Collagen processing and cuticle formation is catalysed by the astacin metalloprotease DPY-31 in free-living and parasitic nematodes.

Gillian Stepek, Gillian McCormack and Antony P. Page*

Division of Infection and Immunity, Institute of Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Bearsden Road, Glasgow, G61 1QH

*Correspondence author: a.page@vet.gla.ac.uk; Telephone: 44 141 330 1997: Fax: 44 141 330 5603

Note: Nucleotide sequence data reported in this paper are available in the GenBank™ database under the accession numbers FJ812517 (*H. contortus nas-35*) and FJ812518 (*B. malayi nas-35*). Supplementary data is associated with this article.

ABSTRACT

The exoskeleton or cuticle performs many key roles in the development and survival of all nematodes. This structure is predominantly collagenous in nature and requires numerous enzymes to properly fold, modify, process and crosslink these essential structural proteins. The cuticle structure and its collagen components are conserved throughout the nematode phylum but differ from the collagenous matrices found in vertebrates. This structure, its formation and the enzymology of nematode cuticle collagen biogenesis have been elucidated in the free-living nematode *Caenorhabditis elegans*. The *dpy-31* gene in *C. elegans* encodes a procollagen C-terminal processing enzyme of the astacin metalloprotease or bone morphogenetic protein class that, when mutated, results in a temperature-sensitive lethal phenotype associated with cuticle defects. In this study, orthologues of this essential gene have been identified in the phylogenetically diverse parasitic nematodes *Haemonchus contortus* and *Brugia malayi*. The DPY-31 protein is expressed in the gut and secretory system of *C. elegans*, a location also confirmed when a *B. malayi* transcriptional *dpy-31* promoter reporter gene fusion was expressed in *C. elegans*. Functional conservation between the nematode enzymes was supported by the fact that heterologous expression of the *H. contortus dpy-31* orthologue in a *C. elegans dpy-31* mutant resulted in the full rescue of the mutant body form. This interspecies conservation was established further when the recombinant nematode enzymes were found to have a similar range of inhibitable protease activities. In addition, the recombinant DPY-31 enzymes from both *H. contortus* and *B. malayi* were shown to efficiently process the *C. elegans* cuticle collagen SQT-3 at the correct C-terminal procollagen processing site.

Keywords: *Caenorhabditis elegans*, *Haemonchus contortus*, *Brugia malayi*, astacin metalloprotease, BMP, development, cuticle

1. Introduction

All nematodes are encased in a protective exoskeleton known as the cuticle. This complex extracellular matrix (ECM) is synthesised repeatedly, through a process called moulting, to form larval- and adult- specific cuticles that permit growth and form a protective barrier. The major components of this structure are the highly cross-linked small collagen-like proteins, which are modified by a variety of biosynthetic enzymes (Page and Winter, 2003; Page and Johnstone, 2007; Thein et al., 2009). The N-terminus of these proteins contains 80-150 amino acids of non-repetitive sequence, preceding the signature Gly-X-Y repetitive domain followed by a conserved subtilisin-like pro-domain cleavage site (Page and Johnstone, 2007). A non-repetitive region is also present at the C-terminus, following the Gly-X-Y repeat domain and contains an astacin-like, bone morphogenetic protein (BMP) processing domain (Novelli, 2006). The C-terminal non-repetitive region and its flanking cysteine residues are highly conserved between *H. contortus* and *C. elegans* collagens and hence equivalent molecules in the two nematodes probably share a similar function (Shamansky et al., 1989). Cuticle collagen genes represent large families in all nematodes examined and have a high degree of similarity in gene size and structure, further suggesting a common function and a similar mechanism for their biogenesis throughout the nematode phylum (Page and Johnstone, 2007).

Proteases are essential for the viability of parasitic nematodes, performing crucial functions, such as cuticle moulting, host tissue penetration and digestion. There are several classes of proteases found in nematodes: cysteine, serine, aspartic and metalloproteases. Important developmental roles performed by the metalloprotease class of enzymes include hatching, cuticle collagen processing and cuticle moulting in

67 *Caenorhabditis elegans* (Hishida et al., 1996; Davis et al., 2004; Novelli et al., 2004),
68 activation of the free-living to parasitic stage of *Ancylostoma caninum* (Hawdon et al.,
69 1995), feeding or host tissue penetration of *Trichuris suis* (Hill et al., 1993), digestion in
70 *Ancylostoma caninum* (Jones and Hotez, 2002), tissue penetration by *Strongyloides*
71 *stercoralis* (Gallego et al., 2005) and ecdysis of *Haemonchus contortus* (Gamble et al.,
72 1989; Gamble, 1996). There are several families of metalloproteases in nematodes, with
73 the astacin metalloproteases belonging to the M12A family (Möhrlen et al., 2003).
74 Astacin metalloproteases are structurally distinct zinc metallo-endopeptidases, include the
75 bone morphogenetic proteins and are characterised by two conserved motifs in the
76 catalytically-active astacin domain: the catalytic active-site
77 (HExxHxxGFxHExxRxDRD), for binding the essential Zn^{2+}, and the methionine-turn
78 (SxMHY), which maintains the enzyme conformation. The active-site zinc is penta-
79 coordinated and is bound by the three histidine residues and the glutamic acid residue in
80 the active-site and the tyrosine residue in the methionine-turn (Stöcker et al., 1993).
81 There are 39 Nematode AStacin (NAS) metalloproteases in the free-living
82 nematode, *C. elegans*, divided into six subgroups (I to VI) depending on the C-terminal
83 domain structure, which is believed to determine the function of the enzyme (Möhrlen et
84 al., 2003). The enzymes in subgroup V (NAS-33 to NAS-38) have a unique nematode-
85 specific domain arrangement, in that they have a signal peptide, a prodomain, the N-
86 terminal catalytic astacin domain, and the C-terminal EGF, CUB and TSP-1 domains.
87 These C-terminal domains are important for the regulation of the proteolytic activity and
88 are commonly found in proteins that play critical developmental roles (Möhrlen et al.,
89 2003). The subgroup V enzymes perform important developmental functions in *C.*

90 *elegans*. Nematodes with mutations in the *nas-34* gene display a delayed HatCHing
91 phenotype, Hch-1 (Hishida et al., 1996). Mutations in the *nas-35* gene result in a severe
92 DumPY appearance, hence the name Dpy-31, and this enzyme has been proven to be an
93 essential procollagen C-peptidase involved in proper cuticle formation (Novelli et al.,
94 2004). Nematodes with mutations in *nas-36* and *nas-37*, but particularly the *nas-37* gene,
95 show defects in the ecdysis step of moulting and subsequent shedding of the cuticle
96 (Davis et al., 2004; Suzuki et al., 2004). An unidentified protease from the exsheathing
97 fluid of *H. contortus* infective larvae has previously been shown to stimulate ecdysis by
98 promoting the escape of the L3 stage from the L2 sheath through the formation of an
99 anterior refractile ring and removable cap structure (Gamble et al., 1989; Gamble, 1996).
100 An identical ring structure was formed on isolated *H. contortus* larval cuticles in the
101 presence of recombinant *C. elegans* NAS-37, suggesting that similar specific astacin
102 substrates are shared between these diverse nematode species (Davis et al., 2004). It is
103 reasonable to predict that the remaining nematode-specific, subgroup V astacin
104 metalloproteases, that have essential roles in the hatching and collagen processing in *C.*
105 *elegans*, may likewise be functionally conserved in parasitic nematodes and may
106 therefore represent potential vaccine and drug targets.
107 In this study, we investigate the role that the procollagen C-peptidase DPY-31
108 (NAS-35) plays in cuticle formation and identify and characterize orthologues of this
109 enzyme from the parasitic nematodes *H. contortus* and *Brugia malayi*. In *C. elegans*,
110 DPY-31 plays an essential role in SQT-3 cuticle collagen processing and normal cuticle
111 formation (Novelli et al., 2004; Novelli, 2006). Mutations in the gene encoding this

enzyme affect post-embryonic viability and have profound effects on the cuticle structure and nematode morphogenesis.

2. Materials and Methods

2.1. Nematode strains

The wild-type Bristol N2 and *dpy-31(e2770)* strains of *C. elegans* were provided by the *Caenorhabditis* Genetics Centre, University of Minnesota and Prof. Jonathan Hodgkin, University of Oxford, respectively. *Haemonchus contortus* adults were provided by Dr. Frank Jackson (Moredun Research Institute) and *Brugia malayi* adults by Prof. Rick Maizels (University of Edinburgh).

2.2. Preparation of genomic DNA, RNA and cDNA from C. elegans, H. contortus *and* B. malayi *worms*

Genomic DNA was isolated from adult nematodes using a standard protocol involving homogenization in Proteinase K, followed by repeated phenol:chloroform extraction. Briefly, worms were homogenised in six volumes of lysis buffer containing 100µg/ml Proteinase K and incubated at 65^{o}C for 4hrs. Worm debris was removed by centrifugation and the DNA was purified by repeated phenol:chloroform and chloroform extractions, ethanol precipitated and resuspended in TE Buffer, pH 8.0. The DNA was treated with a final concentration of 100µg/ml RNAse A for 1hr at 37^{o}C, phenol:chloroform extracted, chloroform extracted, ethanol precipitated and resuspended in TE Buffer, pH 8.0. Total RNA was isolated from adult nematodes following Trizol (Invitrogen) extraction, and cDNA was prepared using the AffinityScript Multiple

Temperature cDNA synthesis kit (Stratagene), following the manufacturer's recommendations, with 1µg RNA per reaction with oligo-dT primer.

2.3. Isolation of the cDNA and genomic DNA of H. contortus *dpy-31*

The *C. elegans* DPY-31 protein sequence was used to BLAST search the *H. contortus* database (http://www.sanger.ac.uk/Projects/H_contortus/), initially using the options labelled "assembled contigs (27/01/06)" and "sequence reads (01/08/05)". The sequence reads haem-479f01.p1k and haem-479f01.q1k, and contigs 049443 and 037116 had the highest homology score, and were put into a Scaffold program designed by Dr. Robin Beech, McGill University, which assembles a group of physically linked sequences using the *Haemonchus* genome project databases. Primers were designed from the reads, haem-479f01.p1k and haem-479f01.q1k, and used to PCR, clone and sequence the genomic DNA of *dpy-31* from *H. contortus*, using *PfuTurbo* or *PfuUltra* polymerases. Primers were then designed to sequence the full 2715bp PCR product. From the Scaffold results, one BAC clone, HaemApoBac 18h16, was identified as containing part of the *dpy-31* gene. Further reads were identified from the initial Scaffold procedure upstream of the 2715bp clone and primers were designed for sequencing the genomic PCR products between the new read pairs. A contig was formed between the three cloned PCR products and the BAC clone, HaemApoBac 18h16, using ContigExpress, and potential intron-exon splice sites were predicted using the GeneWise tool on the ExPASy proteomics site (http://us.expasy.org/). This predicted partial coding sequence was translated using the Translate tool on the ExPASy proteomics site, and then aligned with the *C. elegans* protein sequence using ClustalX and BoxShade

158 (http://www.ch.embnet.org/software/BOX_form.html). Amplification of the 5′ and 3′
159 ends of the *H. contortus dpy-31* cDNA was performed using the Invitrogen 5′- and 3′-
160 RACE systems with the following primers: Hc nas-35 5′RT, Hc nas-35 5′R1, Hc nas-35
161 5′R2, Hc nas-35 3′F1 and Hc nas-35 3′F2. The complete coding sequence was formed
162 from the 5′ and 3′ RACE products and the original partial sequence in ContigExpress.
163 The primers, Hc nas-35 F and Hc nas-35 R were used to PCR the genomic sequence of *H.*
164 *contortus dpy-31*, and the 5′ and 3′ unknown genomic regions were sequenced using the
165 primers Hc nas-35gF2, Hc nas-35gR2 and Hc nas-35gR3. Gene Structure Draw was used
166 to produce a scaled schematic depicting the positions of the introns and exons in the gene.
167 The relationship between the contigs and PCR fragments is depicted in the supplemental
168 Figure 1, together with all primer sequences reported (supplemental Table 1).
169
170 *2.4. Isolation of the cDNA and genomic DNA of* B. malayi *dpy-31*
171 A BLAST search was performed using the *C. elegans* DPY-31 protein in the *B.*
172 *malayi* genome database (http://www.tigr.org/tdb/e2k1/bma1/), and a homologous gene
173 was found on TIGR assembly 14973 (locus identifier: Bm1_41035). Amplification of the
174 5′ and 3′ ends of the predicted coding sequence using the Invitrogen 5′ and 3′-RACE
175 systems and the primers Bm nas-35RT, Bm nas-35R1, Bm nas-35R2, Bm35R3, Bm nas-
176 35F1 and Bm nas-35F2 revealed the actual start codon 4098bp upstream of the *in silico*
177 prediction. The complete sequence was confirmed using PCR and the primers, Bm nas-
178 35F and Bm nas-35R. Intron/ exon site prediction, translations and protein alignment
179 were performed as described above. Primer sequences can be found in supplemental table
180 2.

181

182 *2.5. Generation of* H. contortus *and* B. malayi *dpy-31 rescue constructs*

183 A 2067bp *Sal*I-*Pst*I *C. elegans dpy-31* promoter was generated by PCR from *C.*
184 *elegans* N2 genomic DNA using the primers, Ce35promF and Ce35promR, and cloned
185 into pBlueScript SK-. A 709bp *Not*I-*Sac*II *C. elegans dpy-31* 3´-UTR was generated by
186 PCR from *C. elegans* N2 genomic DNA using the primers, Ce35utrF and Ce35utrR and
187 cloned into the *Ce-dpy-31* promoter-pBlueScript construct to create the plasmid p*nas-35*.

188 The coding sequences of the *H. contortus* and *B. malayi dpy-31* genes (two
189 alternatively-spliced forms for *B. malayi*) were isolated by PCR using the primers, Hc35
190 PstIF and Hc35b NotIR, Bm35 PstIF and Bm35a NotIR or Bm35 PstIF and Bm35b
191 NotIR, and *PfuUltra* polymerase to generate products of 1647bp, 1836bp and 1800bp,
192 respectively. These PCR products were cloned into pCR-TOPO2.1, and a synthetic intron
193 (5´-gtaagtttaaactattcgttactaactaactttaaacatttaaattttcag-3´) was inserted by ligation of a
194 double-stranded oligo into an *Ale*I (*H. contortus*) or *Afe*I (*B. malayi*) blunt-ended
195 restriction site. These products were then cloned into the *Pst*I-*Not*I digested p*nas-35*
196 vector to create *H. contortus* and *B. malayi dpy-31* rescue constructs (Figure 3H). The
197 correct sequences of these plasmids were confirmed over the cloning and synthetic intron
198 junctions using the primers, Ce35p inF, Ce35u inR and Hc35 SIF for the *H. contortus*
199 constructs, and Ce35p inF, Ce35u inR and Bm35 SIF for the *B. malayi* constructs.

200 The *dpy-31* rescue constructs were microinjected into the syncytial gonad of *C.*
201 *elegans dpy-31(e2770)* mutant nematodes at a concentration of 10µg/ml, together with
202 5µg/ml of a p*dpy-7*:GFP fusion construct or 25µg/ml p*cpr-5*:GFP fusion construct and
203 pBlueScript SK- to make the final concentration up to 150µg/ml. Transformants were

selected by GFP (Green Fluorescent Protein) fluorescence. Four transgenic lines (with the *dpy-7* marker) and one transgenic line (with the *cpr-5* marker) were obtained for the *Hc-dpy-31* construct. In addition, single-worm PCR was performed for three GFP positive and three GFP negative worms per line, and for three N2 worms and three *dpy-31(e2770)* mutant worms, using the rescue primers, Hc35 PstIF and Hc35b NotIR. All denoted primer sequences can be found in supplemental Table 3.

The *Bmdpy-31* constructs failed to rescue the *dpy-31(e2770)* mutant phenotype at 10µg/ml with either 5µg/ml of a p*dpy-7*:GFP fusion construct or 25µg/ml p*cpr-5*:GFP fusion construct, and only dead GFP positive animals were obtained. In order to establish if the transgene was being expressed in the transgenic lines the *Bmdpy-31* construct was introduced into the wild type N2 strain at 10µg/ml with 5µg/ml of a p*dpy-7*:GFP and six transgenic lines were established. mRNA was made from the transgenic N2 lines following the protocol described in 2.2 and PCR was performed for GFP positive, N2 worms and *dpy-31(e2770)* mutants to confirm that the rescue plasmid was functional and the transgene was expressed (Supplemental Figure 2). Male nematodes from these lines were then used to introduce the transgene into the *dpy-31(e2770)* mutant background via crossing. Four independent GFP positive transgenic lines were established, of which all failed to rescue the mutant phenotype.

2.6. Generation of B. malayi dpy-31 *promoter-reporter transgenic lines*

Primers, Bm35Fp and Bm35Rp, were designed from the *B. malayi* genomic sequence isolated above and used to amplify a 2130bp promoter fragment from *B. malayi* genomic DNA (-2026 to +5 relative to ATG start). This promoter fragment was cloned

227 into a *Pst*I-*BamH*I digested reporter gene vector, pPD96:04 (Addgene), and sequences
228 over the restriction sites were checked using the primers, Bm35 inFp and Bm35 inRp.
229 This reporter construct was microinjected into the syncytial gonad of *C. elegans* N2
230 nematodes, together with pRF4 (*rol-6*) marker plasmid at 100µg/ml each. Four transgenic
231 lines were identified and examined for reporter gene expression by staining
232 glutaraldehyde-fixed worms for β-galactosidase activity and examining live worms for
233 GFP expression. Primer sequences in supplemental Table 4.
234
235 *2.7. Microscopy*
236 Live nematodes, mounted on 1% agarose/ 0.6% azide pads, or fixed nematodes,
237 were viewed under Differential Interference Contrast (DIC) or fluorescence (GFP) optics
238 on a Zeiss Axioskop2 microscope, and images were taken using an AxioCam camera and
239 Axiovision software.
240
241 *2.8. Antibodies*
242 Anti-TY tag antibodies were applied in Western blots as described previously
243 (Bastin et al., 1996; Thein et al., 2009). The following peptide sequences were used to
244 raise antibodies in rabbits following fusion to keyhole limpet haemocyanin (CovaLabs):
245 *C. elegans* DPY-31 NH_2-CYMDKLNKLADEKHPEEIE- $CONH_2$; *B. malayi* DPY-31
246 NH_2-CEKAKTFGQSAEEIQK- $CONH_2$; *H. contortus* DPY-31 NH_2-
247 CESYNKDSPKNEAYKWRKQ- $CONH_2$. Antibodies were affinity purified and eluted
248 from the corresponding peptide column prior to use.
249

250 *2.9. Antibody staining of* C. elegans *wild-type worms*

251 Wild type *C. elegans* worms were incubated with either a pre-bleed or a rabbit
252 anti-DPY-31 antibody then anti-rabbit Alexa fluor 488 secondary antibody (Molecular
253 probes) described following standard published methods (Thein et al., 2009).

254

255 *2.10. Recombinant expression of* H. contortus *and* B. malayi *DPY-31*

256 The sequences encoding the mature proteins (without signal peptide and
257 prodomain) of *H. contortus* and *B. malayi* DPY-31 were cloned by PCR using the
258 primers rHc35matF and rHc35bmatR, or Bm35matF and Bm35amatR, and *PfuUltra*
259 polymerase, to generate products of 1199bp and 1202bp, respectively. These PCR
260 products were first cloned into pCR-TOPO2.1, then into the *Kpn*I-*Pst*I digested pQE30
261 vector to create protein expression constructs. All constructs were sequenced to confirm
262 authenticity prior to transformation in *E. coli* M15 (pREP4) cells. Expression of the
263 encoded protein was induced by an adapted autoinduction method (Studier, 2005).
264 Briefly, 10ml cultures of MDG medium containing 100μg/ml ampicillin and 25μg/ml
265 kanamycin were inoculated overnight with these clones, and then this non-induced
266 culture was used to inoculate 500ml of ZYM-5052 medium for 40 hrs at 25^{o}C. The cells
267 expressing the protein were collected after centrifugation at 5500rpm, 4^{o}C for 20 mins
268 and resuspended in native lysis buffer, pH 8.0 containing 10mM imidazole and 1mg/ml
269 lysozyme. The lysed cells were sonicated and the soluble protein was collected by
270 centrifugation at 11500rpm, 4^{o}C for 30 mins. Purification of the proteins from the soluble
271 cell lysates was performed using Ni-NTA resin columns at 4^{o}C under native conditions

and the fractions were analysed by SDS-PAGE and western blotting using affinity purified anti-DPY-31 antibodies. Primers sequences are found in supplemental Table 5.

2.11. Astacin activity assay

An astacin assay was performed to determine the activity of recombinant DPY-31 from both *H. contortus* and *B. malayi* using an adapted method (Gamble et al., 1989; Stöcker, 1995; Gamble, 1996). Briefly, 100nM of recombinant enzyme was incubated with 0.3mM $ZnCl_2$ and 0, 1, 2, 5 or 10mM 1, 10-phenanthroline, and the volume was made up to a total of 100µl with 50mM HEPES, pH 8.0 buffer. Incubation occurred for 1hr at 25^{o}C prior to the addition of 250µM Suc-Ala-Ala-Ala-pNA (Bachem) and subsequent incubation at 25^{o}C for a further 2hrs. Absorbance was measured on a plate reader at 450nm. Each sample was performed in triplicate.

2.12. SQT-3 cleavage assay

A SQT-3 cleavage assay was performed to determine the activity of recombinant DPY-31 from both *H. contortus* and *B. malayi*. Primers were designed to amplify the *sqt-3* (F23H12.4) coding sequence from *C. elegans* mixed-stage N2 cDNA (Sqt-3 cDNA PstI F1 and Sqt-3 cDNA XbaI R1) using *Pfu* polymerase. This PCR product was cloned into pCRScript (Stratagene) and then sub-cloned into pVL1392 using *Pst*I and *Xba*I to produce a *sqt-3* construct with no TY tag (plasmid 1). The pCRScript clone was also sub-cloned into the Stratagene vector pBC KSP using *Pst*I and *Xba*I, which was then linearly digested by *Asp*EI prior to insertion of a Ty-tag. The Ty-tag was prepared using primers Ty-sense sqt-3 and Ty-anti sqt-3 and ligated into the *Asp*EI site of the *sqt-3*/pBC plasmid.

This *sqt-3* construct, with an N-term Ty-tag, was finally sub-cloned into pVL1392 using *Pst*I and *Xba*I (plasmid 2). Plasmid 1 above was used as the DNA template for amplification with primers Sqt-3 cDNA PstI F1 and Sqt-3 RTyC with *Pfu* polymerase, and the subsequent PCR product was cloned into pCRScript prior to sub-cloning into pVL1392 using *Pst*I to produce a *sqt-3* construct with a C-term Ty-tag (plasmid 3). These three SQT-3 constructs were then expressed in Sf9 insect cells prior to harvesting and freezing at -80°C by Johanna Myllyharju (University of Oulu, Finland). These constructs had been TY-tagged in different positions: 1) SQT-3 with no TY tag; 2) SQT-3 with the TY tag between the Gly-X-Y repeats; and 3) SQT-3 with the TY tag at the C-terminus in a position after the predicted cleavage site for DPY-31. The SQT-3 proteins were collected following homogenization of the insect cells in 0.1M NaCl, 0.1M glycine, 0.1% Tx100, 10μM DTT in 10mM Tris pH7.7 at 4°C. For the cleavage assay, 4μM of recombinant DPY-31 or BSA was incubated with an equivalent amount of each of the three SQT-3 proteins, in the presence and absence of 0.3mM $ZnCl_2$, and the volume was made up to a total of 20μl with dH_2O. These samples were incubated at room temperature for approximately 24hrs, boiled for 5 mins at 100^oC and then analysed by SDS-PAGE and western blotting using an anti-TY tag antibody. Primer sequences are found in supplemental Table 6.

3. RESULTS

3.1. Identification and characterization of H. contortus *and* B. malayi dpy-31 *homologues*

Using a combination of *H. contortus* database BLAST hits, scaffold assembly, 5′– and 3′–RACE and full length PCR confirmation, a single 1647bp *dpy-31* cDNA was

identified for this species. The BLAST search of the *B. malayi* database revealed that a single homologous gene was present; however, 5´-RACE detected an alternative start site 4098bp upstream of the *in silico* prediction. A comparison of the gene structure of the nematode *dpy-31* homologues is depicted in Table 1 and Figure 1. Predicted transcripts for each nematode *dpy-31* homologue indicated that there are two 3´ alternatively spliced forms (A and B). However, whereas both isoforms were experimentally isolated from *C. elegans* (following exon VII) and *B. malayi* (following exon VIII), the second *dpy-31* isoform (predicted to follow exon XIV) could not be amplified from *H. contortus* adult cDNA, but we have not excluded the possibility that it may be expressed in the larval stages. The *C. elegans dpy-31b* form contains seven exons and *B. malayi dpy-31b* contains eight exons (data not shown). The full-length cDNAs for *dpy-31b* in *C. elegans* and *B. malayi* are 1590bp and 1836bp and encode polypeptides of 529 and 611 amino acids, respectively.

Signal peptide cleavage sites were predicted using SignalP and the prodomain cleavage sites by the ProP program for the nematode DPY-31 homologues (Figure 2). Thus, the predicted mature proteins for DPY-31 consist of 399, 466 and 388 amino acids for *H. contortus*, *C. elegans* and *B. malayi*, respectively (Table 1), having an identity of 66.2% between Ce-DPY-31 and Hc-DPY-31, 51.3% between Ce-DPY-31 and Bm-DPY-31 and 56.2% between Hc-DPY-31 and Bm-DPY-31 (Figure 2). The mature DPY-31 proteins from the three nematodes range in size from 45 to 53kDa, with pIs in the range of 7.22 to 8.17 (Table 1). The mature proteins for DPY-31B are 403 and 401 amino acids for *C. elegans* and *B. malayi*, respectively, and share 55.9% identity.

340 The nematode DPY-31 homologues have an N-terminal catalytic astacin domain
341 (Pfam PF01400), containing the crucial zinc-binding site, which is characterised by the
342 sequence HExxHxxGFxHExxRxDRD, and a conserved methionine-turn (SxMHY), of
343 which the tyrosine residue forms one of the five important ligands essential to bind the
344 active-site zinc. In addition to the essential catalytic domain, all DPY-31 proteins encode
345 successive EGF (Epidermal Growth Factor), CUB (C1r/C1s, embryonic sea urchin
346 protein Uegf, Bmp-1) and TSP-1 (Thrombospondin type-1 repeat) domains, except the
347 homologue from *B. malayi*, which lacks a TSP-1 domain (Figures 1 and 2). There are two
348 putative N-linked glycosylation sites in DPY-31 from the three nematode species: Asn-
349 Ile-Thr and Asn-Ser-Thr (Figure 2).
350
351 *3.2. Complementation of* dpy-31 *mutant worms with a* H. contortus dpy-31 *but not a* B.
352 malayi dpy-31 *construct*
353 The *C. elegans dpy-31(e2770)* mutant is a temperature sensitive larval lethal
354 strain that is inviable at 20°C and is barely viable at 15°C, with all surviving worms
355 exhibiting a strong recessive Dpy phenotype (Novelli et al., 2004). This phenotype arises
356 due to a single GC-AT point mutation that results in a Leu-Pro transition in the CUB
357 domain (Novelli et al., 2004) (Figure 2). This mutant strain was transformed, via
358 microinjection, with a *Hcdpy-31* cDNA rescue construct, together with the GFP-tagged
359 transformation markers *dpy-7* or *cpr-5*, resulting in five independent rescue lines raised at
360 the non-permissive temperature of 20°C. These lines were assessed microscopically for
361 the presence or absence of body morphology defects. Figure 3A shows a wild-type N2
362 worm, while Figure 3B shows the larval lethal phenotype of the *dpy-31(e2770)* mutant

363 worm raised at 20°C. These images were compared to Figures 3C–F to indicate that the
364 *Hcdpy-31* construct was able to rescue the *dpy-31(e2770)* mutant animals at 20°C, a
365 result independent of which reporter transgene (the *dpy-7* hypodermal or the *cpr-5* gut
366 marker) was applied. Single worm PCR demonstrated that only the GFP positive worms
367 for each transgenic line amplified the correct product of 1647bp for the *Hc-dpy-31*
368 transgene (Figure 3G), confirming that rescue of the *dpy-31(e2770)* worms correlated
369 with the presence of the *Hc-dpy-31* transgenic construct.

370 Rescue attempts with both of the *B. malayi dpy-31* constructs (*a* and *b*) failed to
371 complement the *dpy-31(e2770)* mutant. Four independent *Bm-dpy-31* transgenic lines
372 were obtained and the presence of the transgene confirmed by single worm PCR
373 however, restoration of the wild type body form was not achieved and larval lethality was
374 evident (data not shown). In an attempt to establish that the transgene was expressed the
375 *Bmdpy-31* construct was first introduced into a wild type (N2) background, transgenic
376 (GFP positive) males were produced and used to cross the transgene into the *dpy-*
377 *31(e2770)* mutant strain. Following mRNA extraction, PCR confirmed that the transgene
378 was expressed in these N2 transgenic lines (supplementary Figure 2), but the *Bmdpy-31*
379 constructs also failed to rescue the associated lethality when crossed into the *dpy-*
380 *31(e2770)* mutant strain.

381

382 *3.3. Recombinant expression of* H. contortus *and* B. malayi *active DPY-31 enzymes*

383 The mature *H. contortus* and *B. malayi* DPY-31 enzymes, minus the prodomains,
384 were cloned into the pQE30 expression vector and the histidine-tagged recombinant
385 proteins were expressed following an autoinduction method. The Ni-NTA column

386 purified 50kDa proteins were detected by affinity purified anti DPY-31 antibodies for
387 both nematode species (results not shown).

388 The purified recombinant proteins were found to be active in an astacin activity
389 assay against a synthetic substrate (Figure 4). This activity was progressively inhibited by
390 1, 10-phenanthroline (Figure 4) suggesting that DPY-31 from the parasitic nematodes, *H.*
391 *contortus* and *B. malayi*, are indeed specific, functionally-active, zinc metalloproteases.

392 Through genetic characterisation, it was previously demonstrated that *C. elegans*
393 DPY-31 is a procollagen C-proteinase, which cleaves the essential cuticle collagen, SQT-
394 3 (Novelli et al., 2004; Novelli, 2006) (Figure 5C). To examine the functional activity of
395 recombinant DPY-31 enzymes from *H. contortus* and *B. malayi,* a specific assay was
396 developed to measure the cleavage of a recombinantly expressed *C. elegans* cuticle
397 collagen, SQT-3. These constructs were TY-tagged in different positions: untagged SQT-
398 3; SQT-3 with the internal TY tag between the Gly-X-Y repeats (TY1 in Figure 5C); and
399 SQT-3 with a C-terminal TY tag after the predicted cleavage site for DPY-31 (TY2 in
400 Figure 5C). The *H. contortus* DPY-31 recombinant enzyme, but not the negative control
401 protein BSA, was demonstrated to specifically cleave SQT-3 at the predicted C-terminal
402 site (Figure 5A), activity that was found to be dependent on the inclusion of 0.3mM
403 $ZnCl_2$ (data not shown). The *B. malayi* enzyme (Figure 5B) produced identical cleavage
404 profiles of the SQT-3-TY tagged substrates to that of the *H. contortus* recombinant DPY-
405 31. The specificity of this assay is highlighted by the fact that the anti-TY tag antibody
406 does not detect the untagged SQT-3 protein but reacts to the internal TY tagged SQT-3,
407 even after DPY-31 cleavage. DPY-31 cleavage of SQT-3 was confirmed by the removal
408 of the C-terminal TY epitope, as shown by the dramatic decrease in reactivity to this

epitope (Figures 5A and 5B, lane 3) compared to incubation with a control protein (Figure 5A, lane 6). This result supported the previous genetic suppressor screens for *Cedpy-31*, which indicated that the cleavage site for DPY-31 is indeed at the C-terminus of the SQT-3 molecule and demonstrates a conservation of enzyme function between *C. elegans* and parasitic nematodes.

3.4. Heterologous expression of B. malayi dpy-31 *promoter-reporter in* C. elegans

C. elegans wild-type nematodes were transformed with a *Bmdpy-31* promoter-reporter construct, resulting in four semi-stable transgenic lines. This reporter expresses both *gfp* and *LacZ* genes together with a nuclear localization motif that permits promoter fusions to be localized by virtue of nuclear localised GFP and β-galactosidase activity, although the GFP signals tend to diffuse and do not remain nuclear localized. Expression of BmDPY-31 is predominantly restricted to the gut cells of *C. elegans*, as demonstrated by GFP (Figure 6B) and β-galactosidase (Figure 6C) expression patterns. This expression pattern was consistent throughout all life-cycle stages, including embryos (results not shown), larvae (Figure 6C) and adults (Figure 6B). *Bmdpy-31* promoter/reporter was also expressed in the gland cells of the pharynx in *C. elegans* via β-galactosidase staining (arrow in Figure 6D). The *C. elegans* N2 strain was also examined for the expression of DPY-31 following immunolocalization with a *C. elegans* affinity purified anti-DPY-31 peptide antibody (Figures 6E-H). As for the transgenic reporter fusion, the expression of CeDPY-31 was primarily located in the gut cells (arrow in Figure 6H), but was also found in the excretory duct of the excretory system when antibodies were applied to fixed nematodes (arrow in Figure 6F). The antibody expression pattern was confirmed to be

specific following the inability of the rabbit pre-immunization sera to highlight any of the aforementioned tissues (data not shown). The GFP, β-galactosidase and antibody expression patterns are therefore consistent with a nematode enzyme that is expressed in the gut and the glandular excretory system.

4. DISCUSSION

This paper describes the detailed identification and biochemical characterisation of the nematode procollagen C-peptidase DPY-31 from the parasitic nematodes *H. contortus* and *B. malayi*. The highest homology was noted between *C. elegans* and *H. contortus* and is consistent with both species belonging to Clade V, whereas *B. malayi* is a more distantly related, Clade III nematode (Blaxter et al., 1998). The size of the coding sequences between the three nematode species is approximately equivalent, but more numerous and larger introns are present in the *dpy-31* genes from the parasitic nematodes. This presence of a greater number of introns in *H. contortus* compared to the *C. elegans* homologue has previously been noted in a number of genes (Redmond et al., 2001; Stepek et al., 2009) and is contradictory to the estimated genome size of *H. contortus*, being 20Mb smaller than the 100Mb genome of *C. elegans* (Mitreva et al., 2005). This may suggest that there are either smaller intergenic regions or fewer genes present in the *H. contortus* genome and/or a higher level of redundancy of some genes in *C. elegans*. The former is indeed the case for the plant parasitic nematode, *Meloidogyne hapla*, which has a relatively small genome of 54Mb that consequently encodes fewer genes than that of *C. elegans* (Opperman et al., 2008). The latter point is however supported by the analysis of the Cys-loop ligand gated ion channel family (Williamson et al., 2007) where

455 a large, potentially redundant, family is present in *C. elegans* compared to either *B.*
456 *malayi* or *Trichinella spiralis*.

457 The *H. contortus dpy-31* rescue construct successfully complemented the
458 temperature sensitive phenotype of the *C. elegans dpy-31*(*e2770*) mutant allele by
459 returning the wild-type body form at the normally non-permissive temperature, a result
460 that supports the conservation of function for the encoded developmentally-essential
461 metalloprotease between these nematodes. No rescue was obtained with either isoform of
462 the *B. malayi* homologue, an observation that may relate to the increased phylogenetic
463 distance between *B. malayi* and *C. elegans* (Blaxter et al., 1998). However, there is at
464 least one published example of a *B. malayi bli-5* kunitz-domain encoding gene rescuing
465 the corresponding *C. elegans bli-5* mutant (Stepek et al., 2009). Alternatively, the failure
466 of the *B. malayi* homologue to complement the *C. elegans* mutant may be the result of
467 technical difficulties, although we were able to detect the expression of the *B. malayi*
468 transgene in several wild type *C. elegans* lines that were ultimately used to cross the
469 transgene into the *dpy-31* mutant.

470 The *dpy-31(e2770)* mutant allele contains a point mutation in the CUB domain
471 (Novelli et al., 2004), an essential domain that has also been shown to be involved in the
472 secretion and catalytic activity of vertebrate BMP-1 (Hartigan et al., 2003). An intact
473 CUB domain that contained the conserved leucine residue, which is mutant in *dpy-*
474 *31(e2270),* was present in the *Hcdpy-31* rescue construct. It is interesting to note that only
475 one DPY-31 transcript could be amplified from *H. contortus* adult cDNA, although *in*
476 *silico* analysis predicted a potential second form with an alternative final intron/exon
477 splice site. Both isoforms (A and B) are identifiable and amplified from both *C. elegans*

and *B. malayi*, suggesting that both genes are authentic. It may however be the case that the second *dpy-31* isoform is exclusively expressed in one or more of the larval stages, and hence is absent from adult stages of *H. contortus.*

Expression of DPY-31 appears to be similar to that of other secretory proteases, particularly those with roles in nematode development, with prominent expression in the pharyngeal gland cells, the excretory duct of the excretory system and the gut. NAS-37 from *C. elegans* was likewise abundantly expressed in the excretory duct and in the pharynx, both regions consistent with the enzyme's role in nematode ecdysis and moulting (Davis et al., 2004). A prolyl 4-hydroxylase, which is involved in matrix synthesis, is likewise expressed in the pharyngeal gland cells and the excretory duct of *C. elegans* (Keskiaho et al., 2008). Thus, proteins with similar developmental functions appear to be found in the cells involved in secretion from the nematode. The pharyngeal gland cells and the excretory duct are part of the excretory/ secretory system that secrete enzyme-rich exsheathment fluids, that aid the moulting process (Albertson and Thomson, 1976; Nelson et al., 1983; Bird, 1987). Additional proteins involved in moulting are expressed in the pharyngeal glands of *C. elegans* and *Meloidogyne javanica*, and in the hypodermis of *C. elegans*, *Nippostrongylus brasiliensis* and *M. javanica* (Bird, 1987). A reporter construct for *C. elegans* DPY-31 predominantly highlighted the hypodermal cells and some unidentified head neurons in the pharyngeal region (Novelli et al., 2004); this location could in fact correspond to the excretory system and gland cells, localizations highlighted by anti-DPY-31 antibodies applied in this present study. We did not find expression of DPY-31 in the hypodermal cells, either through anti-peptide antibody staining or through *B. malayi* promoter/reporter localisation. The excretory/secretory

501 expression pattern of DPY-31 in nematodes is however consistent with a role for this
502 enzyme in normal development through cuticle matrix remodelling.

503 In *C. elegans*, DPY-31 was identified as being a functional homologue of the
504 vertebrate type I procollagen C-proteinase (PCP) termed BMP-1 (Bone Morphogenetic
505 Protein-1) (Novelli et al., 2004). PCPs are secreted N-glycosylated metalloproteases and,
506 in the case of BMP-1, require calcium for optimal activity (Kessler et al., 1996). PCPs
507 cleave the carboxyl terminus of procollagens in the process of vertebrate extracellular
508 matrix formation, and BMP-1 is involved in bone formation and morphogenesis. Proteins
509 that are synthesised as precursors require proteolytic processing to regulate proper protein
510 function, with metalloproteases of the M12 family having this role in many diverse
511 organisms, such as nematodes, *Drosophila melanogaster*, humans and rodents. The
512 M12A sub-family (e.g. BMP-1 and astacin proteases) have been shown to cleave C-
513 propeptides, while the M12B sub-family (e.g. ADAMTS proteases) are involved in the
514 proteolytic processing of N-propeptides (Ge and Greenspan, 2006). In *C. elegans*, a
515 matrix metalloprotease (GON-1) similar to ADAMTS-1 and bovine procollagen N-
516 protease was found to display procollagen N-protease activity that regulated
517 organogenesis (Blelloch and Kimble, 1999; Moerman, 1999). GON-1 was predicted to
518 control gonad development through its ability to remodel type IV basement membrane
519 collagens (Moerman, 1999). In *C. elegans*, DPY-31 cleaves the C-terminus of the
520 essential cuticle collagen, SQT-3 (Novelli, 2006), and this current study reveals that this
521 function is conserved in distantly related parasitic nematodes. The recombinant DPY-31
522 enzymes in both *H. contortus* and *B. malayi* act as functional PCP's on recombinant
523 SQT-3 from *C. elegans*. The nematode enzymes require zinc for optimal activity, a

known property of PCPs (Kessler et al., 1996), and these enzymes are N-glycosylated at two sites, a property likewise shared with the other PCPs (Kessler et al., 1996). Conserved SQT-3 collagens are present in both *C. elegans* (Van der keyl et al., 1994) and *H. contortus* (Shamansky et al., 1989), and preliminary BLAST searches of other nematode databases (http://www.nematode.net/) indicate that both DPY-31 and SQT-3 homologues are present in a wider range of parasitic nematodes, including *A. suum*, *S. stercoralis*, *M. hapla*, *A. caninum* and *O. ostertagi*. This suggests that both these proteins are essential for the proper development of the cuticle across the nematode phyla.

In summary, DPY-31 is an important BMP-like enzyme in *C. elegans*, involved in proper cuticle formation through the C-terminal cleavage of procollagens to mature collagens (Novelli et al., 2004), an essential precursor to proper matrix assembly and cuticle formation. This study has demonstrated that this essential collagen processing function of DPY-31 may be conserved in the parasitic nematodes, *H. contortus* and *B. malayi*, and may therefore represent a future target for nematode control.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council (BBSRC) to Antony Page. Alan Winter (University of Glasgow) is thanked for comments and suggestions on this manuscript and Johanna Myllyharju (University of Oulu) for expressing and harvesting the SQT-3 proteins from insect cell culture. Jonathan Hodgkin, Frank Jackson and Rick Maizels kindly provided the *C. elegans*, *H. contortus* and *B. malayi* nematodes, respectively.

References

Albertson, D.G., Thomson, J.N., 1976. The pharynx of *Caenorhabditis elegans*. Phil. Trans. Roy. Soc. London B 275, 299-325.

Bastin, P., Bagherzadeh, Z., Matthews, K.R., Gull, K., 1996. A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. Mol. Biochem. Parasitol. 77, 235-239.

Bird, A.F., 1987. Moulting of parasitic nematodes. Inter. J. Parasitol. 17, 233-239.

Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., Vida, J.T., Thomas, W.K., 1998. A molecular evolutionary framework for the phylum Nematoda. Nature 392, 71-75.

Blelloch, R., Kimble, J., 1999. Control of organ shape by a secreted metalloprotease in the nematode *Caenorhabditis elegans*. Nature 399, 586-590.

Davis, M.W., Birnie, A.J., Chan, A.C., Page, A.P., Jorgensen, E.M., 2004. A conserved metalloprotease mediates ecdysis in *Caenorhabditis elegans*. Development 131, 6001-6008.

Gallego, S.G., Loukas, A., Slade, R.W., Neva, F.A., Varatharajalu, R., Nutman, T.B., Brindley, P.J., 2005. Identification of an astacin-like metallo-proteinase transcript from the infective larvae of *Strongyloides stercoralis*. Parasitol. Inter. 54, 123-133.

Gamble, H.R., Purcell, J.P., Fetterer, R.H., 1989. Purification of a 44 kilodalton protease which mediates the ecdysis of infective *Haemonchus contortus* larvae. Mol. Biochem. Parasitol. 33, 49-58.

570 Gamble, H.R., Fetterer, R.H., Mansfield, L.S., 1996. Developmentally regulated zinc
571 metalloproteinases from third- and fourth-stage larvae of the ovine nematode
572 *Haemonchus contortus*. J. Parasitol. 82, 197-202.
573 Ge, G., Greenspan, D.S., 2006. Developmental roles of the BMP-1/TLD
574 metalloproteinases. Birth Def. Res. 78, 47-68.
575 Hartigan, N., Garrigue-Antar, L., Kadler, K.E., 2003. Bone morphogenetic protein-1
576 (BMP-1). Identification of the minimal domain structure for procollagen C-
577 proteinase activity. J. Biol. Chem. 278, 18045-18049.
578 Hawdon, J.M., Jones, B.F., Perregaux, M.A., Hotez, P.J., 1995. *Ancylostoma caninum*:
579 metalloprotease release coincides with activation of infective larvae *in vitro*.
580 Exper. Parasitol. 80, 205-211.
581 Hill, D.E., Gamble, H.R., Rhoads, M.L., Fetterer, R.H., Urban, J.F., 1993. *Trichuris suis*:
582 a zinc metalloprotease from culture fluids of adult parasites. Exper. Parasitol. 77,
583 170-178.
584 Hishida, R., Ishihara, T., Kondo, K., Katsura, I., 1996. *hch-1*, a gene required for normal
585 hatching and normal migration of a neuroblast in *C. elegans*, encodes a protein
586 related to TOLLOID and BMP-1. EMBO J. 15, 4111-4122.
587 Jones, B.F., Hotez, P.J., 2002. Molecular cloning and characterization of *Ac-mep-1*, a
588 developmentally regulated gut luminal metalloendopeptidase from adult
589 Ancylostoma caninum hookworms. Mol. Biochem. Parasitol. 119, 107-116.
590 Keskiaho, K., Kukkola, L., Page, A.P., Winter, A.D., Vuoristo, J., Sormunen, R., Nissi,
591 R., Riihimaa, P., Myllyharju, J., 2008. Characterisation of a novel *Caenorhabditis*

592 *elegans* prolyl 4-hydroxylase with a unique substrate specificity and restricted
593 expression in the pharynx and excretory duct. J. Biol. Chem. 283, 10679-10689.
594 Kessler, E., Takahara, K., Biniaminov, L., Brusel, M., Greenspan, D.S., 1996. Bone
595 Morphogenetic Protein-1: the type I procollagen C-proteinase. Science 271, 360-
596 362.
597 Mitreva, M., Blaxter, M.L., Bird, D.M., McCarter, J.P., 2005. Comparative genomics of
598 nematodes. Trends Genet. 21, 573-581.
599 Moerman, D.G., 1999. Organ morphogenesis:a metalloprotease prepares the way. Current
600 Biol. 9, R701-R703.
601 Möhrlen, F., Hutter, H., Zwilling, R., 2003. The astacin protein family in *Caenorhabditis*
602 *elegans*. Eur. J. Biochem. 270, 4909-4920.
603 Nelson, F.K., Albert, P.S., Riddle, D.L., 1983. Fine structure of the *Caenorhabditis*
604 *elegans* secretory-excretory system. J. Ultra. Res. 82, 156-171.
605 Novelli, J., Ahmed, S., Hodgkin, J., 2004. Gene interactions in *Caenorhabditis elegans*
606 define DPY-31 as a candidate procollagen C-proteinase and SQT-3/ROL-4 as its
607 predicted major target. Genetics 168, 1259-1273.
608 Novelli, J., Page, A.P. and Hodgkin, J., 2006. The C terminus of collagen SQT-3 has
609 complex and essential functions in nematode collagen assembly. Genetics **172**,
610 2253-2267.
611 Opperman, C.H., Bird, D.M., Williamson, V.M., Rokhsar, D.S., Burke, M., Cohn, J.,
612 Cromer, J., Diener, S., Gajan, J., Graham, S., Houfek, T.D., Liu, Q., Mitros, T.,
613 Schaff, J., Schaffer, R., Scholl, E., Sosinski, B.R., Thomas, V.P., Windham, E.,

614 2008. Sequence and genetic map of *Meloidogyne hapla*: a compact nematode
615 genome for plant parasitism. Proc. Nat. Acad. Sci. USA 105, 14802-14807.
616 Page, A.P., Winter, A.D., 2003. Enzymes involved in the biogenesis of the nematode
617 cuticle. Adv. Parasitol. 53, 85-148.
618 Page, A.P., Johnstone, I.L., 2007. The cuticle. In: WormBook. The *C. elegans* research
619 community. (Eds). doi/10.1895/ wormbook.1.138.1
620 Redmond, D.L., Clucas, C., Johnstone, I.L., Knox, D.P., 2001. Expression of
621 *Haemonchus contortus* pepsinogen in *Caenorhabditis elegans*. Mol. Biochem.
622 Parasitol. 112, 125-131.
623 Shamansky, L.M., Pratt, D., Boisvenue, R.J., Cox, G.N., 1989. Cuticle collagen genes of
624 *Haemonchus contortus* and *Caenorhabditis elegans* are highly conserved. Mol.
625 Biochem. Parasitol. 37, 73-86.
626 Stepek, G., McCormack, G., Page, A.P., 2009. The kunitz domain protein BLI-5 plays a
627 functionally conserved role in cuticle formation in a diverse range of nematodes.
628 Mol. Biochem. Parasitol. In Press, 10.1016/j.molbiopara.2009.08.005.
629 Stöcker, W., Gomis-Rüth, F.-X., Bode, W., Zwilling, R., 1993. Implications of the three-
630 dimensional structure of astacin for the structure and function of the astacin
631 family of zinc-endopeptidases. Eur. J. Biochem. 214, 215-231.
632 Stöcker, W., Zwilling, R., 1995. Astacin. Methods in Enzymology 248, 305-325.
633 Studier, F.W., 2005. Protein production by auto-induction in high-density shaking
634 cultures. Prot. Expr. Purif. 41, 207-234.

635 Suzuki, M., Sagoh, N., Iwasaki, H., Inoue, H., Takahashi, K., 2004. Metalloproteases
636 with EGF, CUB, and thrombospondin-1 domains function in molting of
637 *Caenorhabditis elegans*. Biol. Chem. 385, 565-568.
638 Thein, M.C., Winter, A.D., Stepek, G., McCormack, G., Stapleton, G., Johnstone, I.L.,
639 Page, A.P., 2009. Combined extracellular matrix cross-linking activity of the
640 peroxidase MLT-7 and the dual oxidase BLI-3 Is critical for post-embryonic
641 viability in *Caenorhabditis elegans*. J. Biol. Chem. 284, 17549-17563.
642 Van der Keyl, H., Kim, H., Espey, R., Oke, C.V., Edwards, M.K., 1994. *Caenorhabditis*
643 *elegans sqt-3* mutants have mutations in the col-1 collagen gene. Dev. Dyn. 201,
644 86-94.
645 Williamson, S.M., Walsh, T.K., Wolstenholme, A.J., 2007. The cys-loop ligand-gated ion
646 channel gene family of *Brugia malayi* and *Trichinella spiralis*: a comparison with
647 *Caenorhabditis elegans*. Invert. Neurosci. 7.
648

FIGURE LEGENDS

Figure 1. Gene structures of *C. elegans*, *H. contortus* and *B. malayi dpy-31*. The introns are indicated by lines while the exons are open boxes. The translational start and stop codons are indicated by the ATG and TGA or TAG, respectively. The catalytic zinc-binding site in the astacin domain is indicated by *, and the conserved methionine-turn, also in the astacin domain, by +. The EGF domain is highlighted by a black arrowhead, the CUB domain is over-lined and the TSP-1 domain is indicated by a black arrow. A schematic of the DPY-31 protein, showing each identified domain is below the gene structures, with the signal peptide indicated by a white box and the checked box depicting the prodomain. The catalytic zinc-binding site and methionine-turn are represented by black ovals.

Figure 2. Alignment of DPY-31 between *C. elegans*, *H. contortus* and *B. malayi*. Amino acid sequences were aligned using ClustalX and BoxShade. Identical amino acids are shaded black and similar amino acids are shaded grey. The signal peptide domain is shown in italics, while the prodomain is boxed in black. The catalytic zinc-binding site (HExxHxxGFxHExxRxDRD) is indicated by * and the conserved methionine-turn (SxMHY) by ^. The EGF domain is double-underlined, the CUB domain is marked by + and the TSP-1 domain by a grey box. Potential N-glycosylation sites are overlined. The position of the Leu to Pro mutation in *C. elegans dpy-31(e2770)* is indicated thus, #. The percentage sequence identity between *C. elegans* and the parasitic nematodes is shown at the end of the alignment.

672

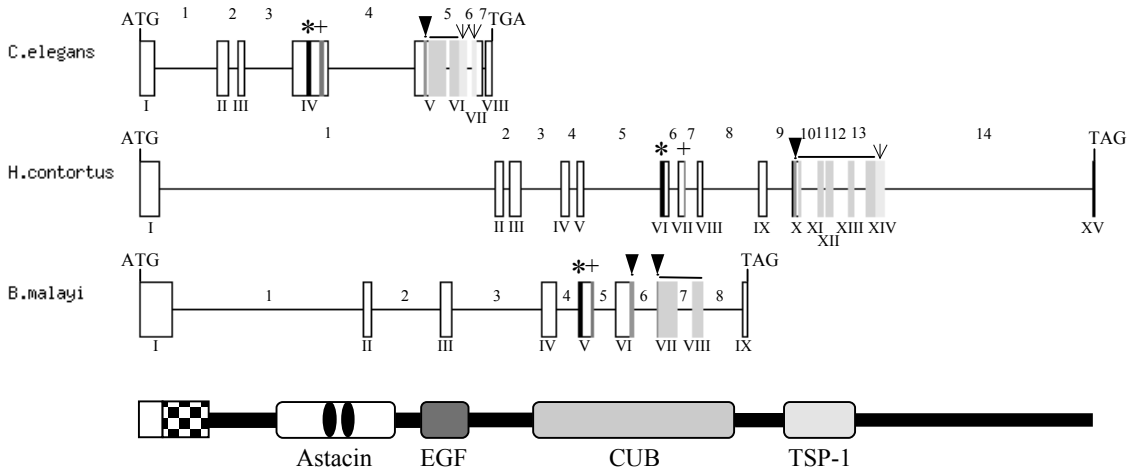
673 **Figure 3.** A *H. contortus dpy-31* homologue complements the *C. elegans dpy-31(e2770)*
674 mutant strain. The construct was co-injected with either *dpy-7*:GFP marker or *cpr-5*:GFP
675 marker. These lines were examined by both DIC and GFP microscopy, and representative
676 images are presented. h, represents the head-end of the nematodes. (A) DIC image of a
677 wild type (N2) worm. (B) DIC image of a *dpy-31(e2770)* mutant worm at 20^{o}C. (C) DIC
678 image of a rescued *dpy-31(e2770)* worm with the *H. contortus dpy-31* cDNA construct
679 and the *dpy-7*:GFP marker at 20^{o}C. (D) GFP image of the worm in (C). (E) DIC image of
680 a rescued *dpy-31(e2770)* worm with the *H. contortus dpy-31* cDNA construct with the
681 *cpr-5*:GFP marker at 20^{o}C. (F) GFP image of the worm in (E). (G) Single worm PCR was
682 performed for three GFP positive (lanes 1-3) and three GFP negative worms (lanes 4-6)
683 for the *dpy-7* lines, for three N2 worms (lanes 7-9) and three *dpy-31(e2770)* mutant
684 worms (lanes 10-12), and for three GFP positive worms (lanes 13-15) and one GFP
685 negative worm for the *cpr-5* line (16), to confirm that rescued *C. elegans dpy-31(e2770)*
686 expressed the 1647bp *Hcdpy-31* cDNA rescue construct. (H) A schematic representation
687 of the *Hcdpy-31* rescue construct. The rescue vector (p*nas-35*) contained the *Cedpy-31*
688 promoter and 3′-UTR sequences, inserted into pBlueScript SK-. The *Hcdpy-31* rescue
689 construct contained the *Cedpy-31* promoter and 3′-UTR sequences, and the *Hcdpy-31*
690 coding sequence from the initiating methionine codon to the stop codon, with a single *C.*
691 *elegans* synthetic intron (SI) inserted.

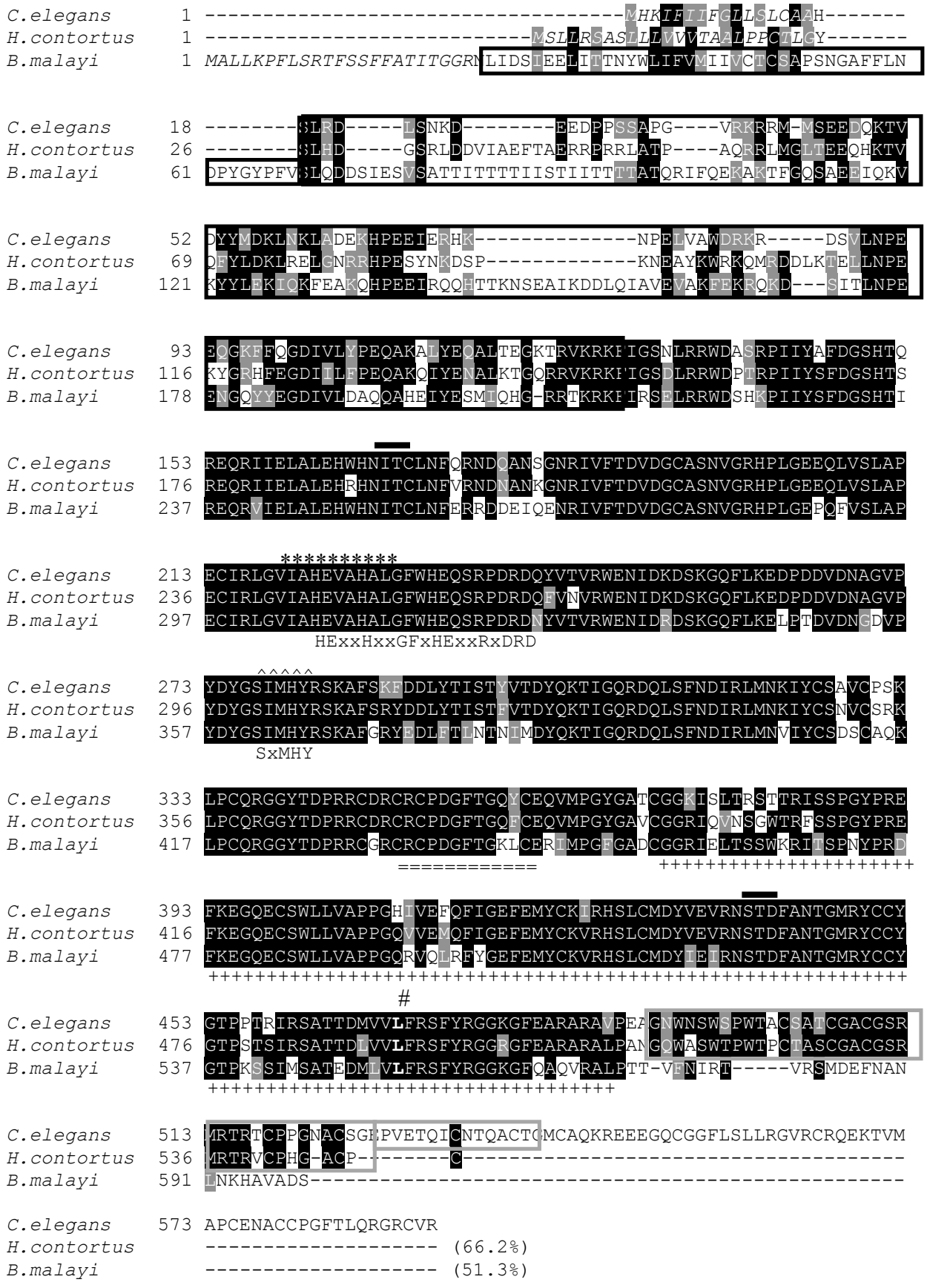
692

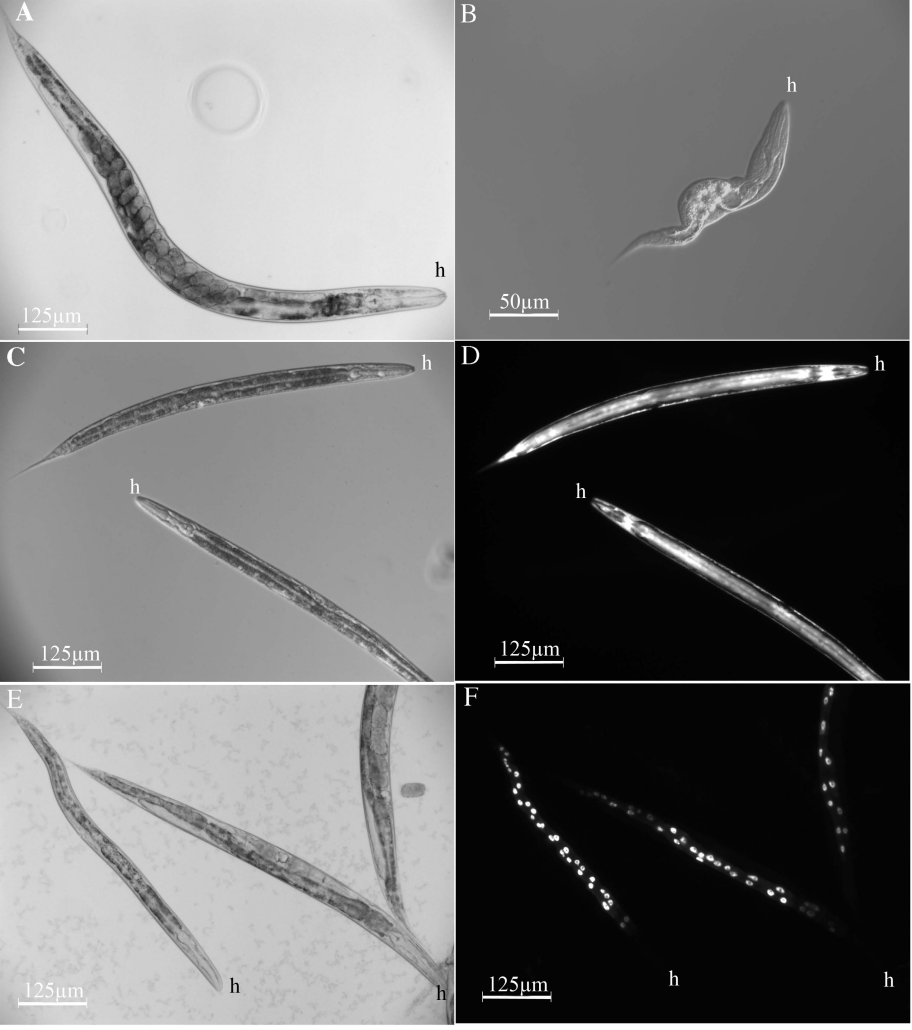
693 **Figure 4.** Astacin activity of recombinant DPY-31 from *H. contortus* and *B. malayi*.
694 100nM of each of the recombinant enzymes was incubated with 0.3mM $ZnCl_2$ and 0, 1, 2,

695 5 or 10mM 1, 10-phenanthroline. Each sample was performed in triplicate and standard
696 error bars are depicted.
697
698 **Figure 5.** Cleavage of *C. elegans* recombinant SQT-3 by *H. contortus* or *B. malayi*
699 recombinant DPY-31 demonstrates the functional activity of DPY-31 from *H. contortus*
700 and *B. malayi*. The three baculovirus expressed SQT-3 recombinant protein constructs
701 were TY-tagged in different positions. 4μM of recombinant enzyme or bovine serum
702 albumin (BSA) was incubated with each of the SQT-3 proteins, in the presence of 0.3mM
703 $ZnCl_2$, and then analysed by western blotting using an anti-TY tag antibody. (A) Lanes 1-
704 3: *H. contortus* DPY-31. Lanes 4-6: BSA. Lanes 1 and 4, untagged SQT-3; lanes 2 and 5,
705 SQT-3 with the TY tag between the Gly-X-Y repeats; and lanes 3 and 6, SQT-3 with the
706 TY tag at the C-terminus in a position after the predicted cleavage site for DPY-31.
707 Marker sizes (kDa) from prestained standards (BioRad) are indicated on the left of the
708 blot. (B) *B. malayi* DPY-31. Lane 1, untagged SQT-3; lane 2, SQT-3 with the TY tag
709 between the Gly-X-Y repeats; and lane 3, SQT-3 with the TY tag at the C-terminus in a
710 position after the predicted cleavage site for DPY-31. Marker sizes (kDa) from prestained
711 standards (BioRad) are indicated on the left of the blot. (C) Schematic diagram of the
712 SQT-3 cuticle collagen constructs. C, conserved cysteine residues; Gly-X-Y, Gly-X-Y
713 repeat regions; TY1, TY-tag added between the Gly-X-Y repeat regions; TY2, TY-tag
714 added in the C-terminal non-repetitive region, after the potential DPY-31 cleavage site;
715 arrow, predicted cleavage site for DPY-31 between amino acids 289 and 290, YCAL/D.
716

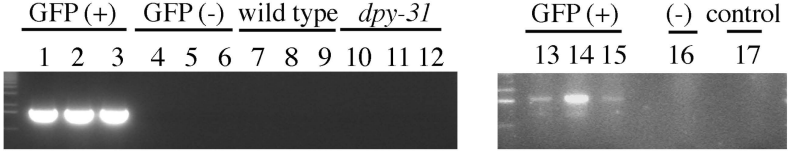
Figure 6. Expression of the *dpy-31* transcript and protein. *C. elegans* wild type (N2) nematodes were transformed with the *B. malayi* promoter-reporter construct and transgenic lines were identified and examined for reporter gene expression in live worms for GFP expression (A and B) and by staining glutaraldehyde-fixed worms for β-galactosidase activity (C and D). p, represents the terminal bulb of the pharynx and In, is the intestine. A is the DIC image and B shows the diffuse fluorescent image of the intestine of the same worm. DPY-31 was also expressed in the gland cells of the pharynx (arrow in D) anterior to the nuclei of the first pair of intestine cells. Worms were also examined for the expression of *C. elegans* DPY-31 by staining with an anti-Ce-DPY-31 peptide antibody (E-H). E is the DIC image and F is the fluorescent image of same worm, depicting the excretory gland and duct (arrowed). G is the DIC image and H is the fluorescent image of same worm, and the gut is arrowed in H. Representative images are depicted. Both staining methods demonstrated that DPY-31 was predominantly expressed in the gut cells (C), but also in the excretory duct of the excretory system (arrow in F).


C.elegans
ATG
TGA
I
II III
IV
V
VI
VII
VIII
H.contortus
ATG
TAG
I
II III
IV V
VI VII VIII
IX
X XI XII XIII XIV
XV
B.malayi
ATG
TAG
I
II
III
IV
V
VI
VII VIII
IX
Astacin
EGF
CUB
TSP-1

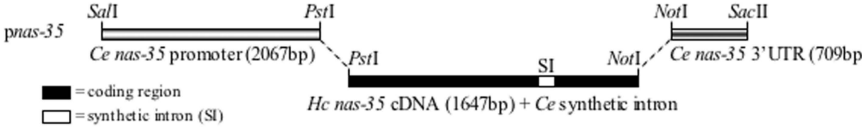

C.elegans
H.contortus
B.malayi
HExxHxxGFxHExxRxDRD
SxMHY
(66.2%)
(51.3%)

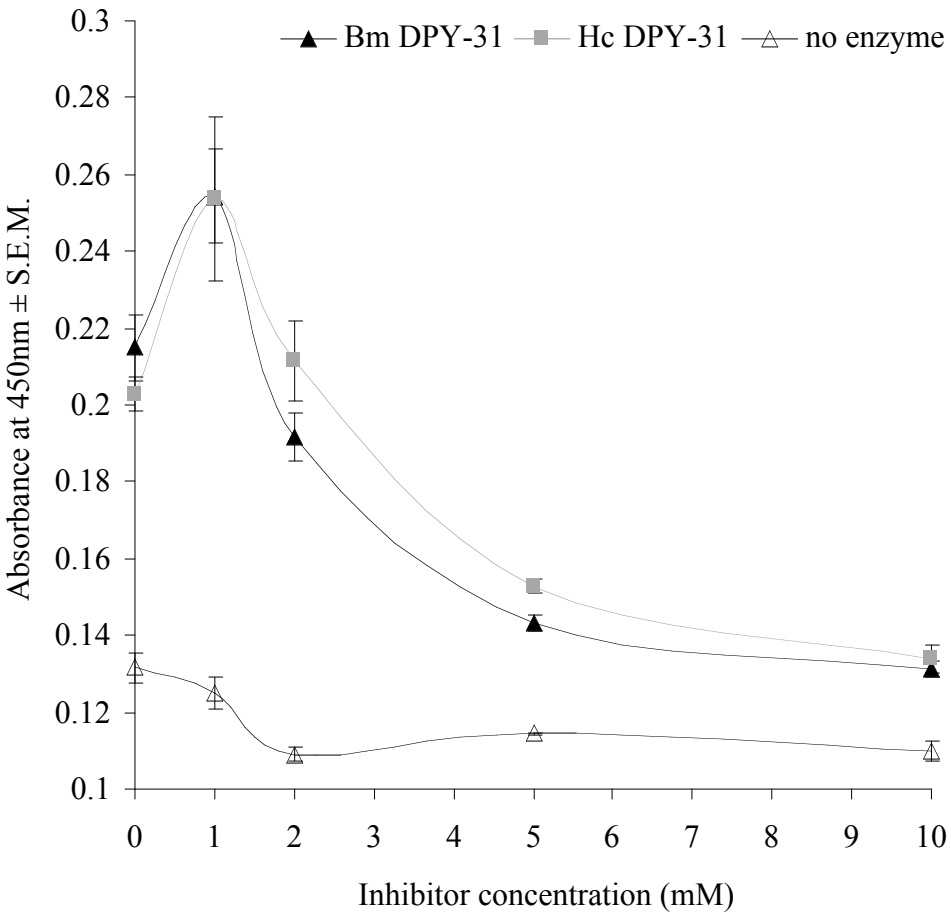

A
h
125µm
B
h
50µm
C
h
h
125µm
D
h
h
125µm
E
125µm
h
h
F
125µm
h
h

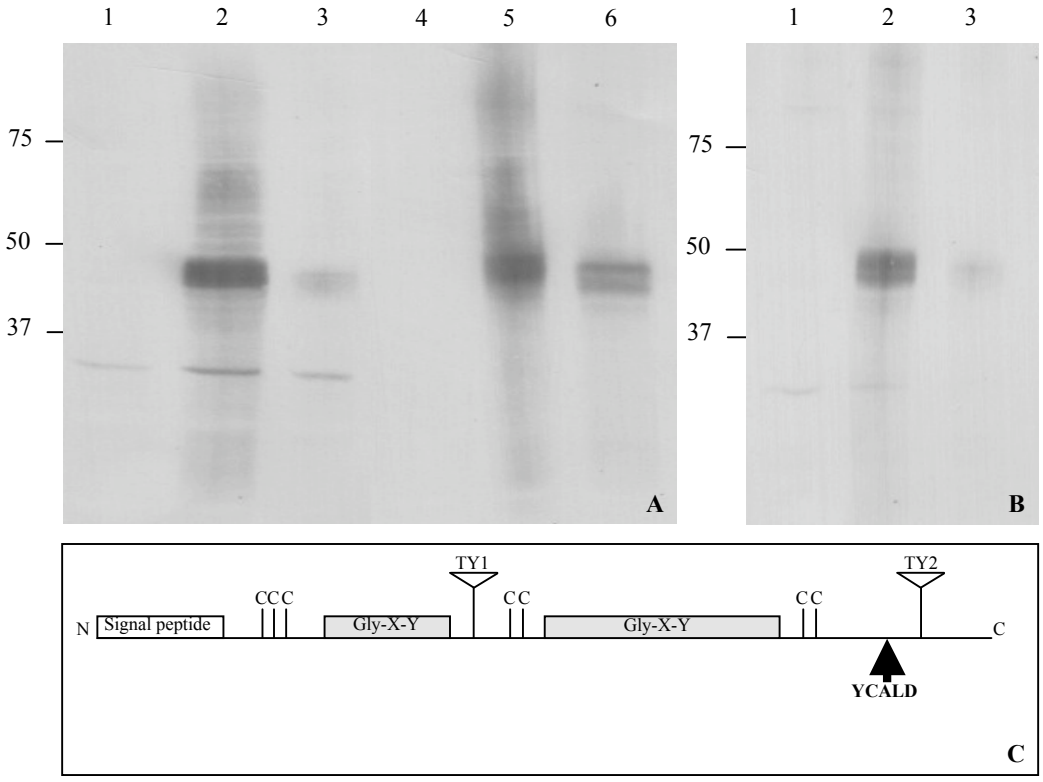
G


GFP (+)
GFP (-)
wild type
dpy-31
1 2 3 4 5 6 7 8 9 10 11 12
GFP (+)
(-)
control
13 14 15
16
17

H


pnas-35
SalI
PstI
Ce nas-35 promoter (2067bp)
PstI
SI
NotI
NotI
SacII
Ce nas-35 3'UTR (709bp
= coding region
= synthetic intron (SI)
Hc nas-35 cDNA (1647bp) + Ce synthetic intron


Bm DPY-31
Hc DPY-31
no enzyme
Absorbance at 450nm ± S.E.M.
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0
1
2
3
4
5
6
7
8
9
10
Inhibitor concentration (mM)


1
2
3
4
5
6
75
50
37
A
1
2
3
75
50
37
B
TY1
TY2
CCC
CC
CC
N
Signal peptide
Gly-X-Y
Gly-X-Y
C
YCALD
C