-of-function mutant, *dpy-11(e224)*, which arises through a glycine substitution, exhibits a Dpy phenotype at the mid-body and tail regions and bifurcated and branched alae (as indicated by SEM imaging) (Thein *et al.*, 2003). The null mutant (*dpy-11(e1180)*) results in a more severe Dpy and slow growing phenotype and is phenocopied by RNA-treatment. Both alleles have male tails which are abnormal in morphology but which are functional (Ko and Chow, 2002).

#### 1.11.5.2 Tyrosine-derived cross-links are catalysed by peroxidases

The di-, tri- and isotri-tyrosine cross-links in the *C. elegans* cuticle are biphenyl linkages of tyrosine molecules (Fujimoto, 1975). The former two have been identified in a number of structural proteins in at least three phyla (arthropoda, nematoda and vertebrate) (Fujimoto, 1975) while isotri-tyrosine cross-links to date, have only been associated with cuticles of nematodes (Fetterer and Rhoads, 1990). Tyrosine cross-linkages are present in the cuticle of *Ascaris suum* (Fetterer and Hill, 1993), in the cuticle of *Ascaris lumbricoides* (Sakura and Fujimoto, 1984)

(Fujimoto *et al.*, 1981), in the sheath of infective *Haemoncus contortus* (Fetterer and Rhoads, 1990) and in cuticulin of *Ascaris lumbricoides* (Fujimoto, 1975). As stated above, ablation of these tyrosine-derived cross-links in the *C. elegans* cuticle is highly disruptive to morphology and can be lethal.

Tyrosine-derived cross-links are catalysed by peroxidases and this was first demonstrated in inhibitor studies in *Ascaris* (Fetterer *et al.*, 1993) and since then, focus has been on the (predominantly alae-residing) cuticulins which are highly cross-linked via tyrosine residues (Parise and Bazzicalupo, 1997) (M. Sapio, personal communication). Lassandro *et al.*, 1994, were able to demonstrate that CUT-2 can be cross-linked *in vitro* by horse radish peroxidase (Lassandro *et al.*, 1994).

### 1.11.5.3 Duox: a dual oxidase that catalyses the formation of tyrosine cross- links

Dual oxidase (Duox) enzymes, identified in humans and in *C. elegans*, have two catalytic domains: an N-terminal NADPH-oxidase domain (homologue of gp91 *phox*) and a C-terminal domain with homology to animal haem peroxidases (they were first characterised for their role as NADPH-oxidases). By activity assays using recombinant isolated domains, the latter domain was identified as being responsible and sufficient for the catalysis of tyrosine-derived cross-links (in a synthetic substrate) while the NADPH-oxidase domain is thought to provide the hydrogen peroxide substrate ( $H_2O_2$ ) for this reaction (Edens *et al.*, 2001). There are two *C. elegans* Duox genes, f53g12.3 and f56c11.1, whose encoded proteins share 94% similarity and which in addition to their two catalytic domains, encode signal peptide and transmembrane motifs. However, f53g12.3 is slightly truncated at its C-terminus and this is believed to abolish its NADPH-oxidative activity (Edens *et al.*, 2001). No other NADPH-oxidase homologues are encoded in the *C. elegans* genome and thus f56c11.1 probably represents the only  $H_2O_2$ -donating species. Disruption of these genes by RNAi, both singly and doubly produces severe morphological defects and lethality. Morphologically mutant phenotypes include mild Bli, severe Bli, Dpy, movement defects, a translucent body, or a combination of these phenotypes (Edens *et al.*, 2001) (these will be discussed in chapter 5). Electron microscopy confirmed that the medial strut regions of RNAi-treated cuticles are disrupted (Edens *et al.*, 2001).

f56c11.1 has recently been attributed to the *bli-3* locus (Simmer *et al.*, 2003). The lesions in the two available *bli-3* alleles (*bli-3(e767)* and *bli-3(n529)*, which are mild and medium Dpy/Bli respectively, are positioned in the peroxidase domain (Simmer *et al.*, 2003).

The sulfation of tyrosine, catalysed by a tyrosyl protein sulfotransferase enzyme, may be a prerequisite for the cross-linking of nematode collagens since RNAi-treatment to target the enzyme results in moult defects and belted worms (Kim *et al.*, 2005). Similar O-sulfation has been shown to be a common modification of secreted proteins (Bundgaard *et al.*, 1997), and has been found to occur in coagulation factors and in proteins involved in inflammation and leukocyte adhesion (Kim *et al.*, 2005).

#### 1.11.5.4 Transglutaminases and PDIs

Transglutaminases (TGases) catalyse the formation of non-reducible  $\epsilon$ -( $\gamma$ -glutamyl)lysine cross-links between substrates that include fibrinogen, actin, myosin, fibronectin, laminin and collagen in many vertebrate and non-vertebrate systems (Griffin *et al.*, 2002). TGases are involved in a diverse number of systems including cell death in *C. elegans* (Madi *et al.*, 1998), blood clot formation (Murthy *et al.*, 1991) and the early development of filarial parasites (Mehta *et al.*, 1992). Aberrant TGase activity is associated with many human diseases (including those related to the epidermis of the skin) and the inability to release microfilaria in filarial parasites. The role of these enzymes in the cuticle was demonstrated by the thinner L4 cuticles of filarial parasites when treated with TGase inhibitors (Chandrashekar *et al.*, 2002).

In *Dirofilaria immitis*, TGase activity was found in a protein exhibiting sequence similarity, not to the previously characterised TGases, but to PDIs of the thioredoxin family (Chandrashekar *et al.*, 1998). Subsequently, a *C. elegans* orthologue of this gene was identified and termed PDI-3 (Eschenlauer and Page, 2003). Eschenlauer *et al.* (2003) confirmed that PDI-3 has TGase-, isomerase- and chaperone- activity and the thioredoxin motif of the protein was implicated in mediating the TGase activity (Blasko *et al.*, 2003; Eschenlauer and Page, 2003). Subsequently, analysis of PDI-1 and PDI-2 confirmed similar activity in these two enzymes (but they have less TGase activity than PDI-3) (Eschenlauer and Page, 2003). Thus, the functions of PDIs (PDI-1, PDI-2, and PDI-3) in the cuticle are multifarious: they catalyse the formation of disulphide bonds via their isomerase activity, they (potentially) form  $\epsilon$ -( $\gamma$ -glutamyl)lysine cross-links and have chaperone activities. In addition, PDI-2 is the  $\beta$  subunit of a collagen hydroxylase. However, PDI-2 is the only PDI that can act as the  $\beta$  subunit in PHY complexes (Winter and Page, 2000).

PDI-3 RNAi has no obvious effect on wild type animals and may reflect a level of redundancy in the cross-linking enzymes (Eschenlauer and Page, 2003). However, in backgrounds of

*dpy-18(e364, bx26 and ok162)* and *sqt-3(e2117)*, *dpy-18* RNAi produces Dpy morphology, egg-laying defects (Egl), tail defects (in male and hermaphrodite), bulging vulva defects and seam cell aberrations. Additionally, larval lethality is observed in *sqt-3(2117)* backgrounds (Eschenlauer and Page, 2003). Inhibition of the *D. immitis* orthologue produces moulting defects (Chandrashekar *et al.*, 2002).

## 1.12 Body length mutants

The body shape mutants described above elicit gross morphological abnormalities by disrupting the normal cuticle structure but do not directly or actively regulate body length (Nystrom *et al.*, 2002). Body length is under the influence of a TGF- $\beta$ -like signalling pathway, the components of which are described below. Mutations to these components can result in long (Lon) and small (Sma) phenotypes.

### 1.12.1 The TGF- $\beta$ pathways

The family that includes transforming growth factor- $\beta$  (TGF- $\beta$ ) and TGF- $\beta$ -like molecules is a group of multifunctional hormones that elicit concentration-dependent responses (Heldin *et al.*, 1997; Padgett *et al.*, 1998) by stimulating a signalling pathway. The TGF- $\beta$ -like signalling pathway is a relatively simple one that includes a ligand, two transmembrane receptors (type I and type II) and primary downstream components (termed Smads/Smas) which translocate to the nucleus and act as transcriptional activators (Padgett *et al.*, 1998).

A number of distinct TGF- $\beta$ -related pathways exist which utilise different TGF- $\beta$ -like ligands, TGF- $\beta$ , the activins, and bone morphogenic proteins (BMPs) in addition to other pathway-specific components. The various roles of these different TGF- $\beta$ -like pathways include many developmental processes such as the inhibition of cell proliferation (Yingling *et al.*, 1996), regulating growth, and patterning of many tissue types (Padgett *et al.*, 1998; Inoue and Thomas, 2000). In vertebrates and insects, BMPs act in a pathway that regulates early dorsoventral patterning of embryos. In *C. elegans*, there are at least two pathways: one controls the transition to the alternative third-stage dauer (Savage *et al.*, 1996; Savage-Dunn *et al.*, 2000) in response to various environmental factors (Inoue and Thomas, 2000), and another controls body length and male tail ray pattern formation (Morita *et al.*, 1999). As would be expected, mutation to components of the pathways that regulate these process result in aberrant dauer formation (*daf*) and altered body size (long or small worms with male tail defects) respectively.

### 1.12.2 The DBL-1 pathway is branched and controls male tail patterning and body length.

The components of the pathway responsible for body length and male tail patterning include: the pathway-specific ligand, DBL-1; the type I receptor (SMA-6); the type II receptor (DAF-4); and the downstream Dwarfing components (SMA-2, SMA-3 and SMA-4) (Heldin *et al.*, 1997). Other components include: the negative regulator, LON-2 (T. Gumienny, personal communication), a collagen (LON-3) (Nystrom *et al.*, 2002) and a regulator of ploidy (LON-1) (Morita *et al.*, 2002). All but DAF-4 are unique components of the body size and male tail-patterning pathway. The DBL-1 pathway has several independent gene targets and consequently, there are separate regulatory pathways for early larval elongation, late larval elongation, ploidy, and male tail patterning. All of the pathway components were identified: through mutagenesis screens searching for new Lon and Sma phenotypes (Savage *et al.*, 1996; Nystrom *et al.*, 2002) and suppressors of Sma phenotypes (Savage *et al.*, 1996), and through differential hybridisation on arrayed cDNAs (Mochii *et al.*, 1999).

Type I (SMA-6) (Yoshida *et al.*, 2001) and type II (DAF-4) receptors, to the latter of which the DBL-1 ligand binds, are serine/threonine kinases. Upon ligand binding, they form a heterotetramer in order to facilitate the type II-mediated phosphorylation of the type I receptor at a region rich in glycines and serines (Heldin *et al.*, 1997). Consistent with the global role of DAF-4 in all TGF- $\beta$ -like pathways, it is ubiquitously expressed in neurons, the intestine, pharynx, hypodermis and body wall muscle (Yoshida *et al.*, 2001). In contrast, SMA-6 is only associated with body size and consequently, its expression is restricted to the hypodermis (Yoshida *et al.*, 2001).

Directly downstream of the receptors are the related yet non-redundant dwarfins, SMA-2, SMA-3 and SMA-4, which are phosphorylated at a C-terminal SSXS motif by the kinase activity of the activated receptor complex. Activated SMAs oligomerise with other SMAs in various combinations, translocate to the nucleus, recruit transcription factors and finally bind DNA in order to modulate transcription of downstream target genes. SMAs can also act negatively and prevent action of positively acting groups (Padgett *et al.*, 1998).

Downstream of SMAs, there are numerous independent branches of the pathway that all result in the alteration of body size. Each branch is mediated by separate SMA targets and regulates late

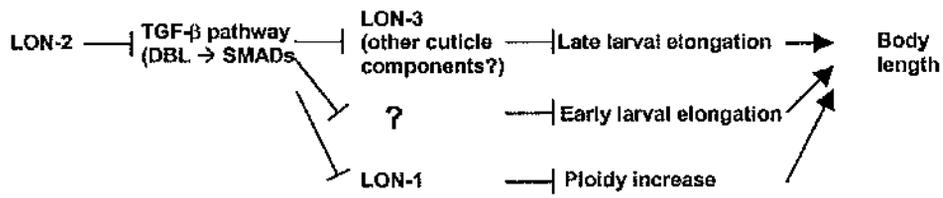
larval elongation, early larval elongation or ploidy (figure 1.5). One branch involves the cuticle collagen LON-3 and regulates late larval elongation. LON-3 itself is regulated by the DBL-1 pathway at a post-transcriptional level through modulation of the processing, degradation or synthesis of either its mRNA or protein. A separate branch downstream of the SMAs is mediated by the pathogenesis-related (PR)-protein homologue, LON-1, a protein that regulates hypodermal ploidy. The fact that these two pathways are distinct or have limited cross talk is illustrated firstly, by the fact that *lon-1* mutants exhibit altered ploidy whilst *lon-3* mutants have normal ploidy and secondly, that *lon-3* and *lon-1* body size mutants have an additive effect (Maduzia *et al.*, 2002). LON-2, a heparan sulfate proteoglycan, is assumed to function upstream of the DBL-1 ligand as a negative regulator of the pathway (T. Gumienny, personal communication).

Hyperactivation of the signalling pathway occurs through the over-expression of *dbl-1* ligand (Nystrom *et al.*, 2002; Suzuki *et al.*, 2002), loss of function of the LON-2 negative regulator (Nystrom *et al.*, 2002), or loss of function of the LON-3 collagen, and results in a Lon phenotype. Conversely a small phenotype is observed when there is over-expression of the *lon-1* negative regulator, loss of function of the ligand, or over-expression of *lon-3*. In this respect, *lon-3* is unique as no other collagen has the reciprocal property whereby it can produce long or small phenotypes according to whether it is under or over-expressed (Nystrom *et al.*, 2002).

Control of body length by DBL-1 is not mediated by dictating the number of somatic cells since Lon and Sma mutants appear to have wild type cell numbers (Morita *et al.*, 1999; Nystrom *et al.*, 2002) and consequently, there is no correlation between cell number and body size. Size, therefore, is thought to be altered by a relative increase or decrease in cell size and cell ploidy (Morita *et al.*, 1999; Suzuki *et al.*, 2002). Ploidy is determined by the number of rounds of endoreduplication that occurs, that is, the number of repetitive nuclear DNA synthesis cycles that occur without division (Flemming *et al.*, 2000).

### 1.12.3 Interaction between components controlling body size and the cuticle

There is some "cross talk" between the processes that determine body size via the TGF- $\beta$ -like pathway and those involved in cuticle composition and body shape. This is illustrated by the complex genetic interactions between the genes of the TGF- $\beta$ -controlled collagen, LON-3, and the cuticle collagens, SQT-1 and ROL-6 (Suzuki *et al.*, 2002). *sqt-1* null mutants are able to suppress both the Lon (loss-of-function) and the Sma (over-expression) phenotypes of *lon-3*,



**Figure 1.5: The body length determining pathway.** Ploidy and cuticle construction are regulated by separate branches. Modified from Suzuki *et al.* (2002).

which suggests that the presence of SQT-1 is required for the manifestation of the Lon phenotype. *rol-6* null mutants, on the other hand, are able to suppress only the Lon phenotype, and actually increase the expressivity of the Sma phenotype of *lon-3* over-expressing worms. Further, *lon-3* nulls are able to suppress the phenotypes of *sqt-1* and *rol-6* morphological mutants. The conclusion of these observations is that not only is SQT-1 required for the Lon phenotype, but that SQT-1 also requires the expression of *lon-3*. This interdependency is mediated in a direct or indirect manner (Nystrom *et al.*, 2002). Thus, functional SQT-1 and ROL-6 are required for DBL-1 to regulate body length (Nystrom *et al.*, 2002). It is also possible that DBL-1 may regulate other cuticular components (Suzuki *et al.*, 2002).

### 1.13 Other factors involved in *C. elegans* morphogenesis and moulting

Morphological, organogenesis and moulting defects can arise as a result of mutations to, or ablation of, enzymes that are not directly involved in the steps of collagen biosynthesis. In the last few sections about morphology, these enzymes are discussed.

#### 1.13.1 Proteases

The N- and C- terminal collagen propeptidases, BLI-4 and DPY-31/NAS-35, described above, are non-viable when null, and are therefore essential components in collagen processing (Thacker *et al.*, 1995; Novelli *et al.*, 2004). Despite being unable to fully compensate for the loss of these proteases, other proteases do exist that have roles in cuticle biogenesis as demonstrated by their mutant morphological phenotypes. The substrates of these proteases have not been identified but are known to be ECM components. However, this is probably a reflection that they have multiple substrates and accordingly, multiple roles. As will be described below, many of the proteases can mediate parallel roles in cuticle biogenesis, moulting and ECM remodelling. The proteases identified to date that are of importance to morphology, fall into one of three metalloprotease families: the astacins (also termed M12A), the adamolysins (ADAMs) (M12B) or the cysteine proteases (M14).

ADAMs are zinc proteases characterised by a catalytic sequence HEXGHXXGXXHD. They also contain a prodomain, an adhesion-mediating disintegrin-like domain, a cysteine-rich domain, an EGF-repeat region, a transmembrane domain and a cytoplasmic tail. Some members of the family contain three thrombospondin type I-like (TSP) motifs instead of the transmembrane regions and are termed ADAMTS (Black and White, 1998). ADT-1 (Kuno *et al.*, 2002), SUP-17 (Wen *et al.*, 1997) and GON-1 (Blelloch *et al.*, 1999) are ADAMTS that have

been assigned roles in organogenesis and ECM morphogenesis on the basis of the respective abnormal male tail, vulva and gonadal structures of null mutants. Rather than mediating the actual organogenesis or forward migration of their respective structures (of the distal tips of the gonad, for example), the main function of the latter two proteases appears to be in remodelling of the ECM in front of the migrating structures. However, ADT-1 is conversely thought to actually remodel ray cuticle (Kuno *et al.*, 2002). It has also been suggested that these ADAMTS may also be involved in the proteolytic modification or activation of transcription factors and signalling components of their respective processes (Porter *et al.*, 2005). Important for the subject matter of this thesis that focuses on the cuticle, mammalian ADAMTS-2 has been identified as having pro-collagen N-proteinase activity (Colige *et al.*, 1997). Thus at least some ADAMs may be directly involved in collagen processing/degradation. In addition, mammalian ADAMTS-2 has a role in development, illustrating the multifarious functions that ADAMs may have.

#### 1.13.1.1 Moulting defects as a consequence of aberrant protease activity

While the ADAMTS mutants display abnormal organ morphology, the consequences of altered astacin (e.g. NAS-37) and cysteine protease (CPL-1 and CPZ-1) activities can be moulting defects and embryonic lethality.

As stated previously, the family of astacins (containing 40 genes) are sub-grouped according to domains (Mohrlen *et al.*, 2003). Extensive RNAi analysis of these genes identified only a subset of these having mutant phenotypes and indicated a level of redundancy between them (Mohrlen *et al.*, 2003). The only identified mutable astacins were DPY-31/NAS-35 (which has been previously described), NAS-37 and HCH-1. It is important to note, when analysing such data, that investigation of DPY-31 demonstrated a significant insusceptibility to RNAi due to the fact that RNAi-treated animals are only mildly Dpy while null mutants are embryonic lethal (Novelli). It was suggested by Mohrlen *et al.* (2003) that RNAi of astacins has reduced penetrance.

NAS-37 and HCH-1 both fall into the same subgroup as DPY-31/NAS-35 due to the presence of TSP domains. Thus the TSP domains, for which this subgroup is characteristic, appears to specialise these enzymes for cuticle/matrix degradation. TSP-1 has previously been implicated in ECM binding (Kuno and Matsushima, 1998). The degradative activity of HCH-1 is required for the proper degradation of the eggshell (Hishida *et al.*, 1996) while NAS-37 is necessary for

moulting (Davis *et al.*, 2004). The latter enzyme is expressed in the head hypodermal cells and is thus believed to degrade a specific substrate collagen set at the anterior end of the worm. NAS-37 null and RNAi mutants fail to complete ecdysis. A constriction along their mid-body is the site of continued attachment of the old cuticle, which can be seen being dragged behind the worm as it moves. This is a typical moult defect. NAS-37 is also thought to participate in the synthesis of the new cuticle as mutants exhibit cuticle aberrations (Davis *et al.*, 2004).

Another group of collagens that are involved in moulting and cuticle synthesis are cysteine proteases which were first linked to moulting in *D. immitis* on the basis of assays using specific protease inhibitors (Richer *et al.*, 1993). Cathepsin-L (CPL) and Z-like (CPZ) proteins of the cysteine protease family have now been characterised and their activities attributed to embryogenesis and moulting. CPLs have been cloned in *C. elegans* and characterised in *Onchocerca volvulus*, *Brugia malayi* and *Brugia pahangi* and all have localisations, temporal expression patterns, and RNAi/inhibitor phenotypes consistent with their role in moulting (Hashmi *et al.*, 2002; Guiliano *et al.*, 2004). *ceCPL-1* (the only *C. elegans* CPL) is expressed cyclically, peaking 4 hours prior to the moult, and is hypodermally expressed. Its native protein accumulates in gut cells of late stage embryos and in the lumen of the yolk vesicles. Localisation here has been linked to the essential role that this protease plays in embryonic development (yolk processing) (Britton and Murray, 2002; Hashmi *et al.*, 2002). This is its primary role and RNAi of *cpl-1* is embryonic lethal (Britton and Murray, 2002). By limiting RNAi-treatment to L3 larvae that have completed embryogenesis, other phenotypes became apparent. Animals thus treated are slow to develop into adults and additionally have reduced egg production (Hashmi *et al.*, 2002). These results, complemented with the fact that *cpl-1*, when expressed 4 hours prior to the moult, becomes localised to both the old and new cuticles, indicate that CPL-1 has an additional role in moulting (Hashmi *et al.*, 2002).

Two CPZs are encoded in *C. elegans*, and *cpz-1* is expressed at the time of each moult (Hashmi *et al.*, 2004). *cpz-1* RNAi-treated animals have moult defects and some degree of embryonic lethality and additionally have abnormal head, tail and gonad morphology. The latter observations implicate CPZ-1 in having a role in cuticle synthesis. Such a role, distinct from the moult, is consistent with the adult expression of a *Toxocara canis* homologue and the role that human forms of this protease have in tumour progression (Hashmi *et al.*, 2004).

### 1.14 Regulators of moulting

Ablating certain non-enzymatic factors can also induce moulting defects. This includes the nuclear hormone receptors, *nhr-23* and *nhr-25* (whose ligands are yet unidentified) as well as the megalin receptor, *lrp-1*. NHR-23/CHR3 is a ligand-dependent transcription factor whose homologue, *Drosophila* DHR3, responds to ecdysone during *Drosophila* metamorphosis where it is thought to “reset” the hormone signalling pathway (Lam *et al.*, 1997; Kostrouchova *et al.*, 1998). Due to the cyclical expression pattern of ceNHR-23/CHR3 (peaking 2 hours prior to each larval moult), its epidermal-specific localisation and an RNAi-induced moulting defect, it must play a role in the moulting process (Kostrouchova *et al.*, 2001). Interestingly, RNAi-treatment results in Dpy phenotypes and a 5- to 11- fold decrease in levels of *dpy-7* expression (Kostrouchova *et al.*, 2001). These observations imply that NHR-23/CHR3-derived disruption of ecdysis may be linked to cuticle defects. Mutation of NHR-25 has also been associated with moulting defects as well as altered cuticular structures and accumulation of vesicles in the cuticle layers. From these observations, it was suggested that among the roles of this nuclear hormone receptor is the secretion of cuticular components (Silhankova *et al.*, 2005). About 270 *C. elegans* NHRs have been identified, several times the number found in other organisms but the roles of most of them have yet to be elucidated (Asahina *et al.*, 2000).

The *C. elegans* megalin receptor, LRP-1, like its mammalian homologue, is thought to mediate the essential process of cholesterol endocytosis since the *lrp-1*(RNAi) moulting phenotypes phenocopy those of carbohydrate starvation (media induced) (Yochem *et al.*, 1999). In insects, cholesterol has been proven to be necessary for the synthesis of ecdysteroid. However, no homologues of ecdysteroid receptors have been identified in *C. elegans*, nor has there been any evidence for synthesis of this hormone (Yochem *et al.*, 1999).

### 1.15 Basement membrane collagens

The basement membrane, a complex collagenous polygonal network, is the second *C. elegans* ECM. It underlies the hypodermis and surrounds the pharynx, gonad, intestine and body wall muscle (Kramer, 1997). It is composed of only two collagens, namely LET-2 and EMB-9 (Guo and Kramer, 1989; Gupta *et al.*, 1997; Norman and Moerman, 2000) and on account of the encoded Gly-X-Y domain being interrupted by ~20 small interruptions and 4 conserved N-terminal cysteine residues, these basement membrane-specific collagens are homologous to the vertebrate type IV collagens (Graham *et al.*, 1997). As would be expected for an ECM composed of a repertoire of only two collagens, both LET-2 and EMB-9 are essential, being

embryonic lethal when mutated or ablated (Guo *et al.*, 1991b; Gupta *et al.*, 1997). This shows they have essential functions during embryonic development (Guo *et al.*, 1991b). The composition of the *C. elegans* basement membrane can, to a certain extent, be altered by the ability of *let-2* to be alternatively spliced to form two isoforms (J. Kramer, personal communication) which are expressed in either the embryonic or larval/adult basement membranes (Sibley *et al.*, 1993). Thus there is developmental stage diversity in the composition of this structure (Sibley *et al.*, 1993; Graham *et al.*, 1997). Accordingly there appear to be specific embryonic and larval/adult roles for the basement membrane. The LET-2 and EMB-9 collagens, akin to their cuticle collagen counterparts, exist as trimers that are cross-linked to form higher order structures. Disruption of the triple helix via Gly substitution is potent and is more so than null mutations (embryos arresting in a temperature-sensitive fashion in the two-fold rather than the three-fold embryo) (Guo *et al.*, 1991a). Higher order structures are maintained by lysyl-derived cross-links which require lysyl residue hydroxylation. The importance of such cross-linking is indicated by the embryonic lethality observed upon inhibiting LET-268, a lysyl hydroxylase which is specific for basement membrane collagens (Norman and Moerman, 2000).

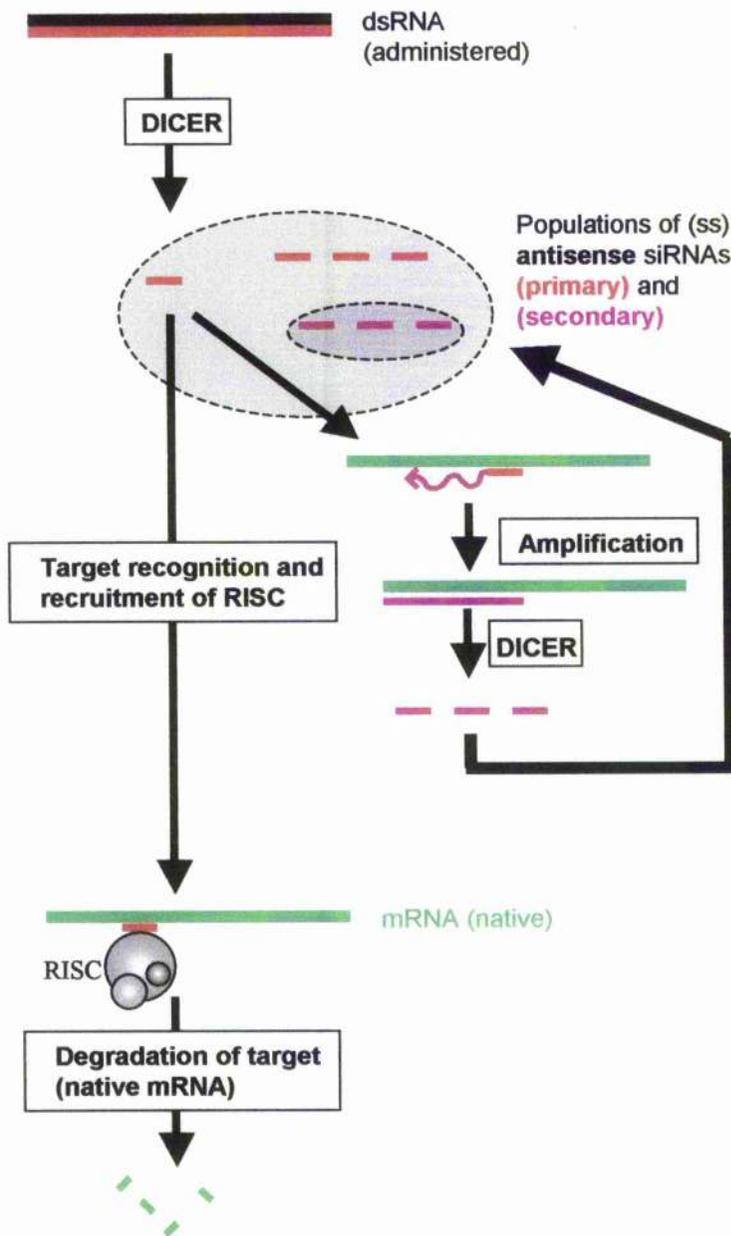
#### 1.16 RNA-mediated interference (RNAi)

RNAi is a technique, first described in *C. elegans* by Fire *et al.* (1998), that uses exogenous double-stranded RNA to temporarily interfere with the function of an endogenous gene. Cognate cytoplasmic mRNAs are targeted for degradation, which effectively results in the silencing of full gene expression at the post-transcriptional level. It is an invaluable reverse genetic tool for determining the functions of specific genes (Kamath *et al.*, 2000). Three methods of administering RNA corresponding to the gene under investigation exist, including microinjection of a solution of double-stranded (ds) RNA into the syncytial gonad of adults, soaking L4 larvae in a similar solution, and feeding worms on a lawn of bacteria expressing target gene dsRNA. Each of these are facilitated by the surprising ability of the RNAi effect to cross cellular boundaries including movement into the germline (Timmons *et al.*, 2003). Early experiments investigating the validity and effectiveness of RNAi in *C. elegans* used the two former methods and showed that interference mediated by dsRNAi is several orders of magnitude more efficient at producing an effect than any single stranded procedures; it was suggested that single strands may be degraded (Fire *et al.*, 1998). The feeding method usually involves the cloning of a 0.5 – 2 kb fragment of the gene of interest into the vector L4440 although no precise size limit has been determined to be optimally efficient. The construct is transformed into an RNaseIII-deficient bacterial strain on which worms are fed. Feeding has been shown to have a

similar potency to RNAi by injection although the F1 generation after feeding exhibits a wider range of phenotypes (Kamath *et al.*, 2000). This is a particularly useful method as it is less labour-intensive and it facilitates the widespread screening of genes by RNAi. Consequently, it was the method used to carry out a systematic functional analysis of the *C. elegans* genome by RNAi in which 86% of the 19,427 predicted genes were specifically targeted (Fraser *et al.*, 2000) (Kamath *et al.*, 2003) (Simmer *et al.*, 2003). Other RNAi screens have been previously described in this chapter.

Small quantities of RNA (only a few molecules of dsRNA per cell) are required in order to elicit an effect (Fire *et al.*, 1998; Montgomery *et al.*, 1998), and it was therefore suggested that there is some amplification of RNAi signal, instead of simple stoichiometric interference by the foreign RNA with endogenous mRNA. Indeed, in the current model that explains how RNAi elicits its potent effect, there are two ways in which amplification can take place. Figure 1.6 illustrates the key steps that occur during RNAi. The administered dsRNA is initially processed into short fragments (20-25 bp) termed short interfering (si) RNAs via the proteolytic activity of DICER. They then become single stranded. These fragments are highly unstable and while sense strands are immediately degraded, antisense strands only become stabilised in the presence of complementary sequences of native mRNA (Plasterk 2002), to which they bind (Baulcombe, 2002). At this position on a native mRNA strand, siRNAs recruit a complex termed the RNA interfering specificity complex (RISC), which degrades the target mRNA. The activity of DICER acts as the initial amplification step because it increases the number of recruiting RNAs by 10 to 20 fold (Baulcombe, 2002). Further amplification comes from the fact that siRNAs, once hybridised to native mRNA, can function as primers and initiate the polymerase-mediated synthesis of more RNA from the template to which they are bound. The RNA-directed RNA polymerases that function in this step are tissue specific, for example, being EGO-1 in germline cells and RRF-1 in somatic cells (Plasterk, 2002). In turn, this so-called secondary RNAi is converted into siRNA, thus amplifying the signal. This not only increases the quantity of siRNA that is present, but also enables "spreading". This occurs because these secondary siRNAs are derived from polymeration in a 5' direction from the initial site of binding of the primary siRNA. They represent sequences not present in the original administered dose. However, this effect does not spread more than a few hundred bases from the original site of binding (Plasterk, 2002).

Via these steps, the RNAi technique exerts its potent effects. However, RNAi is not 100% effective for all genes, in particular, genes with neural function (Timmons *et al.*, 2001), and



**Figure 1.6: Mechanism of dsRNA mediated mRNA silencing.** The administered dsRNAi is processed into primary siRNAs by the activity of DICER. siRNAs may undergo an amplification step in which they act as primers on native mRNA sequences, producing secondary dsRNAs which are processed by DICER. Primary and secondary siRNAs recruit the RISC complexes to native (homologous) mRNAs, thereby targeting them for RISC-mediated degradation. Figure modified from Plasterk (2002), and Novina and Sharp (2004).

consequently not all of the RNAi phenotypes observed to date correlate with the null phenotype of the gene in question. The technique is very specific in that, unless there is a gene (or multiple genes) that share more than 85% homology, the resulting phenotype can almost always be attributed to the targeted locus. In cases where homologies are over the 85% threshold, it must be considered that the phenotype may be the result of the targeting of more than one gene (Fire *et al.*, 1998).

There are over 10 gene products in the *C. elegans* genome that are involved in mediating ds-derived mRNA silencing, indicating the requirement of a natural gene-silencing phenomenon. In *C. elegans*, this is likely to be a system of genome protection which reduces the mobility of transposable DNA elements. Evidence for such a role comes from the fact that mutants, defective in RNAi silencing, have high a frequency of mutation (Baulcombe, 2002). This is an ancient system that has been observed in plants, insects, fungi, protozoa and other nematode species. Interestingly, the potency of the RNAi technique in the laboratory is facilitated by another *C. elegans* system. It is believed that the ability to take up RNAi administered by feeding uses a natural function used by the organism to extract from food the nucleic acids required for replication and transcription (Plasterk, 2002). Vertebrates also have specific responses to the presence of foreign dsRNA. However, in mammalian cells, this is in the form of a global panic response that results in the cessation of *all* (as opposed to homologous) mRNAs (Montgomery *et al.*, 1998). This vertebrate response is believed to be involved in reducing the genomic invasion of viral and transposable elements (Baulcombe, 2002). However, even with such control systems in place, 45% of the human genome is derived from ancient invasion of these elements (Plasterk, 2002).

### 1.17 Project Aims

This study is based on the use of a strain carrying a COL-19::GFP reporter construct

The aims of the project are five-fold:

- 1) To analyse the role of COL-19 in the cuticle and to characterise the wild type expression pattern of this collagen in the cuticle.
- 2) To establish which properties of COL-19 allow it to form a functional GFP-fusion.
- 3) To compare the patterns of COL-19::GFP in wild-type animals (the TP12 strain) and in morphological mutants in order establish the molecular basis of gross morphological phenotypes. This information will also be invaluable in elucidating the roles of different cuticular collagens.
- 4) To utilise the sensitivity of TP12 to detect cuticle defects in an RNAi screen.
- 5) To characterise the putative animal haem peroxidase, HPX-1, and establish whether it has a role in *C. elegans* cuticle synthesis.

**Chapter 2**  
***Materials and Methods***

## 2.1 Chemical abbreviations

2-ME	2-mercaptoethanol
$\beta$ -gal	$\beta$ -galactosidase
Amp	ampicillin
BSA	bovine serum albumin
cDNA	copy/complementary DNA
CIP	calf alkaline phosphatase
CPD	critical point dryer
DABCO	1, 4-Diazabicyclo[2.2.2]octane
DIC	differential interference contrast
DEPC	diethyl pyrocarbonate
DMF	N, N-Dimethylformamide
dNTP	dioxynucleoside triphosphate
dH <sub>2</sub> O	deionised water
EDTA	ethylenediaminetetraacetate disodium salt
FITC	fluorescein isothiocyanate
IPTG	isopropyl $\beta$ -D-thiogalactoside
IDA	iminodiacetic acid
Kan	kanamycin
Klenow	DNA Polymerase I, Large (Klenow) Fragment
LB	Lucia – Bectani
LDS	lithium dodecyl sulphate
Mab	monoclonal antibody
MES	2-Morpholinoethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
NGM	nematode growth medium
PBS	phosphate buffered saline
PMSF	phenyl methane sulfonyl fluoride
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
rNTP	ribonucleoside triphosphate
SDS	sodium dodecyl sulphate
<i>Taq</i>	<i>Thermus aquaticus</i>
TEMED	N, N, N', N'-tetramethyl ethylene diamine

Tet/tetracycline	tetracycline hydrochloride
TMB	3, 5, 3', 5'-tetramethylbenzidine
Tris-HCl	2-amino-2-(hydroxymethyl)-1,3-propanediol-hydrochloride
Tween 20®	polyoxyethylene sorbitan monolaurate
X-Gal	5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactopyranoside

## 2.2 Reagents

Sigma-Aldrich:	ampicillin, tetracycline, kanamycin, DEPC, ethidium bromide, 2-mercaptoethanol, PBS tablets, SDS, boric acid, Tween® 20 bromophenol blue, coomassie blue, phenol, chloroform:isoamylalcohol (24:1), poly-L-lysine (Sigma diagnostics).
Fisher Scientific:	EDTA, ethanol, CaCl <sub>2</sub> , Tris-HCl, NaH <sub>2</sub> PO <sub>4</sub> , KH <sub>2</sub> PO <sub>4</sub> , MgCl, glucose.
Promega:	IPTG, X-gal, dNTPs.
Invitrogen:	agarose (electrophoresis grade), NuPAGE sample buffer (4X), oligonucleotides.
Riedle de Haën:	NaCl, MgSO <sub>4</sub> , glycerol, acetone.
BDH:	Na <sub>2</sub> HPO <sub>4</sub> , NaOH, methanol, acetic acid.
BD Difco:	bacto tryptone, yeast extract, bacto agar, agar, peptone.
New England Biolabs:	BSA, 1 kb ladder mw markers, VentR® polymerase.
Roche:	proteinase K
ICN Biomedical:	urca
Stratagene:	<i>Pfu</i> DNA polymerase
Applied Biosystems:	<i>Taq</i> DNA polymerase
Pierce:	Coomassie Reagent protein assay system

## 2.3 Standard media

Ampicillin (1000X):	100 mg/ml stock in sterile dH <sub>2</sub> O. Filter sterilised and stored at -20°C.
BSA:	10 mg/ml stock. Stored at -20°C.

Coomassie blue solution:	0.25 g of Coomassie brilliant blue R-250, 45% methanol, 10% glacial acetic acid. The solution was filtered through a Whatman no. 1 filter to remove particles.
DEPC H <sub>2</sub> O:	0.1% DEPC (v/v) in sterile dH <sub>2</sub> O mixed overnight and autoclaved. Stored at room temperature.
Destain solution:	45% (v/v) methanol, 10 % (v/v) glacial acetic acid.
DNA sample buffer (10X):	50% (v/v) glycerol, 0.25% (w/v) bromophenol blue.
DTT:	1 M stock. Stored at -20°C.
EDTA:	0.5 M stock, pH 8.0. Autoclaved and stored at room temperature.
Ethidium bromide:	8 mg/ml stock in dH <sub>2</sub> O. Stored at room temperature.
Freezing solution:	5.85 g NaCl, 6.8 g KH <sub>2</sub> PO <sub>4</sub> , 300 g glycerol, 5.6 ml of 1 M NaOH in 1 L dH <sub>2</sub> O. Autoclaved and 3 ml of 0.1 M MgSO <sub>4</sub> added.
IPTG:	1 M stock. Filter sterilised and stored at - 20°C.
Kanamycin:	50 mg/ml stock in dH <sub>2</sub> O. Stored at -20°C.
Lucia -Bectani-(L)-broth:	1% (w/v) bacto tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl in dH <sub>2</sub> O, pH 7.0.
LB-agar:	L-broth + 15 g/L bacto-agar. Autoclaved and stored at room temperature. To make LB-agar plates, solid agar was made molten by heating, allowed to cool slightly, had the appropriate antibiotic added and was poured (using sterile techniques) into 9 cm diameter plates. Allowed to set. Unused plates could be stored for short periods at 4°C.
M9 buffer (10X):	3% (w/v) KH <sub>2</sub> PO <sub>4</sub> , 6% (w/v) Na <sub>2</sub> HPO <sub>4</sub> , 5% (w/v) NaCl, 10 mM MgSO <sub>4</sub> in dH <sub>2</sub> O. 10X stock autoclaved and stored at room temperature.
2-mercaptoethanol:	14.3 M stock.
MES (20X)	1 M MES, 1 M Tris base, 68.3 mM SDS (2% w/v), 20 mM EDTA. pH 7.3. Purchased from Invitrogen.
MOPS (20X):	1 M MOPS, 1 M Tris base, 69.3 mM SDS (2% w/v), 20mM EDTA. Purchased from Invitrogen.
NGM-agar:	0.3% (w/v) NaCl, 1.7% (w/v) agar, 0.25% (w/v) peptone, 0.0005% (w/v) cholesterol (1 ml/L of 5 mg/ml stock in ethanol), in dH <sub>2</sub> O. Autoclaved and 1 ml/L 1 M CaCl <sub>2</sub> , 1 M MgSO <sub>4</sub> and 25 ml/L KPO <sub>4</sub> (pH 6.0) added.

NGM-agarose:	Same as above but with agar substituted by agarose.
PBS:	purchased as tablets. Dissolved 1 tablet in 200 ml dH <sub>2</sub> O to obtain a 137 mM NaCl, 2.7 mM KCl, and 10 mM phosphate buffer solution (pH 7.4). Stored at room temperature.
PBST:	Same as the above with Tween® added to a final concentration of 0.2% (v/v). Stored at room temperature.
ProteinaseK:	20 mg/ml stock in dH <sub>2</sub> O. Stored at -20°C.
Recovery buffer:	2% (w/v) glucose in 1X M9 (made fresh each time).
SDS:	10% (w/v) stock. Stored at room temperature.
SOB medium:	20 g bacto-tryptone, 5 g bacto-yeast extract, 0.5 g NaCl, 10 ml of 250 mM KCl in 1 L dH <sub>2</sub> O (pH 7.0). Stored at room temperature.
SOC medium:	400 µl of filter sterilised solution of 1 M glucose added to 20 ml SOB medium. Aliquots stored at -20°C.
TBE (10X):	0.9 M Tris-HCl, 0.9 M Boric acid, 25 mM EDTA in dH <sub>2</sub> O (pH 8.0). Stored at room temperature.
TE buffer:	10 mM Tris, 1 mM EDTA in dH <sub>2</sub> O (pH 8.0). Stored at room temperature.
Tetracycline (1000X):	12.5 mg/ml stock dissolved in 50% (w/v) ethanol. Stored at -20°C.
X-gal:	2% (w/v) stock dissolved in DMF and stored at -20°C. Kept out of light.

Preparation of any standard reagents or media not detailed can be found in Sambrook *et al.*, Molecular cloning: a laboratory manual (1989). Where stated, solutions were autoclaved using the conditions: 120°C, 15lb/in<sup>2</sup> for 15 minutes.

2.4 *E. coli* strains

Name	Genotype	Notes	Use
OP50		Variant of the uracil-requiring OP50 strain (Brenner, 1974) which was transformed with a plasmid marker containing the Tet <sup>r</sup> marker. A gift from CGC.	Food source of <i>C. elegans</i>
HT115 (DE3)	<i>F<sup>-</sup>, mcrA, mcrB, In(rrD-rrE)1, rnc14::Tn10(DE3 lysogen: lacUV5 promoter-T7 polymerase)</i>	Tetracycline resistant, IPTG inducible promoter, T7 polymerase, RNase III minus (Timmons <i>et al.</i> , 2001). A gift from CGC.	Used in RNAi feeding
XL10-gold™	<i>Tet<sup>R</sup> Δ(mcrA)183Δ(mcrCB-hsdSMR-mmr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacZΔM15 Tn10 (Tet<sup>R</sup>) Amy Cam<sup>R</sup>]</i>	Ultra competent cells, blue white selection, <i>recA</i> and <i>endA1</i> mutations (to give greater stability to transformed DNA and better quality of purified plasmid, multiple genes ( <i>mcrA</i> , <i>mcrCB</i> , <i>mmr</i> , <i>hsdR</i> , <i>mcrF</i> ) that cleave cloned methylated DNA have been deleted. Purchased from Stratagene.	Used in general cloning
OneShot® TOP10 cells	<i>F<sup>+</sup> mcrA Δ (mrr-hsdRMS-mcrB C)Φ 80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ (ara-leu)7697 galU galK rpsL (Str<sup>R</sup>) endA1 nupG</i>	Competent cells that are derived from DH10B™ cells, blue white selection, <i>recA</i> and <i>endA1</i> minus (as above), mutations in <i>mcrA</i> , <i>mcrB</i> and <i>mrr</i> genes (as above), do not require IPTG during cloning. Purchased from Invitrogen.	Used in general cloning
M15(pREP4)	<i>NaI<sup>S</sup>, Str<sup>S</sup>, Rif<sup>S</sup>, Thi<sup>-</sup>, Lac<sup>-</sup>, Ara<sup>+</sup>, Gal<sup>+</sup>, Mtl<sup>-</sup>, F<sup>-</sup>, RecA<sup>+</sup>, Uvr<sup>+</sup>, Lon<sup>+</sup></i>	Contains a repressor plasmid (pREP4) that produces high levels of <i>lac</i> repressor that is highly regulated in order to prevent leaky expression prior to induction. pREP4 must be maintained by selection for kanamycin resistance. Purchased from Qiagen and made chemically competent.	Used for expression of recombinant proteins

Table 2.1: *E. coli* strains

## 2.5 Vectors and clones

Name of vector	Notes and use
PCR-Script	A PCR fragment cloning system used for cloning blunt end PCR fragments. Purchased from Stratagene.
PCR@ 2.1 TOPO TA	A PCR fragment cloning system used for cloning A-tail PCR fragments. Purchased from Invitrogen.
pQE-30	A vector for high-level expression of N-terminally His-tagged recombinant proteins. Purchased from Qiagen.
pPD129.36 (L4440)	Vector with two T7 promoter regions flanking the multiple cloning site. Functions as a template during <i>in vitro</i> RNA synthesis for RNAi microinjection and soaking. Can also be transformed into HT115(DE3) an <i>E. coli</i> strain for RNAi by feeding (Timmons, <i>et al.</i> , 2001). A gift from A. Fire (Carnegie Institution of Washington, Baltimore, USA).
pPD49.78	Was used in this study during the cloning of a Ty-tagged COL-19. A gift from A. Fire
pPD96.04	Used to fuse the promoter of a heterologous gene in-frame to a <i>GFPLacZ</i> marker. A gift from A. Fire.
pPAF207	Allows the in-frame fusion of a promoter sequence of a heterologous gene to a destabilised EGFP marker to allow the kinetics of transient gene regulation to be assayed. A gift from A. Frand (Massachusetts General Hospital, Boston, MA, USA).
Plasmid clone	
BA7.1	Contains the promoter sequence and coding region of <i>col-19</i> fused in-frame to a GFP marker. The <i>col-19</i> sequence was excised from this vector and replaced with promoter and coding sequences of other collagens. A gift from C. Shoemaker (AgResearch Ltd. Upper Hutt, New Zealand)
p76-16B	Contains 10.7kb <i>Xba</i> I wild type genomic fragment of <i>unc-76</i> . Rescues the Unc phenotype of DR96 in transformed lines following microinjection of DNA. Used as a marker. A gift from Laird Bloom (MIT center for Cancer Research, Boston, MA, USA).

Table 2.2: Vectors and clones.

## 2.6 Bacterial cultures

### 2.6.1 Antibiotic concentrations for bacterial culture

Antibiotics in bacterial cultures were used at the following final concentrations:

Ampicillin (Amp):	100 $\mu\text{g}/\text{ml}$
Tetracycline (Tet):	12.5 $\mu\text{g}/\text{ml}$
Kanamycin (Kan):	25 $\mu\text{g}/\text{ml}$
Streptomycin (Strep):	12.5 $\mu\text{g}/\text{ml}$

### 2.6.2 Bacterial culture on solid media

All *E. coli* strains were grown by incubation at 37°C for 16 hours on LB-agar plates (9 cm diameter plates, Greiner Labortechnik, Ltd) supplemented with the appropriate antibiotic.

### 2.6.3 Liquid cultures of bacteria

Using sterile techniques, single colonies were transferred to LB-broth supplemented with an appropriate antibiotic (Amp, Kan, Tet or Strep, used singly or in combination as stated in brackets where appropriate). Cultures were grown for 16 hours at 37°C in an orbital shaker.

### 2.6.4 Making competent M15 cells

25  $\mu\text{l}$  of a 10 ml L-broth cultures of M15 (non-competent) cells that had been grown overnight were used to inoculate 25 mls of L-broth and cultures were grown in an orbital shaker until an OD<sub>600nm</sub> of 0.4 was reached. Cultures were then centrifuged (10 minutes, 3000 rpm, 4°C) to form a cell pellet, which was subsequently resuspended in 0.5 x volumes (of the original culture volume) of cold CaCl<sub>2</sub>. This cell suspension was kept on ice for 30 minutes, centrifuged and resuspended in 0.1 volumes of 0.1M CaCl<sub>2</sub>. Glycerol (sterile) was added to 10% (v/v), the cell suspension was aliquoted and fast frozen on dry ice.

## 2.7 *C. elegans* culture

Most strains and transgenic lines were maintained at 20°C on NGM agar plates (9 cm, 5.5 cm or 3.5 cm diameter plates (Greiner)) seeded with OP50 as a food source, according to standard procedures (Sulston and Hodgkin, 1988). Starved plates (i.e. those devoid of bacteria and containing dauer larvae) were stored at 15°C. Some strains are temperature-sensitive, and these were maintained at 15°C. Worms were picked onto fresh plates when necessary.

2.7.1 *C. elegans* strains

Strain	Allele	Source
N2	wild type	CGC
CB769	<i>bli-1(e769)</i>	CGC
CB768	<i>bli-2(e768)</i>	CGC
MT1144	<i>bli-3(n529)</i>	CGC
CB767	<i>bli-3(e767)</i>	CGC
CB937	<i>bli-4(e937)</i>	CGC
CB518	<i>bli-5(e518)</i>	CGC
BE38	<i>dpy-2(sc38)</i>	CGC
CB1359	<i>dpy-2(e1359)</i>	CGC
CB489	<i>dpy-2(e489)</i>	CGC
CB93	<i>dpy-2(38)</i>	CGC
CB27	<i>dpy-3(e27)</i>	CGC
CB1166	<i>dpy-4(e1166)</i>	CGC
CB61	<i>dpy-5(e61)</i>	CGC
CB14	<i>dpy-6(e14)</i>	CGC
CB88	<i>dpy-7(e88)</i>	CGC
MQ375	<i>dpy-7(qm36)</i>	*
CB1281	<i>dpy-8(e1281)</i>	CGC
CB130	<i>dpy-8(e130)</i>	CGC
BE44	<i>dpy-8(sc44)</i>	CGC
CB12	<i>dpy-9(e12)</i>	CGC
CB128	<i>dpy-10(e128)</i>	CGC
TN64	<i>dpy-10(cn64)</i>	CGC
CB224	<i>dpy-10(e224)</i>	CGC
CB1180	<i>dpy-11(e1180)</i>	CGC
CB458	<i>dpy-13(e458)</i>	CGC
CB188	<i>dpy-14(e188)</i>	CGC
CB164	<i>dpy-17(e164)</i>	CGC
CB364	<i>dpy-18(e364)</i>	CGC
JK2729	<i>dpy-18(ok162)</i>	CGC
CB185	<i>lon-1(e185)</i>	CGC
CB678	<i>lon-2(e678)</i>	CGC
MT3847	<i>lon-2(n1630)</i>	CGC
CB4123	<i>lon-3(e2175)</i>	CGC
DR518	<i>rol-6(su1006)</i>	CGC
CB502	<i>sma-2(e502)</i>	CGC
CB491	<i>sma-3(e491)</i>	CGC
BE1	<i>sqt-1(sc1)</i>	CGC
CB1350	<i>sqt-1(e1350)</i>	CGC
BE13	<i>sqt-1(sc13)</i>	CGC
BE108	<i>sqt-2(sc108)</i>	CGC
BE3	<i>sqt-2(sc3)</i>	CGC
BE63	<i>sqt-3(sc63)</i>	CGC
CB4121	<i>sqt-3(e2117)</i>	CGC

\* gift from Jonathan Ewbank (Centre d'Immunologie de Marseille-Luminy, Marseille, France) and Siegfried Hekimi (McGill University, Montreal, PQ, Canada).

**Table 2.3:** *C. elegans* strains

## 2.7.2 Long term storage of nematodes

Large populations of *C. elegans* strains were washed from NGM agar plates into 15 ml tubes using 1X M9 buffer. Worms were allowed to settle and form a pellet and were washed thoroughly in the same buffer (3 x 15 ml washes). Worms were transferred to microfuge tubes in 400 µl 1X M9 and then a similar volume of freezing solution was added. After mixing well, tubes were slowly frozen at -80°C using an insulated container. Worms were stored at -80°C or in liquid nitrogen. Worm cultures could be recovered by thawing tubes and pipetting worms onto an OP50 NGM agar plate and incubating at 20°C.

## 2.8 Agarose gel

Linear fragments of nucleic acids were separated according to size using agarose gel electrophoresis. Agarose (Invitrogen) was dissolved in 1x TBE buffer to form gels with w/v ratios ranging from 0.6% to 2%. Samples were loaded into gel wells with DNA sample buffer at a final 1X concentration alongside a suitable DNA size standard (1kb ladder or 50kb ladder (both from GibcoBRL) and gels were run in 1X TBE-filled electrophoresis tanks from GibcoBRL with typical voltages of 100-150V.

DNA was visualised by soaking gels in a bath of 0.2 µg/ml of ethidium bromide in H<sub>2</sub>O for 20 minutes (ethidium bromide, which is visible under UV light, chelates to the major groove of nucleic acids). A UV transilluminator (Biorad) was then used to visualise images which were then captured using a QualityOne program (Biorad). Alternatively, ethidium bromide could be added to the molten agarose (2 µl/100 ml).

## 2.9 Purification and synthesis of genomic DNA and cDNA

### 2.9.1 Phenol:chloroform extraction and ethanol precipitation of DNA and RNA

High purity DNA or RNA was obtained using phenol:chloroform extraction. This involved the addition of 1X volume of phenol:chloroform:isoamyl alcohol (24:25:1) (Sigma-Aldrich) to a DNA/RNA solution. It was important that for DNA, phenol at pH 8.0 was utilised while phenol at pH 4.5 was used for RNA purification. The mixture was vortexed for 1 minute and then spun at 16,000 RCF in a bench top centrifuge for 2 minutes. The top layer was removed to a fresh tube and an equal volume of chloroform:isoamyl alcohol (24:1) added to it. This was vortexed, spun and separated as the previous step. DNA/RNA was precipitated using 100% ice-cold ethanol (2 volumes) and either 1/25<sup>th</sup> volume of 5 M NaCl or 1/10<sup>th</sup> volume of 3 M Sodium acetate (C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>Na) (pH 6.0). This mixture was stored for 15 minutes at -80°C and

then pelleted in a bench top centrifuge at full speed for 30 minutes. The pellet was washed using 1 ml of 75% ice-cold ethanol (spun immediately for 10 minutes and then discarded). The DNA/RNA pellet was air dried and resuspended in an appropriate volume of TE buffer.

## 2.9.2 Genomic DNA isolation

Mixed populations of the N2 strain of *C. elegans* were washed off NGM agarose OP50 plates using 1X M9 buffer. After multiple washes with the same buffer, the worms were pelleted by centrifugation and 6 volumes of 1X worm lysis buffer was added. Worm suspensions were disrupted in a glass hand-held homogeniser and subsequently incubated at 65°C for 4 hours. Debris was removed by centrifugation and multiple phenol:chloroform (pH 8.0) and chloroform extractions, followed by an ethanol precipitation. Purified DNA, resuspended in 500 µl of TE buffer was treated with RNase-It™ Ribonuclease Cocktail (Stratagene) (a cocktail of RNase A and RNase T at a final concentration of 80 µg/ml (RNaseA) and 4 Units/ml (RNase T)) for 1 hour at 37°C. Finally, DNA was purified by phenol:chloroform extraction, chloroform extraction and ethanol precipitation and the resulting pellet of genomic DNA was resuspended in 150 µl of TE buffer. Throughout the process of genomic DNA extraction, extreme care was taken to avoid the shearing of DNA. Thus, vortexing or vigorous pipetting was avoided. 100-500 ng of genomic DNA was used in 50µl PCR reactions.

5X worm lysis buffer: 0.25 M Tris-HCl (pH 8.0), 0.5 M NaCl, 0.25 M EDTA.  
pH to 8.0, autoclave and store at 4°C.

1X worm lysis buffer: The above 5X stock solution was used to make 1X solution freshly to result in the final concentrations:  
50 mM Tris-HCl (pH 8.0), 100mM NaCl, 50 mM EDTA, 1% SDS, 30mM 2-mercaptoethanol, 100 µg/ml Proteinase K (Roche).

## 2.9.3 Production of mixed stage cDNA

Production of cDNA required that mRNA be isolated from cultures, purified, and then used as a template for synthesis of first strand cDNA.

### 2.9.3.1 Total RNA isolation

100-1000  $\mu$ l of mixed populations of *C. elegans* were pelleted and then lysed in a 4 x volume of Trizol reagent (GibcoBRL Life Technologies) using vigorous intermittent vortexing over a 20-minute period (in a fume hood). Insoluble material was removed by centrifugation (at 16,000 RCF on a bench top centrifuge (Avanti™ 30- Beckman) for 10 minutes at 4°C ) and the resulting supernatant was subjected to purification using 2 x volumes of chloroform:isoamyl alcohol (24:1) (vortex 15 minutes, incubate at room temperature for 2 minutes and then centrifuge 14,000 RCF for 15 minutes at 4°C). Isopropanol was used to precipitate the RNA from the aqueous phase by mixing, incubating at room temperature for 10 minutes and centrifugation at 16,000 RCF for 10 minutes to form a pellet. The RNA pellet was subsequently washed in 75% (v/v) ethanol (which had been made up in DEPC-treated H<sub>2</sub>O), pelleted by centrifugation at 7,500 RCF for 10 minutes (4°C), and resuspended in 200  $\mu$ l of DEPC-treated H<sub>2</sub>O.

### 2.9.3.2 Purification of mRNA

A Poly (A) Quik® mRNA isolation kit (Stratagene) column kit was used to purify the total mRNA from the total RNA isolated above. Supplied oligo (dT) columns accommodate up to 500  $\mu$ g of total RNA in volumes ranging from 200  $\mu$ l to 1 ml and therefore the concentration of the purified total RNA was calculated and the sample diluted and aliquoted accordingly. Procedures, as detailed in the manufacturer's instructions, were followed with the final pellet being resuspended in 15  $\mu$ l of DEPC-treated H<sub>2</sub>O. mRNA was quantified using a spectrophotometer.

### 2.9.3.3 Synthesis of first strand cDNA

A heterogeneous population of cDNA molecules was generated via first-strand synthesis. In this study, the Stratagene RT-PCR Kit was utilised and the manufacturer's instructions followed. In order to use the amounts of mRNA required in each reaction (50-100 ng), the samples from above were aliquoted into an appropriate number of reactions. The primers utilised were the oligo(dT)primers.

## 2.10 Polymerase chain reaction (PCR)-based cloning procedures

### 2.10.1 PCR conditions and polymerases

All PCR reactions were carried out using a Robocycler Gradient 96 PCR machine (Stratagene) in 50  $\mu$ l volumes, unless otherwise stated. Conditions of the program were altered according to

the requirements of the PCR product and were altered in single reactions in order to optimise the purity and efficiency. In terms of the programme cycle, an initial denaturing step was generally carried out for 5 minutes at 95°C. The cycling program typically consisted of 30 cycles with: denaturation step being 92-95°C for 1 minute, annealing step being 52-56°C for 1 minute and elongation step at 72°C. The length of the elongation step was adjusted according to the length (in bp) of the amplified region and the identity of the DNA polymerase used: *Taq* polymerase (Applied Biosystems) has an activity of approximately 2 kb/minute while *Pfu* (Stratagene) and *Vent* (New England Biolabs) polymerases have extension times of roughly 0.5 kb/minute.

Each of these polymerases were used with suitable buffers at a final concentration of 1X in the 50 µl reaction: *Pfu* (used at 2.5-5 unit of enzyme per reaction) was used with the supplied *Pfu* buffer (10X) while *Vent* (used at 2-4 units per reaction) was used with the supplied Thermopol buffer (10X). Reactions in which *Taq* was the polymerase, used an 11.1X buffer (detailed below) at a final 1X concentration. 5 units of *Taq* were typically used per reaction. In some cases, the reaction was supplemented with additional MgCl<sub>2</sub> (increasing the final concentration in 0.2 mM increments) in order to increase the enzyme efficiency. 50 µl reactions also contained: 100 ng of each oligonucleotide primer (forward and reverse) (purchased from Invitrogen), 100-500 ng of DNA and 90-250 µM of each dNTP (New England Biolabs). The exact PCR conditions and primers used are detailed specifically where appropriate. *Pfu* and *Vent* were utilised when the highest accuracy of amplification was required since these enzymes can proof-read for nucleotide mis-incorporations using their 3' to 5'-exonuclease activity.

11.1X buffer: 45 mM Tris HCl (pH 8.8), 11 mM Ammonium sulphate, 4.5 mM MgCl<sub>2</sub> 6.7 mM 2-mercaptoethanol, 4.4 µM EDTA, 1 mM each dNTP and 113 µg/ml BSA.

### 2.10.2 Purification of PCR products

The PCR products resulting from the above reactions were run on agarose gels and, to separate any unwanted amplification products and primers, the band of the required size was excised. A QIAquick Gel extraction kit was used, following the instructions as directed for using a mini spin column and microcentrifuge. 2 consecutive elutions were carried out in which 15 µl of dH<sub>2</sub>O, that had been incubated at 50°C, was loaded onto the mini spin column and allowed to stand for 1 minute before being centrifuged. The two elution fractions, containing the purified

DNA were pooled. This protocol was also used to separate out DNA fragments from restriction digests (section 2.15).

## 2.11 DNA transfer vectors

All gel purified PCR products were first cloned into a transfer vector before being inserted into the vector of specific function. This method was used because the efficiency of restriction digestion is reduced when sequences flanking the recognition sites are shorter in length. Thus, in some cases, the efficiency of restriction digestion of pure PCR products would be low. The use of transfer vectors is also useful because they provide a means of ligating a DNA insert in to a vector, that is generally much more efficient than many other vectors in terms of ligation. Thus, instead of repeating the PCR reaction as a means of producing more product, the original PCR product can be efficiently ligated into these vectors and can be replicated via transformation into bacteria.

The two transfer vectors utilised were pCRScript (Stratagene) and TOPO T-A (Invitrogen). These differ because the former requires inserted DNA to have blunt ends (for insertion into an *Srf* I site) while the latter requires A-overhangs. *Taq* and other low fidelity polymerases generate T-overhangs so were cloned into TOPO vectors immediately. However, polishing the ends was required in order to prepare them for insertion in to pCRScript. Conversely, fragments generated by *Pfu* and *Vent* are blunt ended and therefore, it was necessary for A-overhangs to be generated for their insertion into TOPO. Protocols for both the polishing of PCR product or the addition of 3' A-overhangs are detailed in the pCRScript and TOPO TA cloning instruction manuals, respectively. Similarly, the methods by which PCR products were ligated into each of these vectors is detailed in the manufacturer's instructions (the only variation is that 0.8  $\mu$ l of TOPO vector was used instead of the stated 1  $\mu$ l).

## 2.12 Transformation of *E. coli*

### 2.12.1 Standard *E. coli* strains

pCRScript ligations were transformed into 40  $\mu$ l of XL10-Gold ultracompetent cells (Stratagene) while TOPO-TA ligations were transformed into 50  $\mu$ l of the supplied TOP10 One Shot® cells (chemically competent). All transformations carried out in this thesis were carried out in the same way, i.e. a heat shock method, unless otherwise stated. Approximately 5-10 ng of plasmid DNA was added to competent cells (on ice). These were left on ice for 30 minutes and then incubated in a water bath at 42°C for exactly 30 seconds. Cells were allowed to

recover for 2 minutes on ice before 400  $\mu$ l of pre-warmed SOC medium was added. Cells were placed in an orbital shaker at 37°C for 1 hour (in which time the antibiotic resistance gene product could be synthesised). Cell suspensions were spread on LB + antibiotic plates, dried and grown overnight at 37°C. For blue/white colour selection, of transformants, plates were supplemented with 100  $\mu$ l of 0.1M IPTG and 100  $\mu$ l of 2% X-Gal.

### 2.12.2 HT115(DE3) strains

The same method as described above was used to transform plasmid DNA into HT115(DE3) cells (which are tetracyclin resistant), the strain used in the RNAi feeding method. However, the heat shock step was carried out at 37°C for 1 minute.

## 2.13 Identification of bacterial transformants

In order to identify which transformants contained the desired insert, blue/white screening, PCR screening, or restriction digestion was utilised.

### 2.13.1 Blue/white selection

Plasmids designed for blue/white selection (such as pCRScript and TOPO) greatly facilitate the identification of positive colonies from transformations. The blue/white selection stems from the multiple cloning site of the plasmid being located within a peptide (*lacZ*) that encodes for the N-terminal region of  $\beta$ -galactosidase, an enzyme that metabolises X-Gal (5-bromo-4-chloro-3-indolyl- $[\beta]$ -D-galactopyranoside) from a colourless to a blue product. In the presence of an insert within the multiple cloning site which interrupts the *lacZ* gene, a functional  $\beta$ -galactosidase is unable to form which removes the ability to mediate a colour change. Consequently, colonies containing successfully ligated plasmids appear white while blue colonies imply the failed ligation into the multiple cloning site of the transformed vector. It must be taken into account that a white colour is not a guarantee that there is an insertion or that, if there is an insertion, that it is the sequence of interest. Thus, this screening method should be used in conjunction with the PCR screening or with a restriction digestion analysis using appropriate enzymes (sections 2.13.2 and 2.15).

### 2.13.2 PCR screening

Transformants resulting from ligation into pCRScript, TOPO or other functionally specific vectors were identified as to whether they contained the desired insert using a PCR method with gene specific- or vector specific- primers. Alternatively a combination of one of each could be used which allowed for screening of direction insertions. For each construct, typically 20

colonies were screened. A toothpick was used to pick a colony and streak some of the cells onto a gridded LB + antibiotic agar master plate. The cells remaining on the toothpick were transferred to separate 50  $\mu$ l aliquots of dH<sub>2</sub>O and lysed by boiling for 5 minutes. 5  $\mu$ l of the resulting cell lysate was then added to a 50  $\mu$ l PCR reaction mixture containing 1X PCR buffer (from 11.1X stock detailed above), 1 unit of *Taq* polymerase and 100 ng of a forward and reverse primer. A simple PCR program was used: 92°C 1 minute, 56°C 1 minute, 72°C 1 minute (30 cycles) and the products run on an agarose gel but, as stated previously, longer extension times were used if the fragment was particularly long. Colonies amplifying fragments of the expected size were assumed to be positive transformants and DNA preparations were made from them by returning to the gridded master plate to start L-broth cultures.

#### **2.14 Plasmid DNA preparation**

The QIAGEN mini-spin columns and Midi purification techniques are based on the alkaline lysis of cells in L-broth (+ antibiotic) cultures (5 ml for mini preps and 100 ml for Midi preps). The protocols supplied by the manufacture were used apart from an adjustment to the Midi prep protocol that introduced an additional DNA precipitation step subsequent to precipitation of DNA with isopropanol. Instead of proceeding immediately with the 70% ethanol wash, the isopropanol-derived pellet was resuspended in 200  $\mu$ l TE buffer, ethanol precipitated (as detailed in section 2.9.1) and washed in 70% ethanol. The dried pellet was resuspended in 200  $\mu$ l of TE buffer. Mini preps were eluted in 100  $\mu$ l elution buffer. MIDI preparation of DNA was used when higher purity of DNA was required. Such preparation also results in higher yields of DNA.

#### **2.15 Restriction endonuclease digestion**

Restriction enzyme digestion was used to verify the inclusion of an insert within a plasmid. However, the methods described here are relevant for all the restriction digestion reactions carried out in this thesis apart from where otherwise stated. For analyses, 200-500 ng of plasmid DNA and approximately 2-5 units of restriction enzyme were used in 50  $\mu$ l reaction volumes of 1X buffer supplied with the enzyme. Many enzymes are compatible with each other and could be used simultaneously in the same reaction mixture, even if it entailed using one of the enzymes at sub-optimal conditions (c.g. 50-75% efficiency). However, when this was not possible, it was necessary to digest using the two enzymes in succession, using one enzyme (2 hours 37°C and then heat inactivate it) and ethanol precipitating the linear DNA

before setting up the restriction digest with the second enzyme in its appropriate buffer. Incubations were typically carried out at 37 °C for 2 hours unless another temperature was specified for the enzyme as recommended by the supplier.

These reactions were scaled up for the production of large amounts of linear fragments required, for example, for ligation reactions (described below). In such cases 5 µg of DNA were used in similar 50 µl reactions. Even larger quantities of DNA could be digested by scaling up the volume of the reaction and using MIDI prep DNA. For purification of specific products of restriction digestion, the reaction mixture was run on an agarose gel (of an appropriate percentage), the correct sized fragment was excised, and the DNA was extracted and purified as described in section 2.10.2.

### 2.16 5' end dephosphorylation of linearised vector DNA

In cases where restriction digestion of a vector had left blunt ends or two similar ends, prior to ligation, it was necessary to prevent recircularisation of the plasmid (without an insert) by dephosphorylation of the 5' end. 10-20 units of calf alkaline phosphatase (CIP) (New England Biolabs) were used in a 50 µl reaction volume that also contained 5 µg of linearised (and gel purified) DNA and any of the New England Biolab standard buffers (buffers 1-4) at a 1X concentration. Reactions were incubated at 37°C for 1 hour, after which the dephosphorylated DNA was purified using the standard phenol:chloroform extraction method (including the chloroform purification and ethanol precipitation steps).

### 2.17 Ligations into other vectors

Plasmid preparations of: the required vector and a (specific function) vector containing the required insert, were restriction digested in separate reactions using similar or compatible restriction enzymes. Typically, approximately 1 µg and 5 µg of vector and insert-containing plasmid DNA respectively were digested. The products were run on an appropriate percentage agarose gel and the correct bands excised and gel extracted. The ideal molar ratio of insert:vector for cloning is considered to be 3:1 and thus, after determining the concentrations of linearised DNA by the running 2 µl of gel extracted DNA on an agarose gel against a DNA ladder of known concentration, the relative amounts of each component required was calculated via the following equation:

$$\frac{\text{ng vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \frac{3}{1} = \text{ng of insert}$$

Typically 100 ng of vector was used in a 10 $\mu$ l ligation reaction that also included T4 DNA ligase (New England Biolabs) and its supplied T4 5X Buffer (added to a 1X concentration). Ligations were incubated at 15°C overnight after which 2  $\mu$ l of the mix was used to transform XL10 *E. coli* cells as detailed in section 2.12.1. Prior to transformation, it was, in some cases, possible to carry out a restriction digestion step (37°C 30-60 minutes) using an enzyme specific for vector sequences that had been incompletely digested and as a consequence had resealed back to its original form (i.e. not containing the desired insert). To optimise the efficiency of this step, the ligation mixture (in 1X ligase buffer), could be supplemented in order to recreate the conditions of the supplied restriction digestion buffer.

### 2.18 Sequence analysis

The sequence of inserted DNA was verified by sequence analysis. Ethanol precipitated DNA samples (2  $\mu$ g) were sent to MWG sequencing along with appropriate oligonucleotide primers. In some cases, primers were required that ensured the sequencing of the entire cloned sequence. However, in other circumstances, the cloning junctions (ligation sites) were the only segments that were required to be sequenced. Returned sequences were analysed using Vector NTi (Informax) and its associated program, Contig Express, and compared against published gene sequences in WormBase (<http://www.wormbase.org>).

### 2.19 Bioinformatic resources

The aforementioned WormBase site was an invaluable tool since it is a central resource of sequence-, RNAi-, expression-, family- and functional- data. The “motif search” was also a key tool in the construction of the list of RNAi targets described in chapter 4. Other bioinformatics tools utilised include the ClustalW tool from ‘Pole Bioinformatique Lyonnais Network of protein sequence analysis (NPS)’ (<http://npsa-pbil.ibcp.fr>). The prediction of ER signal peptide cleave site of proteins (c.g. HPX-1 in chapter 5) was via the Signal P program through the ExpASy molecular biology server (<http://www.expasy.ch>). Secondary structure prediction of HPX-1 was performed by a Predict Protein tool through the EMBL Heidelberg server (<http://www.embl-heidelberg.de>).

### 2.20 Semi quantitative reverse transcriptase PCR (SQ RT-PCR)

A semi quantitative-reverse transcriptase PCR (SQ RT-PCR) approach was used in order to examine the temporal expression patterns of *col-19* and *hpx-1*. This method uses the premise that *ama-1*, that encodes for a subunit of RNA polymerase II, is constitutively expressed throughout the larval and adult stages and that the expression level of a gene of interest can be

thus be expressed in relative terms of *ama-1*'s constant level of expression. A set of staged cDNA samples representative of the mRNA population at every two hours of development (a gift from Iain Johnstone, WCMP, University of Glasgow, UK) (Johnstone and Barry, 1996) was used to amplify up *ama-1* and the gene of interest using the sets of primers listed in table 2.4. It was only necessary that short fragments of each be amplified. However, it was important that the region between the primers contained an intron in genomic DNA that would not be present in the cDNA. This would enable that products amplified from any contaminating genomic DNA in the cDNA fraction, whose levels would not alter temporally, could be separated from products amplified from cDNA.

The conditions of the PCR reaction are based on the protocol described in Johnstone and Barry, 1996 where the importance of not exhausting the reagents was emphasised. For each stage represented by a cDNA sample, a reaction was set up containing each of the four primers (a forward and reverse for both *ama-1* and the gene of interest) in excess (100 pmoles per reaction), appropriate amounts of the 11.1X buffer to give a final 1X concentration, 1-3  $\mu$ l of the mixed stage cDNA sample and 1  $\mu$ l of *Taq* polymerase. Refer to tables 2.4 and 2.7 for details of primers. 30 PCR cycles of 92 °C 1 minute, 56°C 1 minute and 72°C 1 minute were used. 25  $\mu$ l of each reaction mixture was run on a 200 ml 1% agarose gel. After soaking the gel in a solution of ethidium bromide, the DNA fragments were visualised on a UV transilluminator. In the *hpx-1* RT PCR experiments, the image from the transilluminator was saved as a TIFF file which was analysed using ImageQuant software (Amersham Bioscience). However, for the *col-19* RT PCR, a southern analysis method was utilised. For this, the gel was first trimmed to leave the region containing the bands of interest (at this stage it was especially useful to remove the bands corresponding to the primers).

SQ RT PCR experiment	Primers and sequence	cDNA fragment (bp)	Genomic DNA fragment (bp)
<i>col-19</i>	ama 1a2 (cagtggctcatgtcgagttccaga) ama 1b2 (gcgctggaaattgacagattgc)	426	523

SQ RT PCR experiment	Primers and sequence	cDNA fragment (bp)	Genomic DNA fragment (bp)
	semiQ PCR col-19 fwd (atgggcaagctcattgtg) semiQ PCR col-19 rev (agctgggcaaccctcgga)	300	353
<i>hpx-1</i>	ama-1 (1228/744) fwd (gcgctggaaattgacagattgc) ama-1 (1228/744) rev gaatacaattgtccacaagagttgc)	744	1228
	zk430.8 (1831/853) rev (gctctgtgagcttctgagc) zk430.8 (1831/853) fwd (gtccatctcccgactaatcaacc)	853	1831

**Table 2.4: Primers utilised for RT-PCR and the amplified fragment sizes.** Sequences of primers are detailed in table 2.7.

## 2.21 Southern blotting

DNA was transferred to a nylon membrane (Hybond™ N, Amersham Life Science) by first treating the DNA by soaking the gel in the following solutions for 15 minutes each: 0.25 M HCl, denaturing buffer (0.4 M NaOH, 0.8 M NaCl) and finally neutralising buffer (0.5 M Tris, 1.5 M NaCl, pH 7.6). The transfer apparatus was set up with the following layers: 4 wicks cut from 3MM paper soaked in 10X SSC (see below for 20X SSC); denatured gel; nylon membrane; 6 filters cut from 3MM paper (not soaked). The nylon membrane and 3MM had been cut to be the exact size of the gel. Cling film was placed closely around the sides of the gel in order to prevent the transfer of fluid via any route other than through the gel and membrane. A stack of paper towels was placed on top of the blot and weighed down with a heavy object and it was left to blot overnight at room temperature. The nylon membrane was removed from the above blot and soaked briefly in 3X SSC. It was then air dried and the DNA cross-linked using a UV Stratalinker 2400 (Stratagene).

20X SSC:

3 M NaCl, 0.3 M NaCitate, pH 7.5

### 2.21.1 Probing the membrane

The transferred DNA was visualised by hybridisation of the membrane using radioactive probes. These were prepared using a Prime-It® II Random Primer labelling Kit (Stratagene). For synthesis of each probe (*ama-1* and *col-19*), ~100 ng of template DNA, that had digested out of vectors (vectors that had been supplied by Gillian McCormack), was used according to the instructions stated by the manufacturer. A few modifications were made, including that 100 ng of DNA was used as a template (instead of the suggested 25 ng), 3 µl of <sup>32</sup>P αdCTP labelled nucleotide was used and the mix containing the labelled nucleotides was incubated at 37°C for 30 minutes. The probes were purified on microspin S-200 HR columns (Amersham Biosciences) following the instructions of the manufacturer.

The nylon membrane containing the transferred DNA was allowed to prehybridise for 10 minutes in a hybridisation bottle at 65°C in 50 ml of prehybridisation mix (in a rotating hybridisation oven (Stuart Scientific)). Radiolabelled probes were boiled for 5 minutes (95 °C) in order to denature them and 10 µl (approximately 12.5 ng of labelled DNA) of each probe was added to the solution. Hybridisation was allowed to continue overnight at 65°C. The membrane was washed with prewarmed 2X SSC/0.1% SDS in two 10 minute incubations (65°C) and then quickly rinsed in 0.2X SSC. The membrane was allowed to air dry and then exposed to a phosphoimage screen (Fuji) at room temperature for between 1 and 24 hours. The screen was then scanned using a Typhoon 8600 phosphoimager and the results were analysed using ImageQuant software.

Church and Gilbert mix:	100 ml of 1 M NaPO <sub>4</sub> (71 g anhydrous Na <sub>2</sub> HPO <sub>4</sub> with 4 mls of 85% orthophosphoric acid in 1 L of H <sub>2</sub> O), 70 ml 20% SDS, 0.4 ml 0.5M EDTA (pH 8), 26.6 H <sub>2</sub> O.
Prehybridisation mix:	50 ml Church and Gilbert buffer with 200-500 µg herring sperm DNA (denatured by incubation at 95°C for 5 minutes).

## 2.22 Standard *C. elegans* techniques

### 2.22.1 Generation of males

Males arise as a consequence of spontaneous X chromosome non-disjunction during meiosis that results in animals having a nullo-X genotype. In wild type populations, this occurs at a

frequency of approximately 0.2%. To increase this frequency, a heat shock method was used. This involved incubating a number of L4 hermaphrodites (N2 or TP12) on NGM agar OP50 plates at 30°C for 8 hours. In order to maintain high populations of males, plates were continually set up containing 6 L4 males (from the F1 population of the heat-shocked worms) and 2 L4 hermaphrodites from the same strain.

### 2.22.2 Crosses

To obtain morphologically mutant strains expressing the COL-19::GFP construct, 6 TP12 males were placed on a plate with 2 hermaphrodites of the morphologically mutant strain. Because the COL-19::GFP is a dominant trait the F1 generation derived from a cross could be identified based on the expression of the fluorescent marker. 6 such GFP positive F1 hermaphrodites were placed onto a fresh plate with 2 TP12 males. Since most of the lesions of the morphological mutants are recessive, successfully transformed animals could be selected for based on the expression of the GFP marker and the mutant phenotype in the progeny of this cross ('F2' generation). Such single animals were picked onto individual plates (1 worm per plate) and allowed to self-fertilise. The progeny resulting from this self-fertilisation ('F3' generation) were checked for 100% of the progeny being morphological mutant. This was especially necessary if the morphological mutant lesion was dominant. 100% of the progeny from a self-fertilisation represented that a stable line had been created.

### 2.22.3 Synchronising populations

Large populations of worms were grown on at least five 9 cm NGM agar OP50 plates. It was important that these populations contain a large percentage of adult animals. Such populations were washed into falcon tubes (Grenier) (on ice) using 1X M9 buffer and washed with the same buffer 3 times (each time, allowing the worms to pellet by gravity, removing the supernatant and refilling the tube with fresh buffer). Subsequent to washing, 1ml of bleaching solution (1% (v/v) bleach, 0.25 M KOH) was added to each 100 µl of packed worms. Bleaching was allowed to continue for only 3-5 minutes and immediately washed with 1X M9. The bleached worms were washed 3 times with 1X M9 and then were transferred to a conical flask in 20 ml final volume of 1X M9. The bleaching process killed all animals but the unhatched embryos were still viable. Consequently, after being left overnight at 25°C, the embryos released from dead adults were able to hatch but were arrested at the L1 larval stage due to the absence of food. In order to reset the life cycle, aliquots of L1 arrested larvae could be transferred onto NGM agar plates supplemented with OP50 bacteria. For synchronous populations of L1 larvae,

animals could be washed off plates after 2 hours of feeding. For L3 populations, plates were incubated at 20°C overnight before being washed off. Such a method was used for transgenic lines carrying the *hpx-1* promoter-LacZ construct and the *hpx-1*-PEST construct. After the L3 stage, synchronous populations generated in such a manner lost some synchrony. An alternative method to obtain synchronous populations of later larval stages was used that involved allowing populations of animals to starve and revert to the essentially, arrested L3, dauer stage. Re-establishment of a food source enabled these populations of worms to re-enter the life cycle as a synchronous population of L4s, which could be collected. At 25°C it took approximately 8 hours for dauer animals to start feeding again and a further 4 hours to reach the dauer-L4 moult while it took longer at 15°C and 20°C. In this way, synchronous populations of adults could also be obtained. This was the method used to observe the COL-19::GFP pattern as adults aged. The onset of the adult stage could be identified by the expression of the adult-specific COL-19::GFP marker.

#### 2.22.4 Transformation of *C. elegans*

#### 2.22.5 Microinjection

Transformed lines of *C. elegans* were obtained via microinjection following a method described by C. Mello and A. Fire, 1995 that details the use of co-injection of experimental DNA and a selectable marker. Both plasmid DNA constructs were based on the pPD series of vectors (supplied from the Fire lab). The selectable marker used was plasmid p76-16B (a gift from R. Horvitz, MIT, Massachusetts, USA) that rescues the Unc phenotype of the *unc-76* mutant strain DR96 (*unc-76(e911)*) because it encodes a wild-type copy of the gene (Bloom and Horvitz, 1997). DNA solutions were prepared using a standard Midi preparative (QIAGEN) method and mixed to give a final concentration of 20 µg/ml of experimental plasmid and 100 µg/ml of rescue plasmid (in dH<sub>2</sub>O). The mixture was centrifuged for 30 minutes to remove any debris which might block the needle and the supernatant from this step centrifuged for a further 15 minutes as an extra precaution. DNA was loaded into needles that had been pulled on a computer-controlled electrode puller model 773 (Campden Instruments) using 1.2 mm O.D (outside diameter) x 0.69 mm I.D. (inside diameter) borosilicate glass capillaries with standard wall and inner filament (Clark Electromedical Instruments). Young adult nematodes were immobilised under a drop of liquid paraffin (BDH) on 2% (w/v) agarose pads (on 64 mm x 22 mm coverslips (Thickness 1, BDH)) that had been flattened with a second coverslip and subsequently thoroughly dried at 80°C. The gonads of hermaphrodites were injected using an

Axiovert-100 inverted microscope (Zeiss) equipped with a flat, free-sliding glide stage with centred rotation with DIC/Nomarski optics, a micromanipulator guided needle, and pressurised nitrogen that was controlled via a foot pedal. In order to obtain transformed lines, it was necessary that the DNA be injected into the gonad arms. Recovery buffer (2% glucose, 1X M9) was used to remove worms from pads and a large pool of it was placed on a fresh NGM OP50 plate which provided an environment for the worms to recover in (at 20°C). Transformed progeny were identified from the F1 generation of injected animals by the *Unc* rescue phenotype (essentially wild type). Plates containing 10 rescue phenotype worms from the F1 generation were prepared and the resulting F2 generation analysed.

#### **2.22.5.1 Behaviour of injected DNA**

Formation of stable transformed lines occurs because it is possible that in 1-10% of the F1 population, the two pPD vectors can, due to regions of homology within the vector backbones, recombine to form long tandem arrays. Arrays that attain a large size can become heritable as extrachromosomal elements which are transmitted at a frequency of between 5% and 95% in subsequent generations.

#### **2.22.6 Maintenance of transgenic strains carrying a free array**

Lines of transgenic animals were maintained by picking approximately 10 rescued hermaphrodites onto fresh plates.

#### **2.22.7 TP12 integration of extrachromosomal arrays**

Integrated lines carrying the COL-19::GFP construct were obtained by subjecting a number of lines expressing the COL-19::GFP extrachromosomal array (populations of >200 L4 animals on 4 x 9 cm NGM agar OP50 plates) to irradiation (38 Gy) from a <sup>60</sup>Co source (25 minute cycles). For each line, worms were allowed to recover for approximately 2 hours before being separated onto 24 individual plates containing 7 irradiated young adults each. These were allowed to grow for 10-14 days, until starved, after which a small section of the NGM agar was excised and transferred onto fresh plates. From each of the 24 plates, 20 GFP expressing worms, escaping from the dauer stage, were picked onto individual (3.5 cm) NGM agar plates. Those plates with 100% transgenics in the F1 generation were chosen and from these, individual animals were picked and allowed to self. The generations produced as a result of the self cross were checked in order to verify 100% expression of the integrated construct. The best plates were picked for each transgenic line and maintained. Animals from these lines were backcrossed 4 times against wild type nematodes to remove deleterious mutations. This involved over the course of

4 generations, taking animals derived from the integration and setting up a cross (as described in section 2.22.2) with N2 males. A single worm PCR was performed in order to verify the presence of the correctly integrated DNA.

### 2.23 Single Worm PCR

Single worms were transferred to 2.5  $\mu$ l of single worm lysis buffer, placed on dry ice for > 15 minutes and then immediately incubated at 65°C for 90 minutes. Finally, they were incubated at 90°C for 20 minutes and allowed to cool. 22.5  $\mu$ l of premix was added to each reaction and used in the following PCR programme: 30 cycles at 94°C 30 seconds, 58°C 1 minute and 72°C 1 minute. Primers used for verification of the presence of the COL-19::GFP construct were 'col-19 fwd' and 'gfpc3rev' as described in table 2.7.

Lysis buffer: 100  $\mu$ g/ml proteinaseK, 10mM Tris (pH8), 50mM KCl, 2.5mM MgCl<sub>2</sub>, 0.45% (v/v) Tween@20 (Sigma-Aldrich), 0.05% (w/v) gelatin.

Premix: 25-30 pmol of each primer, buffer (11.1X at a final concentration of 1X), 2.5 units of *Taq* polymerase.

### 2.24 Microscopy

Nematodes were viewed using an Axioskop 2 microscope (Carl Zeiss, Oberkochen, FRG) with lenses: 10x/0.3 Plan-Neurofluor and 63x/1.4 and 100x/1.4 Plan-Apochromat. GFP fluorescence was viewed with a mercury lamp (240V 100W, Osram) UV source and filter set 09. A Hamamatsu C4742-95 monochrome digital camera (Hamamatsu photonics, Hamamatsu City, Japan) and Openlab 2.2.0. software were used for image capture. Images were processed and pseudocoloured, where appropriate, using Improvision Openlab software. This apparatus was used for visualisation of live- and fixed- specimen (the latter of which is described in sections 2.29.2).

### 2.25 Preparation of whole worms

2% agarose + 0.1% sodium azide pads, which had been flattened with a coverslip, were made on glass twin frosted microscope slides. Animals were picked onto pads which had been supplemented with a small amount of 1X M9. A cover slip, whose edges had been coated with a light film of soft paraffin BP, was carefully placed on the pad. The paraffin acted as a seal. Prior to visualisation, slides were kept in a humid chamber at 4°C and in the dark. However, best results were obtained when samples were analysed immediately.

## 2.26 Preparation of freeze-cracked specimens for Immunolocalisation

Specimen for immunolocalisation assays were prepared using a freeze cracking method. Slides were prepared by the application of 2 sequential coats (each approximately 100  $\mu$ l spread over the whole surface) of poly-L-lysine (0.1% w/v in H<sub>2</sub>O) (Sigma diagnostics, MO, USA) and incubated at 80°C until dry. Poly-L-lysine is an adhesive solution for adhering tissue sections to glass slides. Large populations of worms, prepared by washing either mixed or synchronous populations from NGM OP50 plates, were washed three times with 1X M9 buffer and finally allowed to settle and form a loose pellet. Enough buffer was left around the pellet to enable that worms be pipetted using a glass Pasteur. Approximately 20  $\mu$ l of worm pellet was spread onto the dried poly-L-lysine slides. A 64 mm x 22 mm glass coverslip was placed on top. Slides were immediately placed at -80°C for 20 minutes. Freeze cracking was then carried out using a scalpel blade to prise the coverslip off. Slides were then transferred to ice-cold 100% methanol (20 minutes) and then to ice-cold 100 % acetone (20 minutes), after which they were left to air dry at room temperature. Slides could be stored at -20°C.

## 2.27 Immunolocalisation procedure for whole worm staining

Immunolocalisation was carried out in microfuge tubes rather than on slides and therefore using PBST (0.1% Tween), freeze cracked worms, prepared by the above method, were washed from slides into tubes. Worms were allowed to settle by leaving on ice and the supernatant was removed. The first blocking step was carried out in approximately 150  $\mu$ l of 1% milk (Marvell) in PBST or in 1% bovine serum albumin (BSA) (96% electrophoresis grade -Sigma-Aldrich) in PBST for 2 hours at room temperature on a shaker. Blocking solution was removed by allowing a loose pellet to settle and removing the supernatant and then the pellet was then washed 3 times with 1 ml volumes of PBST. The washed pellet was resuspended in 150  $\mu$ l to 200  $\mu$ l of primary antibody (raised against COL-19, MH27, DPY-7 or Ty) that had been diluted to the appropriate concentration in PBST and placed on a shaker for 2 hours at room temperature or at 4°C overnight. The primary antibody was removed in a similar manner from the previous step and the pellet was washed three times with PBST. The pellet was resuspended in a volume of secondary antibody (150  $\mu$ l - 200  $\mu$ l) that had been diluted to the correct concentration in PBST. For DPY-7, MH27 and TY staining, the secondary antibody was AlexaFluor 594 goat anti mouse (Molecular probes). This was left at room temperature for 2 hours after which the pellet was washed 3 times. To put immunostained worms back onto slides, the pellet, left in ~20  $\mu$ l solution with a drop of mount solution (50% glycerol, 0.5X PBS, 2.5% DABCO (Sigma-Aldrich)) added, was pipetted onto a fresh (non poly-L-lysine)

slide using a glass Pasteur pipette. An even and fine spread of animals was achieved by using the long flat side of the pipette. A 64 mm x 22 mm glass coverslip was placed on and the excess liquid removed from underneath by using the capillary action of a piece of tissue placed at the side of the coverslip. When as much liquid as possible had been removed the slide was sealed using clear nail varnish. Slides were stored in humid chambers at 4°C and then viewed via epifluorescence as described in section 2.24.

Antibody	Block	Primary	Secondary
DPY-7	1% milk (PBST)	1:50	1:100 goat anti-mouse
MH27	1% milk (PBST)	1:100	1:100 goat anti-mouse
Anti-Ty	1% milk/1% BSA	1:50	1:200 goat anti-mouse
COL-19	Tried 1%, 2% milk and 2%, 3% BSA	Tried 1:50, 1:100, 1:200, 1:500, 1:5000, 1:10000 and pre-absorbed sera On: -methanol/acetone-fixed worms -methanol/acetone/formaldehyde-fixed worms -cuticle extracts	Tried 1:100, 1:200 goat anti-rabbit

**Table 2.5: Conditions of antibody staining.**

### 2.28 Optimisation of antibody staining using anti COL-19 antibodies

COL-19 antibodies were raised against a 16-residue peptide (APGPDAGYCPCPSRAA) corresponding to the C-terminal, non Gly-X-Y, region of COL-19 by using a standard immunisation protocol in rabbits (Sigma). As described in table 2.5, a number of attempts were made to make the immunolocalisation using this antibody as specific as possible. However no technique proved successful. Initial trials used the standard protocol as detailed above trying primary antibody concentrations of 1:50, 1:100, 1:200, 1:500, 1:1000, 1:5000. There was no middle ground between samples that stained non-specifically and those that exhibited no staining at all.

The next stage attempted to use worms that had been fixed using the “heavy” formaldehyde method described below. Such animals were treated with antibodies in standard fashion. A final means of preparation was to make cuticle extracts. Such cuticle extracts were treated with antibodies according to the standard procedures and were also treated with samples of antibodies that had been pre-absorbed with acetone-extracted preparations from populations of worms that had been treated with *col-19* RNAi. Details of these processes are described below.

### **2.28.1 ‘Heavy fixation’ of freeze-cracked slides for immunolocalisation**

The ‘harder’ fixation protocol, using an additional formaldehyde fixation step, was attempted for antibody staining with COL-19 antibodies. Slides were fixed using the standard methanol and acetone solutions as described above and were subsequently transferred to a chamber containing 4% formaldehyde in PBS for 30 minutes at room temperature (an 8% formaldehyde stock was made by heating and adding drops of NaOH until dissolved). Slides were washed thoroughly 3 times in PBS buffer and then used as normal. This is a modified version of that by Janet Duerr as described in WormAtlas ([www.wormatlas.org](http://www.wormatlas.org)). Worms prepared in such a way were treated with primary antibodies at the same concentrations of COL-19 antibodies as in table 2.5.

### **2.28.2 Cuticle extraction**

Worms from large mixed populations (from 8 x 9 cm diameter NGM plates that had relatively high adult populations) were washed off NGM plates using 1X M9 buffer which was then replaced with 1 ml of TBS + 0.1% NP<sub>4</sub>O and the following proteases at their final concentrations: 100 mM PMSF (serine protease inhibitor), 100 mM EDTA (metalloprotease inhibitor), 0.1 mM Pepstatin (acid protease inhibitor) and 2 mM E64 (cysteine protease inhibitor). A pestle and mortar (surrounded by a polystyrene insulator) were pre-cooled using liquid nitrogen. The pestle was filled with liquid N<sub>2</sub> and the 1 ml mixture of worms and TSB solutions was administered using a glass Pasteur. Worms were ground into a fine powder and then transferred to a 1 ml homogeniser in which grinding was continued (on ice for 10 minutes). The ground worms were incubated on ice for a further 30 minutes. Cuticle extracts were pelleted by centrifugation at 10,000 g in a 1.5 ml microfuge on a Avanti (Beckman) benchtop centrifuge (for 20 minutes).

Cuticle extracts prepared in this manner were subjected to immunolocalisation using standard methods (2% milk (PBST) block and primary antibody concentrations of 1:50, 1: 500, and 1:5000). They were also treated with pre-absorbed antisera (section 2.28.3).

### 2.28.3 Pre-absorption of antisera

The method of pre-absorption of antisera was used with the aim of reducing the background staining pattern when using COL-19 antibodies. Extracts containing all native proteins except for COL-19 were used. These extracts were obtained from large populations of worms that had been grown on *col-19* RNAi plates in order to remove this protein from the pre-absorption mixture. Worms were washed off 10 large NGM *col-19* RNAi feeding plates using 1X M9. The pellet was washed with 0.1 M NaCl and the worms lysed on ice by sonication (40 % amplitude, 3 minutes: pulses of 5 minutes on, then 5 minutes off). 5 volumes of ice-cold acetone were added to the sonicated liquid. A pellet was obtained by centrifugation for 15 minutes and was resuspended in a small amount of 100% acetone. This mixture was dropped onto the centre of a Millipore disc (Millipore size 0.025  $\mu$ M) that had been placed on a piece of 3MM filter paper. This was allowed to dry at 37°C. The region of the membrane to which the protein extract had been applied was cut into (9) pieces to effectively aliquot it. One such aliquot was soaked in 50  $\mu$ l of PBS and agitated at room temperature overnight. For pre-absorption, a 1:3 ratio of volume of packed worms to volume of neat antibody was required and appropriate volumes of antibody were added to the 50  $\mu$ l of protein extract. Protein extract and antibody were allowed to pre-absorb for 2 hours at room temperature (it was then centrifuged for 10 minutes). This pre-absorbed antibody was used on cuticle extracts at concentrations of 1:500, 1:1000, 1:5000, and 1:10000 (blocked with 1% milk).

## 2.29 Promoter analyses

### 2.29.1 The *hpx-1* / *lacZ* promoter/reporter gene construct

In order to characterise the spatial expression pattern of *hpx-1* the reporter gene vector pPD96.04 (A. Fire) was used. This allows *cis*-acting sequences of a gene of interest to be inserted at a 5' cloning site in order to be fused in-frame to the *lacZ* coding region which encodes for the enzyme  $\beta$ -galactosidase ( $\beta$ -gal). Sites of promoter-driven expression are visualised via an X-Gal staining method on the basis that X-Gal is the substrate for  $\beta$ -gal. At sites of reporter expression, X-gal is converted into a clearly visible blue product. This is a commonly used reporter mechanism that can be used for expression assays in many *C. elegans* tissues. A variety of similar pUC19-derived reporters are available from the laboratory of A. Fire.

The pPD96.04 vector used here has an eight amino acid SV40-derived nuclear localisation signal (NLS) at the junction between the inserted promoter and the *lacZ* reporter. This is a useful element to have in a reporter construct because it directs the synthetic transport of the fusion molecule into the nucleus, thereby sequestering the reporter moiety into a limited and punctate organelle and making the identification of the cell in which expression is occurring easier. A number of synthetic introns are included in the vector sequence, one upstream of the NLS and 12 within the *lacZ*, as the presence of introns has been demonstrated to improve the level of expression. The *lacZ* coding sequence is followed by a 3' multiple cloning site containing the *C. elegans unc-54* gene. An additional means of visualisation is provided for in the pPD96.04 vector by the inclusion of a GFP moiety 5' of, and in frame to, the *lacZ* sequence. The efficiency/sensitivity of both the GFP and *lacZ* reporter constructs is increased by the addition of a 5' decoy sequence upstream of the 5' multiple cloning site whose encoded termination codon effectively prevents read-through from background expression.

A 2115 bp putative promoter region of the *hpx-1* gene was amplified (30 cycles) from N2 genomic DNA using Vent polymerase and the primers 'zk430.8 promoter fwd' (*Sph* I) and 'zk430.8 promoter rev' (*Xba* I) (see table 2.7). The amplified sequence includes approximately 2000 bp of 5' promoter sequence as well as protein coding sequence encoding for the ATG start site, the predicted signal peptide, the signal peptide cleavage site and 4 residues which lie 3' to the cleavage site. An extra base was also included in order to keep the 5' *lacZ* gene sequence in-frame. PCR products were ligated into a PCRScript transfer vector which was used as source from which the amplification product could be excised using *Sph* I and *Xba* I. The excised fragment was ligated into a similarly restriction digested pPD96.04 vector and the cloning junctions of the resulting construct were sequenced using the primers 'M13 reverse', 'zk430.8 prom seq rev', 'zk430.8 prom seq fwd' and '96.04 sequencing rev'.

After verification that the correct insert had been cloned, and most importantly that the *lacZ* had been kept in-frame, transgenic lines were obtained by microinjection of the construct and the *unc-76* rescue plasmid into DR96 animals as detailed in section 2.12. Two transgenic lines were obtained in this way.

### 2.29.2 Staining for $\beta$ -galactosidase activity

Two synchronous populations of L1 and L3 worms for each of the transgenic lines were obtained by the method described in section 2.22.5. These were washed from plates using 1X

M9 buffer (3 washes of 1X M9). The volume after the final wash was reduced to 100  $\mu$ l and one volume of 2.5% glutaraldehyde (Sigma-Aldrich) was added. Suspensions were incubated for 20 minutes at room temperature with gentle agitation and were subsequently washed 2 times with 1X M9. The volume was reduced to 100  $\mu$ l and pipetted onto a slide. The slide was dried by allowing the liquid to evaporate at room temperature for 30 minutes and subsequently, the worms were dehydrated by transferring the slide to ice cold acetone for 20 minutes. Slides were air-dried and a reservoir was painted around worms using nail varnish. 50  $\mu$ l of stain mix was added to the reservoir, a coverslip was applied and sealed with nail varnish. Worms were checked every hour for staining or left overnight.

Staining mix: 500  $\mu$ l 0.4M NaPO<sub>4</sub>, 1  $\mu$ l 1 M MgCl<sub>2</sub>, 100  $\mu$ l Redox buffer (50 mM potassium ferricyanide, 50 mM potassium ferrocyanide), 4  $\mu$ l 1% SDS, 400  $\mu$ l dH<sub>2</sub>O, 12  $\mu$ l of 2% (w/v) X-gal.

### 2.29.3 Cloning of the *hpx-1* promoter sequence into a vector with a destabilised EGFP

The vector, pPAF207 (a gift from A. Frand, Massachusetts General Hospital, Boston) is a modification of pPD95.81 from Andrew Fire's lab that was designed in order to create a vector tool with the capacity to assay the kinetics of transient gene regulation. This is facilitated by the incorporation of pd1EGFP-N1, a destabilised enhanced green fluorescent protein, between the last coding codon and the stop codon of the pPD95.81 GFP. pd1EGFP-N1, originally from mouse ornithine decarboxylase (Li *et al.*, 1998), was obtained from Clontech and has a half-life of approximately 1 hour. Consequently the GFP reporter, under the influence of the promoter in question, is degraded after an hour of being expressed, thus ensuring that no residual reporter product remains. The vector backbone, pPD95.81, is similar to pPD96.04 in having synthetic introns incorporated into the coding sequence of the reporter gene (only 5 SI are inserted into GFP instead of 12 in pPD96.04), a decoy 5' sequence and *unc-54* at the 3' end. However it differs as reporter constructs are not sequestered to the nucleus by a synthetic NLS.

2089 bp of the putative promoter of *hpx-1* as well as 14 bases of coding sequence, encoding 4 residues plus two bases that keep the construct in frame, were cloned into pPAF207 using the primers 'zk430.8 prom fwd' (*Sph* I) and 'pest vector rev zk430.8' (*Bam*H I). These were used in a reaction containing 1  $\mu$ l of 12.5 mM dNTPs, 1  $\mu$ l genomic DNA (~200 ng), 1  $\mu$ l of 25 mM

MgCl<sub>2</sub> and 2 units of Vent polymerase in a 50 µl volume at 1X concentration of ThermoPol buffer. 25 cycles of PCR were utilised to amplify the fragment which was then cloned into TOPO. Using the restriction enzymes *BamH* I and *Sph* I, the fragment was excised from this vector and ligated into a similarly linearised pPAF207 vector with a 3:1 insert:vector ratio (100 ng of vector).

Positive clones were identified using a restriction digestion method with the *Sph* I and *BamH* I enzymes and were subsequently sequenced using the primers: 'paf207 seq fwd', 'paf207 seq rev', 'zk430 prom seq fwd2' and 'zk430. prom seq rev2'

This construct was co-injected into DR96 animals with the Unc rescue plasmid, p76-16B. Successfully transformed animals were identified on their Unc rescue phenotype. Two lines were obtained which were used in a time course analysis using synchronous populations as described in section 2.22.3. A number of arrested L1 animals were placed onto OP50 in order to induce their entry into the life cycle. Since only a percentage of worms in transformed lines express the construct, the animals transferred to food-supplemented plates were chosen on the basis of their movement as L1 arrested larvae in 1X M9. Plates were incubated at 25°C and at 2-hour intervals, were monitored and the number of animals expressing the marker construct was counted. This was continued until the animals reached the adult stage.

## 2.30 RNA interference (RNAi)

### 2.30.1 Construction of plasmids

Prior to the release of the RNAi library (Kamath *et al.*, 2003), plasmids required for RNAi experiments had to be constructed. The primers, listed in table 2.6 and detailed in table 2.7, were used to amplify fragments of the genes of interest from N2 cDNA or genomic DNA using *Taq* polymerase (30 cycles of the PCR programme). Amplified products, whose sizes are detailed in table 2.6, were cloned into either TOPO or PCRScript. Fragments were excised from their transfer vectors using restriction enzymes sites, situated within the vector and which are specified in table 2.6. Fragments were ligated into a similarly restriction digested L4440 vector at a site between the 2 T7 promoters in the latter vector. During these standard cloning procedures, transformation of *E.coli* strains was into XL10-gold cells. The *col-19*, *sqt-1* and *sqt-3* RNAi constructs were constructed by Gillian McCormack (University of Glasgow).

Gene	Primers	Fragment Size (bp)	Restriction sites used	TOPO/PCR Script
zk430.8	zk430.8 fwd zk430.8 rev	781 (cDNA)	<i>Pst</i> I/ <i>Not</i> I	PCRScript
<i>sqt-1</i>	<i>sqt-1</i> fwd <i>sqt-1</i> rev	975 (cDNA)	<i>Not</i> I/ <i>Xho</i> I	PCRScript
<i>sqt-3</i>	<i>sqt-3</i> fwd <i>sqt-3</i> rev	906 (cDNA)	<i>Not</i> I/ <i>Xho</i> I	PCRScript
<i>col-19</i>	<i>col-19</i> F <i>col-19</i> r	870 (cDNA)	<i>Not</i> I/ <i>Nco</i> I	pGEM
co5d11.1	c05d11.2 fwd c05d11.2 rev	1207 (cDNA)	<i>Pst</i> I/ <i>Not</i> I	PCRScript
<i>dpy-7</i>	<i>dpy-7</i> fwd <i>dpy-7</i> rev (no stop)	1590 (gDNA)	<i>Hind</i> III/ <i>Not</i> I	PCRScript
<i>dpy-13</i>	<i>dpy-13</i> fwd <i>dpy-13</i> rev (no stop)	1084 (gDNA)	<i>Hind</i> III/ <i>Not</i> I	PCRScript
f56g4.5	f56g4.5 fwd f56g4.5 rev	1153 (cDNA)	<i>Pst</i> I/ <i>Not</i> I	PCRScript
c17g1.6	c17g1.6 fwd c17g1.6 rev	1220 (cDNA)	<i>Xho</i> I/ <i>Kpn</i> I	TOPO
c33h5.9	c33h5.9 fwd c33h5.9 rev	898 (cDNA)	<i>Pst</i> I/ <i>Not</i> I	PCRScript
f08c6.1	f08c6.1 fwd f08c6.1 rev	1118 (cDNA)	<i>Hind</i> III/ <i>Not</i> I	PCRScript
f45g2.5	f45g2.5 fwd f45g2.5 rev	420 (cDNA)	<i>Hind</i> III/ <i>Not</i> I	PCRScript

**Table 2.6: Primers used for amplification of RNAi constructs.** For each gene, the pair of primers utilised is given. In addition, the size of the amplification product, the transfer vector utilised and the restriction sites used for ligation into the L4440 plasmid, are given

The *dpy-5* RNAi construct (McMahon *et al.*, 2003), was obtained from I. Johnstone. All other RNAi feeding clones were obtained from a library supplied by Julie Ahringer (Kamath *et al.*, 2003). This library contains 16, 757 bacterial clones that represent the RNAi feeding constructs

of 86% of the *C. elegans* genome. Clones were retrieved from the library by streaking LB Agar plates (Amp + Tet) with glycerol stock supplied in the feeding library. Plates were incubated at 37°C overnight.

### 2.30.2 RNAi by feeding

Purified plasmid DNA of the clones that were constructed were transformed into *E. coli* HT115 (DE3) cells (Timmons *et al.*, 2001) by the method described above 2.12.2 in section. Single colonies from transformation plates (Tet and Amp) or from plates containing colonies from the RNAi feeding library, were used to inoculate 10 ml liquid LB cultures (Amp) that were incubated at 37°C overnight with shaking. It has been noted previously that, despite HT115 (DE3) cells being resistant to tetracycline, the addition of this antibiotic into cultures may be detrimental to the subsequent RNAi effect. Approximately 200  $\mu$ l of culture was pipetted onto 4.5 cm diameter NGM Agar plates that had been supplemented with 1mM IPTG and 100  $\mu$ g/ml of ampicillin. These were incubated at room temperature overnight in order to induce T7 RNA polymerase. The following day, typically 5 L4 animals were picked onto each plate. Plates were incubated at 20°C and/or 25°C. However, *sqt-3* targeted RNAi experiments were carried out at 15°C. After 24 hours the P0 worms had laid eggs, and they were observed for any phenotypes and were then picked off. The plates containing the F1 larvae were monitored for phenotypes. This method is based on the protocol described by Kamath *et al.* (2000).

### 2.30.3 *In vitro* transcription of RNA

*In vitro* synthesis of double stranded RNA was required in order to administer RNAi by microinjection and soaking, techniques that were used to specifically target *col-19*. A construct containing *col-19* within an L4440 vector (section 2.30.1) was used as a template for RNA synthesis by T7 RNA polymerase. As described above, the L4440 construct had been designed in order that the 2 T7 promoter regions encoded by the vector flanked the inserted gene. 10  $\mu$ l of plasmid DNA was restriction digested in two separate reactions to create two populations of linearised DNA (*Not* I digested and *Nco* I digested). Each had been digested appropriately in order that T7 polymerase would transcribe the inserted gene from opposing directions. It was important that complete digestion occurred. DNA was precipitated using the ethanol precipitation technique described in section and was diluted in 40  $\mu$ l of DEPC-treated dH<sub>2</sub>O. RNA was synthesised using a RiboMAX™ Large Scale RNA production system from Promega. The T7 reaction was set up and the subsequent removal of the DNA template steps were carried out as describe in the manufacturer's instructions. The modifications to the

directions were that ethanol precipitation was chosen over isopropanol precipitation and that 1  $\mu$ l of RNase inhibitor was added to the final elution fraction. RNA was quantified with a spectrophotometer.

dsRNA was obtained by annealing the 2 RNA products together. This was carried out by incubating a mixture, containing 1 mg/ml of each RNA strand (in a 50  $\mu$ l reaction (DEPC-treated H<sub>2</sub>O)), at 37°C for 30 minutes. The RNA synthesis described in this thesis was carried out by Gillian McCormack (WCMP, University of Glasgow, UK).

#### **2.30.4 dsRNA microinjections**

Double-stranded RNAi was administered by microinjection at a concentration of 0.4mg/ml. Prior to injection, as described previously, samples were centrifuged to remove debris. A microinjection technique identical to that described in section 2.22.5 was performed. However, it was not essential that the gonads were injected with the RNA solution as the RNAi effect spreads through the tissues. The microinjection of dsRNA described in this thesis was carried out by Gillian McCormack (WCMP, University of Glasgow, UK).

#### **2.30.5 RNAi by soaking (*col-19*)**

RNA was annealed as previously described (section 2.30.3) and was used to perform RNAi by soaking. Two different RNAi mixtures were used, containing either spermidine or lipofectin. For lipofectin reactions, 15-20 L4 larvae were transferred to tubes containing 10  $\mu$ l of 2X M9, 0.4 mg/ml ds RNA, and lipofectin 0.067 g/ml. In spermidine reactions, the lipofectin was substituted with 1  $\mu$ l of 60 mM spermidine. Tubes were set up so that each of the different solutions was incubated at 20°C and 25°C. Worms were soaked for between 1-3 day before the RNAi mixture was pipetted onto a fresh OP50 NGM agar plate. The adult animals were allowed to recover for 6 hours before being transferred singly to a fresh plate. Adults were allowed to egg lay and were then removed. The soaking reactions described in this thesis were carried out by Gillian McCormack (WCMP, University of Glasgow, UK). F1 progeny on these plates were examined for visible phenotypes.

#### **2.31 Construction of a Ty-tagged COL-19**

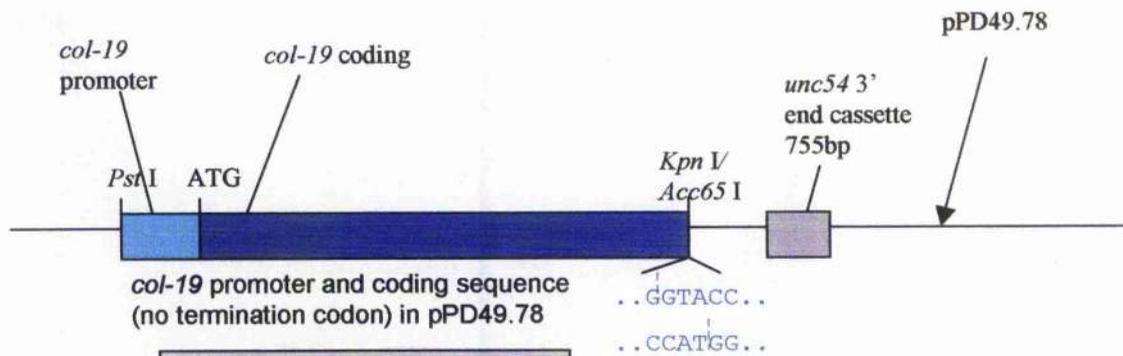
The initial step to produce a Ty-tagged COL-19, was to ligate the promoter and coding sequences of *col-19* into the vector pPD49.78 which is another in the series of vectors produced by the laboratory of Andrew Fire. pPD49.78 is a vector traditionally used for heat-shock inducible expression. However, the heat-shock function was not utilised in these experiments.

The *col-19* promoter and coding sequences had previously been cloned into the BA7.1 vector to form the COL-19::GFP fusion construct (supplied by Shoemaker and described in Thein *et al.* (2003). However, the *col-19* coding sequence did not include its stop codon in order to allow read-through into the GFP marker moiety. Flanking *Pst* I and *Kpn* I restriction sites of the *col-19* promoter-coding region were used to excise this region from the *col-19::gfp* encoding vector. Similar restriction sites are present in pPD49.78 in multiple cloning sites (MCS) I and II respectively and restriction digestion of the vector removed the synthetic intron between the two multiple cloning sites. The *col-19* promoter plus coding sequences was ligated between these sites. Positioning the construct here meant that it was flanked at its 3' end by the *unc54* 3' end cassette (755 bp).

The Ty-tag, derived from the major structural protein of the *Saccharomyces cerevisiae* Ty1 virus-like particle, is a 10-residue long epitope tag, encoded by a 30 bp sequence (Bastin *et al.*, 1996). It is a useful marker because of its small size (less disruptive) and the fact that it has been proven to be immunologically recognised in various environments (Bastin *et al.*, 1996). The Ty-tag epitope in this case was designed to be inserted at the C-terminus of the *col-19* coding sequence at an identical site to where the GFP moiety attached. Because of the manner in which the COL-19::GFP construct was formed originally in BA7.1, this site of attachment is defined by a *Kpn* I recognition site. A PCR amplification step was not utilised during the construction, rather, the Ty-tag was ligated into the vector in the form of a blunt-ended, double stranded (sense-strand and anti sense-strand) primer. Although, for such a ligation step, restriction sites are not required, the primers were designed to re-establish a *Kpn* I restriction site at the 5' end (of the sense strand). In order that the Ty-tag could be excised out at a later stage, a 3' restriction site (*Xba* I) was also included. Further 3' to this *Xba* I site, a termination codon was encoded. This ensured that when the Ty-tag was removed, the termination codon (in frame) would be left. Primers were designed in order to maintain correct read-through by using a combination of the choice of restriction sites and the addition of inserted bases (figure 2.1).

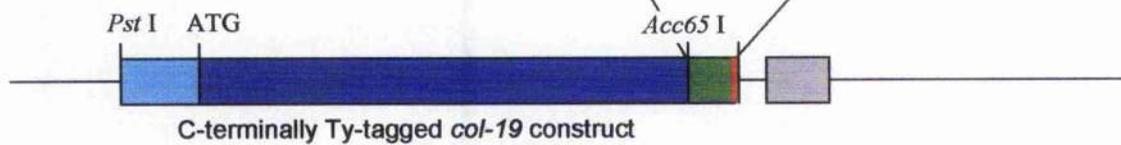
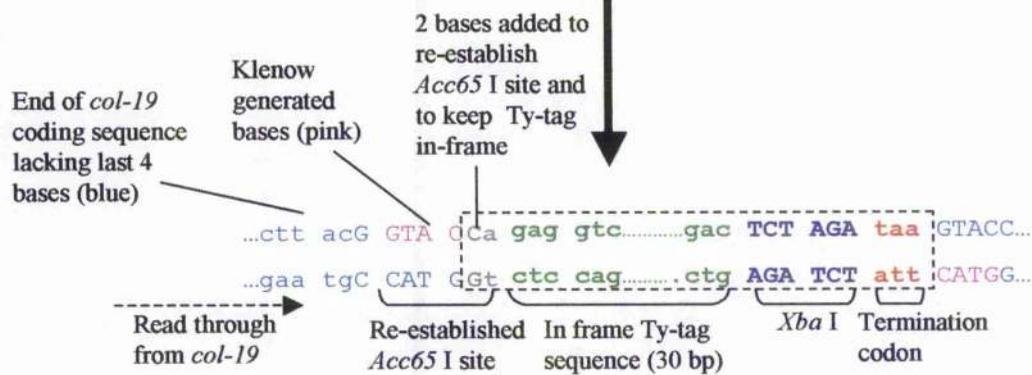
The sequence of the Ty-tag primers ('ty-tag sense' and 'ty-tag antisense' (HPLC grade) are detailed in table 2.7 and show in figure 2.1. These primers, each at 1 µg/ml in TE buffer, were annealed in a 100 µl volume at 55°C for 5 minutes and allowed to cool slowly. The final concentration of annealed primers was therefore 1 µg/ml. This blunt ended insert was ligated into the *col-19*-containing pPD49:78 that had been prepared via the following steps: 10 µg of plasmid DNA was restriction digested for 2 hours at 37°C using 20 units of *Ace* 651, an

**Figure 2.1: Construction of a C-terminally Ty-tagged *col-19*.** **A)** Using the restriction enzymes *Pst* I and *Kpn* I, the promoter (light blue) and genomic coding sequence (dark blue) of *col-19* derived from BA7.1 (containing the full length sequence of *col-19* minus 4 bp), were ligated into the multiple cloning site of the pPD49.78 plasmid which contains a 3' UTR sequence from *unc54* of *Schistosoma mansoni* (grey). The resulting construct is shown. **B)** The *col-19*-containing vector was restriction digested at the 3' end of *col-19* by *Acc65* I, an isoschizomer of *Kpn* I that produces 5' overhangs. Klenow treatment produced blunt ends (pink) that could be ligated to blunt ended and annealed primers that contained: 30 bp of the Ty moiety (green); two 5' end bases to ensure in-frame read through into the Ty-tag and to re-establish the *Acc65* I site (grey); an *Xba* I site (purple); and a 3' end in-frame termination codon (red). Restriction digest sites are in capital letters. Sequences derived from the Ty-tag primers are bold and are boxed.



**A:** (Resulting construct is shown). Ligation of the promoter and coding sequence of *col-19* into pPD49.78 (*Pst* I/*Kpn* I). These were obtained from BA7.1 and thus lack the termination codon.

**B:** Restriction digestion with *Acc65* I and Klenow treatment to generate blunt ends (pink). Blunt-ended ligation of annealed primers. The Ty-tag sequences are boxed and bold.



isoschizomer of *Kpn* I, in a 50µl reaction volume of 1X buffer. After 2 hours, the reaction was terminated by incubating the reaction at 65°C for 20 minutes and then 1 µl of 12.5 mM dNTPs and 10 units of DNA Polymerase I, Large (Klenow) fragment (New England Biolabs) were added to the mixture and incubated at 25°C for 15 minutes. The reaction was terminated by adding EDTA to a 10mM final concentration and heating at 75°C for 10 minutes. During such treatment (figure 2.1), 5' overhangs were generated by the *Acc* 65I linearisation step which were subsequently blunt ended using the polymerase activity of Klenow, a proteolytic product of *E. coli* DNA Polymerase I which has lost its 5'-3' exonuclease. A ligation reaction was set up in 1X ligase buffer with 20 ng of treated vector, 1 ng of annealed Ty primers and 400 units of T4 DNA ligase enzyme (New England Biolabs). Ligations were then transformed into XL10-Gold cells following the method described in section, DNA preparations were made from individual colonics and these were restriction digested with *Pst* I and *Xba* I to identify positive clones. Subsequently, positive clones were subjected to a PCR method using the primers 'antisense Ty' and 'col-19 prom fwd' in order to identify into which of these clones, the Ty-tag insert had been ligated in the correct orientation. Clones were sequenced using the primers 'Seq col-19 minus 117 bp' and HSPR1.

Transgenic lines containing this construct were obtained through microinjection. Purified Ty-tag construct and the *unc-76*-rescue plasmid at 20 µg/ml and 100 µg/ml respectively were injected into DR96 worms and transformants were selected on the basis of Unc rescue in the F1 population. Worms were then subjected to staining by anti-Ty antibodies which were a gift from Keith Gull (University of Oxford, UK) (Bastin *et al.*, 1996; McMahon *et al.*, 2003). The protocol for such immunolocalisation assays is detailed in section 2.27.

### 2.32 Re-ligation of the Ty-tag into the *col-19* coding sequence

It was of interest to observe the effects that the Ty moiety had when inserted into the coding sequence of the *col-19*. Thus, a scheme as depicted in figure 2.2 was carried out. During this process, the Ty-tag coding sequence was removed from the original Ty construct and a new Ty-tag moiety was ligated at a new position. The sequence of the original Ty-tag (which encoded only a partial fragment of the *Kpn* I site) was designed in order that there was re-establishment of the full recognition *Kpn* I/*Acc* 65I sequence in successful ligations. Restriction digestion (2 hours 37°C) at this site (using *Acc* 65I) and at the 3' *Xba* I site, enabled the excision of the Ty-tag (from a 2 µg sample of DNA). Similar Klenow treatment as described above was then carried out, the only difference being that this step was followed by

**Figure 2.2: Ligation of a Ty-tag into the *col-19* coding sequence.**  
**A)** To firstly remove the Ty-tag from the C-terminus of *col-19* but leave the (synthetic) termination codon (red) the construct containing the C-terminally tagged construct was restriction digested with *Acc65* I and *Xba* I (the *Xba* I site had been included in the Ty-tag, 5' to the termination codon). After treatment with Klenow made blunt ends, which were annealed, leaving the termination codon in frame. **B)** In order to insert a Ty-tag in frame into the *col-19* coding sequence (blue) the plasmid resulting from part 1 was restriction digested with *Nsi* I. Ty-tag primers were designed containing the 30bp of the Ty moiety (green), *Nsi* I sites at either end, and three bases (grey) that ensured the correct frame. Restriction sites are in capitals. Sequences from Ty-tag primers are bold.

Facing page 96



**A) Removal of the Ty-tag from the C-terminus of *col-19* to leave the in-frame stop codon**

Original Ty-tag at the C-terminus of *col-19*

```

...g cag ctt acG|GTA CCa gag gtc.....gac TCT AGA taa C..
...c gtc gaa tgc|CAT Ggt ctc cag.....ctg AGA TCT att G..
                Acc65 I  Ty-tag (30bp)  Xba I
  
```

↓ *Acc65 I*+*Xba I* restriction digest

```

...g cag ctt acG                               CT AGA taa C..
...c gtc gaa tgc|CAT G                          att G..
  
```

↓ Klenow polymerase treatment to make blunt ended and ligate

Regeneration of *Acc65 I* site (not the *Xba I* site) keeping *col-19* in frame with stop codon (red)

```

...g cag ctt acG GTA C CT AGA taa C..
...c gtc gaa tgc|CAT G, GA TCT att G..
                Acc65 I
  
```



**B) Insertion of the Ty-tag into the *col-19* coding sequence**

```

...A TGC AT...
...T ACG TA...
      Nsi I
  
```

+ Base added to keep Ty-tag in frame      2 bases added to keep *col-19* in frame

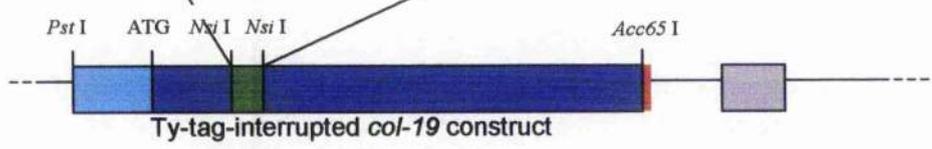
```

Tc gag gtc...gac ggA TGC A
ACG TAG ctc cag...ctg ccT
Nsi I      Ty tag (30bp)      Nsi I
  
```

↓ Ligation of *Nsi I*-digested vector and annealed Ty-tag primers containing *Nsi I* overhangs. Blue writing is *col-19* coding sequence. Green/grey, bold is Ty-tag insert.

```

...A TGC ATc gag gtc...gac ggA TGC AT...
...T ACG TAG ctc cag...ctg ccT ACG TA...
      Nsi I      Ty tag (30bp)      Nsi I
  
```



gel extraction to purify the vector. Such treatment left a termination codon (which was encoded by the initial Ty-tag primers) in-frame at the 3' end of the *col-19* coding sequence. Clones were sequenced using the primers 'Seq col-19 minus 117 bp' and 'HSPR1'. This was an important step because Klenow reactions, when incubated at longer than the suggested incubation time, may result in exonuclease activity which would have resulted in the stop codon being out of frame.

A pair of sense and anti sense primers 'Ty-tag *nsi-1* sense' and 'Ty-tag *nsi-1* antisense' (HPLC grade) (table 2.7 and figure 2.2) were designed which allowed the in-frame addition of the new Ty-tag into an *Nsi* site within the Gly-X-Y regions. In order to generate the cohesive ends for ligation into the construct, annealed primers were initially ligated into the intermediate vector, pPCRScript. This circumvented the problem of too few bases around the restriction site for efficient recognition by the restriction enzyme. Approximately 20 µg of pPCRScript DNA was restriction digested using the enzyme *Nsi*I. It was not possible to extract the excising band from standard agarose gels because of the small size (45 bp) of the fragment. Thus a polyacrylamide gel was utilised to purify the Ty-tag insert. This is detailed below.

### 2.32.1 Polyacrylamide gel purification of small DNA fragments

A 5% w/v non-denaturing acrylamide gel was prepared. The glass plate sandwich was assembled as per manufacturer's instructions (using 1.5 mm spacers). A mixture was made containing: 1.25ml of a 30% stock of 1:29 acrylamide:bisacrylamide, 6.68ml of H<sub>2</sub>O and 2ml of 5X TBE, and, prior to casting, 70 µl of 10% ammonium persulphate and 7 µl of TEMED were added (followed by mixing and administering the mixture into the plate sandwich). A 15 well comb was inserted and the gel was allowed to set. The entire restriction digest (from the step in section 2.32), after being mixed with DNA sample buffer to a final volume of 1X, was loaded in 2 wells and the gel was run in a mini-PROTEAN® electrophoresis cell system at 100v in 1X TBE until the dye front reached the bottom of the gel. DNA was visualised by soaking the gel in 50ml of 1X TBE + 5 µl ethidium bromide.

The band was excised using a scalpel and transferred to a microfuge tube. A 'crush and soak' method, as described in Sambrook *et al.*, 1989, was used to extract the Ty-tag DNA. The first step of this process was to crush the gel fragment against the side of the tube using a pipette tip. 1-2 volumes of elution buffer were added to the gel fragment and this was incubated at 37 °C on a rotary platform for 4 hours. The remnants of the gel fragments were subsequently pelleted via

centrifugation for 1 minute at 12,000 g and the supernatant was transferred to a fresh tube. To ensure that as much DNA as possible was extracted from the gel, an additional 500  $\mu$ l of elution buffer were added to the gel remnants to wash it further (vortex briefly, re-centrifuge). The supernatant was pooled with the first elution volume. The DNA was precipitated by adding 2 volumes of ethanol, mixing well, and allowing DNA to precipitate on ice for 30 minutes. The DNA was pelleted via centrifugation at 12,000 g for 10 minutes and the pellet was resuspended in 200  $\mu$ l of TE buffer. The DNA was precipitated for a second time and was washed with 1ml of 70% ethanol. The DNA was resuspended in a final volume of 10  $\mu$ l dH<sub>2</sub>O. The only modifications from the protocol in Sambrook *et al.*, 1989, was that the suggested filtration step was excluded and that the DNA was resuspended in dH<sub>2</sub>O rather than TE buffer since it was possible that the TE may interfere with subsequent reactions.

Elution buffer:            0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA (pH 8.0), 0.1% SDS.

### 2.32.2 Ligation of the Ty-tag into the Gly-X-Y sequence of COL-19

Prior to ligation of the Ty-tag into the *Nsi* I site of the *col-19* sequence, the *Nsi* I cohesive ends were treated with CIP in order to dephosphorylate the 5' end and prevent recircularisation of the vector. ~10  $\mu$ g of the plasmid DNA obtained as described in section 2.31 was digested (37°C 2 hours) with 20 units of *Nsi* I and 0.5  $\mu$ l BSA in a 50  $\mu$ l volume of 1X buffer. After termination of this reaction by heat inactivation (65°C 20 minutes), 1  $\mu$ l of CIP was added to 10  $\mu$ l (~2  $\mu$ g DNA) of this reaction mixture. This was incubated at 37°C for 1 hour, after which the dephosphorylated vector was phenol:chloroform extracted, precipitated and resuspended in 10  $\mu$ l of dH<sub>2</sub>O (section 2.9.1). A ligation reaction was set up with 10  $\mu$ l of the acrylamide gel purified DNA (insert extract), 1  $\mu$ l of the dephosphorylated vector, 400 units of T4 ligase, in 13  $\mu$ l total volume of 1X T4 ligase buffer. This ligation was left at 15°C overnight before being transformed into XL10 cells. A colony PCR technique was used to confirm the presence of the Ty-tag. This used the primers 'Ty-tag col-19 rev seq' and 'M13 reverse'. Plasmid DNA resulting from successful ligations was used to transform DR96 (*unc-76*) worms using the standard microinjection procedure. Transformed lines were used in immunological assays using an anti-Ty antibody and the methods described above (section 2.27).

## 2.33 Promoter swap construction

### 2.33.1 GFP fusion collagens under the *col-19* promoter

The scheme depicted in figure 2.3 (using the example of *sqt-1*) was used to create a construct, in a BA7.1-derived backbone, in which the *col-19* promoter was fused to the coding sequences of a number of collagens, which in turn were fused in-frame to *gfp-c3*. To achieve the read-through, it was necessary to design each of the reverse primers for amplification of coding sequences so that the termination codon was omitted. Additionally, the final coding base was omitted to maintain the reading frame.

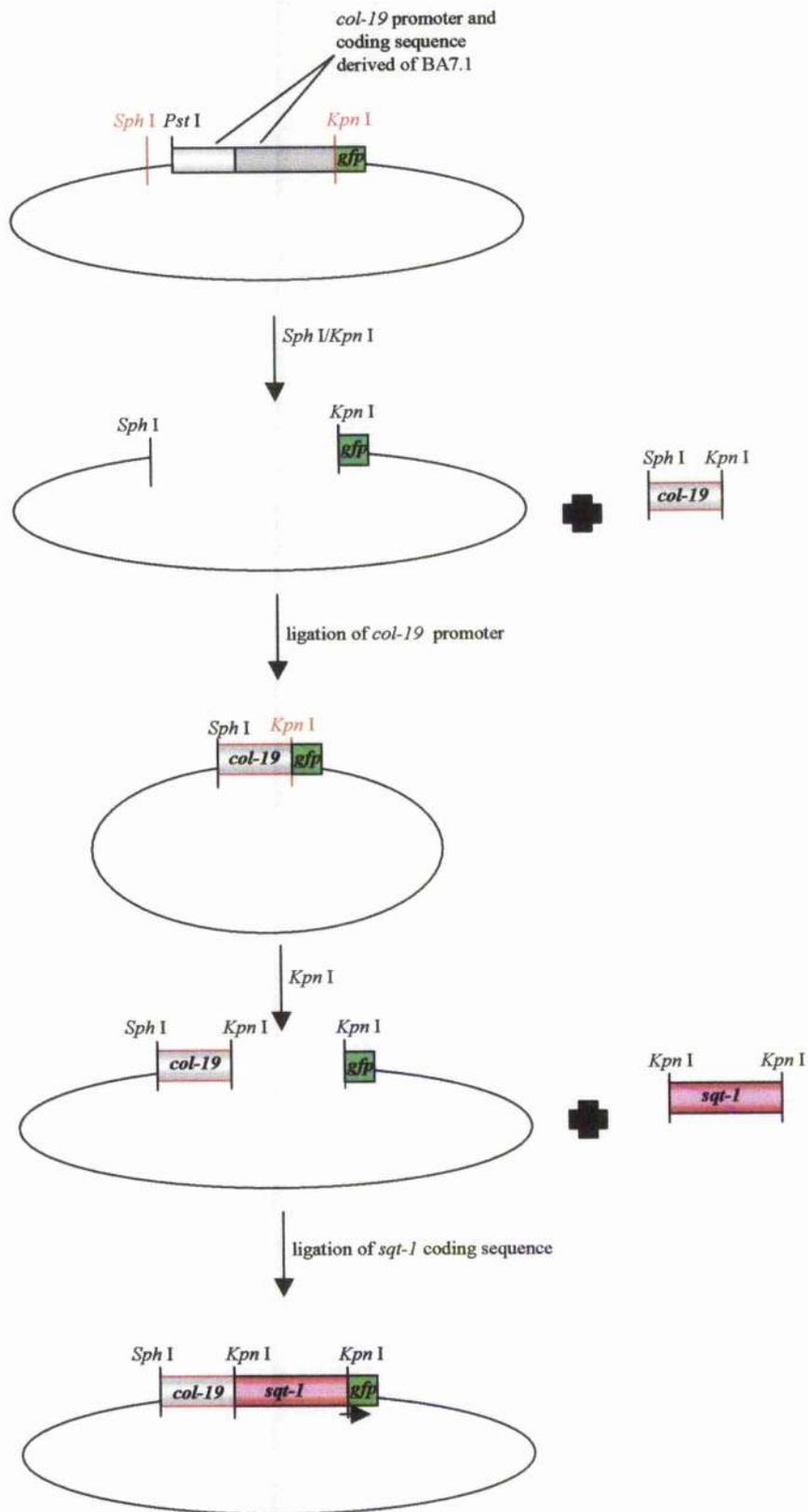
A new *col-19* promoter, flanked by *Sph* I and *Kpn* I sites, was amplified using the primers 'prom col-19 fwd' and 'prom col-19 rev' (*Taq* polymerase, 30 cycles), and subsequently cloned into pPCRScript and then using standard ligation, transformation and colony screening protocols, the *col-19* promoter was ligated into a linearised pCHUCK vector that had been restriction digested with *Sph* I and *Kpn* I to excise the *col-19* promoter and coding sequences (figure 2.3).

The resulting construct encoded a *Kpn* I site at the junction of the promoter and *gfp* moiety. This vector was restriction digested at this *Kpn* I site and 1 µg of linearised vector was then subjected to standard CIP treatment (refer to section 2.16). The coding sequences of the collagens were ligated into this *Kpn* I-*Kpn* I site. The coding sequences were amplified (vent, 27 cycles, genomic DNA) using the primers: 'sqt-1 code fwd' and 'sqt-1 code rev (no stop)'; 'dpy-5 fwd' and 'dpy-5 rev (no stop)'; 'dpy-7 fwd' and 'dpy-7 rev (no stop)' and 'dpy-13 fwd' and 'dpy-13 rev (no stop)'. All the primers were designed with engineered *Kpn* I sites and all PCR reactions utilised Vent polymerase, genomic DNA and 30 cycles. The amplification products were cloned into either pPCRScript or TOPO and were subsequently excised using *Kpn* I and ligated into the dephosphorylated vector using standard ligation and transformation methods. Positive clones were identified by colony PCR with the 'prom col-19 fwd' primer and the gene-specific reverse primer. Transformed lines carrying these constructs were obtained by microinjection of plasmid DNA and an Unc rescue marker (p76-16B) into DR96.

### 2.33.2 GFP fusion collagens under the *sqt-1* promoter

A slight modification of the above procedure was used to clone a construct containing the *col-19* coding sequence, fused in-frame to *gfp*, under the influence of the *sqt-1* promoter (figure

**Figure 2.3: Creation of a promoter swap construct: *sqt-1* coding sequence under the influence of the *col-1* promoter.** The *col-19* promoter and coding sequences were excised from BA7.1 using *Kpn* I and *Sph* I. The *col-19* promoter sequence, flanked with synthetic *Sph* I and *Kpn* I sites, was amplified and ligated into empty BA7.1 vector. This construct was then restriction digested with *Kpn* I (and dephosphorylated). The coding sequence of *sqt-1* was amplified by PCR and ligated into the *Kpn* I site, using synthetic flanking restriction sites. The last 4 bases of the *sqt-1* sequence were excluded during amplification to remove the termination codon and to allow for in-frame read through into *gfp*. Not to scale. Note that *Sph* I and *Pst* I sites are adjacent to each other in the vector but are separated in the figure for ease of visualisation. Horizontal arrow denotes read-through into *gfp*. The amplified *col-19* product has been outlined in red to distinguish it from the original BA7.1 *col-19* coding sequence.



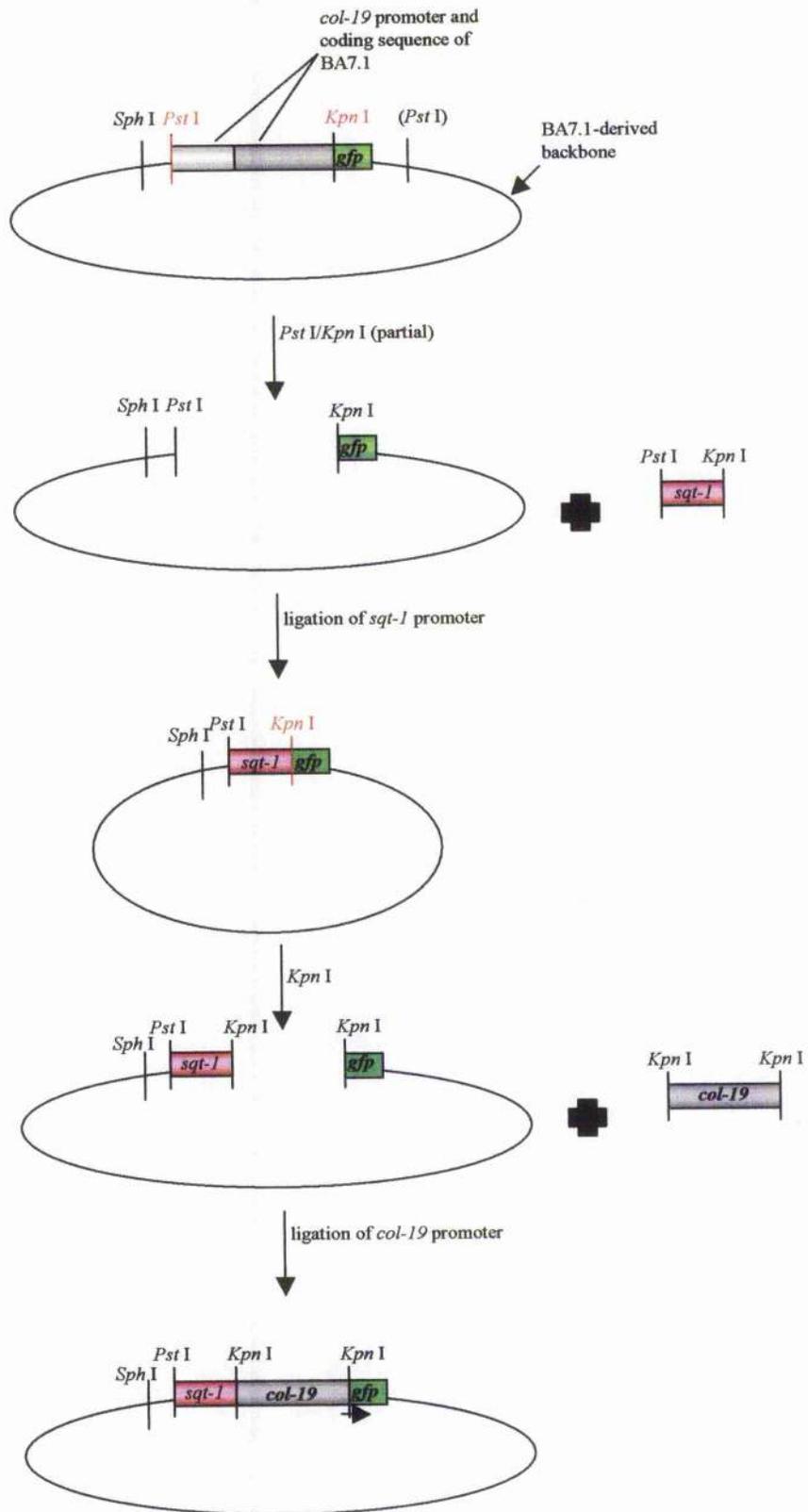
2.4). An alternative means of construction was required due to the fact that the *sqt-1* promoter contains an *Sph* I site. As an alternative, the *Pst* I site that lies 5' to the start of the *col-19* promoter in pCHUCK, was used as a site for cloning. However, because there is a second *Pst* I site in the vector backbone, the restriction digestion step was required to be a partial digest. This was carried out by first linearising 10 µg of the vector using *Kpn* I. Subsequently, the DNA was purified by phenol:chloroform extraction and ethanol precipitation. This DNA was aliquoted between three restriction digestion reactions each with different amounts of *Pst* I (1 unit, 0.5 unit, 0.1 unit). In addition 10 units of *Nst* I were added. The reactions were incubated at 37°C for 30 minutes and the correctly sized product was excised from a gel and purified. The *sqt-1* promoter, which had been amplified (with primers 'sqt-1 prom rev' (*Sph* I) and 'sqt-1 prom fwd' (*Pst* I) and *Pfu* polymerase, 27 cycles) with engineered *Sph* I and *Pst* I ends (via a TOPO transfer vector) was ligated into the vector using standard ligation methods.

The resulting plasmid was restriction digested with *Kpn* I. This was used in a ligation mixture containing the cloned *col-19* coding sequence that had been amplified using the primers 'col-19 code fwd' (*Kpn* I) and 'col-19 code rev (no stop)' (*Kpn* I) (*Pfu* polymerase, 27 cycles) and as an intermediate step, had been cloned into TOPO (from which it was excised with *Kpn* I). The constructs resulting from successful ligations were sequenced using the primers 'sqt-1 prom seq rev', 'sqt-1 seq prom fwd', 'col-19 code seq fwd', 'col-19 code seq rev', 'seq col-19 minus 117bp', 'col-19 code seq rev 2' and 'gfpc3 rev'. Transformed lines were obtained via microinjection as described in section 2.22.5.

### 2.34 Fusing *sqt-1* and its promoter to a GFP marker

The BA7.1 backbone was prepared in the same manner as described above (section 2.33.2), by restriction digestion with *Kpn* I and partial digestion using *Pst* I. This vector was used in a ligation reaction with an amplification product containing the promoter and coding sequence of *sqt-1*. This had been cloned (30 cycles, genomic DNA, Vent) using the primers 'sqt-1 prom fwd' and 'rev sqt-1 code' which engineer the *Kpn* I and *Pst* I sites at the 5' and 3' ends respectively. Successfully ligated clones containing the *sqt-1* promoter and coding sequences were sequenced using the primers: 'sqt-1 prom seq rev', 'sqt-1 prom seq fwd', 'gfpc3 rev', 'sqt-1 ISP F1', 'sqt-1 ISP R2', 'sqt-1 ISP F2', and 'gfpc3 rev'. Transformed lines were obtained by microinjection of plasmid DNA as described previously (section 2.22.5).

**Figure 2.4: Creation of a promoter swap construct: *col-19* coding sequence under the influence of the *sqt-1* promoter.** The *col-19* promoter and coding sequences were excised from BA7.1 using *Kpn* I and *Pst* I (partial digestion to avoid cleavage 3' of *gfp*). The *sqt-1* promoter sequence, flanked with synthetic *Pst* I and *Kpn* I sites, was amplified and ligated into empty BA7.1 vector. This construct was then restriction digested with *Kpn* I (and dephosphorylated). The coding sequence of *col-19* was amplified by PCR and ligated into the *Kpn* I site, using synthetic flanking restriction sites. The last 4 bases of the *col-19* sequence were excluded to remove the termination codon and to allow for in-frame read through into *gfp*. Not to scale. Note that *Sph* I and *Pst* I sites are adjacent to each other in the vector but are separated in the figure for ease of visualisation. Horizontal arrow denotes read-through into *gfp*.



### 2.35 Preparation of samples for scanning electron microscopy (SEM)

A method of SEM sample preparation using the strong fixative,  $O_2O_4$ , is necessary for *C. elegans* due to the cuticle of the animal slowing the diffusion of the reagents used in other preparations. This work was carried out at the Integrated Microscopy Facility at the University of Glasgow. 13 9-cm plates were used to produce a large mixed population of *hpx-1* (zk430.8) RNAi-treated animals (section 2.30.2). Nematodes were washed off plates using 1X M9 buffer, rinsed 3 times in the same buffer and transferred to a microfuge tube. 2.5% (w/v) glutaraldehyde in PBS (pH 7.4) was added, the tube chilled on ice for 90 minutes with occasional disruption of the pellet. The glutaraldehyde supernatant solution was discarded and the worms were washed 3 times in PBS + 2% (w/v) sucrose. After the final wash, the volume was reduced to 50  $\mu$ l and an equal volume of 2% (w/v)  $O_2O_4$  was added. The tube was left at room temperature (in a fume hood) for 1 hour. Subsequently, the  $O_2O_4$  was discarded and the pellet was thoroughly rinsed with 3 x 10-minute washes of PBS and 3 x 10-minute washes with  $dH_2O$ . Supplementary fixation to increase the stability of the tissues and increase the conductivity of the sample, was carried out using 0.5% (w/v) uranyl acetate (aqueous) (1 hour in foil in fume hood) after which worms were washed 3 x 5-minute in  $dH_2O$ . Worms were then dehydrated in a stepwise manner using increasing concentrations of acetone: 30% (v/v) acetone, 50% (v/v) acetone and 70% (v/v) acetone. At this stage, animals were transferred to a critical point dryer (CPD) filter chamber containing 1  $\mu$ m polycarbonate filters. The dehydration step was continued by washing the worms once in 90% (v/v) acetone (15 minutes), twice in absolute acetone (15-minutes each) and finally in dried absolute acetone. Samples were then processed in a critical point dryer (CPD). Worms were mounted onto copper double-sided tape by picking individual worms with a toothpick with an eyelash end and were viewed using a Phillips 500 scanning electron microscope.

### 2.36 Protein techniques

#### 2.36.1 SDS polyacrylamide gel electrophoresis (SDS PAGE)

A NuPage XCell Sure Lock system (Invitrogen), was set up as described in the manufacturer's instructions and was used with 4-12% gradient Bis-Tris high performance pre-cast gels to separate proteins. In preparation for loading onto the gels, 4X LDS sample buffer (Invitrogen) was added to samples for a final 1X concentration. For reducing conditions, the 4X sample buffer was made to 20% (v/v) final concentration of 2-ME. Broad range molecular weight markers (not-prestained) (New England Biolabs) were run alongside samples. Gels were run at

200 V in 1X MOPS buffer (Invitrogen) until the dye front reached the bottom of the gel. Gels were removed from the gel cases and were rinsed in H<sub>2</sub>O and were then transferred to a Coomassie blue solution (0.25 g of Coomassie brilliant blue R-250, 45% methanol, 10% glacial acetic acid) in order to visualise proteins. Gels were stained for 1 hour on a rocker, were transferred to detain solution (45% methanol, 10% glacial acetic acid) and left on a rocker overnight at room temperature.

### 2.36.2 Transfer of protein to PVDF membranes

A NuPAGE XCell II system (Invitrogen) was used for the transfer of proteins to a Hybond-P PVDF membrane (Amersham Pharmacia Biotech). The membrane was briefly soaked in methanol and 1X MES buffer and then the apparatus was set up as per manufacturer's instructions. The transfer was run for 1 hour at 25 volts using 1X MES buffer, which is recommended for the transfer of small molecular weight proteins. Subsequently, the membranes were washed briefly in H<sub>2</sub>O and then transferred to a milk + PBST blocking solution, washed in PBST (3 x 10-minute washes) and then transferred to a solution containing primary antibody. Membranes were washed again and then either transferred to a solution containing secondary antibody or were immediately developed. The specific milk concentrations of the block solutions, the primary and secondary antibody concentrations and the details of protein visualisation are explained more specifically in the following sections.

### 2.36.3 Construct preparation for expression of *col-19* in a baculovirus system

The baculovirus system presents an approach in which high levels of a target gene can be expressed and was therefore the system of choice for the production of recombinant COL-19 because the heterologous gene can be co-expressed with the  $\alpha$  and  $\beta$  subunits of P4H in order that the collagen be sufficiently hydroxylated for folding (Lamberg *et al.*, 1996). The primers 'col-19BVF' and 'col-19BVR' were used to amplify (Vent, 30 cycles) the full-length coding sequence of *col-19*, from mixed stage cDNA. The PCR product was first cloned into pPCRScript from which it was excised using the engineered *Pst* I and *Xba* I restriction sites. A ligation was then set up containing a 3:1 molar ratio of this excised product and a similarly restriction digested pPVL1392. The positive clones resulting from this ligation were sequenced using the T3 and T7 primers.

These constructs were given to Alan Winter (WCMP, University of Glasgow) who used a Pharmingen system to produce COL-19 recombinant protein extracts (this work was carried out

at The Collagen Research Unit, Oulu, Finland). Recombinant protein expression was carried out following the methods detailed in the instruction manual provided by the BV expression vector system manual (Crossen and Gruenwald, 1998). In this system, the non-essential gene, polyhedrin, of the baculovirus *Autographa californica* nuclear polyhedrous virus (AcNPV), is replaced by a heterologous gene of interest via homologous recombination from the transfer vector. This is performed by co-transfection of insect cells (*Spodoptera frugiperda*) with the transfer vector in which the MCS is flanked by sequences that direct the homologous recombination. Subsequently, expressed protein was purified using a method described by Lamberg, 1996.

#### **2.36.4 Trypsin treatment of BV expression-derived recombinant COL-19 protein**

20 µl of Triton X-100 extracted recombinant protein samples were subjected to trypsin treatment in 22.2 µl reaction mixtures using trypsin buffer at 1X concentration. Reactions were carried out at room temperature for 2-, 4- and 8- minutes, stopping the reaction by the addition of 5 µl of 4X non-reducing buffer (NuPAGE LDS Sample buffer) and by immediately boiling the sample for 5 minutes. These samples were loaded on a 4-12% gradient gel alongside reduced (1X SDS sample buffer with 5% 2-mercaptoethanol) and non-reduced (1X SDS sample buffer) samples that had not been treated with trypsin (and that had been boiled for 5 minutes prior to loading). A molecular weight marker (Broad Range, New England Biolab) was also loaded after denaturation.

Trypsin buffer: 0.4 M NaCl, 0.1 M Tris pH 7.4, 0.1mg/ml trypsin.

After running the gel and blotting the protein onto a PVDF membrane via the methods described in section 2.36.2, the membrane was treated with Ponceau S Staining solution in order to visualise the molecular weight marker bands. This was carried out as per manufacturer's instructions. However, because the bands did not become immediately obvious upon such treatment, the molecular marker band was cut off using a scalpel and left in the Ponceau S Staining solution overnight while the rest of the membrane was transferred to 20 ml of blocking solution (5% milk PBST) and left overnight at 4°C on a rocker. The following morning, the molecular weight markers had developed and could be permanently marked on the membrane with a pencil. The blocking solution was replaced with a 10 ml solution of primary antibody (anti COL-19) at a 1:250 dilution in 2.5% milk PBST (2 hours at room temperature, rocking). The membrane was subsequently washed 3 times with 2.5% milk PBST and then a streptavidin

alkaline phosphatase conjugate (Calbiochem) secondary antibody solution was added at a 1:15,000 dilution in 2.5% milk PBST. The membrane was left for 2 hours at room temperature, was then washed twice in PBST and once in PBS and was finally developed in a solution prepared by dissolving 2 Sigma Fast™ BCIP/NBT tablets (Sigma-Aldrich) in 20 ml of H<sub>2</sub>O.

### 2.36.5 Expression and purification of recombinant His-tagged HPX-1 protein

Purified recombinant protein was made using a cloning system (pQE30) that enables high expression of a N-terminal His-tagged heterologous gene product. A 1772 bp region of HPX-1 that encodes for the peroxidase domain was amplified using the primers 'zk430 fwd exon 5' and 'zk430 pqe30 rev' that bind to a region in exon 5 and the end of the gene respectively. This was amplified from mixed stage cDNA with *Pfu* polymerase using primers at 200 ng in PCR reactions (this is more than in the standard PCR protocol). The resulting PCR product was cloned into a TOPO transfer vector which was subsequently restriction digested at *Pst* I and *Hind* III sites that had been engineered into the PCR primers. This insert was cloned via the standard ligation reaction into a similarly digested pQE30 vector (Qiagen). The restriction sites had been engineered in order to maintain the reading frame in the His-tag encoding region. After transformation into XL10 cells, positive clones were identified via a colony PCR method using the primers 'm13' and 'pgm-1'. The identified clones were sequenced using the primers 'pqe30F', 'zk430 perox domain seq f2', 'zk430 perox domain seq rev', 'zk430 peroxdomseqf3', 'zk430 perox dom seq r2' and 'zk430 perox domain seq fwd', 'pqe30R'. Finally, pure plasmid preparations of positive clones were transformed into chemically competent M15 cells.

### 2.36.6 Purification of recombinant HPX-1 protein

The colonies derived from the transformation of M15 cells with recombinant protein constructs were used the same day to inoculate 20 ml liquid L-broth cultures plus antibiotics - ampicillin at 100 µg/ml and kanamycin at 25 µg/ml. Cultures were incubated in a rotary shake overnight at 37°C. The following day all 25 ml of culture was used to inoculate a 500 ml liquid culture (Amp 100 µg/ml and Kan 25 µg/ml) and this was incubated at 37°C in a rotary shaker until an OD<sub>600</sub> of 0.6 had been reached. At this stage, a 1 ml aliquot was removed and centrifuged for 1 minute to form a cell pellet and the rest of the culture was inoculate with IPTG to a final 1mM concentration. Cultures were grown for a further 2 hours before cells were pelleted in 250 ml Nalgene flasks (1 ml of post-induction culture was kept separate and pelleted in a microfuge tube to be run on an SDS PAGE gel at a later stage). Pellets were frozen at -20°C overnight.

The buffers utilised in the next steps are those described in the manual for the pQE vectors. The frozen pellet was allowed to thaw slowly and then was resuspended in 20 ml of 'buffer B'. This was placed at  $-80^{\circ}\text{C}$  for 30 minutes before again, being slowly thawed in  $\text{H}_2\text{O}$ . The cell slurry was sonicated (10 x 10 second pulses with 10 second pauses) using a wide tip probe. Great care was taken during the sonication step to ensure that the slurry was on ice at all times and that during sonication, foam did not form. The cellular debris were pelleted by centrifugation (30 minutes at 10,000g) and afterwards the supernatant (containing soluble protein) was removed to a fresh tube while the pellet (containing insoluble proteins) was resuspended in 5 ml of buffer B. 4 ml of ProBond resin (Invitrogen) that comes as a 50% slurry in 20% ethanol (to give a column with a volume of 2 ml) was loaded into a disposable polypropylene 5 ml column (Pierce) as per manufacturer's instructions. This resin is a nickel-charged affinity resin that uses the chelating ligand iminodiacetic acid (IDA) coupled to a highly cross-linked 6% agarose resin. The column was equilibrated with 20 ml of buffer B before the soluble protein-fraction from the centrifugation step was loaded onto it (20  $\mu\text{l}$  of the soluble protein-fraction was kept to be run on a SDS PAGE gel at a later stage). Proteins not binding to the column were washed out using 25 ml of 'buffer C' and finally, any proteins bound to the column were eluted into a fresh tube with 5 ml of 'buffer E'.

To prepare samples for the SDS PAGE gel, pre-induction and post-induction pellets were diluted in 50  $\mu\text{l}$  and 100  $\mu\text{l}$  of 1X 2-ME sample buffer respectively while 5  $\mu\text{l}$  of 2X SDS 2-ME sample buffer was added to 5  $\mu\text{l}$  of the crude extracts (the pellet and the supernatant from the centrifugation step). 5  $\mu\text{l}$  of 4X 2-ME SDS buffer was added to 20  $\mu\text{l}$  of each of the flow through, wash, and elution fractions. Samples were boiled for 5 minutes and centrifuged for 1 minute. 10  $\mu\text{l}$  of the cell pellet fractions and the crude sample fractions and 20  $\mu\text{l}$  of fraction samples were loaded onto a gel alongside 15  $\mu\text{l}$  of pre-boiled Broad Range marker. Two additional lanes were loaded, one with a Broad Range marker and the other with the elution fraction in preparation for Western analysis detection. The gel was run at 200 V after which the lanes loaded for Western analysis were cut off using a scalpel. The proteins on the rest of the gel were visualised using Coomassie blue and destained as described previously.

Buffers:

Buffer B:	100 mM $\text{NaH}_2\text{PO}_4$ , 10 mM TrisCl, 8 M urea (pH 8.0)
Buffer C:	same as above (pH 6.3)
Buffer E:	same as above (pH 4.0)

### 2.36.7 Determination of protein concentration

It was necessary to determine the yield of recombinant protein resulting from the purification step. To do this, a Coomassie Reagent protein assay system (Pierce) was used as per manufacturer's instructions using the BSA standards at working ranges between 100 – 1,500 µg/ml

### 2.36.8 Western Analysis of HPX-1 recombinant proteins

The proteins from the fragment of gel containing the elution fraction and the molecular weight marker were transferred to a PVDF membrane using the NuPage protocol. Novex 25X Tris Glycine buffer was used at a 1X concentration. The PVDF membrane was subsequently blocked using 2.5% milk PBST (2 hours at room temperature and was then replaced with Anti-His (C-terminal) Horseradish Peroxidase primary antibody (Invitrogen) at a 1:2000 concentration (2 hours at room temperature). A secondary antibody is not required in this protocol because of the conjugated HRP that can be detected using an ECL plus detection system (Amersham). 2 ml of the supplied ECL buffer A was mixed with 50 µl of the buffer B, the mixture was allowed to rest for 2 minutes and then was transferred to the PBST-washed membrane that had been placed protein-side-up on a piece of Saran wrap. The reaction was allowed to continue for 10 minutes at room temperature after which the detection solution was removed, the membrane wrapped up and exposed to X-ray film (Kodak) for between 1 and 4 minutes. The film was then developed.

### 2.36.9 Peroxidase activity detection

Since the eluted protein was in a buffer that was 8 M urea and pH4.5, both conditions that were suspected to be unfavourable for enzymatic activity, 1 ml aliquots of elution fractions were dialysed using a 10,000 kDa molecular weight cut off Slide-a-lyser (Pierce). The urea concentration was decreased incrementally, reducing the concentration by half in each step. It was then attempted to increase the pH incrementally to pH 6 but event at pH 5, the protein precipitated.

The method used for peroxidase activity detection was described in Edens *et al.*, 2001 which is based on the use of 3, 5, 3', 5'-tetramethylbenzidine (TMB) as a substrate (Holland, 1974, Edens, 2001). A volume of dialysed protein solution containing 100 µg of protein was added to a 1 ml aliquot of the TMB substrate system (Sigma-Aldrich) and mixed well. A similar volume of H<sub>2</sub>O was added to 1 ml of TMB as a control reaction. Another reaction was set up containing 100 µg of protein and a 30 µM final concentration of the peroxidase inhibitor aminobenzyhydrazide (Sigma-Aldrich). The reactions, carried out at room temperature, were initially allowed to

continue for 2 minutes before the optical density was read at 655 nm. However, when no differences between the control, experimental and inhibited reactions were observed, the reactions were allowed to continue overnight and again monitored at OD 655 nm.

**Table 2.7: Primer list.** Names and sequences of the primers utilised.

The experiment in which primers were used is noted. Sequences are given in a 5'-3' direction. Engineered regions are denoted with small case letters. Restriction site sequences are underlined and are in lower case letters. In the "Use/s" column, each promoter swap experiment has been allocated a number: refer to key at the bottom of the table.

Oligonucleotide primer	Sequence (5'-3')	Use/s
96.04 seq rev	TCTGAGCTCGGTACCCFCCAAGGG	Sequencing ( <i>LacZ</i> )
ama 1a2	CAGTGGCTCATGTGAGTTCCAGA	SQRT-PCR ( <i>col-19</i> )
ama 1b2	CGACCTTCTTTCCATCATTTCATCGG	SQRT-PCR ( <i>col-19</i> )
ama-1 (1228/744) fwd	GCGCTGGAAATTTGACAGATTGC	SQRT-PCR ( <i>hpx-1</i> )
ama-1 (1228/744) rev	GAATACAATTGTCCACAAGAGTGC	SQRT-PCR ( <i>hpx-1</i> )
c05d11.2 fwd	ATGGGGGAGTCTCATGCTCC	RNAi
c05d11.2 rev	GGTTGCCACGTGTACATGCAGTGG	RNAi
c17g1.6 fwd	CGGATTAGTCAGTATAGTATCTACAGCG	RNAi
c17g1.6 rev	CGGTAAACTACAACATAGTCGTGCG	RNAi
c33h5.9 fwd	CTTGGTGATCAACTTGAATCCG	RNAi
c33h5.9 rev	GTTTCTTGAAGAATCGAAATGGCG	RNAi
col-19 code seq fwd 1	GAGCTATCACCCAAGACC	Sequencing (2)
col-19 code seq rev 1	GCTCGAATCCAGCTTGACA	Sequencing (2)
col-19 code seq rev 2	GTAAGCTGCACGAGATGG	Sequencing (2)
col-19 coding fwd	<u>ggtacc</u> ATGGGCAAGCTCATTGTGG ( <i>Kpn</i> I)	Cloning (2)
col-19 coding rev (no stop)	<u>ggtacc</u> GCCTTGTAAAGCTGCACG ( <i>Kpn</i> I)	Cloning (2)
col-19 fwd	ATGGCAAGCTCATTGTGGTTG	SW PCR
col-19bvf	<u>gcgctgcaga</u> TGGGCAAGCTCATTGTGGTTGG ( <i>Pst</i> I)	BV/trypsin
col-19bvr	<u>gcgtctag</u> ATTAAGCCTTGTAAAGCTGCACGA ( <i>Xba</i> I)	BV/trypsin
col-19f	ATGGGCAAGCTCATTGTGGTTG	RNAi

Oligonucleotide primer	Sequence (5'-3')	Use/s
		SW PCR
col-19r	TTAAGCCTTGTAAGCTGCACG	RNAi
<i>dpy-13</i> fwd	cggggtaccATGGACAITGACACTAAAATCAAG ( <i>Kpn</i> I)	RNAi Cloning (6)
<i>dpy-13</i> rev (no stop)	cggggtaccGGCGAGTTCCGTCCTCGAAGAAG ( <i>Kpn</i> I)	RNAi Cloning (6)
<i>dpy-5</i> fwd	gcggtaccATGGTAAAGGCCGTCGTCGGA ( <i>Kpn</i> I)	RNAi Cloning (4)
<i>dpy-5</i> rev (no stop)	gcggtaccACGCGTCTGCGCTTCTCTCTG ( <i>Kpn</i> I)	RNAi Cloning (4)
<i>dpy-7</i> fwd	cggggtaccATGGAGAAGCCAGTTCGGGGGCC ( <i>Kpn</i> I)	RNAi Cloning (5)
<i>dpy-7</i> rev (no stop)	cggggtaccTTCTTTCCATAACCACCACCAGAA ( <i>Kpn</i> I)	RNAi Cloning (5)
f08c6.1 fwd	AATGCAGATGCCTTTCTTCCG	RNAi
f08c6.1 rev	ACGTGGCTTTTCCAGGACCC	RNAi
f45g2.5 fwd	ATGTGTTTCGTGGAAGATGCCG	RNAi
f45g2.5 rev	AGTAATCCTCACAATGTGCG	RNAi
f56g4.5 fwd	ATGCCGGTAACGGAAGTTGGC	RNAi
f56g4.5 rev	CCATGAGCTCACAGACCCGGCGGACGGC	RNAi
gfpc3rev	TTCTCCTTTGCTAGCCAT	SW PCR Sequencing (1, 2, 3)
HPSRI	GGACTTAGAAGTCAGAGGCACGG	Sequencing (Ty)
M13 Reverse	AACAGCTATGACCATGATTA	Sequencing ( <i>LacZ</i> )
PAF 207 seq fwd	GCTCACTCATTAGGCAC	Sequencing (Pest)
PAF 207 seq rev	CATCACCTTACCCTCTCC	Sequencing (Pest)
pCHUCK seq fwd	CCGTCAGGGCGCGTCAGCGG	Sequencing (3)
pest vector rev (zk430.8)	gcggtaccTGGAGTCTTTCATGGTGTTC ( <i>Bam</i> HI)	Cloning Pest
pQe30 fwd	TTTGCTTTGTGAGCGGATAAC	Sequencing (HPX-1 recombinant)
PQE30R	AAGCTAGCTTGGATTCTCACC	Sequencing (HPX-1 recombinant)
prom col-19 fwd	cgccatgcCATTGAAAATTGCACC ( <i>Sph</i> I)	Cloning (1, 4, 5, 6) Sequencing (1)
prom col-19 rev	cggggtaccGTGATGAACTGATGTCTTTC ( <i>Kpn</i> I)	Cloning (1, 4, 5, 6)
semiQ PCR	ATGGGCAAGCTCATTTGTG	SQ RT-PCR ( <i>col-19</i> )

Oligonucleotide primer	Sequence (5'-3')	Use/s
col-19 fwd		
semiQ PCR col-19 rev	AGCTGGGCAACCCTCGGA	SQ RT-PCR ( <i>col-19</i> )
Seq col-19 minus 117 bp	GACCAGGACCACCAGGGCAT	Sequencing (Ty-tag, 2)
sqt-1 code fwd	<u>cgcggtacc</u> ATGTCTGTAAAACCTTGCGTG ( <i>Kpn</i> I)	Cloning (1)
sqt-1 code rev (no stop)	<u>cgcggtacc</u> ATATTTCTGTATCCACGGTTGG ( <i>Kpn</i> I)	Cloning (1,3)
sqt-1 IPS R2	TCAAGTCCTGGAACCTCCGTCG	Sequencing (1, 3)
sqt-1 ISP F1	CCAAATGGAGCCAATGGAAATG	Sequencing (3)
sqt-1 ISP F2	ATCCAATCGGAAGACCAGGAC	Sequencing (1, 3)
sqt-1 prom fwd	<u>gcgctgcag</u> GAGAACACITTTCTACGTGTCC ( <i>Pst</i> I)	Cloning (2, 3)
sqt-1 prom rev	<u>Ggtacc</u> CTCTGAAATTA'CAAATTTAGG ( <i>Sph</i> I)	Cloning (2)
sqt-1 prom seq rev	CGCAAACITTCAGTGAGATACTTGCG	Sequencing (2, 3)
sqt-1 rev	<u>gccctgag</u> TTAGATATTTCTGTATCCACG	RNAi
sqt-3 fwd	<u>cgcggccgc</u> ATGGGAAACTGACGGTAGGCTC	RNAi
sqt-3 rev	<u>gccctgag</u> TTATCGTCTGGTTCCGTCC	RNAi
sqt-fwd	<u>ccgcgccgc</u> ATGTCTGTAAAACCTTGCGTG	RNAi
T3	AATTAACCCTCACTAAAGGG	Sequencing (COL-19BV)
T7	GTAATAGACTCACTATAGGGC	Sequencing (COL-19BV)
Ty-tag antisense	<u>tatctagag</u> tcaagtggatcctggttagtatggacctctg ( <i>Xba</i> I/ <i>Kpn</i> I) taa correspond to a synthetic stop codon	Ty-tag (C-terminus) (HPLC grade)
Ty-tag <i>col-19</i> rev seq	GAGAAAGAGCATGTAGGG	Colony PCR (Ty-tag (Gly-X-Y))
Ty-tag nsi-1 antisense	<u>atgcatc</u> cgtaagtggatcctggttagtatggacctgatgcat	Ty-tag (Gly-X-Y) (HPLC grade)
Ty-tag nsi-1 sense	atgcatcagggccataactaaacttgaccaggatgcatgcat	Ty-tag (Gly-X-Y) (HPLC grade)
Ty-tag sense	<u>cagaggtc</u> ataactaaaccaggatccacttgactctagatga (( <i>Kpn</i> I)/ <i>Xba</i> I) taa is a synthetic stop codon	Ty-tag (C-terminus) (HPLC grade)
zk430.8 (183/1853) rev	GCTCTGTGAGCTTCTGAGC	SQRT-PCR ( <i>hpx-1</i> )
zk430.8 (1831/853) fwd	GTCCATCTCCCGACTAATCAACC	SQRT-PCR ( <i>hpx-1</i> )

Oligonucleotide primer	Sequence (5'-3')	Use/s
zk430.8 fwd	TTGTAAGACTTGTGACCCGG	RNAi
zk430.8 fwd exon 5	<u>gccctgcag</u> TAATGTCCATCTCCCG ( <i>Pst</i> I)	Cloning (HPX-1 recombinant)
zk430.8 perox dom seqfw	TTCACCCGCTCGCAAATG	Sequencing (HPX-1 recombinant)
zk430.8 peroxdom seq rev	CCGTCAAGAGTACGGTAC	Sequencing (HPX-1 recombinant)
zk430.8 peroxdom seqf2	GAAGTGTCCAACTGG	Sequencing (HPX-1 recombinant)
zk430.8 peroxdom seqf3	GCCGGAGATATTCGAGC	Sequencing (HPX-1 recombinant)
zk430.8 peroxdom seqR2	CTCTGTGAGCTTCTGAGC	Sequencing (HPX-1 recombinant)
zk430.8 pqc30 rev	<u>gcgaagctt</u> CTAGAGAGCTC'CC ( <i>Hin</i> D III)	Cloning (HPX-1 recombinant)
zk430.8 prom fwd	<u>cgccatgc</u> GGCAGGTGCAGGTAAAATGAAGG ( <i>Sph</i> I)	Cloning LacZ Cloning Pest
zk430.8 prom rev	<u>gcgtctagat</u> GGAGTCTTCTCATGGTGTCC ( <i>Xba</i> I)	Cloning LacZ
zk430.8 prom seq fwd	GCCTCTCTCTCTTACTTCTTGG	Sequencing ( <i>LacZ</i> )
zk430.8 prom seq fwd 2	CCCATATACTC'TTCTACA	Sequencing (Pest)
zk430.8 prom seq rev	GTCGATGCAATCGTGCTCTG	Sequencing ( <i>LacZ</i> )
zk430.8 prom seq rev 2	CACAGAACTCACTCGGTAG	Sequencing (Pest)
zk430.8 rev	ACGATGTGGTCGTTTGACGGC	RNAi
<b>Key to promoter swap clones:</b> 1) <i>col-19</i> promoter + <i>sqt-1</i> coding, 2) <i>sqt-1</i> promoter + coding, 3) <i>sqt-1</i> promoter + <i>sqt-1</i> coding, 4) <i>col-19</i> promoter + <i>dpy-5</i> coding, 5) <i>col-19</i> promoter + <i>dpy-7</i> coding, 6) <i>col-19</i> promoter + <i>dpy-13</i> coding.		

**Chapter 3**

***The exoskeleton collagen, COL-19, as an adult-specific  
marker in the C. elegans cuticle***

### 3.1 Introduction

The complexity of the ECM structure is illustrated by the fact that over 154 cuticle collagen genes have been identified in the *C. elegans* genome, a number that far outweighs the number in other organisms. Complex collagen interactions occur and the roles of individual collagens range from being essential to redundant. Markers that aid in the examination of this complex ECM structure as well as enable the dissection of the collagen synthesising machinery have long been sought after. These markers will firstly, be useful in finding the wild type expression patterns of individual collagens within the cuticle and secondly, will be tools to give detailed characterisation at the collagen level of the disruptions that occur in worms with abnormal morphology. These include worms with the typical Dpy, Lon, Sqt, Sma and Bli phenotypes that are associated with mutation of, or RNAi targeting of, certain collagens and ECM-processing genes. McMahon *et al.* (2003) characterised *dpy-2*, *dpy-5*, *dpy-7*, *dpy-10* and *dpy-13* mutants at the cuticle level and showed that morphologically similar mutants, and even alleles of the same locus, can differ in terms of their cuticle structure. This will be discussed further in chapter 4. Being able to describe phenotypes in terms of the cuticle structure and collagen expression rather than simply overall phenotype will be useful to further understand the molecular basis of the ECM and will complement and clarify data from the genome-wide RNAi screen (Kamath *et al.*, 2003).

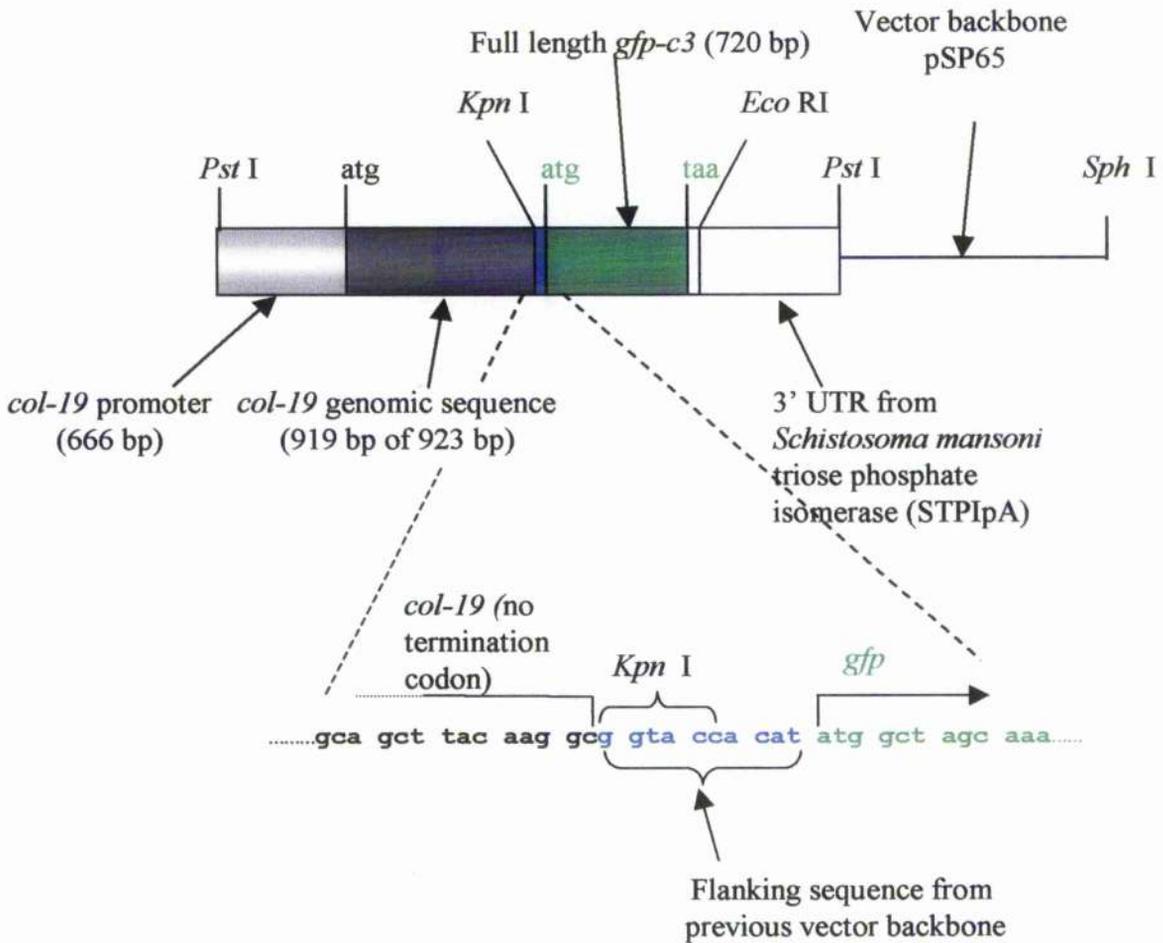
To elucidate collagen expression patterns, it was initially attempted to raise antibodies against individual collagens. However, due to the sequence similarity of collagens and their tightly folded nature that prevents access by antibodies, to date, effective and specific antibodies have only been raised against DPY-7 (McMahon *et al.*, 2003), SQT-1 (Yang and Kramer, 1999) and SQT-3 collagens (A. Page unpublished data). In light of this, epitope tagging of collagens at the N-terminus was tried. During collagen processing, collagen trimers are cleaved by N- and C-terminal proteases to remove the propeptide domains and it was thought that such cleavage would likely lead to the removal of any marker that was translationally fused. N-terminal cleavage occurs in all *C. elegans* collagens at a highly conserved subtilisin-like cleavage site while only SQT-3 has been proven to be proteolytically processed at its C-terminus (Novelli *et al.*, 2004). In Ty-tag experiments, the ten amino acid Ty epitope tags derived from the major structural protein of the *Saccharomyces cerevisiae* Ty1 virus-like particle were fused to the amino-termini of collagen genes, 3' to the proposed subtilisin-like cleavage site, and were detected using antibodies raised against Ty (Bastin *et al.*, 1996) (McMahon *et al.*, 2003).

Positioning of the Ty-tag 3' to the site of N-terminal cleavage ensured its presence in the mature peptide even after N- and C- terminal processing. Both methods described above are useful yet require fixation and immunolocalisation, which are extra steps in the protocol, and which importantly mean that live worms cannot be observed. Production of fluorescently-tagged collagen markers was therefore proposed as a way of circumventing these issues.

Although small Ty epitopes have successfully been added to the N-terminal of collagens, 3' to the cleavage site in order to avoid proteolytic cleavage, it was predicted that similar positioning of a GFP epitope would not be possible because of the mutational sensitivity of this region; mutations around this N-terminal subtilisin-like cleavage site typically lead to aberrant collagen function (Johnstone, 1994). Consequently, it was predicted that the relatively large size of the GFP epitope (240 residues) added at this position could lead to similar altered function or aberrant collagen cleavage.

Since positioning of the marker at the N-terminus was deemed unlikely to be successful, the C-terminus became a target. It was proposed that a short C-terminus could potentially reduce the susceptibility of a collagen to proteolytic cleavage, thereby preventing the loss of the marker. The collagens COL-19, ROL-6 and C39E9.9, which have short C-terminal domains (14, 25 and 15 residues respectively), were therefore chosen as potential candidates for forming GFP-tagged collagen markers (Thein *et al.*, 2003). The GFP variant, GFP-C3 (Crameri *et al.*, 1996), was cloned in-frame to the full-length coding and promoter sequences of these genes and transgenic strains carrying the constructs were obtained through microinjection. These experiments revealed that only COL-19 was able to form GFP-positive functional fusions. This is interesting because COL-19 has the shortest C-terminal domain and is the only one of the three genes whose expression is restricted to the L4-adult and L2-dauer moults as determined by promoter/reporter gene constructs (Liu *et al.*, 1995). This chapter will describe the spatial and temporal expression pattern and trimer association of COL-19 as well as the effects when it is targeted by RNAi.

The *col-19::gfp* fusion construct, BA7-1, was supplied by C. Shoemaker (AgResearch, New Zealand) and is depicted in figure 3.1. It includes 666 bp of promoter sequence and 919 bp of the genomic sequence fused to the full-length sequence of *gfp-c3*. The 919 bp of genomic *col-19* sequence is the full-length gene, minus 4 bp -the stop codon plus one extra- in order to allow in-frame read through into the *gfp*. The vector also carries a 3' UTR from *Schistosoma mansoni*



**Figure 3.1: *col-19::GFP* fusion construct of plasmid BA7.1** (not drawn to scale). The promoter region (light grey) and coding sequence (dark grey) of *col-19* (genomic) were cloned in-frame to a GFP marker (green) via *Pst* I and *Kpn* I restriction sites. The expansion of the fusion region shows how the in-frame read-through from the C-terminus of *col-19* to the *gfp* marker moiety was obtained by omitting the termination codon as well as a fourth base. The pSP65 backbone contains a 3' UTR from *Schistosoma mansoni* (white box).

triose phosphate isomerase (STP1pA), which may or may not be utilised in this construct. The backbone is pSP65 (Promega).

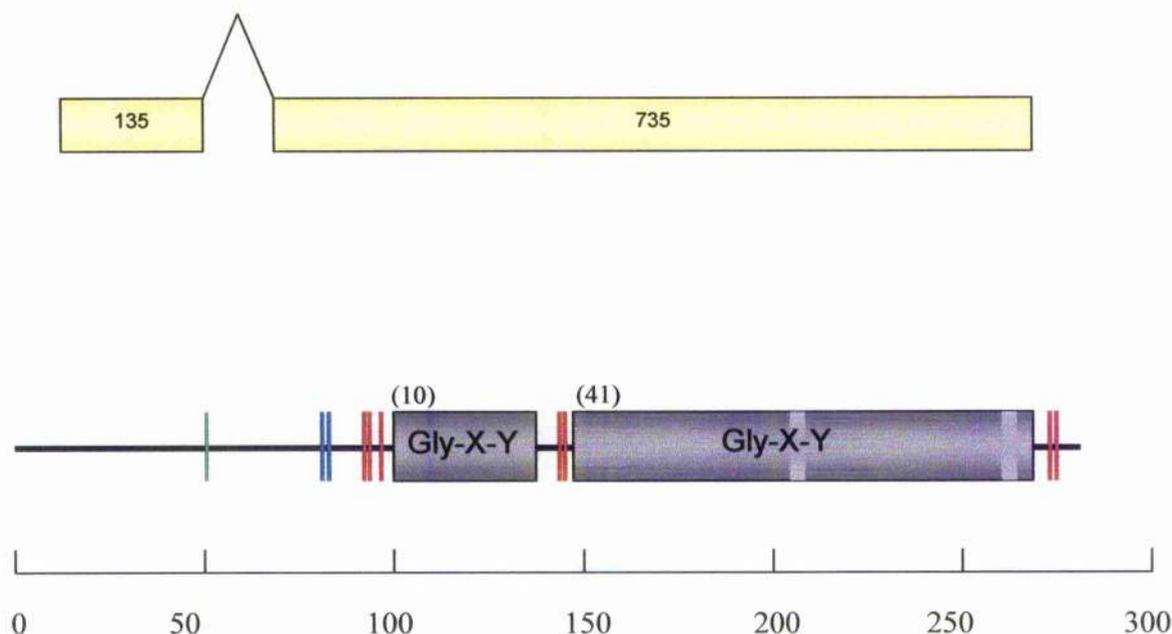
The simple gene structure of *col-19* (chromosome X) that contains only one 197 bp intron within a small 1067 bp gene is typical of *C. elegans* cuticle collagens and is shown in figure 3.2A. The encoded protein also contains features characteristic of cuticle collagens (figures 3.2 B and C). It consists of two stretches of Gly-X-Y repeats (10 and 41 repeats respectively), the second of which contains two small (3 residue long) interruptions. The pattern of conserved cysteines places COL-19 into the same collagen group as COL-8 (group I) (Johnstone, 2000). The N-termini of cuticle collagens typically have four short stretches of homology, termed homology blocks (HBA-HBD) (Kramer, 1994b). The N-terminus of COL-19 contains the conserved Arg-X-X-Arg motif in HBA at which proteolytic cleavage occurs and also contains the conserved Trp residue within HBB. However, COL-19, like COL -8, COL-35, COL-39 and DPY-7, does not exhibit homology to HBD or HBC (Kramer, 1994b). COL-19 has the shortest C-terminal non-Gly-X-Y regions of all examined *C. elegans* cuticle collagens, namely 8 residues. As stated above, the site at which processing occurs in this C-terminal, if at all, is unknown.

## 3.2 Results

### 3.2.1 The wild type pattern of COL-19::GFP in the strain TP12

Transgenic lines expressing COL-19::GFP were subjected to gamma irradiation (38 Gy) to form the integrated strain TP12:kaIs(*col-19::gfp*) as well as two additional integrated lines. From here onwards, the strain will be referred to as TP12. All lines are GFP-positive and exhibit wild type morphology, suggesting that the addition of the GFP epitope does not cause severe disruption to the cuticle or to gross body morphology.

As predicted from the adult specific expression of COL-19, the GFP-tagged form is only expressed in the adult cuticle. COL-19::GFP localises to discrete circumferential bands (figure 3.3 A) and these reflect the regularly spaced 1.2  $\mu\text{m}$  wide annulae of wild type animals. COL-19::GFP therefore appears to be localised to the cortex of these structures whilst being excluded from the annular furrows (figure 3.3 B). These bands appose the tri-laminate lateral alae, which COL-19::GFP also localises to. Staining of both these structures suggests that COL-19 is expressed in both forms of hypodermally-derived cuticle, namely that overlying the



1 MGKLIIVVGS CGVLVCVLASLYTIGNLLNEIEELQVEFQDGMVEFRAITQD

51 T<sup>W</sup>ARMVSKHINPTGHTNAPPTFETLFGARTAR<sup>Q</sup>AGFEQ<sup>C</sup>N<sup>C</sup>GPKSEG<sup>C</sup>PA

(10) (41)

101 GPPGPPGEGGQKGNPFGHDGDDGKPGAPGVI<sup>V</sup>AITHDI<sup>P</sup>GG<sup>C</sup>IK<sup>C</sup>PPGRPG

151 PKGPSGLPGSAGPAGGNGRRGPPGPVGGPGEQGPQGDAGRPGAAGRPGPA

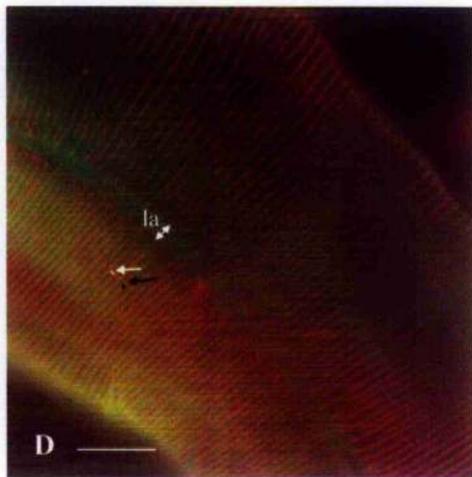
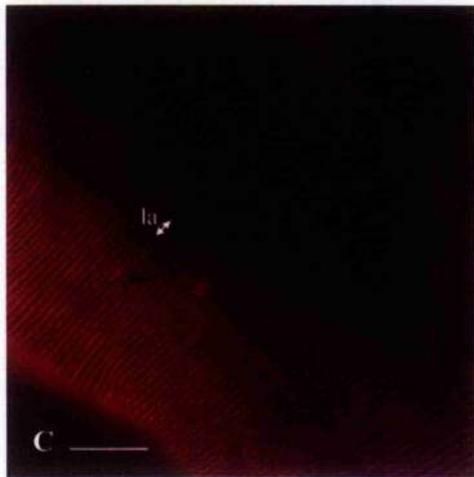
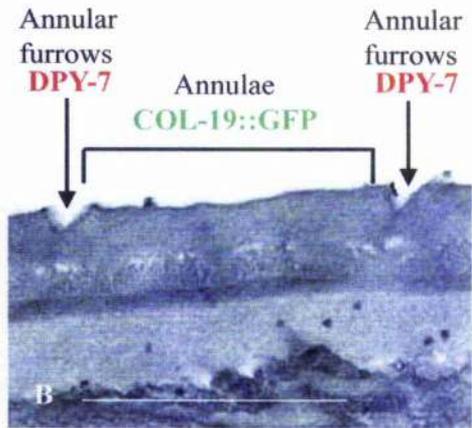
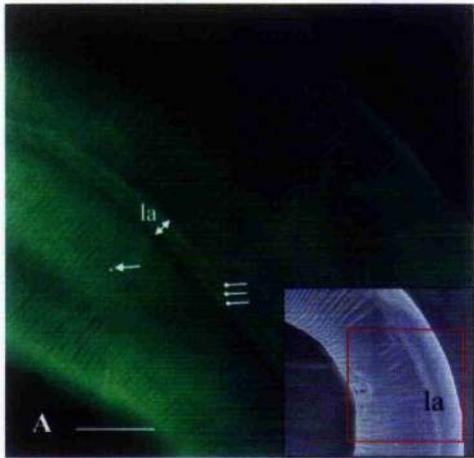
201 GPRGEPGTEYKPGQPGRPGPQGRGETGPAGNPGAPGNDGEAGKNGNAGR

250 PGPPGHPGKNGVPGQKGEDAAPGPDAGY<sup>C</sup>PC<sup>C</sup>PSRAAYKA\*

8 residues

**Figure 3.2: Gene and protein structures of *col-19*/COL-19.** A) Gene structure of *col-19*. Including the one intron, the genomic DNA is 1067 bp in length. Excluding it, the two exons result in a cDNA of 870 bp. B and C) Diagrammatic scheme (B) and amino acid sequence (C) of the COL-19 protein. Similar to other cuticle collagens, there are two blocks of Gly-X-Y repeats (10 repeats and 41 repeats respectively) (coloured/highlighted grey: numbers of repeats shown in brackets). There are two small interruptions in the second Gly-X-Y repeat region (light grey). The conserved cysteines are depicted by red lines (figure B) or highlighted in red (figure C), the two arginine residues believed to be the site of procollagen cleavage are depicted with a blue line (B) or are highlighted with blue (C) and the conserved tryptophan is depicted with a green line or highlighted in green (C). COL-19 has the shortest C-terminus of all examined collagens (8 residues).

**Figure 3.3:** A) The COL-19::GFP localisation exhibited in wild type (TP12) adult animals. An example of which section of the adult cuticle is being depicted is shown by the inset. The red box highlights a similar section on a SEM image of an adult worm in which the trilaminar alae are labelled. The annulae are not in focus. (The vulva is also observed.) Referring to the main image, the GFP-tagged COL-19 collagen localises in a regular-spaced circumferential pattern that corresponds with the annulae (single headed arrow). It also localises to the trilaminar alae (la) (blunt-ended lines). The annulae appose the alae in wild type (TP12) animals (double headed arrow). The width of the observed circumferential bands (the width of a single annulus is depicted by a small white line) suggests that the COL-19::GFP is localising to the cortex of the annulae and is excluded from the annular furrows. B) A transmission photomicrograph of a longitudinal transverse section of the adult cuticle that shows the structure of a single annulus. The region to which COL-19::GFP localises is labelled. Also depicted are the annular furrows (arrows) where COL-19::GFP is excluded. The cuticle collagen, DPY-7 localises in these furrows (arrows). C) DPY-7 immunolocalisation in an equivalent image to A shows DPY-7 localising in bands (black arrow) that are less wide than the bands of COL-19::GFP. The annular furrows can be seen to appose the alae (double headed arrow). However, the DPY-7 collagen does not localise to these trilaminar structures. D) Merged image of COL-19::GFP and DPY-7 staining shows an alternating pattern of COL-19::GFP and DPY-7 collagens and illustrates the relative positions of the collagens in the cortex of the annulae and the annular furrows respectively. The COL-19::GFP localisation is labelled with a white single headed arrow while DPY-7 localisation is labelled with a black single headed arrow. The relative widths of the two structures are shown (white line and black spot). The apposed nature of the annulae to the alae are also indicated with a double headed arrow. In A, C, and D the lateral alae are labelled 'la'. Scale bars = 10µm in A, C and D, 1.2µm in B.



ventral and dorsal hypodermal cells and that overlying the seam cells. In addition, COL-19::GFP localises to the male-specific tail circular ray and fan structures and to the hermaphrodite specific vulva (not shown).

That the GFP pattern reflects the endogenous COL-19 pattern is corroborated by a number of observations: Firstly, the adult-specific expression confirms the GFP fluorescence is specific to COL-19 expression. Further, this expression and the position of the annulae coincide and this suggests that the COL-19::GFP is indeed localising to the cuticle. Finally, similar patterns of expression are observed in immunolocalisation assays using rabbit polyclonal antibodies raised against total adult cuticle on wild type animals (Cox *et al.*, 1980) and using anti-TY antibodies to stain worms expressing the Ty-tagged and annular cortex-localised DPY-13 transgene (McMahon *et al.*, 2003). This suggests that the GFP pattern observed is one that is consistent with collagen expression.

McMahon *et al.* (2003) confirmed that two discrete and functionally distinct yet co-operating cuticle substructures exist. One group, which includes the collagens DPY-5 and DPY-13, forms the substructure that makes up the cortex of the annulae while a separate group, consisting of DPY-2, DPY-7, DPY-10 and DPY-13 localises to the annular furrows (McMahon *et al.*, 2003). The observed COL-19::GFP expression supports that COL-19 localises in a similar pattern as the former group. This was further exemplified when mouse antibodies, raised against the unique carboxy-terminal domain of the annular furrow-localised DPY-7, were used to co-stain TP12 worms (figure 3.3 C and D). As expected, these show the alternating pattern of COL-19 and DPY-7 corresponding to the annular cortex and annular furrows respectively. The DPY-7 staining observed on TP12 worms correlates exactly to that of wild type worms, and this is another indication which suggests that the addition of the GFP epitope is not disrupting the cuticle structure and that the expression observed is indicative of wild type COL-19. Tagging of COL-19 with a Ty epitope as further proof of wild type expression is discussed in section 3.2

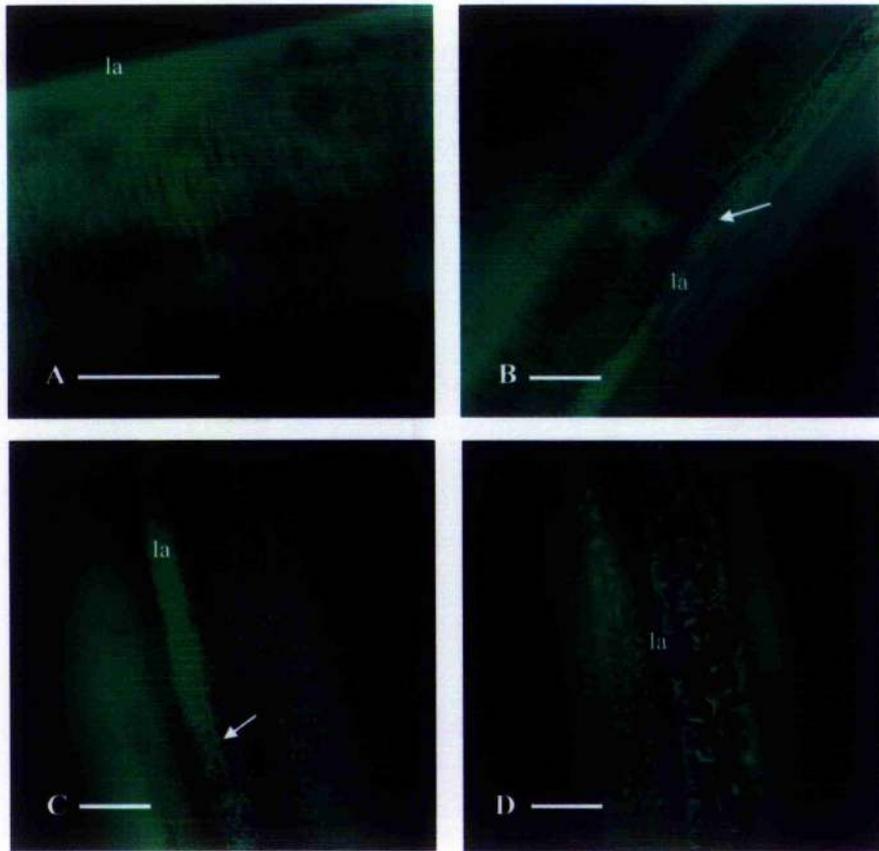
It became obvious that the pattern of COL-19::GFP expression alters as the adult worms age, first losing the defined annular pattern and, at the most extreme, appearing branched and as small wispy fragments. These changes are likely to be representative of the overall dynamic progression of the whole cuticle structure. Understanding the wild type nature of the ageing cuticle is important to distinguish between actual *bona fide* aberrations in the cuticle and the wild type behaviour of the collagen. A time course was carried out in order to characterise these

changes. 35 dauer larvae were placed onto seeded plates and monitored at 15°C, 20°C and 25°C. After an initial 12 hours during which the dauer larvae start feeding and reach the L4 moult (at 25°C), members of the synchronous population were taken at intervals over an 8-day period and their COL-19::GFP pattern observed. Figure 3.4 shows a number of stages during the time course, beginning with the L4-adult moult when COL-19::GFP expression first appears. The pattern described as the wild type TP12 COL-19::GFP pattern localising to the annular cortex and the trilaminar alae is the dominant pattern until 4 days post L4-adult moult (figure 3.4 A and B). However, at this latter stage, slight fragmentation in the collagenous cuticular structures overlying the alae becomes evident (Figure 3.4 B). After 8 days of incubation at 20°, there is very little COL-19::GFP localisation in the annulae and the dominant pattern is fragmented (figure 3.4 D). This is especially obvious in the cuticle overlying the alae. As a consequence of the reduced annular staining, the intensity of GFP fluorescence is very low, and suggests either a reduced amount of the tagged collagen in the cuticle or GFP decay. As the COL-19::GFP fluorescence decreased, there is a simultaneous increase in auto-fluorescence derived from the granules in the small intestine. Such fluorescence can be distinguished from GFP fluorescence, when looking down the viewfinder, by the more yellow, and less green, intensity.

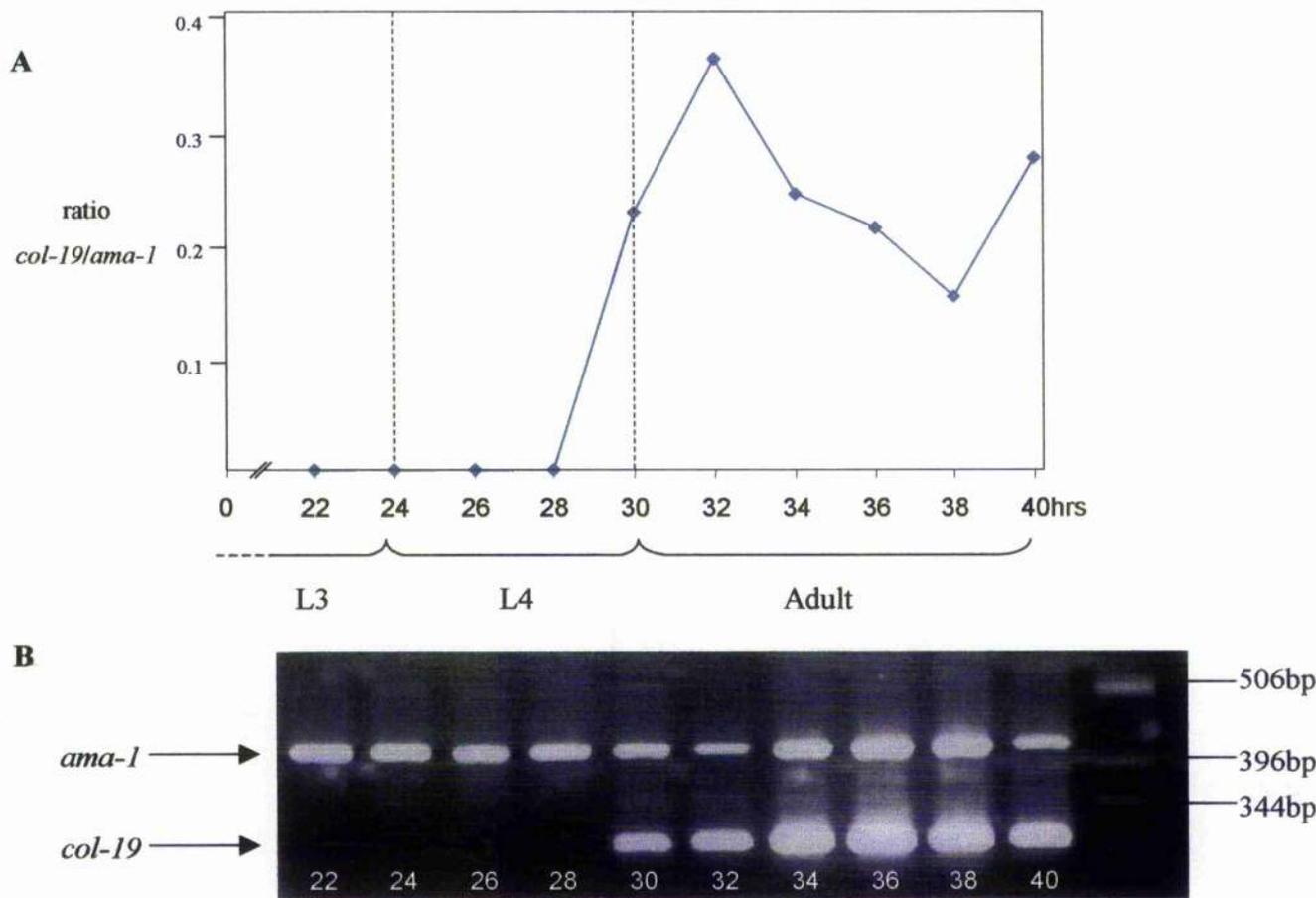
### 3.2.2 *col-19* expression coincides with *col-12*

Between hatching and the adult stage, the worm undergoes four moults during which the cuticle from the previous larval stage is shed and a new and structurally distinct cuticle is synthesised. During these moults, the expression of collagens and enzymes associated with cuticle synthesis peak to produce a tightly controlled sequence of synthesis. Collagens have been described as being expressed in one of at least four distinct waves of expression; ranging from early (4 hours prior) to late (concurrent) with the moult (Johnstone and Barry, 1996), and it is believed that the timing dictates the interactions that an individual collagen makes, and consequently substructures to which it localises (McMahon *et al.*, 2003).

A semi-quantitative reverse transcriptase PCR (SQ RT-PCR) approach (Johnstone and Barry, 1996) was used to examine exactly when, during the L4-adult moult, *col-19* is expressed. Only samples corresponding to 20-40 hours post hatch, were used in the SQ RT-PCR assays due to specificity of COL-19::GFP expression to the L4-adult moult. Of these, sample 24 represents the L3-L4 moult and sample 30 represents the L4-adult moult. The results from SQ RT-PCR analysis, presented in figure 3.5, clearly show that there is no expression of *col-19* between 20 and 28 hours post hatching and confirm the lack of expression in L4 worms. After this, and



**Figure 3.4: The changing expression of COL-19::GFP in aging adult TP12 worms (at 20°C).** Patterns observed 2 days (A), 4 days (B), 6 days (C) and 8 days (D) after the L4-adult moult. Position of lateral alae are marked 'la'. As adults age COL-19::GFP localisation around the alae becomes increasingly fragmented (white arrows in figures B and C). In addition, the regular annular staining becomes less evident and instead becomes fragmented (black arrows in figures C and D). In accordance with a decreased amount of COL-19::GFP in the cortex of the annulae, the fluorescence decreases. However, any reduction in GFP intensity has been compensated for and thus the relative levels of fluorescence cannot be compared (this is particularly true for figures C and D). Scale bars = 10µm.



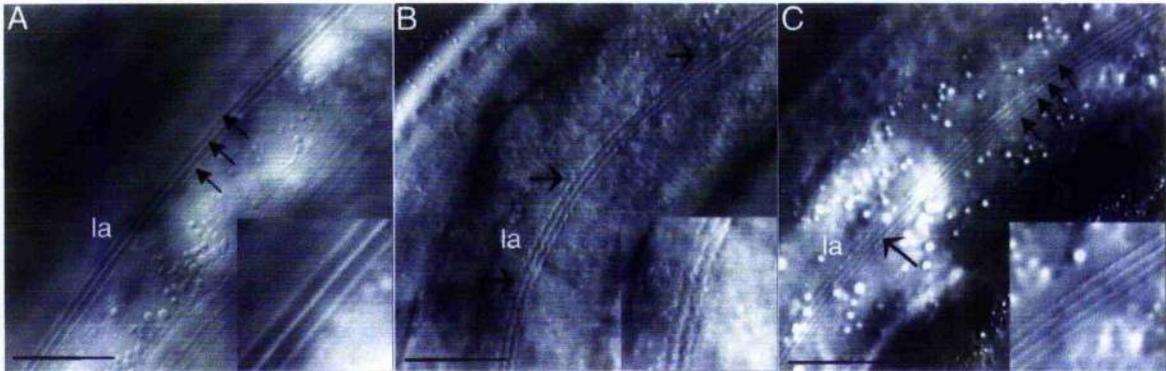
**Figure 3.5:** Temporal expression of *col-19* via a SQ-RT-PCR assay. **A)** Graph showing the expression levels of *col-19* during the course of the *C. elegans* life cycle expressed as a ratio of *ama-1* expression. The time of the larval moults (L3-L4 and L4-Adult) are indicated with the dotted lines. *col-19* expression is initiated 30 hours post hatch which coincides with the L4-Adult moult. The x-values on the graph represent the number of hours after the life cycle is restarted by returning L1 arrested larvae to OP50-containing plates. The y-axis values show the expression level of *col-19* being expressed in terms of its ratio to the constitutively expressed *ama-1*. **B)** The corresponding agarose gel showing *col-19* and *ama-1* (control) products from the SQ-RT-PCR from staged cDNA (hours post moult are indicated at the bottom).

corresponding to the time of the L4-adult moult, expression is initiated. *col-19* expression appears to peak at 32 hours, which is unique from the previously assayed collagens that all exhibit troughs in expression at 32 hours (Johnstone and Barry, 1996; McMahon *et al.*, 2003). However, experiments that detailed collagen expression at 15-minute intervals during a moult (Johnstone and Barry, 1996) showed that the slope of the graph does not necessarily represent for induction of expression. Instead, it appears that the slope is largely the result of asynchrony in the culture and that real induction in any one animal is virtually instantaneous, occurring within a 15-minute interval (Johnstone and Barry, 1996). In light of this, COL-19 expression can probably be assumed to occur 30 hours after hatching, when its expression first appears. This coincides with 'late' expressing collagen COL-12, which is described as occurring concurrently with the moult (Johnstone and Barry, 1996).

### **3.2.3 COL-19 is non essential but its RNAi-mediated depletion leads to adult-specific alae defects**

The removal of many collagens, such as DPY-3, DPY-4, DPY-5, DPY-7, DPY-8, DPY-9, DPY-10, DPY-13 and LON-3 via RNAi or in null alleles, is sufficient to disrupt the cuticle and give abnormal phenotypes (reviewed in Page and Winter, 2003). In order to test whether the removal of COL-19 had a similar effect on the cuticle, an RNA interference (RNAi) approach was applied to N2 and TP12 worms.

When COL-19 is targeted by RNAi in N2 and TP12 worms, no severe gross morphological effects are observed. This suggests that this collagen is not essential for a properly formed adult cuticle. It is interesting to note, therefore, that RNAi targeting of *col-19* results in subtle alae defects, as illustrated in figure 3.6, since alae are specific to the adult, dauer and L1 stage cuticles. Instead of the normal wild-type triple-track alae structure, RNAi-treated animals present alae that are still evenly spaced (i.e. not branched) but are discontinuous in areas. Animals injected with *col-19* dsRNA exhibit extra alae compared to wild type animals. Annulae and annular furrows in RNAi-treated animals (all methods) are wild type in appearance. There are some cases in which certain genes are less susceptible to RNAi, in particular it has been observed that genes in the mature nervous system exhibit partial or complete resistance to dsRNA-mediated silencing (Timmons *et al.*, 2001). However, confirmation that the COL-19 signal was effectively reduced in these experiments came from the loss of GFP fluorescence in TP12 animals.



**Figure 3.6: Nomarski images of alae in untreated TP12 animals and in animals targeted by *col-19* RNAi.** Insets show 2x magnifications of alae. Alae are labelled 'la'. A) Wild type alae are trilaminar and unbroken (arrows depict each alae). B) Administering RNAi by feeding results in alae that exhibit numerous breaks (arrows). C) Administering RNAi by injection is more potent and in addition to alae breaks (large arrow), results in multiple alae (4 instead of 3) (each marked with a small arrow).

Although, this data suggests that COL-19 is not essential for maintaining gross morphology, COL-19 may play a role in the proper formation of the alae. COL-19 is indeed expressed in the alae, as GFP localisation exhibits. Although COL-19::GFP localises to the annulae, it appears that COL-19 is not an essential component of these structures. This conclusion was inferred from the fact that RNAi targeting this collagen does not lead to defects in the annulae.

### 3.2.4 Confirming the wild-type expression of COL-19

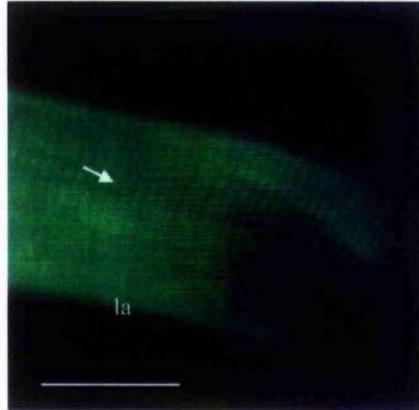
GFP is a 239-residue moiety, which is relatively large in terms of the small *C. elegans* cuticle collagens. It is therefore plausible that the GFP marker that has been fused to COL-19 in the strain TP12, has a disrupting effect on the collagen structure and consequently its localisation in the cuticle. In such a case the fluorescent pattern observed in TP12 worms would not be a reflection of the wild type localisation of COL-19. To resolve this anomaly, a COL-19-Ty-tag fusion was constructed.

#### 3.2.4.1 Immunofluorescence of COL-19

A Ty-tag fusion method was chosen after unsuccessful attempts were made at whole worm immunolocalisation using antibodies raised against the C-terminal, non-Gly-X-Y region of COL-19. The inability to successfully immunologically label worms in this way was not surprising because of the lack of access for antibodies due to the tightly folded nature of processed collagens, and the lack of specificity of collagen antibodies due to the high level of homology. The main contributor to the failure of the immunolocalisation assays was the former explanation, as it was later shown that COL-19 antibodies are specific for the collagen (section 3.2.5).

#### 3.2.4.2 Tagging the C-terminus of COL-19 with a Ty-tag

The Ty-tag epitope, being 10 residues in size (Bastin *et al.*, 1996), is much smaller than GFP, and is therefore potentially less of a steric hindrance. A construct encoding a Ty-tagged COL-19 was produced as described in section 2.31 in which a 30 bp Ty-tag moiety was cloned in-frame to the C-terminus of the coding sequence of *col-19*. Two transgenic lines expressing this construct were obtained via microinjection, and Unc rescue adult worms were subjected to immunolocalisation using anti-Ty antibodies. The results are shown in figure 3.7. This clearly depicts COL-19-TY localising to wide circumferential bands that correspond with the annulae. COL-19::GFP and COL-19::TY thus display similar expression patterns. Since the Ty moiety



**Figure 3.7: Expression pattern of Ty-tagged COL-19 in the cuticle of adult worms.** Immunolocalisations using anti-Ty antibodies. The pattern of annular staining is observed which is similar to the COL-19::GFP pattern of the strain TP12 presented in figure 3.3A. Annulae are labelled with an arrow and lateral alae are denoted “la” (but alae are not in the plane of focus). Scale bar = 10 $\mu$ m.

represents the wild type deposition of COL-19, this data suggests that the pattern of expression observed for COL-19::GFP in TP12 is reflective of the wild type collagen. Staining was only observed in adult worms, which provides proof that anti-Ty antibodies were specific to the adult-expressed COL-19.

#### 3.2.4.3 Inserting the Ty-tag to interfere with the Gly-X-Y repeat region

The construct described above was designed in order that the Ty-coding epitope could be removed by restriction digestion, leaving an in-frame stop codon at the end of the *col-19* sequence. New sets of sense and anti-sense primers were designed that could be inserted into the coding sequence of COL-19 in order that the effects of disrupting the Gly-X-Y repeat region of the collagen could be observed.

Data obtained from such experiments, together with the *col-19* RNAi data, had the potential to be used to explore the role of COL-19 in the cuticle. RNAi data suggests that COL-19 is not essential for the proper formation of the cuticle. However, many morphological mutants arise from mutations in the cuticle collagens, particularly, mutations in the Gly-X-Y motif, which are vital for the proper helical folding of the collagen (Johnstone, 2000; Page and Winter, 2003). Since such alleles are not available for *col-19*, we alternatively disrupted the Gly-X-Y repeat region through the addition of a Ty-tag moiety. This method is preferable to simple site-directed mutagenesis of the gene since it provides a method to visualise the location of the collagen.

Constructs were made as described in section 2.32.2 in which the Ty-tag epitope was cloned in-frame at a site within the coding sequence of the collagen. There were only two restriction sites (*Nsi* I and *Pml* I) that were identified into which the Ty-tag could possibly be cloned. An unsuccessful attempt was made to ligate the marker moiety into the *Pml* I site. However, a clone in which the Ty-tag was ligated into the *Nsi* I site was made and used to generate transformed lines via microinjection. Transformed animals were identified on the basis of the rescue of the *unc-76* phenotype, as previously described. A single line was created in which worms were rescued. Animals in this line show wild type body morphology and thus from preliminary data, it appears that the integration of the Ty-tag is not disruptive. Despite the *Unc* rescue phenotype, when immunological assays were attempted using Ty-tag antibodies, no staining was observed. The ability to analyse the reasons for such a failure for antibodies to bind, is made difficult by the fact that only one line was created. However, a few preliminary suggestions may be formed to explain the lack of staining. One possibility is that only the marker plasmid was integrated

into extra chromosomal array or alternatively, that there could be a high level of mosaicism. It must also be considered, however, that one possibility is that the Ty-tagged COL-19 is being excluded from the cuticle by the protein sorting machinery, which recognises the collagen as being aberrant. It is known from RNAi studies that the removal of the collagen is not severely detrimental to cuticle formation but results in alae defects. No obvious alae defects were observed from initial analyses. Despite being positioned at a site within the region of the coding sequence of *col-19*, the *Nsi* I site actually falls within the interrupted region separating the two Gly-X-Y regions and consequently is not positioned at a site within a helical region. This reduces its potential to be disruptive to COL-19 folding or the cuticle integrity. The most important step to begin elucidating the effects of the Ty-tag in this construct would be to create more transgenic lines, which could be compared.

### 3.2.5 Trypsin digestion of recombinant COL-19 from a BV expression system

The specific nature of interactions between collagens is still unclear. To date, there has been little evidence as to whether nematode collagen triple helix units exist as homotrimers or heterotrimers, or both. It has been suggested that similarly temporally expressed collagens might act as interacting partners (McMahon *et al.*, 2003); however, there has been no further evidence to prove this. Early data investigating the *sqt-1* and *rol-6* collagen-encoding genes, which are simultaneously expressed, pointed to direct interactions of these collagens because of the genetic interactions displayed. Their sequence similarity made their interaction within triple helices probable, since glycine substitution-induced disruption of trimer association suggested that heterotrimerisation could only occur between the most similar collagens (Johnstone, 1994; Page and Winter, 2003). However, direct interactions between these collagens were dismissed after studies, using SDS PAGE analysis, showed that in *rol-6* N-terminal processing mutants (*rol-6(su1006)*), multiple oligomer bands containing SQT-1, were not altered in size. Recent studies using the baculovirus expression system have also confirmed this because expression of SQT-1 and ROL-6 in different combinations failed to produce functional trimers (A. Page, unpublished data).

The baculovirus system (Crossen and Gruenwald, 1998) is a method that facilitates the high-level co-expression of multiple proteins via co-infection by more than one recombinant virus. Heterologous collagen genes, under the influence of a strong late viral promoter, can be co-expressed with the subunits of recombinant P4H in order that they be hydroxylated, an

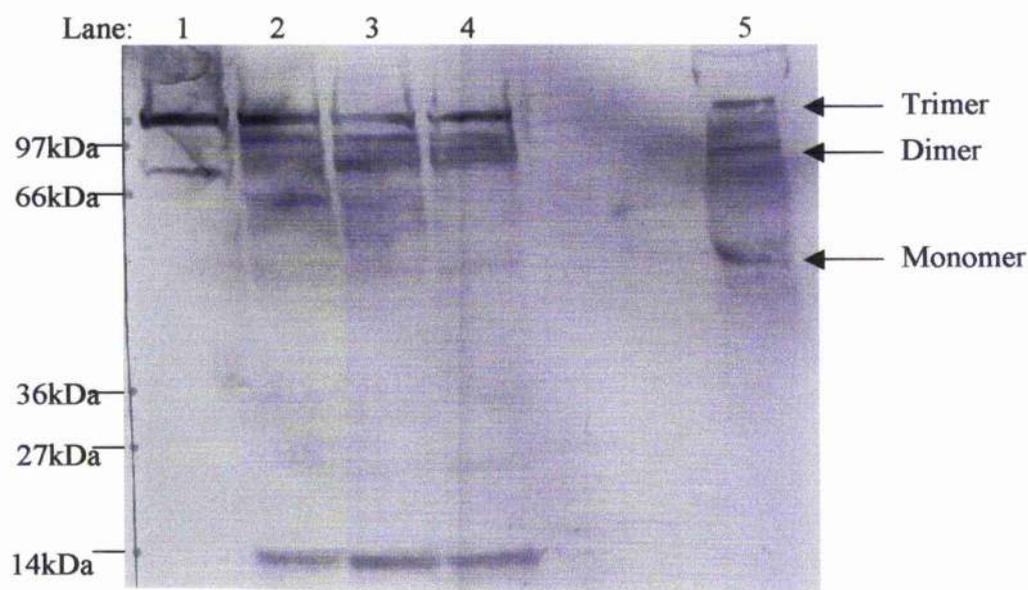
essential step for collagen folding. Most recombinant systems do not have sufficient amounts of P4H activity.

Aliquots of both Triton X-100 and SDS extracts prepared by A. Winter were subjected to 2-, 4-, and 8- minutes of trypsin digestion. These trypsin treated samples were run on a 4-12% gradient acrylamide gel alongside samples of non-trypsin treated extracts in reducing and non-reducing loading buffers. After transfer onto PVDF membranes, western blotting using antibodies raised against COL-19, was performed. The resulting gel from triton X100-extracted samples is shown in figure 3.8.

Samples were loaded in both reducing and non-reducing conditions in order to obtain evidence of the presence of reducible (e.g. disulphide) and non-reducible (e.g. tyrosine-derived) bonds. The predicted size of the protein encoded by the 870 bp *col-19*, is around 31.9kDa; however, the nature of the gel (and probably glycosylations) result in the COL-19 running higher on the gel (~55kDa). A band this size is evident in lane 5, the reduced 2-mercaptoethanol sample, where it is marked 'monomer'. This appears as one of the three bands resulting from such reduction in which the S-S bonds have been chemically broken. Confirmed by the predicted sizes: slightly larger than 31.9, 63.8 and 95.7 kDa. These correspond to the monomer, dimer and trimer forms respectively. The fact that a trimer form is still evident upon 2-mercaptoethanol-mediated reduction is evidence that the trimer is cross-linked by non-reducible bonds.

The trimer band is more apparent when the sample is non-reduced (compare lane 5 and lane 1) which suggests that the presence of additional reducible bonds in the trimer. A dimer band is also observed in lane 1. Lanes 2, 3 and 4 show non-reduced samples that have been subjected to trypsin-mediated collagenase activity for 2-, 4-, and 8- minutes respectively. Due to the maintenance of the trimer band even after 8 minutes of treatment, these results show that the tightly folded nature of COL-19 within a triple helix makes it resistant to trypsin.

These results show that COL-19 is able to form homotrimers that are resistant to reduction. This is very interesting since the only collagen to date that exhibits a similar pattern is SQT-3 (A. Page, unpublished data).



**Figure 3.8: Western analysis of BV-expressed recombinant COL-19 after trypsin treatment and under reducing and non-reducing conditions.** Triton-X100 extracts were trypsin-treated and run on a native PAGE gel alongside non-trypsin treated samples. Transferred onto a membrane and visualised using anti-COL-19 antibodies, lanes 1-4 represent non-reduced bands while lane 5 is run under reducing conditions. Three bands are evident in this reduced lane (lane 5), representing the monomer, dimer and trimer forms of COL-19. In samples not treated with trypsin (lane 1), the dominant band is that of the trimer. A second band representing the dimer is also observed. Upon trypsin digestion for 2, 4 and 8 minutes (represented by lanes 2,3 and 4 respectively), the trimer band remains as the dominant band. However, the dimer band disappears and a faint monomer band appears.

### 3.2.6 SQT-1 cannot form a functional GFP fusion

Since there is no direct evidence, to date, to show that proteolytic cleavage to remove the C-terminal pro-collagen actually occurs, SQT-1, a collagen that is expressed in all larval and adult stages was chosen as a target of a GFP marker epitope. A *sqt-1::gfp* construct was cloned through PCR amplification of, firstly, 1000 bp of the *sqt-1* promoter and secondly 1018 bp of the full length coding sequence (genomic DNA) that omitted the last 4 bases in order to remove the termination codon and facilitate in-frame read through to the GFP moiety. These were ligated into the BA7-1 vector, from which the *col-19* promoter and coding sequence had been digested out. This construct was injected at 20µg/ml into DR96 worms along with rescue plasmid p76-16B at 100µg/ml. Successfully transformed worms were chosen on the basis of Unc rescue. Two lines were obtained, none of which exhibited fluorescence, even when viewed at high power. This indicates that the GFP moiety may be cleaved from SQT-1 and therefore this is preliminary evidence to demonstrate that SQT-1 does appear to be processed at its C-terminus. It also shows that the processing events of COL-19 and SQT-1 are distinct.

### 3.2.7 Investigating the role of collagen promoters

COL-19::GFP is the first example of a functional collagen-GFP fusion and two postulations have been proposed as to what facilitates the apparent failure of the collagen to be cleaved at its C-terminal, thus allowing the read through to the GFP epitope. Firstly the short C-terminus of the collagen could potentially reduce the chances of proteolytic cleavage of its pro-collagen domain. Secondly, it is possible that COL-19 is not cleaved due to adult-specific processing during which the extent of proteolytic cleavage is altered. The successful cloning of a functional GFP-fusion with COL-7 (Gillian McCormack, WCMP, University of Glasgow unpublished results), another adult specific collagen, has recently strengthened the latter hypothesis. COL-7 has a longer C-terminus than COL-19, being 44 residues long, and localises to the annular furrows when fused to the GFP marker. COL-7, similar to COL-19, is expressed coincident with the moult.

Studies were carried out to dissect the relative influences of the promoter and the collagen sequence in directing temporal and spatial expression and controlling the interactions that a collagen makes. Two different kinds of constructs were made. Firstly, a number of collagens were placed under the control of the *col-19* promoter and, secondly, and conversely, *col-19* was placed under the influence of the *sqt-1* promoter (figure 3.9). The aim of these constructs was to

both identify the conditions under which GFP can be fused to a collagen and characterise any expression pattern resulting from the fusion.

### 3.2.7.1 Collagen fusions under the influence of the influence of the *col-19* promoter

Regions of 1018 bp, 1056 bp, 851 bp, 1013 bp, 1132 bp respectively of coding sequence of *sqt-1*, *sqt-3*, *dpy-5*, *dpy-7* and *dpy-13* (minus the last 4 bp) were cloned into a prepared BA7.1 vector. Between 1 and 3 transformed lines were obtained via microinjection for all except the *sqt-3*-containing clones, and were tested for the formation of functional fluorescent fusions. Figure 3.9 depicts the constructs made and whether GFP expression was observed in the resulting transgenic lines.

Constructs with *dpy-5*, *dpy-7* and *dpy-13* under the influence of the *col-19* promoter failed to produce functional GFP fusions (figure 3.9 A). It appears that these collagens are not able to form functional fusions, even when their expression is dictated by the *col-19* promoter.

Interestingly, however, when *sqt-1* was placed under the influence of the *col-19* promoter, a functional fusion resulted. The pattern of expression is shown in figure 3.9. As expected from the *col-19* promoter, this expression was observed only in adult worms. The fusion protein expression appears to be to localise to the annular furrows.

### 3.2.7.2 *col-19* under the influence of the *sqt-1* promoter

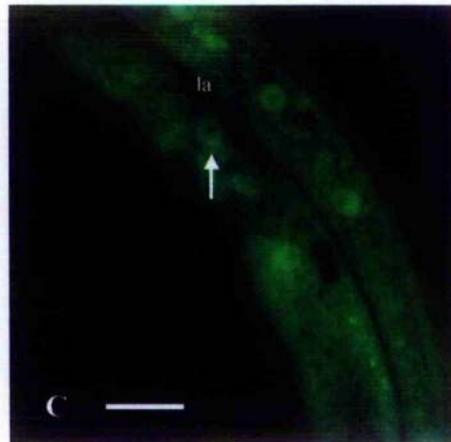
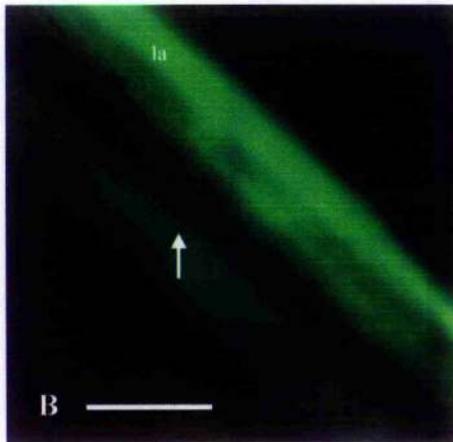
*col-19* was placed under the influence of the *sqt-1* or *sqt-3* promoters by removing the promoter and coding sequence of *col-19* from BA7.1 via restriction digest and ligating firstly, the promoter of either *sqt-1* or *sqt-3* (1000 bp and 1133 bp respectively) and secondly, 866 bp of *col-19* coding sequence (i.e minus the last 4 bp, to allow for read through into the GFP moiety), into the vector. The promoters were cloned into the *Pst* I-*Kpn* I site of the BA7.1 vector and the coding sequences were cloned in between two *Kpn* I sites. Transgenic lines carrying the *sqt-1*-containing construct were obtained and tested for fluorescence.

*col-19* under the influence of the *sqt-1* promoter forms a functional fusion whose fluorescence is observed in larval and adult animals. However, it appears that the fusion protein is not being fully processed into the cuticle and is instead being localised non-specifically, appearing as if it has 'leaked into' the body cavity of the animal (figure 3.9).

**3.9: Results from promoter swap experiments.** A) Combinations of collagen promoters and coding sequences that were fused. Tick mark denotes a functional GFP fusion was made, a cross denotes no GFP fluorescence was observed in transformed lines. A horizontal line denotes that these were not attempted and NA represents that the constructs have been made but have not yet been analysed. B) GFP localisation observed when *sqt-1* is under the influence of the *col-19* promoter. An annular staining pattern is observed as denoted with the arrow. Staining is also observed in the alae ('la') but these are not in the plane of focus. C) Staining observed when *col-19* coding sequence is placed under the *sqt-1* promoter. Non-specific localisation can be seen. GFP appears to accumulate in nuclei which are denoted with an arrow. Scale bars = 10  $\mu$ m

Facing page 138

A		Promoter		
		<i>col-19</i>	<i>sqt-1</i>	<i>sqt-3</i>
Coding sequence	<i>col-19</i>	✓	✓	NA
	<i>sqt-1</i>	✓	X	—
	<i>dpy-5/7/13</i>	X	—	—
	<i>sqt-3</i>	NA	—	—



Some annular patterning is observed, although it is difficult to establish whether this is annular furrow- or cortex- localised. It is also not possible to confirm that this is a result of the depositing of the fusion moiety into the cuticle or whether it is from non-specific localisation around the annulae/annular furrows. Further investigation is required.

The fact that COL-19 can form a GFP fusion under these conditions provides evidence that even when expressed outside of the L4-adult moult, this collagen is apparently not enzymatically cleaved at its C-terminus. The similar phenomenon observed in TP12 cannot solely be attributed to the specificity of adult processing.

### 3.3 Chapter 3 Discussion

#### 3.3.1 COL-19::GFP as a *C. elegans* cuticular marker

The *C. elegans* cuticle collagen, COL-19, has been investigated for its ability to form a functional translational fusion with a GFP marker moiety. Herein is described the wild type expression pattern of the tagged collagen which localises to the annular cortex in young adults and which becomes more disrupted as the worm ages. The temporal expression of this collagen has been investigated through a semi quantitative reverse transcription (SQ RT-PCR) approach and has confirmed that *col-19* expression is limited to the L4-adult moult and more specifically, that this expression begins at a time coincident with the moult. This is in comparison to collagens such as *dpy-7* and *dpy-5*, for example, whose expressions peak at 4- and 2- hours prior to the moult respectively. When *col-19* was targeted by RNAi, no gross morphological abnormalities were observed apart from alae being broken and sometimes multiple. The phenotypes were dependent on the method of administration with microinjection of dsRNA being more potent than feeding animals on bacterial lawns expressing dsRNA or soaking in a solution of dsRNA. By fusing a Ty-tag moiety to COL-19 and immunologically determining its localisation in whole worms, it has also been demonstrated here that the expression of COL-19::GFP reflects that of the wild type collagen. Preliminary attempts were made to target the coding sequence of the collagen with a moiety that was potentially disruptive to collagen folding and to the cuticle structure.

A number of properties of COL-19 make its GFP-tagged form an ideal candidate as a marker for the analysis of the *C. elegans* cuticle. Foremost is the fact that this collagen is actually able to form a translational fusion. Three collagens, ROL-6, COL-132 (Thein *et al.*, 2003) and SQT-1

(this study) have been analysed for their ability to do the same but have proven unsuccessful. Proteolytic removal of the C- and N- prodomains from collagen trimers has been demonstrated to be an essential step during vertebrate collagen biosynthesis since it removes the globular termini. Proof that similar steps occur at the N-terminus of *C. elegans* collagens has come from morphological mutants with lesions in either N-terminal proteases (BLI-4) (Thacker *et al.*, 1995) or in the N-terminal collagen subtilisin-like proteolytic cleavage site (Yang and Kramer, 1994), which is conserved among all *C. elegans* collagens. C-terminal processing was recently identified as being an essential step during the processing of SQT-3 (Novelli *et al.*, 2004) but the extent of processing of other collagens is still unknown. However, the inability to form functional translational fusions was attributed to C-terminal proteolytic-mediated removal of the marker. The C-terminal non-Gly-X-Y domains of *C. elegans* range from 14 to 400 residues in COL-19 and BLI-1 respectively and therefore COL-19 was chosen as a potential target on the basis that its short C-terminal domain could reduce the chances of its proteolytic cleavage. In fact, the potential region where cleavage of this collagen could occur is only 8 residues if the essential Cys residues that are required for intermolecular disulphide bond formation are taken into account: because they are essential, proteolytic cleavage would be required to occur N-terminal to them. The ability of COL-19::GFP to express the marker moiety suggests that cleavage of this collagen does not occur. Another distinguishing feature of *col-19* is its adult-specific expression pattern, which may facilitate the decreased degree of processing.

COL-19 is typical for *C. elegans* collagens: within the relatively short collagen molecule two Gly-X-Y repeats regions are encoded. The first region is 10 repeats long and the second has 41 repeats. As with other collagens, the larger Gly-X-Y region has a number of small interruptions; in the case of COL-19, there are two. The conserved pattern of cysteine residues places it in a group of collagens with COL-8. It also encodes the RXXR subtilisin-like proteolytic cleavage site. 919 bp of cDNA sequence encoding for this collagen as well as 666 bp of 5' sequence was fused in-frame to a GFP variant, GFP-C3. Exclusion of the last 4 bp facilitated the in-frame read-through into the marker moiety. Despite 235 bp of promoter sequence being sufficient to drive *col-19* expression, 666 bp were used because of the likelihood that elements controlling expression levels were contained in the more 5' region (Liu *et al.*, 1995).

The COL-19::GFP construct was integrated to form the strain, TP12. Such a strain is more useful than transgenic strains because of its ease of maintenance, the ability to perform crosses and the 100% penetrance (compared to ~20% of transgenic lines). An integrated line is also

more advantageous than the commonly utilised immunolocalisation using antibodies raised against the cuticle collagen DPY-7 since live worms can be visualised and there is no requirement for multiple processes. The three integrated strains obtained and the transgenic lines all expressed similar patterns of COL-19::GFP. No changes in gross morphology were observed, the first line of evidence to suggest that the translational fusion is not disruptive to the cuticle.

### 3.3.2 COL-19::GFP localisation represents that of wild type COL-19

In wild type (TP12) animals, COL-19::GFP localises to the annular cortex (figure 3.3 A) and is observed in adult animals only, the latter observation being in accordance with the adult-specific expression that was demonstrated using SQ RT-PCR (figure 3.5). The exclusion of COL-19::GFP from the annular furrows was confirmed by immunolocalisation using antibodies raised against the furrow localised collagen, DPY-7 (figure 3.3 C). As was expected, an alternating pattern of COL-19::GFP localising to the annular cortex and DPY-7 in the furrows was observed (figure 3.3 D). Each of these substructures has previously been demonstrated to have different collagenous compositions with individual collagens being sequestered to either one or the other. It is no surprise therefore, that COL-19::GFP is excluded from the furrows. The pattern of DPY-7 observed in the cuticle of TP12 animals in these experiments was the same as that of wild type animals, further evidence that no cuticle disruptions occur as a consequence of the introduction of the relatively large GFP moiety fusion construct. Concrete evidence that COL-19::GFP is reflective of the wild type expression of COL-19 came from the fact that identical patterns of expression were observed when a Ty-moiety, which is only 10 residues in size and of significantly less steric hindrance, was fused at the same site in the C-terminus as GFP, and visualised via immunolocalisation using anti-Ty antibodies.

The changing cuticle composition that accompanies aging of adult worms, has also been demonstrated here. The COL-19::GFP pattern becomes increasingly more disrupted, with the first aberrations being fragmented strands which appear around the alae. The dorsoventral region then progressively becomes more fragmented. The disruptions observed may be a consequence of the growth of the animal and the requirement for the cuticle to stretch in order to accommodate it. Between the larval moults, the growth of the animal proceeds in linear fashion, with each subsequent larval stage having increasing rates (Knight *et al.*, 2002). Growth during the adult stage is implied by the fact that *col-19* (from the SQ RT-PCR data, presented here) and other collagens, such as *col-12* and *sqt-1* (Johnstone and Barry, 1996), and collagen modifying enzymes, e.g. *pdi-2* (Winter and Page, 2000), are expressed after the final moult.

### 3.3.3 The role of COL-19 in the cuticle

A SQ RT-PCR approach confirmed previous data that *col-19* expression is limited to the L4-adult moult. Such expression is also illustrated by the adult specific expression of the COL-19::GFP marker. *col-19* has previously been shown to be expressed in dauer cuticles as well (Cox and Hirsh, 1985), but no COL-19::GFP expression was observed in starved populations or when TP12 was crossed with a dauer-inducing strain (*daf-2(e1370)*). The absence of this collagen in larval stages shows that this it is not essential for their cuticles. The essential nature of this collagen in adult cuticles is less obvious. However, the lack of a gross morphological RNAi phenotype described here indicates that the cuticle is able to form in the absence of the collagen; the effective removal of the collagen by this method was confirmed by the loss of the fluorescence when the technique was performed on TP12 strain. Defects in the trilaminar alae were the only aberrations observed in RNAi-treated animals and these were in the form of broken and/or multiple alae, dependent on the mode of RNAi administration, micro injection of ds RNA being the most potent. This suggests that a role of COL-19 may be in the formation of the alae, which is consistent with the adult- and dauer- specific expression of COL-19 since these (and the L1-stage) are the only stages at which alae are present. Components of the alae (e.g. CUT-1) (Sebastiano *et al.*, 1991) are likely to be the substrates of extensive non-reducible cross-linking. By the resistance of baculovirus-expressed COL-19 trimers to reducing conditions, COL-19 has also been shown to be cross-linked by similar non-reducible bonds. However, COL-19 is by no means exclusive to *C. elegans* collagens in being cross-linked and these bonds must be prevalent due to the resistance of the cuticle to reducing agents and the fact that collagens exhibiting lesions in cross-linking residues produce morphologically mutant phenotypes (e.g. *sqt-1(sc13)*) (Yang and Kramer, 1999).

The passive effect of introducing the COL-19::GFP moiety into the cuticle is an insight of the role of this collagen. Many of the cuticle structures are mutationally sensitive, with small aberrations being able to produce severe defects in the cuticle. This is exemplified best by improperly N-terminally processed SQT-1 molecules (e.g. *sqt-1(e1350)*) that impart huge cuticular disruptions. In contrast, no disruption to gross morphology or to the annulae is observed, even with the addition of the much larger GFP moiety onto the C-terminus of COL-19. The wild type phenotype implies that structurally important collagens such as DPY-7, DPY-5 and SQT-1 are localised normally and therefore that COL-19 has little influence on the rest of the components of the cuticle. The lack of influence of this collagen on the localisation of DPY-7, DPY-5 and SQT-1 can be explained by its late expression in relation to the larval moult.

It has been shown herein that COL-19 is late expressed, peaking coincident with the moult in a similar manner as COL-12 (Johnstone and Barry, 1996). By this time, the aforementioned collagens have been expressed and have presumably been incorporated correctly into the cuticle.

Another reason for the mild RNAi phenotype is the fact that, as shown here by expression of COL-19 in a baculovirus system, COL-19 is able to form homotrimers. Assuming this is the only form would mean that the effects of removing the collagen only affect one type of trimer molecule.

#### **3.3.4 The properties of COL-19::GFP that optimise it for being a useful marker**

TP12 was produced as a means of observing the cuticle aberrations in morphological mutants in order to dissect the effect that the lesions have on the cuticle structure. Previous similar studies have been immunological, using antibodies against DPY-7 or Ty-tagged collagens to observe the expression patterns. These were also able to show that the cuticles of different morphological mutants are different (McMahon *et al.*, 2003). However, in addition to the aforementioned facts that TP12 animals require no immunological processing and the fact that live worms can be viewed, this strain has advantages over immunological methods since COL-19::GFP localises to the alae as well as to the seam cell-derived cuticle, thereby allowing alae defects to be observed. As will be discussed in the next chapter, the seam cell-derived cuticle is distinct from the dorsoventral region of the cuticle and therefore, use of TP12 as a means of observing the cuticular pattern in this area, is useful.

The late expression of COL-19 is also a useful property to have as a molecular marker because by the time that COL-19::GFP is expressed, the majority of the collagens will have been incorporated in the cuticle. Thus, COL-19::GFP localisation is a reflection of the cuticle in terms of all of its components. In addition, COL-19::GFP does not get excluded from the cuticle as a consequence of mutations to other mutable collagens. COL-19::GFP has been shown to be present both in the absence of DPY-7 and DPY-5 (this is discussed in chapter 4). This is unlike the obligate partners for these two collagens, as removal of DPY-7 results in the exclusion of the annular furrow-localised collagens from the cuticle. Similarly, removal of DPY-5 results in the absence of DPY-13 in the cuticle (McMahon *et al.*, 2003). COL-19::GFP patterns being observed in these mutants implies that this collagen does not require these collagens for its secretion into the cuticle. It does, however, require them for wild type localisation and this will be examined in the next chapter. It has been demonstrated here, through the use of baculovirus

expression system, that COL-19 forms homotrimers *in vitro*. These results suggest that even if COL-19 can interact with other collagens, in their absence, this collagen might have the ability to homotrimerise, thus possibly becoming incorporated successfully into the cuticle. Moreover, of note, is that the obligate partners, if any, that COL-19 may have, are not mutable to a morphological phenotype, since all such collagens which lead to a mutant phenotype have been tested using COL-19::GFP (described in chapter 4), and all express COL-19::GFP in their cuticle. A reason for the lack of mutable obligate partners is likely due to the ability of COL-19 to form homotrimers. SQT-3 is the only other collagen identified to date, that can homotrimerise.

The reasons described above exemplify how TP12 is an extremely useful tool for the dissection of the cuticle in morphological mutants, which is described next.

### 3.3.5 Investigation into the role of the promoter in collagen processing

The ability of a collagen to be fused to a marker moiety is relatively rare, and this has been attributed to the proteolytic processing that may occur at the C- and N-terminals of most collagens. It was of interest to understand the reasons why COL-19 is apparently not cleaved when fused to a GFP marker. Two properties of COL-19 were suggested as being factors that prevent its proteolytic processing, namely the fact that COL-19 is only expressed during the L4-adult moult (confirmed by SQ RT-PCR assays (figure 3.5)) and the fact that COL-19 has the shortest C-terminal non-Gly-X-Y domain. It was suggested that adult-specific processing may occur or that proteolytic enzymes might not recognise the short C-terminal as a site of cleavage.

The presence of a short propeptide domain being sufficient to prevent proteolytic removal of the marker was ruled out because ROL-6 and COL-132 (c39e9.9) were shown previously to be unable to form functional fusions, despite having relatively few C-terminal residues (Thein *et al.*, 2003). It has also shown, in this study, that SQT-1 cannot form a functional fusion either.

In order to analyse the effect of the temporal expression on collagen processing, a number of constructs were made in which different collagens were put under the influence of a foreign promoter, observing, in the resulting transgenic lines, whether altering the time at which a collagen is expressed changes the extent of C-terminal processing. Figure 3.9 shows the constructs made and whether functional marker fusions were observed as a result.

An exciting finding was that *sqt-1* could not form a functional fusion, when expressed in its native period of expression, but could, when placed under the influence of *col-19*. Because *sqt-1*, under its own promoter, is expressed in the adult stages, it was suggested that the unique processing that was preventing proteolytic cleavage was more specifically associated with a limited time period *within* a moult i.e. at the late stage when *col-19* is expressed. These ideas were further supported by the discovery that COL-7, another collagen that is expressed late within the adult moult, could also form a functional GFP fusion (Gillian McCormack, WCMP, University of Glasgow). The SQT-1::GFP expressed under the influence of the *col-19* promoter appears to localise to the annular furrows and strongly stains the alae. The staining is quite blurred and could indicate that some of the protein is 'leaking out'. This may be a consequence of the lack of obligate interacting partners for the collagen. Where SQT-1 localises, when under its own promoter, has not yet been identified, so it cannot be established whether the SQT-1 fusion molecule is localised in a wild type fashion or whether the late expression has resulted in it being deposited into different cuticle substructures. The observation that SQT-1 cannot homotrimerise in baculovirus systems (A. Page and A. Winter, WCMP, University of Glasgow) is quite interesting because it suggests that this collagen interacts with other collagens within the cuticle and strengthens the case for the requirement for obligate partners for this collagen. It might be useful to investigate whether SQT-1 can interact with other collagens that are expressed late in the moult.

It appears that the sequestration of a collagen to this temporal period is not sufficient to prevent processing because *dpy-5*, *dpy-13*, and *dpy-7* fail to produce functional fusions when fused with the *col-19* promoter. Thus, there appears to be sequence specificity as to which collagens are processed during the late stages of the adult moult, a phenomenon that may indicate the presence of a limited set of proteolytic enzymes, with specificities for collagen substrates. It must also be taken into account, however, that a reason for the absence of a fluorescently-tagged protein in the cuticle of the transgenic lines containing these constructs may be a result of the GFP-tagged collagens being excluded from the cuticle after being recognised by the protein sorting machinery as being aberrant.

In light of the results that suggested that processing of SQT-1 may occur during the early stages of the adult moult, but not later stages, it was of interest to observe whether proteolytic processing of COL-19 could be induced by expressing it earlier, i.e. at the time of *sqt-1* expression. Thus *col-19* was placed under the influence of the *sqt-1* promoter. Fluorescence is

observed in transgenic lines carrying this strain. However, the staining is non-specific and is not contained within the cuticle. There may be a number of reasons for this: there is high-level expression of the collagen fusion, which cannot be accommodated for in the cuticle, thus leaking out (proteolytic processing may have been prevented by the lack of enzymes specific for COL-19); the lack of obligate collagens to which it can bind means that the COL-19 cannot be incorporated into the cuticle (despite the fact that COL-19 has been shown to be to homotrimerise in BV systems); or the GFP marker has been proteolytically removed from the COL-19 and it is the marker alone which is being secreted non-specifically, perhaps because of the high levels of expression, the excess GFP cannot be turned over and thus accumulates. Further analysis is necessary to analyse these data and this is discussed in section 3.4.

### 3.4 Future work

Further analysis will likely concentrate on creating more transgenic strains of the Ty-tagged COL-19 and investigating the promoter swap experiments further.

The construct containing the Ty-tag ligated into the coding sequence of *col-19* was created with the aim of analysing the effect of aberrantly processed COL-19 collagens in the cuticle. Preliminary observations have not been conclusive because antibody staining experiments have been unsuccessful, presumably because of the failure of the transgenic line to carry the construct. It will be vital to create more transgenic lines and observe whether a similar phenomenon occurs. It will also be interesting, even in the obtained line, to carefully observe the alae on the premise that another explanation for the lack of Ty-tagged collagen in the cuticle might be that the (presumably) mis-folded collagen is being excluded from the cuticle. It has been shown here that such removal of COL-19, via RNAi, results in alae defects, and similar observations would be expected if the Ty-tagged collagen was retained in the ER. It has been suggested that COL-19 may be highly cross-linked in a manner similar to other alae-localising collagens. Such aberrantly folded collagens would be especially disruptive to such highly cross-linked structures, so it could be envisaged that a mechanism exists to exclude these collagens.

In the future, it will also be interesting to continue with investigations into the roles of the promoters from the 'promoter swap' experiments. To first confirm the annular localisation of the *col-19:sqt-1* fusion molecule, it will be useful to use DPY-7 antibodies in immunolocalisation assays. As described above, it was interesting to note that SQT-1 could be incorporated into the cuticle because this collagen has been demonstrated to be unable to

homotrimerise in BV systems. Consequently, expression of this collagen in a period when its presumed usual obligate partners are not being expressed might mean that SQT-1 has the ability to interact with collagens that are late expressed, which it does not normally interact with. It may be interesting to co-express *sqt-1* and late expressing collagens such as *col-19* or *col-7* in baculovirus systems to observe whether trimers can be formed

The *sqt-1* promoter was fused to the *col-19* coding sequence. The fluorescent marker did not become incorporated into the cuticle and instead was localised non-specifically. Since one of the possible reasons for the non-specific staining is that the expression of the fusion construct is too high, a new construct could be produced containing a shorter *sqt-1* promoter sequence. Gilleard *et al.* (1997) demonstrated that, for collagen expression, short promoter sequences are sufficient, with more 5' sequences only enhancing the levels of expression. Thus, by reducing the length of the promoter sequence in the constructs, this enhanced expression could be minimised and could increase the chances of the fluorescent marker to be incorporated into the cuticle. Such a marker would be extremely useful as a means of observing cuticle aberrations in larvae.

It may also be interesting to create a greater array of fusion constructs, using, for example, the promoters of *dpy-7* and *dpy-5* with the coding sequences of other collagens. Two constructs have already been prepared containing *sqt-3* and *col-19* under the influence of each other's promoters. Results from these will be interesting because SQT-3 is the only collagen that has been proven to be proteolytically cleaved at its C-terminus, by DPY-31. It will be interesting to observe whether altering its temporal expression, such processing could be prevented. SQT-3 has also been shown to homotrimerise in BV systems and consequently, it may be possible for it to be incorporated into a cuticle without the requirement for obligate partners.

## **Chapter 4**

***Analysis of the cuticles of morphological mutants and an RNAi screen to find genes involved in cuticle synthesis.***

## 4.1 Introduction

In the previous chapter, the hypodermally expressed GFP-tagged COL-19, was shown to form a functional fusion that localises to the cortex of the circumferential annulae. It was designed as a tool for visualising the cuticle without the need for the complicated experimental procedures required by antibody staining, the traditional means of observing the cuticle structure. This chapter describes how the COL-19::GFP marker strain, TP12, has been exploited through two separate approaches, namely, a series of crosses with morphological mutants, and the RNAi-targeting of a variety of cross-linking enzymes, proteases, collagens and other gene products. A wide range of definable cuticular disruptions have been made evident as a result of both the crosses and RNAi experiments. Consequently, TP12 has been successfully shown to provide a means of efficiently describing cuticle aberrations in terms of COL-19::GFP disruption. This has facilitated firstly, the characterisation of morphological mutant cuticles and secondly, the identification of genes whose products are involved in dictating the body shape morphology of *C. elegans*.

Morphological mutants exhibiting Dpy, Sqt, Bli, Rol, Lon and Sma phenotypes are vital in the study of the *C. elegans* cuticle as these mutants result from the disruption of cuticle collagens, the malfunction of the cuticle processing machinery, or from alterations of components controlling body size. The cuticle is a complex structure with multiple layers (Cox *et al.*, 1981a). Evidence from ultra-structural analysis has shown that Rol and Bli phenotypes are a result of aberrations to or loss of fibrous/intermediate layers and struts respectively (Cox *et al.*, 1981a). On the other hand, Dpy mutants do not exhibit such obvious gross loss or alterations to cuticle layers but do display an overall increase in body width. Through SEM analysis and antibody staining against the DPY-7 cuticle collagen, at least two different classes of Dpy mutants have been identified, as discussed in section 1.10.1 (McMahon *et al.*, 2003). Dpy mutants, whose affected loci are expressed four hours prior to the moult (*dpy-7*, *dpy-8*, *dpy-2* and *dpy-10*), exhibit smooth surfaces that are devoid of annulae, while those expressed two hours before the moult (*dpy-5* and *dpy-13*) exhibit annulae with mild branching (McMahon *et al.*, 2003). This demonstrates the variation in cuticle structure exhibited by animals sharing gross morphological phenotypes.

Due to the structural complexity of the cuticles of morphological mutants, the effectiveness of TP12 to define cuticle disruption has aided the elucidation of how the lesions affect the overall

cuticle structure. After crossing the TP12 strain with morphological mutants, the resulting pattern of COL-19::GFP was compared to that of the pattern of regular circumferential bands corresponding to the annular cortex that is displayed by wild type (TP12) animals. The range of COL-19::GFP patterns observed as a result of these crosses is described in detail in this chapter. The data have enabled the nature of the cuticle aberrations to be observed and has provided a structural basis for the morphological defects. It has furthered our knowledge of the overall composition of the cuticle, the structural and enzymatic role of each component within the system, as well as the complex collagen interactions. Mutations to different components of the system controlling body shape have also enabled the characterisation of distinct regions of the cuticle, namely that overlying the seam cell hypodermis and that overlying the dorsoventral hypodermis. Overall, these crosses have helped to investigate the relationship between the genes and their effect on body morphology.

Certain molecular lesions affect distinct genetic loci differently and this is particularly evident in cuticle collagen mutants. For example, the Dpy phenotype of the cuticle collagen locus, *dpy-7*, is a result of the collagen's exclusion from the cuticle, either through RNAi/null allele (e.g. *dpy-7(qm63)*) or by the removal of the aberrantly folded glycine substitutions by the protein sorting machinery (e.g. *dpy-7(e88)*). In *sqt-1* mutants, on the other hand, such removal of the collagen does not affect the gross morphology of the worm since null alleles (e.g. *sqt-1(sc103)*) and RNAi-treated worms appear wild type. The Sqt phenotypes of this genetic locus are exhibited by alleles whose lesions prevent the proper N-terminal processing (*sqt-1(e1350)*) or C-terminal cross-linking (*sqt-1(sc13)*). The resulting aberrantly folded collagens are not excluded from the cuticle by the processing machinery, and the Sqt phenotype results as a consequence of the incorporation of such misfolded collagens into the cuticle. A mild phenotype is exhibited by *sqt-1* glycine substitution alleles (e.g. *sqt-1(sc107)*). It must be noted that despite the wild type appearance of *sqt-1(sc103)* null worms, ultra structural analysis has revealed that these worms do indeed lack one of the fibrous layers, in a similar manner to their Rol phenotype counterparts. The experiments described in this chapter have shown that COL-19::GFP of TP12-crossed animals has the ability to reveal cuticle defects that are not evident from gross morphological phenotype and consequently facilitate a more detailed analysis of the cuticles of different alleles.

Mutations of some loci that encode for cuticle processing enzymes can also produce allele-specific gross morphologies in a similar manner to cuticle collagen mutants. One such example of such an enzyme-encoding locus is that of the subtilisin-like serine endoprotease locus, *bli-4*. All of the known *bli-4* alleles identified to date are embryonic or larval lethal, apart from *bli-4(e937)* which is viable and displays an adult-specific Bli phenotype. The allele-specific viability of *bli-4(e937)* has been explained by the fact that it is the only lesion that does not result in the removal of all 12 exons of the gene product (Thacker *et al.*, 1995).

*dpy-18/phy-1* (a prolyl 4-hydroxylase), *dpy-11* (a thioredoxin), *bli-3* (a dual oxidase) and *bli-5* (a serine protease inhibitor) are also mutants whose affected loci encode for enzymes or regulators associated with collagen processing. These mutants have been crossed with TP12 in order to visualise their cuticle disruption.

In addition to these loci, RNAi-targeting of additional enzyme loci with potential roles in cuticle synthesis has been shown to lead to characteristic Bli, Dpy and Rol phenotypes, as demonstrated by the genome-wide RNAi screen by Kamath *et al.* (2003). In the second part of this chapter, the use of TP12 in a similar, but more selective, RNAi screen to find genes with potential roles in cuticle synthesis, will be described. The process of cuticle synthesis, as described in section 1.6.1, involves multiple steps that include hydroxylation of proline and lysine residues, *cis-trans* isomerisation, proteolytic processing of the pro-collagen N- and C- termini, and numerous cross-linking steps. Many of the enzymes with potential roles in such steps were the targets of this RNAi screen and as a result, a number of them were identified as having important roles.

## 4.2 Results 1: Crosses between TP12 and *C. elegans* morphological mutants

Crosses were set up in which six TP12 males were plated with a single hermaphrodite of the morphological mutant in question. F1 homozygous hermaphrodites were chosen on the basis of their wild type phenotype and F2 worms chosen on the basis of GFP and morphology as assessed via microscopy.

### 4.2.1 Five classes of COL-19::GFP disruption are observed

The various TP12-carrying morphological mutants exhibit COL-19::GFP patterns that range from the wild type pattern of regular circumferential stripes, to severe global disruption in the

annulae and alae. Each TP12-carrying strain was scored with a value of between 0 to 3 on the relative disruption of COL-19::GFP pattern in the annulae and in the alae. In terms of annulae, the designations were as follows: wild type (narrow or wide) (0); abnormal branched (lateral hypodermis) (1); missing or amorphous (2); and severe disruption (3). Scoring of the alae included the categories of: wild type (0), multiple or discontinuous (1), abnormal and branched (2) and severe disruption with branching (3).

As a consequence, five distinct classes of cuticle disruption were defined and are described in table 4.1 and illustrated in figure 4.1. Class I mutants exhibit a COL-19::GFP pattern indicating a relatively wild type annular banding pattern and regularly spaced trilaminar alae (annulae and alae scores between 0 and 1). Class II mutants display an amorphous pattern of COL-19::GFP in all regions of the cuticle with no apparent annular banding patterns (annulae score of 2). Alae are relatively wild type albeit occasionally multiple (alae score of between 0 and 1). In class III mutants, worms exhibit the regular wild type pattern of circumferential annulae in the cuticle overlying the dorsoventral hypodermis but adopt a branched and thread-like COL-19::GFP pattern in the cuticle overlying the seam cells (annular score of 1). The alae, around which the branched pattern extends, are wild type in nature (alae score of 0). Class IV is characterised by severe global disruption and exhibits both dramatic branching of the annulae on all regions of the cuticle and large loops in the alae (annular score of 3, alae score of between 2 and 3). Finally, Class V mutants are a less severe form of Class IV, and are described as having branched or amorphous annulae and branching alae (annular score of 2, alae score of between 1 and 2).

The identity of the loci that fall into each class of COL-19::GFP disruption is shown in table 4.1. It is also shown that the severity of the gross morphological phenotype does not necessarily correspond to the extent of COL-19::GFP disruption, nor, in collagen mutants, is there any correlation between the class of disruption and the predicted *C. elegans* pro-collagen structural group. However, a general pattern appears, in which: *Rol*, *Sma* and *Lon* mutants fall into class I; *Dpy* mutants fall into class II; *Dpy* and *Bli* mutants fall into class III; *Sqt* mutants and *dpy-17*, which is most similar to *sqt-1*, comprise class IV while; class V consists of enzyme mutants (refer to figure 4.1). This illustrates, for example, that the *Sqt* collagen mutants exhibit a distinct cuticle structure from that of *Dpy* and *Bli* mutants.

Another observation is that *Dpy* mutants, even though exhibiting similar gross morphological phenotypes (i.e. *Dpy*), in fact exhibit one of two extremely distinct cuticle structures; exhibiting

**Table 4.1: Definitions of Classes of COL-19::GFP Adult Cuticle Collagen Disruption.**

Genotype	Identity (Cosmid)	Lesion	Phenotype(adult)	Alac <sup>1</sup>	Annulae <sup>2</sup>
<b>I Mild Disruption or Wild Type</b>					
TP12 (N2;kals12[col-19::gfp])	NA		Wild type	0	0
<i>bli-4(e937)</i> (TP29) II	Subtilisin protease (k04f10.4)	[Exon deletion]	Bli	0	0
<i>rol-6(su1006)</i> (TP11) II	Cuticle collagen (t01b7.7)	[N-protease R71C]	Rol	0	0
<i>sma-2(e502)</i> (TP27) III	Dwarf1 (zk370.2)	?	Small (mta)	0	0 <sup>N</sup>
<i>lon-3(e2175)</i> (TP31) V	Cuticle collagen (zk836.1)	?	long	1	1 <sup>W</sup>
<i>lon-1(e185)</i> (TP42) III	PR protein (f48c8.1)	?	long	1	1 <sup>W</sup>
<i>lon-2(e678)</i> (TP44) X	Not yet known (c39e6.1)	?	long	1	1 <sup>W</sup>
<i>lon-2(n1630)</i> (TP52) X	Not yet known (c39e6.1)	?	long	1	1 <sup>W</sup>
<b>II Amorphous Annulae</b>					
<i>dpy-7(e88)</i> (TP17) X	Cuticle collagen (f46c8.6)	[Gly156Arg]	medium Dpy	0	2
<i>dpy-7(qm63)</i> (TP36) X	Cuticle collagen (f46c8.6)	[null]	medium Dpy	0	2
<i>dpy-3(e27)</i> (TP22) X	Cuticle collagen (egap7.1)	?	medium Dpy	1	2
<i>dpy-10(e128)</i> (TP19) II	Cuticle collagen (t14b4.7)	[g766a splice]	small Dpy	1	2
<i>dpy-10(cn64)</i> (TP43) II	Cuticle collagen (t14b4.7)	[Arg92Cys]	medium Dpy Rol	1	2
<i>dpy-8(e130)</i> (TP24) X	Cuticle collagen (c31h2.2)	?	medium Dpy	1	2
<i>dpy-8(sc44)</i> X	Cuticle collagen (c31h2.2)	?	medium Dpy	1	2
<i>dpy-9(e12)</i> (TP35)IV	Cuticle collagen (t21d12.2)	[Gly149Glu]	large Dpy	1	2
<i>dpy-2(e489)</i> (TP37) II	Cuticle collagen (t14b4.6)	[Gly247Arg]	Dpy Rol	1	2
<i>dpy-2(sc38s)</i> (TP49)II	Cuticle collagen (t14b4.6)	[Gly129Glu]	Dpy Rol	1	2
<i>dpy-2(e1359)</i> (TP47)II	Cuticle collagen (t14b4.6)	?	Dpy Rol	1	2
<b>III Branched Seam Cell Annulae</b>					
<i>dpy-13(e458)</i> (TP21) IV	Cuticle collagen (f30b5.1)	[deletion of Gly-X-Y]	medium Dpy	0	1
<i>dpy-4(e1166)</i> (TP25) IV	Cuticle collagen (y41e3.2)	[frameshift@392]	large Dpy	0	1
<i>bli-1(e769)</i> (TP28)II	Cuticle collagen(c09g5.6)	?	Bli (ts) (mta)	0	1
<i>bli-5(e518)</i> (TP18) III	Serine protease inhibitor	[splice donor site]	Dpy/Bli (mta)	0	1
<i>dpy-5(e61)</i> (TP14) I	Cuticle collagen (t27c1.8)	[Gly202opal]	strong Dpy	1	1
<b>IV Severe Global Disruption</b>					
<i>sqt-3(sc63)</i> (TP23) V	Cuticle collagen (f23h12.4)	[Gly227Glu]	medium Dpy (ts)	2	3
<i>sqt-3(e2117)</i> (TP20)V	Cuticle collagen(f23h12.4)	[Gly114Arg]	extreme Dpy (ts)	2	3
<i>sqt-1(e1350)</i> (TP26) II	Cuticle collagen (b0491.2)	[N-protease R69C]	medium Dpy (mta)	3	3
<i>sqt-1(sc1)</i> (TP40) II	Cuticle collagen (b0491.2)	[N-protease R66C]	medium Dpy (mta)	3	3
<i>sqt-1(sc13)</i> (TP33) II	Cuticle collagen (b0491.2)	[C-domain C303Y]	medium Dpy Rol	2	3
<i>sqt-2(sc3)</i> (TP45) II	Cuticle collagen (c01b12.1)	[N-protease R76C]	medium Dpy	3	3
<i>sqt-2(sc108)</i> (TP48) II	Cuticle collagen (c01b12.1)	[N-protease R79C]	medium Dpy	3	3
<i>dpy-17(e164)</i> (TP32) III	Cuticle collagen (f54d8.1)	[Gln188Amber]	medium Dpy	3	3
<i>dpy-6(e14)</i> (TP46) X	Actin interacting? (f16f9.2)	[Thr5number]	Dpy	2	3
<b>V Branching Alac and Annulae</b>					
<i>dpy-18(e364)</i> (TP13) III	Prolyl 4-hydroxylase (t28d6.1)	[Trp76amber]	medium Dpy (mta)	2	1
<i>dpy-18(ok162)</i> (TP41) III	Prolyl 4-hydroxylase (t28d6.1)	[multiple lesions]	medium Dpy	2	1
<i>dpy-11(e224)</i> (TP16) V	Thioredoxin (f46e10.9)	[G76E partial]	medium Dpy (mta)	2	1
<i>dpy-11(e1180)</i> (TP30) V	Thioredoxin (f46e10.9)	[null]	medium Dpy (mta)	2	2

Dpy = dumpy. Bli = blister. Mta = male tail abnormal. Rol = roller. Ts = temperature sensitive.

**COL-19::GFP expression**

<sup>1</sup> Alac

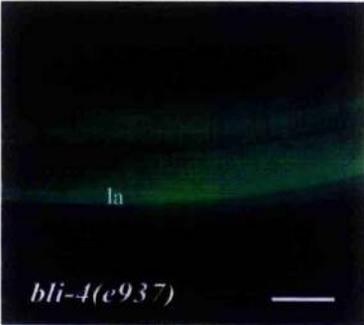
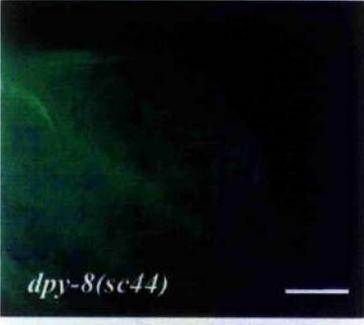
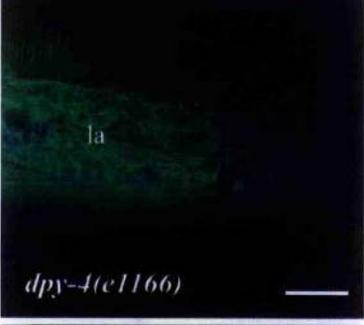
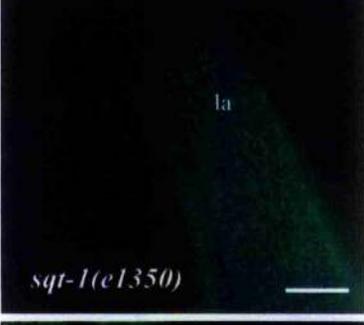
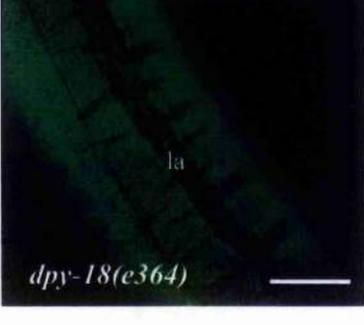
- 0. Wild type.
- 1. Multiple or discontinuous alac.
- 2. Abnormal branched alac.
- 3. Severe disruption with branching.

<sup>2</sup> Annulae

- 0. Wild type, (<sup>N</sup> narrow annulae, <sup>W</sup> wide annulae).
- 1. Abnormal branched annulae (lateral hypodermis)
- 2. Missing or amorphous annulae.
- 3. Severe disruption.

**Figure 4.1: The 5 classes of COL-19::GFP disruption in morphological mutants.** The COL-19::GFP pattern of each morphological mutant was scored on the basis of disruption to the annulae and the alae (see key). The figures show an example of the disruption observed and the morphological mutant from which it was produced is denoted. Lateral alae are denoted with "la" (but in class II are not in plane of focus). The gross morphological phenotype associated with each class is also shown. In addition, the DPY-7 immunolocalisation pattern and the associated loci are described. Scale bars = 10µm.

Facing page 155

Class	Example of COL-19::GFP pattern	Appearance of DPY-7 in cuticle	Associated Loci
<p><b>I</b>  <b>Mild disruption or wild type</b>  annulae: 0-1  alae: 0-1</p>	 <p><i>bli-4(e937)</i></p>	wild type	Lon Sma Rol
<p><b>II</b>  <b>Amorphous annulae</b>  annulae: 2  alae: 0-1</p>	 <p><i>dpy-8(sc44)</i></p>	none	Dpy
<p><b>III</b>  <b>Branched seam cell annulae</b>  annulae: 1  alae: 0</p>	 <p><i>dpy-4(e1166)</i></p>	prematurely terminated	Dpy Bli
<p><b>IV</b>  <b>Severe global disruption</b>  annulae: 3  alae: 2-3</p>	 <p><i>sqt-1(e1350)</i></p>	branched	Sqt
<p><b>V</b>  <b>Branching alae and annulae</b>  annulae: 1-2  alae: 2</p>	 <p><i>dpy-18(e364)</i></p>	variable	Dpy (enzyme)

Scores: Alae: 0) Wild type; 1) multiple or discontinuous alae; 3) severe disruption with branching. Annulae: 0) wild type; 1) abnormal branched annulae (lateral hypodermis); 2) missing or amorphous annulae; 3) severe disruption

class II or class III disruption. This is interesting and reiterates the value of the TP12 strain as a tool for looking at cuticle structure. For the morphological mutant strains tested, all of the alleles of a single locus exhibit the same category of disruption e.g. all *dpy-2* alleles exhibit type II disruption. This is also true for the RNAi-phenotypes of *dpy-5*, *dpy-7*, *dpy-13* and *dpy-10* (*dpy-10* and *dpy-13* RNAi not shown). However, this list is not exhaustive and importantly does not include, for example, the null allele *sqt-1* which is wild type in nature and which consequently would most likely not exhibit the same type IV disruption as the alleles exhibiting the Sqt/Rol phenotype. The allele specificity of *sqt-1* and *sqt-3* will be discussed in a later section.

Similar observations of distinct cuticle structures have been described by McMahon *et al.* (2003), who used SEM and immunolocalisation techniques (DPY-7 and anti-Ty antibodies) to observe the cuticle structure of cuticle collagen mutants. *dpy-7*, *dpy-2*, *dpy-3*, *dpy-8*, *dpy-10*, mutants, whose encoded collagens are expressed simultaneously approximately 4 hours prior to the moults, exhibit smooth cuticle surfaces, devoid of annulae as well as a reduced amount, (or total loss of, of DPY-7 antibody staining. Where staining is visible, the pattern deviates from the regular annular pattern and is instead fragmented. Conversely, the mutants in the later expressing collagen genes, *dpy-13* and *dpy-5*, have cuticles which are patterned by the annular furrows. The conclusions were that at least two distinct sets of collagens exist, each with a characteristic temporal expression, localisation in cuticle substructures, and mutation-associated cuticle structure (McMahon *et al.*, 2003). Figure 4.1 lists the DPY-7 staining pattern resulting from immunolocalisation assays in the present study and represents the DPY-7 localisation in each of our defined classes. To date, the mutants falling into the groups defined by McMahon *et al.* (2003) according to their temporal expression, coincide exactly with the members of our groups, which we have defined according to the COL-19::GFP pattern. This lends further support that substructure localisation as well as temporal expression dictate the structure of the corresponding mutant cuticle.

It is important to note that varying intensities of GFP fluorescence are observed in different mutants. However, it is difficult to quantify the staining levels due to practical variables that include the age of the worm (discussed in section 3.2.1), the mounting techniques and the general microscope set up. Thus, the intensity of the fluorescence has not been taken into account when analysing the data even though the fact that a reduction in fluorescence may be a good indication that COL-19::GFP is being excluded from the cuticle.

#### 4.2.2 The COL-19::GFP expression pattern in Long and Sma mutants

Lon and Sma phenotypes result from mutations in the components of the TGF- $\beta$ -like pathway that dictate the body length of the worm. This regulation comes through a number of means, including control of hypodermal cell ploidy and cuticle composition.

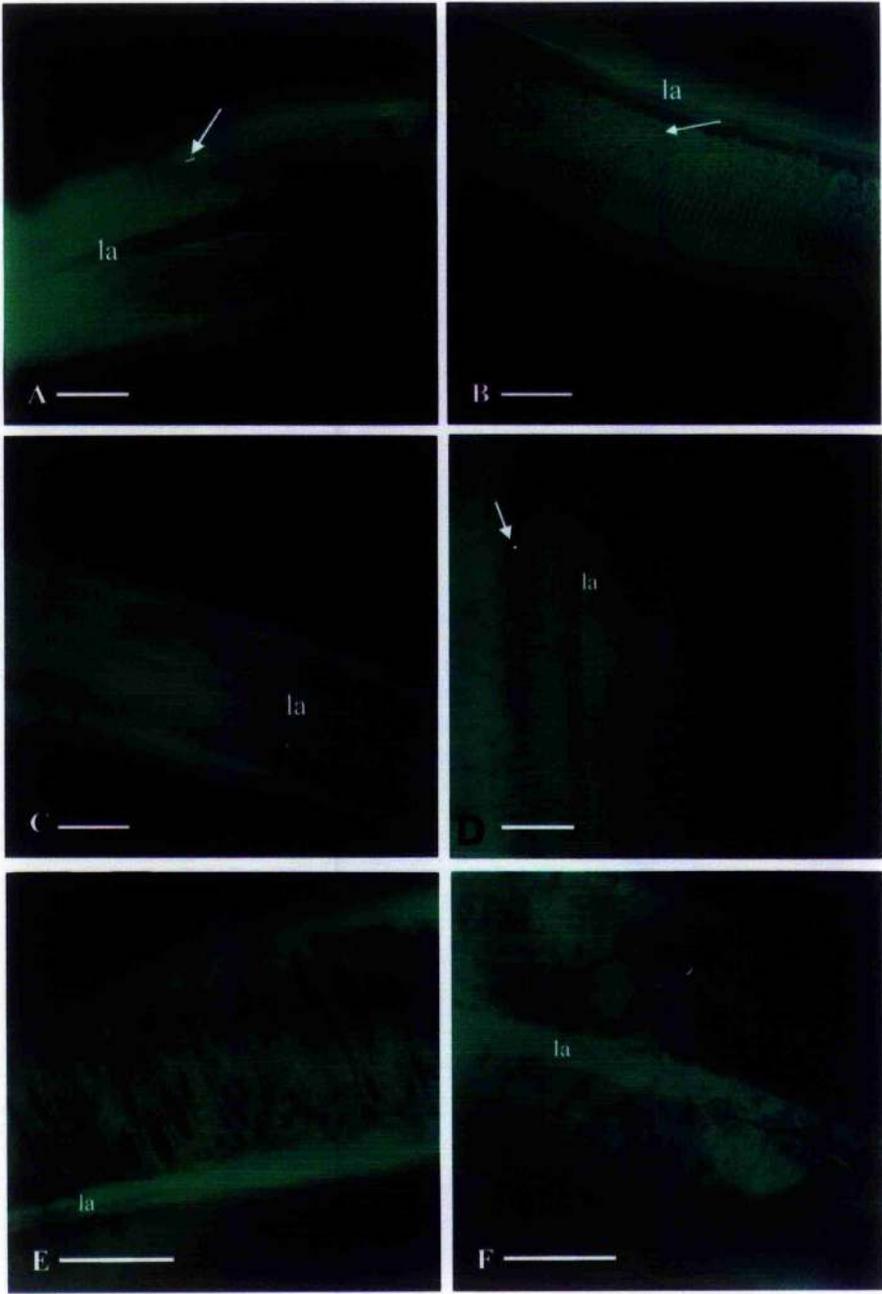
The Lon mutants *lon-1(e185)*, *lon-2(e678)*, *lon-2(n1630)* and *lon-3(e2175)* and the Sma mutant *sma-2(e502)*, are genes involved in size determination, and in this study, were crossed with TP12. The collagen encoding gene, in *lon-3(e2175)* mutants, lacks 2.4 kb of its sequence and has an additional Tc5 transposon insertion (Nystrom *et al.*, 2002; Suzuki *et al.*, 2002). Despite the fact that the molecular lesions in the other genes are not known, the morphological mutant alleles of these loci phenocopy the RNAi effect at the level of gross morphology (Suzuki *et al.*, 2002). Whatever the lesions might be, they are sufficient to cause an alteration in body size presumably by modifying the downstream components of the TGF- $\beta$ -like pathway.

The resulting dominant COL-19::GFP pattern in these TP12-crossed Lon mutants appears as regular circumferential bands along the length of the worm, reminiscent of the wild type pattern (figure 4.2). All the Lon mutants exhibit mild annular disruption and some examples of multiple alae are apparent. *lon-3(e2175)* in particular shows worms with quite distinct COL-19::GFP fragmentation around the alae (figure 4.2B). However, these mutants have been classed as type I disruption according to the dominant and most commonly observed COL-19::GFP pattern that is wild type. A defining feature of these mutants is the increased or decreased width of the annular bands in the Lon and Sma mutants respectively. This is exemplified in figures 4.2 A and D in *lon-2(n1630)* and *sma-2(e502)* mutants respectively, where the width of the annulae can be compared. Attempts at counting the number of annulae were unsuccessful, but the altered width of annulae suggests a similar number of annulae in these mutants and in wild type animals. It has previously been noted that, instead of increasing the number of annulae in order to facilitate a longer body length, the annuli stretch in response to growth during adult life (Costa *et al.*, 1997).

Referring back to the branched pathway presented by Suzuki *et al.* (2002) (figure 1.5), body length determination is controlled by the TGF- $\beta$ -like pathway via regulation of both hypodermal ploidy and control of cuticle components. Hypodermal ploidy is mediated via LON-1 while the collagen, LON-3, plays a role in regulating the cuticle composition. These are separate and

**Figure 4.2: COL-19::GFP expression of Lon, Sma and Rol mutants.**

A) *lon-2(n1630)* mutants exemplify the increased width of annulae in Lon animals (white arrow) compared to wild type (refer to figure 3.3). The line in figure A denotes the width of one annulus. B) In addition to showing a similar increased annular width, *lon-3(e2175)* mutants show differing levels of annular disruption around the alae: the figure exemplifies the most severe form observed (which is denoted by a white arrow). The latter observation indicates that there is active alteration of the cuticle in these mutants. C) In *lon-1(e185)* mutants however, there is less disruption to the cuticle, indicative that the cuticle is most likely exhibiting mechanical stretching in response to the increased body length from increased ploidy of these mutants. D) *sma-2(e502)* mutants show a predominantly wild type pattern of annulae. Consistent with the Sma phenotype, the annulae are reduced in width (arrow) compared to wild type (the white dot depicts the width of a single annulus in the mutant). E) COL-19::GFP of a *rol-6* N-protease cleavage site mutant allele *rol-6(su1006)*. Helical twisting of the alae (la) can be seen. Annulae appear slightly amorphous but are predominantly wild type. F) Adult F1 progeny from targeted RNAi interference of *rol-6* in TP12 animals (feeding at 20°C). An amorphous COL-19::GFP pattern in both the annulae and the alae is observed. In all figures, lateral alae, are labelled (la) (in A and C, alae are not in focus). Scale bars = 10µm



independent control processes, and the LON-3 and LON-1 pathways exhibit little/no cross talk. The results presented here confirm this hypothesis; while *lon-3* mutants exhibit cuticular disruption around the alae (figure 4.2 B), *lon-1* mutants do not (figure 4.2 C). The wide, yet wild type, annulae of *lon-1* mutants are thus most likely to be the consequence of the cuticle passively stretching to accommodate for the increased ploidy. The slight branching pattern of the annulae observed in these *lon-1* mutants could be a subtle effect of disruption due to mechanical stretching.

The dwarfin/Sma mutant, *sma-2(e502)* exhibits regular bands corresponding to annulae but these are thinner than in wild type animals. Only very subtle branching annulae are observed (refer to figure 4.2 D). Due to the upstream position of SMA-2 in the TGF- $\beta$ -like pathway, it was expected that *sma-2* mutants affect both the LON-3 and the LON-1 pathways. A similar situation is likely for *lon-2* mutants since the alleles tested result in relatively wide but wild type annulae.

#### 4.2.3 The COL-19::GFP disruption in *rol-6* mutants

The cuticle collagen mutant, *rol-6(su1006)*, exhibits a lesion in its N-protease subtilisin-like cleavage site and was crossed with TP12. The COL-19::GFP pattern in crossed worms (figure 4.2 E) is relatively wild type since individual annulae can be identified. Alae also appear to be wild type apart from being helically twisted along the body length, as would be expected for the Rol phenotype (refer to figure 4.2 E). As *rol-6(su1006)* is not a null allele, RNAi of *rol-6* was performed on TP12 worms. RNAi-treated animals exhibited no obvious gross morphological defects but had an amorphous COL-19::GFP pattern (figure 4.2 F). The fact that the COL-19::GFP disruption in the null phenotype (RNAi) is more severe than that in the N-terminal protease site mutant, indicates that the collagen is essential and that the lesion in the collagen encoded by the *rol-6(su1006)* mutant is not sufficient to (completely) prevent its incorporation into the cuticle or to cause aberrations in the COL-19::GFP expression pattern.

It must be noted that *rol-6(su1006)* mutants, despite having a wild type COL-19::GFP pattern, must clearly have some form of aberration due to the exhibited helical twist. This observation exemplifies a precaution that has to be taken when analysing COL-19::GFP disruption, which is that aberrations in the underlying layers are not always visible at the level of the GFP marker.

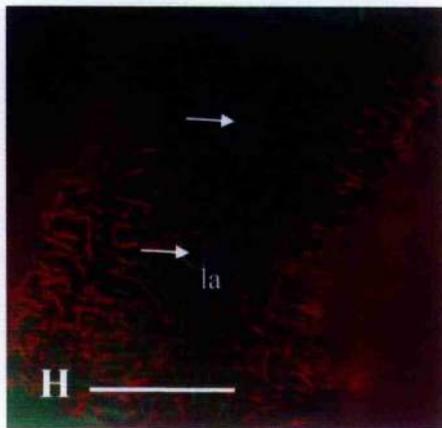
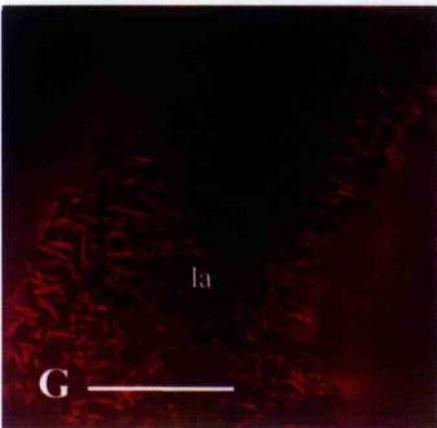
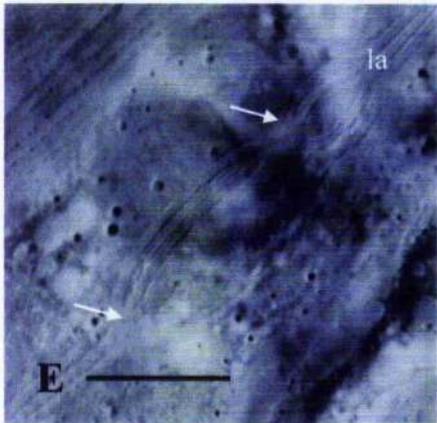
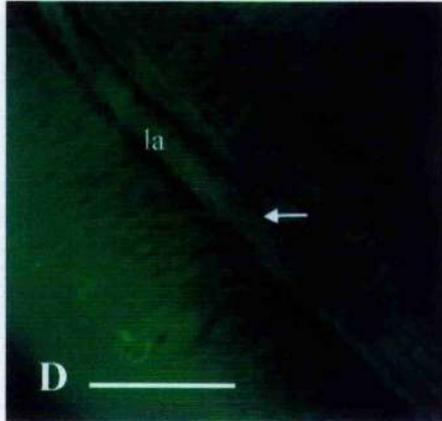
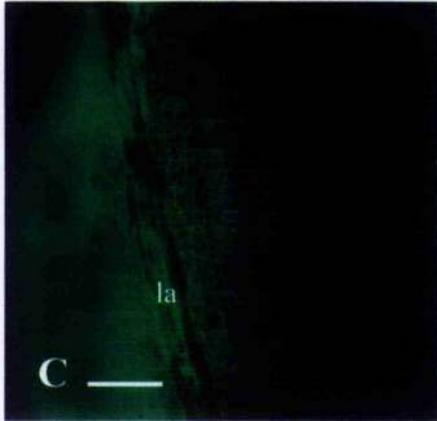
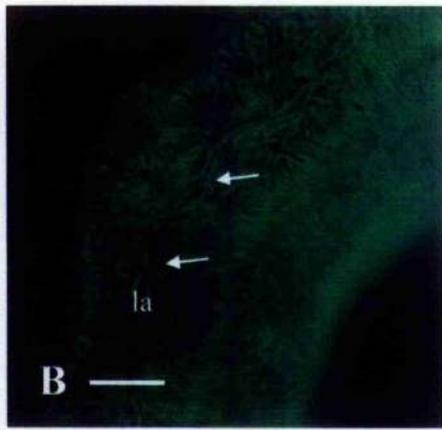
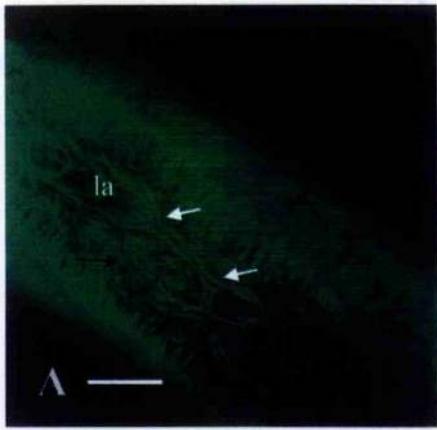
#### 4.2.4 The COL-19::GFP disruption in Sqt mutants

The Sqt phenotype results from mutations in a number of collagen loci and in this study the Sqt gross morphological phenotype has been associated with a class of COL-19::GFP disruption, termed severe global disruption. Extensive branching of the annulae in the ventrodorsal-derived and seam cell-derived cuticles as well as severely branched and multiple alae, are characteristic features of type IV disruption. Some, but not all, alleles of the mutant collagen loci *sqt-1*, *sqt-2*, *sqt-3* and *dpy-17* display such a COL-19::GFP phenotype. Some alleles of these loci, however, are wild type at the COL-19::GFP and gross morphological level.

The *sqt-1* mutants *sqt-1(sc1)*, *sqt-1(e1350)*, and *sqt-1(sc13)* were crossed with TP12 (figures 4.3 A-C). The former two of these gene lesions are substitutions in the N-terminal protease domain while the latter is a cysteine substitution in the C-terminus.

The N-terminal processing mutants, *sqt-1(e1350)* and *sqt-1(sc1)* (figures 4.3 A and B), have COL-19::GFP patterns with distinctive branching of the alae and severely fragmented annulae. These phenotypes are more severe than mutants (e.g. *sqt-1(sc13)*) whose encoded SQT-1 cannot correctly form tyrosine cross-links. *sqt-1(sc13)* animals show only limited amounts of branching annulae and alae (figure 4.3C) and also have multiple alae (which are not shown in the figure). The Rol phenotype exhibited by these worms is also apparent with directional twist of the alae. The differences in severity between the N-terminal processing mutants and the cross-linking mutant might be explained by the fact that even though *sqt-1(sc13)* collagens have lost their ability to form cross-links, their actual monomeric collagen molecules are wild type. These collagens would therefore probably be less sterically disruptive than the N-terminal processing mutant collagen monomers of *sqt-1(e1350)* and *sqt-1(sc1)*, because they differ from the wild type monomer in having a relatively large 126 residue pro-collagen domain, which turns the monomer length from its normal ~475 residues to ~601 residues. Nomarski imaging shows that a branching pattern of the alae (figure 4.3 E) which corroborates the COL-19::GFP data and indicates that the observed COL-19::GFP branching in these mutants is a true representation of the alae. The Nomarski image is not able to show the pattern of the annulae and thus TP12-crossed *sqt-1(e1350)* worms were probed using an antibody against the cuticle collagen DPY-7 that localises to the annular furrows. Figure 4.3 G shows the fragmented pattern of annulae as visualised by immunolocalisation with this antibody. The merged COL-19::GFP (figure 4.3 F) and DPY-7 image shows the patterns coincide (figure 4.3 H). This illustrates that the COL-19::GFP pattern reflects the actual aberrant (fragmented) annular pattern.

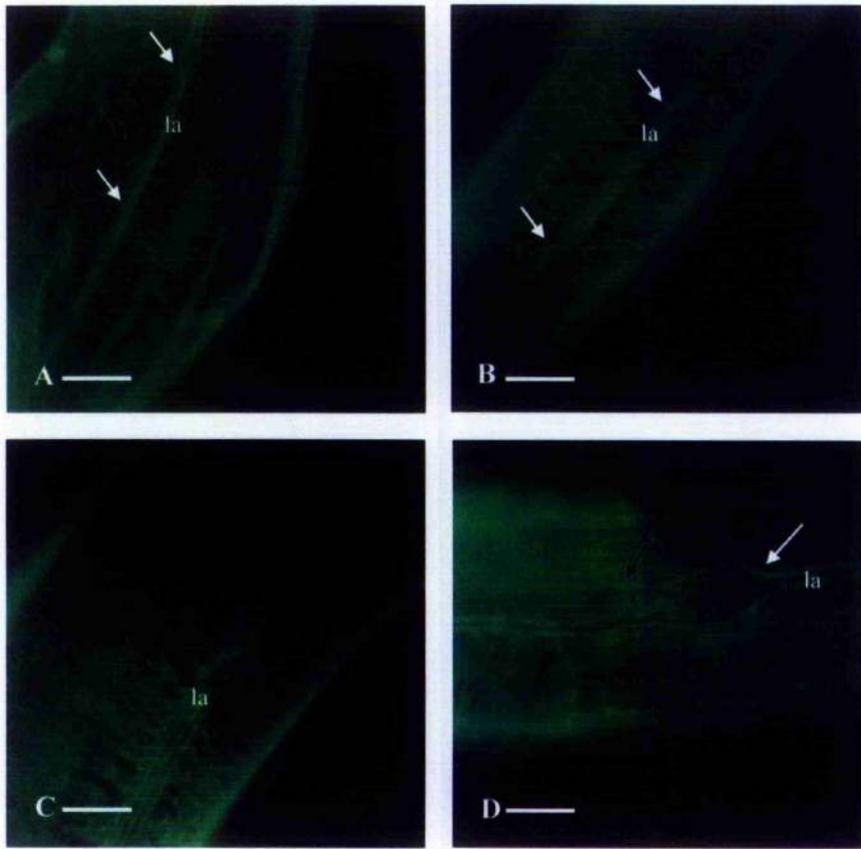
**Figure 4.3: *sqt-1* collagen mutants exhibit type IV COL-19::GFP disruption.** A, B and C) The COL-19::GFP disruption pattern of three alleles of *sqt-1*: (*sqt-1(sc1)*, *sqt-1(e1350)* and *sqt-1(sc13)*) respectively when crossed with TP12. Alae are indicated (la). *sqt-1(sc1)* (A) and *sqt-1(e1350)* (B) are N-terminal processing mutants and exhibit more severe branching in the annulae and alae than the *sqt-1(sc13)* mutant (a C-terminus cross-linking residue mutation) (C). *sqt-1(sc13)* mutants have a Rol phenotype and this is reflected in the altered angle in which the alae run along the body. D) The COL-19::GFP pattern of F1 progeny of *sqt-1* RNAi treated L4 TP12 animals (by feeding). RNAi-treated animals have a wild type gross morphological phenotype (not shown) and this is reflected in the milder form of COL-19::GFP disruption in these animals: only slight branching of the annulae and alae is observed. E) The branching COL-19::GFP pattern observed does represent branching alae, as observed by Nomarski imaging. F-H) The COL-19::GFP pattern (F) is similar to that observed upon immunolocalisation using antibodies against DPY-7 (G) (see H for merged image) and thus confirms the fragmented pattern of the annulae. In all figures, where present, branching alae are indicated with white arrows and branching annulae are depicted with black arrows. Lateral alae are marked 'la'. Scale bars = 10µm.



RNAi of *sqt-1* was also performed on TP12 worms and, as expected, a wild type gross morphological phenotype was produced. However, at the COL-19::GFP level, slight gross morphological abnormalities (figure 4.3D) were observed, with RNAi-treated animals exhibiting slightly branched annulae and alae. These findings are consistent with the fact that the null mutant, *sqt-1(sc103)*, when observed at the ultra-structural level, does in fact exhibit an aberrant fibrous layer. Thus, while these fibrous layer aberrations are not sufficiently disruptive to result in a Rol phenotype, they do cause slight alterations in overlying layers which are reflected in the slight COL-19::GFP disruption (figure 4.3 D). The less severe cuticle disruption in these RNAi-treated worms strengthens the argument that it is the incorporation of cuticle 'poisons' into the cuticle that are the factors that cause the abnormal body morphology in N- and C-terminal processing and cross-linking mutants.

A Gly substitution mutant, *sqt-1(sc99)*, has also been analysed for its COL-19::GFP cuticle pattern; these animals display wild type phenotypes at both the COL-19::GFP and gross morphological levels (data not shown). The weak phenotype may arise because glycine substitutions in this collagen are not disruptive to the cuticle structure, or because aberrant collagens are excluded from the cuticle, making this strain an effective null.

SQT-3 is a unique collagen with an essential function in embryogenesis. It is a complex locus with an array of mutant phenotypes that include larval lethal, extreme Dpy and wild type, and there are, additionally, a number of temperature- and cold- sensitive alleles. Because of the complexity of the SQT-3 locus, this study has not attempted to fully characterise it. Instead, for a limited number of *sqt-3* mutants, the results obtained from crosses with TP12 can 1) be used to shed some light on the cuticle structure and 2) be used comparatively to analyse correlations between cuticle structure at the COL-19::GFP and morphological levels. Two temperature-sensitive glycine substitution alleles of *sqt-3*, *sqt-3(sc63)* and *sqt-3(e2117)*, which are lethal at 25°C, have been crossed with TP12 and the resulting COL-19::GFP patterns are shown in figures 4.4 A and B. The gross morphology of these *sqt-3* alleles at the permissive temperature (15°C) is more severely Dpy than the *sqt-1(e1350)* or *sqt-1(sc1)* mutants. However, they exhibit a weaker phenotype at the COL-19::GFP level, displaying less branching alae and annulae than the *sqt-1* N- and C- terminal processing mutants. This seemingly contradictory observation between gross morphological and COL-19::GFP phenotype is likely to be derived from the limited expression of *sqt-3* in embryos and L1 larvae only and the consequent



**Figure 4.4: *sqt-3* collagen mutants exhibit type IV COL-19::GFP disruption.** A and B) represent the COL-19::GFP pattern of two temperature sensitive alleles of *sqt-3* mutants, namely *sqt-3(sc63)* (A) and *sqt-3(e2117)* (B) when crossed with TP12. These are C- and N-terminal glycine substitutions respectively and both exhibit branching to the lateral alae (la) and the annulae. *sqt-3*-specific RNAi (feeding) was performed on L4 worms. Figure C depicts an adult from the F1 generation that, at the gross morphological level is wild type but, at the COL-19::GFP level, shows branching annulae. D) *dpy-17* shares highest sequence homology with *sqt-1* and the mutant allele, *dpy-17(e164)*, when crossed with TP12, shares a similar COL-19::GFP pattern with branching alae and annulae. In each figure, where present, lateral alae are marked "la", branching alae are denoted with a white arrow and branching annulae are denoted with a black arrow. Scale bars = 10  $\mu$ m.

presumption that SQT-3 is not a necessary component of the (COL-19::GFP-containing) adult cuticle.

The COL-19::GFP expression pattern exhibited by the wild type animals resulting from *sqt-3* RNAi targeting (feeding) is shown in figure 4.4 C. These results are obtained from animals incubated at 15°C because above this permissive temperature, the RNAi phenotype is embryonic lethal. Consistent with the wild type gross morphological phenotype exhibited by RNAi-treated animals at 15°C, the COL-19::GFP localises in a pattern of regularly spaced annulae and similarly wild type alae.

The *sqt-2* and *dpy-17* mutants and *dpy-17*(RNAi) display disruption patterns similar to the other Sqt mutants. The COL-19::GFP pattern of *dpy-17(e164)* is shown in figure 4.4 D. DPY-17 is a cuticle collagen that at the sequence level is most similar to SQT-1. It is thus not surprising that both collagens exhibit similar COL-19::GFP expression when mutant. The COL-19::GFP patterns of *sqt-2(sc3)* and *sqt-2(sc108)* are not shown but are similar to that of *sqt-1(e1350)* and *sqt-1(sc1)* (severe branching alae and annulae).

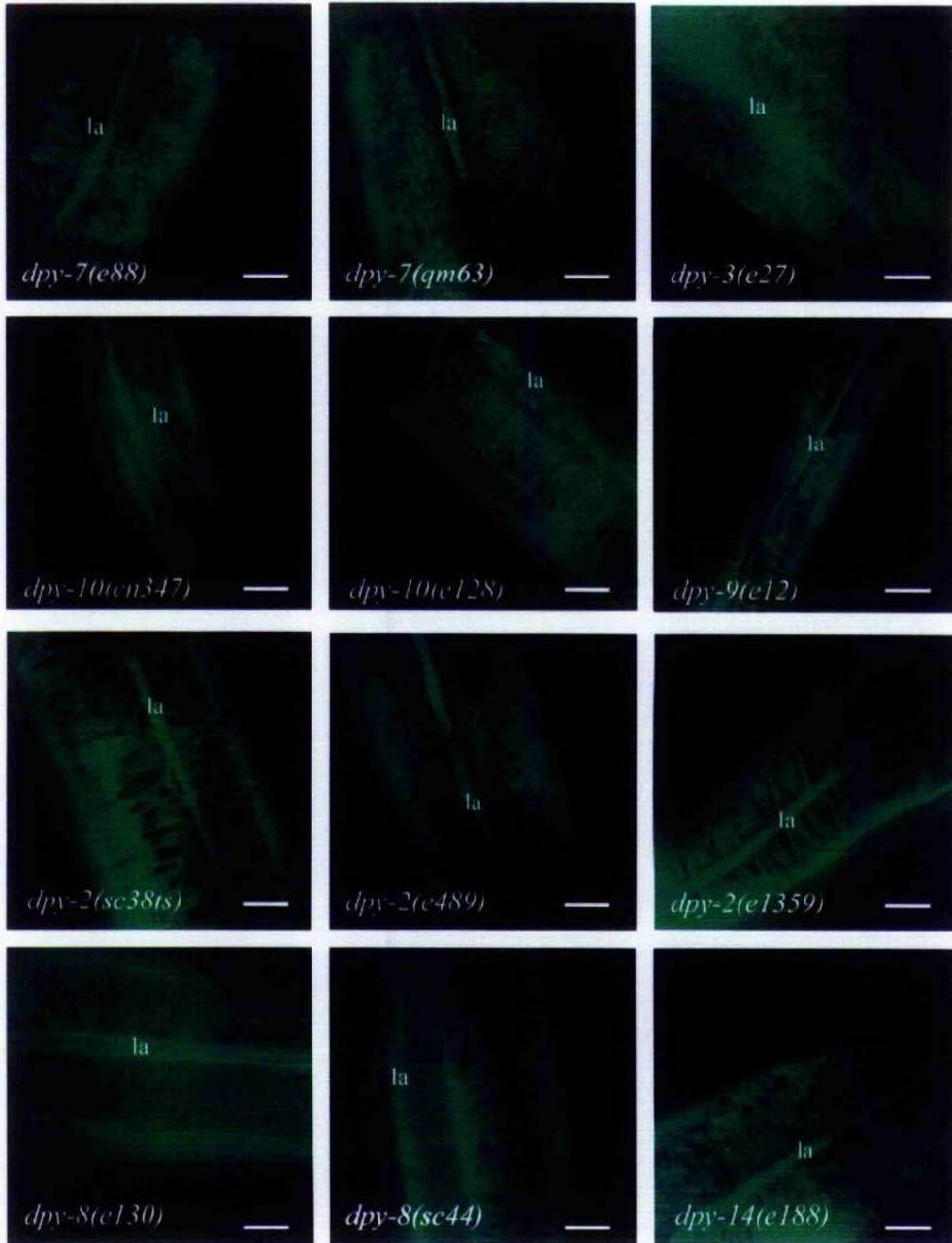
#### 4.2.5 The COL-19::GFP disruption in cuticle collagen Dpy mutants

Numerous alleles of eleven Dpy collagen loci were crossed with TP12 in order to observe the effect of the lesion on the cuticle composition and structure. The resulting COL-19::GFP patterns exhibited by these mutants made it evident that two types of Dpy exist (figure 4.5): ones that have an amorphous pattern of COL-19::GFP (class II disruption) (*dpy-2*, *dpy-8*, *dpy-10*, *dpy-7* and *dpy-3* mutants) (figure 4.5 A), and ones that exhibit branched seam cell annulae (class III disruption) (*dpy-5*, *dpy-4* and *dpy-13* mutants) (figure 4.5 B). During this study, *dpy-9* and *dpy-4* were also identified as being cuticle collagens that fall into the type II class and type III class respectively (described in section 4.3.7.1). These data again exemplify the variety in cuticle structure exhibited by different loci even when appearing similar at the gross morphological level. However, all examined alleles of the same gene produce a similar phenotype. It is interesting to note that from the mutants tested so far, an amorphous pattern of COL-19::GFP with wild type alae is unique to mutants whose loci are cuticle collagens. In the future, it will be interesting to discover whether any enzyme mutants result in this COL-19::GFP phenotype. The COL-19::GFP disruption patterns observed in these collagen mutants demonstrate that the disruption or ablation of one collagen, i.e. the collagen affected in the morphological mutant, may be sufficient to alter the expression pattern of another, i.e. COL-19::GFP.

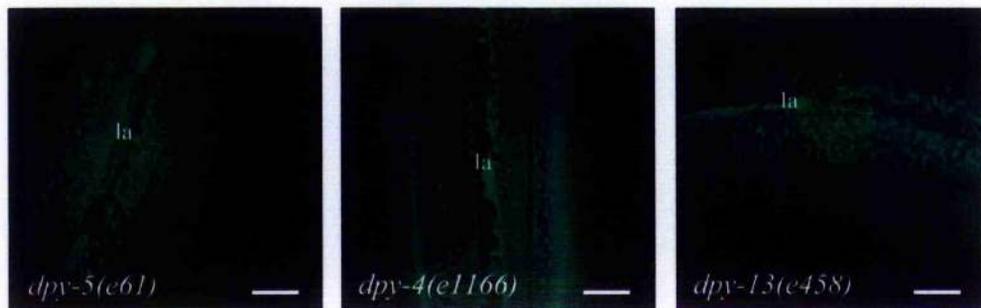
**Figure 4.5: Dpy collagen mutants exhibit one of two distinct types of COL-19::GFP pattern.** A) Alleles of *dpy-2*, *dpy-3*, *dpy-7*, *dpy-8*, *dpy-9*, *dpy-10* and *dpy-14* all exhibit amorphous annulae (class II COL-19::GFP disruption) and some show some levels of aberrant alae (labelled 'la'). B) Alleles of *dpy-5*, *dpy-4* and *dpy-13* show type III COL-19::GFP disruption. Alae are indicated (la). In images of *dpy-8(sc44)*, *dpy-4(e1166)* and *dpy-13(e428)* alae are not in the plane of focus. Scale bars = 10µm.

Facing page 168

**A: Class II Dpy collagen mutants**



**B: Class III Dpy collagen mutants**



As mentioned above, the collagens grouped in class II and class III in this screen coincide with the two groups, the DPY-7 group and the DPY-5 group, classified by McMahon *et al.* (2003), which have distinct temporal regulation (four hours and two hours prior to the moult respectively) and substructure localisation (annular furrow and annular cortex respectively). We might therefore be able to extrapolate from our data to provide information on collagens that were not part of the McMahon study. For example, since *dpy-9(e12)* mutants exhibit a COL-19::GFP pattern similar to that of the *dpy-7* class of temporally expressed mutants, it is likely that this locus also encodes for a cuticle collagen and that the collagen 1) is temporally expressed ~4 hours prior to the moult, 2) that it localises to the annular furrows, and 3) that it has a role as an obligate partner for the other collagens in the class.

The amorphous localisation of COL-19::GFP in the cuticles of TP12-crossed *dpy-7*, *dpy-2*, *dpy-3*, *dpy-10*, *dpy-8* and *dpy-9* mutants, suggests that, similar to the DPY-5 group of collagens (McMahon *et al.*, 2003), COL-19 is another collagen whose localisation is determined by the annular furrow structural scaffolds that are comprised of the DPY-7 set of collagens. Alternatively, COL-19::GFP's localisation might be dictated/guided by the presence of the *dpy-5* group of annular cortex-localised collagens. However, this is unlikely since in null alleles for these collagen, COL-19::GFP is still localised in an annular pattern in the dorsoventral region. It thus appears that the presence of correctly positioned annular furrows is sufficient for the proper localisation of COL-19::GFP in the cortex of the annulae. The *col-19* expression coincident with the moult, as demonstrated by SQ RT-PCR described in section 3.2.2, suggests that, in temporal terms, it is expressed after both the *dpy-7* and the *dpy-5* sets of collagens.

The class III and class II disruptions observed in collagen mutants are described below using *dpy-5* and *dpy-7*, respectively, as examples.

#### 4.2.5.1 Class III mutants are exemplified by *dpy-5* mutants

Class III Dpy mutants namely, "branched seam cell annulae", include the cuticle collagen mutants *dpy-5(e61)* and *dpy-13(e458)*. *dpy-4(e1166)* mutants also fall into this class. On the basis of the similar cuticle structure of this mutant compared to the *dpy-5* and *dpy-13* mutants, *dpy-4* was predicted, from these results, to encode a similar cuticle collagen. This was confirmed in subsequent studies that are detailed in section 4.3.7.1.

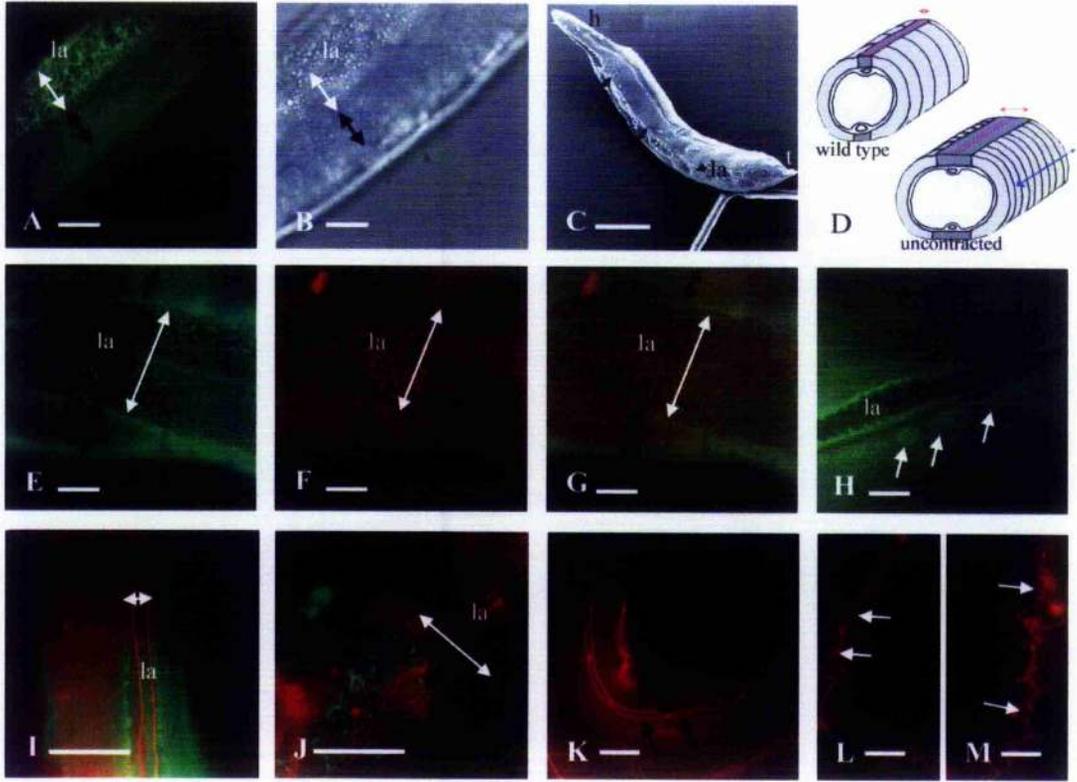
The COL-19::GFP pattern was examined in the cuticle of *dpy-5(e61)*, a Dpy mutant with a Gly-Opal mutation which most likely acts as a null mutant because of the position of the premature stop codon so close to the N-terminus of the collagen. This idea is strengthened by the similar gross morphological and COL-19::GFP phenotypes observed in this allele and in RNAi-treated animals (section 4.2.5.3). Mutants within this class exhibit two distinct COL-19::GFP patterns in different regions of the cuticle, and illustrate that the cuticle overlying the seam cells and that overlying the dorsoventral hypodermis have distinct compositions and structures. Figure 4.6 A clearly demonstrates that while the cuticle overlying the dorsoventral hypodermis exhibits regular circumferential bands consistent with narrow, but otherwise wild type, annulae, the cuticle overlying the seam cells and surrounding the alae is expanded in width and exhibits a branched COL-19::GFP pattern. When the cuticle is viewed by Nomarski (figure 4.6 B), the expanded region of COL-19::GFP disruption correlates with a bumpy cuticle surface while the dorsoventral region is relatively smooth. (Annulae are not in the plane of focus). COL-19::GFP fluorescence shows that in these mutants, alae are wild type in nature. The narrow nature of the annulae (approximately 0.75  $\mu\text{m}$  compared to 1.2  $\mu\text{m}$ ) in the dorsoventral cuticle of the worm is consistent with the shortened body length of dumpy worms.

The fact that the seam cell cuticle is uncontracted means that, in contrast to wild type worms, the annulae do not appose the alae. The uncontracted nature of the cuticle is confirmed by SEM imaging that shows a region where the dorsoventral cuticle has failed to join above the seam cells (figure 4.6 C). A COL-19::GFP pattern similar to, but less severe than that of the *dpy-5(e61)* mutants implying the existence of two distinct regions of the cuticle, is observed in very old wild type (TP12) adults, which become slightly Dpy (refer back to figure 3.4). This suggests that even in wild type worms, the compositions of the two regions of the cuticle are distinct and that in older worms, the seam cell cuticle is weakened by age (rather than mutation) and loses its ability to fully contract. Thus, failure of the seam cell cuticle to contract, whether due to mutation or to age, is consistent with the widened body of the worm in Dpy animals, as illustrated in 4.6 D.

The wild type COL-19::GFP pattern of the cuticle overlying the dorsoventral hypodermis can be explained by the fact that *dpy-5* worms still express wild type forms of the *dpy-7* class of temporally expressed collagens (DPY-7, DPY-8, DPY-2, DPY-10, DPY-3) (McMahon *et al.*, 2003). Consequently, *dpy-5* mutants retain the structural scaffold of the annular furrows. This

**Figure 4.6: *dpy-5(e61)* exemplifies type III COL-19::GFP disruption.** A) Two distinct regions of the cuticle become apparent at the COL-19::GFP level when TP12 is crossed with *dpy-5(e61)*. A regular annular pattern is observed in the cuticle overlying the dorsoventral region (black double headed arrow). Conversely, the pattern of COL-19::GFP overlying the seam cells is branched (white double headed arrow). The latter region represents the uncontracted seam cell cuticle. B) Distinctions between the two regions of the cuticle are also apparent via Nomarski imaging: The cuticle appears bumpy in the region corresponding to the COL-19::GFP disruption (double headed arrow). C) The uncontracted nature of the cuticle is observed via scanning electron microscopy (SEM) of *dpy-5(e61)* adult animals. At this magnification, a notch where contraction has failed is evident (double headed arrows) and corresponds to the position of the lateral alae ('la' and single headed arrow). Further evidence of the failure of the cuticle to contract specifically in regions of the seam cell-derived cuticle is shown by the fact that these mutants exhibit wild type head ('h') morphology since this region is not seam cell-derived. D) A schematic representation of the uncontracted seam cell-derived cuticle compared to wild type. The two distinct regions of the cuticle are depicted (seam cell-derived = dark grey, dorsoventral hypodermal-derived = light grey). The dark grey region represents the region of COL-19::GFP branching that overlies the alae (alae are shown by pink lines). The uncontracted nature thus results in the increased body width of the worm and thus accounts for the *Dpy* phenotype. E) COL-19::GFP pattern of a *dpy-5(e61)* worm which, as depicted in F) has been additionally labelled via immunolocalisation using antibodies against DPY-7. The figure shows aberrant DPY-7 localisation: wild type annulae overlying the dorsoventral hypodermis (black double headed arrows) with a branched pattern overlying the seam cell hypodermis (white double headed arrow) is observed. G) Merged image of COL-19::GFP and DPY-7 localisation shows that the regions of disruption coincide. H) Another characteristic of the COL-19::GFP pattern of *dpy-5(e61)* mutants is the accumulation of the tagged collagen in the periphery of vesicles around the seam cell cords (arrows). I-M) Comparison of MH27 immunolocalisation on wild type (I) and *dpy-5(e61)* (J) worms. MH27 antibodies localise to the hypodermal junctions. Adult *dpy-5(e61)* animals have seam cells that are fused but that exhibit branches. The distance between the junctions (black double headed arrows) does not coincide with the width of cuticle disruption (white double headed arrow). K-M) Comparison of MH27 immunolocalisation on wild type (K) and *dpy-5(e61)* larvae (L and M). Larvae of TP12s that show a pattern of regular and similar shaped cells (K) (arrows). This is in contrast to *dpy-5(e61)* larvae which exhibit seam cells which are aberrant and/or multiple (arrows). This is consistent with the branched pattern observed in adults.

In all figures, where present, lateral alae are denoted 'la' but in H, I and J are out of the plane of focus. Scale bars = 10µm in A, B, and E-J, 1.2µm in H and 15 µm in K-M.



was confirmed by the wild type annular furrow localisation of anti DPY-7 antibodies observed in *dpy-5(col-19::gfp)* worms after immunolocalisation (figures 4.6 F and G). In wild type animals, the annular furrows dictate the deposition of later expressed collagens, such as DPY-5 and DPY-13, and it is likely that once a wild type pattern of annulae has been set, even mutant forms, or ablation of, *dpy-5* and *dpy-13* collagens do not disturb the circumferential annular furrow pattern. It is interesting to imagine that there is a reduction in the amount of collagen that is deposited between the furrows, and consequently a decreased width of the annulae. This is consistent with the narrower annulae and the reduced length of the worm.

While the dorsoventral hypodermal region of the cuticle appears to be wild type apart from having a reduced collagen content in the annular cortex, the cuticle overlying the seam cells appears to be much more susceptible to the apparent reduction of the DPY-5. It is conceivable that, due to the lack of properly formed annulae in this seam cell-derived cuticle (corroborated by the fragmented pattern of DPY-7 staining in this area of the cuticle), this region lacks the regularly spaced circumferential structural scaffold that is present in the dorsoventral cuticle. In such a region lacking such a scaffold, the exclusion of a collagen may be much more disruptive. The weakening of the seam cell-derived cuticle and its failure to contract has been attributed to the wide body shape (figure 4.6 D) and, therefore, this region has a likely role in dictating circumferential contraction. On the other hand, the annulae of the dorsoventral hypodermis are probably involved in the lateral contraction (figure 4.6 D). The requirement for correct seam cell-derived cuticle to direct diametric shrinkage was proven previously when ablation of individual seam cells resulted in animals with greatly weakened cuticles (and gaps in their alae) (Singh and Sulston, 1978).

Cuticular weakness being restricted to seam cell-derived cuticle is confirmed by SEM analysis which reveals that *dpy-5(e61)* adult worms exhibit relatively wild type head morphology (refer to figure 4.6 C). The lateral seam cell cords terminate posterior to the head region while the anterior region is composed, not of seam cells, but of hypodermal cells, *hyp 3-6*.

Another interesting feature of the *dpy-5* mutants and the RNAi-treated worms is the accumulation of the COL-19::GFP marker molecule in the periphery of the seam cell cords (figure 4.6 H). It appears that the absence of one mutant collagen (e.g. DPY-5) might also result in the retention of other collagens (e.g. COL-19::GFP) in the ER. Such a reduction in a number of cuticle collagens may be another factor involved in the overall weakening of the cuticle as it

may have an additive effect. Thus, the decreased width of the annulae might be a consequence of preventing the insertion of, not only the DPY-5 class of collagens into the cuticle, but also an array of other collagens, into the cuticle.

These data show that DPY-5 is a collagen that plays a critical role in the control of morphology and the width of the annulae. It is not essential for the formation of the annular furrows but does “pad out” or at least regulate this “padding out” of the cortex of the annulae.

#### 4.2.5.2 MH27 staining in *dpy-5(e61)* mutants

In order to determine whether the disruption observed in *dpy-5* mutants in the cuticle overlying the seam cells is reflected in the hypodermal cells themselves, a monoclonal antibody raised against a protein in the adherin junctions around the hypodermal cells, MH27, was used (figures 4.6 I-M). Such staining in adult *dpy-5(e61)* worms reveals that the seam cells are able to fuse, resulting in the adherin junctions being a similar distance apart as in wild type (TP12) animals (compare figures 4.6 I and J). These observations reveal that the width of the band of seam cells does not correspond to the width of the disrupted seam cell-derived cuticle. However, the fused cells in *dpy-5(e61)* mutants are aberrant in that they exhibit branching (Figure 4.6 J). Consistent with the abnormal seam cells in adults, the MH27 staining in larvae is also altered (figures 4.6 L-M). Instead of appearing as a regular pattern of long rectangular cells (4.6 K), *dpy-5* mutant (L3s) show cells alternating between long and round cells, the latter of which are sometimes duplicated (figures 4.6 L and M). It is interesting to note such aberrations since it appears that mutations to the cuticle are sufficient to alter the seam cells. Seam cell disruptions are also observed upon *pdi-3* RNAi treatment in a *dpy-18(e364)* background. PDI-3 is an enzyme associated with cuticle synthesis via its dual cuticle cross-linking activities: TGase activity for the formation of bonds between glutamine and lysine, and protein disulphide isomerase activity for the formation of disulphide linkages. The disruption observed in *dpy-5*(RNAi) larvae is less severe than those treated with *pdi-3* (Eschenlauer and Page, 2003).

#### 4.2.5.3 The effect of treating different stage larvae with *dpy-5* RNAi

The general understanding of how a new cuticle is formed is that it is synthesised beneath the old cuticle which is subsequently shed. In order to test whether cuticle disruption of a previous stage has any implications on the synthesis of the new cuticle, TP12 embryos and different stage larvae worms were subjected to *dpy-5* RNAi by feeding. If disruption in a cuticle does affect later cuticles, it would be expected that RNAi-treated embryos and early larvae would exhibit

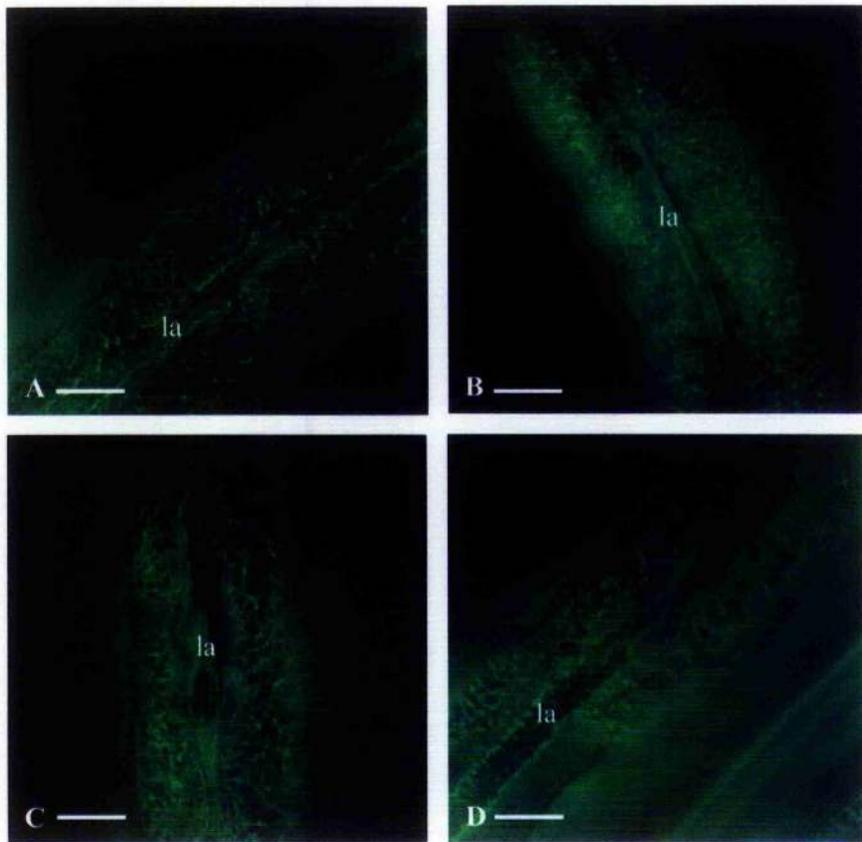
more disruptions in their adult cuticle as result of a series of disrupted cuticles being synthesised upon each other. Conversely, RNAi-treatment of IAs would be expected to result in less disruption as the cuticle upon which the final cuticle is synthesised, would be wild type.

The data resulting from these RNAi experiments (figure 4.7) reveal that there are no severe differences between the early-treated and L4-treated worms, and illustrate that there is little or no amplification in the cuticle disruption.

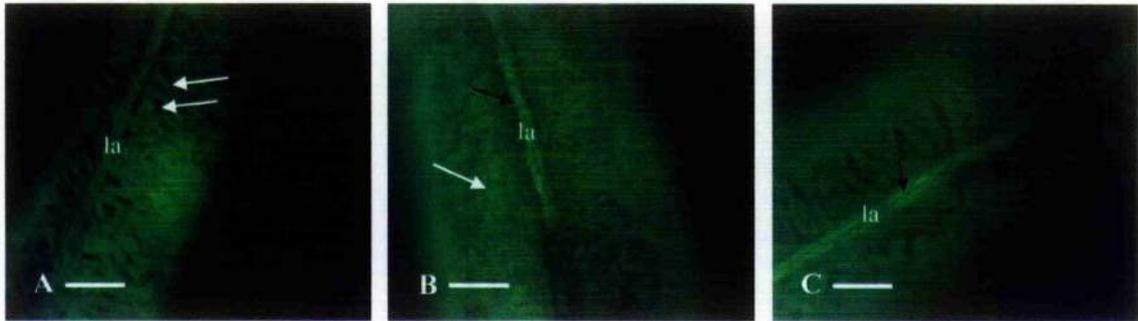
#### 4.2.5.4 Class II COL-19::GFP disruption is exemplified by *dpy-7* mutants

The cuticle collagen *dpy-7* mutants exemplify the class II type of COL-19::GFP disruption. All of the mutants tested to date that fall into this class (*dpy-7*, *dpy-2*, *dpy-3*, *dpy-10*, *dpy-8* and *dpy-9* mutants) are collagen lesions (refer to table 4.1 for a description of the molecular lesions). The *dpy-7* alleles, *dpy-7(e88)* and *dpy-7(qm63)*, a glycine substitution and a null mutant respectively, were crossed with TP12. *dpy-7* was also targeted through RNAi feeding. Both the glycine substitution and the RNAi-treated worms show a ubiquitous pattern of COL-19::GFP disruption, covering both the cuticle overlying the seam cell and dorsoventral hypodermis, giving an appearance of what we have described as “tiger stripes”. The two different regions of the hypodermis cannot be distinguished due to the lack of annulae. Instead of a disruption pattern of branched strands observed in the aberrant seam cell cuticle in class III mutants, in *dpy-7(e88)* mutants the pattern appears as amorphous finger-shaped projections that point in multiple directions (figure 4.8 A). A more severe manifestation of this type of COL-19::GFP disruption occurs in the *dpy-7* null allele (*dpy-7(qm63)*), when the discrete fingers are no longer apparent and the entire cuticle has an amorphous COL-19::GFP pattern (figure 4.8 B). There is an increased severity of the COL-19::GFP disruption in the null mutant (*dpy-7(qm63)*) compared to the RNAi-treated animals (figure 4.8 C). This suggests that the *dpy-7* signal is not being completely ablated upon RNAi-treatment, and that there is potentially a small amount of the collagen in the cuticle. Despite this, the similarity in the patterns of the *dpy-7* alleles and the RNAi-treated animals suggests that the phenotype of *dpy-7* glycine substitutions is due to the retention of most of the mutant glycine substituted collagen in the ER. Immunofluorescence of the DPY-7 protein using anti-DPY-7 antibody failed to show any/minimal cuticular staining.

Loss of the DPY-7 collagen in the cuticle prevents the deposition of the annular furrows; consequently, the surface is smooth. With no structural scaffold to define their deposition, the gene products of the later expressed collagens such as *dpy-5*, *dpy-13* and, as shown here,



**Figure 4.7: Effects of treating different larval stages of TP12 animals with *dpy-5* RNAi feeding.** The aim was to specifically target *dpy-5* in order to observe any additive effects from each moult. A) represents adult worms that have developed from embryos placed on feeding plates while B, C and D) represent adult worms whose RNAi feeding was initiated when they were at L1-L2, L3 and L4 stages respectively. No obvious increase in severity is obtained by feeding the worms from earlier stages. Lateral alae are denoted by 'la'. Scale bars = 10 $\mu$ m.



**Figure 4.8: *dpy-7* mutants crossed with TP12 exemplify class II COL-19::GFP disruption.** A) *dpy-7(e88)* is a glycine substitution. The amorphous pattern resembles fingers/tiger stripes that point in multiple directions (white arrows). Lateral alae ('la') appear slightly amorphous (black arrow). B) The null allele, *dpy-7(qm63)*, exhibits the more severe form of the COL-19::GFP phenotype, showing less defined projections and a more ubiquitous amorphous pattern (white arrows). The alae ('la') appear to be more amorphous than in the *dpy-7(e88)* animals (black arrow). C) Upon specific targeting of *dpy-7* by RNAi feeding, a similar amorphous pattern is observed in adult F1s as for *dpy-7(e88)* animals. Lateral alae ('la') are also amorphous (black arrow). Immunolocalisation using antibodies against DPY-7 was carried out but consistent with the lack of DPY-7 in the cuticle, no staining was observed (not shown). Scale bars = 10µm.

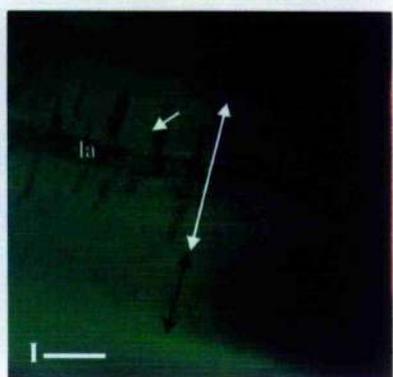
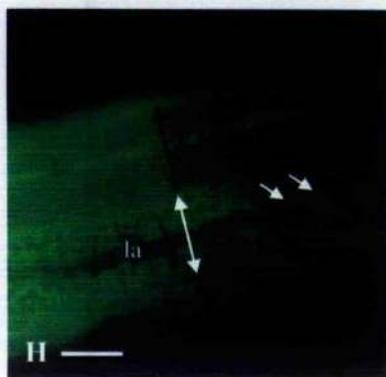
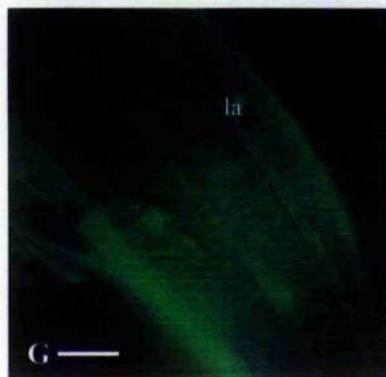
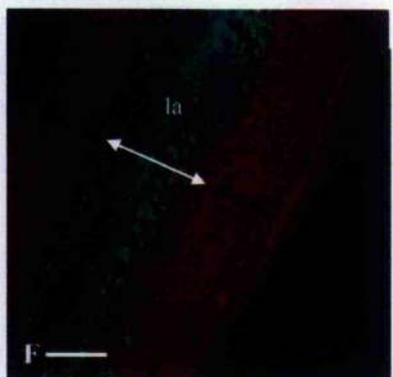
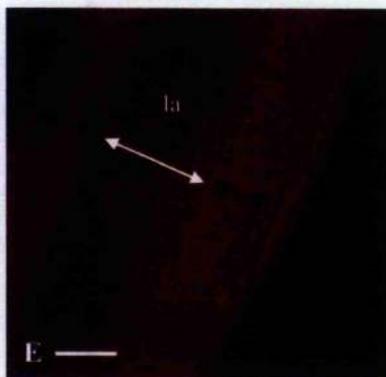
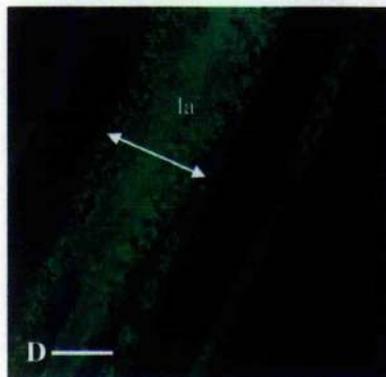
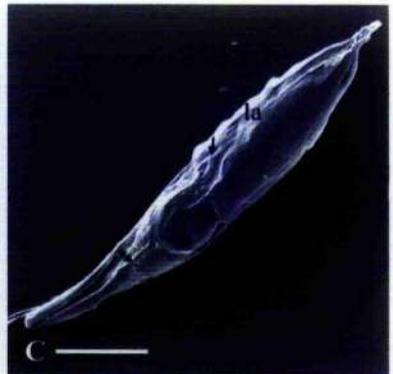
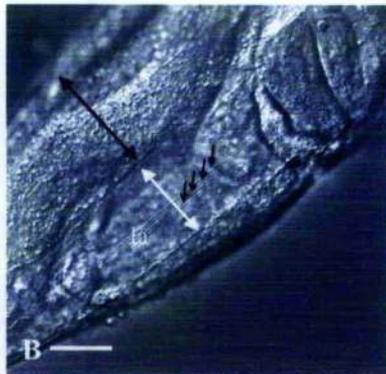
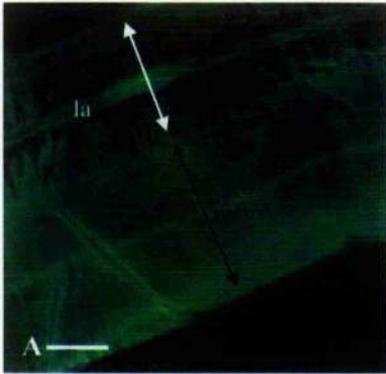
COL-19, are localised aberrantly. There are two possible explanations as to why there is an increased body width in these mutants. The first is that increased body width would result in a similar manner as described for *dpy-5* mutants, namely, that the decrease in DPY-7 collagen weakens the seam cell-derived and diametric shrinkage-controlling cuticle. However, the most likely reason for increased body width is that loss of annulae, as well as the improper localisation of other collagens (e.g. DPY-5 (McMahon *et al.*, 2003) and COL-19::GFP), weakens the dorsoventral region of the cuticle sufficiently to allow it to stretch in response to the hydrostatic pressure of the worm.

#### **4.2.6 The COL-19::GFP disruption resulting from mutations in the collagen processing enzymes display type V COL-19::GFP disruption**

##### **4.2.6.1 COL-19::GFP disruption in *dpy-11* mutants**

The *dpy-11* locus encodes for a novel thioredoxin that is involved in regulation of body- and male tail- morphology (Ko and Chow, 2002). The medium Dpy *dpy-11(e224)* is a partial loss of function allele, while the severe Dpy *dpy-11(e1180)* is a complete null. COL-19::GFP phenotypes of these alleles are illustrated in figures 4.9 A-G. *dpy-11(e224)* animals display a similar but not identical pattern to that of class III mutants, that is, they exhibit two distinct regions of the cuticle. Overlying the dorsoventral regions, a regular, yet slightly constricted, annular pattern is apparent while overlying the seam cells the cuticle is uncontracted and has branched annulae (figure 4.9 A). Similar to *dpy-5(e61)* animals, two distinct regions can be observed via Nomarski imaging. The region above the seam cells appears smooth. What distinguishes *dpy-11* mutants from those in class III (and puts them into class V) is the fact that there are alae defects. This is illustrated by the multiple alae in the Nomarski image of figure 4.9 B and the branching alae in the SFM image of an adult *dpy-11(e224)* hermaphrodite (figure 4.9 C). The SEM image also depicts the uncontracted nature of the cuticle above the seam cells. Upon immunolocalisation using DPY-7 antibodies, the fragmented and prematurely terminated nature of the annulae above the seam cells is again illustrated (Figure 4.9 E). The regions of COL-19::GFP and DPY-7 disruption co-localise (figure 4.9 F). In the corresponding COL-19::GFP figure (figure 4.9 D), the disruption overlying the seam cells is branched, and is also slightly amorphous.

**Figure 4.9: *dpy-11* and *dpy-18* mutant phenotypes.** A) Partial loss of function mutant animals of the *dpy-11* locus (*dpy-11(e224)*) exhibit a COL-19::GFP pattern similar to type III COL-19::GFP disruption. They are placed into a separate class (class V) due to their branching and amorphous alae. However these are not depicted in A. B) In a similar manner to *dpy-5* mutants, the two distinct regions can also be distinguished upon Nomarski imaging. In this instance, lateral alae ('la') are not branched but are multiple, as indicated by the single headed arrows (4 instead of 3). C) Scanning electron microscope image of a *dpy-11(e224)* adult hermaphrodite. The mid body ('m') region is particularly Dpy. Branching alae are indicated with an arrow. Regions of uncontracted seam cell-derived cuticle are observed and indicated by a double headed arrow. D-F) Immunolocalisation using DPY-7 antibodies was carried out. Similar to that seen with COL-19::GFP (D), the extended region of fragmented alae is observed in figure E. In the dorsoventral region of the cuticle, annulae are more regular and wild type. F shows the merged COL-19::GFP and DPY-7 images and indicates that the regions of COL-19::GFP disruption coincide with regions of DPY-7 disruption. G) Another allele of *dpy-11*, *dpy-11(e1180)* is a null mutant. The COL-19::GFP pattern is more amorphous than as seen for the *dpy-11(e224)* mutant. Lateral alae are amorphous. H) COL-19::GFP disruption of *dpy-18(e364)*, which is a null allele. Two distinct regions of the cuticle are apparent. In addition to the branching annulae in the region overlying the seam cells, projections similar to those observed in *dpy-7* mutants are observed (white arrows). I) COL-19::GFP localisation of another *dpy-18* null allele, *dpy-18(ok162)*. Two distinct regions are again observed. The annulae are extremely fragmented in the region overlying the seam cells. In all figures, lateral alae are denoted with 'la' but in figures D, E, F, H and I, they are not in the plane of focus. In A, B, D, E, F and H, the two distinct regions of the cuticle are denoted with double headed arrows: the uncontracted regions overlying the seam cells are shown with white arrows while the wild type dorsoventral-derived regions are denoted with black arrows. Scale bars in C= 150µm. In the rest, scale bars = 10 µm.



The null allele, *dpy-11(e1180)*, exhibits an amorphous pattern of COL-19::GFP disruption (figure 4.9 G). It also differs from the partial loss of function allele since the pattern of disruption is ubiquitous over all regions of the cuticle. *dpy-11(e1180)* animals also exhibit branching alae (not shown).

These data demonstrate that the thioredoxin encoded by the *dpy-11* locus is essential for proper cuticle formation, and mutation to or ablation of the gene affects both the alae and the annular structures. The ability to distinguish between the two distinct regions of the cuticle in *dpy-11(e224)* mutants and the similarity to the *dpy-5* collagen mutants is interesting since it is possible that the reduced efficiency of this partial loss-of-function form of the enzyme affects DPY-5 and its obligate partners that reside in the cortex of the annulae, more severely than the class of annular furrow localised collagens (e.g. DPY-7). In such a scenario, the wild type DPY-7 collagens can form the structural scaffolds of the annular furrows and thus produce the regular pattern of COL-19::GFP observed in these animals. Conversely, in *dpy-11(e1180)* animals, the regular annular pattern is not observed and thus it is expected that complete ablation of the enzyme activity does indeed prevent the proper positioning of the class of DPY-7 collagens. This is evidence that both types of collagen may be the targets for the enzyme. However, the enzyme may have different specificities for different collagens, as demonstrated by the *dpy-11(e224)* mutant, whose G76E lesion within the thioredoxin-like (TRX) domain, specifically affects collagens that are localised in the seam cell-derived cuticle while not affecting the set of DPY-7 collagens.

#### 4.2.6.2 COL-19::GFP disruption in *dpy-18* mutants

*dpy-18*, also known as *phy-1*, is one of the three PHY-encoding genes in the *C. elegans* genome. Together with the other PHY ( $\alpha$ ) subunit, PHY-2, and two PDI-2 ( $\beta$ ) subunits, an active mixed tetramer is formed ( $\alpha_2\beta_2$ ). It has now been established that deletion or disruption of *dpy-18* in these morphological mutants leads to the consequential loss of the tetramer form of the enzyme (Myllyharju *et al.*, 2002). This loss, however, is compensated for by the up-regulation of active PHY-2/PDI-2 dimers. As a result, *dpy-18* mutants and *dpy-18* RNAi lead to viable yet Dpy animals rather than the larval lethality that is caused by the ablation of the only encoded  $\beta$  (PDI-2) subunit.

*dpy-18* mutant alleles, *dpy-18(e364)* and *dpy-18(ok162)* are both considered to be functionally null (Friedman *et al.*, 2000). Both mutants exhibit similar COL-19::GFP disruption patterns

which are reminiscent of type III mutants but differ from this class in the fact that they have branched alae and are thus placed into the class of type V COL-19::GFP disruption (figures 4.9 H and I). *dpy-18* worms are wild type in the dorsoventral region of the cuticle but show severe disruption over the seam cell hypodermis. It is interesting to note that the *dpy-18(ok162)* mutants (figure 4.9 I) exhibit a much more fragmented pattern (shorter fragments) around the alae than their *dpy-18(e364)* counterparts (figure 4.9 H), despite both alleles being null.

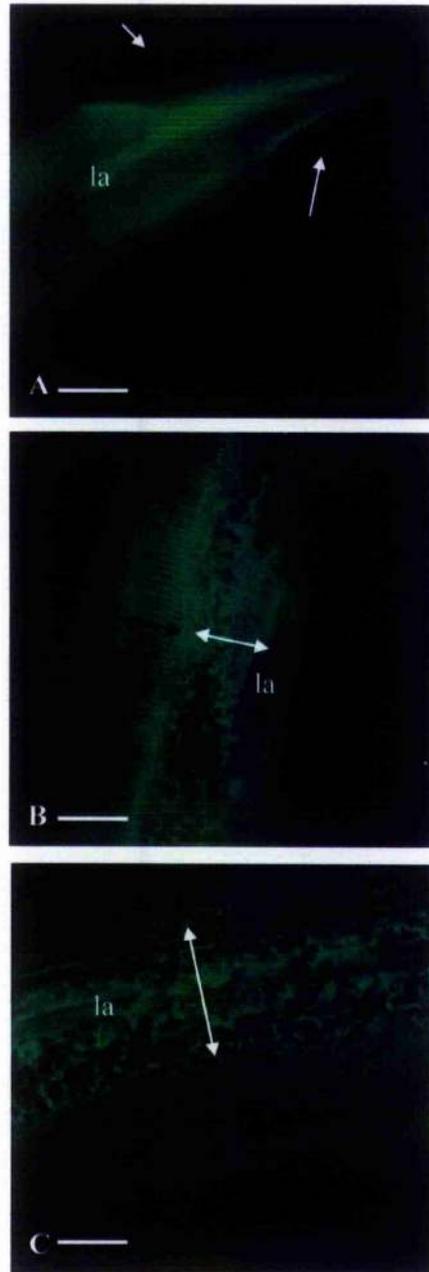
#### 4.2.7 The COL-19::GFP expression pattern in Bli mutants

Bli morphological mutants are a result of lesions in genes encoding for cuticle collagens (*bli-1* and *bli-2*) (Crew and Kramer, personal communication), enzymes involved (or predicted to be involved) with cuticle synthesis (*bli-4*) (Thacker *et al.*, 1995) or in loci probably involved in the regulation of cuticle processing (e.g. *bli-5*) (Simmer *et al.*, 2003). Despite all exhibiting a blister phenotype, these mutants fall into either type I or type III COL-19::GFP classes of disruption (figure 4.10)

##### 4.2.7.1 BLI-4 subtilisin protease mutants exhibit a wild type COL-19::GFP expression pattern

The subtilisin-like protease mutant, *bli-4(e937)* exemplifies cuticle disruptions that lead to gross morphological phenotypes but which are not reflected in the COL-19::GFP-containing cuticle layer (figure 4.10 A). Despite these mutants exhibiting adult specific blisters, indicative of strut aberrations, they display a wild type COL-19::GFP pattern in overlying layers.

It has already been established that alleles which completely ablate all of the 9 isoforms (A-I) (and thus the function) of this protease are lethal, but that the lesion in *bli-4(e937)* affects only the A, E, F, G and H isoforms which have been implicated in adult cuticle formation (Thacker *et al.*, 1995). From the wild type COL-19::GFP pattern displayed by these mutants (figure 4.10 A), it appears that despite the loss of this isoform, annulac form properly, a fact which implies that the DPY-7 set of collagens are being properly processed. This TP12 x *bli-4* cross further reveals therefore, that the deleted isoforms in this allele are vital for the processing of only a subset of the cuticle components and confirms previously speculated data that only strut associated collagens, BLI-1 and BLI-2, are affected by the loss of the 5 isoforms in *bli-4(e937)* mutants. The above evidence suggests that BLI-1 and BLI-2 may be the only collagens that form aberrant structures in the absence of the lost isoforms, but it cannot definitely prove that the other collagens are not processed by these isoforms at all. It is possible that the lost isoforms do



**Figure 4.10: Bli mutants display either wild type or type III COL-19::GFP patterns.** A) *bli-4(e937)* mutants show a regular pattern of circumferential annulae. COL-19::GFP does not localise to the regions of the cuticle that are blistered (arrow). B and C) *bli-1(e769)* and *bli-5(e518)* mutants both display type III COL-19::GFP localisation and the distinct seam cell- and dorsoventral- derived regions of the cuticle are denoted with white and black double-headed arrows respectively. Alae are indicated ('la'). Scale bars = 10 $\mu$ m.

indeed have roles in processing the DPY-7 class of collagens, but that the presence of the N-propeptide in the cuticle is not disruptive enough to alter their proper deposition (no N-terminal processing mutant alleles are available for *dpy-7*).

#### 4.2.7.2 The COL-19::GFP expression pattern in BLI-1 mutants

The cuticle collagen mutant *bli-1(e769)* exhibits blisters at the gross morphological level (as do RNAi-treated worms) (Kamath *et al.*, 2003), and when crossed with TP12, displays type III COL-19::GFP disruption (figure 4.10 B). This illustrates that this strut-localised collagen is also essential for the proper formation of the cuticle. The disruption observed in the seam cell-derived cuticle in these mutants might be an indication of the presence of BLI-1 in this region of the cuticle. This is interesting since, to date, this collagen has been assumed to be restricted to the struts, which are presumably only associated with the dorsoventral region. Alternatively, the disruption in the hypodermis overlying the seam cells may result from interactions between the strut-localised BLI-1 collagen with other collagens which do localise above the seam cells. Annulae in these mutants do appear to be wild type, despite the strut aberrations. This suggests that in adult cuticles, proper strut formation is not a prerequisite for the placement of the annular furrows, and that the aberrations in the strut layers are not transferred to the overlying ones.

#### 4.2.7.3 The COL-19::GFP expression pattern in *bli-5* mutants

A lesion in *bli-5*, a serine protease inhibitor (Simmer *et al.*, 2003), results in the Bli phenotype. *bli-5(e518)* was crossed with TP12 in order to visualise the effects that mutation to such a potential regulator of cuticle synthesis has on the cuticle. *bli-5(e518)* animals exhibit a COL-19::GFP pattern that is wild type over the dorsoventral hypodermis and disrupted and branched over the seam cell hypodermis (figure 4.10 D). Such a pattern is reminiscent of type III disruption, and is a pattern that is shared by the *bli-1* cuticle collagen mutant (figure 4.10 B). RNAi of *bli-5* was also performed, and such treated worms exhibited a similar pattern as the mutant strain (not shown). This suggests that despite being uncharacterised, the lesion in the mutant strain has a similar effect as the RNAi phenotype. The data presented here suggest that firstly, BLI-5 must have a role (direct or indirect) in the processing of BLI-1 and BLI-2 due to the Bli phenotype exhibited. Due to the aberrations observed in the cuticle overlying the seam cells, it must also have a role in the processing of other collagens. Because of the similarities of the COL-19::GFP phenotypes to alleles of *dpy-5* and *dpy-13*, it is likely that this inhibitor has a role in the processing of these collagens. On the other hand, annulae are normally formed in the

dorsoventral region in these mutants, which implies that the DPY-7 family of collagens are being processed into the cuticle normally, despite the aberrant form or loss of the inhibitor. There is a similar discrepancy as described in *bli-4* mutants i.e. whether the wild type localisation of these collagens is due either 1) to these collagens being properly processed because of the inhibitor not having a role in their processing, or 2) whether the inhibitor does have a role in the processing of these collagens but that the effects of preventing its proper functioning does not have a severe effect on the incorporation of these annular furrow-localising collagens into the cuticle.

### 4.3 Results 2: An RNAi screen to identify components in cuticle synthesis

An array of enzyme- and collagen- encoding genes have been identified as the loci responsible for the morphological mutants described above, and these enzymes have consequently been implicated in playing important roles in cuticle biogenesis or moulting. These include: ones with cross-linking roles, such as thioredoxins like DPY-11 that catalyse the formation of disulphide bonds; ones with roles in hydroxylation, such as DPY-18 that encodes one of the  $\alpha$  subunits of a prolyl 4-hydroxylase (PHY); and ones implicated in N- and C-terminal proteolytic processing e.g. BLI-4 subtilisin-like N-terminal protease (Thacker *et al.*, 1995) and DPY-31/NAS-35 (an astacin C-terminal protease) (Novelli *et al.*, 2004).

In addition to the loci identified from morphological mutants, other enzymes have been implicated in the formation of the cuticle due to their RNAi/deletion phenotypes and biochemical characterisation. These include: 1) PDIs that catalyse the formation of disulphide cross-links (via thioredoxin activity) and  $\epsilon$ -( $\gamma$ -glutamyl)lysine links (via transglutaminase activity) as well as act as chaperones (and PDI-2 being the necessary  $\beta$  subunit of prolyl 4-hydroxylase); 2) peroxidases, such as BLI-3, that catalyse the formation of tyrosine-derived cross-links; 3) tyrosinases, which are hypothesised to exhibit cross-linking activities in the *C. elegans* cuticle (Blaxter personal communication, M. Sapio personal communication); and 4) peptidyl-prolyl *cis-trans* isomerase (PPI) enzymes that have a role in the formation of the cuticle, the latter having an essential isomerase step required for helix propagation.

Finally, certain proteases are also involved in cuticle synthesis, the main protease families implicated in having such morphology-determining roles namely ADAMs/ADAMTSs, astacins and cysteine proteases. Via proteolytic activity on different substrates, these can function to either degrade (e.g. GON-1) (Blelloch *et al.*, 1999) or synthesise the ECM (e.g. DPY-31/NAS-37) (Novelli *et al.*, 2004). The former role is related to tissue remodelling during

body/organ morphogenesis and to the moulting process, while the latter role involves the processing of collagens to facilitate their proper incorporation into the cuticle (Kuno and Matsushima, 1998; Blelloch *et al.*, 1999; Novelli *et al.*, 2004). Based on morphological phenotypes, only two (DPY-31/NAS-35 and NAS-37) of the 40 members of the astacin family have, to date, been positively attributed to having roles in cuticle synthesis/moulting. DPY-31/NAS-35 is a C-terminal collagen protease that is necessary for removing the globular collagen propeptides of SQT-3 (and presumable other members of the SQT-3 family). NAS-37 was first identified for its role in ecdysis of the old cuticle (Davis *et al.*, 2004) but due to the cuticle defects exhibited by RNAi-treated animals, this protease is also thought to contribute to cuticle synthesis by processing a subset of collagens. The proteases of the ADAM/ADAMTS family (a disintegrin-like and metalloproteinase/with thrombospondin type I motif), are believed to be involved in tissue remodelling. Mutants of ADT-1 and GON-1, show male tail and gonadal defects respectively (Blelloch *et al.*, 1999). The mechanism by which GON-1 regulates gonadogenesis is, as yet, unknown. However, it is hypothesised that its action involves the degradation of components of the ECM in order to permit the expansion and morphogenesis (Blelloch *et al.*, 1999).

Yet another potentially interesting family of proteases is the carboxypeptidases which are characterised by the presence of an M14 protease domain. 4 of the 37 proteins with M14 domains also encode cysteine protease domains. One such cysteine (M14) protease is cathepsin L-like protease of *C. elegans*, ceCPL-1, which has been demonstrated to be necessary for embryogenesis (Britton and Murray, 2002). On the basis of *O. volvulus*, *Brugia pahangi* and *Brugia malayi* ceCPL-1 homologues which are secreted coincident with moults and produce moulting defects upon treatment with specific inhibitors (Guiliano *et al.*, 2004), ceCPL-1 is also likely to be involved in moulting. Cathepsin Z-like proteases (CPZs) have also been assigned to having a role in moulting due to their temporal, spatial and moult defect RNAi phenotypes (Hashmi *et al.*, 2004). Also of note is the fact that two M14 encoding enzymes within this family have additional DNA-J motifs, which are associated with chaperone function.

In addition to enzymes that synthesis the cuticle, there are many factors that contribute to proper body morphology. This is especially true for components of systems that regulate the establishment of the vermiform shape of the worm during embryogenesis since these occurs prior to the synthesis of the first cuticle. As described in section 1.3, the final elongated shape of the animal results from hypodermal remodelling, which involves the intercalation of hypodermal

cells to form a dorsal surface, enclosure along the ventral surface and then fusion of the hypodermal cells to form multinucleate syncytia. Finally, elongation occurs. This is via a combination of actin-mediated processes, involving the polymerisation and subsequent constriction of actin filaments that are orientated circumferentially at regular intervals along the length of the embryo (Priess and Hirsh, 1986), and myosin-mediated processes, which alter the shape of the seam cells (Shelton *et al.*, 1999). Interruption of many of these steps results in mutant embryonic, and larval and adult body morphology or lethality. Thus, actin, myosin and other actin/myosin binding proteins are all required for, or play a part in, regulating body morphology (Simske and Hardin, 2001). In addition, as already demonstrated, removal of some ECM components (collagens or cuticulins) can also lead to morphological defects.

The *C. elegans* genome encodes for homologues of many of these morphology-determining components, and it is therefore of great interest to observe whether any have similar enzymatic roles in the cuticle. A forward genetics approach has been undertaken using TP12 in an RNAi screen to target these morphology/moulting/morphogenesis-dictating enzymes or their homologues in order to investigate the effects of their presumed down-regulation on the COL-19::GFP pattern.

A genome-wide RNAi screen has previously been completed by Kamath *et al.* (2003), and some of the genes that were identified as producing Bli, Dpy, Sma or Sqt phenotypes, were taken as targets in our screen for a more detailed analysis using the COL-19::GFP marker. Similar to that described for morphological mutants, the ability to visualise the COL-19::GFP pattern in the RNAi-treated animals provides information, specifically at the cuticle level, and allows the observation of *how* the cuticle is affected. In other words, it can illustrate, for example, which type of Dpy (class II or class III) a morphological mutant is. The use of TP12 in an RNAi screen effectively provides a more sensitive screen than the Kamath *et al.* (2003) screen to identify genes with roles in the ECM, since disruption in the COL-19::GFP pattern can be observed even in the absence of a gross morphological phenotype. Thus, using TP12, genes that were dismissed in the Kamath *et al.* (2003) screen as not having a visible RNAi effect, have now been identified as actually producing cuticle defects when targeted by RNAi.

#### 4.3.1 Selection of gene targets

Genes chosen to be targets of RNAi feeding were identified by a number of methods. The genes are listed in table 4.2.

**Table 4.2: Genes targeted by RNAi feeding and the resulting phenotypes.** Genes are organised by the severity of COI.-19::GFP disruption and are coloured accordingly. Red = severe disruption, orange= medium disruption, yellow= slight/mild disruption, white = wild type or extremely mild disruption. Bli= blister, Gro= growth defects, Unc= uncoordinated, Eml= embryonic lethal, Lvl = laval lethal, Dpy= dumpy, S= severely expressed phenotype, W= weak phenotype. ADAM refers to the subgroup of proteases as does ZC (zinc carboxypeptidase). The expression mountains as defined by Kim *et al.*, 2003, in which each gene resides is also noted (roman numerals in brackets).

Animal Haem Peroxidases		
Gene/identity	Gross morphological	COL-19::GFP pattern
f56c11.1 (DUOX) (7/16)	Bli, Gro, Unc (S)	Multiple alae, branched alae
f53g12.3 (DUOX) (3)	Bli, Gro, Unc (S)	Multiple alae, branched alae
zk430.8 (16)	Bli, Gro, Unc (S)	Amorphous
f09f3.5 (9)	WT	Amorphous
c16c8.2 (3/14)	Dpy	Amorphous
t06d8.10 (14)	Dpy (W)	WT

Proteases		
Gene/identity	Gross morphological	COL-19::GFP pattern
<i>nas-38</i> (f57c12.1) (14)	Dpy (Rol)	Branching/ disrupted alae
<i>nas-28</i> (f42a10.8) (17)	Unc	Branching alae, amorphous, wispy
f08c6.1 (ADAM) (16)	Dpy	Amorphous
<i>sup-17</i> (dy3.7) (ADAM) (11)	WT	Crosshatched and amorphous
t21d11.1 (ADAM) (14)	WT	Amorphous
<i>nas-5</i> (t23h4.3) (0)	WT	Amorphous, localises to vacuoles
<i>nas-35</i> (r151.5/toh-2) (1)	WT	Amorphous
<i>nas-37</i> (c17g1.6) (38)	Mit (S)	Amorphous, multiple alae, branching annulae
<i>nas-36</i> (c26c6.3) (29)	WT	Amorphous
<i>nas-39</i> (f38e9.2) (3)	WT	Amorphous
w01a8.6 (ZC) (16)	WT	Wispy
r11a5.7 (ZC) (1)	Dpy	Wispy
c42d8.5 ( <i>acn-1</i> ) (14)	Dpy (W)	Amorphous
ze434.9 (ZC) (16)	Dpy (W)	Amorphous
w09d12.1 (ADAM) (8)	Dpy (W)	WT
<i>nas-14</i> (f09e8.6) (Z) (14)	WT	WT
<i>nas-15</i> (t04g9.2) (14)	WT	WT

Thioredoxins		
Gene/identity	Gross morphological	COL-19::GFP pattern
y44e3a.3 (2)	WT	Amorphous
f56g4.5 (2)	WT	Amorphous and nuclear staining
c05d11.5 (18)	WT	Amorphous
t10h10.2 (16)	dpy (W)	Amorphous
f29b9.5 (18)	dpy (W)	Amorphous
c30h7.2 (2)	WT	Amorphous
t06c6.11 (8)	WT	Amorphous
t05f7.11 (8)	WT	WT

Misc.		
Gene/identity	Gross morphological	COL-19::GFP pattern
r160.1 ( <i>dpy-23</i> ) (adaptin) (1)	Sma	Wispy and amorphous (similar to <i>dpy-7</i> )
c34g6.6 (hypothetical <i>cut-1</i> homologue) (14)	Lvl	Wispy, fragmented
c18e9.2 (ER translocation) (20)	Lon	Wispy
f169.2 ( <i>dpy-6</i> ) (actin interacting) (14)	Dpy (S)	Amorphous (similar to <i>dpy-7</i> )
f52b10.1 ( <i>my-1</i> ) myosin (1)	Dpy (W)	Branching annulae (similar to <i>dpy-5</i> )
r07g3.3 (myosin heavy chain) (11)	Dpy (W)	Amorphous
c33h5.9 (myosin) (2)	Dpy	Amorphous
b0511.1 (PPI) (6)	WT	Amorphous
zk484.7 (PPI) (4)	Sma (W)	Amorphous
c02b10.5 (p transport from ER-golgi) (5)	Eml (W)	Amorphous
c02c2.1 (p transport from ER-golgi) (8)	WT	Amorphous
zk783.1 (serine protease/laminin)	Sma (W)	WT
f18c12.1 (DNA-J like domain) (13)	WT	WT
c37c3.3 (SNF-7 domain) (protein sorting?) (1)	Unc	WT
f42a8.1 (ds target of DBL pathway) (17)	WT	WT

Collagens		
Gene/identity	Gross morphological	COL-19::GFP pattern
<i>dpy-9</i> (t21d12.2) (16)	Dpy	Amorphous
<i>dpy-7</i> (f46c8.6) (14)	Dpy	Amorphous
<i>dpy-17</i> (f54d8.1) (12)	Dpy (W)	Branching annulae & alae (sim to <i>sgt-1 N-term mutants</i> )
<i>dpy-10</i> (t14b4.7)	Dpy	Branching annulae
<i>dpy-4</i> (y41e3.2) (16)	Dpy	Branching annulae
<i>dpy-5</i> (f27c1.8) (16)	Dpy	Branching annulae
<i>rol-6</i> (t01b7.7) (16)	Dpy (W)	Amorphous
<i>col-35</i> (c15a11.1) (17)	Dpy (W)	Amorphous
<i>col-38</i> (f54c9.4) (35)	Dpy (W)	WT
<i>sgt-1</i> (b04g1.2) (16)	WT	WT
<i>sgt-3</i> (f23h12.4) (14)	WT	WT

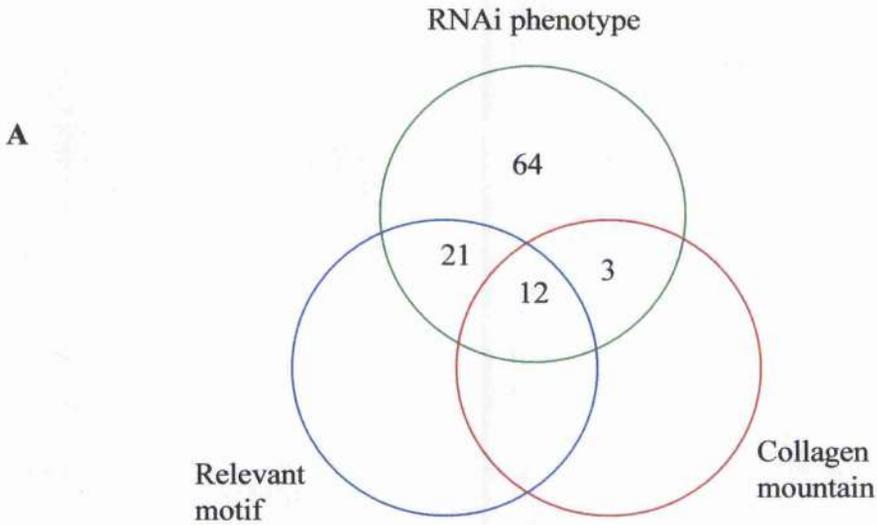
  

Tyrosinase		
Gene/identity	Gross morphological	COL-19::GFP pattern
f21c3.2 (6)	WT	Wispy
k08c3.1 (16)	WT	WT
y44e3b.2 (6)	WT	WT

The first method was a motif-based search using the Wormbase database in order to identify homologues of the proteins described in section 4.3 on the basis of characteristic domains. Thus, motif searches were carried out in order to identify uncharacterised proteins encoding collagen-, cuticulin-, transglutaminase-, thioredoxin-, peroxidase-, tyrosinase-, PPI-, PDI-, astacin-, ADAM/ADAMTS-, M14-containing-, actin-, and myosin- domains. For example, a variety of zinc-dependent proteases were chosen on the basis of the presence of astacin, CUB and thrombospondin-1 like domains.

From this list of potential genes, prioritisation was given to genes that give a Dpy, Sqt, Bli or Sma RNAi phenotype according to the genome-wide RNAi screen. This was particularly true for the collagens. The majority of the collagens do not have RNAi phenotypes due to redundancy while the ones that do, have almost all been previously characterised. Only four collagens remained that had not been previously allocated to a locus and that produce a phenotype in the genome-wide screen. Also, only a subset of the myosins, kinesins, lectins, laminins, dyneins and actins identified by the motif searches, were chosen for further RNAi analysis, being selected on the basis of their RNAi phenotype.

There was much overlap between the genes identified through motif searches and genes exhibiting RNAi phenotypes (figure 4.11). However, a number of genes identified in the genome-wide RNAi screen encode products not previously implicated in having a role in the cuticle and these were also chosen as targets. Therefore, the second means for identifying RNAi targets for the screen described in this thesis was on the basis of RNAi phenotype alone. These include genes encoding for: adaptins (*dpy-23*); proteins with protein-protein interacting domains (*c34g6.6*); proteins with putative roles in translocation (*c18e9.2*), transport between the endoplasmic reticulum (ER) and the Golgi apparatus (*c02b10.5* and *c02c2.1*); and proteins involved in protein folding (*f18c12.1*). Once identified according to their morphological phenotype, the role of these genes can be hypothesised in terms of their role in cuticle biogenesis (for example, *c02b10.5* and *c02c2.1* may have roles in transporting cuticular components). However, these could not have been implicated in cuticle synthesis on the basis of their roles alone. This exemplifies how invaluable the data obtained from the RNAi screen was for our selection of RNAi targets.



**B**

	No RNAi phenotype, not expressed in collagen mountain	Have RNAi Phenotype	Expressed in collagen mountain	Both RNAi and expressed in collagen mountain
Collagens	36	2	55	7
Animal haem peroxidases	68	8	5	8
Tyrosinases	78	11	11	0
Proteases	75	6	14	5
Thioredoxins	84	9	5	2

**Figure 4.11: Data utilised to select genes for RNAi targeting.** A) The genome-wide RNAi screen carried out by Kamath *et al.* (2003), identified 138 genes which resulted in body morphology defects (Dpy, Lon, Bli, Rol) and moult defects, when targeted by RNAi. 29 the 138 genes (potentially) encode tyrosinase, thioredoxin, collagen, tyrosinase, and peroxidase motifs, termed 'relevant motifs' in figure A. The allocated expression mountains of these genes were also analysed. 20 of the 138 genes fall into the expression mountains enriched by collagens. The Venn diagram shows the percentages of genes with RNAi phenotypes, relevant motifs and expression in collagen mountains. This shows that over 50% of the genes with RNAi phenotypes are probably not associated with cuticle synthesis. B) Table to observe correlations between encoded motifs, RNAi phenotype and expression mountain. Motif searches were carried out and the identified genes analysed for expression mountain and RNAi phenotype.

Subsequent to the publication of the Kamath *et al.* (2003) screen, a similar screen was undertaken using the RNAi-sensitive strain, *rrf-3* (Simmer *et al.*, 2003). Despite this identifying many genes within the genome that had not previously been associated with RNAi phenotypes, no further genes were highlighted as targets for our screen that had not already been chosen.

As a third way of selecting targets, microarray data from Kim *et al.* (2001) was analysed. The data from this paper was obtained via high-throughput microarray assay on the whole genome and resulted in 43 expression mountains being defined, with each mountain containing genes with similar expression profiles. Five of the mountains have high concentrations of collagen genes. Since it is a long established fact that collagens are expressed in a highly regulated and cyclical manner, the presence of distinct “collagen” mountains suggested the likelihood that genes with similar expression profiles might be involved in their modification. Each of the collagen mountains was searched to identify genes with RNAi phenotypes or with interesting motifs.

Once identified as potential targets for RNAi, feeding constructs were either obtained from a library of RNAi feeding clones provided by the Ahringer lab (Kamath *et al.*, 2003), or were cloned via PCR amplification from either genomic or coding DNAs (refer to methods for details).

### 4.3.2 General trends from RNAi data

Genes tested in this screen give both wild type and mutant RNAi phenotypes. On the gross morphological level, wild type, Dpy and Bli phenotypes were observed, while in terms of the COL-19::GFP pattern, a similar variety as that observed in the crosses result were produced (table 4.2).

The majority of the RNAi phenotypes for the genes tested have been termed amorphous; however, some exhibit this phenotype to a weak degree while others express it strongly. In general, a weak amorphous phenotype is characterised when the dominant pattern of COL-19::GFP is the annular staining but which is blurred. A strong phenotype is one resembling type II staining where no annular pattern is apparent and the entire cuticle is amorphous. In the weaker form of the phenotype, it is likely that only a subset of cuticle collagens are affected (i.e. the DPY-7 set of collagens are not affected due to the fact that annulae are discernible).

### 4.3.3 Correlation between gross body morphology and COL-19::GFP pattern

An interesting phenomenon is the presence of RNAi-induced COL-19::GFP disruption in the absence of a gross morphological phenotype, as exemplified by many of the protease genes (e.g. *dy3.7* and *t21d11.1*). In such cases, wild type body morphology is associated with a COL-19::GFP phenotype that is amorphous and similar to class II mutants. From the amorphous phenotype, it could be assumed that RNAi-treatment has ablated the annular furrows. Since annulae are predominantly associated with mediating contractions laterally rather than circumferentially, ablation of these structures in these mutants does not necessarily need to correspond to circumferential expansion or a Dpy phenotype. On the other hand, no examples of wild type worms exhibiting type III COL-19::GFP disruption have been observed. This can be explained by the fact that in class III mutants, such as *dpy-5*, the cuticle disruption is confined to the area above the seam cells and the consequential lack of contraction in this region is the main reason for the increased body width. In other words, this seam cell cuticle disruption and the Dpy phenotype go hand in hand; if RNAi is sufficient to disrupt this area in order to prevent contraction, a Dpy phenotype would be expected to naturally follow.

As well as wild type worms exhibiting type II COL-19::GFP disruption, the converse is observed, since some Dpy RNAi-treated animals show a wild type COL-19::GFP pattern (e.g. *w01a8.6* (a zinc metalloprotease) and *col-38* (a collagen)); this is only exhibited by animals that are weak Dpy. It is unlikely that the slight dumpiness results from the weakening of the cuticle overlying the seam cells since, as explained above, such aberrations that lead to uncontracted cuticle can only really result in a Dpy phenotype. In these Dpy animals, DPY-7 and its associated collagens must also be assembling properly because of the presence of wild type annular furrows. However, perhaps there is overall (relatively mild) weakening of the cuticle to allow it to stretch in response to the internal hydrostatic pressure.

Apart from these few Dpy RNAi-treated animals that show wild type COL-19::GFP patterning, most Dpy RNAi-treated animals exhibit either class II or class III disruption. All RNAi-derived Bli phenotypes observed in this screen are associated with a COL-19::GFP phenotype showing either an amorphous phenotype, or a novel COL-19::GFP phenotype of extreme branching alae and annulae. Other genes exhibit moulting defects, uncoordinated movements (Unc), Sma phenotypes or levels of lethality when targeted by RNAi (table 4.2). However, there is no direct correlation between these gross phenotypes and the type and severity of COL-19::GFP disruption.

#### 4.3.4 Trends in the data

Although there is no connection between gross morphological phenotype and type and severity of the COL-19::GFP pattern, there are correlations between gene identity and the disruption. Most notable is that many of the protease and peroxidase targets do result in COL-19::GFP disruption (this will be discussed in sections 4.3.6). Conversely, most of the tyrosinases and thioredoxins do not cause severe effects on the cuticle when targeted by RNAi. Within the category containing the cuticle collagens, a range of COL-19::GFP phenotypes are observed. This will be discussed in section 4.3.7.

The screen has provided a useful way of identifying general classes of genes, such as the proteases and the peroxidases which produce more severe effects on the cuticle when targeted by RNAi than other classes. It has also highlighted a number of individual genes that appear to have significant roles in cuticle production. Further characterisation of these genes and their role in the cuticle will be extremely interesting.

#### 4.3.5 Tyrosinases and thioredoxins do not have severe RNAi effects

A number of reports have implicated tyrosinases and thioredoxins in cross-linking. While the thioredoxin mutant, *dpy-11*, results in a Dpy phenotype and male tail defects when mutated, of the thioredoxins, only t10h10.2 and f29b9.5 exhibit morphological and COL-19::GFP disruption. However, these are only weakly Dpy and amorphous. y44e3a.3, f56g4.5, c05d11.5, t06g6.11 and c30h7.2 exhibit weak (or very weak) amorphous COL-19::GFP patterns in the absence of a gross morphological phenotype (data are not shown). The only COL-19::GFP disruption to be exhibited upon RNAi of a tyrosinase is a very weak wispy appearance in f21c3.2.

It appears that either 1) these enzymes have redundant roles, 2) cross-linking by these enzymes is not vital for cuticle synthesis, 3) RNAi is inefficient at targeting these genes, and/or 4) despite encoding for a domain that is homologous to a thioredoxin domain, they exhibit little or no activity. The latter point has been partially addressed, for at least one of the tested tyrosinases, Gerrits and Blaxter, 1995 personal communication, investigated two tyrosinases (*cup-4* and k08e3.1) and found only low activities. This may be the reason for the fact that RNAi of the latter tyrosinase is not effective.

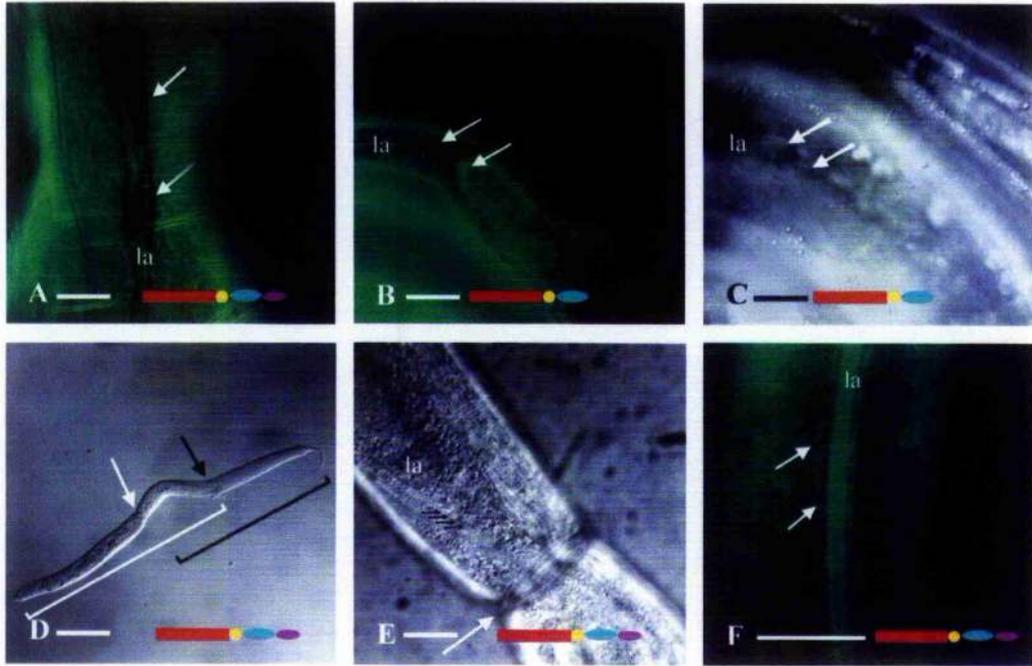
### 4.3.6 RNAi of metalloproteases

#### 4.3.6.1 RNAi of astacins

17 metalloproteases were examined: nine astacins, four ADAM/ADAMTS-like proteins, three zinc carboxypeptidases and an additional angiotensin converting enzyme (with an M12 protease active site) that is required for adult morphogenesis. As described previously, 40 astacins are encoded by the *C. elegans* genome and therefore the genes tested by RNAi in this screen represents less than a quarter of all astacins. However, in the sample group there is a representative from all but one of the astacin subgroups defined by Mohrlen *et al.* (2003) (according to domain structure). All the encoded astacins contain presequence, prosequence and astacin protease domains while the presence of SXC, TSP-1 and EGF-like domains are subset-specific (Mohrlen *et al.*, 2003). Subgroup V, for instance, in which the characterised DPY-31/NAS-35, NAS-36 and NAS-37 fall, encode CUB, EGF-like and TSP domains. At the time of the RNAi screen, neither DPY-31/NAS-35, nor NAS-37 had been characterised and therefore these are included in the results. The other astacins targeted by RNAi were: *nas-5* (t23h4.3), *nas-14* (fo9e8.6), *nas-15* (t04g9.2), *nas-28* (f42a10.8), *nas-38* (f57c12.1) and *nas-39* (f38e9.2). Despite being previously characterised, *hch-1/nas-36* was included in order to observe cuticular RNAi-mediated defects.

The majority of the metalloproteases examined for an RNAi phenotype show some sort of COL-19::GFP disruption upon their targeting. However, many display only a subtle amorphous appearance of the annule (refer to table 4.2). While there is no clear trend as to whether ADAM-like or astacins give more severe phenotypes, it must be noted that the two most severe RNAi phenotypes result from targeting astacins, *nas-38* and *nas-28*. The disruption observed is shown in figures 4.12 A – C. Targeted RNAi of both genes result in branched alae and an amorphous pattern of COL-19::GFP (refer to figures). These astacins fall into separate subgroups, however, and this shows that the presence of certain domains does not predispose an astacin to having a severe RNAi-induced cuticle disruption. Other proteases tested exhibited differing levels of amorphous COL-19::GFP (refer to table 4.2).

Concurrent with the analysis of *nas-37* by Davis *et al.* (2004), data from my RNAi screen identified the moult phenotype upon RNAi of this astacin. Davis *et al.* (2004) later confirmed it as having a role, not only in the process of moulting (ecdysis specifically), but also in synthesis of the new cuticle (Davis *et al.*, 2004). The phenotypes observed from the TP12 screen are



**Figure 4.12: Phenotypes resulting from RNAi of astacins.** A) COL-19::GFP from animals RNAi targeted for *nas-38* (f57c12.1) (astacin). Branching alae are indicated with white arrow. B) COL-19::GFP from RNAi of *nas-28* (f42A10.8) (astacin). Branching alae are indicated with white arrows. C) Corresponding Nomarski image of figure B. Branching alae are indicated with white arrows. D) Nomarski image of animals treated with RNAi of *nas-37* (c17g1.6) (astacin). White arrow indicates the site of constriction at the mid body of the worm where the old cuticle is attached. Black arrow indicates the tail end of the worm. The unshed old cuticle is dragged behind the worm (translucent appearance) (length of worm is indicated by a white bracket and old cuticle indicated with black bracket). E) Nomarski image focusing on the constriction around the mid body region (white arrow) in animals fed with *nas-37* RNAi. Alae are also in the plane of focus and are multiple (4 instead of 3) (black arrows). F) COL-19::GFP pattern after targeted RNAi of *nas-37*. Branching annulae are indicated with white arrows. In all figures, where present, lateral alae are indicated with 'la'. Diagrammatic domain structures of each of the encoded astacins are shown (modified from Möhrlen *et al*). Red rectangle = astacin domain, yellow circle = EGF-like domain, blue oval = CUB domain, purple oval = TSP domain. Scale bars = 10 $\mu$ m in A, B, C, E, and F. 100 $\mu$ m in D.

illustrated in figures 4.12 D and E. Figure 4.12 D depicts an adult hermaphrodite that has failed to shed the cuticle from the previous moult. The old cuticle appears to be attached at the mid body region of the worm where a constriction has formed and the rest of the old cuticle is dragged behind the worm. Figure 4.12 E shows the constriction at the mid body of an adult hermaphrodite at a higher magnification and also shows another characteristic phenotype of targeted RNAi of this gene, multiple alae. RNAi-treated worms exhibit disruption at the COL-19::GFP level, as depicted in figure 4.12 F. The dominant COL-19::GFP pattern is slightly amorphous annulae (These are not in the plane of focus in figure 4.12 F). In addition, the annulae around the alae are branched (refer to figure 4.12 F). Alae also appear slightly amorphous. Comparison of COL-19::GFP patterns in RNAi-treated animals (from the current study) and the published images of the null mutant, *nas-37(tm410)* (Davis *et al.*, 2004) suggests that the null phenotype is more severe at the COL-19::GFP level. In such deletion mutants, the COL-19::GFP appears more amorphous and instead of displaying only multiple alae, branching alae are also observed. These data show that the cuticle structure of RNAi-treated worms is aberrant and seems to strengthen the hypothesis that the enzyme has supporting roles in the steps outwith ecdysis (e.g. cuticle synthesis) i.e. if it were only involved with ecdysis, disruption would be expected to be restricted to only the cuticle that is being shed.

Again, concurrent with analysis by Novelli *et al.* (2004) of *dpy-31/nas-35*, which is now known to encode a homologue of BMP-1 and to catalyse the C-terminal processing of cuticle collagens, *nas-35* was included in the TP12 RNAi screen. As described in table 4.2, targeted RNAi of this gene results in neither a mutant gross morphological phenotype nor an aberrant COL-19::GFP expression. This is consistent with the findings of Novelli *et al.*, 2004. It is important to note, however, that Novelli *et al.* (2004) observed embryonic lethality in null mutants for this gene. This indicates that RNAi by feeding for this gene is not fully penetrant. It has been suggested that RNAi may similarly be an ineffective method for many of the astacins (Novelli *et al.*, 2004). In light of this, care must be taken when interpreting the results of the other astacins. As described in table 4.2, the other astacins targeted by RNAi, *nas-5*, *nas-36* and *nas-39*, result in wild type phenotypes and mild COL-19::GFP disruption. However, because of the hypothesised reduced effectiveness of RNAi on these genes, it is possible that these mild phenotypes are not true reflections of null alleles. It is also worth noting that despite the severe COL-19::GFP phenotypes of *nas-38* and *nas-28*, the true null phenotypes could be even more severe.

Targeted RNAi of two of the four ADAMs analysed produces moderate cuticle aberrations in the shape of amorphous COL-19::GFP expression. This suggests a role for these proteases in the cuticle. SUP-17 (*dy3.7*) has previously been characterised to be involved in the lin-12/NOTCH pathway that decides cell fates, including those of the developing vulva and gonad (Wen *et al.*, 1997). LIN-12/NOTCH is believed to be the substrate of SUP-17 during these processes (Wen *et al.*, 1997). Based on the fact that RNAi-treated animals have amorphous cuticles, this ADAM may have multiple roles including processing of the cuticle.

#### 4.3.7 RNAi of cuticle collagens

Only 19 of the 160 *C. elegans* cuticle collagens have been attributed to a mutant locus. However, four as yet uncharacterised collagen genes (*col-38*, *col-35*, *col-100* and *col-134*) were chosen for RNAi analysis using TP12 as they produced mutant RNAi phenotypes (according to the screen by Kamath *et al.*, 2003). In addition, known collagens *sqt-1*, *rol-6*, *dpy-5*, *sqt-1*, *dpy-10* and *dpy-7* were targeted by RNAi (the results of these have been described in previous sections). Utilisation of TP12 to visualise the cuticle structure of these mutants has confirmed the presence of at least three distinct groups of collagens with characteristic temporal, spatial, structural and interaction patterns (COL-19::GFP disruption groups II, III and IV), and RNAi of cuticle collagens has facilitated their designation into one of these groups on the basis of their COL-19::GFP disruption pattern. This is exemplified by three collagens: COL-100 (*t21d12*), COL-134 (*y41e3.2*) and DPY-17. *col-100* RNAi-treated worms exhibit an amorphous pattern (type II disruption), *col-134* RNAi animals display a type III disruption and targeting *dpy-17* results in animals showing type IV disruption (figure 4.13 shows *y41e3.2* and *t21d12.2*). Consequently, these can be predicted to interact with and share temporal patterns with DPY-7, DPY-5 and SQT-1 respectively. Other collagens tested include *col-35* (*c15a11.1*) and *col-38* (*f54c9.4*), which produce little or no COL-19::GFP disruption upon RNA interference (not shown). This is not unexpected for *C. elegans* cuticle collagens because of the great redundancy in the 160 collagens that are encoded in the genome.

RNAi can be a problematic method for the analysis of cuticle collagens as there is a huge amount of similarity at the sequence level (due to the Gly-X-Y repeat region). Consequently, it is highly possible that the RNAi targets many collagens non-specifically. In the four uncharacterised collagens described in this screen, this does not seem to be the case, as we have observed at least

three different phenotypes. Further proof that the RNAi results observed are genuine is detailed in the following section.

#### 4.3.7.1 Identification of two novel collagen genes

As already stated, some of the genetic loci responsible for the morphological mutants are unidentified. It was therefore of great interest to observe that RNAi treatment of *col-100* (t21d12.2) and *col-134* (y41e3.2) produced Dpy phenotypes with type II (figure 4.13 D) and type III COL-19::GFP (figure 4.13 A) disruption respectively. Since the genes mapped close to the predicted genomic location of the (as yet uncharacterised) morphological mutant loci *dpy-4(e1166)* (type II disruption) and *dpy-9(e12)* (type III disruption) respectively, the observations from the RNAi feeding presented here made them ideal candidates for these uncharacterised Dpy mutants. Indeed, the COL-19::GFP disruption patterns of the RNAi-treated animals and the morphological mutants are similar (figure 4.13). Concurrent with these predictions being made, Simmer *et al.* (2003) published a paper describing their RNAi screen and confirmed our predictions. The TP12 strain has thus been a useful tool in accurately predicting the genetic loci of two collagen genes.

#### 4.3.8 RNAi of peroxidases

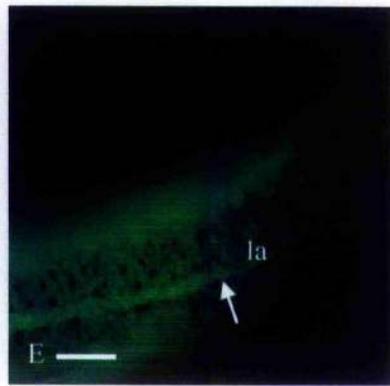
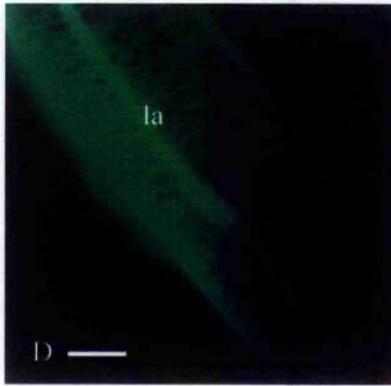
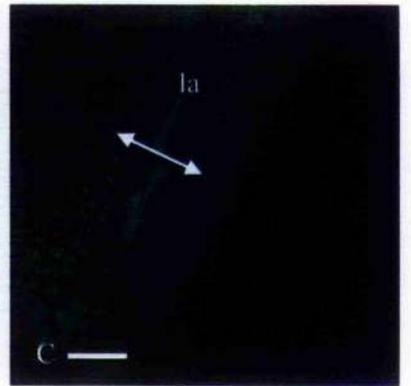
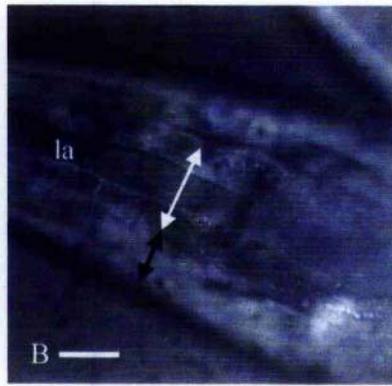
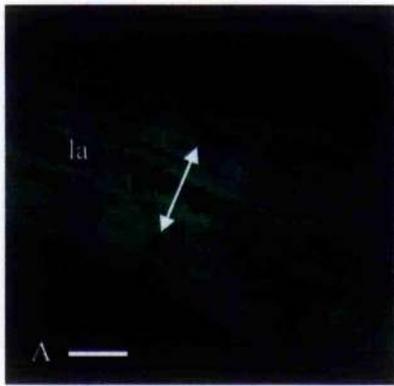
The genes encoding for the two Duox enzymes identified by Edens *et al.*, 2001, (BLI-3 and f53g12.3), as well as four other animal haem peroxidase-containing proteins (zk430.8, f09f3.5, c16c8.2 and t06d8.10) were targeted by RNAi (table 4.2). 4 of the 6 peroxidase genes tested resulted in severe COL-19::GFP disruption and thus made the further study of the peroxidase genes very interesting. ZK430.8 was of special interest as it displays an RNAi phenotype similar, albeit weaker, to that of the Duoxs. On this basis, *zk430.8* was chosen from all the genes tested, to be the target of further detailed analysis: This will be discussed in chapter 5.

#### 4.3.9 RNAi of other genes

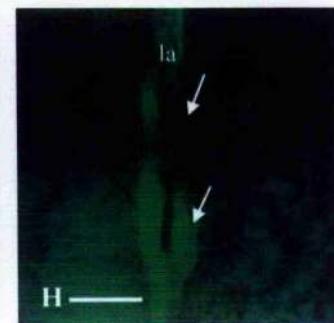
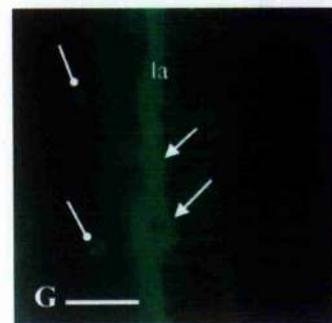
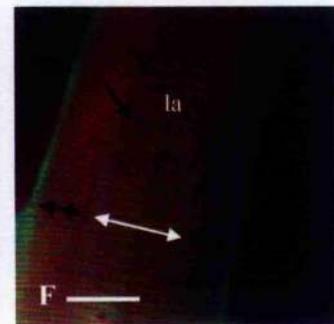
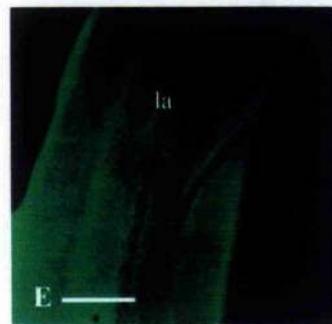
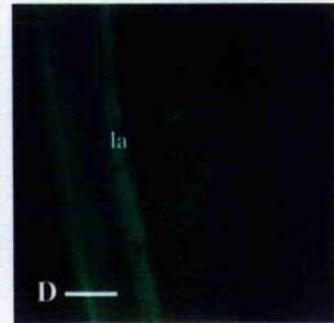
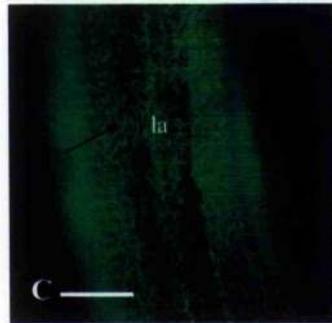
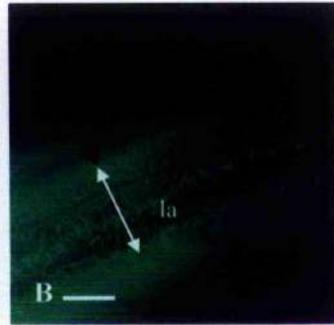
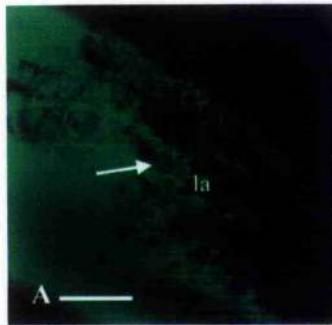
A number of other genes were targeted according to their mutant RNAi phenotypes. The resulting COL-19::GFP patterns of the most interesting RNAi-treated animals are depicted in figure 4.14 and are described below. The less severe phenotypes are described in table 4.2.

c34g6.6 was not included in the RNAi screen by Kamath *et al.* (2003), but the screen by Simmer *et al.* (2003) subsequently showed that its RNAi phenotype is body morphology defects and lethality. A Blast search against this gene in the *C. elegans* genome highlights this as encoding a potential cuticulin. In this study, a predominantly larval lethal (Lvl) phenotype has been

**Figure 4.13: Comparison of COL-19::GFP phenotypes in RNAi-treated animals and uncharacterised morphological mutants.** A) COL-19::GFP phenotype as a result of feeding with *y41e3.2*. Two distinct patterns are observed: disruption in an uncontracted region overlying the seam cells (white double headed arrow) and wild type in the region overlying the dorsoventral hypodermis (black double headed arrow). B) The same regions (denoted by white and black double headed arrows) are observed when viewed via Nomarski imaging. C) COL-19::GFP phenotype of the morphological mutant (crossed with TP12), *dpy-4(e1166)*. The pattern is similar to the RNAi (white and black double headed arrows) and thus assigns *y41e3.2* as the *dpy-4* locus. D) COL-19::GFP phenotype as a result of feeding with *t21d12.2*. An amorphous pattern is observed. E) A similar amorphous COL-19::GFP pattern is observed in the cuticle of the morphological mutant, *dpy-9(e12)*. (The amorphous alae are denoted with an arrow.) The similar amorphous pattern implicates *t21d12.2* as the mutant locus in *dpy-9* animals. Lateral alae are labelled 'la'. Scale bars =  $\mu\text{m}$ .



**Figure 4.14: Cuticle aberrations resulting from RNAi of miscellaneous genes.** A-G) Patterns of COL-19::GFP in A) *r160.1* (adaptin) RNAi B) *c34g6.6* (hypothetical cuticulin homologue) RNAi; C) *c18e9.2* (ER protein translocation) RNAi; D) *f52b10.1* (myosin) RNAi; E) *c33h5.9* (myosin) RNAi. F) Merged image of COL-19::GFP and DPY-7 expression patterns (after immunolocalisation with DPY-7 antibodies) in *c33h5.9* RNAi animals. G) *dpy-6* (actin-interacting) (RNAi). The pattern observed from RNAi treatment of *dpy-6* is similar to that of the morphological mutant, *dpy-6(e14)* which is shown in H. In all figures, where present, branched or prematurely terminated annulae are denoted with a single headed black arrow, branched alae are denoted with a single-headed white arrow, seam cell-derived uncontracted cuticle is marked with a white double headed arrow, dorsoventral-derived cuticle is shown with a black double-headed arrow. In figure G, the spots of intense COL-19::GFP amorphous material are indicated with blunt arrows. Lateral alae are indicated ('la') but in figures A, B, C, E and F, are out of the plane of focus. Scale bars = 10µm.



observed. Escapers exhibit a severely fragmented cuticular pattern type III COL-19::GFP (figure 4.14 B). These observations confirm this putative cuticulin as being an essential (structural) component of the cuticle (or basement membrane).

It is interesting to note the high predominance of myosin-related genes which produce severe forms of COL-19::GFP disruption. *f52b10.1/nmy-1* (myosin) (figure 4.14 D), *r07g3.3* (myosin heavy chain) and *c33h5.9* (myosin) exhibit cuticles that have amorphous COL-19::GFP patterns (figures 4.14 E). *c33h5.9* RNAi-treated animals are novel in that in the dorsoventral region of the cuticle, they exhibit an amorphous pattern at the COL-19::GFP level but show a pattern of DPY-7 in which annulae appear properly formed (figures 4.14 E and F). This is in contrast to most other amorphous mutants previously tested, which all display reduced or severely fragmented patterns of DPY-7. It had been assumed that the amorphous phenotype is a result of aberrant formation of annular furrows and the consequential loss of annulae. In this case, it appears that COL-19::GFP is being expressed in an amorphous pattern, despite the presence of properly laid down structural scaffolds. RNAi of *dpy-6*, which encodes for an actin interacting protein, produces a COL-19::GFP phenotype similar to that of the *dpy-6(e14)* allele (compare figures 4.14 G and H). In both cases, the cuticle is amorphous and additionally, in *dpy-6(e14)*, accumulation of COL-19::GFP can be seen in vesicles. The aberrant cuticles exhibited by animals targeted for putative proteins involved in the actin and myosin contractile system herein have been shown to produce aberrant cuticles, and provide further evidence of the influential roles played by these proteins in the establishment of the embryonic vermiform and the cuticle furrows. It can also be extrapolated from this data that the yet unidentified lesion in *dpy-6(e14)* may be an effective null, due the similarity in COL-19::GFP expression after RNAi treatment.

Other interesting RNAi phenotypes include the severely wispy and amorphous COL-19::GFP pattern of the adaptin (*r160.1/dpy-23*) (figure 4.14 A) and the wispy pattern of a potential component of an ER protein translocation complex (*c18e9.2*) (figure 4.14 C). Adaptins are components of the intracellular transport vesicles and certain *C. elegans* homologues have previously been identified that result in embryonic lethality and, more importantly for this study, Dpy phenotypes when targeted by RNAi (Boehm and Bonifacino, 2001). Their function in transport and their mutability to form aberrant cuticles suggests that these could have a role in secretion of cuticle components. The similarity of the cuticle of *r160.1/dpy-23*(RNAi)-animals and *dpy-7* mutants suggests that a substantial proportion of the cuticle components are lost upon targeting of this protein. C18E9.2 may have a similar role in the secretion of cuticle components

into the cuticle. However, its RNAi-derived COL-19::GFP expression pattern is distinct from that of r160.1/*dpy-23*(RNAi), which suggests different set of collagens/cuticle components are transported by the two systems.

#### 4.3.10 Searching for trends within similarly expressed genes

A paper published by Kim *et al.* (2001) used microarray data in order to group co-regulated genes together in “mountains”, five of which are highly enriched in collagen genes. It was of interest to see whether genes with similar expression profiles as collagens would have a more prominent role in the collagen-containing cuticle than those genes with different expression patterns. Three of the mountains were searched for genes with potentially interesting roles (e.g. peroxidases, proteases and tyrosinases) which could be used as targets for RNAi feeding. In table 4.2 the mountains in which each of the RNAi targets appears is noted. However, there is no obvious correlation between severity of RNAi phenotype and mountain.

#### 4.4 Discussion

The aim of this chapter was to use the COL-19::GFP marker strain, TP12, in order to analyse the effect that a number of lesions have on the *C. elegans* cuticle. Crosses between TP12 and morphological mutants have shown that aberrations to the cuticle are expressed in terms of differing degrees of disruption to annulae and alae. Five distinct patterns of COL-19::GFP disruption were thereby defined: wild type/mild disruption (class I), amorphous annulae (class II), broken seam cell annulae (class III), severe global disruption (class IV) and branching alae and annulae (class V). It has been established that an amorphous COL-19::GFP pattern is a result of the loss of definition of the annular furrows and consequently the annulae. The majority of the morphological mutants tested have lesions within genes encoding for the few mutable cuticle collagens. We report here that the severity of the resulting COL-19::GFP aberrations, and, in some cases, the nature of the gross morphological phenotype, are dictated by the identity of the lesions. In addition, the groups of collagens defined here correlate with the groups defined by McMahon *et al.* (2003) on the basis of temporal- and spatial- expressions and obligate partners. This confirms that the type of COL-19::GFP defined cuticular disruption that a mutant collagen generates, is dictated by its expression patterns.

In addition, TP12 has been used with an RNAi approach to target genes with potential roles in cuticle synthesis. With data from the crossed strains and the from the RNAi screen, a number of combinations of gross morphological and COL-19::GFP phenotypes have been observed: *Dpy*

animals have been observed with annulae defects, amorphous cuticles or with wild type COL-19::GFP expression patterns, while some BliS have been shown to display either aberrant- or wild type- COL-19::GFP expression. SqtS have very severe annular disruptions while RolS, LonS and wild types can be amorphous or wild type at the COL-19::GFP level. This demonstrates that animals exhibiting similar gross morphological phenotypes can have very different cuticle structures. Thus TP12 provides a means for defining what type of Dpy, Bli or Rol a mutant is. Consequently, this enables the dissection of the roles of the cuticular component (collagenous and enzymatic) in ECM synthesis.

The complexity of the *C. elegans* cuticle, being composed of over 160 collagens in multiple layers, means that it is easily mutable to a variety of morphological phenotypes when its components are altered. Lesions in collagen loci can produce Bli, Dpy, Sqt, Rol and Lon phenotypes. This demonstrates that each collagen, and the structure within which it resides, has a distinct role. However, only 19 of the 160 collagens have been found to be mutable to mutant body morphology phenotypes, indicating that some collagens play more of an essential role than others.

#### 4.4.1 Examination of data from Rol mutants

Lesions in the N-terminal protease subtilisin-like cleavage site of the cuticle collagen, ROL-6, results in a Rol phenotype but a relatively wild type, albeit helically twisted, pattern of annulae, as seen with the *rol-6(su1006)* mutants. Conversely, animals treated with RNAi to target this collagen are wild type but have an aberrant and amorphous pattern of COL-19::GFP. Because of the differences exhibited by *rol-6* RNAi-treated and N-terminal processing mutants, it can be extrapolated that the latter mutant is not acting as an effective null.

As *rol-6* null mutants (*rol-6(n1178)*) are wild type at the ultra-structural level (Peixoto *et al.*, 1998), ROL-6 is not required for the proper formation of the cuticular layers. However, because of the amorphous pattern of COL-19::GFP observed in *rol-6* RNAi-treated animals, the function of this collagen appears to be required to direct the proper localisation of at least one annular localised collagen (COL-19) into the cuticle. How the absence of ROL-6 in the cuticle affects other collagens (e.g. DPY-5 and DPY-7) remains to be established. However, it appears that the annular furrow structures, the localisation of the DPY-7 class of collagens, and the seam cell-derived cuticle must be relatively wild type in composition or localisation because neither *rol-6* RNAi-treated animals nor *rol-6* nulls (*rol-6(n1178)*) exhibit an obvious Dpy phenotype.

These results show that neither the annulae nor its components, are sufficient to dictate the annular localisation of COL-19::GFP, i.e. that the marker collagen does not simply localise to spaces between the furrows and ROL-6 has an active role in its localisation.

ROL-6 is likely to be a component of the fibrous layers, but the helical orientation of these layers suggests that they are presumably not patterned to reflect the localisation of the annulae or annular furrows. Consequently, since ROL-6 is unlikely to be positioned in a circumferential pattern, it is interesting to speculate, how ROL-6 could dictate the localisation of collagens from its position in these layers. It is possible that ROL-6 is among a number of collagens that interact and that define the position of the annulae. These structural cues may stem from DPY-7 collagens in the annulae or from the furrowed surface of the hypodermis. The role of ROL-6, especially in the latter scenario, would be to transmit the information of the annulae, through the fibrous layer.

Ultra-structural studies by Peixoto *et al.* (1998) demonstrated that N-terminal processing mutants (*rol-6(su1006)*) have wild-type cuticle structures apart from the intermediate layer being filled with a granular material. In wild type animals this layer appears to be an empty space, apart from the presence of the struts. It appears, therefore, that in *rol-6(su1006)* animals, aberrantly processed ROL-6 collagen, and possibly other collagens, might be secreted into the intermediate layer. Exclusion of much of the mutant collagen from its predicted location in the fibrous layer, is supported by the fact that this layer, rather than being disrupted by the mutant collagen, is wild type. However, on the basis of the COL-19::GFP, in these mutants, which localises in a predominantly wild type but slightly amorphous annular pattern, it appears that ROL-6 may still be able to form the hypothesised interactions, whether it be via reduced amounts in the fibrous layer or from its position in the intermediate layer, and thus maintains its ability to direct the localisation of the marker collagen.

#### **4.4.2 Examination of data from Sqt mutants**

##### **4.4.2.1 *sqt-1* mutants**

SQT-1 and ROL-6 were considered to be similar on the basis that, at the gross morphological level, both null mutants are wild type and both become morphologically mutant upon the introduction of an aberrantly folded form of their collagen into the cuticle. As described above, ultra-structural studies revealed that *rol-6* null mutants do not exhibit aberrations in the fibrous

and intermediate layers of the cuticle. However, *sqt-1* nulls do; their cuticles are devoid of struts in the intermediate layer and have additional alterations in the fibrous layers, having only a single fibrous layer that is positioned parallel to the length of the worm (Peixoto *et al.*, 2000). The use of COL-19::GFP to analyse the cuticles of these two null mutants has further demonstrated that there are differences between these loci despite their similar wild type phenotype at the gross morphological level. Using COL-19::GFP, we have shown, for example, that *rol-6* null mutants have an amorphous COL-19::GFP phenotype while *sqt-1* nulls are more wild type, displaying only limited branching annulae and alae. As described above, the presence of ROL-6 in the cuticle is not required for the proper formation of the cuticle layers. However, it plays a vital role in transmitting the position of the annular furrows to overlying layers. The data presented here argues against SQT-1 having such a role due to the fact that annular furrows are defined in its absence. As these *sqt-1* nulls are assumed to have wild type ROL-6, it may be possible that the information about the position of the annulae is mediated via this collagen (ROL-6), even in the absence of the struts or the honeycomb element (refer to figure 4.14). This exemplifies that these ROL-6 "connections" are not sufficient to transmit the Rol phenotype which would be expected from the presence of only a single fibrous layer. It must also be considered that the non-Rol phenotype is derived because of the fibrous layer lying parallel to the alae. However, *sqt-1(sc13)* animals also have a similarly orientated fibrous layer but do exhibit a Rol phenotype (Peixoto *et al.*, 1998).

These data showing the different null phenotypes of SQT-1 and ROL-6 also imply that these collagens are not obligate partners since removal of one of them does not appear to inhibit the incorporation of the other into the cuticle.

*sqt-1* and *rol-6* loci also appear to differ in terms of the effect of introducing aberrantly processed forms of their encoded collagens into the cuticle since *sqt-1* mutants have cuticles which are extremely disrupted; this is in contrast to *rol-6* mutants which are relatively wild type. SQT-1 N-terminal proteolytic cleavage site mutants (*sqt-1(e1350)* and *sqt-1(sc1)*) display severely branching alae and fragmented annulae, and are believed to be the result of the steric hindrance imposed by the relatively large N-termini to either the triple helices or to tightly packed higher order structures. On the other hand, the addition of an improperly cross-linked collagen into the cuticle in *sqt-1(sc13)* mutants is disruptive but to a lesser extent than the N-terminal processing mutants. These *sqt-1(sc13)* mutants display aberrations to the cuticle layers (having a single, parallel-orientated fibrous layer, no struts and a honeycomb element

filling its intermediate layer), a Rol phenotype as well as multiple alae, a twisted cuticle and limited amounts of branching annulae (as observed with COL-19::GFP). It is possible that relatively severe aberrations in the fibrous layer, resulting as a consequence of steric hindrance, are transmitted to the overlying layers via the material that builds up in the intermediate layer.

The more severe *sqt-1* N-terminal processing site mutants have as yet not been characterised at the ultra structural level making it more difficult to analyse. However, they are also likely to be due to severe disruptions occurring in the fibrous and intermediate layers. What the COL-19::GFP phenotypes have been able to illustrate is the increased severity of *sqt-1(sc1)* cuticle aberrations compared to *sqt-1(e1350)* mutants, a difference which is not easily discernible at the gross morphological level. This indicates that the second Cys of the Cys-X-X-Cys motif (the mutated residue in *sqt-1(sc1)*) may be more critical for proteolytic cleavage.

Another question arising from these mutants is: how do the alae become branched as a consequence of cuticle aberrations. Such cuticle defects could arise simply by the aberrant cuticular composition. However, another hypothesis is that cuticle aberrations could somehow affect the seam cells.

The patterns observed with COL-19::GFP being a reflection of the actual annular patterns were confirmed since the pattern obtained from immunolocalisation using antibodies against DPY-7 coincide exactly with that of COL-19::GFP.

#### 4.4.2.2 *sqt-3* mutants

SQT-3 is a unique collagen as it is essential for embryonic development. The *sqt-3* locus is complex and, as stated previously, the data obtained in this study is by no means sufficient to fully analyse this locus. The two *sqt-3* alleles analysed in this study were Gly substitution mutants (*sqt-3(e2117)* and *sqt-3(sc63)*), which are both extreme Dpy at 15°C and lethal at 25°. Similarly, RNAi treatment at 25°C is embryonic lethal. However, below this critical temperature, RNAi-treated animals are wild type. On the basis of the gross morphological data alone, the distinction in phenotypes between the null and the glycine substitution mutants is indicative that the aberrant collagen is being incorporated into the cuticle and is acting as a structural poison to disrupt the cuticle.

Upon analysis of the COL-19::GFP of these *sqt-3* mutants, it is observed that in spite of the severe Dpy phenotype exhibited by the glycine substitution mutants, the corresponding cuticles are surprisingly un-disrupted, showing only a certain degree of branching in the annulae and alae and being much less severe than the aberrations observed in *sqt-1* N-terminal processing mutants. The reason for this phenomenon may be attributed to the fact that *sqt-3* is predominantly expressed in embryonic and L1 stages, and is consequently not a likely critical component of the adult cuticle. Because of this, the adult cuticle is probably being synthesised in an almost wild type manner in these mutants, with the extreme gross morphological phenotype observed, being predominantly a result of the severe cuticular disruption in early stages of its life cycle. The ability for post L1 cuticles to form in a wild type fashion in the absence of *sqt-3* is demonstrated by the fact that wild type animals can be obtained from null mutant strains even above permissive temperatures as long animals are maintained below the threshold during embryogenesis. Post-embryonic increases in temperature do not result in the lethality generally associated with these temperature-sensitive mutants.

#### 4.4.3 Examination of Dpy mutants

Two distinct groups of Dpy mutants were defined according to the classes of COL-19::GFP in this study and by McMahon *et al.* (2003), who defined their collagen mutants on the basis of expression patterns (temporal and spatial) and obligate partners. The collagens falling into the groups defined in both studies were identical and demonstrated that the cuticle pattern observed via COL-19::GFP is determined by the temporal and spatial expression patterns of the collagens. Herein, the different Dpy mutants have been characterised on the basis of their COL-19::GFP localisation and DPY-7 localisation.

*dpy-7*, *dpy-3*, *dpy-2*, *dpy-10*, *dpy-8* and *dpy-9* are the collagen mutants that give rise to a pattern of type II COL-19::GFP localisation, while *dpy-5*, *dpy-13* and *dpy-4* are mutable to type III disruption. On the basis of their RNAi Dpy phenotypes, all these are essential collagen loci for the synthesis of a properly formed cuticle. Most, if not all, members falling into the same category of disruption appear to be obligate partners such that removal of one of these collagens from the cuticle also results in exclusion of the others (it cannot be stated definitely that they *all* act in this manner because the loci responsible for the *dpy-9* and *dpy-4* loci were identified after the McMahon study and thus their obligate partners and interactions with the distinct sets of collagens is not yet defined).

The molecular lesions of all the alleles tested are described in table 4.1. However, based on the similarities between the RNAi- and the mutant- phenotypes of the *dpy-10*, *dpy-5*, *dpy-7* and *dpy-13* loci at both the gross morphological and COL-19::GFP level, it has been suggested that many of these alleles are effectively null mutants in which the aberrantly folded collagens are excluded from the cuticle by being retained by the protein sorting machinery. Because of the relationship with their obligate partners within the DPY-5 and DPY-7 groups, null alleles therefore lack at least two and six collagens respectively.

The information obtained from crosses of the two groups of morphological mutants with TP12 have further characterised their cuticles, and has allowed the molecular basis of their Dpy morphology to be analysed.

#### 4.4.3.1 Class III mutants

Class III mutants exhibit a distinct COL-19::GFP pattern that demonstrates the presence of two distinctly composed cuticles: that overlying the dorsoventral hypodermis and that overlying the seam cells. The COL-19::GFP pattern indicates that the former type of cuticle is patterned by evenly spaced annulae, while the latter exhibits severe branching annulae. Annulae are properly formed in the dorsoventral region of the cuticle, since DPY-7 collagens are secreted in a wild type manner. This was verified by immunolocalisation using anti-DPY-7 antibodies that localised in the annular furrows. These mutant loci (whose alleles are wild type) have been placed into the type III category of COL-19::GFP expression, namely, branched seam cell annulae. The two distinct cuticular lesions are distinguished by their differing susceptibilities to mutation, most likely derived from unique cuticular compositions in each cuticle. DPY-7, for example, is one collagen that has previously been demonstrated to be limited to the dorsoventral cuticle (McMahon *et al.*, 2003).

Since *dpy-5* and *dpy-13* mutants are null for the encoded collagens as well as their obligate partner(s), the collagenous content of the cuticle is reduced dramatically. This accounts for the “fat” nature of Dpy mutants; removal of many of its component collagens weakens the seam cell-derived cuticle, which stretches. This region of the cuticle in non-Dpy animals normally contracts in order to mediate the diametric shrinkage of the animal. As a consequence of its weakening in these mutants, however, there is failure to contract and a corresponding increase in the body circumference ensues (refer back to figure 4.6 D). The disrupted nature of the cuticle overlying the seam cells can be seen by the branching COL-19::GFP pattern. The “short” nature

of Dpy mutants of these loci can also be accounted for by the loss of a proportion of essential collagens. The loss is in the dorsoventral region, since shortness is a consequence of the annulae being thinner (rather than being fewer in number along the length of the worm) than in wild type animals. Since DPY-5 and DPY-13 localise to the annulae in wild type animals, their exclusion in mutant alleles obviously results in the reduction of collagenous material accumulating between the furrows. The data presented here has further demonstrated that the decreased collagen content may not be restricted to collagens that are obligate partners of DPY-5 and DPY-13. In *dpy-5(RNAi)* animals, COL-19::GFP was observed to accumulate in the periphery of the seam cell cords, indicating that some of this molecular marker was being excluded from the cuticle. However, this is not a consequence of COL-19::GFP being an obligate partner of DPY-5 or DPY-13, since most/some of this marker does get processed into the cuticle in these mutants. These observations suggest that the removal of one collagen from the cuticle is sufficient, not only to prevent the secretion of its obligate partner(s), but also to exclude possibly an array of collagens, thus further weakening the cuticle.

#### 4.4.3.2 Class II mutants

Class II mutants exhibit an amorphous COL-19::GFP pattern. This shows that in the absence of annular furrows, COL-19::GFP does not localise correctly, indicating that the marker collagen uses DPY-7 as templates for its localisation. The degree of amorphousness can differ, ranging from finger-like projections that point in multiple directions in *dpy-7(e88)* and RNAi animals, to a ubiquitous covering in *dpy-7(qm63)* null mutants. This indicates that in the Gly substitution alleles and in RNAi-treated animals, a small amount of the collagen is indeed secreted into the cuticle. Unlike the type III mutants, the two distinct regions of the cuticle are not distinguishable. There are thus two possibilities to account for the Dpy body shape of class II mutants: the diametric stretching could be caused by the failure of the seam cell-derived cuticle to contract, or it could be due to the weakening of the dorso ventral cuticle such that it cannot contain the hydrostatic pressure. The latter reason is the most likely since there is no DPY-7 in the seam cell-derived cuticle, and crosses between DPY-7 and DPY-13 animals result in “doubly Dpy” animals (McMahon *et al.*, 2003). All this suggests that the accountable cuticle defects for each of these alleles’ Dpy phenotype are individually mutable. The weakened dorsoventral region of the cuticle in *dpy-7* mutants is likely to be the result of a significant number of collagens being aberrantly localised or missing from the cuticle; all of the obligate partners of this collagen are lost and, in addition, DPY-5, DPY-13 and COL-19::GFP are aberrantly localised. These three collagens are likely to represent only a subset of molecules, outside the

DPY-7 group, that are affected. These data also demonstrate that the annular furrows, which are normally associated with driving predominantly lateral contraction, also have a role in regulating the diametric shrinkage of the animal.

#### 4.4.4 Examination of *bli-1* mutants

BLI-1 is necessary for the proper formation of the struts and in the *bli-1(e769)* mutant allele (the genetic lesion of which is uncharacterised) the COL-19::GFP pattern observed for the mutants falls in to type III disruption. While the disruption in the seam cell-derived cuticle of these animals is indicative that BLI-1 may have a role in the composition of this layer, the wild type pattern in the dorsoventral region indicates that annulae are properly formed. At the ultra structural level, these mutants have lost their struts. However, they do have an intermediate layer that is filled with a honeycomb element that might facilitate such transmittal of cues about the localisation of the annulae to the external layers. The fact that the struts are missing suggests that BLI-2, another collagen required for the struts, cannot compensate for the loss of BLI-1 in order to circumvent the expression of the Bli- or mutant cuticle- phenotypes. This may be evidence that these collagens are obligate partners. Because the intermediate layer is material-filled, it is difficult to extrapolate from these data, the real function of BLI-1 in the cuticle.

#### 4.4.5 Overview of the roles of collagens in the *C. elegans* cuticle

These data obtained from the COL-19::GFP analysis of collagen mutants has shown that different collagens have different roles in the cuticle according to the substructure within which they reside. Described here is a hypothetical scenario based on the inferred roles of these collagens. Because components of the most external layer are predicted to be secreted first, with underlying layers being secreted underneath ("top to bottom synthesis"), the steps have been listed in the order of cuticle layers that the collagens are hypothesised to be expressed in (for example, *dpy-7* being expressed first because its gene product is possibly located in the external most cuticle layer).

*-dpy-7* and its obligate partners are expressed 4 hours prior to the moult and localise to the actin-defined furrows on the surface of the hypodermis. DPY-7 and its related collagens only localise in an annular furrow pattern in the dorsoventral region of the cuticle.

*-dpy-5* and its obligate partners are expressed 2 hours prior to the moult and localise between the deposits of DPY-7 to "pad out" the annulae. The prior definition of the annulae by the DPY-7 group of collagens is a prerequisite for the proper localisation of DPY-5 and DPY-13 in the cuticle. DPY-5 and its partner collagens also localise to the seam cell-derived cuticle, and facilitate diametric shrinkage.

*-col-19* is late expressed. Its late temporal expression (i.e. after that of *dpy-7* and *dpy-5*) has also been demonstrated by the fact that the localisation of COL-19::GFP reflects aberrations in the cuticles of both *dpy-7* and *dpy-5* classes of mutant. The SQ RT-PCR data also shows that COL-19::GFP is being expressed coincident with the moult (i.e. late).

*-bli-1* is expressed and forms the struts, probably to maintain the attachment between the fibrous and cortical layers.

*-sqt-1* and *rol-6* are expressed and form the fibrous layers. ROL-6 possibly interacts with other cuticle layers in order to maintain the furrows in the external layers.

There are some anomalies in this scheme however. This is because, as demonstrated by SQ RT-PCR, *sqt-1* is expressed temporally at a time between that of *dpy-7* and *dpy-5* but is apparently expressed in underlying layers. Thus, it is possible that step 5 comes between steps 1 and 2. Such deviation from the "top to bottom" sequence of collagen secretion may be helpful in explaining another anomaly in the sequence above. This lies in the fact that the Rol mutants cannot be explained by this series; for the overlying cuticles to display the twist of the Rol phenotype, a consequence of the removal of one of the fibrous layers, it follows that these aberrantly formed layers must be deposited first, since secretion afterwards could not really cause the overlying layers to alter their conformation so drastically. If the cuticle is synthesised from "top to bottom", the fibrous layers would not be able to dictate the twist of previously deposited annulae. Translation of the Rol phenotype could only occur, if, as demonstrated by its temporal expression, SQT-1 (and ROL-6) are secreted to form the (abberant) fibrous layers prior to the completion of the external layer.

Rol mutants are useful in illustrating another property of the *C. elegans* cuticle, that is, the relationship between the cuticle, the hypodermis, and underlying body. Mutations that lead to

the RoI phenotype are not only manifested in the cuticle but also have twisted body musculature, internal organs, ectodermal cells and ventral nerve cord. Thus, the aberrations from the cuticle must somehow induce these changes. It has long been established that interactions between the cuticle exist. Further feedback between the cuticle and hypodermis has been described in this study in the discovery that *dpy-5* mutants exhibit seam cell that are aberrantly localised (but which fuse in a wild type manner). This ability of cuticle mutations to affect the hypodermis has also been observed when *dpy-18* mutants are treated with PDI-3 RNAi (Eschenlauer and Page, 2003). This interaction with the hypodermis may be a factor in how branched alae are formed.

Here we have demonstrated, the presence of two distinct regions of the cuticle. The roles of each cuticle have been dissected, and we have shown that the seam cell- and dorsoventral- derived cuticles have distinct but overlapping roles. This observation is a reflection of the different roles of their integral collagens.

#### 4.4.6 Examination of cuticle processing enzyme mutants

Defining the morphological mutants into the groups of COL-19::GFP disruption resulted in the grouping of the enzymatic *dpy-11* and *dpy-18* loci (which encode for a thioredoxin and a prolyl 4-hydroxylase catalytic subunit respectively) into a distinct category. This was on the basis that in such mutants, both alae and annulae are aberrant. *dpy-11* point mutants (*dpy-11(e224)*) display aberrant seam cell-derived cuticles and branched alae; *dpy-11* null mutants (*dpy-11(e1180)*) have amorphous COL-19::GFP expression patterns and branching alae; and *dpy-18* mutant alleles (*dpy-18(e364)* and *dpy-18(ok162)*) have fragmented COL-19::GFP in the seam cell-derived cuticle and also have branching alae. This is, to an extent, expected since it is assumed that the enzymes that fall into this category are part of a pathway responsible for the processing of multiple collagens. However the observations described here show that not all collagens are affected in these mutant alleles. This is indicated by the fact that *dpy-11(e224)* and both *dpy-18* mutants have annulations in their dorsoventral cuticle, the presence of which indicates that the DPY-7 set of collagens, at least, are being deposited in a wild type fashion.

In terms of the null *dpy-18* mutants, the presence of annulae and properly processed DPY-7 suggests that there is partial redundancy for this enzyme's action. This observation supports published data which demonstrates that PHY-2 can replace DPY-18 in the multi subunit P4H complexes (Winter and Page 2000). The region of disruption in the seam cell-derived cuticle is indicative, however, that PHY-2 is not sufficient to completely replace the DPY-18 functions.

The observation of annulae formation in *dpy-11(e224)* mutants is more complicated. This is because the amorphous pattern of *dpy-11* null mutants (*dpy-11(e1180)*) argues against there being redundancy for this enzyme's action. *dpy-11(e224)* represents a partial loss-of-function in which the cross-linking efficiency is reduced. The observations described here therefore indicate that different collagens may have differing requirements for being cross-linked. It appears that DPY-7 and its associated collagens are able to form normally, even if the degree of cross-linking, assumingly, is diminished. The cuticle overlying the seam cells, however, does appear to be more reliant on cross-links since in their absence, it is weakened and fails to contract. This susceptibility of the seam cell-derived cuticle to being weakened is likely to be a consequence of the lack of supporting structures, such as the annular furrows, and has been demonstrated previously by its mutability in *dpy-5* mutants.

In contrast to DPY-18 and DPY-11 that have a large repertoire of substrates, the ablated isoforms in *bli-4(e937)* mutants have only a limited set of collagen substrates, this being reflected in the wild type COL-19::GFP patterns. This observation, which suggests that both DPY-7 and DPY-5 sets of collagens are processed into the cuticle effectively supports previously published analyses; that in the absence of the A, E, F, G and H isoforms, it is only the strut-localised collagens, BLI-1 and BLI-2, that are aberrantly processed. Thus at least one of the aforementioned isoforms is specific for the processing of these collagens. The adult-specific expression of blisters supports this because BLI-1 and BLI-2 are limited to the adult cuticle (Crew and Kramer, personal communication). *bli-4(e937)* is the only viable *bli-4* allele, the others being mostly embryonic lethal (Thacker *et al.*, 1995).

#### 4.4.7 Overview of cuticular enzyme mutants

The observations from the enzyme morphological mutants have shown the complexity and variety that exist in regard to the processing of collagens. It has been established that different sets of collagens require differing degrees of cross-linking (exemplified by *dpy-11* alleles) and that PHY-2 can replace only a proportion of DPY-18s role in cuticle collagen biosynthesis. In addition, the fact that distinct processing occurs at different times of the life cycle has been shown. This was demonstrated by the adult-specific processing exhibited by some BLI-4 isoforms. However, the fact that *dpy-11* mutants exhibit the Dpy phenotype at all stages, implies that certain processing steps/enzymes are required at every larval/adult stage.

#### 4.4.8 Examination of body size mutants

The causative effect of Lon and Sma body size mutants are alterations to the homeostasis of the TGF- $\beta$ -like signalling pathway. Regulation of body size is mediated via individually mutable pathways that control hypodermal ploidy and cuticle composition. Thus these mutants are distinct from the body *shape* altering lesions. In this study, the COL-19::GFP expression of *lon-3(e2175)* and *lon-1(e185)* mutants have been compared. This was on the basis that LON-3 and LON-1 have roles in mediating cuticle composition and hypodermal ploidy respectively. The distinguishing feature of all Lon mutants is the increased width of the annulae. Some *lon-3(e2175)*, but not *lon-1(e185)*, animals additionally show significant cuticle aberrations in the form of branching annulae in the seam cell-derived cuticle. The aberrant expression pattern of COL-19::GFP indicates that the composition of the cuticle is being altered in the latter mutant. Thus this is evidence that the components of the LON-3 branch of the pathway actively alter the cuticle composition, and this is consistent with *lon-3* encoding for a cuticle collagen. On the other hand, the lack of such distinctive cuticle aberrations in *lon-1* mutants suggests that this branch of the pathway, as previously predicted, has no effect on the cuticle structure. The Lon body morphology in these mutants is thus purely derived from the increase in ploidy with the resulting increased body size being accommodated for by the passive stretching of the cuticle. There is no active alteration of the cuticle composition in this process. Despite the absence of severe COL-19::GFP disruption, there are subtle cuticle aberrations in *lon-1* mutants, and this is the likely consequence of mechanical stress from the stretching.

As LON-2 (the identity of which is not firmly established) and SMA-2 (a dwarfin) are upstream components of the TGF- $\beta$ -like pathway, they most likely regulate both the LON-3 and LON-1 pathways (Suzuki *et al.*, 2002). This is not reflected in *lon-2(n1630)* or *sma-2(e502)* mutants, however, since, apart from their characteristic increased and decreased width of the annulae, no severe cuticle aberrations are observed.

#### 4.4.9 Analysis of data from the RNAi screen

A number of enzyme and collagen encoding mutants were chosen as targets for analysis via RNAi with the aim of identifying novel genes with roles in cuticle biosynthesis or moulting. Candidate genes were chosen on the basis of encoded motifs, RNAi phenotype (from previous genome-wide RNAi screens) and expression patterns (data obtained from microarray data). Based on this criteria, the list of targets included: zinc metalloproteases (astacins, ADAM/ADAMTSs and zinc carboxypeptidases), peroxidases, collagens, cuticulins, tyrosinases,

thioredoxins, components of myosin and actin motors, PPIs and proteins with putative roles in protein transport, chaperone function and protein sorting. Not surprisingly, many of the genes fit into more than one of the criteria for selection. This can be seen in figure 4.11. Before the results are discussed, the validity of RNAi as a technique must be analysed.

#### 4.4.9.1 Validity of the RNAi technique for this screen

The validity of RNAi feeding as a reliable technique can be questioned by a few arguments. Firstly, the low penetrance of RNAi of some genes could result in the RNAi technique being less effective, meaning the RNAi phenotype would not reflect the importance of the gene in the cuticle. The second issue for consideration is that of gene redundancy in which the down-regulation of one of the genes sharing similar roles is not sufficient to result in an observable phenotype. Both these properties could lead to negative or weaker-than-null results. It is important, however, to note that this screen is useful for identifying positive results only and by no means is a method sufficient to dismiss genes as having no role in the cuticle. Another consideration in the analysis of the results is non-specific RNAi of homologous genes. In such a case, an injected RNAi may target both the intended gene as well as homologous genes. Such a phenomenon is rare in *C. elegans* RNAi, and normally occurs only in cases where there is >85% identity.

After RNAi by feeding, more reliable techniques must be carried out for full characterisation of a gene. RNAi by soaking and injection are often more potent than RNAi by feeding and may increase the penetrance. However, the only way of ensuring complete ablation of signal is using a genetic null. In the absence of a visible phenotype, a western blot utilising antibodies against the protein in question is an invaluable technique in determining whether there is a decrease in protein expression; when comparing RNAi-untreated and -treated samples, a decrease in signal would reflect effective targeting of the gene. Armed with such knowledge, it can be determined whether the lack of phenotype is due to inefficient RNAi targeting of the gene, or if there is knockdown in the signal and the true RNAi phenotype is wild type.

When an RNAi phenotype is observed, the specificity of the RNAi has to be proven. Because of the resulting phenotype, efficient interference of the target gene is assumed, but to what extent the phenotype is caused by RNAi of other genes has to be investigated. Feasible only if there are limited number of homologues in the genome, Western analysis using specific antibodies can be used to observe whether there is down-regulation of homologues. However, RNAi-targeting of

another (non-homologous) region of the gene is a simple alternative to this technique. In addition, once a genetic null is available, the phenotypes of the two can be compared; if the RNAi effect is more severe than the null, other genes are presumably being targeted.

Despite the questions that are brought about, RNAi by feeding is an excellent system for high-throughput screening in the search for genes that give observable phenotypes when targeted.

#### 4.4.9.2 Results from RNAi

A variety of combinations of gross morphological and COL-19::GFP phenotypes have been observed in the screen. This includes wild type animals with amorphous COL-19::GFP patterns, weak Dpy animal with wild type COL-19::GFP, Dpy phenotypes with both amorphous and branched seam cell annulae and Bli animals with severe COL-19::GFP disruption. In addition, RNAi of some genes resulted in uncoordinated movement (Unc), moulting defects, altered body shapes and lethality. However, no correlation was found between these phenotypes and their COL-19::GFP expression patterns.

Some of the above combinations are interesting in light of the conclusions made for mutant strains on the molecular basis of their Dpy phenotypes. This includes the fact that Dpy animals were observed with wild type patterns of COL-19::GFP when treated with RNAi for t06d8.10, w09d12.1 and *col-38*, which encode for an animal haem peroxidase, an ADAM and a collagen respectively. This implies that there is no significant weakening of the seam cell cuticle in these mutants in the same way as there is in *dpy-5*, and its associated collagens, mutants. Neither, however, is the molecular basis of the Dpy phenotype of these RNAi-treated worms due to the lack of DPY-7 localising to the cuticle. These observations show that in such cases, despite most of the cuticle components being secreted into the cuticle successfully, there is a degree of disruption. This probably weakens the dorsoventral cuticle in order that it stretches in accordance with the internal body hydrostatic pressure. Only weak Dpys are observed in these cases.

Quite frequently, the opposite effect is observed, with wild type animals being associated with COL-19::GFP patterns that are amorphous. However, apart from f09f3.5 (an animal haem peroxidase), dy3.7 (an ADAM) and f21c3.2 (a tyrosinase), the amorphous pattern was only minimal, with annular structures still being defined. Despite representing cuticles with differing

degrees of disruption, these RNAi phenotypes indicate that the aberrations are not sufficient to translate into a morphological phenotype. This is most likely due to the fact that the RNAi-targeted gene has only a minor role, but it must also be considered that a situation similar to that observed in *rol-6* nulls may be occurring, i.e. that relatively severe cuticular aberrations are occurring in the cuticle but, due to the lack of interlayer connections, neither information regarding the annulations nor the gross morphological phenotype can be transferred.

Although there is no correlation between gross morphological phenotype and type and severity of COL-19::GFP expression, there are correlations between gene identity and the cuticle pattern. It has been demonstrated here that none of the thioredoxins or tyrosinases resulted in severe COL-19::GFP disruption. However, there are a number of factors that could explain the lack of a phenotype, including 1) the fact that there may be high degrees of redundancy with similar proteins, 2) that the activity of these enzymes is not required for proper cuticle synthesis, 3) that they are not active, and 4) that they are resistant to RNAi. This reiterates the fact that the negative results from this screen are by no means conclusive.

If an animal displays a mutant phenotype after RNAi treatment, it suggests that the target gene has some activity. The degree of disruption cannot be confirmed to be that of a null making the findings inconclusive. A proportion of genes "positive" for an RNAi-induced phenotype have been identified, and these fall into the collagen, protease and animal haem peroxidase categories. However, not all the members of these families exhibit phenotypes, suggesting either differing roles, different levels of redundancy within the groups or differing susceptibilities to RNAi. The latter possibility was demonstrated here by the conflicting results from the weak level of COL-19::GFP disruption as a consequence of *dpy-31/mas-35* RNAi (targeting an astacin) and the lethality of the DPY-31/NAS-35 mutant presented by Novelli *et al.* (2004).

RNAi of collagens is quite complex since the removal of some collagens, such as *sqt-1*, from the cuticle results in cuticle defects in fibrous layers but no distinguishing gross morphological or COL-19::GFP phenotype. However, two collagens, *col-100* (t21d12.2) and *col-134* (y41e3.2), that were targeted in this study have, based on their RNAi COL-19::GFP been attributed to be the mutant loci of *dpy-9(e12)* and *dpy-4(e1166)* respectively. This was on the basis of the amorphous phenotypes that *dpy-9* morphological mutants and *col-100* (t21d12.2) animals exhibit. Likewise, the *dpy-4* and *col-134* (y41e3.2) cuticle patterns are similar (branched seam cell annulae). Subsequent to this observation, the chromosomal locations of the collagen genes

and the morphological mutants were analysed and were found to coincide. Coincident with these findings, Simmer *et al.* (2003) identified the same loci from an RNAi screen using an RNAi-sensitive strain.

In terms of the proteases, there is no evidence to suggest that the presence of certain functional domains predispose proteases to function in cuticle processing. This was demonstrated by the fact that an astacin of subgroup III, an astacin of subgroup V and two ADAMs were among the four most disrupting genes. The data obtained from RNAi targeting of proteases has also confirmed that both NAS-37 (astacin subgroup V) and SUP-17 (ADAM) have roles in cuticle synthesis distinct to their more characterised respective moulting and gonadogenesis processes. Evidence for this came from the aberrantly formed cuticle (COL-19::GFP) of animals RNAi-treated for these genes. Also, the previously uncharacterised NAS-38, NAS-28 and f08c6.1 have been implicated in cuticle synthesis based on their severe COL-19::GFP phenotypes when mutated. It should be taken into account that, because of the ineffective nature of RNAi, especially on astacins, the phenotypes presented here may not necessarily resemble the null form of this astacin and ADAM.

Other interesting COL-19::GFP phenotypes were observed for r160.1 (*dpy-23*), c34g6.6, c18c9.2 which encode for an adaptin, a hypothetical cut-1 homologue and a protein involved in ER translocation respectively. From these analyses it is proposed that r160.1 (*dpy-23*) and c18c9.2 are involved in the transport of cuticle components into the cuticle while c34g6.6 is likely to be an essential component of one of the ECMs.

The most dramatic COL-19::GFP disruption was observed after the Duox enzymes and zk430.8 were targeted by RNAi. RNA-treated animals were severely blistered and had severe body morphology defects and at the COL-19::GFP level, they displayed severely branched alae and fragmented annulae. The phenotype of ZK430.8 being similar to that of the cross-linking Duox enzymes suggested that this enzyme, now termed HPX-1, might have a similar role. It was chosen as a candidate gene for further analysis and this is described in chapter 5.

#### 4.4.9.3 Overview of the applications of the RNAi screen

This screen has established TP12 as a useful and sensitive marker strain that can be used to identify cuticular aberrations resulting from RNAi. Here, it has been used to show that aberrant cuticles can occur in the absence of an obvious gross morphological defect and has shed some

light on the potential bases of morphological defects. In this screen, a number of genes have been highlighted due to their RNAi-induced COL-19::GFP disruption, even in the absence of a gross morphological phenotype. These have subsequently become targets for further characterisation in this study. Examinations of *hpx-1*, in this study (refer to chapter 5), and *t10h10.2* by A. Birnie (unpublished results), have subsequently confirmed their function in cuticle synthesis. The data collected here can also be analysed on a more global level, and it has been suggested that, due to the lack of RNAi phenotypes of tyrosinases, there is possible redundancy between these genes. However, it must be reiterated that care must be taken when analysing negative RNAi data because of the factors described above which may minimise the efficiency of the RNAi effect. On the other hand, analysing the data globally has shown that there is less redundancy between different proteases, since many of the protease genes analysed resulted in RNAi phenotypes. Thus, TP12 has been shown to be an invaluable preliminary tool for the study of cuticular components.

#### 4.5 Future work

Since the work in this chapter has approached cuticle synthesis from a global angle, comparing the cuticles of different collagen, enzyme mutants and RNAi-treated animals, the next natural step would be to return to the cuticular components that have been identified and continue more specific studies. This could include further immunolocalisation assays, making use of other resources available such as the Ty-tagged DPY-13 construct and the anti DPY-7 antibody. Such studies could provide more information on the localisation of other collagens in the cuticle, which in this study have been, to some extent, presumed. One of the most interesting mutants to analyse would be *rol-6* null/RNAi-treated animals in order to observe whether the DPY-7 collagens localises in discrete bands that correspond with the expected location of the annulac.

As described in chapter 3, it would be useful to create a GFP fusion marker (e.g. the *gfp*-tagged *col-19* under the influence of the *sqt-1* promoter) that can be utilised in a similar screen but which additionally enables the cuticle of larval stages to be visualised. This is especially interesting in the light that different processing occurs at different times.

**Chapter 5**  
***Analysis of HPX-1 (ZK430.8), a novel cross-linking enzyme***

## 5.1 Introduction

### 5.1.1 Aim and introduction to animal haem peroxidases

The RNAi screen described in the previous chapter highlighted *zk430.8* as a gene that, upon targeting by RNAi, produces severely blistered worms. Such a phenotype is indicative of cuticular defects since the fluid-filled blisters result from the separation of the medial layer due to aberrations to, or complete loss of, the struts. We have termed the gene *hpx-1* because of the presence of an animal haem peroxidase domain and will refer to it as such from here onwards.

Two animal haem peroxidase domain-containing proteins that have roles in cuticle synthesis have previously been characterised in *C. elegans*. These hypodermally expressed genes, *f56c11.1* and *f53g12.3*, which arose through a gene duplication, were termed *Duox-1* and *-2* (dual oxidase) on the basis of the presence of two oxidase motifs: a peroxidase domain, followed by a flavin NADPH-oxidase domain (Edens *et al.*, 2001). *f56c11.1* was later attributed to the *bli-3* mutant locus (Simmer *et al.*, 2003) and will be referred to as such henceforth. The presence of a peroxidase domain in these proteins, a domain firmly established as being responsible for catalysing the formation of di-, tri- and isotyrosine cross-links (Fetterer *et al.*, 1993) was strong evidence that the Duox enzymes were involved in the cross-linking of the *C. elegans* cuticle (Edens *et al.*, 2001). This idea was consistent with the morphologically mutant *bli-3* alleles (*bli-3(e767)* and *bli-3(n529)*) and the strong Bli phenotype of RNAi-treated worms (Edens *et al.*, 2001).

The haem-containing family of peroxidases is made up of animal haem peroxidases and a distinct set of peroxidases found in fungi, bacteria and plants. These two subfamilies differ greatly in sequence identity but have a common function, which arose via convergent evolution, to result in a common enzymatic mechanism (Taurog, 1999). Despite the members of the distinct families being highly divergent at the sequence level, crystal structure analyses of the active sites have demonstrated that catalytic residues may be conserved spatially around the active site haem (Taurog, 1999). This catalysed oxidative reaction mechanism centres on the use of hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) as an electron acceptor and the numerous states of the iron haem centre of the enzyme for a wide variety of biological processes. The former subfamily includes myeloperoxidase (MPO), eosinophil peroxidase (EPO), lactoperoxidase (LPO), thyroid peroxidase, prostaglandin H synthase and peroxidasin, while the latter includes yeast cytochrome

c peroxidase, ascorbate peroxidase, bacterial catalase peroxidase, lignin peroxidase and secretory plant peroxidase (Daiyasu and Toh, 2000). The definition "animal haem peroxidase" has recently been found to be seemingly irrelevant since linolate diol synthase from fungus and an oxygenase from plants have been identified as homologues of animal peroxidases, which therefore places these fungal and plant enzymes into a class historically associated only with animal systems (Daiyasu and Toh, 2000). Peroxidases are ubiquitous and multifarious enzymes and constitute a vast subject that is impossible to fully describe here. The primary, and most characterised, role of peroxidases is in the defence against organismal invasion, wounding and toxic peroxides. This process is facilitated by the ability of peroxidases to utilize a large number of substrates including chlorine, nitrate, and thiocyanate to yield cytotoxic hypohalous acids.

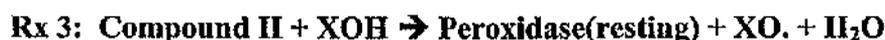
In the peroxidative reaction, the ferric enzyme can oxidise the substrate by the direct transfer of 2 electrons via an oxidised intermediate termed "compound I". Scheme 1:



(Where compound I contains 2 oxidising equivalents more than the resting enzyme) (Furtmuller *et al.*, 1998; Tien, 1999)

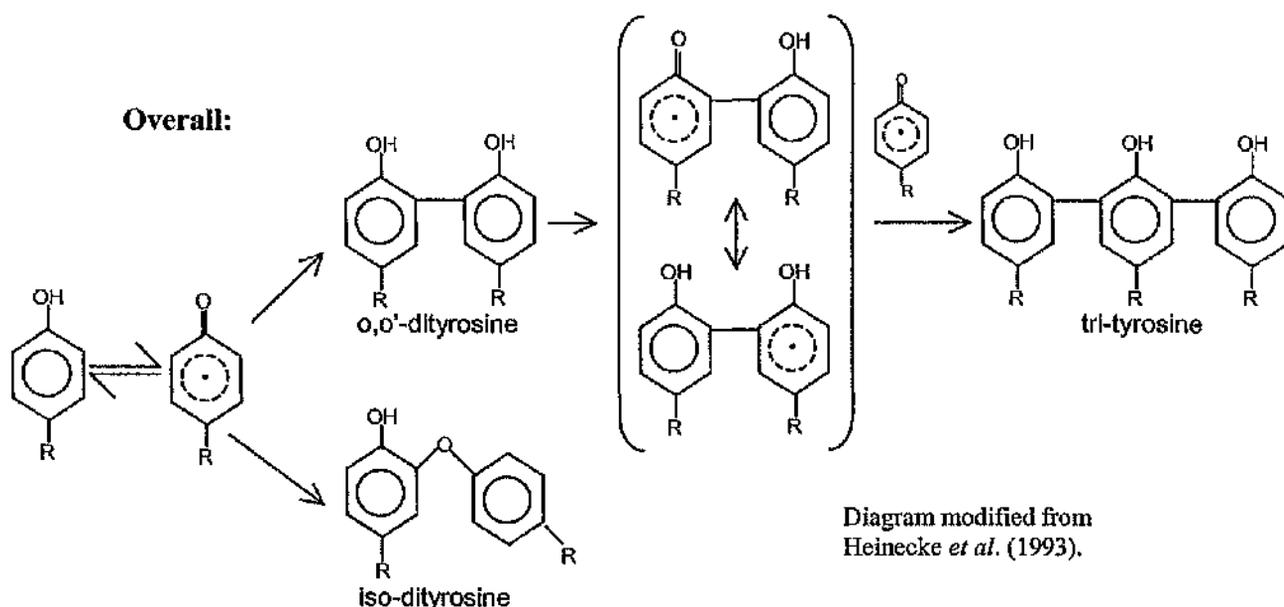
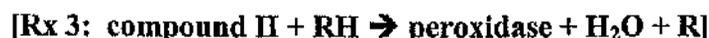
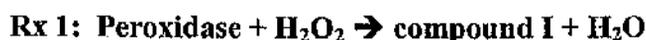
Via this reaction, chloride, nitrate and thiocyanate substrates produce their corresponding hypohalous acids (HOX) or the hypohalite ions (OX<sup>-</sup>). In aqueous media, an HOX/OX<sup>-</sup> equilibrium is reached (Furtmuller *et al.*, 1998).

Alternatively, the two electrons can be transferred to the substrate sequentially, a process that results in the formation of two free radicals. Scheme 2:



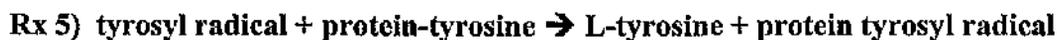
(Where compound II is a one electron oxidised intermediate) (Furtmuller *et al.*, 1998; Tien, 1999; Furtmuller *et al.*, 2001)

It is the above reaction with a L-tyrosine substrate that produces the tyrosyl radicals that partake in the peroxidase-mediated formation of the di-amino acid, *o,o'*-dityrosine. While this is the major product, isodityrosine and trityrosine are also produced (see overall reaction) (Heinecke *et al.*, 1993; Heinecke, 2002). Scheme 3:



Alternatively, the free tyrosyl radicals can diffuse away to interact with protein-bound tyrosine residues (reaction 5 below). The resulting protein-bound tyrosyl radicals are key in the formation of internal tyrosine cross-links (reaction 6 below). Such tyrosine protein cross-linking is found within atherosclerotic plaques, thyroid hormone and in sea urchin eggs (Nomura *et al.*, 1990). Important for the theme of this thesis, tyrosine cross-linking has also been characterised in the cuticle of *Ascaris suum* nematodes (Fujimoto, 1975) and *H. contortus* (Fetterer and Rhoads, 1990).

Scheme 3b: (Tien, 1999) (Heinecke, 2002):



Regardless of the substrate, the prerequisite for any peroxidase reaction is the presence of the  $H_2O_2$  donor which serves as the electron acceptor. Thus, the source of the  $H_2O_2$  must be firstly be described.

### 5.1.2 Duox enzymes are $H_2O_2$ generating enzymes with potential peroxidase activity

$H_2O_2$  is released extracellularly by neutrophils in response to TGF- $\beta$ 1 and various cytokine and growth factor stimuli (e.g. epidermal growth factor (EGF) and platelet derived growth factor (PDGF)) (Lambeth *et al.*, 2000) and due to its localisation at sites of damage, it is associated with playing a central role in tissue injury and repair (Larios *et al.*, 2001).  $H_2O_2$  is also produced by activated neutrophils and eosinophils during the process of host defence at sites of organismal invasion (Lambeth *et al.*, 2000). The  $H_2O_2$  release, and generation of other reactive oxygen species (ROS) is mediated by a membrane-associated and FAD-containing NADPH-oxidase enzyme which utilises the 4 position of the nicotinamide ring of NADH and NADPH as a carrier of hydride ions (Thacker *et al.*, 1995; Lambeth *et al.*, 2000; Larios *et al.*, 2001). Superoxide produced by this calcium dependent enzyme subsequently dismutates (enzymatically or spontaneously) into  $H_2O_2$  (Turner *et al.*, 1985; Heinecke *et al.*, 1993).

The NADPH-oxidase in leukocytes is gp91phox and this enzyme requires the presence of p22phox, p47phox and p67phox (De Deken *et al.*, 2002). Human gp91phox homologues include Nox-1 – Nox-5 as well as two enzymes called Duox/Thox-1 and -2 (Lambeth *et al.*, 2000; Edens *et al.*, 2001; De Deken *et al.*, 2002) .

The latter two enzymes, the Duox enzymes, exhibit ~26% identity (and 53% and 47% similarity) with gp91phox at the amino acid level (Edens *et al.*, 2001) and were originally identified as being expressed predominantly in a thyroid/thyocyte-specific manner (Edens *et al.*, 2001; De Deken *et al.*, 2002; Pachucki *et al.*, 2004). However, more recently these have also been discovered to be localised on mucosal surfaces such as the gastrointestinal tract and the lung (Geiszt *et al.*, 2003). Expression has also been observed in the placenta, testis, prostate, pancreas and heart (Edens *et al.*, 2001). Despite exhibiting sequence similarity to gp91phox, these enzymes are unable to replace gp91phox activity in leukocytes (De Deken *et al.*, 2002). Thus, these  $H_2O_2$  generating systems are analogous but not identical. This is confirmed by the fact that human thyrocytes lack homologues to p47phox and p67phox which are essential to

leukocytes (De Deken *et al.*, 2000). De Deken *et al.* (2002) showed that H<sub>2</sub>O<sub>2</sub> production by the Duox enzymes is not possible upon cDNA-transfection of cell lines, which indicates that Duox enzymes require (thyroid) specific factors for H<sub>2</sub>O<sub>2</sub> production (Wang *et al.*, 2005). *C. elegans* homologues of the Duox enzymes exist (BLI-3 and F53G12.3) and their roles will be discussed below (section 5.1.5).

The H<sub>2</sub>O<sub>2</sub>-generating capability of these gp91phox homologues and their co-localisation with peroxidases e.g. TPO and Duox in the apical membrane of human thyroid cells (De Deken *et al.*, 2002), LPO and Duox 2 in salivary glands and mucosal surfaces (Geiszt *et al.*, 2003) and MPO and gp91phox in neutrophils (Lambeth *et al.*, 2000) is evidence of their role as H<sub>2</sub>O<sub>2</sub> generators for their partner peroxidases.

The human Duox enzymes are unique in having an additional peroxidase domain encoded on their N-terminus, this domain sharing 43% similarity with TPO (De Deken *et al.*, 2002). However, such a role seems rather redundant since it would be a duplication of the activity of the co-localised LPO/MPO or TPO.

A scenario is established in which during host defence, NADPH-oxidases and their dependent peroxidases (separate or coupled) accumulate at sites of tissue damage to produce both ROI, by NADPH-oxidase, or cytotoxic hypohalous acids, by peroxidases, in order to facilitate the bacteriocidal and cross-linking activities (sections 5.1.3 and 5.1.4 respectively).

It must be noted that despite the presence of the peroxidase domain in the Duox enzymes, much of the focus of these enzymes has purely been on their role in H<sub>2</sub>O<sub>2</sub> donation. Their isolated peroxidase domains are able to catalyse the cross-linking of tyrosine residues *in vitro*, as was demonstrated by Edens *et al.* (2001). However, in humans, the biological significance of such activity has not been concluded (De Deken *et al.*, 2002). It has been suggested that tyrosine cross-linking possibly plays a role in the early stages of elastin biogenesis but this has yet to be proven (Edens *et al.*, 2001). In *C. elegans*, on the other hand, the peroxidase activity has been positively attributed to cross-linking the cuticle via the mechanism described above (scheme 3). Despite this, in humans at least, much of the importance of these enzymes does lie in their role as H<sub>2</sub>O<sub>2</sub> donors rather than their peroxidase activity. Duox is not sufficient to overcome the loss of an exogenous peroxidase. This is indicated by the fact that bacterial clearance from airways in sheep is LPO-dependent (Gerson *et al.*, 2000).

Whether due solely to their H<sub>2</sub>O<sub>2</sub> generation activity or to any biologically significant peroxidase activity as well, these Duox enzymes do have an essential role in human and *C. elegans* systems. This is demonstrated by the severe morphological defects seen in *C. elegans* upon their disruption (Edens *et al.*, 2001) and by the hypothyroidism of human patients with biallelic inactivation mutation of the Duox genes (Heinecke *et al.*, 1993).

### 5.1.3 Peroxidase roles that do not involve cross-linking

Each of the characterised animal haem peroxidases, MPO, EPO, LPO and TPO, has multiple functions due to the fact that the products from their oxidation of halides can have a number of fates. The most characterised role of the array of peroxidases is in the host armory system which is both bacterocidal and antiparasitic (Lambeth *et al.*, 2000). Secondly, peroxidases are able to protect against self-inflicted tissue damage from the system's own production of toxic by-products from mitochondrial and chloroplast electron transport systems and from oxidative stress, since products such as hypothiocyanate are milder oxidants and are less harmful to the host than some other free radical species (Daiyasu and Toh, 2000).

With regard to their role in host defence, MPO and EPO are used by neutrophils (Furtmuller *et al.*, 1998) and eosinophils (defence against cancer and larger invading metazoans (antiparasitic)) (O'Brien, 2000) respectively. LPOs are associated with most mucosal surfaces (c.g. gastrointestinal tissue, the lung and brochium) as well as being secreted in milk, saliva, tears and mucus. TPO is antibacterial and is involved in clearance systems in airways (Gerson *et al.*, 2000). The specialised role of each of the peroxidases comes from their different specificities for halide ion substrates (Morrison and Schonbaum, 1976). In physiological conditions, the only substrates that MPO, EPO and LPO can each catalytically oxidise are chloride, bromide and iodide respectively (Domigan *et al.*, 1995; Furtmuller *et al.*, 2000; Gerson *et al.*, 2000). However, their substrate specificities can be altered by varied conditions. At only acidic pH, for example, MPO is able to oxidise all three substrates while EPO can use chloride (Furtmuller *et al.*, 1998). EPO, MPO and LPO can all also catalyse the oxidation of thiocyanate. The resulting hypohalous acids, hypochlorous acid, hypobromous acid and hypothiocyanate, are extremely cytotoxic. Because of the potency of their products at sites of invasion and wounding, a side-effect of peroxidases is oxidative tissue damage (Furtmuller *et al.*, 2001). However, the resulting damage is less than would be expected from other ROS such as O<sub>2</sub><sup>-</sup>. The damaging effect is shown by the accumulation of peroxidases in pathogenic disorders, such as atherosclerotic tissue

and certain tumours, and being attributed to the tissue injuries associated with hypersensitivity reactions (e.g. asthma), and allergic inflammation disorders (Heinecke, 2002).

Peroxidases can also be biosynthetic, as exemplified by the role of TPO in the production of thyroid hormone. During this process, TPO first catalyses the formation of mono- and di-iodotyrosine from tyrosine and iodide and subsequently catalyses their coupling to form thyroxine and triiodotyrosine (Gavaret *et al.*, 1981).

MPO is the most characterised peroxidase to date, and additional roles for the enzyme have therefore been identified. Firstly, and of interest in light of the subject matter of this thesis, MPO has been implicated in controlling the activity of matrix metalloproteases (MMP) (e.g. MMP-7). These proteases are implicated in membrane degradation/dynamics and some MMPs even exhibit true collagenase activity (Woessner and Nagase, 2000). MPO's peroxidation product, HOCl, can both activate and subsequently inactivate these metalloproteases. Inactivation is achieved by HOCl's ability to alter internal cross-linking and peptide structure (Fu *et al.*, 2004).

Another role of MPO is in the regulation of inflammation via the catalysis of nitric oxide. The production of inflammatory nitryl chloride ( $\text{NO}_2\text{Cl}$ ) and nitrogen dioxide ( $\cdot\text{NO}_2$ ) from this oxidative reaction as well as the MPO-mediated nitrosylation of tyrosines in fibronectin are factors in the inflammatory response (Eiserich *et al.*, 2002) (Balduis *et al.*, 2001). LPO and HRP can also form  $\cdot\text{NO}_2$ ; however, their roles are less characterised (vanderVliet *et al.*, 1997).

MPO is the only animal haem peroxidase able to catalyse the production of HOCl, a strong oxidant that reacts readily with certain proteins to denature and change the adhesive properties of tissues (e.g. in human umbilical veins) during inflammation (Vissers and Thomas, 1997). It has been postulated that HOCl oxidises certain ECM components such as type IV collagens, laminin, and thrombospondin, changing their conformation and thus altering the matrix composition (Vissers and Thomas, 1997). There are also reports of HOCl's ability to oxidise pyridinium bonds (Daumer *et al.*, 2000), bonds that are associated with the cross-linking of human type I and III collagens; these bonds are also found in *C. elegans* cuticles (Davis *et al.*, 1982). In the presence of MPO-derived HOCl, loss of basement membrane matrix integrity was observed in bone collagen (Daumer *et al.*, 2000). Interestingly, Davis *et al.* (1982) observed an age-specific increase in these pyridinoline cross-links in the cuticles of *C. elegans* which, together with the

MPO-derived lipid peroxidation, tyrosine- and lipid cross-linking, were deemed to result in a cuticle with decreased elasticity and solubility (Davis *et al.*, 1982). Conversely, HOCl has been observed to result the stiffening of the cuticle of adult desert locusts, *Schistocerca gregaria* (Andersen, 2004). This is mediated by the chlorinated tyrosine derivatives that are formed as by-products of MPO activity (Andersen, 2004). Such halogenated tyrosine derivatives have previously been identified in proteins, particularly structural proteins, from marine invertebrate. The high degree of chlorination is thought to increase the strength and decrease the flexibility of the protein chains by stabilising non-covalent inter-chain interactions (Andersen, 2004). Thus MPO appears to have conflicting roles, since it associated with both the degradation of possibly stabilising collagen cross-links and also the strengthening of the cuticle.

On the same subject of HOCl's ability to alter the matrix, MPOs are also able to alter the composition of the basement membrane indirectly via their ability to activate proteolytic processing of the membrane by elastase (Daumer *et al.*, 2000). MPOs also mediate low density lipoprotein peroxidation and can also alter adhesion ability (although this latter function is not mediated by the peroxidase domain) (Sritunyalucksana *et al.*, 2001).

#### 5.1.4 Tyrosine cross-linking activity of peroxidases

When L-tyrosine, the most easily oxidised aromatic amino acid, is the target of MPO oxidation, tyrosine cross-links result (refer to schemes 3 and 3b) (Heinecke *et al.*, 1993). Due to the restricted access of MPO's active site, tyrosine cross-links are formed via a "mediation mechanism". During this process, the free and diffusible tyrosyl radicals are formed and are used to convey MPO's oxidative action to protein-bound tyrosyl residues which are not accessible to the catalytic domain of the enzyme. This indirect oxidative action produces protein-bound tyrosyl radicals which, upon encountering similar species, form protein-bound dityrosine carbon-carbon cross-links (scheme 3 and 3b) (Heinecke *et al.*, 1993; Tien, 1999; Heinecke, 2002). Dityrosine formation is dependent on the presence of H<sub>2</sub>O<sub>2</sub>, the rate of radical formation, and the half-life of both the free and protein-bound tyrosyl radicals, which are more stable relative to OH· radicals (Tien, 1999; Larios *et al.*, 2001). While dityrosine (o, o'-dityrosine) is the major product of oxidation, trityrosine can also be produced when o, o'-dityrosine itself becomes the substrate for MPO activity (scheme 3). A dityrosyl radical is produced which is able to react with another tyrosyl radical (Heinecke, 2002). In addition to the formation of carbon-carbon bonds of di- and tri-tyrosine, carbon-oxygen cross-links of

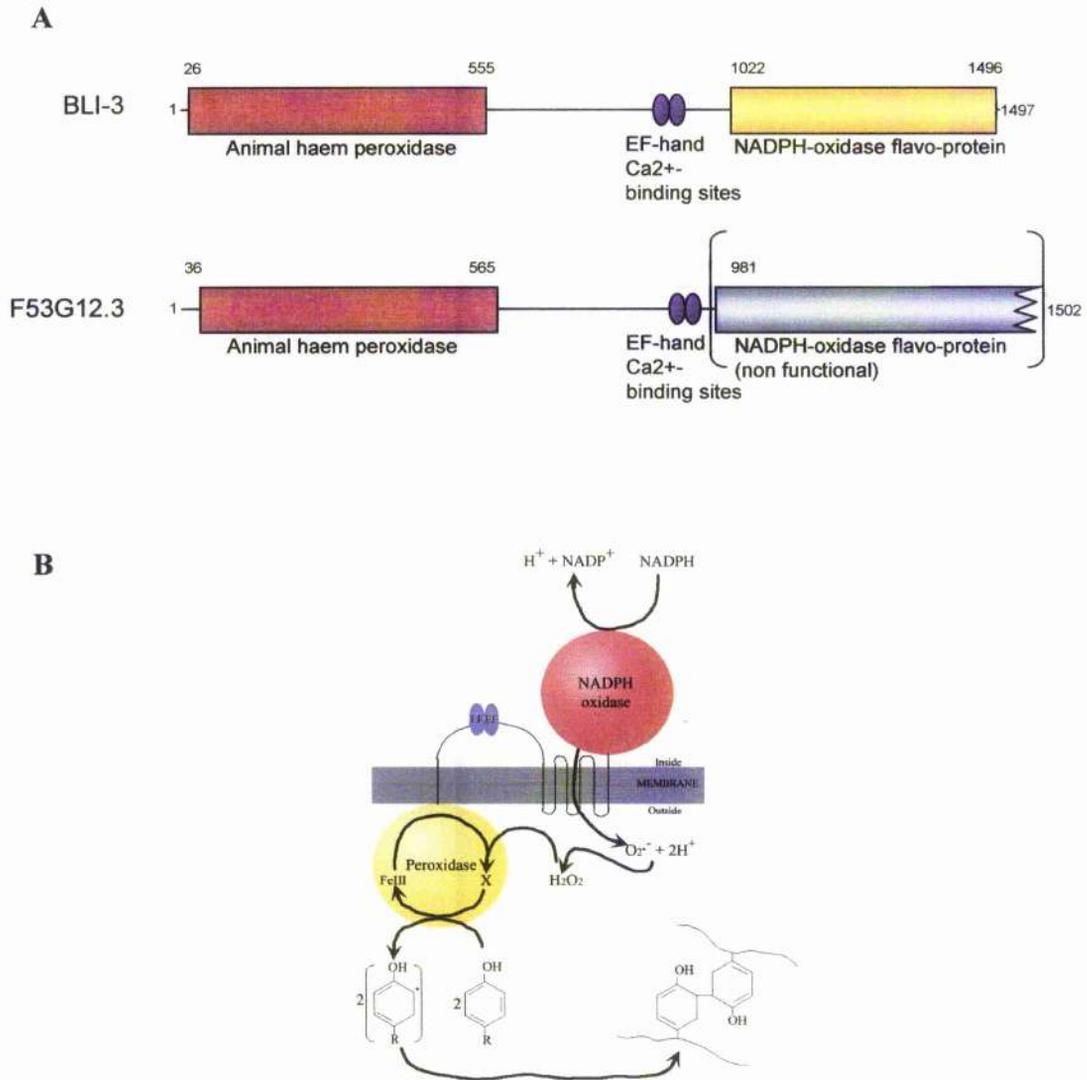
pulcherosine and isodityrosine are also produced via the peroxidation reaction (Heinecke, 2002) scheme 3).

OPO (the enzyme responsible for the post-fertilisation maturation of the sea urchin egg), LPO, EPO, and HRP are also able to catalyse the formation of dityrosine protein cross-links. However they differ from MPO in the fact that they are able to directly access and oxidise protein-bound tyrosyl residues without the requirement for an intermediate redox mediator to convey the oxidative potential (Furtmuller *et al.*, 2001; Heinecke, 2002).

As stated above, collagen, cuticulin and additionally elastin, have been proven to be the targets of protein tyrosine cross-linking a variety of nematodes (Fujimoto *et al.*, 1981; Sakura and Fujimoto, 1984; Fetterer and Hill, 1993; Larios *et al.*, 2001). Additionally, cuticulins, which are found in high levels in the dauer larvae and also localise to the alae, have been demonstrated to be cross-linked. This is in contrast to ECMs of vertebrates in which the level of tyrosine-derived cross-links is low. Nematode cross-linking is also unique due to the fact that these are the only organisms to exhibit cuticular isotryrosine linkages (Fetterer and Hill, 1993). It is such peroxidative action on collagens that is proposed to be mediated by the *C. elegans* Duox proteins (Edens *et al.*, 2001).

### 5.1.5 *C. elegans* Duox enzymes

As described above, *C. elegans* Duox enzymes were confirmed to exhibit tyrosine cross-linking activity that is implicated in stabilising the cuticle (Edens *et al.*, 2001). These enzymes are expressed in the hypodermal cells below the cuticle of larval animals (Edens *et al.*, 2001). Excluding the C-terminal region, BLI-3 (Duox-1) and Duox-2 share ~94% identity to each other and ~30% identity with their human counterparts. Duox-2 differs from *bli-3* in having a truncated NADPH-oxidase domain as a result of a premature stop codon (Edens *et al.*, 2001) (figure 5.1). The similarity of these genes and their juxtaposed position on the end of chromosome I indicate that these genes probably arose as a result of a gene duplication. The proposed topology of the Duox enzymes is shown in figure 5.1 and demonstrates the efficient system of coupling of the H<sub>2</sub>O<sub>2</sub>-donating and H<sub>2</sub>O<sub>2</sub>-utilising moieties. It should be noted here that the peroxidase domain of BLI-3 has been attributed to belonging to a novel peroxidase family due to the fact that three of the active site residues that are conserved in other peroxidases, are not similarly conserved in BLI-3. BLI-3 also differs in the pattern of cysteines, indicating that this domain folds in a novel manner.



**Figure 5.1: Domain structure and topology of the Duox enzymes** A: The protein domains of BLI-3 and F53G12.3. Both proteins contain a peroxidase domain (red) (that catalyses the formation of protein-bound tyrosine cross-links) and an EF-hand motif (purple ovals). While BLI-3 has an active NADPH-oxidase domain (yellow), f53g12.3 has a premature stop codon that eliminates the extreme C-terminus of the NADPH oxidase domain, rendering it inactive (grey). B: The coupling of an NADPH-domain to the peroxidase domain is an efficient way of providing the peroxidase with its  $\text{H}_2\text{O}_2$  substrate during the formation of cross-links in the *C. elegans* cuticle. A and B are adapted from Edens *et al.* (2001).

Edens *et al.* (2001) described the severe Bli and Dpy phenotypes observed upon RNAi-treatment using both *bli-3* and *f53g12.3* RNAi. The blisters are the result of the loss of and distension of the struts. The RNAi phenotype is more severe than the *bli-3* alleles, *bli-3(e767)* and *bli-3(n529)*. While *bli-3(n529)* worms do exhibit severe dumpiness and some blistering, *bli-3(e767)* are mildly Dpy and lack any obvious Bli phenotype. Both these mutants have lesions in their peroxidase domain: *bli-3(e767)* has a G246D change while the lesion in *bli-3(n529)* is D393N. The severity of the RNAi phenotype is not unexpected for a number of reasons. Firstly, RNAi results in an effective “knock-out/knock-down” which is in contrast to the point mutations of the *bli-3* alleles. It has not yet been established how severely these mutations affect the activity of the peroxidase domain. No matter how drastically they do so, the NADPH-donating function of these mutants alleles is presumably still active. Given that in mammalian systems, this is considered to be the dominant role of these enzymes, it is not surprising that the peroxidase domain mutants exhibit weaker phenotypes. Of great importance to this hypothesis is the fact that BLI-3 is the only gp91*phox* homologue in the *C. elegans* genome. Thus the effect of ablating its H<sub>2</sub>O<sub>2</sub>-donating function could potentially affect all the peroxidases in the *C. elegans* genome by removing their sole H<sub>2</sub>O<sub>2</sub> source. The resulting cuticle defects in RNAi-treated worms may therefore be in part, or even wholly, due to the inactivation of downstream peroxidases, including any peroxidases that might be involved in the cross-linking of the cuticle, that rely on *bli-3* to provide their substrate. One such peroxidase whose activity might be affected by the abolition of an NADPH-oxidase, is HPX-1

Since the RNAi phenotypes of *bli-3* and *f53f12.3* are similar, another reason for the severity of the RNAi effect might be that the high homology between the two Duox enzymes results in the simultaneous down-regulation of both genes. In such a scenario, the severity of the targeted RNAi of *f53g12.3* RNAi would be due, not to its loss, but to the indirect targeting of *bli-3* and its H<sub>2</sub>O<sub>2</sub> donating function.

Further proof of the role of Duox enzymes in the cuticle is in the cyclical pattern of their expression, coinciding temporally with the larval moults and specifically with *bli-1*, *bli-2*, *sqt-1* and *rol-6* expression (Edens *et al.*, 2001). Lesions in *bli-1* and *bli-2*, as their names suggest, also result in cuticle blisters. Expression during the period of cuticle synthesis is thus evidence that Duox are vital catalysts of the *C. elegans* cuticle.

## 5.2 Results

### 5.2.1 Targeting of peroxidase containing genes by RNAi

The establishment of a peroxidase domain-containing enzyme, BLI-3, being involved in the formation of di-, tri- and iso-dityrosine cross-links in the *C. elegans* cuticle (Edens *et al.*, 2001), made other peroxidase-containing enzymes interesting targets for RNAi analysis and therefore, four other peroxidases were chosen to be targeted in the RNAi screen described in chapter 4. As stated previously, zk430.8, which we refer to as *hpx-1* (animal haem peroxidase), was chosen as a target for further investigation. Its RNAi phenotype, which is similar to that of the Duox enzymes, is described in following sections.

In order to analyse the role of HPX-1 in the cuticle, it was important to establish how many peroxidases are encoded by the *C. elegans* genome and whether any of these might have parallel roles to HPX-1 and BLI-3. Using the WormBase motif search engine, 13 animal haem peroxidase domain-containing genes have been identified (6 of which had already been analysed by RNAi in the initial screen) and are listed in table 5.1. This list again exemplifies the many peroxidases (table 5.1 A) yet only one functional NADPH-oxidase (BLI-3) (table 5.1 B) encoded by the *C. elegans* genome. As suggested in previous sections, this may imply that BLI-3 is the sole H<sub>2</sub>O<sub>2</sub> donor for all the encoded peroxidases. The animal haem peroxidase genes (apart from f32a5.2 and f49e12.1) were all targeted by RNAi in the TP12 strain, and the results are listed in the table 5.1. These experiments demonstrate that the Duox genes and *hpx-1* are unique in this class in giving a severe RNAi phenotype since all of the other genes have wild type RNAi phenotypes. However, these experiments were able to establish that the COL-19::GFP pattern of f09f3.5 is slightly amorphous. c16c8.2 RNAi-treated animals have an even less pronounced amorphous phenotype and this is not thought to be significant. The lack of RNAi phenotypes may be due to these genes not having a role in the cuticle, a level of redundancy between them, or because they are inefficiently targeted by the RNAi technique.

Peroxidases that do not fall into the animal haem peroxidase category also exist in the *C. elegans* genome; there are homologues of the plant/fungal/bacterial class of peroxidases (table 5.1) as well as acid phosphatase/vanadium-dependent haloperoxidases, and glutathione peroxidases. The plant/fungal/bacterial class of peroxidases could also require a H<sub>2</sub>O<sub>2</sub> source and may also possibly be Duox-dependent. Many of the plant/fungal/bacterial haem peroxidases appear to have functions distinct from their peroxidase activity, or at least, do not seem to have roles that

A

Animal haem peroxidase	RNAi phenotype (COL-19::GFP disruption in brackets)
[ <i>bli-3</i> (f56c11.1/ <i>duox 1</i> )]	Bli, Mit, Lvl (severe GFP)
[f53g12.3/ <i>duox2</i> ]	Bli, Mit, Lvl (severe GFP)
<i>hpx-1</i> (zk430.8)	Bli, Mit, Lvl (severe GFP)
r08f11.7	WT
t06d8.10	WT
f09f3.5	WT (slightly amorphous)
c16c8.2	WT
c46a5.4	WT
f32a5.2a/b	WT
pxn-1	WT
pxn-2	WT
f49e12.1	WT
k10b4.1	WT
<b>Haem peroxidase</b>	
<i>ced-7</i>	WT
<b>Plant/fungal/bacterial haem peroxidase</b>	
t22e5.6	NA
f47b8.10 orthologue to glucose phosphate transferase	Gro
t13b5.3	WT
t22d1.11 carboxyl esterase	WT
h04j21.1	WT
srd-71 serpentine receptor (chemoreceptor)	WT
y54e10br.6 RNA polymerase	Emb, Lvl,
egl-2 K <sup>+</sup> voltage gated channel	WT
sri-17 serpentine receptor	WT
sri-12 serpentine receptor	WT
c01b10.4 carboxyl esterase	WT

B

(functional) NADPH-oxidase/gp91phox homologues
[ <i>bli-3</i> ]

**Table 5.1: List of different classes of haem peroxidases.** A) A number of haem peroxidases are encoded by the *C. elegans* genome. The animal haem peroxidases were tested in this study for the effect of RNAi-treatment on the cuticle. None except for the Duoxs and *hpx-1*(zk430.8) show significant disruption. The Duoxs are in square brackets because of their second function as an NADPH-oxidase. The RNAi phenotype derived from the Kamath *et al.* 2003, screen are shown for haem peroxidases and plant/fungal/bacterial haem peroxidases (Kamath, 2003). B) There is only one NADPH-oxidase (*bli-3*) encoded by the *C. elegans* genome. This may represent the sole H<sub>2</sub>O<sub>2</sub>-donating moiety and may be required for the functioning of some, if not all the peroxidases.

would implicate them in cuticle synthesis. The Kameth *et al.* (2003) screen shows RNAi-treatment against these genes does not result in aberrant morphology. Three of these genes encode serpentine receptors that either function as chemoreceptors (*srd-71*) or as olfactory G-protein coupled receptors (*sri-12* and *sri-17*). Two peroxidase-containing putative proteins have been implicated as being carboxyl esterases (c01b10.4 and t22d1.11) and other proteins have been associated with being an RNA polymerase (y54e10br.6), an orthologue to glucose-6-phosphate transporter-1 (f47b8.10). On this basis, these peroxidases were not used as targets for our RNAi screen.

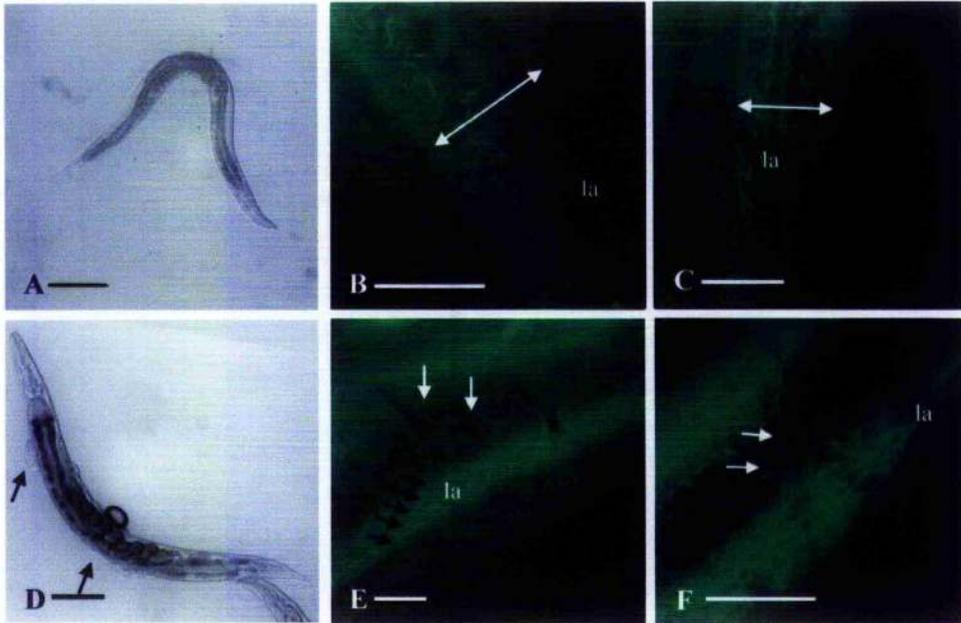
Thus, these data suggest that there are no other peroxidases with functions overlapping with those of BLI-3, f53g12.3 and HPX-1.

### 5.2.2 Analysis of *bli-3* mutant alleles using TP12

The *bli-3* alleles *bli-3(e767)* and *bli-3(n529)*, which are medium Dpy-non-Bli and severe Dpy-Bli respectively (figures 5.2 A and D), were crossed with the COL-19::GFP marker strain, TP12, in order to visualise the pattern of this adult-specific collagen in the mutant cuticles. The disruption observed is depicted in figure 5.2. *bli-3(e767)* worms, containing a G246D lesion within the peroxidase domain, show a disrupted and branched COL-19::GFP pattern overlying the seam cell-derived cuticle and a wild type annular pattern in the dorsoventral cuticle (class III COL-19::GFP disruption according to the classes described in chapter 4) (figures 5.2 B and C). Such a pattern is similar to that of the other Bli loci, *bli-1* and *bli-5* (figure 4.10) and is indicative that the lesion in *bli-3* allows the relatively wild type construction of the annular furrows. On the other hand, *bli-3(n529)* worms, which have a D393N lesion in their peroxidase domain, exhibit an amorphous pattern of annulae, multiple alae and weaker intensity of the COL-19::GFP marker (figures 5.2 E and F), illustrating that this allele severely disrupts the cuticle and prevents the proper formation of the annulae. It is interesting to note that the *bli-3* locus is unlike any that have been previously tested using TP12 due to the fact that its alleles exhibit different classes of COL-19::GFP disruption. However, other Bli loci have not been tested using multiple alleles.

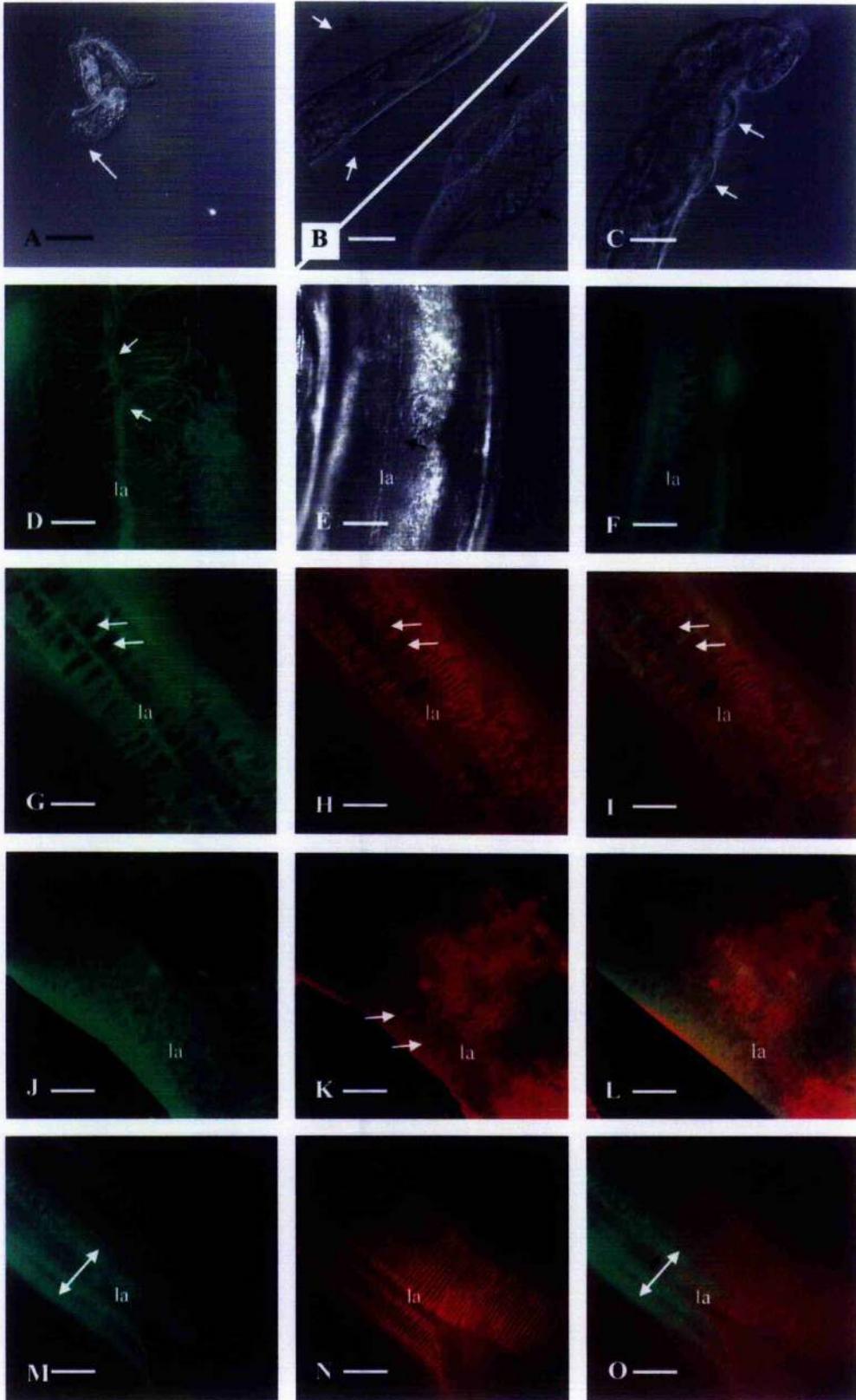
### 5.2.3 *bli-3* RNAi using TP12

RNAi treatment of *bli-3* in a TP12 background, using the feeding clones from Kameth *et al.* (2003), result in a variety of phenotypes ranging from extremely malformed and extensively blistered (figure 5.3 A) to relatively wild type and a similar range of COL-19::GFP disruption. The range in phenotypes is likely derived from the variable penetrance of the RNAi technique.



**Figure 5.2: Gross morphological and COL-19::GFP phenotypes of the *bli-3* alleles *bli-3(e767)* and *bli-3(n529)*.** A) *bli-3(e767)* animals are mildly Dpy and do not exhibit blisters. B and C) At the cuticle level, they exhibit a pattern of COL-19::GFP disruption that resembles that of *dpy-5(e61)*. An uncontracted region overlying the seam cell-derived cuticle is observed to have fragmented and branched annulae (white double headed arrow). The region overlying the dorsoventral hypodermis is wild type (black double headed arrow). The fragmentation in figure B is more severe than in figure C. D) *bli-3(n529)* animals are more Dpy than *bli-3(e767)* and exhibit fluid-filled blisters which are marked by arrows. E) At the COL-19::GFP level, annulae appear amorphous (denoted with white arrows). This animal also exhibits multiple alae (5 instead of 3)(denoted with black arrows). F) In addition to the amorphous phenotype, this strain also exhibits branching annulae (white arrows). In all figures, lateral alae are denoted with 'la'. Scale bars = 100 $\mu$ m in A and D. 10 $\mu$ m in B, C, E and F.

**Figure 5.3: Gross morphological and cuticle phenotypes resulting from targeted RNAi of *bli-3*.** A-C) Nomarski images of gross morphological phenotypes: A) severe body shape defects and blisters. Arrow denotes a large blister in the head region. B) Blisters in the head region are either filled with granular material or are clear and fluid filled. C) The head regions appear lumpy (arrows). D) The most severe form of COL-19::GFP disruption shows thread-like alae that branch into the region normally associated with the presence of annulae (white arrows). Annulae themselves are few in number and are severely fragmented (black arrows). E) RNAi-treated animals also display aberrant seam cells. Enlarged cells can be seen to coincide with the normal position of the alae (arrow). Seam cells in wild type animals are fused at the adult stage. F) The region of aberrant seam cells coincides with an amorphous pattern of COL-19::GFP. G – O) Differing cuticle disruptions observed in RNAi-treated animals viewed by COL-19::GFP and DPY-7 localisation (and merged images). G) An amorphous pattern is the extreme form of the finger like projections (previously observed in *dpy-7(e88)* mutants). This results in regions of the cuticle being devoid of COL-19::GFP staining (arrows). H) The corresponding DPY-7 pattern shows the annulae to be severely truncated (single headed arrows) and to be completely absent in the region surrounding the seam cells (double headed arrow). I) Merged image shows the COL-19::GFP and DPY-7 patterns coincide. J) An amorphous and ubiquitous COL-19::GFP expression similar to *dpy-11(e1180)*. K) The corresponding DPY-7 pattern is fragmented (arrows). L) merged image. M) Example of type III COL-19::GFP pattern exhibited by RNAi-treated animals. The uncontracted region overlying the alae is denoted with a white double headed arrow. N) Interestingly, the annulae, as viewed by DPY-7, appear to appose the alae (black double headed arrow). This is in contrast to the pattern of COL-19::GFP. O) Merged images. In figures D-O, lateral alae are denoted with 'la'. Scale bars = 100  $\mu$ m in A-C and 10 $\mu$ m in D-O.



Figures 5.3 B and C illustrate other interesting phenotypes at the morphological level. The blistered worms shown in figure 5.3 B illustrate that blisters can be filled with material or can be clear and fluid-filled. Figure 5.3 C shows the severe “lumpy” body morphology of other RNAi-treated worms. There is also a level of embryonic lethality as only 78% and 76% of F1 embryos develop into larvae upon targeting by *f53g12.3* and *bli-3* RNAi respectively. As expected from the findings by Eden *et al.* (2001), RNAi targeting of either gene results in similar phenotypes (also illustrated by the same degree of embryonic lethality). In the most severely morphologically mutant “grubs” which lack any kind of body definition and which are highly blistered (figure 5.3 A), it is not possible to visualise COL-19::GFP firstly, because of the limitations of the method of visualisation (e.g. worms burst on the slide) and secondly, since the reduction in COL-19::GFP probably represents a cuticle so disrupted that the collagen content is extremely low/altered. In blistered worms a unique pattern of COL-19:GFP is observed. The most severe form of COL-19::GFP in the RNAi-treated worms appears as a number of long intertwined strands that run parallel with the normal positioning of the alae (figure 5.3 D). These branch out to cover the dorsoventral hypodermal derived cuticle. The pattern is not at all reminiscent of annular patterning. In addition to these alae defects, these worms also exhibit a fragmented and “wispy” pattern in a region where the annulae normally reside (figure 5.3 D).

Other phenotypes observed after RNAi treatment include those illustrated in figures 5.3 E and F where an amorphous pattern of COL-19::GFP is observed localising above apparently aberrant seam cells (the seam cells appear swollen (see arrow)). Differing levels of disruption are observed, and are exhibited in figures 5.3 G and J. Figure 5.3 G shows COL-19::GFP in long and defined finger-like projections, while a more ubiquitous amorphous pattern is illustrated in figure 5.3 J. Figure 5.3 M also demonstrates that COL-19::GFP staining patterns, reminiscent of that seen in DPY-5 mutants, are observed.

#### 5.2.4 DPY-7 staining of *bli-3* RNAi-treated worms

An immuno-cytochemical approach was used to observe the pattern of the cuticle collagen, DPY-7, in the cuticle of the *bli-3* RNAi-treated animals. As expected, a range of phenotypes arise; however, these do not vary as much as the COL-19::GFP phenotypes. In general, the DPY-7 localises in a pattern corresponding to the disrupted annular furrows. A severe form of annular disruption is shown in figure 5.3 K where the pattern of DPY-7-stained annulae appears as short fragments indicating that these structures are extremely interrupted. Figure 5.3 H

illustrates similar fragmented annulae but these are additionally prematurely terminated. These data suggest that annulae, although being fragmented, are present, and it is therefore extremely interesting to note that the corresponding COL-19::GFP pattern is an amorphous one.

Another observation that seems to contradict previous conclusions is illustrated by figures 5.3 M-O. The COL-19::GFP pattern in the depicted animal is similar to that of *dpy-5* mutants, displaying a region of relatively wild type expression over the dorsoventral hypodermis, and being disrupted around the seam cells. In *bli-3* (RNAi) animals, similar to *dpy-5* mutants, wild type DPY-7-stained annulae are present. However, there is a crucial difference because in *dpy-5* mutants, the annulae are prematurely terminated as a result of the seam cell cuticle being uncontracted. In *dpy-5* mutants, therefore, the COL-19::GFP disruption occurs in the uncontracted seam cell-derived cuticle. In *bli-3*(RNAi) animals, however, annulae appose the alae (figure 5.3 N), implying that the seam cell-derived cuticle is fully contracted and that the aberrant localisation of the marker collagen occurs in the dorsoventral cuticle. This suggests that there are two distinct staining patterns in the cuticle overlying the dorsoventral hypodermis, which is a novel observation.

### 5.2.5 Isolation of *bli-3* deletion mutant

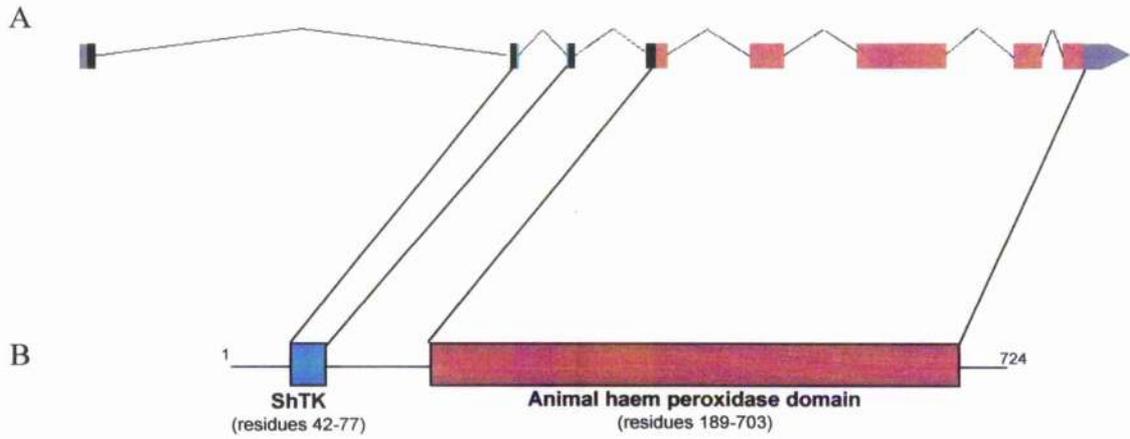
The RNAi-effect is transient and possibly only partially penetrant and therefore attempts were made to isolate a *bli-3* deletion mutant from a mutant library. A phenotype-independent mutational screen using a PCR approach was carried out using the protocol described in (Barstead, 1999). These attempts were unsuccessful.

### 5.2.6 Sequence analysis of *hpx-1*/HPX-1

The gene structure of *hpx-1*, located on chromosome II, is illustrated in figure 5.4 A, which shows the 8 predicted exons. The long first intron may encode for regulatory factors that control its expression. The genomic DNA is 11891 bp while the cDNA is 2175 bp.

The gene encodes for a mature 724 residue peptide containing two distinct motifs: a 536 residue peroxidase domain that makes up most of the 724 residue protein, as well as an ShTK domain (figure 5.4 B). It does, however, lack the NADPH-oxidase domain that is encoded by the *Duox* genes. From its length, it is predicted to encode a protein of approximately 80 kDa.

The peroxidase domains of: HPX-1; human and mouse MPO, OPO, TPO and LPO; sea urchin



**Figure 5.4: Gene and protein structure of *hpx-1*/HPX-1.** A: gene structure of *hpx-1* (gDNA=11,891 bp, cDNA=2,175 bp) and B: the ShTK and animal haem peroxidase domains encoded in the 724 residue protein.

OVO; as well as *C. elegans* homologues of TPO and OVO, were aligned along their peroxidase domains using a ClustalW tool (figure 5.5). The multiple alignment reveals that, unlike Duox, HPX-1 does indeed possess all the conserved catalytic and  $\text{Ca}^{2+}$  binding residues characteristic of the other animal haem peroxidases, such as the active site distal and proximal histidines and the distal arginine (figure 5.5: red lettering). The  $\text{Ca}^{2+}$  binding region is also highly conserved and HPX shares 90% similarity with OVO in this region (red lettering). In addition, within the aligned region, HPX-1 has all 13 of the cysteines that are highly conserved among the animal haem peroxidase family; however, in some cases, the position of these are slightly shifted (figure 5.5: green highlight). In MPO, these cysteine residues have been demonstrated to participate in the formation of six inter-chain disulphide bond as well as in dimerisation (Nomura *et al.*, 1999). This suggests that HPX-1 has the potential to fold properly and also to dimerise. This observation coupled with the conservation of important catalytic residues further strengthens the argument that this enzyme is indeed an active peroxidase. These important residues aside, analysis of the peroxidase domain as a whole using as a ClustalW tool shows that it shares between 24% and 47% identity (47%-67% similarity) with other animal haem peroxidases, depending on which peroxidase. Its highest homology is with human TPO. This might be relevant since, during thyroid hormone-synthesis, TPO uses protein-bound tyrosine as a co-substrate. This tyrosine-related role is the main function of TPO, unlike the other peroxidases whose predominant role is in the oxidation of halide ions, and consequently, its sequence probably facilitates the efficient targeting of tyrosine. For example, it may optimise the access of protein-bound tyrosine to its active site (unlike MPO). This is relevant for the analysis of HPX-1 since if this putative peroxidase is indeed involved with the catalysis of tyrosine cross-links, such access would be vital.

In alignments, HPX-1 differs from all other aligned peroxidase domains at one residue. It is the only peroxidase not to encode a proline at a site 15 residues away from the proximal histidine; an alanine is encoded instead. The significance of this change is not known. A final observation from the alignment is that there are two residues that are conserved in nematode peroxidases and ovoperoxidase only (figure 5.5: dark yellow highlight). This might suggest certain species-specific specialisations exist.

A predict-protein tool was used to identify regions of secondary structure and, as would be expected from the high level of homology between the peroxidases, the predicted secondary structure coincides well with the other peroxidases (not shown).

**Figure 5.5: ClustalW alignment of the animal haem peroxidase domains of human (h), mouse (m), *C. elegans* (C. e) and sea urchin animal haem peroxidases.** HPX-1 is in blue text. Residues conserved in all peroxidases are highlighted in black. Residues conserved in >7/13 of the peroxidases are highlighted in dark grey. Residues highlighted in light grey represent conservation of residues with similar properties (properties defined according to EMBL ClustalW). Cysteines are highlighted in green. The residues implicated in catalytic activity are in red text and are denoted with a star. These include the active site distal His, distal Arg, and proximal His and the residues of the Ca<sup>2+</sup> binding site. BLI-3 is the most variant at sequence level. The two residues that are changed in the morphological mutant BLI-3 alleles are denoted with a black arrow. It is interesting to note the Gly residue that is mutated in *bli-3(e767)* (also highlighted with a red ring): In all other peroxidases, a His is found at this position. Residues that are specific to all *C. elegans* and sea urchin peroxidases are highlighted in dark yellow. HPX-1 has one residue that varies from all other peroxidases: This is highlighted in red.



In addition to its peroxidase domain, this predicted 80 kDa protein possesses an ShTK domain. The presence of this short domain, which has also been termed SXC on the basis of the presence of six, probably di-sulphide bonded, conserved cysteine residues, is interesting as these domains 1) have previously been implicated in extracellular matrix-specific protein-protein interactions, 2) are associated with proteins containing tyrosinase-, astacin-, CUB-, matrixin-, peroxidase-, zinc carboxypeptidase- and serine protease- domains, and 3) are found in proven cuticle-modifying enzymes such as mucins and some MMPs (Blaxter, 1998).

Unlike TPO and the Duox enzymes, Kyte and Doolittle hydropathy plot calculations (ProtScale) reveal that HPX-1 does not appear to be membrane-bound (not shown).

### 5.2.7 RNAi of *hpx-1*

An RNAi-feeding construct was cloned that incorporated 802 bp of genomic DNA of *hpx-1*. Additionally, the HPX-1 construct supplied in Julie Arhinger's RNAi library was used. Both these constructs target exon 5 of the *hpx-1* gene. In order to ensure that no non-specific RNAi was taking place, a BlastN search was carried out against the Wormbase EST database. The gene showing highest homology shares 66% identity. This is below the standard 85% homology cut-off under which the RNAi effect is believed to be specific.

Both RNAi-feeding clones were used in experiments using the COL-19::GFP strain TP12 and both were sufficient to produce the severe Bli phenotype that was expected, according to the Kamath *et al.* (2003) observations. It is possible to induce the plated L4 worms to develop into phenotypically blistered adults and on average, about half of the L4 worms themselves go on to exhibit blisters after two days of exposure to the RNAi (at both 20°C and 25°C). This shows the potency of the RNAi and shows that the phenotype is not only caused by maternal inheritance. The F1 generation was analysed for embryonic/larval lethality and gross morphological defects. The results below are based on two separate RNAi experiments in order to account for varied RNAi feeding conditions. Results from the individual experiments varied from each other between 2 % and 10 % but were similar in terms of the proportions of phenotypes. The results below were calculated by combining the numbers of worms from both experiments.

There is no significant embryonic lethality upon RNAi-treatment (based on 1350 and 2280 embryos respectively). However, there is a level of larval lethality since only ~45% (at 20°C)

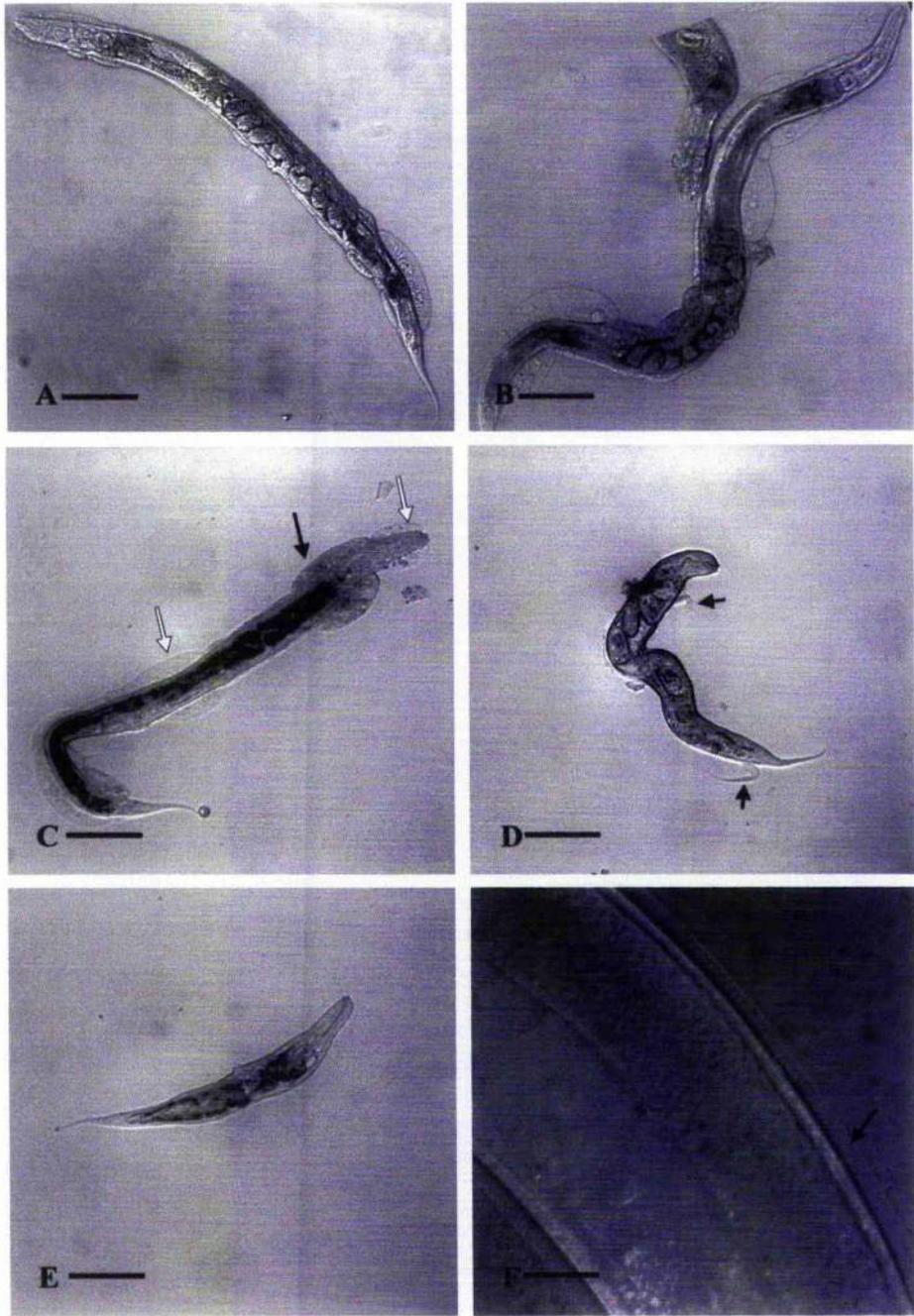
and ~47% (at 25°C) of hatched larvae develop into adults (based on 1445 and 2294 larvae). On the gross morphological level, 17% and 32% of adult-stage RNAi-treated worms show any kind of defect at 20°C and 25°C respectively. The gross morphological phenotypes are most obvious in adults and the majority of larvae appear relatively wild type at this level; however, it was often difficult to tell the age of the worm due to the cuticular disruption. A contributing factor to the lack of phenotype in larvae, is the fact that the easily distinguishable Bli phenotype is only expressed in adult cuticles. Because of these facts and the adult-specificity of the COL-19::GFP marker with which we are concerned, the results below are derived from the scoring of the adult population only.

The resulting phenotypes are illustrated in figure 5.6. These include mild Bli (exhibiting three or fewer small blisters) (figure 5.6 A), severely Bli (showing large or >3 blisters) (figures 5.6 B and C), “dissolved”/grub-like (when the cuticle appears weak, there is a lack of clearly defined body pattern and the worms appear opaque) (figures 5.6 D and E) and Dpy (this includes worms whose dumpiness is restricted to only portions of the body length) (also in figure 5.6 E). At 20°C, the major phenotype is mild Bli (~8% of adult worms) (figure 5.6 A). In addition ~2% of the worms reaching adulthood (661 worms) exhibit severe blistering while ~5% appear dissolved and <2% are Dpy. At 25°C, about half of the morphologically mutant worms show the “dissolved” appearance, thus making it the most dominant morphological defect: ~6% of all adults are mild Bli, ~7% are severely Bli, and <1% are Dpy. The resulting RNAi phenotypes are thus temperature-dependent, with 25°C being sufficient to cause the severe degradation of the entire cuticle as the predominant phenotype. Figure 5.6 F also shows that even in the absence of an obvious Bli phenotype, there may be some degree of separation of the cuticle layers.

In accordance with the varied gross morphological phenotypes, the COL-19::GFP pattern of RNAi-treated worms ranges from being mildly disrupted to severe and amorphous (figure 5.7). In many worms, only very faint patterns of fluorescence are observed, suggesting that collagen might not be being fully incorporated into the cuticle. In the mildest affected cuticle (not pictured) the annular pattern of wild type COL-19::GFP is still apparent, albeit being slightly amorphous. Higher degrees of amorphous COL-19::GFP staining are shown in figures 5.7 and it is in worms exhibiting these COL-19::GFP patterns, that the intensity of the marker is reduced. Differing levels of the amorphous pattern can be seen to range from the uniform pattern pictured in 5.7 A to the pattern of finger-like projections (figure 5.7 B) that is reminiscent of *dpy-7(e88)* mutants. Some of the animals displaying the amorphous phenotype also exhibit an interesting

**Figure 5.6: Examples of gross morphological phenotypes observed in the F1 generation of animals fed with *hpx-1* RNAi** A) Blisters can be relatively small and limited in number or, as depicted in B and C, can be larger and more numerous. In C, examples of both fluid- and granule- filled blisters are observed. More severe phenotypes are depicted in D and E which are termed 'dissolved'. Animals are very small even though being adults. D also shows traces of unshed cuticle from previous moults (black arrows). E shows a slightly Dpy animal. Scales in A-E are similar for comparison of relative sizes. F) At higher magnification, the cuticle of a non-Bli animal can be seen to be slightly separated. This implies that cuticle defects are present even in the absence of discernable Bli phenotypes. Scale bars in A-E = 100  $\mu$ m and in F = 10 $\mu$ m.

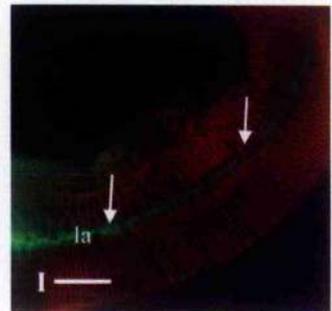
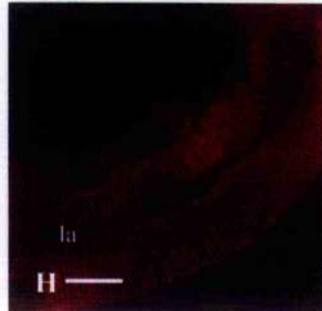
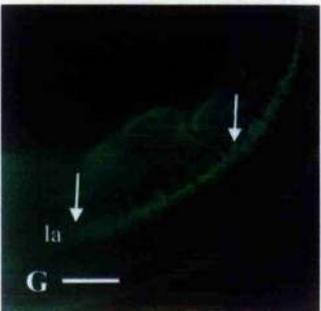
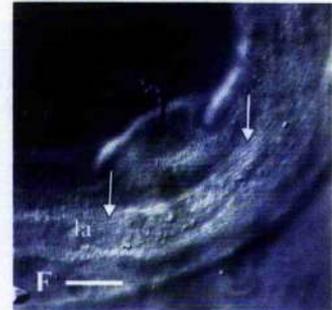
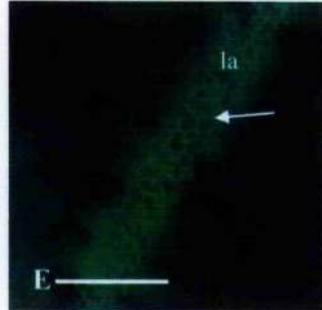
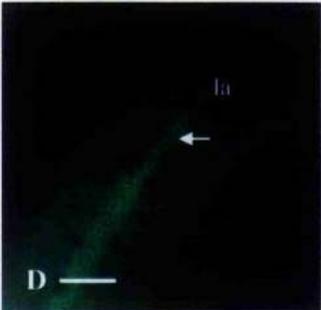
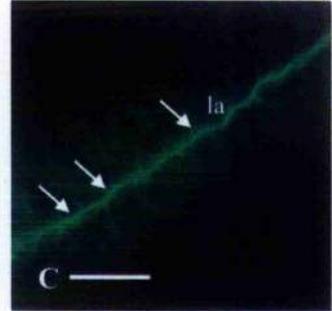
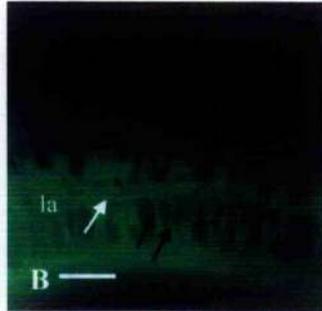
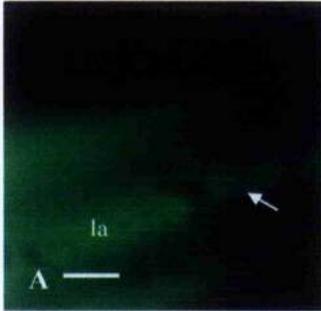
Facing page 250



**G**

	Bli	Severe Bli	Dissolved	Dpy
20°C	8	2	5	<2
25°C	6	7	50	1

**Figure 5.7: Differing levels of cuticular disruption in animals treated with *hpx-1* (RNAi).** A-E are COL-19::GFP images. A) shows the amorphous phenotype which appears mesh like. Alae are also amorphous and are denoted with a white arrow. B) An amorphous pattern (similar to that seen in *dpy-7(e88)* mutants) is also observed with regions of the cuticle being devoid of COL-19::GFP (black arrows). Again, amorphous alae are denoted with a white arrow. C) A similar but mild form of that seen in *bli-3* RNAi. The alae appear like intertwined threads and are often (severely) branched (white arrows). The COL-19::GFP content in the rest of the cuticle is severely reduced. D) A novel pattern of COL-19::GFP overlying the alae is observed as well as cuticle that is amorphous. E) shows an enlargement of D and shows that the pattern observed is most likely derived from the presence of multiple alae that are branched. F-I) are corresponding images viewed with Nomarski (F), COL-19::GFP (G), DPY-7 immunolocalisation (H), and merged COL-19::GFP/DPY-7 (I). In F) the broken and fragmented nature of the alae are depicted with white arrows. Annulae, whilst appearing almost wild type in some regions of the cuticle, are in some places fragmented (black arrow). H also shows the fragmented nature of these annulae. The DPY-7 pattern and the COL-19::GFP patterns do not correspond because the former is mostly wild type and annular while the latter is amorphous. The two images are superimposed in figure I. In all images, lateral alae are indicated with 'la'. Scale bars = 10um.



pattern overlying the alae, which is illustrated in figures 5.7 D and E. What appears as a net-like pattern overlying the alae is most likely caused by the presence of multiple alae, which are branched. In some cases, alae do not appear multiple but are severely branched (figure 5.7 C) and are reminiscent of the twisted alae seen in *bli-3*(RNAi) worms (compare figures 5.7 C and 5.3D). The phenotype is obviously a milder form of the *bli-3*-derived disruption, but this dominant alae expression, compared to annular expression, is quite a unique feature of these peroxidases. The fact that there is similarity between *bli-3* and *hpx-1* RNAi is further evidence that cross-linking activity is exhibited by HPX-1. The increased severity of the RNAi phenotype of *bli-3* is interesting in itself because it could indicate that multiple peroxidases, instead of just HPX-1, are being down regulated as a consequence of the loss of the BLI-3 NADPH-oxidase.

### 5.2.8 DPY-7 staining of *hpx-1* RNAi-treated worms

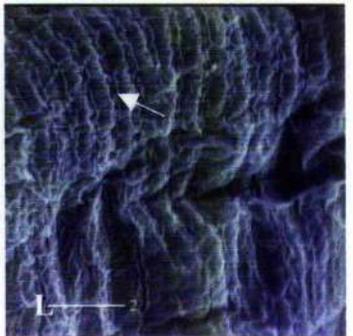
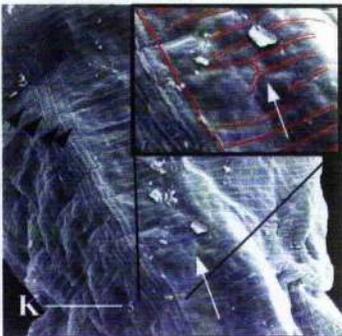
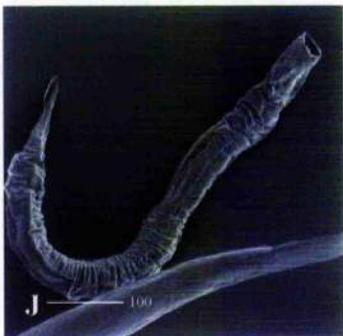
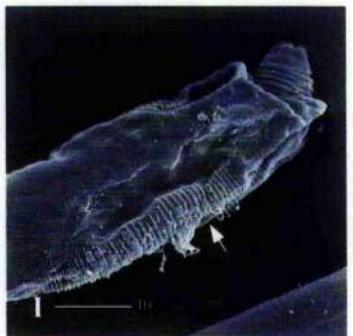
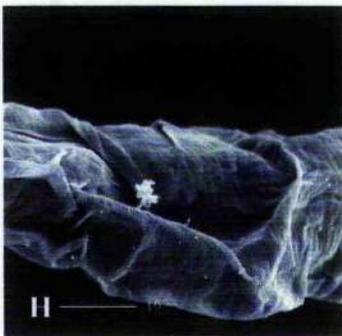
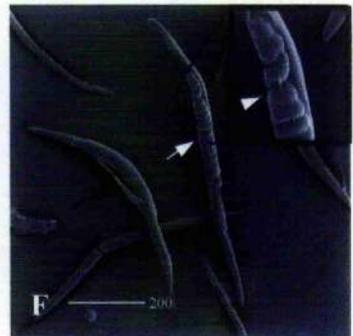
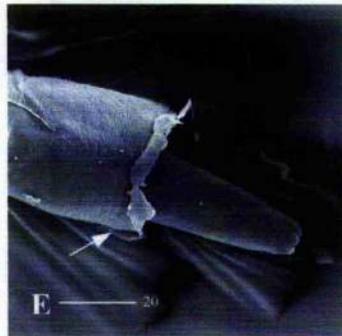
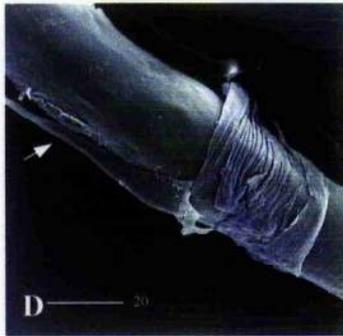
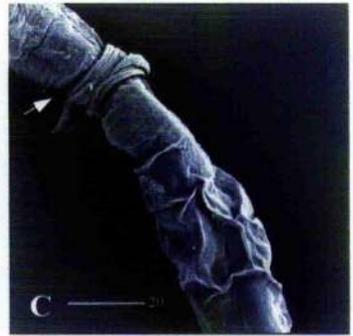
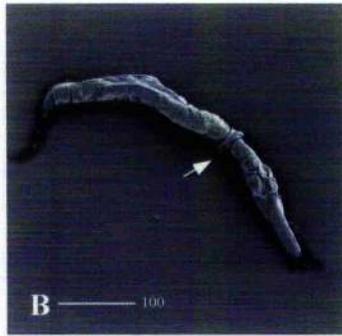
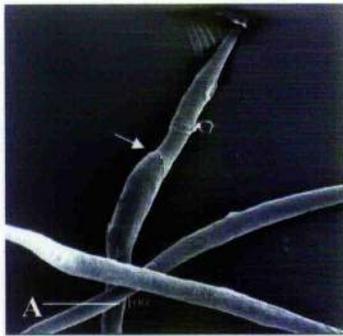
When anti-DPY-7 antibodies are used to stain the cuticle of RNAi-treated worms, a fragmented and truncated pattern of annulae becomes evident (figure 5.7 H). Compared to the COL-19::GFP disruption (figure 5.7 G) of the same worm, however, this pattern is relatively wild type. This is similar to as described for the *bli-3* RNAi-treated animals. In both circumstances, animals exhibit an amorphous COL-19::GFP pattern in the same region as relatively wild type annular furrows (figure 5.7 I exhibits this in *hpx-1*(RNAi) animals). In some animals with more severe gross morphological defects, DPY-7 staining is severely reduced. These latter animals probably exhibit a more global disruption of cuticle collagens with the consequence that DPY-7 and other collagens are not incorporated properly. Alternatively, more severely affected animals and their severely weakened cuticles may not have been able to withstand the preparative techniques required for immunolocalisation.

### 5.2.9 SEM of *hpx-1* RNAi-treated animals

Scanning electron microscopy was performed on worms following *hpx-1* RNAi. Figure 5.8 shows the cuticular disruptions observed under high magnification that vary from mild to severe and that are apparent in all larval stages, albeit at a lesser extent in early larval stages. These larval defects have not previously been described.

The majority of the worms show some form of morphological disruption at this magnification. Most of the worms appear to have moulting defects in which the old cuticle is not properly shed (the variety of such moulting phenotypes is illustrated in figures 5.8 A-F). The constrictions around the centre of the worm (figure 5.8 B, C and D), as well as the attachment around the head

**Figure 5.8:** SEM images of *hpx-1* RNAi-treated worms. A-F) Animals exhibit moulting defects: A) Mild moulting defects consist of the old cuticle not being shed, which restricts the growth. This can be seen from the bulge of the growing animal where the old cuticle has loosened (white arrow). B-F) Severe moulting defects: the old cuticle is not shed and constricts the centre of the worm (arrows in B and C) and it is here that the tail of the old cuticle is often still attached (D: arrow). E) The old cuticle is also seen around the head region (arrow) where during normal ecdysis, the worm breaks the old cuticle and crawls out of the old cuticle. F) Since the old cuticle is not shed, embryos become trapped between the two cuticles of the hermaphrodite (white arrow and enlarged inset). Also shown is another worm with the old cuticle tail still attached (black arrow). G-J: The cuticle is very baggy and probably corresponds to the blisters. Blisters are usually observed round the head region and different levels of bagginess are observed (G, I & J). Figure I also shows build up of bacteria on the annulae (arrow). K) Multiple alae are observed (four instead of three) (black arrows). Annulae do not appose the alae and are branched (white arrow) (inset: the annulae and one of the alae have been coloured red for clarity. The white arrow points to the annulae. L) an example of more severely branched annulae.



region (figure 5.8 E) are typical indications of such moulting defects, the latter defect being the failure of the worm to successfully complete ecdysis and loosen itself from the old cuticle. Mild moulting defects consist of the failure of the old cuticle to be shed, which consequently restricts growth. This can be seen from the bulge of the growing animal where the old cuticle has loosened (figure 5.8 A). Yet another observation that implies a moulting defect is that embryos are observed to be trapped between the old and the unshed cuticle of the hermaphrodite (figure 5.8 F).

The cuticle in some worms appears to be extremely baggy and presumably corresponds to areas of blistering (figures 5.8 G-K). Blisters are most commonly found around the head region and this corresponds with the areas where much of the baggy cuticle phenotype is observed. Figure 5.8 L shows that at x 10,000 magnification, the composition of the cuticle is aberrant in that there are multiple alae and the annular furrows are less pronounced and are truncated. This could indicate that the patterns of COL-19::GFP correspond with actual discernible features on the surface of the cuticle. Another observation from the SEM photographs is that bacteria build-ups are observed around the annulae (figure 5.8 D), which could be an indication that particular components of the cuticle, to which the bacteria adhere, become exposed on the surface of the worm due to the aberration of the cuticle. In summary, SEM shows the clear moulting and growth defects exhibited by RNAi-treated animals. In addition, the baggy appearance of the cuticle and the observation of aberrant annulae suggest a weakened, and possibly weakly cross-linked cuticle.

#### 5.2.10 Isolation of *hpx-1* deletion mutant

In the same manner as described in section 5.2.5, it was attempted to isolate an *hpx-1* deletion mutant from populations of animals treated with chemical mutagens. These attempts, however, were unsuccessful.

#### 5.2.11 Double targeting of *bli-3* and *hpx-1*

In order to investigate the functional relationship between BLI-3 and HPX-1, *hpx-1* was targeted by RNAi in two morphologically mutant *bli-3* strains: *bli-3(e767)* and *bli-3(n529)*.

Since *bli-3* alleles exhibit lesions in their peroxidase domain rather than their NADPH-oxidase domain, these experiments enable the examination of the relative roles of BLI-3's and HPX-1's peroxidase domains.

#### 5.2.11.1 RNAi-targeting of *hpx-1* in the strain *bli-3(e767)*.

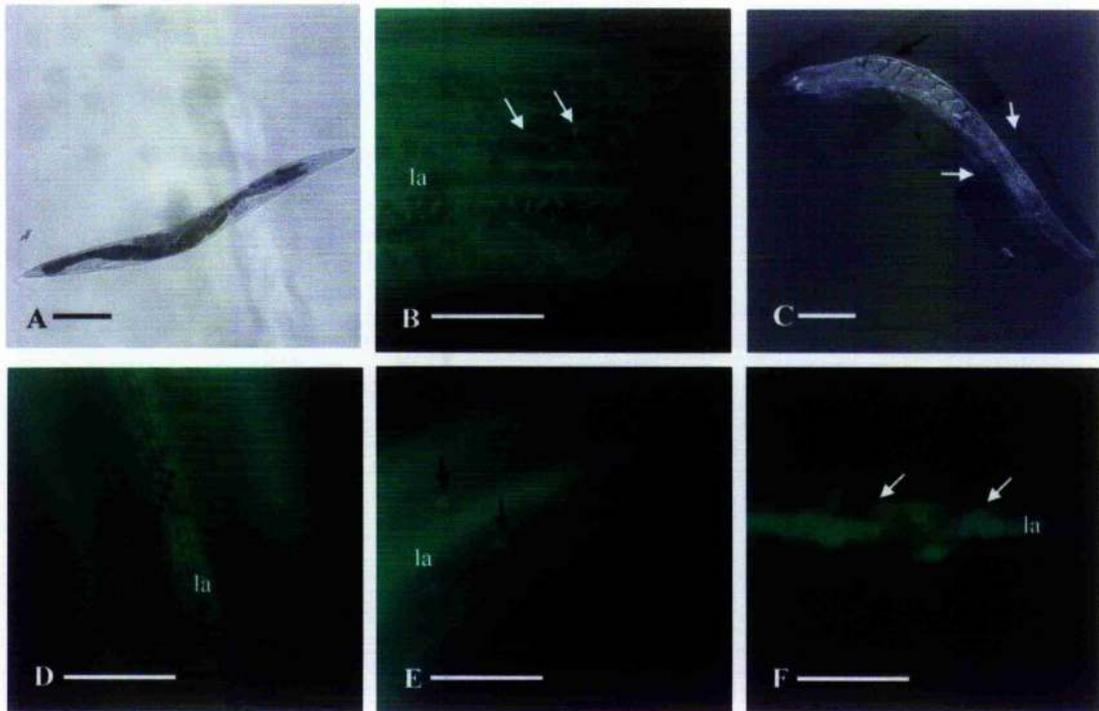
Figure 5.9 shows the gross morphological and COL-19::GFP phenotypes exhibited by *bli-3(e767)*(COL-19::GFP) animals subjected to RNAi feeding of *hpx-1*. The gross morphological phenotypes observed from such experiments are similar to those resulting from RNAi on wild type (TP12) animals: Severe Bli, mild Bli, dissolved and Dpy animals are observed (figure 5.9 G). These data are based on scores from separate RNAi experiments (1083 and 1007 F1 adult animals at 20°C and 25°C respectively). Based on the scored phenotypes, it appears that targeting of *hpx-1* in *bli-3(e767)* animals is fractionally less potent since the percentages of animals showing phenotypes are lower. At both temperatures less larval lethality is displayed by *bli-3(e767)* animals than by wild type (TP12) animals. A slightly higher percentage of animals reaching the adult stage display mutant phenotypes at 20°C compared to wild types. However, this is likely to be a consequence of the reduced larval lethality. At 25°C, equivalent percentages of adult animals display phenotypes compared to wild type. Also of note is the fact that at 20°C and 25°C, the percentages of treated *bli-3* animals showing phenotypes are similar, indicating that these RNAi feeding assays are not temperature-sensitive.

At the COL-19::GFP level, RNAi against *hpx-1* in a *bli-3(767)* background results in an amorphous annular pattern (figures 5.9 D and F) as well as aberrations in the alae which are amorphous, branched (figure 5.9 F) and multiple (5 instead of 3) (figure 5.9 D). Another observation is the accumulation of the GFP-tagged collagen in organelles, which implies that the incorporation of COL-19::GFP into the cuticle is being partially prevented (figure 5.9 E).

These results show that in the assumed absence of *hpx-1*, the *bli-3(e767)* peroxidase domain displays activity.

#### 5.2.11.2 RNAi targeting of *hpx-1* in *bli-3(n529)*

A very different scenario than that observed in *bli-3(e767)* animals is exhibited by *bli-3(n529)* animals. In the absence of RNAi-treatment, as discussed in section 5.2.2 this strain is severely Dpy and medium Bli at the gross morphological level (figure 5.10 A), and compared to RNAi of *bli-3*, shows moderately severe COL-19::GFP disruption (figures 5.10 B and C). However, upon RNAi-treatment to target *hpx-1*, 100% lethality of the P0 L4 animals is observed. One day after being placed onto plates, animals become completely dissolved and expire.



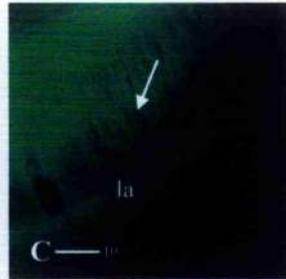
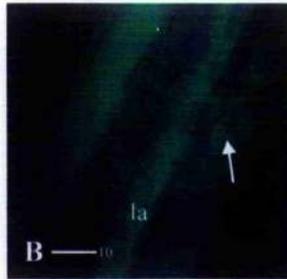
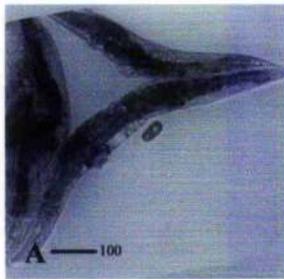
**G**

	20°C	Eml	Lvl	sev Bli	weak Bli	dissolved	Dpy
Wild type		7	55	2	8	5	2
<i>bli-3(e767)</i>		8	40	6	10	11	<1
	25°C	Eml	Lvl	sev Bli	weak Bli	dissolved	Dpy
Wild type		0	53	6	7	18	<1
<i>bli-3(e767)</i>		9	47	6	7	14	<1

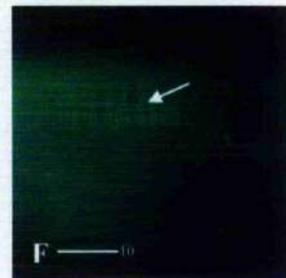
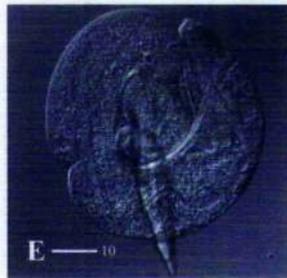
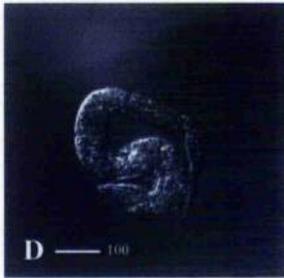
**Figure 5.9: Gross morphological and COL-19::GFP phenotypes resulting from targeted RNAi of *hpx-1* in the strain *bli-3(e767)*(COL-19::GFP) compared to untreated animals.** A-B) Untreated *bli-3(e767)* animals are slightly Dpy (A) and have branched annulae (arrows) (B). C-F) Gross morphological and COL-19::GFP phenotypes displayed by F1 animals from *hpx-1* RNAi experiments incubated at 25°. C) RNAi-treated animals display severe blisters (white arrows) and slightly lumpy/Dpy body morphology (black arrow). COL-19::GFP phenotypes include: D) amorphous annulae and multiple alae (5 instead of 3) (black arrows); E) accumulation of the tagged collagen in organelles (black arrows); and F) branching and amorphous alae (white arrows). Alae are marked 'la'. Scale bar: in in A = 100 μm; in C= 30μm; in B, D, E and F = 10μm. Figure G is a table showing the percentage of animals (surviving to adult) that display each phenotype when TP12 and *bli-3(e767)* strains are subjected to feeding of *hpx-1* at 20°C and 25°C. Emb and Lvl are embryonic- and larval- lethal respectively. Sev/weak Bli = severe/weak Bli. Dpy = dumpy.

**Figure 5.10: Gross morphological and COL-19::GFP phenotypes resulting from targeted RNAi of *hpx-1* in the strain *bli-3(n529)(COL-19::GFP)* animals (compared to untreated animals).**

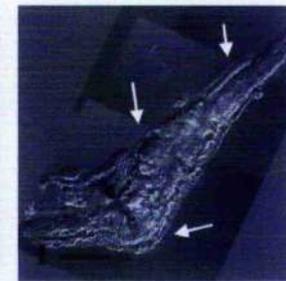
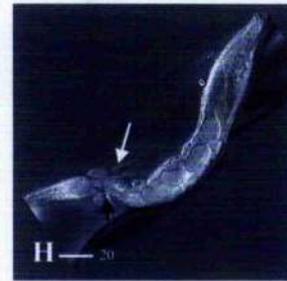
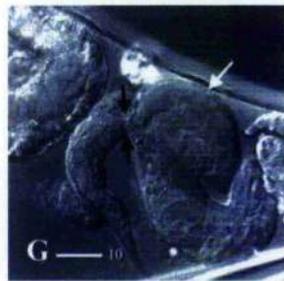
A-C) As shown previously in figure 5.2, untreated *bli-3(n529)* animals are Dpy-Bli (A) and display amorphous annulae (white arrows in figures B and C). D) Example of phenotypes displayed by P0 *bli-3(n529)* animals that were placed on feeding plates as embryos at 15°C. E and F) Gross morphological and COL-19::GFP phenotypes displayed by P0 *bli-3(n529)* animals that were placed on feeding plates as L2s at 15°C. Branching and truncated annulae are denoted with an arrow. The rest of the cuticle is devoid of COL-19::GFP staining. G-I) Severe gross morphological defects displayed by P0 *bli-3(n529)* animals that were placed on feeding plates as embryos and incubated at 25°. Figures show severe granule-filled blisters (G), unshed cuticle attached at the head region due to a moulting defect (arrow in figure H) and severe lumpy body morphology (I). Constrictions around the body circumference are shown by the animals in G, H and I and are denoted with black arrows. Such treatment is 100% larval lethal so there are no corresponding COL-19::GFP images. J-L) Severe gross morphological and COL-19::GFP phenotypes displayed by P0 *bli-3(n529)* animals that were placed on feeding plates as L2s and incubated at 25°. Protrusions from the body are indicated with arrows (J). In K and L, annulae defects observed at the COL-19::GFP level are denoted with white arrows and alae defects with black arrows. Lateral alae are marked 'la'.



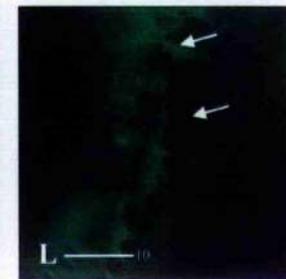
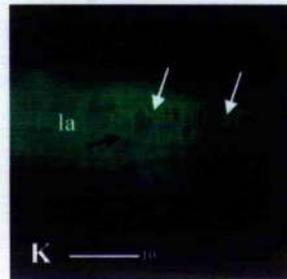
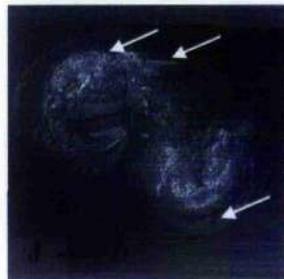
untreated



15°C  
(embryos/L2)



25°C  
(embryos)



25°C  
(L2)

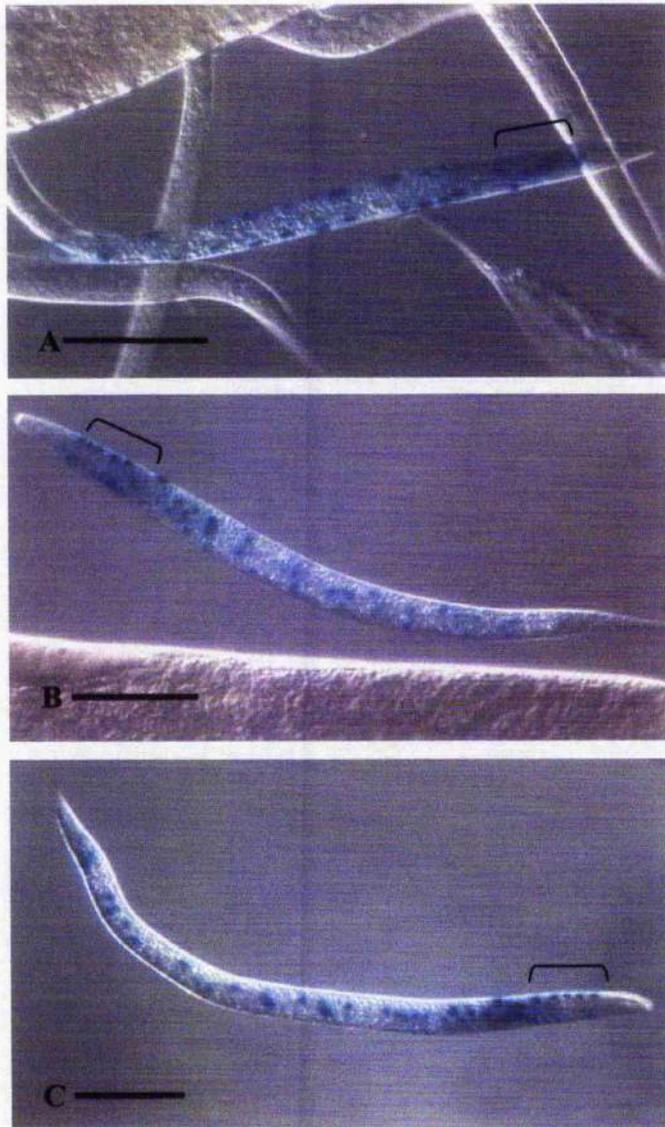
To further characterise the effects of RNAi on the *bli-3(n529)* strain, RNAi of embryos and L2 animals were assayed at 15°C, 20°C and 25°C. These results show that by using L2-stage animals, it is possible to delay lethality in order to obtain scorable phenotypes (figures 5.10 E and F (15°C) and 5.10 J-L (25°C)). However, only a handful of animals survive to adulthood (e.g. 2 of 30 animals at 25°C) and of these almost all show a completely dissolved phenotype (figure 5.10 J). The COL-19::GFP pattern of these animals at both temperatures appears as an extreme form of branched alae. The overall appearance is reminiscent of a loosely-formed network of fibres (figures 5.10 K and L) and figure 5.10 F demonstrates that the amount of collagen being processed into the cuticle is severely compromised. The depicted worm exhibits a severely aberrant cuticle, devoid of staining; only a few strands of COL-19::GFP covering only part of the cuticle are visible (figure 5.10 F).

The placing of embryos on feeding plates is more potent: only 40%, 60% and 30% of embryos hatch at 15°C, 20°C and 25°C respectively (based on between 120 and 200 plated embryos). There is practically 100% larval lethality with those that do survive being barely viable grubs having such minimal amounts of defined body morphology that they cannot be moved. All RNAi-treated embryos exhibit severe gross morphological phenotypes during their larval stages and these are illustrated in figures 5.10 D (15°C) and 5.10 J-L (25°C). These show that animals have no resemblance to normal worm morphology, have extreme blisters, moult defects and have many protrusions and lumps.

### 5.2.12 Spatial expression analysis of *hpx-1*

In order to characterise the spatial expression pattern of *hpx-1*, a 2115 bp putative promoter region of the *hpx-1* gene was cloned into the reporter gene vector pPD96.10. The amplified region included approximately 2000 bp of 5' sequence as well as sequence encoding the predicted signal peptide and 4 residues 3' to the signal peptide cleavage site (the site of which had been predicted using a Signal P prediction program). This facilitated the observation of cellular sites of *hpx-1* expression. Staining was observed in all larval stages, localising to the nuclei of the hypodermal cells, and predominantly in the head, tail and vulva regions. Two stably inherited lines were obtained and X-gal staining was similar in both (figure 5.11, which shows the staining of one line).

Since the hypodermis is the site of expression of the cuticle collagens as well as a number of enzymes of the cuticle processing machinery (e.g. *dpy-18* (Winter and Page, 2000)), the *hpx-1*



**Figure 5.11: Tissue specific expression of *lacZ* under the influence of *hpx-1* promoter sequence in L3 larvae.** Driven by a synthetic NLS signal encoded by the vector, the encoded  $\beta$ -galactosidase is sequestered to the nuclei of the cells where expression occurs. In this case, there is clear staining of the hypodermal cells. Such hypodermal expression is characteristic for other enzymes required for cuticle synthesis. The head hypodermal cells are denoted with brackets. Scale bars = 10 $\mu$ m

expression observed here strengthens the suggestion that HPX-1 has a role in cuticle production. *C. elegans* Duoxs are also expressed in the hypodermal cells (Edens *et al.*, 2001), and such a similarity in spatial expression of Duox and *hpx-1* is significant if Duox is the H<sub>2</sub>O<sub>2</sub> donor.

#### 5.2.13 SQ RT-PCR of *hpx-1*

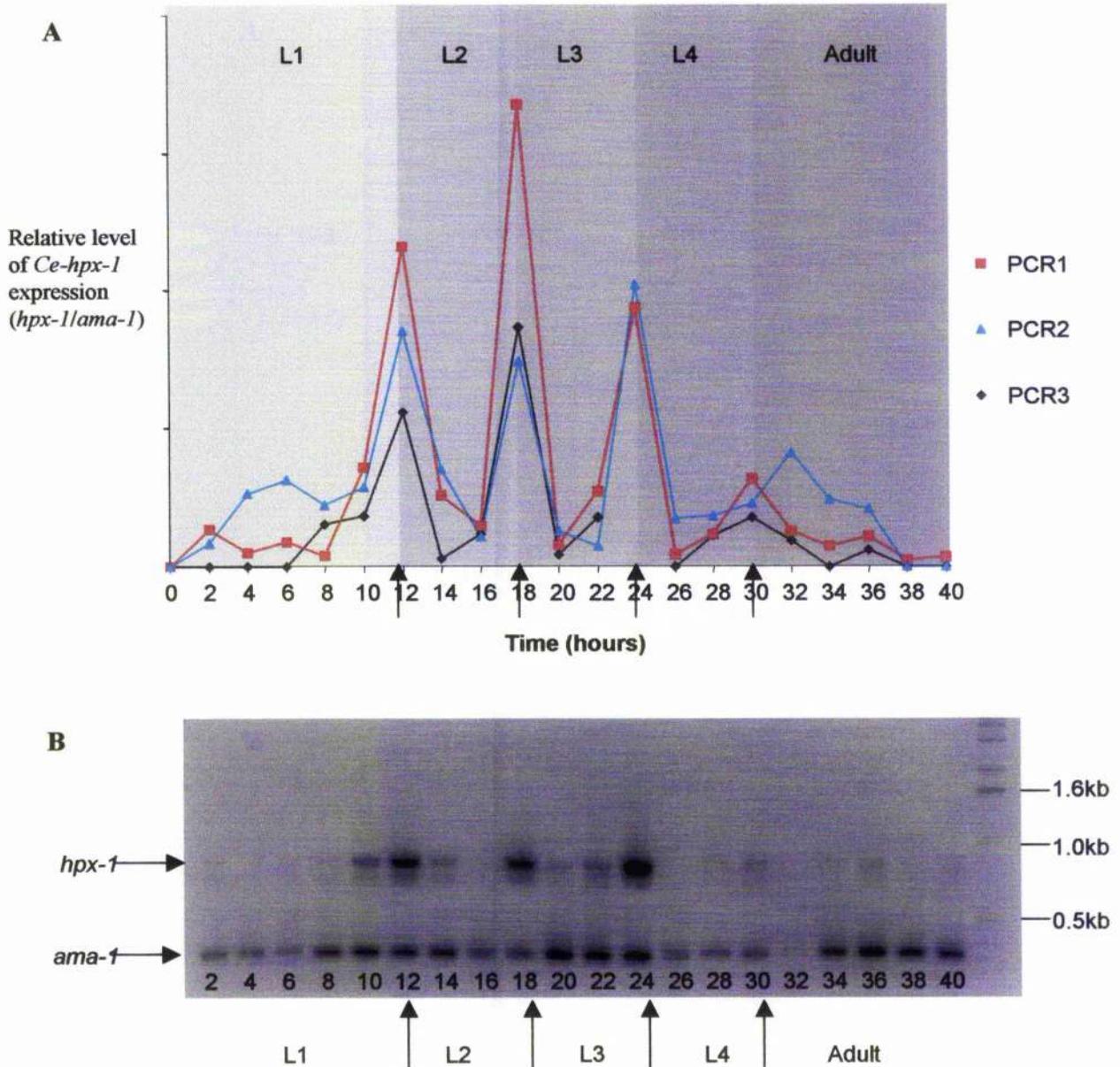
A semi quantitative-reverse transcriptase PCR (SQ RT-PCR) approach was used in order to examine the temporal expression pattern of *hpx-1*. This method uses the premise that *ama-1* is constitutively expressed throughout the larval and adult stages and that the expression level of a gene of interest can thus be expressed in relative terms of *ama-1*'s constant level of expression. A set of cDNAs representative of the mRNA population at every two hours of development was used to amplify fragments of *ama-1* and *hpx-1*. The results, which are based on triplicate experiments, are shown in figure 5.12.

Both the graph (figure 5.12 A) and the gel picture (figure 5.12 B) clearly show peaks of *hpx-1* expression that coincide with the larval moults (at 12, 18, 24 and 30 hours post-hatch). This is substantial evidence that HPX-1 has a function in the processing of the cuticle, either the formation of the new cuticle, or the degradation and shedding of the old one. The latter role is supported by SEM data that show a clear moulting defect while the former role is shown by loss of integrity to the cuticle, especially the struts; hence the blister phenotype of RNAi-treated worms.

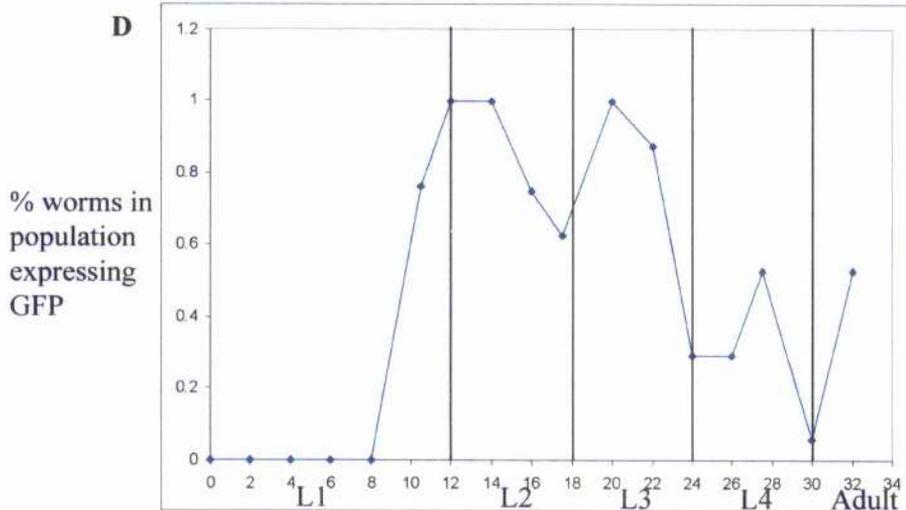
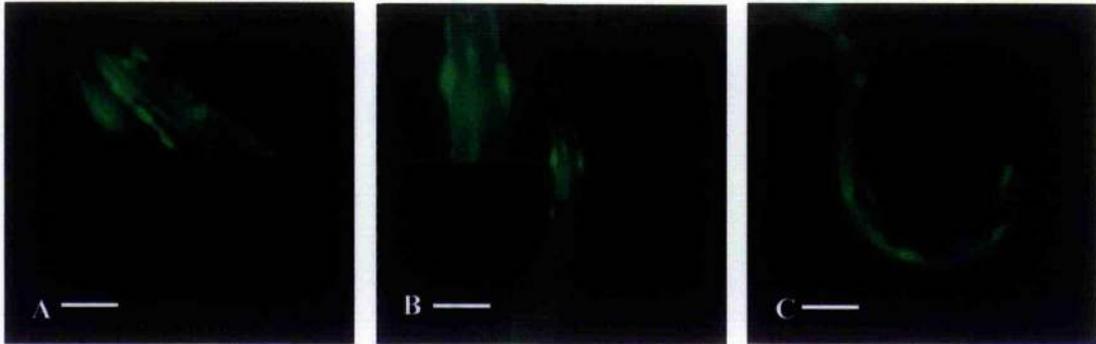
#### 5.2.14 Temporal expression of *hpx-1* determined by a destabilised EGFP marker

2089 bp of the putative promoter of *hpx-1* as well as 14 bases of coding sequence (encoding 4 residues plus two bases that keep the construct in frame) were cloned into pPAF207. This vector (a gift from A. Frand) has been designed in order to create a vector tool with the capacity to assay the kinetics of transient gene regulation. This is facilitated by the incorporation of a destabilised enhanced green fluorescent protein (pd1EGFP-N1), which has a half-life of approximately 1 hour. Consequently the GFP reporter, under the influence of the promoter in question, is degraded after an hour of being expressed, thus ensuring that no residual reporter product remains.

The GFP reporter is not sequestered to the nucleus and consequently, the GFP signal in transformed animals is extremely strong and can easily be visualised under a low-power dissection microscope. Staining is predominantly observed in the head and tail regions.



**Figure 5.12: Results from SQ-RT PCR of *hpx-1* using *ama-1* as a control.** A) Graph representing data from three separate PCR experiments. Level of *hpx-1* expression at each 2 hour intervals of the life cycle is depicted. Each larval stage is denoted by a different shade of grey and moults are indicated with vertical arrows. Peaks of *hpx-1* expression coincide with these moults. For each PCR experiment, the *hpx-1/ama-1* ratios were calculated. The relative efficiencies of the *hpx-1* and *ama-1* amplifications varied between experiments and consequently the *hpx-1/ama-1* ratios differed in scale in different PCR experiments. Thus, for ease of presentation, in graph A, the data from the separate experiments have been normalised. B) Original agarose gel from one PCR experiment (PCR1). The *hpx-1* and *ama-1* bands are indicated as are times of larval moults (vertical arrows) and the different larval and adult stages. Final lane is the 1 kb ladder.



**Figure 5.13: Expression of a destabilised GFP marker under the control of the *hpx-1* promoter.** *hpx-1* putative promoter in a pPAF reporter construct enables temporal and spatial expression to be assessed in relation to each other. A-C: Reporter GFP localisation in L3 (A) and L2 (B and C) larvae from a synchronous population. The inset in figure B is an enlargement. The marker localises to hypodermal cells but, due to the high expression level, leaks out in a non-specific manner. Scale bars = 10  $\mu$ m. D: Temporal expression over a 34-hour time course analysis. Times of moults are denoted with a vertical line. Expression of the marker is cyclical. The broad nature of the peaks is due to the  $\sim$ 2 hour half-life of the marker.

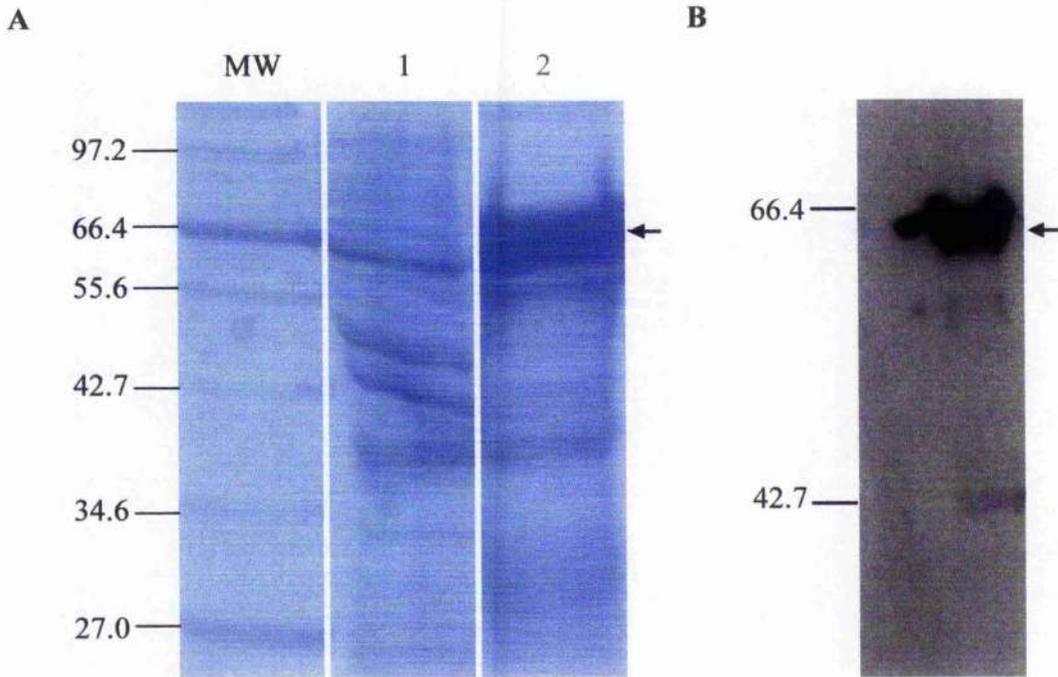
When viewed under higher power, it becomes apparent that there is hypodermal expression of the reporter construct. This is illustrated in figure 5.13 A where the GFP reporter can be seen localising to the perinuclear space of the hypodermal cells of the head. In addition, the pharynx is a major site of GFP localisation. Another interesting phenomenon is that the pattern is reminiscent of the cuticular localisation of collagens (figure 5.13 B). It is extremely unlikely that this reporter construct would be incorporated into the cuticle, specifically in what appears to be the annular furrows, since only 3 residues of the coding sequence are included in the construct. These would not be sufficiently long in length to mediate specific interactions with cuticular components. It is more feasible that the pattern observed, is in fact, non-specific annular-looking staining from the high expression level of the reporter construct which would cause the reporter to leak out of the cell at the site of expression and into surrounding tissues (in the case of figure 5.13 B, leakage might be into the intermediate filaments). A more extreme example of this non-specific GFP localisation is exhibited in figure 5.13 C. In worms expression such a phenotype, it appears that the only cells from which the reporter molecule is being excluded, are the seam cells.

A time course analysis to observe the temporal expression pattern of the reporter was carried out. Expression of the GFP reporter was monitored at 2 hour intervals post hatch of a synchronous population at 25°. The graph in figure 5.13 D illustrates that a cyclical pattern of *gfp* expression is observed and supports the findings of the SQ-RTPCR analysis. The greater peaks of expression are wider, due to the 1 hour half-life of the reporter. However, the last two moults appear to be out of synchrony.

#### 5.2.15 Expression of recombinant HPX-1

A 1772 bp fragment of *hpx-1* cDNA spanning from exon 5 to the termination codon was cloned into the vector pQE-30 (Qiagen) which adds an N-terminal His-tag to the encoded protein. This is a sufficient amount to encode for the full-length animal haem peroxidase domain but excludes the ShTK domain. M15 cells transformed with this construct and supplemented with IPTG enabled the induction of a ~65 kDa peptide to be produced (at 37°C). Figure 5.14 A shows an SDS PAGE gel in which the band corresponding to the encoded protein (of the expected size) is shown to be induced.

Recombinant protein was insoluble in the standard phosphate buffers described in the Qiagen protocol and consequently, urea-containing buffers were used during the purification procedures



**Figure 5.14: Recombinant HPX-1.** A) SDS PAGE showing protein content in 1 ml uninduced cells (lane 1) and elution fraction (lane 2). Molecular weight standards are shown in kDa. Arrow indicates band of induced and purified recombinant HPX-1 peptide (fragment of HPX-1 that includes the peroxidase domain ~66 kDa). B) PDVF membrane blotted from elution fraction and probed with anti-His antibodies. Arrow indicates band of interest.

which used a nickel-nitrilotriacetic (Ni-NTA) column. Elution was carried out by reducing the pH of the 8M urea buffer to pH 4.5. A single major band was eluted. Confirmation that the eluted purified protein was that encoded by the cloned fragment was achieved by Western analysis using antibodies against the His tag (figure 5.14 B). This shows the comparable sizes of the bands appearing upon IPTG-induction and labelled by anti-His antibodies.

pH 4.5 is a sub-optimal pH for the previously characterised peroxidases whose activities were shown to peak at pH 6 (Qian *et al.*, 1997). Consequently, a series of dialyses were performed to increase the pH of the protein solution. However, increasing the pH, even to pH 5.0, resulted in precipitation of the protein. Reducing the urea concentration via dialysis was also tried. Both dialysed solutions were assayed for peroxidase activity using a TMB substrate system, as described in Edens *et al.* (2001). Peroxidase activity in such an assay is accompanied by a spectral change. However, neither of the protein solutions tested were able to show peroxidase activity.

Further work is necessary and a number of parameters need to be changed. These include alteration of pH, urea concentration, and using different length fragments of the peptide. The decision to clone only a fragment rather than the full-length peptide was made, since cloning of the latter was not possible even after alterations to the primers and PCR conditions. The peroxidase assays described in Edens *et al.* (2001) were successfully performed using a peptide containing only the animal haem peroxidase domain of BLI-3. These experiments thus showed that for BLI-3, at least, the isolated peroxidase domain is sufficient for activity.

### 5.3 Discussion

The initial RNAi screen indicated that animal haem peroxidase-containing gene products have the capacity to be highly disruptive to the cuticle upon RNAi targeting. One such gene, zk430.8 that encodes for such a peroxidase, was chosen as a candidate for further analysis on the basis of the severity of its RNAi phenotype at both the COL-19::GFP and gross morphological levels. This gene has now been termed *hpx-1* (animal haem peroxidase). The extreme RNAi-induced Bli phenotype and COL-19::GFP pattern exhibited by *hpx-1*(RNAi) animals is reminiscent of that of the previously characterised BLI-3 and its f53g12.3 homologue, enzymes that are collectively called the Duoxs and that have been positively attributed to functioning as cross-linking catalysts in the cuticle. The term Duox is derived from the fact that these enzymes have two distinct catalytic oxidative domains, namely an animal haem peroxidase domain and an

NADPH-oxidase domain. This coupling is unique, as the latter domain provides the  $H_2O_2$  substrate required for the peroxidative reaction. Interestingly, the Duox enzymes represent the only such domains encoded by the *C. elegans* genome and it is therefore possible that the  $H_2O_2$  requirements of other peroxidases may be Duox-derived. Consequently, in the study of HPX-1, it is essential that its functional relationship with BLI-3 be analysed. This is especially important in the light that, as shown here, RNAi-targeting of HPX-1 homologues does not result in drastic cuticle disruption. On the assumption that such RNAi-treatment is a true representation of the “knock down” effect of the homologues, this suggests that Duox and HPX-1 are the key peroxidases. In order to do this BLI-3 and HPX-1 have been compared at the sequence level; their respective RNAi-phenotypes have been analysed using the cuticle markers COL-19::GFP and DPY-7; *hpx-1*(RNAi) animals have been analysed at the SEM level; morphological mutants of the *bli-3* locus have been characterised; and these mutants have also been targeted by *hpx-1* RNAi. In addition, to further characterise *hpx-1*, its temporal and spatial expressions have been determined, and detection of peroxidase activity of recombinant forms of the protein has been attempted.

### 5.3.1 Analysis of BLI-3, f53g12.3 and HPX-1 at the sequence level

*bli-3* and f53g12.3 which, as described above, encode for a peroxidase domain, an NADPH-oxidase domain and EF-hand  $Ca^{2+}$ -binding sites, are 94% identical (and are 30% identical with their human counterparts (Edens *et al.*, 2001)) and are thought to be derived from a recent gene duplication. Both enzymes encode for a C-terminal peroxidase domain and a N-terminal NADPH-oxidase domain with EF-hand domains  $Ca^{2+}$ -binding domains in between. However, F53G12.3 is N-terminally truncated which renders the NADPH domain inactive. Thus BLI-3 encodes the only truly functional form of this oxidase.

An alignment of peroxidase domains of 13 animal haem peroxidases is shown in figure 5.5. It shows that a number of residues are conserved between all peroxidases and many of these are the active site residues or ones that are involved in  $Ca^{2+}$  binding. However, BLI-3 has been placed at the top of the list. This is because the peroxidase domains of BLI-3 and F53G12.3 and the human Duox enzymes, diverge from other characterised peroxidases to such an extent, that they is regarded as a functionally distinct family of haem peroxidases. This is on the basis that a number of the essential active site residues and disulphide bond-forming cysteines, which are conserved among the other peroxidases, are not encoded by *bli-3*. The lack of some of the conserved cysteine residues suggests that the folding topology of the active site differs between

the Duox enzymes and the other peroxidases. This makes it possible that active site residues might not be conserved in terms of their positions in the sequence but may be conserved in terms of their spatial position in the active site. This has been previously observed with plant/bacterial/fungal peroxidases which show diversity to animal haem peroxidases at the sequence level but which show similarity in superimpositions of active site topologies (Taurog, 1999).

*hpx-1* encodes for a peroxidase domain that has all the active site and cysteines residues that are conserved in the other peroxidases. It is also predicted to have similar secondary structure. These observations suggest that its functional reaction and folding topology are equivalent to the other peroxidases. The peroxidase domain of HPX-1 is most similar to that of TPO, an observation that may be significant due to the fact that the main function of the latter enzyme is in mediating tyrosine cross-links during the process of thyroid hormone production. TPO uses protein-bound tyrosine as its major cofactor and is therefore distinct from other peroxidase which, because of the fact that proteins cannot access their active site, use a "long range" effect to target proteins. Thus the homology of HPX-1 and TPO may indicate that protein-bound tyrosines can access the active site of HPX-1, a set-up that would be fundamental to the efficiency of its cross-linking activity.

### 5.3.2 *hpx-1* temporal and spatial expression

The temporal expression of *hpx-1* has been analysed using a SQ RT-PCR approach which clearly shows that synthesis of its gene product is cyclical, peaking coincident with the larval moults. This is a typical expression profile for components and modifying enzymes of the cuticle and therefore this data provides some evidence of the role that HPX-1 plays.

The spatial expression of *hpx-1* has been determined in two separate approaches, in which *lacZ* or a destabilised GFP were put under the influence of the predicted *hpx-1* promoter sequence. These approaches differ because, while the *lacZ* approach results in the encoded  $\beta$ -galactosidase marker being sequestered to the nucleus, the GFP is not nuclear localised. However, the use of a destabilised GFP means that the marker is turned over at high rate, ensuring that marker is present only as a consequence of recent promoter activation. Thus such a method potentially allows for the spatial and temporal expressions to be observed in relation to each other. Both methods of analysis suggest that *hpx-1* is hypodermally expressed.

### 5.3.3 Assaying the peroxidase activity of recombinant HPX-1

The peroxidase domain alone of BL1-3 was shown previously to be sufficient for activity (Edens *et al.*, 2001) so when attempts to amplify the full-length sequence of *hpx-1* failed, making recombinant protein of the peroxidase domain alone was considered to be a credible option. The His-tagged peroxidase domain of *hpx-1* was expressed as a recombinant protein, purified using a buffer containing urea and dialysed to 0 M urea and pH 4.5.

Attempts at increasing the pH further (even to pH 5) resulted in the precipitation of the protein. The recombinant protein was used in assays for peroxidase activity using a TMB substrate system in which activity is accompanied by a change in spectral properties. No such change was observed from the recombinant peroxidase. A number of reasons are suggested for the apparent lack of peroxidase activity. These include: that pH 4.5 is a sub-optimal pH for the reaction to occur in (MPO, for example, has an optimal activity of around pH 6.0 (Qian *et al.*, 1997)); that the presence of the peroxidase domain alone is not sufficient for HPX-1 activity (it may be folded incorrectly or not include important residues); or that undetectable levels of peroxidase activity are occurring. Another reason is based on the findings by Qian *et al.* (1997), who established that very low concentrations (e.g. <10  $\mu$ M) of cyanate (which forms spontaneously in solutions of urea) irreversibly inhibits the peroxidase activity of MPO in the presence of H<sub>2</sub>O<sub>2</sub>. Thus, despite being dialysed to remove urea from the solution, since such low levels of cyanate are sufficient to be inhibitory, there may have been residual urea from the dialysis step that could have inhibited any peroxidase activity by the recombinant protein. (The final reduction in urea concentration via dialysis was from a 0.5 M to zero molar solution).

In future experiments, it will be necessary to optimise the conditions for the peroxidase reaction. A final dialysis step, into a fresh solution of zero molar urea be included that may reduce the urea concentration further. As the recombinant protein cannot be dialysed into solutions with pHs above 4.5, it may be necessary to re-clone a new recombinant protein that would allow this. This may also allow critical residues involved in folding or peroxidase activity to be included into the recombinant protein.

### 5.3.4 RNAi of *bli-3* and *hpx-1*

The severe gross morphological phenotype of *bli-3*(RNAi) and f53g12.3 (RNAi by feeding) worms was previously characterised by Edens *et al.* (2001) and found to be equivalently disruptive. As also demonstrated here (at 25°C), targeting of both Duox enzymes results in

extremely aberrant body morphology, severe blistering, lumpy body morphology and a degree of embryonic lethality. Here, the cuticular defects associated with such phenotypes have been examined and have also shown that the phenotypes for the two genes are identical. At the COL-19::GFP level, as would be expected with the variably penetrant RNAi technique, there is a range of *bli-3*(RNAi) phenotypes including wild type and amorphous. The most extreme pattern of COL-19::GFP is more severe than any previously observed in all the tested morphological mutants and RNAi-treated animals. This disruption is in the form of long intertwined COL-19::GFP-stained strands running parallel with the normal position of the alae. Wispy fragments of annulae are also present. An additional phenotype observed in some worms after *bli-3*(RNAi) treatment is aberrant seam cells.

*hpx-1*(RNAi feeding) animals do not exhibit significant embryonic lethality but only about 50% of larvae reach adult stages, suggesting that there is ~50% larval lethality. Gross morphological phenotypes are exhibited by 20% and 30% of adult animals at 20°C and 25°C respectively and include mild Bli (exhibiting three or fewer small blisters), severe Bli (showing larger or >3 blisters), “dissolved”/grub-like (when the cuticle appears weak, there is lack of clearly defined body pattern and the worms appear opaque), and Dpy. Bli is the most predominant phenotype at 20°C while at 25°C, most of the animals display are “dissolved”. COL-19::GFP expression is also varied and includes differing degrees of the amorphous pattern, multiple and slightly twisted and branched alae. The alae aberrations are similar to, but a less severe, than those seen in *bli-3*(RNAi) worms. Branching and broken alae are also visible via Nomarski (figure 5.7 F).

*hpx-1* RNAi-treated animals also have moult defects that are not observed at either the light microscope or Nomarski level. However, using SEM to visualise RNAi-treated animals, abnormal moulting appears to be the dominant phenotype. The characteristic phenotypes for moulting defects are observed and include constrictions around the mid-body of animals, the attachment of the old cuticle around the head of the worm and the inability to lay eggs. It is quite interesting to note that moulting defects occur as *hpx-1* is a “late” expressed gene whose expression peaks coincident with the moult, when the cuticle is being shed. This timing and the moulting defects observed suggest that HPX-1 may function in the shedding processes. Many of the animals exhibit extremely baggy cuticles in which the cuticle appears much too big for the body size of the animal. This is especially apparent in figure 5.8 J. This could be due to the animal being prevented from “growing into” its cuticle because its growth is restricted by the smaller old cuticle or, by its feeding being inhibited by the presence of the old cuticle. It could

be that the secreted cuticle is weakened sufficiently for it to hold itself together. The lumpy phenotype of the RNAi-treated worms is also visible by SEM.

### 5.3.5 DPY-7 localisation in *hpx-1* and *bli-3* RNAi-treated animals

Another observation from the SEM analysis is the presence of annulae on at least a proportion of the animals. This was confirmed by immunolocalisation using antibodies raised against the annular-furrow localising collagen DPY-7. An interesting observation from the analysis of *bli-3*(RNAi) and *hpx-1*(RNAi) animals is that in these animals, COL-19::GFP and DPY-7 expression patterns do not correspond. It has been previously established that COL-19::GFP is susceptible to mutation of both DPY-7 and DPY-5 and that its expression reflects the cuticular disruption when either of these collagens are ablated. Consequently, in the absence of DPY-7, COL-19::GFP localises in an amorphous pattern. Conversely, if the annular furrows are defined, COL-19::GFP localises between them. Upon DPY-7 staining of *hpx-1* and *bli-3* RNAi-treated animals, however, it appears that despite the DPY-7-containing annular furrows being formed normally, the marker collagen cannot localise in a wild type fashion and is instead amorphous. This is exhibited in figures 5.3 G-O (*bli-3*) and 5.7 G-I (*hpx-1*). These observations suggest that the loss of function of these enzymes affect the localisation of different collagens to differing degrees. This has also been observed in *dpy-11* partial-loss-of-function mutants in which ablation of the DPY-11 cross-linking enzyme disrupts the seam cell-derived cuticle but still enables the proper formation of annulae. It thus appears that the annular furrow-comprising collagens are relatively insensitive to the reduction of higher order cross-linking. However, on the basis of the amorphous phenotype upon the complete abolition of DPY-11 activity in DPY-11 null mutants, a degree of cross-linking is necessary. In this light, because of the formation of annulae after *hpx-1* and *bli-3* targeted RNAi treatment, some cross-linking is likely to be taking place. Despite being present, the annulae are disrupted, being fragmented and broken. The fact that COL-19::GFP is amorphous suggests that the marker collagen is acting differently in these mutants and that the annular furrows no longer have the capability to define COL-19::GFP localisation. This may represent a loss of cross-linking between different sets of collagens which allows them to be positioned correctly in relation to each other. It is therefore quite interesting that the expression profile of *hpx-1* indicates that its expression peaks coincident with the moults. This is considered to be a "late" stage of the moulting process and occurs after the expression of many of the cuticle collagens. DPY-7 and its obligate partners are expressed 4 hours prior to the moult, while DPY-5 and DPY-13 are expressed 2 hours prior. Even SQT-1 collagens appear to be laid down into the cuticle before moulting of the old cuticle occurs.

Considering that the DPY-7 set of collagens are secreted up to 4 hours prior to the expression of HPX-1 occurs, this might suggest that these collagens are not substrates for the peroxidase. This would explain the fact that annular furrows appear to be forming in a wild type fashion in *hpx-1(RNAi)* animals. Conversely, and consistent with the disruption observed in the COL-19::GFP expression pattern, is the fact that *col-19* is a "late"-expressed collagen. The co-expression of the collagen and the peroxidase could therefore indicate that COL-19 is a substrate for *hpx-1*-derived peroxidase activity.

Another interesting discrepancy observed when analysing the COL-19::GFP and DPY-7 staining patterns is seen in figure 5.3 O. This shows that in some *bli-3(RNAi)* worms, a COL-19::GFP pattern reminiscent of DPY-5 animals is observed in which the seam cell-derived cuticle is apparently weakened and uncontracted. However, the results obtained here suggest that this is not the causative effect in *bli-3(RNAi)* animals. This is because DPY-7 antibody staining reveals that the annulae appose the alae in a wild type manner. Since annulae are not present in the seam cell-derived cuticle, the COL-19::GFP disruption must be occurring in the dorsoventral cuticle rather than the seam cell-derived cuticle as in *dpy-5* mutants.

### 5.3.6 *bli-3* mutant strains

Two *bli-3* morphological mutants were analysed at the cuticle level using COL-19::GFP. Both mutants exhibit lesions in the peroxidase domain: *bli-3(e767)*, which is mild Dpy-non-Bli is a G246D lesion, while *bli-3(n529)*, which is Dpy-(Bli) is a D393N lesion. The former allele exhibits type III COL-19::GFP disruption, similar to other Bli loci (e.g. *bli-1* and *bli-5*). Akin to RNAi treatment, therefore, annulae are forming in a wild type fashion in the dorsoventral cuticle. However, it is likely these mutants have disruptions to the seam cell-derived cuticle. Because the anomalies described above suggest the possibility that the disruptions are not in the seam cell cuticle but in the dorsoventral cuticle, immunolocalisation with DPY-7 will be required before it can be confirmed in which cuticle the disruption is occurring. *bli-3(n529)* animals exhibit more severe COL-19::GFP than *bli-3(e767)*. The dominant pattern is amorphous and animals have multiple alae.

The lesions in neither allele appear to affect the residues characterised to be involved in peroxidative enzymatic activity and this makes it difficult to determine why *bli-3(n529)* has the more severe phenotype. It is possible that the lesion in *bli-3(n529)* may alter the protein activity in an indirect way, for example, by altering the state of domain folding. It is also possible that,

since the peroxidase domain of BLI-3 is distinct from the well characterised animal haem peroxidases, the residue affected in *bli-3(n529)* does indeed have a role in enzyme activity.

It is quite interesting to note that one of the residues that is conserved in all peroxidases apart from BLI-3 is the one that is mutated in the strain *bli-3(e767)*. The wild type form of BLI-3 encodes a glycine at this position instead of a conserved histidine. Because of its mutability, the deviation of this residue (in wild type BLI-3) from the other peroxidases may signify that the activity of BLI-3 is unique. On a similar line, it is interesting to note the severity of *bli-3(n529)* since there is no apparent conservation in other peroxidases at this residue. The functional significance is less clear and its mutability may be a sign that, in BLI-3 only, this residue is of functional importance.

### 5.3.7 Double targeting of *bli-3* and *hpx-1*

By performing *hpx-1* RNAi on *bli-3* mutant alleles, a clearer understanding of the relative roles of each peroxidase has begun to be established (figure 5.15). This study has revealed that the disruption observed upon subjection of *bli-3(e767)* animals to *hpx-1*(RNAi) is comparable to that when TP12 animals are similarly treated. In fact, larval lethality was reduced in the former case. However, this is likely to be an artefact of reduced penetrance of RNAi.

The case is very different when *bli-3(n529)* animals are subjected to *hpx-1* RNAi; when L4 larvae are placed on *hpx-1* RNAi feeding plates, 100% die within 24 hours. When larvae and embryos are treated, it is possible to delay lethality in order to obtain scorable phenotypes. The effect of RNAi on *bli-3(n529)* embryos and L2 larvae was carried out at 15°C, 20°C and 25°C. RNAi-treatment of embryos at 25°C had the most potent effect, being 100% lethal. Only a few treated embryos survived at 15°C but they were so extremely dissolved that COL-19::GFP could not be observed. Larvae treated in this manner (at all temperatures) produced animals that completely lost the vermiform shape, and had multiple protrusions, moulting defects, lumpy body morphology and severe blisters. By feeding L2s, however, more viable animals were obtained in order to analyse the cuticular pattern. At 15°C, both larvae and adults were morphologically mutant and exhibited similar gross morphological phenotypes as the treated embryos. As adults, a pattern of severe COL-19::GFP disruption being highly branched and appearing like a stretched net, was apparent at both 15°C and 25°C (figures 5.10 F, K, L). The COL-19::GFP intensity was weak in these animals, suggesting that there was significant reduction in the cuticle components.

**Figure 5.15: Investigating the activities of BLI-3 and HPX-1 in morphological mutants and in RNAi experiments.** Peroxidase domains of BLI-3 and HPX-1 are coloured in red while the NADPH-oxidase domain of BLI-3 is in yellow. This domain is potentially the sole H<sub>2</sub>O<sub>2</sub> source for all peroxidases and this H<sub>2</sub>O<sub>2</sub>-donating activity is denoted with dashed lines. A red cross indicates when the NADPH domain is ablated by RNAi, with the presumed loss of the H<sub>2</sub>O<sub>2</sub> -donating activity. Loss of activity is shown when the domain is shaded in grey, with lighter shades denoting more loss of activity. The tick marks indicate which domains are active. They are bracketed if this activity is compromised by mutation. The final column denotes the resulting severity. It is scored from 0 – 5 where 0 is wild type and 5 is lethal.

Facing page 277

		BLI-3 NADPH-oxidase	BLI-3 peroxidase	HPX-1 peroxidase	Severity
WT		✓	✓	✓	0
<i>bli-3(e767)</i>		✓	[✓]	✓	1
<i>bli-3(n529)</i>		✓	[✓]	✓	2
RNAi <i>bli-3</i>				[✓]	4
RNAi <i>hpx-1</i>		✓	✓		3
<i>bli-3(e767)</i> + RNAi <i>hpx-1</i>		✓	[✓]		3
<i>bli-3(n529)</i> + RNAi <i>hpx-1</i>		✓	[✓]		5

### 5.3.8 Overview of *bli-3* and *hpx-1*

Figure 5.15 shows the list of scenarios that have been tested here. It shows which domains are active upon each treatment and the resulting severity (based on gross morphological and COL-19::GFP phenotypes). For ease of description, henceforth, the relative degree of severity will be described as a score ranging from 0 (wild type) to 5 (100% lethality). A few assumptions are made: that RNAi is 100% effective, and that BLI-3 is the sole H<sub>2</sub>O<sub>2</sub> donor. Upon *bli-3* (RNAi) treatment therefore, it is assumed that peroxidase function of f53g12.3 (its NADPH-oxidase-deficient homologue), HPX-1 and any other active peroxidase is inactive. A number of observations are made:

- 1) HPX-1 is an essential peroxidase. This is shown by its severe (severity=3) RNAi phenotype and the fact that removal of both HPX-1 and BLI-3 is sufficient to induce 100% lethality.
- 2) The peroxidase domain of BLI-3 is weakly active. This is shown by the fact that *hpx-1*(RNAi) is severe (severity=3) indicating that the peroxidase that remains (i.e. from BLI-3, F53G12.3 and any other peroxidase) is not sufficient to compensate for the loss of HPX-1.
- 3) Despite being unable to compensate for HPX-1, BLI-3 is an active peroxidase with some function: Edens *et al.* (2001) first proved this by showing the activity of recombinant peroxidase domains. This is also demonstrated by the fact that *bli-3* alleles with lesions in the peroxidase domain exhibit aberrant cuticles and body morphology, albeit, relatively mildly (severity=1 for *bli-3(e767)* and =2 for *bli-3(n529)*).
- 4) BLI-3 peroxidase activity is essential. This is demonstrated most clearly by the lethality when *bli-3(n529)* animals are treated with *hpx-1* RNAi. In such a scenario, the peroxidase domain of BLI-3 is compromised (it cannot be said to what extent), and HPX-1 function is ablated. These factors are sufficient for the extreme severity (severity = 5) and there are no peroxidases (including F53G12.2) with functions that can replace the activity of BLI-3.

- 5) The lesion in *bli-3(e767)* animals does not reduce the peroxidase activity of BLI-3 to a large degree. This can be determined because of the weak phenotype of this allele (severity =1) and the fact that when this strain is treated with *hpx-1* RNAi, the effect is not worsened compared to *hpx-1* RNAi treatment of wild type (TP12) animals.

It is interesting to note that *bli-3* RNAi does not result in 100% lethality since the resulting combination would be the ablation of BLI-3 and any peroxidase activity dependent on its H<sub>2</sub>O<sub>2</sub>-donation. This was proven to be lethal when *bli-3(n529)* animals were treated with *hpx-1* RNA. There are a few suggestions to explain the reduced level of lethality. These include that treatment may not be 100% effective, that there is an alternative source of H<sub>2</sub>O<sub>2</sub> and/or that some peroxidase function can occur in the absence of H<sub>2</sub>O<sub>2</sub>. However, as described above in point 4, it appears that there are no moieties that can replace the function of these enzymes. It is also possible that the 100% lethality observed upon *hpx-1* RNAi treatment of *bli-3(n529)* is caused by non-specific RNAi effects that inhibit other peroxidases. As discussed above, such non-specific effects are thought to be unlikely because of the relative levels of homology shared by HPX-1 and its homologues are below the 80% identity threshold (below which RNAi is believed to be specific). From the above observations, it appears that HPX-1 and BLI-3 are the main mediators of peroxidase activity and are both essential. However, HPX-1 is more essential than BLI-3. This latter point may be associated with the fact that, on the basis of lack of sequence conservation, BLI-3 appears to be functionally distinct from HPX-1 and other peroxidases. These peroxidases may be performing separate functions, or may have distinct targets.

### 5.3.9 Function of HPX-1

The data above have suggested at least two functions of HPX-1. Most of the data supports the fact that HPX-1 is integral for cuticle synthesis, most likely in cross-linking. However, it has also been established that it is likely that only a subset of collagens are substrates for this peroxidase. This conclusion was drawn from the fact that despite the formation of properly formed annular furrows (as demonstrated by DPY-7 immunolocalisation), in (*hpx-1*) RNAi-treated animals, the COL-19::GFP pattern is amorphous. In normal circumstances the localisation of COL-19::GFP is dictated by the position of the annulae. Since COL-19::GFP patterns have been shown to reflect the cuticular disruption of both DPY-7 and DPY-5 mutants and the fact that it is late expressed, COL-19 is proposed to be a "downstream" collagen. Therefore, to find a situation like the one observed in *hpx-1* (RNAi) animals, in which the

COL-19::GFP disruption occurs independently of the annular pattern, suggests that the affinity of HPX-1 is limited to a subset of collagens (which includes COL-19) or that the cuticle disruption is occurring "downstream" of the DPY-7 and DPY-5 secretions. In fact, both appear to be true in the light that *hpx-1* is expressed "late" in the moult. This late expression would mean that the majority of the cuticle components are laid down before the loss of HPX-1 could be disruptive to them. It also means that HPX-1 has a limited subset of substrate collagens, i.e. those that are expressed coincident with the moult. It is interesting to note that one of its putative substrates, COL-19, has been shown, in this study, to be a component of the alae, which, due to their high cuticulin content are likely to be highly cross-linked. Further, it has been confirmed, via expression in a baculovirus system, that COL-19 is cross-linked by non-reducing bonds. These data are strong evidence for this alae-composing collagen, COL-19, as being a substrate for HPX-1 cross-linking activity. Aberrant formation of the alae in these mutants is demonstrated by the severe defects of these structures in (*hpx-1*) RNAi-treated animals.

It was of interest to note whether collagens differed in the number of tyrosines they encoded, since it could follow that COL-19, for example, might have more potential sites of tyrosine cross-linking. However, COL-19, SQT-1, ROL-6, DPY-5 and DPY-7 have 4, 5, 6, 3 and 4 tyrosine residues respectively so there is little variation between collagens. However, an interesting observation is that none of the tyrosine residues in DPY-7 are adjacent to cysteine residues, which in SQT-1 mutants have been demonstrated to be vital for proper cross-linking. This may be the distinguishing feature, perhaps a control mechanism, that results in DPY-7 having fewer cross-links, and consequently being less susceptible to the loss of HPX-1 and BLI-3.

*hpx-1* expression at the time of ecdysis and the moulting defects observed in *hpx-1* RNAi-treated worms suggests that HPX-1 might also be playing a role in the moulting process. It is interesting to think how a cross-linking enzyme would be useful in such a degradative process. There is considerable evidence for peroxidase-derived strengthening of the cuticle, and this has been shown to be a function of the BLI-3 enzymes. However, it has also been previously noted that peroxidases are able to degrade cuticles. MPO-derived HOCl, for example, has been shown to be involved in matrix dynamics since it oxidises ECM components such as laminins, thrombospondins and type IV collagens. This is very interesting as HPX-1 may have similar abilities. It has also been reported that HOCl oxidises pyridinoline cross-links (which are present in *C. elegans* cuticles), an action which was shown to decrease basement membrane matrix

integrity in bone collagen (Daumcr *et al.*, 2000). Yet another observation is that in *C. elegans*, the cuticle of old adults is associated with increased tyrosine and lipid cross-linking (Davis *et al.*, 1982). Finally, MPO has been shown to activate and subsequently inactivate matrix metalloproteases, which as their name suggests, have roles in matrix proteolysis, thus providing another way in which peroxidases can modulate the matrix. These exemplify a number of ways that HPX-1 could function in the moulting process. Therefore, it appears that HPX-1 may have roles in both cuticle synthesis and degradation of the cuticle.

#### 5.4 Future work

The main priority of future work will be to establish whether HPX-1 catalyses the formation of tyrosine cross-links in the *C. elegans* cuticle. Optimisation of the purification process of the already cloned recombinant HPX-1 protein may be possible. However, it may be necessary to create a new construct from which recombinant protein can be purified without the use of urea and whose pH can be raised suitably. Even after confirming the ability of recombinant HPX-1 to catalyse a peroxidative reaction, it will be vital to establish whether such activity results in cuticle cross-links and whether the disrupted cuticle in RNAi-treated animals is due to their loss. For this, an approach described by Edens *et al.* (2001) could be utilised in which cuticle extracts from wild type and RNAi-treated worms would be analysed and compared by HPLC monitoring fluorescence. Identifiable peaks for dityrosine and trityrosine become apparent in wild type worms (Edens) and loss of these peaks in spectra of (*hpx-1*) RNAi-treated worms would be an indication that cross-linking activity had been lost.

Recombinant protein might also be used to raise antibodies against HPX-1. These would be invaluable in order to establish where the protein resides (e.g. maintained in hypodermal cells or secreted). It may also be useful to alter the construct in which the *col-19* promoter was fused to destabilised EGFP. Transgenic lines carrying this construct had quite non-specific staining which could possibly be reduced by using a shorter promoter sequence containing the minimal sequence required for expression but lacking 5' enhancing sequences. It might also be useful to integrate these constructs, in order to eliminate any mosaicism.

It will also be interesting to examine another *bli-3* allele (*bli-3(gk141)*) using the TP12 strain (in crosses) and also treat it with *hpx-1* RNAi. This will be useful because this allele has an 830 bp deletion in the NADPH-oxidase domain and may therefore be able to provide information on the role of BLI-3 in producing H<sub>2</sub>O<sub>2</sub>. Despite this being an exciting prospect, the *bli-3(gk141)* strain

is wild type at the morphological level, which may indicate that the deletion does not affect its activity.

As described above, attempts were made to produce null mutants of both *hpx-1* and *bli-3* and in the future these will be extremely useful.

**Chapter 6**  
***General Discussion***

## 6.1 Aim

The aim of this thesis was to investigate the biosynthesis of the *C. elegans* cuticle from a number of perspectives. Much of the study has stemmed from the use of a COL-19::GFP-containing strain, TP12.

## 6.2 Wild type expression of *col-19*

It was necessary to examine the wild type expression pattern of COL-19::GFP. These investigations demonstrated that *col-19* is hypodermally expressed late during the adult moult and that its encoded product localises to the seam cell-derived cuticle, annular cortex and is an integral component of the alae. Similar localisation to these structures is observed when COL-19 is tagged with either a GFP or a Ty-tag marker, indicating that the expression pattern observed in TP12 animals is reflective of the wild type localisation of COL-19. The localisation of this collagen is dependent on the presence of the annular furrow-localised collagens and thus, when expressed in mutants, the marker collagen is able to reflect the cuticular and alae aberrations of the *dpy-7* and *dpy-5* classes of morphological mutants. Secretion of COL-19::GFP into the cuticle of these strains indicates that none of the morphologically mutable collagens are obligate partners of COL-19, which itself is able to homotrimerise.

## 6.3 Processing of collagens

Collagens were placed under the influence of the alternative collagen promoters in order to investigate the properties of COL-19 that allow it to apparently avoid C-terminal proteolytic cleavage and form a functional fusion. These studies have shown that there appears to be a limited repertoire of late-expressed proteases which themselves have specific collagen substrates. This was concluded on the basis that *sqt-1*, but not *dpy-7*, *dpy-5* or *dpy-13*, can form a functional fusion when placed under control of the *col-19* promoter.

## 6.4 The molecular bases of morphological mutants

The analysis of the COL-19::GFP expression in the cuticles of *Sqt*, *Rol*, *Lon*, *Sma*, *Dpy* and *Bli* morphological mutants and RNAi-treated animals has enabled the molecular bases of certain morphological phenotypes to be further analysed.

### 6.4.1 Lon mutants

Data from TP12-crossed *Lon* strains have confirmed that the cuticle in *Lon* animals can either be actively altered by modulation of the collagenous content (*lon-3* mutants) or can passively stretch as a consequence of increased ploidy (*lon-1* mutants).

### 6.4.2 Dpy mutants

The Dpy phenotype has been shown to be the consequence of either the overall weakening of the dorsoventral cuticle, or the weakening of the seam cell-derived cuticle and its consequential inability to contract. These data have demonstrated both the different mutabilities of the two distinct cuticular regions and the essential role of the annular furrows in establishing a structured cuticular framework. Both Dpy-related phenotypes are the result of a substantial decrease in the collagenous content in the cuticle. These data have shown that while DPY-7 and similarly temporally expressed collagens are crucial components of the annular furrows, DPY-5 collagens localise between them and in addition, contribute to the collagenous content of the seam cell-derived cuticle.

### 6.4.3 Bli mutants

The blisters of Bli animals are a direct result of the separation of cuticle layers (Crew and Thacker, personal communication), but observations from this study have enabled the molecular basis of the slightly Dpy phenotype of Bli mutants (de Melo *et al.*, 2003) to be postulated. It has been inferred that in *bli-1* and *bli-5* mutants, the seam cell-derived cuticle is altered as a consequence of the loss, or alteration of, interactions between the strut-localised BLI-1 and BLI-2 collagens and collagens residing in the seam cell-derived cuticle.

### 6.4.4 Sqt and Rol mutants

The roles that the collagens, encoded by the *sqt* loci, have in the cuticle have been investigated. These studies have provided a means of visualising the severely branched annulae and alae of N-terminal processing- and cross-linking- mutants and have also demonstrated that the apparently wild type RNAi-treated animals, actually display limited amounts of cuticle disruption. These data support that the role of these collagens, which are presumed to be localised in the fibrous layers, are not entirely defunct. The data also complement previous studies which determined the disruptive effect that mutant forms of these collagens have on the cuticle (Yang and Kramer, 1999).

These studies have also shown that *rol-6* and *sqt-1*, although interacting genetically (Kramer *et al.*, 1990) and being similar in terms of their mutant alleles (Kramer and Johnson, 1993), may encode collagens that are functionally distinct. This is on the basis that *rol-6* and *sqt-1* RNAi COL-19::GFP phenotypes are significantly different. The apparent lack of annular localisation of COL-19::GFP in the former mutant has been attributed to ROL-6 playing an essential role in

dictating the cuticular localisation of at least one collagen but not being an integral component of any of the cuticular layers. SQT-1, on the other hand, has been demonstrated to be an essential component of the fibrous layer, while not being necessary for the localisation of COL-19::GFP.

### 6.5 Enzymes in cuticle synthesis and RNAi

The sensitivity of the COL-19::GFP marker to show cuticular defects, even in the absence of gross morphological phenotypes, was exploited in an RNAi screen in which a number of enzymes with potential roles in *C. elegans* morphogenesis and cuticle processing were targeted. The marker has proven to be a useful tool, demonstrated by the fact that it has implicated at least one gene (*t10h10.2*), with a COL-19::GFP disruption pattern but wild type phenotype, as having a role in cuticle synthesis.

Using the information gained from the cuticle morphological mutants has enabled the collagen substrates for enzymes to be identified. This has been useful in the RNAi screen and in the investigation of enzyme-related mutants. Such studies have demonstrated that the DPY-7 set of collagens are able to form apparently wild type annular structures in the absence of DPY-18, indicating that these collagens may not be the target of this enzyme or that such processing is not a prerequisite for their secretion into the cuticle. DPY-11, on the other hand, is likely to have a wider range of collagenous substrates, based on its amorphous COL-19::GFP phenotype of the null allele. However, in *dpy-11* partial loss of function mutants, the seam cell-derived cuticle is aberrant while the dorsoventral region appears wild type. This implies that the DPY-7 set of collagens within the annular furrows are normally processed while those in the seam cell-derived region are not. These observations may reflect the differing collagen substrate-specificities of the partially functioning enzyme, or the differing susceptibilities of the two regions to cuticle aberrations. Similar examination of the cuticular substrates of enzymes, especially in the RNAi screen, has proven to be particularly facilitated by using the COL-19::GFP data in conjunction with immunolocalisation with DPY-7 antibodies.

### 6.6 HPX-1

Based on the fact that the RNAi phenotypes of *hpx-1* and *bli-3* are similar, the animal haem peroxidase-encoding gene, *hpx-1* was chosen as a target for further investigation. BLI-3 (Duox) contains a peroxidase domain, which catalyses the formation of di- and tri- tyrosine cuticle cross-links, and an NADPH-oxidase domain, which functions as a H<sub>2</sub>O<sub>2</sub> source for the

peroxidative reaction. The ability of HPX-1 to catalyse similar cross-linking reactions has been investigated. Evidence for such activity in cuticle synthesis include the facts that this gene is expressed hypodermally and coincident with the moults, and that its RNAi-targeting produces Bli and Bmd phenotypes and severe COL-19::GFP disruptions. Its RNAi-induced Mlt defect and expression at a late stage during the moult also implicate the *hpx-1* gene product in being involved in the moulting process. On the basis of its late expression and the fact that RNAi-treated animals exhibit COL-19::GFP disruption yet wild type patterns of DPY-7, the collagen substrates of HPX-1 are likely to be late expressed collagens, such as COL-19. It has been suggested that it may have a role in catalysing cross-links in the alae since these cuticle structures have high levels of cross-links, contain COL-19, and are severely disrupted upon RNAi targeting of *hpx-1*.

The functional relationship between BLI-3 and HPX-1 has been investigated. These have shown that peroxidase activity of both enzymes are necessary and are likely to represent the main mediators of tyrosine cross-linking in the cuticle. The NADPH-oxidase domain of BLI-3 is likely to be the source of H<sub>2</sub>O<sub>2</sub> for HPX-1 peroxidative activity. However, its peroxidase domain is also active.

These studies have demonstrated that HPX-1 potentially joins a repertoire of multifunctional proteins, including BLI-3 (Edens *et al.*, 2001), NAS-37 (Davis *et al.*, 2004), NHR-23/CHR (Kostrouchova *et al.*, 2001), NHR-25 (Asahina *et al.*, 2000), and CPZ-1 (Hashmi *et al.*, 2004), which are involved in both cuticle biosynthesis and the moult.

## **References**

- Ambros, V. (2000). Control of developmental timing in *Caenorhabditis elegans*. *Current Opinion in Genetics & Development* **10**, 428-433.
- Andersen, S. O. (2004). Chlorinated tyrosine derivatives in insect cuticle. *Insect Biochemistry and Molecular Biology* **34**, 1079-1087.
- Asahina, M., Ishihara, T., Jindra, M., Kohara, Y., Katsura, I., and Hirose, S. (2000). The conserved nuclear receptor Ftz-F1 is required for embryogenesis, moulting and reproduction in *Caenorhabditis elegans*. *Genes to Cells* **5**, 711-723.
- Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Ruvkun, G. (2003). Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* **421**, 268-272.
- Bächinger, H. P. (1987). The influence of peptidyl-prolyl *cis-trans* isomerase on the *in vitro* folding of type III collagen. *Journal of Biological Chemistry* **262**, 17144-17148.
- Baird, S. E., and Emmons, S. W. (1990). Properties of a class of genes required for ray morphogenesis in *Caenorhabditis elegans*. *Genetics* **126**, 335-344.
- Baldus, S., Eiserich, J. P., Mani, A., Castro, L., Figueroa, M., Chumley, P., Ma, W. X., Tousson, A., White, C. R., Bullard, D. C., Brennan, M. L., Lusic, A. J., Moore, K. P., and Freeman, B. A. (2001). Endothelial transcytosis of myeloperoxidase confers specificity to vascular ECM proteins as targets of tyrosine nitration. *Journal of Clinical Investigation* **108**, 1759-1770.
- Barstead, R. (1999). Reverse Genetics. In "C. elegans A practical approach" (I. A. Hope, Ed.).
- Bastin, P., Bagherzadeh, A., Matthews, K. R., and Gull, K. (1996). A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **77**, 235-239.
- Baulcombe, D. (2002). RNA silencing. *Current Biology* **12**, R82-R84.
- Blasko, B., Madi, A., and Fesus, L. (2003). Thioredoxin motif of *Caenorhabditis elegans* PDI-3 provides Cys and His catalytic residues for transglutaminase activity. *Biochemical and Biophysical Research Communications* **303**, 1142-1147.
- Black, R.A., and White, J.M. (1998). ADAMs: focus on the protease domain. *Current Opinion in Cell Biology* **10**, 654-659.
- Blaxter, M. (1998). *Caenorhabditis elegans* is a nematode. *Science* **282**, 2041-2046.
- Blelloch, R., Santa Anna-Arriola, S., Gao, D. L., Li, Y. J., Hodgkin, J., and Kimble, J. (1999). The gon-1 gene is required for gonadal morphogenesis in *Caenorhabditis elegans*. *Developmental Biology* **216**, 382-393.
- Bloom, L., and Horvitz, H. R. (1997). The *Caenorhabditis elegans* gene *unc-76* and its human homologs define a new gene family involved in axonal outgrowth and fasciculation. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 3414-3419.
- Boehm, M., and Bonifacino, J. S. (2001). Adaptins - The final recount. *Molecular Biology of the Cell* **12**, 2907-2920.
- Bond, J. S., and Beynon, R. J. (1995). The Astacin Family of Metalloendopeptidases. *Protein Science* **4**, 1247-1261.
- Bornstein, P. (2003). Covalent cross-links in collagen: a personal account of their discovery. *Matrix Biology* **22**, 385-391.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
- Britton, C., and Murray, L. (2002). A cathepsin L protease essential for *Caenorhabditis elegans* embryogenesis is functionally conserved in parasitic nematodes. *Molecular and Biochemical Parasitology* **122**, 21-33.
- Brodsky, B., and Ramshaw, J. A. M. (1997). The collagen triple-helix structure. *Matrix Biology* **15**, 545-554.

- Bulleid, N. J. (1996). Novel approach to study the initial events in the folding and assembly of procollagen. *Seminars in Cell & Developmental Biology* **7**, 667-672.
- Bulleid, N. J., and Freedman, R. B. (1988). Defective co-translational formation of disulfide bonds in protein disulfide-isomerase-deficient microsomes. *Nature* **335**, 649-651.
- Bundgaard, J. R., Vuust, J., and Rehfeld, J. F. (1997). New consensus features for tyrosine O-sulfation determined by mutational analysis. *Journal of Biological Chemistry* **272**, 21700-21705.
- Cabral, W. A., Chernoff, E. J., and Marini, J. C. (2001). G76E substitution in type I collagen is the first nonlethal glutamic acid substitution in the alpha 1(I) chain and alters folding of the N-terminal end of the helix. *Molecular Genetics and Metabolism* **72**, 326-335.
- Cariello, L., Velasco, P. T., Wilson, J., Parameswaran, K. N., Karush, F., and Lorand, L. (1990). Probing the Transglutaminase-Mediated, Posttranslational Modification of Proteins During Development. *Biochemistry* **29**, 5103-5108.
- Chandrashekar, R., Devarajan, E., and Mehta, K. (2002). *Dirofilaria immitis*: further characterization of the transglutaminase enzyme and its role in larval molting. *Parasitology Research* **88**, 185-191.
- Chandrashekar, R., Tsuji, N., Morales, T., Ozols, V., and Mehta, K. (1998). An ERp60-like protein from the filarial parasite *Dirofilaria immitis* has both transglutaminase and protein disulfide isomerase activity. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 531-536.
- Chou, M. Y., and Li, H. C. (2002). Genomic organization and characterization of the human type XXI collagen (COL21A1) gene. *Genomics* **79**, 395-401.
- Colige, A., Li, S. W., Sieron, A., Nusgens, B. V., Prockop, D. J., and Lapiere, C. M. (1997). cDNA cloning and expression of bovine procollagen I N-proteinase. A new member of the superfamily of zinc-metalloproteinases with binding sites for cells and other matrix components. *Matrix Biology* **16**, 124-124.
- Colige, A., Sieron, A. L., Li, S. W., Schwarze, U., Petty, E., Wertelecki, W., Wilcox, W., Krakow, D., Cohn, D. H., Reardon, W., Byers, P. H., Lapiere, C. M., Prockop, D. J., and Nusgens, B. V. (1999). Human Ehlers-Danlos syndrome type VIIC and bovine dermatosparaxis are caused by mutations in the procollagen IN-proteinase gene. *American Journal of Human Genetics* **65**, 308-317.
- Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-2018.
- Costa, M., Draper, B. W., and Priess, J. R. (1997). The role of actin filaments in patterning the *Caenorhabditis elegans* cuticle. *Developmental Biology* **184**, 373-384.
- Costa, M., Raich, W., Agbunag, C., Leung, B., Hardin, J., and Priess, J. R. (1998). A putative catenin-cadherin system mediates morphogenesis of the *Caenorhabditis elegans* embryo. *Journal of Cell Biology* **141**, 297-308.
- Cox, G. N. (1992). Molecular and biochemical aspects of nematode cuticle collagens. *Journal of Parasitology* **78**, 1-15.
- Cox, G. N., and Hirsh, D. (1985). Stage-Specific Patterns of Collagen Gene-Expression During Development of *Caenorhabditis-Elegans*. *Molecular and Cellular Biology* **5**, 363-372.
- Cox, G. N., Kusch, M., Denevi, K., and Edgar, R. S. (1981a). Temporal regulation of cuticle synthesis during development of *Caenorhabditis elegans*. *Developmental Biology* **84**, 277-285.
- Cox, G. N., Kusch, M., and Edgar, R. S. (1981b). Cuticle of *Caenorhabditis elegans*: its isolation and partial characterisation. *Journal of Cell Biology* **90**, 7-17.
- Cox, G. N., Laufer, J. S., Kusch, M., and Edgar, R. S. (1980). Genetic and phenotypic characterisation of roller mutants of *Caenorhabditis elegans*. *Genetics* **95**, 317-339.

- Cramer, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnology* **14**, 315-319.
- Crossen, R., and Gruenwald, S. (1998). Baculovirus expression vector system instruction manual. PharMingen, San Diego, CA.
- Daiyasu, H., and Toh, H. (2000). Molecular evolution of the myeloperoxidase family. *Journal of Molecular Evolution* **51**, 433-445.
- Dauner, K. M., Khan, A. U., and Steinbeck, M. J. (2000). Chlorination of pyridinium compounds - Possible role of hypochlorite, N-chloramines, and chlorine in the oxidation of pyridinoline cross-links of articular cartilage collagen type II during acute inflammation. *Journal of Biological Chemistry* **275**, 34681-34692.
- Davis, B. O., Anderson, G. L., and Dusenbery, D. B. (1982). Total Luminescence Spectroscopy of Fluorescence Changes During Aging in *Caenorhabditis-Elegans*. *Biochemistry* **21**, 4089-4095.
- Davis, M. W., Birnic, A. J., Chan, A. C., Page, A. P., and Jorgensen, E. M. (2004). A conserved metalloprotease mediates ecdysis in *Caenorhabditis elegans*. *Development* **131**, 6001-6008.
- De Deken, X., Wang, D. T., Dumont, J. E., and Miot, F. (2002). Characterization of ThOX proteins as components of the thyroid H<sub>2</sub>O<sub>2</sub>-generating system. *Experimental Cell Research* **273**, 187-196.
- De Deken, X., Wang, D. T., Many, M. C., Costagliola, S., Libert, F., Vassart, G., Dumont, J. E., and Miot, F. (2000). Cloning of two human thyroid cDNAs encoding new members of the NADPH oxidase family. *Journal of Biological Chemistry* **275**, 23227-23233.
- de Melo, J. V., de Souza, W., and Peixoto, C. A. (2003). Immunocytochemical and freeze-fracture characterization of the *Caenorhabditis elegans* DR 847 bli-1(n361) mutant which produces abnormal cuticle blisters. *Cell and Tissue Research* **312**, 229-235.
- Doerge, K. J., and Fessler, J. H. (1986). Folding of Carboxyl Domain and Assembly of Procollagen-I. *Journal of Biological Chemistry* **261**, 8924-8935.
- Domigan, N. M., Charlton, T. S., Duncan, M. W., Winterbourn, C. C., and Kettle, A. J. (1995). Chlorination of Tyrosyl Residues in Peptides by Myeloperoxidase and Human Neutrophils. *Journal of Biological Chemistry* **270**, 16542-16548.
- Edens, W. A., Sharling, L., Cheng, G. J., Shapira, R., Kinkade, J. M., Lee, T., Edens, H. A., Tang, X. X., Sullards, C., Flaherty, D. B., Benian, G. M., and Lambeth, J. D. (2001). Tyrosine cross-linking of extracellular matrix is catalyzed by Duox, a multidomain oxidase/oxidoreductase with homology to the phagocyte oxidase subunit gp91 *phox*. *Journal of Cell Biology* **154**, 879-891.
- Eiserich, J. P., Baldus, S., Brennan, M. L., Ma, W. X., Zhang, C. X., Tousson, A., Castro, L., Luscis, A. J., Nauseef, W. M., White, C. R., and Freeman, B. A. (2002). Myeloperoxidase, a leukocyte-derived vascular NO oxidase. *Science* **296**, 2391-2394.
- Engel, J., and Prockop, D. J. (1991). The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annual Review of Biophysics and Biophysical Chemistry* **20**, 137-152.
- Epstein, A. C. R., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., Mukherji, M., Metzen, E., Wilson, M. I., Dhanda, A., Tian, Y. M., Masson, N., Hamilton, D. L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P. H., Pugh, C. W., Schofield, C. J., and Ratcliffe, P. J. (2001). *C. elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54.
- Eschenlauer, S. C. P., and Page, A. P. (2003). The *Caenorhabditis elegans* ERp60 homolog PDI-3 has disulphide isomerase and transglutaminase-like cross-linking activity and is involved in the maintenance of body morphology. *Journal of Biological Chemistry* **278**, 4227-4237.

- Fessler, L. I., and Fessler, J. H. (1974). Protein assembly of procollagen and effects of hydroxylation. *Journal of Biological Chemistry* **249**, 7637-7640.
- Fetterer, R. H., and Hill, D. E. (1993). The Occurrence of Phenol Oxidase Activity in Female *Trichuris*- Suis. *Journal of Parasitology* **79**, 155-159.
- Fetterer, R. H., and Rhoads, M. L. (1990). Tyrosine-derived cross-linking amino acids in the sheath of *Haemonchus contortus* infective larvae. *Journal of Parasitology* **76**, 619-624.
- Fetterer, R. H., Rhoads, M. L., and Urban, J. F. (1993). Synthesis of tyrosine-derived cross-links in *Ascaris suum* cuticular proteins. *Journal of Parasitology* **79**, 160-166.
- Fire, A., Xu, S. Q., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.
- Flemming, A. J., Shen, Z. Z., Cunha, A., Emmons, S. W., and Leroi, A. M. (2000). Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5285-5290.
- Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408**, 325-330.
- Freeman, R. S., Hasbani, D. M., Lipscomb, E. A., Straub, J. A., and Xie, L. (2003). SM-20, EGL-9, and the EGLN family of hypoxia-inducible factor prolyl hydroxylases. *Molecules and Cells* **16**, 1-12.
- Friedman, L., Higgin, J. J., Moulder, G., Barstead, R., Raines, R. T., and Kimble, J. (2000). Prolyl 4-hydroxylase is required for viability and morphogenesis in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4736-4741.
- Fu, X. Y., Kao, J. L. F., Bergt, C., Kassim, S. Y., Huq, N. P., d'Avignon, A., Parks, W. C., Mecham, R. P., and Heinecke, J. W. (2004). Oxidative cross-linking of tryptophan to glycine restrains matrix metalloproteinase activity - Specific structural motifs control protein oxidation. *Journal of Biological Chemistry* **279**, 6209-6212.
- Fujimoto, D. (1975). Occurrence of dityrosine in cuticulin, a structural protein from *Ascaris* cuticle. *Comparative Biochemistry and Physiology* **51B**, 205-207.
- Fujimoto, D., Horiuchi, K., and Hirama, M. (1981). Isotryrosine, a new crosslinking amino acid isolated from *Ascaris* cuticle collagen. *Biochemical and Biophysical Research Communications* **99**, 637-643.
- Furthemuller, P. G., Burner, U., Jantschko, W., Regelsberger, G., and Obinger, C. (2000). Two-electron reduction and one-electron oxidation of organic hydroperoxides by human myeloperoxidase. *Febs Letters* **484**, 139-143.
- Furthemuller, P. G., Burner, U., and Obinger, C. (1998). Reaction of myeloperoxidase compound I with chloride, bromide, iodide, and thiocyanate. *Biochemistry* **37**, 17923-17930.
- Furthemuller, P. G., Jantschko, W., Regelsberger, G., and Obinger, C. (2001). Spectral and kinetic studies on eosinophil peroxidase compounds I and II and their reaction with ascorbate and tyrosine. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1548**, 121-128.
- Gallagher, L. A., and Manoil, C. (2001). *Pseudomonas aeruginosa* PAO1 kills *Caenorhabditis elegans* by cyanide poisoning. *Journal of Bacteriology* **183**, 6207-6214.
- Garcia Castineiras, S., Dillon, T., and Spector, A. (1978). Detection of bityrosine in cataractuous human lens protein. *Science* **199**, 897-899.
- Gavaret, J. M., Cahmann, H. J., and Nunez, J. (1981). Thyroid-Hormone Synthesis in Thyroglobulin - the Mechanism of the Coupling Reaction. *Journal of Biological Chemistry* **256**, 9167-9173.

- Geiszt, M., Witta, J., Baffi, J., Lckstrom, K., and Leto, T. L. (2003). Dual oxidases represent novel hydrogen peroxide sources supporting mucosal surface host defense. *FASEB Journal* **17**, U362-U375.
- Gerson, C., Sabater, J., Scuri, M., Torbati, A., Coffey, R., Abraham, J. W., Lauredo, I., Forteza, R., Wanner, A., Salathe, M., Abraham, W. M., and Conner, G. E. (2000). The lactoperoxidase system functions in bacterial clearance of airways. *American Journal of Respiratory Cell and Molecular Biology* **22**, 665-671.
- Gilleard, J. S., Barry, J. D., and Johnstone, I. L. (1997). *cis* regulatory requirements for hypodermal cell-specific expression of the *Caenorhabditis elegans* cuticle collagen gene *dpy-7*. *Molecular and Cellular Biology* **17**, 2301-2311.
- Gönczy, P., Echeverri, C., Oegema, K., Coulson, A. R., Jones, S. J. M., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S. C., Flohrs, K., Goessen, A., Leidel, S., Alleaume, A. M., Martin, C., Ozlu, N., Bork, P., and Hyman, A. A. (2000). Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* **408**, 331-336.
- Graham, P. L., Johnson, J. J., Wang, S. R., Sibley, M. H., Gupta, M. C., and Kramer, J. M. (1997). Type IV collagen is detectable in most, but not all, basement membranes of *Caenorhabditis elegans* and assembles on tissues that do not express it. *Journal of Cell Biology* **137**, 1171-1183.
- Greenberg, C. S., Birckbichler, P. J., and Rice, R. H. (1991). Transglutaminases - multifunctional cross-linking enzymes that stabilize tissues. *FASEB Journal* **5**, 3071-3077.
- Griffin, M., Casadio, R., and Bergamini, C. M. (2002). Transglutaminases: Nature's biological glues. *Biochemical Journal* **368**, 377-396.
- Guiliano, D. B., Hong, X. Q., McKerrow, J. H., Blaxter, M. L., Oksov, Y., Liu, J., Ghedin, E., and Lustigman, S. (2004). A gene family of cathepsin L-like proteases of filarial nematodes are associated with larval molting and cuticle and eggshell remodeling. *Molecular and Biochemical Parasitology* **136**, 227-242.
- Guo, X. D., Johnson, J. J., and Kramer, J. M. (1991a). Embryonic lethality caused by mutations in basement membrane collagen of *C. elegans*. *Nature* **349**, 707-709.
- Guo, X. D., Johnson, J. J., and Kramer, J. M. (1991b). Embryonic Lethality Caused by Mutations in Basement-Membrane Collagen of *C. elegans*. *Nature* **349**, 707-709.
- Guo, X. D., and Kramer, J. M. (1989). The two *Caenorhabditis elegans* basement membrane (type IV) collagen genes are located on separate chromosomes. *Journal of Biological Chemistry* **264**, 17574-17582.
- Gupta, M. C., Graham, P. L., and Kramer, J. M. (1997). Characterization of alpha 1(IV) collagen mutations in *Caenorhabditis elegans* and the effects of alpha 1 and alpha 2(IV) mutations on type IV collagen distribution. *Journal of Cell Biology* **137**, 1185-1196.
- Hartigan, N., Garrigue-Antar, L., and Kadler, K. E. (2003). Bone morphogenetic protein-1 (BMP-1) - Identification of the minimal domain structure for procollagen C-proteinase activity. *Journal of Biological Chemistry* **278**, 18045-18049.
- Hashmi, S., Britton, C., Liu, J., Guiliano, D. B., Oksov, Y., and Lustigman, S. (2002). Cathepsin L is essential for embryogenesis and development of *Caenorhabditis elegans*. *Journal of Biological Chemistry* **277**, 3477-3486.
- Hashmi, S., Tawc, W., and Lustigman, S. (2001). *Caenorhabditis elegans* and the study of gene function in parasites. *Trends in Parasitology* **17**, 387-393.
- Hashmi, S., Zhang, J., Oksov, Y., and Lustigman, S. (2004). The *Caenorhabditis elegans* cathepsin Z-like cysteine protease, Ce-CPZ-1, has a multifunctional role during the worms' development. *Journal of Biological Chemistry* **279**, 6035-6045.
- Heinecke, J. W. (2002). Tyrosyl radical production by myeloperoxidase: a phagocyte pathway for lipid peroxidation and dityrosine cross-linking of proteins. *Toxicology* **177**, 11-22.

- Heinecke, J. W., Li, W., Daehnke, H. L., and Goldstein, J. A. (1993). Dityrosine, a Specific Marker of Oxidation, Is Synthesized by the Myeloperoxidase-Hydrogen Peroxide System of Human Neutrophils and Macrophages. *Journal of Biological Chemistry* **268**, 4069-4077.
- Heldin, C. H., Miyazono, K., and tenDijke, P. (1997). TGF- $\beta$  signalling from cell membrane to nucleus through SMAD proteins. *Nature* **390**, 465-471.
- Higgins, B. J., and Hirsh, D. (1977). Roller mutants of the Nematode *Caenorhabditis elegans*. *Molecular and General Genetics* **150**, 63-72.
- Hill, K. L., Harfe, B. D., Dobbins, C. A., and L'Hernault, S. W. (2000). *dpy-18* encodes an  $\alpha$ -subunit of prolyl-4-hydroxylase in *Caenorhabditis elegans*. *Genetics* **155**, 1139-1148.
- Hishida, R., Ishihara, T., Kondo, K., and Katsura, I. (1996). *hch-1*, a gene required for normal hatching and normal migration of a neuroblast in *C. elegans*, encodes a protein related to TOLLOID and BMP-1. *EMBO Journal* **15**, 4111-4122.
- Holmgren, S. K., Bretscher, L. E., Taylor, K. M., and Raines, R. T. (1999). A hyperstable collagen mimic. *Chemistry & Biology* **6**, 63-70.
- Hulmes, D. J. S. (2002). Building collagen molecules, fibrils, and suprafibrillar structures. *Journal of Structural Biology* **137**, 2-10.
- Inoue, T., and Thomas, J. H. (2000). Targets of TGF- $\beta$  signalling in *Caenorhabditis elegans* dauer formation. *Developmental Biology* **217**, 192-204.
- Jenkins, C. L., and Raines, R. T. (2002). Insights on the conformational stability of collagen. *Natural Product Reports* **19**, 49-59.
- Johnstone, I. L. (1994). The cuticle of the nematode *Caenorhabditis elegans*: a complex collagen structure. *Bioessays* **16**, 171-178.
- Johnstone, I. L. (2000). Cuticle collagen genes: expression in *Caenorhabditis elegans*. *Trends in Genetics* **16**, 21-27.
- Johnstone, I. L., and Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development. *EMBO Journal* **15**, 3633-3639.
- Johnstone, I. L., Shafi, Y., and Barry, J. D. (1992). Molecular analysis of mutations in the *Caenorhabditis elegans* collagen gene *dpy-7*. *EMBO Journal* **11**, 3857-3863.
- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231-237.
- Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G., and Ahringer, J. (2000). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biology* **2**, 1-10.
- Kapyla, J., Jaalinoja, J., Tulla, M., Ylostalo, J., Nissinen, L., Viitasalo, T., Vehvilainen, P., Marjomaki, V., Nykvist, P., Saamanen, A. M., Farndale, R. W., Birk, D. E., Ala-Kokko, L., and Heino, J. (2004). The fibril-associated collagen IX provides a novel mechanism for cell adhesion to cartilaginous matrix. *Journal of Biological Chemistry* **279**, 51677-51687.
- Kassner, A., Tiedemann, K., Notbohm, H., Ludwig, T., Morgelin, M., Reinhardt, D. P., Chu, M. L., Bruckner, P., and Grassel, S. (2004). Molecular structure and interaction of recombinant human type XVI collagen (vol 339, pg 835, 2004). *Journal of Molecular Biology* **342**, 1675-1675.
- Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Bizinger, A., Wylie, B. N., and Davidson, G. S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* **293**, 2087-2092.

- Kim, T. H., Hwang, S. B., Jeon, P. Y., Lee, J. H., and Cho, J. W. (2005). Requirement of tyrosylprotein sulfotransferase-A for proper cuticle formation in the nematode *C. elegans*. *FEBS Letters* **579**, 53-58.
- Kingston, B. I. (1991). Nematode collagen genes. *Parasitology Today* **7**, 11-15.
- Kivirikko, K. I. (1998). Collagen biosynthesis: a mini-review cluster. *Matrix Biology* **16**, 355-356.
- Kivirikko, K. I., and Myllyharju, J. (1998). Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biology* **16**, 357-368.
- Kivirikko, K. I., and Myllylä, R. (1982). Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. *Methods in Enzymology* **82**, 245-304.
- Kivirikko, K. I., Myllylä, R., and Pihlajaniemi, T. (1992). Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In, *Post translational modification of proteins* (Harding, J.J. and Crabe, M. J. C., eds) CRC Press, Boca Raton, pp. 1-51.
- Kivirikko, K. I., and Pihlajaniemi, T. (1998). Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. *Advances in Enzymology and Related Areas of Molecular Biology* **72**, 325-400.
- Knight, C. G., Patel, M. N., Azevedo, R. B. R., and Leroi, A. M. (2002). A novel mode of ecdysozoan growth in *Caenorhabditis elegans*. *Evolution & Development* **4**, 16-27.
- Ko, F. C. F., and Chow, K. L. (2002). A novel thioredoxin-like protein encoded by the *C. elegans dpy-11* gene is required for body and sensory organ morphogenesis. *Development* **129**, 1185-1194.
- Koh, K., and Rothman, J. H. (2001). ELT-5 and ELT-6 are required continuously to regulate epidermal seam cell differentiation and cell fusion in *C-elegans*. *Development* **128**, 2867-2880.
- 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.
- Kostrouchova, M., Krause, M., Kostrouch, Z., and Rall, J. E. (2001). Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7360-7365.
- Kramer, J. M. (1994a). Genetic-analysis of extracellular-matrix in *C. elegans*. *Annual Review of Genetics* **28**, 95-116.
- Kramer, J. M. (1994b). Structures and functions of collagens in *Caenorhabditis elegans*. *FASEB Journal* **8**, 329-336.
- Kramer, J. M. (1997). Extracellular matrix. In, *C. elegans II*, (Riddle, D. L., Blumenthal, T., Meyer, B. J., Priess, J. R., eds) Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, pp. 471-500.
- Kramer, J. M., Cox, G. N., and Hirsh, D. (1985a). Expression of the *Caenorhabditis elegans* collagen genes *col-1* and *col-2* is developmentally regulated. *Journal of Biological Chemistry* **260**, 1945-1951.
- Kramer, J. M., Cox, G. N., and Hirsh, D. (1985b). Expression of the *Caenorhabditis elegans* Collagen Genes *Col-1* and *Col-2* Is Developmentally Regulated. *Journal of Biological Chemistry* **260**, 1945-1951.
- Kramer, J. M., French, R. P., Park, E. C., and Johnson, J. J. (1990). The *Caenorhabditis elegans rol-6* gene, which interacts with the *sqt-1* collagen gene to determine organismal morphology, encodes a collagen. *Molecular and Cellular Biology* **10**, 2081-2089.
- Kramer, J. M., and Johnson, J. J. (1993). Analysis of mutations in the *sqt-1* and *rol-6* collagen genes of *Caenorhabditis elegans*. *Genetics* **135**, 1035-1045.

- Kramer, J. M., Johnson, J. J., Edgar, R. S., Basch, C., and Roberts, S. (1988). The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. *Cell* **55**, 555-565.
- Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991). Mutations in Collagen Genes - Causes of Rare and Some Common Diseases in Humans. *FASEB Journal* **5**, 2052-2060.
- Kuno, K., Baba, C., Asaka, A., Matsushima, C., Matsushima, K., and Hosono, R. (2002). The *Caenorhabditis elegans* ADAMTS family gene *adt-1* is necessary for morphogenesis of the male copulatory organs. *Journal of Biological Chemistry* **277**, 12228-12236.
- Kuno, K., and Matsushima, K. (1998). ADAMTS-1 protein anchors at the extracellular matrix through the thrombospondin type I motifs and its spacing region. *Journal of Biological Chemistry* **273**, 13912-13917.
- Kusch, M., and Edgar, R. S. (1986). Genetic studies of unusual loci that affect body shape in the nematode *Caenorhabditis elegans* and may code for cuticle structural proteins. *Genetics* **113**, 621-639.
- Kypreos, K. E., Birk, D., Trinkaus-Randall, V., Hartmann, D. J., and Sonenshein, G. E. (2000). Type V collagen regulates the assembly of collagen fibrils in cultures of bovine vascular smooth muscle cells. *Journal of Cellular Biochemistry* **80**, 146-155.
- Lam, G. T., Jiang, C. A., and Thummel, C. S. (1997). Coordination of larval and prepupal gene expression by the DHR3 orphan receptor during *Drosophila* metamorphosis. *Development* **124**, 1757-1769.
- Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T., and Kivirikko, K. I. (1996). Characterization of human type III collagen expressed in a baculovirus system. Production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit. *Journal of Biological Chemistry* **271**, 11988-11995.
- Lambeth, J. D., Cheng, G. J., Arnold, R. S., and Edens, W. A. (2000). Novel homologs of gp91phox. *Trends in Biochemical Sciences* **25**, 459-461.
- Larios, J. M., Budhiraja, R., Fanburg, B. L., and Thannickal, V. J. (2001). Oxidative protein cross-linking reactions involving L-tyrosine in transforming growth factor-beta 1-stimulated fibroblasts. *Journal of Biological Chemistry* **276**, 17437-17441.
- Lassandro, F., Sebastiano, M., Zei, F., and Bazzicalupo, P. (1994). The role of dityrosine formation in the cross-linking of *cut-2*, the product of a second cuticlin gene of *Caenorhabditis elegans*. *Molecular and Biochemical Parasitology* **65**, 147-159.
- Lee, S. S., Lee, R. Y. N., Fraser, A. G., Kamath, R. S., Ahringer, J., and Ruvkun, G. (2003). A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. *Nature Genetics* **33**, 40-48.
- Levy, A. D., Yang, J., and Kramer, J. M. (1993). Molecular and genetic analyses of the *Caenorhabditis elegans* *dpy-2* and *dpy-10* collagen genes: a variety of molecular alterations affect organismal morphology. *Molecular Biology of the Cell* **4**, 803-817.
- Li, S. W., Sieron, A. L., Fertala, A., Hojima, Y., Arnold, W. V., and Prockop, D. J. (1996). The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 5127-5130.
- Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fang, C., Huang, C.-C., and Kain, S. R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. *Journal of Biological Chemistry* **273**, 34970-34975.
- Liu, L. X., Spoerke, J. M., Mulligan, E. L., Chen, J., Reardon, B., Westlund, B., Sun, L., Abel, K., Armstrong, B., Hardiman, G., King, J., McCague, L., Basson, M., Clover, R., and Johnson, C. D. (1999). High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Research* **9**, 859-867.

- Liu, Z. C., Kirch, S., and Ambros, V. (1995). The *Caenorhabditis elegans* heterochronic gene pathway controls stage-specific transcription of collagen genes. *Development* **121**, 2471-2478.
- Madi, A., Punyiczki, M., Di Rao, M., Piacentini, M., and Fesus, L. (1998). Biochemical characterization and localization of transglutaminase in wild-type and cell-death mutants of the nematode *Caenorhabditis elegans*. *European Journal of Biochemistry* **253**, 583-590.
- Maduzia, L. L., Gumienny, T. L., Zimmerman, C. M., Wang, H., Shetgiri, P., Krishna, S., Roberts, A. F., and Padgett, R. W. (2002). lon-1 regulates *Caenorhabditis elegans* body size downstream of the dbl-1 TGF beta signaling pathway. *Developmental Biology* **246**, 418-428.
- Martin, J. L. (1995). Thioredoxin - a Fold for All Reasons. *Structure* **3**, 245-250.
- Matsui, M., Oshima, M., Oshima, H., Takaku, K., Maruyama, T., Yodoi, J., and Taketo, M. M. (1996). Early embryonic lethality caused by targeted disruption of the mouse thioredoxin gene. *Developmental Biology* **178**, 179-185.
- Matsuki, M., Yamashita, F., Ishida-Yamamoto, A., Yamada, K., Kinoshita, C., Fushiki, S., Ueda, E., Morishima, Y., Tabata, K., Yasuno, H., Hashida, M., Iizuka, H., Ikawa, M., Okabe, M., Kondoh, G., Kinoshita, T., Takeda, J., and Yamanishi, K. (1998). Defective stratum corneum and early neonatal death in mice lacking the gene for transglutaminase 1 (keratinocyte transglutaminase). *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1044-1049.
- Mazzorana, M., Cogne, S., Goldschmidt, D., and Aubert-Foucher, E. (2001). Collagenous sequence governs the trimeric assembly of collagen XII. *Journal of Biological Chemistry* **276**, 27989-27998.
- Mazzorana, M., Gruffat, H., Sergeant, A., and Vanderrest, M. (1993). Mechanisms of Collagen Trimer Formation - Construction and Expression of a Recombinant Minigene in HeLa-Cells Reveals a Direct Effect of Prolyl Hydroxylation on Chain Assembly of Type-Xii Collagen. *Journal of Biological Chemistry* **268**, 3029-3032.
- McLaughlin, S. H., and Bulleid, N. J. (1998a). Molecular recognition in procollagen chain assembly. *Matrix Biology* **16**, 369-377.
- McLaughlin, S. H., and Bulleid, N. J. (1998b). Thiol-independent interaction of protein disulphide isomerase with type X collagen during intra-cellular folding and assembly. *Biochemical Journal* **331**, 793-800.
- McMahon, L., Muriel, J. M., Roberts, B., Quinn, M., and Johnstone, I. L. (2003). Two sets of interacting collagens form functionally distinct substructures within a *Caenorhabditis elegans* extracellular matrix. *Molecular Biology of the Cell* **14**, 1366-1378.
- Mehta, K., Rao, U. R., Vickery, A. C., and Bireckbichler, P. J. (1990). Significance of transglutaminase-catalyzed reactions in growth and development of filarial parasite, *Brugia malayi*. *Biochemical and Biophysical Research Communications* **173**, 1051-1057.
- Mehta, K., Rao, U. R., Vickery, A. C., and Fesus, L. (1992). Identification of a novel transglutaminase from the filarial parasite *Brugia malayi* and its role in growth and development. *Molecular and Biochemical Parasitology* **53**, 1-15.
- Mei, D., Woo, W. M., and Chisholm, A. D. (2004). The cytoskeleton and epidermal morphogenesis in *C. elegans*. *Experimental Cell Research* **301**, 84-90.
- Melacini, G., Bonvin, A., Goodman, M., Boclens, R., and Kaptein, R. (2000). Hydration dynamics of the collagen triple helix by NMR. *Journal of Molecular Biology* **300**, 1041-1048.
- 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.

- Mochii, M., Yoshida, S., Morita, K., Kohara, Y., and Ueno, N. (1999). Identification of transforming growth factor- $\beta$ -regulated genes in *Caenorhabditis elegans* by differential hybridization of arrayed cDNAs. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 15020-15025.
- Mohrlen, F., Hutter, H., and Zwillig, R. (2003). The astacin protein family in *Caenorhabditis elegans*. *European Journal of Biochemistry* **270**, 4909-4920.
- Montgomery, M. K., Xu, S. Q., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 15502-15507.
- Morita, K., Chow, K. L., and Ueno, N. (1999). Regulation of body length and male tail ray pattern formation of *Caenorhabditis elegans* by a member of TGF- $\beta$  family. *Development* **126**, 1337-1347.
- Morita, K., Flemming, A. J., Sugihara, Y., Mochii, M., Suzuki, Y., Yoshida, S., Wood, W. B., Kohara, Y., Leroi, A. M., and Ueno, N. (2002). A *Caenorhabditis elegans* TGF- $\beta$ , DBL-1, controls the expression of LON-1, a PR-related protein, that regulates polyploidization and body length. *EMBO Journal* **21**, 1063-1073.
- Morrison, M., and Schonbaum, G. R. (1976). Peroxidase-catalyzed halogenation. *Annual Review of Biochemistry* **45**, 861-888.
- Muriel, J. M., Brannan, M., Taylor, K., Johnstone, I. L., Lithgow, G. J., and Tuckwell, D. (2003). M142.2 (cut-6), a novel *Caenorhabditis elegans* matrix gene important for dauer body shape. *Developmental Biology* **260**, 339-351.
- Murthy, S. N. P., Wilson, J., Guy, S. L., and Lorand, L. (1991). Intramolecular Cross-Linking of Monomeric Fibrinogen by Tissue Transglutaminase. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10601-10604.
- Myers, J. C., Li, D. Q., Amenta, P. S., Clark, C. C., Nagaswami, C., and Weisel, J. W. (2003). Type XIX collagen purified from human umbilical cord is characterized by multiple sharp kinks delineating collagenous subdomains and by intermolecular aggregates via globular, disulfide-linked, and heparin-binding amino termini. *Journal of Biological Chemistry* **278**, 32047-32057.
- Myllyharju, J. (2003). Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biology* **22**, 15-24.
- Myllyharju, J., and Kivirikko, K. I. (2001). Collagens and collagen-related diseases. *Annals of Medicine* **33**, 7-21.
- Myllyharju, J., and Kivirikko, K. I. (2004). Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends in Genetics* **20**, 33-43.
- Myllyharju, J., Kukkola, L., Winter, A. D., and Page, A. P. (2002). The exoskeleton collagens in *Caenorhabditis elegans* are modified by prolyl 4-hydroxylase with unique combinations of subunits. *Journal of Biological Chemistry* **277**, 29187-29196.
- Nagai, N., Hosokawa, M., Itohara, S., Adachi, E., Matsushita, T., Hokawa, N., and Nagata, K. (2000). Embryonic lethality of molecular chaperone Hsp47 knockout mice is associated with defects in collagen biosynthesis. *Journal of Cell Biology* **150**, 1499-1505.
- Nguyen, C. Q., Hall, D. H., Yang, Y., and Fitch, D. H. A. (1999). Morphogenesis of the *Caenorhabditis elegans* male tail tip. *Developmental Biology* **207**, 86-106.
- Nomura, K., Hoshino, K., and Suzuki, N. (1999). The primary and higher order structures of sea urchin ovoperoxidase as determined by cDNA cloning and predicted by homology modeling. *Archives of Biochemistry and Biophysics* **367**, 173-184.
- Nomura, K., Suzuki, N., and Matsumoto, S. (1990). Pulcherosine, a Novel Tyrosine-Derived, Trivalent Cross-Linking Amino-Acid from the Fertilization Envelope of Sea-Urchin Embryo. *Biochemistry* **29**, 4525-4534.

- Norman, K. R., and Moerman, D. G. (2000). The *let-268* locus of *Caenorhabditis elegans* encodes a procollagen lysyl hydroxylase that is essential for type IV collagen secretion. *Developmental Biology* **227**, 690-705.
- Novelli, J., Ahmed, S., and Hodgkin, J. (2004). Gene interactions in *Caenorhabditis elegans* define DPY-31 as a candidate procollagen C-proteinase and SQT-3/ROL-4 as its predicted major target. *Genetics* **168**, 1259-1273.
- Nystrom, J., Shen, Z. Z., Aili, M., Flemming, A. J., Leroi, A., and Tuck, S. (2002). Increased or decreased levels of *Caenorhabditis elegans* lon-3, a gene encoding a collagen, cause reciprocal changes in body length. *Genetics* **161**, 83-97.
- O'Brich, P. J. (2000). Peroxidases. *Chemico-Biological Interactions* **129**, 113-139.
- Pace, J. M., Chitayat, D., Atkinson, M., Wilcox, W. R., Schwarze, U., and Byers, P. H. (2002). A single amino acid substitution (D1441Y) in the carboxyl-terminal propeptide of the pro alpha (1) chain of type I collagen results in a lethal variant of osteogenesis imperfecta with features of dense bone diseases. *Journal of Medical Genetics* **39**, 23-29.
- Pachucki, J., Wang, D., Christophe, D., and Miot, F. (2004). Structural and functional characterization of the two human ThOX/Duox genes and their 5'-flanking regions. *Molecular and Cellular Endocrinology* **214**, 53-62.
- Padgett, R. W., Das, P., and Krishna, S. (1998). TGF- $\beta$  signaling, Smads, and tumor suppressors. *BioEssays* **20**, 382-390.
- Page, A. P. (1997). Cyclophilin and protein disulphide isomerase genes are co-transcribed in a functionally related manner in *Caenorhabditis elegans*. *DNA and Cell Biology* **16**, 1335-1343.
- Page, A. P. (2001). The nematode cuticle: synthesis, modification and mutants. In *Parasitic nematodes* (Kennedy, M.W. and Harnett, W., eds) CABI Press, Wallingford, Oxon, UK, pp. 167-193.
- Page, A. P., MacNiven, K., and Hengartner, M. O. (1996). Cloning and biochemical characterization of the cyclophilin homologues from the free-living nematode *Caenorhabditis elegans*. *Biochemical Journal* **317**, 179-185.
- Page, A. P., and Winter, A. D. (2003). Enzymes involved in the biogenesis of the nematode cuticle. In "Advances in Parasitology, Vol 53", Vol. 53, pp. 85-148.
- Parise, G., and Bazzicalupo, P. (1997). Assembly of nematode cuticle: role of hydrophobic interactions in *cut-2* cross-linking. *Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology* **1337**, 295-301.
- Park, Y. S., and Kramer, J. M. (1994). The *C. elegans* *sqt-1* and *rol-6* collagen genes are coordinately expressed during development, but not at all stages that display mutant phenotypes. *Developmental Biology* **163**, 112-124.
- Peixoto, C. A., Alves, L. C., de Melo, J. V., and de Souza, W. (2000). Ultrastructural analyses of the *Caenorhabditis elegans* *sqt-1(sc13)* left roller mutant. *Journal of Parasitology* **86**, 269-274.
- Peixoto, C. A., de Melo, J. V., Kramer, J. M., and de Souza, W. (1998). Ultrastructural analyses of the *Caenorhabditis elegans* *rol-6 (su1006)* mutant, which produces abnormal cuticle collagen. *Journal of Parasitology* **84**, 45-49.
- Peters, K., McDowall, J., and Rose, A. M. (1991). Mutations in the *bli-4* (I) locus of *Caenorhabditis elegans* disrupt both adult cuticle and early larval development. *Genetics* **129**, 95-102.
- Plasterk, R. H. A. (2002). RNA silencing: The genome's immune system. *Science* **296**, 1263-1265.
- Poole, C. B., Jin, J. M., and McReynolds, L. A. (2003). Cloning and biochemical characterization of blisterase, a subtilisin-like convertase from the filarial parasite, *Onchocerca volvulus*. *Journal of Biological Chemistry* **278**, 36183-36190.

- Porter, S., Clark, I. M., Kevorkian, L., and Edwards, D. R. (2005). The ADAMTS metalloproteinases. *Biochemical Journal* **386**, 15-27.
- Pothof, J., van Haften, G., Thijssen, K., Kamath, R. S., Fraser, A. G., Ahringer, J., Plasterk, R. H. A., and Tijsterman, M. (2003). Identification of genes that protect the *C. elegans* genome against mutations by genome-wide RNAi. *Genes & Development* **17**, 443-448.
- Priess, J. R., and Hirsh, D. I. (1986). *Caenorhabditis elegans* morphogenesis: the role of the cytoskeleton in elongation of the embryo. *Developmental Biology* **117**, 156-173.
- Prockop, D. J. (1998). What holds us together? Why do some of us fall apart? What can we do about it? *Matrix Biology* **16**, 519-528.
- Prockop, D. J., and Kivirikko, K. I. (1984). Heritable Diseases of Collagen. *New England Journal of Medicine* **311**, 376-386.
- Prockop, D. J., and Kivirikko, K. I. (1995). Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry* **64**, 403-434.
- Prockop, D. J., Sieron, A. L., and Li, S. W. (1998). Procollagen N-proteinase and procollagen C-proteinase. Two unusual metalloproteinases that are essential for procollagen processing probably have important roles in development and cell signaling. *Matrix Biology* **16**, 399-408.
- Qian, M. W., Eaton, J. W., and Wolff, S. P. (1997). Cyanate-mediated inhibition of neutrophil myeloperoxidase activity. *Biochemical Journal* **326**, 159-166.
- Rattenhol, A., Pappano, W. N., Koch, M., Keene, D. R., Kadler, K. E., Sasaki, T., Timpl, R., Burgeson, R. E., Greenspan, D. S., and Bruckner-Tuderman, L. (2002). Proteinases of the bone morphogenetic protein-1 family convert procollagen VII to mature anchoring fibril collagen. *Journal of Biological Chemistry* **277**, 26372-26378.
- Reichenberger, E., and Olsen, B. R. (1996). Collagens as organizers of extracellular matrix during morphogenesis. *Seminars in Cell & Developmental Biology* **7**, 631-638.
- Richer, J. K., Hunt, W. G., Sakanari, J. A., and Grieve, R. B. (1993). *Dirofilaria immitis* - Effect of Fluoromethyl Ketone Cysteine Protease Inhibitors on the 3rd-Stage to 4th-Stage Molt. *Experimental Parasitology* **76**, 221-231.
- Riihimaa, P., Nissi, R., Page, A. P., Winter, A. D., Keskiäho, K., Kivirikko, K. I., and Myllyharju, J. (2002). Egg shell collagen formation in *Caenorhabditis elegans* involves a novel prolyl 4-hydroxylase expressed in spermatheca and embryos and possessing many unique properties. *Journal of Biological Chemistry* **277**, 18238-18243.
- Rougvie, A. E., and Ambros, V. (1995). The Heterochronic Gene *Lin-29* Encodes a Zinc-Finger Protein That Controls a Terminal Differentiation Event in *Caenorhabditis elegans*. *Development* **121**, 2491-2500.
- Sadek, C. M., Jimenez, A., Damdimopoulos, A. E., Kieselbach, T., Nord, M., Gustafsson, J. A., Spyrou, I., Davis, E. C., Oko, R., van der Hoorn, F. A., and Miranda-Vizuete, A. (2003). Characterization of human thioredoxin-like 2 - A novel microtubule-binding thioredoxin expressed predominantly in the cilia of lung airway epithelium and spermatid manchette and axoneme. *Journal of Biological Chemistry* **278**, 13133-13142.
- Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y., and Prockop, D. J. (1973). Synthesis of (Pro-Hyp-Gly)<sub>n</sub> of defined molecular weights. Evidence for the stabilisation of collagen triple helix by hydroxyproline. *Biochimica Et Biophysica Acta* **303**, 198-202.
- Sakura, S., and Fujimoto, D. (1984). Absorption and Fluorescence Study of Tyrosine-Derived Crosslinking Amino-Acids from Collagen. *Photochemistry and Photobiology* **40**, 731-734.
- Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sun, C. Y., Baird, S. E., and Padgett, R. W. (1996). *Caenorhabditis elegans* genes *sma2*, *sma-3*, and *sma-4* define a conserved family

- of transforming growth factor beta pathway components. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 790-794.
- Savage-Dunn, C., Tokarz, R., Wang, H., Cohen, S., Giannikas, C., and Padgett, R. W. (2000). SMA-3 Smad has specific and critical functions in DBL-1/SMA-6 TGF- $\beta$  -related signaling. *Developmental Biology* **223**, 70-76.
- Schofield, D. S., Uitto, J., and Prockop, D. J. (1974). Formation of interchain disulphide bonds and helical structure during biosynthesis of procollagen by embryonic tendon cells. *Biochemistry* **13**, 1801-1806.
- Sebastiano, M., Lassandro, F., and Bazzicalupo, P. (1991). *cut-1* a *Caenorhabditis elegans* gene coding for a dauer-specific noncollagenous component of the cuticle. *Developmental Biology* **146**, 519-530.
- Shaw, L. M., and Olsen, B. R. (1991). FACIT collagens: diverse molecular bridges in extracellular matrices. *Trends in Biochemical Sciences* **16**, 191-194.
- Shelton, C. A., Carter, J. C., Ellis, G. C., and Bowerman, B. (1999). The nonmuscle myosin regulatory light chain gene *mlc-4* is required for cytokinesis, anterior-posterior polarity, and body morphology during *Caenorhabditis elegans* embryogenesis. *Journal of Cell Biology* **146**, 439-451.
- Sibley, M. H., Johnson, J. J., Mello, C. C., and Kramer, J. M. (1993). Genetic identification, sequence, and alternative splicing of the *Caenorhabditis elegans*  $\alpha 2(IV)$  collagen gene. *Journal of Cell Biology* **123**, 255-264.
- Silhankova, M., Jindra, M., and Asahina, M. (2005). Nuclear receptor NHR-25 is required for cell-shape dynamics during epidermal differentiation in *Caenorhabditis elegans*. *Journal of Cell Science* **118**, 223-232.
- Simmer, F., Moorman, C., van der Linden, A. M., Kuijk, E., van den Berghc, P. V. E., Kamath, R. S., Fraser, A. G., Ahringer, J., and Plasterk, R. H. A. (2003). Genome-wide RNAi of *C. elegans* using the hypersensitive *rrf-3* strain reveals novel gene functions. *PLOS Biology* **1**, 77-84.
- Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J., and Plasterk, R. H. A. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Current Biology* **12**, 1317-1319.
- Simske, J. S., and Hardin, J. (2001). Getting into shape: epidermal morphogenesis in *Caenorhabditis elegans* embryos. *BioEssays* **23**, 12-23.
- Singh, N., and Sulston, J. E. (1978). Some observations on molting in *C. elegans*. *Nematologica* **24**, 63-71.
- Sritunyalucksana, K., Wongsuebsantati, K., Johansson, M. W., and Soderhall, K. (2001). Peroxinectin, a cell adhesive protein associated with the proPO system from the black tiger shrimp, *Penaeus monodon*. *Developmental and Comparative Immunology* **25**, 353-363.
- Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Developmental Biology* **56**, 110-156.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology* **100**, 64-119.
- Sulston, J.E., and Hodgkin, J. (1988). Methods "The Nematode *Caenorhabditis elegans*. *Wood WB and the community of C. elegans Researchers (ed)*, Cold Spring Harbor Laboratory, 587-606.
- Suzuki, N., Labosky, P. A., Furuta, Y., Hargett, L., Dunn, R., Fogo, A. B., Takahara, K., Peters, D. M. P., Greenspan, D. S., and Hogan, B. L. M. (1996). Failure of ventral body wall closure in mouse embryos lacking a procollagen C-proteinase encoded by *Bmp1*, a mammalian gene related to *Drosophila* *tolloid*. *Development* **122**, 3587-3595.

- Suzuki, Y., Morris, G. A., Han, M., and Wood, W. B. (2002). A cuticle collagen encoded by the *lon-3* gene may be a target of TGF- $\beta$  signaling in determining *Caenorhabditis elegans* body shape. *Genetics* **162**, 1631-1639.
- Swan, K. A., Curtis, D. E., McKusick, K. B., Voinov, A. V., Mapa, F. A., and Cancilla, M. R. (2002). High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Research* **12**, 1100-1105.
- Takahara, K., Lyons, G. E., and Greenspan, D. S. (1994). Bone morphogenetic protein-1 and a mammalian tolloid homolog (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues. *Journal of Biological Chemistry* **269**, 32572-32578.
- Taurog, A. (1999). Molecular evolution of thyroid peroxidase. *Biochimie* **81**, 557-562.
- Thacker, C., Peters, K., Srayko, M., and Rose, A. M. (1995). The *bli-4* locus of *Caenorhabditis elegans* encodes structurally distinct kex2/subtilisin-like endoproteases essential for early development and adult morphology. *Genes & Development* **9**, 956-971.
- Thacker, C., and Rose, A. M. (2000). A look at the *Caenorhabditis elegans* kex2/subtilisin-like proprotein convertase family. *BioEssays* **22**, 545-553.
- Thacker, C., Srayko, M., and Rose, A. M. (2000a). Mutational analysis of *bli-4/kpc-4* reveals critical residues required for proprotein convertase function in *C. elegans*. *Gene* **252**, 15-25.
- Thacker, C., Srayko, M., and Rose, A. M. (2000b). Mutational analysis of *bli-4/kpc-4* reveals critical residues required for proprotein convertase function in *C. elegans*. *Gene* **252**, 15-25.
- Thein, M. C., McCormack, G., Winter, A. D., Johnstone, I. L., Shoemaker, C. B., and Page, A. P. (2003). *Caenorhabditis elegans* exoskeleton collagen COL-19: An adult-specific marker for collagen modification and assembly, and the analysis of organismal morphology. *Developmental Dynamics* **226**, 523-539.
- Tien, M. (1999). Myeloperoxidase-catalyzed oxidation of tyrosine. *Archives of Biochemistry and Biophysics* **367**, 61-66.
- Timmons, L., Court, D. L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**, 103-112.
- Timmons, L., Tabara, H., Mello, C. C., and Fire, A. Z. (2003). Inducible systemic RNA silencing in *Caenorhabditis elegans*. *Molecular Biology of the Cell* **14**, 2972-2983.
- Trent, C., Tsung, N., and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. *Genetics* **104**, 619-647.
- Turner, E., Somers, C. E., and Shapiro, B. M. (1985). The relationship between a novel NAD(P)H oxidase activity of ovoperoxidase and the CN<sup>-</sup>-resistant respiratory burst that follows fertilization of sea urchin eggs. *Journal of Biological Chemistry* **260**, 13163-13171.
- van der Keyl, H., Kim, H., Espey, R., Oke, C. V., and Edwards, M. K. (1994). *Caenorhabditis elegans* *sqt-3* mutants have mutations in the *col-1* collagen gene. *Developmental Dynamics* **201**, 86-94.
- vanderVliet, A., Eiserich, J. P., Halliwell, B., and Cross, C. E. (1997). Formation of reactive nitrogen species during peroxidase-catalyzed oxidation of nitrite - A potential additional mechanism of nitric oxide-dependent toxicity. *Journal of Biological Chemistry* **272**, 7617-7625.
- Vissers, M. C. M., and Thomas, C. (1997). Hypochlorous acid disrupts the adhesive properties of subendothelial matrix. *Free Radical Biology and Medicine* **23**, 401-411.

- Vitagliano, L., Berisio, R., Mazzarella, L., and Zagari, A. (2001). Structural bases of collagen stabilization induced by proline hydroxylation (vol 58, pg 459, 2001). *Biopolymers* **59**, 465-465.
- von Mende, N., Bird, D. M., Albrct, P. S., and Riddle, D. L. (1988). *dpy-13*: a nematode collagen gene that affects body shape. *Cell* **55**, 567-576.
- Walmsley, A. R., Batten, M. R., Lad, U., and Bulleid, N. J. (1999). Intracellular retention of procollagen within the endoplasmic reticulum is mediated by prolyl 4-hydroxylase. *Journal of Biological Chemistry* **274**, 14884-14892.
- Wang, D. T., De Deken, X., Milenkovic, M., Song, Y., Pirson, I., Dumont, J. E., and Miot, F. (2005). Identification of a novel partner of Duox - EFP1, a thioredoxin-related protein. *Journal of Biological Chemistry* **280**, 3096-3103.
- Wardle, F. C., Welch, J. V., and Dale, L. (1999). Bone morphogenetic protein 1 regulates dorsal-ventral patterning in early *Xenopus* embryos by degrading chordin, a BMP4 antagonist. *Mechanisms of Development* **86**, 75-85.
- Wen, C. H., Metzstein, M. M., and Greenwald, I. (1997). SUP-17, a *Caenorhabditis elegans* ADAM protein related to *Drosophila* KUZBANIAN, and its role in LIN-12/NOTCH signalling. *Development* **124**, 4759-4767.
- White, J. (1988). The anatomy. In, *The nematode Caenorhabditis elegans* (Wood, W. B., ed) Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, pp. 81-122.
- Wilson, R., Freddi, S., and Bateman, J. F. (2002). Collagen X chains harboring Shmid metaphyseal chondrodysplasia NCI domain mutations are selectively retained and degraded in stably transfected cells. *Journal of Biological Chemistry* **277**, 12516-12524.
- Winter, A. D., and Page, A. P. (2000). Prolyl 4-hydroxylase is an essential procollagen-modifying enzyme required for exoskeleton formation and the maintenance of body shape in the nematode *Caenorhabditis elegans*. *Molecular and Cellular Biology* **20**, 4084-4093.
- Woessner, J. F., and Nagase, H. (2000). "Matrix Metalloproteinases and TIMPs." Oxford University Press,
- Xu, Y. J., Bhate, M., and Brodsky, B. (2002). Characterization of the nucleation step and folding of a collagen triple-helix peptide. *Biochemistry* **41**, 8143-8151.
- Yang, J., and Kramer, J. M. (1994). *In vitro* mutagenesis of *Caenorhabditis elegans* cuticle collagens identifies a potential subtilisin-like protease cleavage site and demonstrates that carboxyl domain disulfide bonding is required for normal function but not assembly. *Molecular and Cellular Biology* **14**, 2722-2730.
- Yang, J., and Kramer, J. M. (1999). Proteolytic processing of *Caenorhabditis elegans* SQT-1 cuticle collagen is inhibited in right roller mutants whereas cross-linking is inhibited in left roller mutants. *Journal of Biological Chemistry* **274**, 32744-32749.
- Yingling, J. M., Das, P., Savage, C., Zhang, M., Padgett, R. W., and Wang, X. F. (1996). Mammalian dwarfs are phosphorylated in response to transforming growth factor beta and are implicated in control of cell growth. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 8940-8944.
- Yochem, J., Gu, T., and Han, M. (1998). A new marker for mosaic analysis in *Caenorhabditis elegans* indicates a fusion between *hyp6* and *hyp7*, two major components of the hypodermis. *Genetics* **149**, 1323-1334.
- Yochem, J., Tuck, S., Greenwald, I., and Han, M. (1999). A gp330/megalin-related protein is required in the major epidermis of *Caenorhabditis elegans* for completion of molting. *Development* **126**, 597-606.
- Yoshida, S., Morita, K., Mochii, M., and Ueno, N. (2001). Hypodermal expression of *Caenorhabditis elegans* TGF- $\beta$  type I receptor SMA-6 is essential for the growth and maintenance of body length. *Developmental Biology* **240**, 32-40.