



University
of Glasgow

Winter, A. D. et al. (2007) *Differences in collagen prolyl 4-hydroxylase assembly between two Caenorhabditis nematode species despite high amino acid sequence identity of the enzyme subunits*. *Matrix Biology*, 26 (5). pp. 382-395. ISSN 0945-053X

<http://eprints.gla.ac.uk/4946/>

Deposited on: 20 February 2009

DIFFERENCES IN COLLAGEN PROLYL 4-HYDROXYLASE ASSEMBLY BETWEEN TWO *CAENORHABDITIS* NEMATODE SPECIES DESPITE HIGH AMINO ACID SEQUENCE IDENTITY OF THE ENZYME SUBUNITS.

Alan D. Winter¹, Katriina Keskiäho², Liisa Kukkola², Gillian McCormack¹, Marie-Anne Felix³, Johanna Myllyharju² and Antony P. Page^{1*}

¹Institute of Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Bearsden Road Glasgow, G61 1QH, Scotland. ²Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, FIN-90014 Oulu, Finland ³Institut Jacques Monod, Tour 43, 2 pl. Jussieu, 75251 Paris Cedex 05, France.

Running title: Nematode Prolyl 4-Hydroxylase Assembly

Corresponding Author: a.page@vet.gla.ac.uk

Fax 0044 1413305603; Tel 0044 1413301997.

The collagen prolyl 4-hydroxylases (P4Hs) are essential for proper extracellular matrix formation in multicellular organisms. The vertebrate enzymes are $\alpha_2\beta_2$ tetramers, in which the β subunits are identical to protein disulfide isomerase (PDI). Unique P4H forms have been shown to assemble from the *Caenorhabditis elegans* catalytic α subunit isoforms PHY-1 and PHY-2 and the β subunit PDI-2. A mixed PHY-1/PHY-2/(PDI-2)₂ tetramer is the major form, while PHY-1/PDI-2 and PHY-2/PDI-2 dimers are also assembled but less efficiently. Cloning and characterization of the orthologous subunits from the closely related nematode *Caenorhabditis briggsae* revealed distinct differences in the assembly of active P4H forms in spite of the extremely high amino acid sequence identity (92-97%) between the *C. briggsae* and *C. elegans* subunits. In addition to a PHY-1/PHY-2/(PDI-2)₂ tetramer and a PHY-1/PDI-2 dimer, an active (PHY-2)₂(PDI-2)₂ tetramer was formed in *C. briggsae* instead of a PHY-2/PDI-2 dimer. Site-directed mutagenesis studies and generation of inter-species hybrid polypeptides showed that the N-terminal halves of the *Caenorhabditis* PHY-2 polypeptides determine their assembly properties. Genetic disruption of *C. briggsae phy-1* (*Cb-dpy-18*) via a *Mos1* insertion resulted a small (short) phenotype that is less severe than the dumpy (short and fat) phenotype of the corresponding *C. elegans* mutants (*Ce-dpy-18*). *C. briggsae phy-2* RNA interference produced no visible phenotype in the wild type nematodes but produced a severe dumpy phenotype and larval arrest in *phy-1* mutants. Genetic complementation of the *C. briggsae* and *C. elegans phy-1* mutants was achieved by

injection of a wild type *phy-1* gene from either species.

Introduction

The collagen prolyl 4-hydroxylases (P4Hs) are endoplasmic reticulum (ER) luminal enzymes that hydroxylate prolines in -X-Pro-Gly- triplets of collagen chains (1-3). This essential reaction requires Fe²⁺, 2-oxoglutarate, oxygen and ascorbate, and precedes the folding of the polypeptide chains into stable triple-helical collagen molecules. Collagen P4Hs are $\alpha_2\beta_2$ tetramers in vertebrates, the β subunits of which correspond to the enzyme and chaperone protein disulfide isomerase (PDI). These PDI/ β subunits are required to maintain the α subunits in an active, soluble conformation and for retaining the tetramer within the ER. The α subunit contains the catalytic site with conserved residues required for the binding of the Fe²⁺ and 2-oxoglutarate (4, 5) and the peptide-substrate-binding domain (6, 7). Three α subunit isoforms have been characterized from human and mouse that assemble into [α (I)]₂ β_2 , [α (II)]₂ β_2 and [α (III)]₂ β_2 tetramers (8-12) and display distinct tissue distributions (11, 13, 14). The *Drosophila melanogaster* genome contains about 20 genes encoding collagen P4H α subunit-like polypeptides, the one characterized in detail also assembling into an active $\alpha_2\beta_2$ tetramer (15, 16).

The nematode *Caenorhabditis elegans* has two conserved genes encoding collagen P4H α subunits, named *phy-1* (also known as *dpy-18*) and *phy-2* that are involved in the synthesis of cuticle collagens (17-20). Based on *in vivo* and *in vitro* studies these subunits assemble into unique collagen P4H forms with a β subunit encoded by the *pdi-2* gene (17, 21-23). The main collagen P4H form in *C. elegans* is a

PHY-1/PHY-2/(PDI-2)₂ tetramer, but a PHY-1/PDI-2 dimer and a PHY-2/PDI-2 dimer can also be assembled albeit at a much lower efficiency (20-22). A collagen P4H tetramer containing two different catalytically active subunits has not been found in any other organism, and insect cell co-expression studies have shown that the human α (I) and α (II) subunits do not assemble into mixed tetramers of this type (10). Formation of the *C. elegans* mixed tetramer is species-specific as the *C. elegans* PHY-1, PHY-2 or PDI-2 could not be replaced in insect cell coexpression experiments by a recombinant collagen P4H α or β subunit from human or *Drosophila* (20). Detailed studies with hybrid recombinant PHY-1/PHY-2 polypeptides revealed that PHY-1 residues Gln121-Ala271 and PHY-2 residues Asp1-Leu122 are essential for the assembly of the mixed tetramer (20).

The *C. elegans* *phy-1*, *phy-2* and *pdi-2* genes are expressed in the hypodermis (17, 20), a tissue whose main role is to secrete the collagenous exoskeleton known as the cuticle (23). All three genes are expressed temporally in a cyclical pattern that matches the expression of the exoskeleton collagens, namely the molting cycle (17). In agreement with the assembly properties of these subunits, mutations in their genes result in distinct phenotypes. The *phy-1* null mutant (*dpy-18*) is dumpy (short and fat) with a defective exoskeleton (17-19). The loss of the PHY-1/PHY-2/(PDI-2)₂ and PHY-1/PDI-2 forms in these mutants can be partially compensated by assembly of the PHY-2/PDI-2 dimer, which is not detected in the wild type nematodes (20). In contrast, the *phy-2* null mutants are wild type in appearance, an attribute that is a direct consequence of the PHY-1/PDI-2 dimer, the formation of which becomes highly up-regulated in this mutant strain (20). A *phy-2* RNA interference (RNAi) knockdown in the *dpy-18* mutant background leads to loss of all collagen P4H forms and the resultant nematode embryos fail to hatch. These embryos die at the point where the exoskeleton usually assumes the role of maintaining the body shape, thereby confirming an essential function for both PHY-1 and PHY-2 in the synthesis of a proper collagenous cuticle (17). Two additional genes encoding collagen P4H α subunit-like polypeptides exist in the *C. elegans* genome but are not involved in the synthesis of cuticle collagens (24); (Keskiaho et al unpublished).

The unusual collagen P4H forms

assembled by the *C. elegans* PHY-1, PHY-2 and PDI-2 polypeptides and their important roles in the establishment of the collagenous exoskeleton led us to examine the orthologous genes in the closely related free-living bacterivorous soil nematodes *C. elegans* and *Caenorhabditis briggsae*. Complete genome sequences are available for both species allowing detailed comparative genome analysis (25). We report here the cloning of the collagen P4H subunit orthologues in *C. briggsae*, their biochemical and genetic characterization, and relate these results to the *C. elegans* data. Surprisingly, despite the high identity at the amino acid level, some of the assembly properties of the *C. briggsae* collagen P4H α subunits were found to differ distinctly from those of *C. elegans*. In addition to the PHY-1/PHY-2/(PDI-2)₂ tetramers and PHY-1/PDI-2 dimers formed in both species, *C. briggsae* PHY-2 readily associated with PDI-2 to form an active (PHY-2)₂/(PDI-2)₂ tetramer, which is not found in *C. elegans*. This allowed us to identify regions and individual amino acids in PHY-1 and PHY-2 that are important in the assembly of the different enzyme forms. We also studied the *in vivo* roles of the *C. briggsae* collagen P4H subunits in the formation of the cuticle by RNA interference and through the generation of a *Mos1 Cb-dpy-18* mutant, and by examining the ability of the *C. elegans* and *C. briggsae* PHY polypeptides to rescue the *dpy-18* phenotype in both species.

Experimental Procedures

Nematode Strains and Generation of *C. briggsae dpy-18* Mutant - The *C. elegans* strains N2 (wild type), CB364 (*dpy-18 e364*) and *dpy-18 e1096*, and the *C. briggsae* strain AF16 (wild type) were received from the *Caenorhabditis* Genetics Center. To generate the *Cb-dpy-18* mutant strain, mutagenesis using the heterologous *Drosophila Mos1* transposon was performed in *C. briggsae* basically as previously described for *C. elegans* (26, 27). A marked transgenic strain bearing the *Mos1* transposon, JU826 (*mfEx24[pPD9; Ce-myo-3::GFP]*), was obtained by injection of *C. briggsae* AF16 with 10 μ g/ml pPD9 (*Mos1*) and 100 μ g/ml pPD93.97 (*Ce-myo-3::GFP*). A marked transgenic strain bearing the *Mos1* transposase under a heat-shock promoter, JU786 (*mfEx18[pJL44.2; Ce-myo-2::GFP]*), was obtained by injection with 90 mg/ μ l pJL44.2 (*Ce-hsp16.48::GFP::glh-2 3'UTR*) and 10

$\mu\text{g/ml}$ pPD118.33 (*Ce-myo-2::GFP*). For mutagenesis, males bearing *mfEx18* were crossed to hermaphrodites bearing *mfEx24* and double transgenics were amplified for two generations. Double transgenic adults were then heat-shocked (2h40 at 37°C, 35 min at 20°C, 60 min at 37°C and overnight at 25°C). *Mos1* insertion efficiency was monitored by isolating random F1 progeny and non-Myo-3::GFP F2 and F3 progeny and scoring the latter by PCR for transposon presence using oligonucleotides oJL102 and oJL103, and for *mfEx24* absence using oJL103 and oJL104 (26). Three out of 46 random F1 animals had a *Mos1* insertion. The *mf104* mutation was found by visual inspection of the F2 progeny and was outcrossed five times to the wild type AF16 strain, yielding strain JU929. Using inverse PCR and BLAST on the *C. briggsae* genome in Wormbase, the *mf104 Mos1* insertion was found to be located at the very end of exon 4 of *Cb-phy-1* (CBG13195), before the sequence TATgtgagtcgattgtttgatcgg.

Cloning of C. briggsae Collagen P4H Subunits and Generation of Recombinant Baculoviruses - Putative genes encoding *C. briggsae* collagen P4H subunits were identified by homology to the *C. elegans phy-1*, *phy-2* and *pdi-2* genes from sequence data generated by the *C. briggsae* genome sequence project (<http://www.wormbase.org/db/seq/browse/briggsae/>). The cDNAs were cloned and sequenced from mRNA derived from a mixed stage culture of wild type *C. briggsae* nematodes using Trizol reagent (Invitrogen) and a StrataScript cDNA Synthesis kit (Stratagene). The following cloning primers were used (restriction sites underlined):

Cbphy1For (NotI) 5' - cgcgccggccgcATGCGTCTAGCACTTCTTGTGCTG-3', Cbphy1Rev (XbaI) 5' - gcgtctagaTTAAATGGTCTCCCAAACATCTTTTC-3', Cbphy2For (NotI) 5' - cgcgccggccgcATGAGAACAGTAGCGATTGTCTGC-3', Cbphy2Rev (XbaI) 5' - gcgtctagaCTATGGATCGTTGGCGTATGGAGAC-3', Cbpdi2For (NotI) 5' - cgcgccggccgcATGTTCCGGCTCGTCGGTCTG-3', Cbpdi2Rev (XbaI) 5' - gcgtctagaTTAGAGCTCAGTGTGTCCCTC-3'. The products were cloned into the pVL1392 baculovirus expression vector (Pharmlingen).

For the expression of hybrid PHY-2 polypeptides in insect cells constructs were generated in which sequences coding for the

N- and C-terminal halves were swapped between the *Caenorhabditis* species. A cDNA fragment encoding the N-terminal half of Cephy-2 was amplified by PCR using the primers Cephy2BVF (NotI) 5' - gacgcccgcATGAGAGCAGTTTTGCTAGTC-3' and Cephy2N-termR 5' - TCAACAATTCCATCGTACTC-3' and that encoding the C-terminal half using the primers Cephy2C-termF 5' - GAGAGATGCATATGAAGC-3' and phy2BVR (XbaI) 5' - gcgtctagaCTATGGATCATTGGCATATGGGGAC-3', and the plasmid pVL1392Cephy2 (20) as a template. A cDNA fragment encoding the N-terminal half of Cb-phy-2 was generated using primers Cbphy2For (NotI) (see above) and Cbphy2N-termR 5' - TCAACGATTCCATCGAATTC-3', and that encoding the C-terminal half using primers Cbphy2C-termF 5' - AAGAGACGCTTATGAGGC-3' and Cbphy2R (XbaI) (see above), and the pVL1392Cbphy2 (see above) as a template. The amplified fragments were co-ligated into NotI-XbaI digested pVL1392 and named pVL1392-CeN/CbCphy-2 and pVL1392-CbN/CeCphy-2.

Site-directed mutagenesis was carried out using the QuikChange Kit (Stratagene) following the manufacturers recommendations. The single mutants Ce-phy-1 T220A, Cb-phy-1 A220T, Ce-phy-2 W190S and Cb-phy-2 S190W (numbers of the amino acid positions refer to mature protein sequences after cleavage of the signal peptides) were generated in the first mutagenesis round using complementary sense and antisense mutagenic primers (the sense primers are shown, the bases introducing the desired mutation are underlined): Ce-phy-1 T220A 5' - CTCTACAAAATGAACCCAGCTCACCCA CGCGCCAAAG-3', Ce-phy-2 W190S 5' - CACCAACAATTGAGGAATCGGAAATTC TCGAGTATC-3', Cb-phy-1 A220T 5' - GTTGTACAAAATGAACCCA ACTCATCC ACGTGCCAAAG-3', Cb-phy-2 S190W 5' - CCACCAAGTGTCGAAGAATGGGAAATC CTTGAATATCTGG-3'. The plasmids pVL1392Cephy1 (21), pVL1392Cephy2 (20), and pVL1392Cbphy1 and pVL1392Cbphy2 described above were used as templates in the mutagenesis reactions. The double mutants Ce-phy-1 Q174K/T220A, Cb-phy-1 K174Q/A220T, Ce-phy-2 V177N/W190S and

Cb-phy-2 N177V/S190W were generated similarly in the second mutagenesis round using complementary sense and antisense mutagenic primers (the sense primers are shown): Ce-phy-1 Q174K 5'-GTGGATGGAGGAGGCGAAAGAGAAGAC TCGGCGAC-3', Ce-phy-2 V177N 5'-GATGCAGGTTGCGTTGAATAAGATCGA AAACGAG-3', Cb-phy-1 K174Q 5'-GTGGATGGAAGAGGCTCAGAGAAGAA TGGATGATG-3', Cb-phy-2 N177V 5'-GATGCAAGTTGCTTTGGTCAAATCGA GAACGAG-3' and the plasmids derived from the first round of mutagenesis as templates.

The recombinant baculovirus constructs were co-transfected into *Spodoptera frugiperda* Sf9 cells with a modified *Autographa californica* nuclear polyhedrosis virus DNA using the BaculoGold transfection kit (PharMingen) and the resultant virus pools were collected and amplified (28).

Expression and Analysis of Recombinant Proteins in Insect Cells - Insect cells (Sf9, Invitrogen) were cultured in TNM-FH medium (Sigma) supplemented with 10% insect cell qualified fetal bovine serum (Gibco) as monolayers at 27°C. Cells seeded at a density of 5×10^6 per 100-mm plate were infected at a multiplicity of five with different combinations of the viruses. The cells were harvested 72 h after infection, washed with a solution of 0.15 M NaCl and 0.02 M phosphate, pH 7.4, homogenized in a solution of 0.1 M glycine, 0.1 M NaCl, 10 μ M dithiothreitol, 0.1% Triton X-100, and 0.01 M Tris, pH 7.8, and centrifuged at 10,000 x g for 20 min. Samples of the resulting supernatants were analyzed by 8% SDS-PAGE under reducing conditions or 8% nondenaturing PAGE followed by Western blotting and assayed for enzyme activity by a method based on the hydroxylation-coupled decarboxylation of 2-oxo[1-¹⁴C]glutarate (29).

The molecular weight of the P4H assembled from *C. briggsae* PHY-2 and PDI-2 was analyzed by applying the Triton X-100-soluble fraction of insect cells coexpressing these recombinant polypeptides to a calibrated HiPrep Sephacryl S-200 HR gel filtration column (Amersham Biosciences), equilibrated and eluted with 0.1 M NaCl, 0.1 M glycine, 10 μ M dithiothreitol, and 0.01 M Tris buffer, pH 7.4, and P4H activity was assayed in the eluted fractions. Triton X-100 extracts from insect cells coexpressing *C. elegans* PHY-1 and human PDI and purified recombinant human

collagen P4H-I (30) were used as controls in gel filtration experiments.

Antibodies - The polyclonal antibodies used in the study were generated against synthetic peptides representing the C-termini of the *C. elegans* PHY-1, PHY-2 and PDI-2 (20). The amino acid sequences of these peptide antigens are highly conserved in the corresponding *C. briggsae* polypeptides and the antibodies recognize the orthologous subunits in both species. Western blot analysis of baculovirus-generated recombinant proteins was performed by transferring samples from non-denaturing PAGE to an Immobilon-P membrane (Millipore), followed by incubation with the antibodies described above and a subsequent incubation with an alkaline phosphatase-conjugated secondary antibody (Sigma).

Characterization of Collagen P4Hs from Nematode Lysates - Native worm extracts were prepared by extraction from mixed stage nematodes as described previously (20) and the proteins were analyzed by 8% non-denaturing PAGE followed by Western blotting as described above.

Rescue of the dpy-18 Mutants - *Cb-phy-1* and *Cb-phy-2* rescue constructs were amplified from wild type *C. briggsae* genomic DNA (2416 and 3384 bp respectively) using *Pfu* Turbo DNA Polymerase (Stratagene) and the following primers: Cbphy1ResF (XbaI) 5'-gctctagaGATGCGTCTAGCACTTCTTGTGCTGGCA-3', Cbphy1ResR (SacII) 5'-gaccgaggTTAAATGGTCTCCCAAACATCTTTTC-3', Cbphy2ResF (XbaI) 5'-gctctagaGATGAGAACAGTAGCGATTGCTGTC-3', Cbphy2ResR (SacII) 5'-gaccgaggCTATGGATCGTTGGCGTATGGAGAC-3'. A *C. elegans phy-2* rescue construct was generated by PCR (4775bp) from *C. elegans* N2 genomic DNA using the following primers: phy-2FLF (XbaI) 5'-gctctagaGATGAGAGCAGTTTTGCTAGTC-3', phy-2FLR (SacII) 5'-gaccgaggCTATGGATCATTGGCATATG-3'. All products were cloned into the vector pAW1 (31), that contains 2760 bp of the *C. elegans phy1* promoter and 205 bp of the 3' untranslated region, via the cloning vector pCRScript (Stratagene). Generation of the genomic *C. elegans phy-1* rescue construct has been described previously (17). Rescue constructs were coinjected at 5 μ g/ml (exception being injection of *Ce-phy-1* into the *C. briggsae dpy-18* strain which was at 1 μ g/ml) with a *dpy-7::gfp* (5 μ g/ml)

transformation marker and a pBluescript DNA carrier in a final total DNA concentration of 120 µg/ml, into the syncytial gonad of the appropriate *Caenorhabditis* species following published methods (17). Positive transformants were isolated under epifluorescence microscopy, maintained clonally and assessed for phenotype rescue via microscopy.

Transgene expression was confirmed by RT-PCR following RNA purification and cDNA synthesis procedures as outlined above. *phy-1* and *phy-2* transgene-specific products from both species were amplified using the ResF primers listed above (with the exception of *Ce-phy-1ResF*; 5' - ATGCGCCTGGCACTCCTTGAC-3'), in combination with the vector specific primer *TransgeneRev* 5' - GAAAACCAATGCTTAGAG-3' from RNA derived from the recipient *dpy-18* strains and the transformed lines.

RNA Interference - The effects of *phy* gene depletion were examined using standard RNAi injection protocols, and double-stranded RNA for injection was produced *in vitro* as described previously (17). RNA was synthesised from full-length coding sequence corresponding to *Cb-phy-1* and *phy-2* genes minus the signal peptide encoding sequence from the recombinant expression cloning plasmids described above. Ten young adults of wild type and *dpy-18* (*e364* and *e1096*) *C. elegans* and wild type and *dpy-18* (*mf104*) *C. briggsae* strains were microinjected, allowed to recover overnight, and transferred singly to fresh plates, and their progenies scored.

Microscopy - All nematodes were viewed either via differential interference contrast (DIC) or epifluorescence on a Zeiss Axioskop 2 microscope and images were taken with a Zeiss MrS digital camera and assembled in Adobe Photoshop.

Results

Cloning of *Caenorhabditis briggsae* Collagen P4H α and β Subunits - Analysis of the *C. briggsae* database by BLAST with the *C. elegans* PHY-1, PHY-2 and PDI-2 sequences revealed the presence of predicted genes encoding the corresponding homologous proteins. Cloning and sequencing of the cDNAs confirmed the expression and splicing pattern of the predicted genes, which are available under the accession numbers AM260973 (*Cb-phy-1*), AM260974 (*Cb-phy-2*) and AM260975 (*Cb-pdi-2*). A very high

degree of amino acid sequence identity is found between the *C. briggsae* and *C. elegans* collagen P4H subunits (Fig. 1), that of PHY-1 and PHY-2 sequences between the two species being 92% and 96%, respectively, and that between the PDI-2 polypeptides being 97%. The cysteine residues required for intrachain disulfide bonding (4, 32) and the catalytic residues required for binding of Fe²⁺ and 2-oxoglutarate (4, 33) are all conserved (Fig. 1). The PHY-1 polypeptides from both species have an 18-amino-acid C-terminal extension (Fig. 1), which is also found in the PHY-1 polypeptide of the filarial nematode *Onchocerca volvulus* (34), but not in that of *Brugia malayi* (31), nor in the *Caenorhabditis* PHY-2 polypeptides or vertebrate and *Drosophila* α subunits (9-12, 15, 17, 30).

Generation of a *C. briggsae phy-1* (*dpy-18*) Mutant and Its Rescue - A *C. briggsae phy-1* mutant was isolated following a *Mos1* insertion event. Like the *C. elegans phy-1* (*dpy-18*) mutants (Fig. 2F), the *C. briggsae phy-1* mutants were shortened in length relative to the wild type (Fig. 2A, B), the length reduction in both mutant species being comparable. The *C. briggsae phy-1* mutants differed from the *C. elegans dpy-18* in that the former did not display the associated circumferential increase (compare Fig. 2B, F). Therefore, the *phy-1* mutation in *C. briggsae* causes a small (Sma) phenotype rather than the more severe dumpy (Dpy, short and fat) phenotype, but for consistency in nomenclature the *C. briggsae phy-1* mutants will also be referred to as *C. briggsae dpy-18*.

All possible variations of gene rescue experiments were carried out on the *C. elegans* and *C. briggsae dpy-18* mutants using genomic *phy-1* and *phy-2* rescue constructs containing wild type copies of the genes from both species. The rescue constructs were co-injected into the mutant strains with a visible marker construct, the *dpy-7::gfp* cuticle collagen. As previously reported (17) *Ce-phy-1* could completely rescue the *C. elegans dpy-18* phenotype (Table 1). In contrast, the *Ce-phy-1* rescue of the *C. briggsae dpy-18* phenotype was only partial (Table 1) as was *Cb-phy-1* rescue of the *C. briggsae dpy-18* mutant (Table 1, Fig. 2C, D). *Cb-phy-1* was however able to fully rescue *C. elegans dpy-18* mutant (Table 1, Fig. 2G, H). *Ce-phy-2* partially rescued the *C. elegans dpy-18* mutant (Table 1). This may be explained by the over-expression of PHY-2 that could be expected to produce high levels of the unusual PHY-

2/PDI-2 dimer, which is normally present only in very low abundance (20). *Ce-phy-2* could not rescue the *C. briggsae dpy-18*, however, and *Cb-phy-2* failed to rescue either mutant (Table 1). Expression of the transgenes in the *C. elegans* and *C. briggsae dpy-18* nematodes was confirmed by RT-PCR using transgene-specific primers (Fig. 2I, J). All transformed *gfp* positive lines expressed the transgene (Fig. 2I, lanes 1, 3 and 5 and J, lanes 1, 3, 5 and 7), whereas the non-transformed control nematodes did not (Fig. 2I, lanes 2, 4 and 6 and J, lanes 2, 4, 6 and 8).

RNA Interference to Examine the Combined Function of the Cb-phy-1 and Cb-phy-2 Genes - The combined essential function of *Ce-phy-1* and *Ce-phy-2* was previously proven following *Ce-phy-2* RNAi in a *C. elegans dpy-18* mutant background, which resulted in an embryonic lethal (Emb) phenotype (17). This synthetic lethal phenotype has been confirmed following construction of a *dpy-18/phy-2* combined genetic mutant (19); Winter & Page, unpublished observations). To address this issue further in both species, a range of RNAi experiments were carried out in wild type *C. elegans* and *C. briggsae* and in the *C. briggsae dpy-18* mutants (Table 2, Fig. 3). *Cb-phy-1* RNAi in the wild type *C. elegans* and *C. briggsae* strains resulted in slight to medium Dpy and Sma phenotypes, respectively (Table 2). *Cb-phy-2* RNAi in either wild type strain failed to produce any visible phenotype, whereas it produced in the *C. briggsae dpy-18* mutant a severe Dpy and larval arrest with coiled larval phenotypes (Table 2, Fig. 3). *Cb-phy-1* RNAi in the *C. briggsae dpy-18* mutant background failed to increase the severity of the *dpy-18* phenotype (Table 2), thereby supporting the premise that this is a null allele. In direct agreement with previous *Ce-pdi-2* RNAi experiments (17), *Cb-pdi-2* caused similar severe Dpy and embryonic lethal phenotypes irrespective of whether the RNA was injected into a *C. elegans* or *C. briggsae* wild type background (Table 2).

Recombinant C. briggsae PHY-2 Assembles Into a Unique Active Collagen P4H Form with PDI-2 - The *Cb-phy-1* genetic null and RNAi results indicated similarities and highlighted subtle differences between the functional properties of the collagen P4Hs assembled from the highly conserved *C. elegans* and *C. briggsae* PHY-1, PHY-2 and PDI-2 polypeptides. To investigate this at a biochemical level, the recombinant *C.*

briggsae subunits and their collagen P4H forms were characterized following the infection of insect cells with viruses coding for *C. briggsae* PHY-1 and PHY-2, individually or in combination, along with *C. briggsae* or *C. elegans* PDI-2. In control experiments similar infections using viruses coding for the corresponding *C. elegans* PHY polypeptides were performed. The cells were harvested 72 h after infection, homogenized in a buffer containing 0.1% Triton X-100, and centrifuged. The Triton X-100-soluble extracts were analyzed by nondenaturing PAGE followed by Western blotting with antibodies against PHY-1 and PHY-2 (Fig. 4). P4H activity of Triton X-100-soluble extracts was analyzed with an assay based on the hydroxylation-coupled decarboxylation of 2-oxo-[1-¹⁴C]glutarate using a synthetic peptide (Pro-Pro-Gly)₁₀ as a substrate (Table 3). The ability of PHY-1 to form an active dimer with PDI-2 was shared between *C. briggsae* and *C. elegans* and was independent of the *Caenorhabditis* source of PDI-2 (Fig. 4, lanes 1, 2, 5 and 6). The *C. briggsae* PHY-1 and PHY-2 polypeptides were similar to those of *C. elegans* in that they readily assembled into an active PHY-1/PHY-2/(PDI-2)₂ mixed tetramer, the assembly in both cases occurring with PDI-2 from both species (Fig. 4, lanes 9-12 and 21-24, Table 3). Furthermore, a mixed tetramer was also assembled when the PHY-1 and PHY-2 polypeptides were from the different *Caenorhabditis* species (data not shown). Interestingly, a *C. briggsae* PHY-1/PDI-2 dimer was not detected when insect cells expressed both *C. briggsae* PHY-1 and PHY-2 together with PDI-2, although it formed readily when only PHY-1 was present with PDI-2 (Fig. 4, lanes 5, 6, 11 and 12). Based on the recombinant expression data, assembly of the *C. briggsae* PHY-1/PHY-2/(PDI-2)₂ mixed tetramer is strongly favoured over that of the PHY-1/PDI-2 dimer.

The *C. briggsae* PHY-2 differed distinctly from that of *C. elegans* in that a broad PHY-2 reactive band was seen when insect cells were infected with viruses coding for *C. briggsae* PHY-2 and either of the two *Caenorhabditis* PDI-2 polypeptides (Fig. 4, compare lanes 19 and 20 to lanes 15 and 16). The mobility of the complex formed by *C. briggsae* PHY-2 and PDI-2 was between those of the mixed tetramer and the PHY-1/PDI-2 dimer (Fig. 4, compare lanes 19 and 20 to lanes 5, 6, 9, 10, 23 and 24). This indicates that *C. briggsae* PHY-2, unlike that of *C. elegans*, can readily

assemble into a unique soluble complex with PDI-2. Furthermore, P4H activity assay showed that this complex was a highly active P4H (Table 3). The same PHY-2 immunoreactive band was present in addition to the band corresponding to the mixed tetramer in the samples from cells coexpressing *C. briggsae* PHY-1 and PHY-2 and either PDI-2 (Fig. 4, lanes 23 and 24). P4H activity in these samples was 2.5-3-fold higher than in those from cells expressing only *C. briggsae* PHY-2 and PDI-2, suggesting that both enzyme forms contributed to the total activity. Interestingly, based on the nondenaturing PAGE analysis (Fig. 4) and P4H activity assays (Table 3), the *C. briggsae* PHY polypeptides seemed to assemble into active PHY-1/PDI-2 dimers and PHY-1/PHY-2/(PDI-2)₂ mixed tetramers more efficiently than their *C. elegans* counterparts. Based on SDS-PAGE analysis, the expression levels of the PHY polypeptides from both species were equal (not shown), indicating that the more efficient assembly of the *C. briggsae* PHY-1, PHY-2 and PDI-2 to the various active collagen P4H forms may be a true feature of these polypeptides.

As the mobility of the novel recombinant collagen P4H form assembled from *C. briggsae* PHY-2 and PDI-2 was between those of the mixed tetramer and the PHY-1/PDI-2 dimer, we analyzed its molecular weight by gel filtration in a calibrated HiPrep Sephacryl S-200 HR column. In control experiments, either a Triton X-100-soluble extract from insect cells coexpressing *C. elegans* PHY-1 and human PDI or a purified recombinant human collagen P4H was analyzed. The P4H assembled from *C. briggsae* PHY-2 and PDI-2 eluted in exactly the same position as the *C. elegans* PHY-1/PHY-2/(PDI-2)₂ mixed tetramer and the human collagen P4H-I tetramer (Fig. 5) and before the PHY-1/PDI dimer indicating that the novel P4H form is a (PHY-2)₂(PDI-2)₂ tetramer.

Comparison of the Collagen P4H Forms Assembled in Wild Type and Mutant C. briggsae Nematodes to the Recombinant Forms - The differences noted between *C. elegans* and *C. briggsae* in the assembly of active recombinant P4H forms were confirmed following characterization of native extracts prepared from wild type and *dpy-18* mutant *C. briggsae* worms by non-denaturing PAGE and Western blotting. Triton X-100-soluble extracts from insect cells co-expressing *C. briggsae* PHY-1, PHY-2 and PDI-2 or PHY-1

and PDI-2 were run as controls in the same gels. As described for *C. elegans* (20), PHY-1/PHY-2/(PDI-2)₂ tetramers and PHY-1/PDI-2 dimers but not PHY-2/PDI-2 dimers were simultaneously present in native extracts from wild type *C. briggsae* (Fig. 6, lanes 1 and 5). As the recombinant expression experiments showed that assembly of the recombinant *C. briggsae* PHY-1/PHY-2/(PDI-2)₂ mixed tetramer is clearly preferred over that of the PHY-1/PDI-2 dimer (Fig. 4, lanes 11 and 12 and Fig. 6, lane 3), the *in vivo* observation of simultaneous existence of both forms may therefore reflect a limitation in the *in vitro* availability of the PHY-1 polypeptide.

As expected, the *C. briggsae dpy-18* mutant was devoid of any enzyme forms containing PHY-1 (Fig. 6, lane 2) further confirming that it is a null mutant. This mutant did however possess a readily identifiable PHY-2 immunoreactive band that migrated slower than the PHY-1/PDI-2 dimer in the wild type but slightly faster than the mixed tetramer (Fig. 6, compare lane 6 to lanes 1 and 5) and had a similar mobility as the band corresponding to the recombinant (PHY-2)₂(PDI-2)₂ tetramers in insect cell expressions (Fig. 6, compare lanes 6 and 7). The mobility of the *C. briggsae dpy-18* PHY-2 immunoreactive band differs distinctly from that of the PHY-2/PDI-2 dimer present in *C. elegans dpy-18*, the mobility of which is identical to that of the PHY-1/PDI-2 dimer (20). The lack of the (PHY-2)₂(PDI-2)₂ tetramers in the lysates from wild type *C. briggsae* (Fig. 6, lane 5) may again reflect the relative abundance of the PHY polypeptides *in vivo*, *i.e.* limited availability of the PHY-2 polypeptides, and the favoured formation of the mixed tetramer relative to the other forms.

Double Mutagenesis of the C. elegans and C. briggsae PHY-1 and PHY-2 Tetratricopeptide Repeat Motifs Changes Their Assembly Properties Towards Those of the Other Species - A surprising result from the aforementioned *dpy-18* mutant analysis, RNAi, rescue and recombinant expression experiments was that the assembly properties of the *C. briggsae* collagen P4H subunits are not totally identical with those of *C. elegans* despite the extremely high amino acid sequence identity of the subunits. The major difference between *C. elegans* and *C. briggsae* was the ability of CbPHY-2 to efficiently form active (PHY-2)₂(PDI-2)₂ tetramers in addition to the PHY-1/PHY-2/(PDI-2)₂ tetramers and PHY-1/PDI-2 dimers found in both species

(Fig. 7). Furthermore, the assembly of the *C. briggsae* subunits into active P4H forms seemed to be more efficient in general than those of *C. elegans*. Comparison of the amino acid sequences of PHY-1 and PHY-2 from the two species indicated limited amino acid differences (Fig. 1), the majority of which reside in the N-terminal non-catalytic region of the proteins. A search for protein motifs in the PHY sequences using the InterProScan (EMBL-EBI) (35) indicated that the N-terminal regions contain a tetratricopeptide repeat (TPR)-like helical region between residues 134-237 in PHY-1 and 144-226 in PHY-2, with a 34-residue TPR motif between residues 188-221 and 190-223 of the PHY-1 and PHY-2 polypeptides, respectively (Fig. 1). An additional TPR motif was predicted to be present in this region between PHY-1 residues 146-179 and PHY-2 residues 148-181 (personal communication, Malcolm Walkinshaw, University of Edinburgh). TPR motifs mediate protein-protein and protein-peptide interactions, and the assembly of multiprotein complexes (36). The structure of the peptide-substrate-binding domain of human collagen P4H-I, comprising residues 144-244 of the 517-amino-acid α (I) subunit, has been recently shown to consist of two TPR motifs and a solvating helix (37). To study whether amino acid differences in the TPR motifs of the *C. elegans* and *C. briggsae* PHY polypeptides could explain the differences in their assembly properties, site directed mutagenesis was performed on two separate divergent amino acids in each of the PHY-1 and PHY-2 TPR motifs to change the *C. elegans* sequences to a more *C. briggsae*-like sequences and *vice versa*, resulting in Ce-PHY-1-T220A, Cb-PHY-1-A220T, Ce-PHY-2-W190S and Cb-PHY-2-S190W single mutants, and Ce-PHY-1-T220A/Q174K, Cb-PHY-1-A220T/K174Q, Ce-PHY-2-W190S/V177N and Cb-PHY-2-S190W/N177V double mutants (Fig. 1).

The mutant PHY-1 and PHY-2 polypeptides were coexpressed in insect cells together with *C. elegans* or *C. briggsae* PDI-2 and analyzed by non-denaturing PAGE followed by Western blotting and P4H activity assay (Fig. 8, Table 4). The double mutant *C. elegans* PHY-1 polypeptides assembled into a dimer with PDI-2 more efficiently than the wild type PHY-1, the increase in assembly being 3-4-fold based on the P4H activity obtained, while no significant increase was obtained with the single mutant (Fig. 8A,

Table 4). The assembly efficiency of the double mutant *C. elegans* PHY-1 with PDI-2 did not reach that of wild type *C. briggsae* PHY-1, however, the P4H activity obtained being only about 28% of that of the *C. briggsae* PHY-1/PDI-2 dimer (Table 4). The double mutation in the *C. briggsae* PHY-1 decreased its assembly with PDI-2 by about 2-fold relative to that of the wild type polypeptide, despite the single mutation first leading to an increase rather than a decrease in the assembly and activity (Fig. 8A, Table 4). Based on the activity assays, the assembly efficiency of the double mutant *C. briggsae* PHY-1 with PDI-2 was still about 5-fold relative to that of *C. elegans* PHY-1, however (Table 4).

The double mutant *C. elegans* PHY-2 assembled into a soluble complex with PDI-2 that was visible in non-denaturing PAGE after Western blotting and had a mobility similar to that of the complex formed by *C. briggsae* PHY-2 and PDI-2 (Fig. 8B). The P4H activity obtained with the double mutant *C. elegans* PHY-2 with PDI-2 was about 8-fold relative to that obtained with the wild type, but still only about 11% of that obtained with *C. briggsae* PHY-2 and PDI-2 (Table 4). The activity values obtained with the single mutant *C. elegans* PHY-2 with PDI-2 were between those obtained with the wild type and double mutant *C. elegans* PHY-2 with PDI-2 (Table 4). In contrast, the single and double mutations in *C. briggsae* PHY-2 did not change its assembly properties with PDI-2 (Fig. 8B), nor had any marked effect on the resulting P4H activity (Table 4).

The Ability of C. briggsae PHY-2 to Assemble Into a (PHY-2)₂/(PDI-2)₂ Tetramer Depends On Its N-Terminal Half - The results obtained above indicate that the two residues of the TPR motifs subjected to site directed mutagenesis are involved in the assembly of active P4H forms from the *C. elegans* and *C. briggsae* PHY polypeptides, but differences in these two amino acids did not fully explain the differences observed between the assembly properties of the subunits from the two species. We therefore extended these studies by replacing the N-terminal 262 amino acids of the *C. briggsae* and *C. elegans* PHY-2 polypeptides with those from the other species. The exchange of this N-terminal region of *C. briggsae* PHY-2 with that of *C. elegans* PHY-2 resulted in a dramatic loss of the ability to form the (PHY-2)₂/(PDI-2)₂ tetramer (Fig. 9, lane 3) and a 35-45-fold decrease in the P4H

activity (Table 4). In contrast, replacement of the N-terminal 262 amino acids of *C. elegans* PHY-2 with those of *C. briggsae* PHY-2 resulted in efficient formation of the (PHY-2)₂/(PDI-2)₂ tetramer (Fig. 9, lane 4) and an essentially identical P4H activity relative to that obtained with wild type *C. briggsae* PHY-1 and PDI-2 (Table 4).

Discussion

The collagen P4Hs are essential enzymes as hydroxylation of X-Pro-Gly sequences in collagen chains is necessary for the formation of collagen molecules with stable triple-helical structures (2). These enzymes thus have important implications in tissue remodelling and as potential drug targets for the treatment of human fibrotic diseases caused by excessive collagen accumulation (3). Previous biochemical and genetic characterization of the *C. elegans* cuticle collagen P4Hs indicated that their subunit composition differed from those described in other species, including man (17, 20, 21). The main enzyme form in *C. elegans* is a PHY-1/PHY-2/(PDI-2)₂ tetramer that contains two distinct catalytic PHY subunits. In addition, assembly of PHY-1/PDI-2 and PHY-2/PDI-2 dimers is possible but much less efficient. RNA interference studies and analysis of *C. elegans* strains with gene deletions have shown that inactivation of *phy-1* and *phy-2* jointly, or *pdi-2* singly, leads to embryonic lethality due to abnormal cuticle collagen synthesis. In the present work we have further studied the *Caenorhabditis* collagen P4Hs by characterizing the orthologous genes and encoded proteins from the closely related free-living nematode *C. briggsae*. The combined essential function of the *phy-1* and *phy-2* genes was found to be a consistent feature shared between the two *Caenorhabditis* species thus supporting their vital role in the formation of a proper exoskeleton and in nematode survival beyond embryogenesis. This makes these collagen P4Hs potential novel drug targets for the control of medically and veterinary significant nematode species (23).

In spite of the extremely high amino acid sequence conservation between the catalytic PHY subunits of *C. elegans* and *C. briggsae*, distinct differences were noted in the assembly properties of the PHY subunits. In addition to assembly of PHY-1/PHY-2/(PDI-2)₂ tetramers and PHY-1/PDI-2 dimers in both species, the *C. briggsae* PHY-2 formed an active (PHY-2)₂/(PDI-2)₂ tetramer instead of the PHY-

2/PDI-2 dimer that is formed in *C. elegans*. This observation was made with both recombinant and native *C. briggsae* PHY subunits by analyzing co-expression experiments in insect cells and protein extracts from wild type and *C. briggsae dpy-18* nematodes (Figs. 4 and 6). The subunit composition of the active (PHY-2)₂/(PDI-2)₂ *C. briggsae* tetramer represents that commonly present in man, insects, birds and rodents and therefore reflects the consensus subunit composition of animal collagen P4Hs, the mixed tetramer and dimers being unique to nematodes. Based on recombinant expression experiments, assembly of the *C. briggsae* PHY subunits into active P4H forms was in general more efficient than that of the *C. elegans* ones, as reflected by the amount of the different P4H forms observed in Western blot analysis (Fig. 4) and the P4H activity generated (Table 1). It is especially noteworthy that the *C. briggsae* PHY-2 readily assembled into an active (PHY-2)₂/(PDI-2)₂ tetramer in insect cell expression experiments (Fig. 4 and Table 1). In contrast, assembly of recombinant *C. elegans* PHY-2 into an active PHY-2/PDI-2 dimer is very inefficient and the dimer is only observed at Western blot level in *C. elegans dpy-18* nematodes where its assembly is enhanced (Fig. 4, Table 1 and (20)).

The *C. elegans* genome contains over 170 cuticle collagen genes, a small subset of which results in morphological phenotypes when mutated, an even smaller subset of these genes being essential as evidenced by lethality caused by their mutation (23). Although the PHY-1/PHY-2/(PDI-2)₂ mixed tetramer is the major cuticle collagen P4H in the two *Caenorhabditis* species, both species can survive normally in the absence of PHY-2, as assessed by the RNAi depletion experiments (Fig. 2 and (17)) and by the deletion of the *C. elegans phy-2* gene (19). The *C. elegans phy-2* null deletion mutants can compensate for the lack of the mixed tetramer by markedly increasing the assembly of the PHY-1/PDI-2 dimer resulting in retention of 54-57% of the P4H activity (20) and 64% of the 4-hydroxyproline content (19) relative to the wild type, these values being sufficient to maintain a wild type appearance. A similar compensation mechanism is likely to exist in *C. briggsae* as recombinant expression experiments showed that the *C. briggsae* PHY-1 subunit can also assemble into an active dimer with PDI-2. Furthermore, as lack of PHY-2 has no effect on cuticle synthesis, it

can be deduced that the major collagen P4Hs of wild type *Caenorhabditis* nematodes, the PHY-1/PHY-2/(PDI-2)₂ tetramers, and the PHY-1/PDI dimers present in *phy-2* mutant nematodes have identical substrate specificities.

Unlike the lack of PHY-2, lack of PHY-1 affects the cuticular extracellular matrix and leads to a deformed body shape in both *Caenorhabditis* species as indicated by the RNAi experiments and by the characterization of *phy-1* null mutants (Fig. 2, (17, 19, 20)). The resulting *phy-1* mutant phenotypes were not exactly identical, however, being small (Sma) in *C. briggsae* and a combination of small and fat (Dpy) in *C. elegans* (Fig. 2). The nematode cuticle has two distinct regions, namely those derived from the lateral hypodermal seam cells and the dorso-ventral hypodermis (38, 39). The structural properties of the two regions differ and mutations in individual collagen genes can affect one or both of these regions and impart distinct region-specific phenotypes (39). It can be extrapolated that the Sma phenotype is less severe than the Dpy phenotype and is a result of milder disruption affecting only the lateral cuticle structures. The Dpy phenotype, on the other hand, has a more global affect due to disruption of both the lateral and the circumferential dorso-ventral cuticle. The less severe phenotype in the *phy-1* mutant *C. briggsae* may reflect the fact that the remaining PHY-2 isoform can readily assemble into an active (PHY-2)₂(PDI-2)₂ tetramer as evidenced by recombinant coexpression studies (Fig. 4 and Table 3), while the corresponding *C. elegans* subunits associate into a PHY-2/PDI-2 dimer much less efficiently (Fig. 4 and Table 3), the remaining P4H activity in *phy-1* mutant *C. elegans* nematodes being only 0.7-2.5% relative to the wild type (20). It is also possible that the substrate specificities of the *C. briggsae* (PHY-2)₂(PDI-2)₂ tetramers and *C. elegans* PHY-2/PDI-2 dimers are not identical and they may therefore modify different subsets of the cuticle collagens, which in turn are incorporated into the two structurally distinct regions of the cuticle.

The orthologous nature of the *C. elegans* and *C. briggsae* *phy-1* and *phy-2* genes was studied here by detailed analysis of *dpy-18* (*phy-1*) mutants of both species and the ability of intra-species and inter-species *phy-1* and *phy-2* genes to rescue the associated body morphology defects. As expected, the

complementation experiments revealed that the *phy-1* transgenes from the same species were able to rescue the corresponding mutant phenotypes, namely *Ce-phy-1* can rescue *Ce-dpy-18* and *Cb-phy-1* can partially rescue *Cb-dpy-18*. In addition, the *phy-1* genomic rescue constructs were able to completely or partially rescue the mutant phenotype of the other species and thus showed that the encoded polypeptides are highly similar in their structural and functional properties. The reason why only partial rescue was obtained by injection of *Ce-phy-1* into *Cb-dpy-18* may be attributed to the fact that this construct was toxic at the standard injection concentration and was subsequently introduced at a five-fold lower concentration. Although expression of the transgene was detected in the partially rescued lines (Fig. 2J), its expression level may have been lower than that of the endogenous gene in wild type *C. briggsae*. The partial rescue of *Ce-dpy-18* by a *Ce-phy-2* transgene was unexpected, as no rescue has been obtained in previous attempts (Winter and Page, unpublished observations). This may be explained by the use of genomic *phy-2* rescue constructs in the current experiments instead of cDNA-derived constructs used in the previous studies. The most likely explanation for the partial rescue of *Ce-dpy-18* by a *Ce-phy-2* transgene is a further increase in the amount of the PHY-2/PDI-2 dimer. The lack of rescue in the remaining experiments, namely that of *Ce-dpy-18* with *Cb-phy-2* and *Cb-dpy-18* with *C. briggsae* or *C. elegans phy-2*, should be taken as an indication rather than a conclusive lack of compensation in spite of the fact that RT-PCR controls confirmed the expression of the transgenes. It is important to note that all *C. briggsae* rescue constructs were made as a fusion to the *C. elegans phy-1* promoter and 3' UTR regulatory sequences and this may have exceeded the limits of the transgenic complementation system, a point supported by the fact that the *Cb-phy-1* completely rescued the *Ce-dpy-18* mutant but only partially rescued by the *Cb-dpy-18* mutant. In the case of the *phy-2* experiments, lack of rescue may also be compounded by differences in expression levels or the temporal control of expression relative to the endogenous genes. Lack of inter-species *phy-2* rescue may also be a true phenomenon, however, and reflect differences in substrate specificities and/or assembly efficiencies of the P4H forms that contain only PHY-2 as the catalytic subunit.

The fact that despite the extremely high amino acid sequence identity between the *C. briggsae* and *C. elegans* PHY-2 polypeptides, they assembled with PDI into different active P4H forms when no PHY-1 was present, i.e. tetramers and dimers, respectively, enabled us to identify the regions of the PHY-2 polypeptides that are responsible for this difference. Motif searches identified two TPR-like domains between residues 148-223 of the PHY-2 polypeptides. TPR motifs are involved in many protein-protein interactions and assembly of multi-protein complexes (36) and indeed, a 2.3Å crystal structure of the peptide-substrate-binding domain of the α (I) subunit of human type I collagen P4H has shown that it is a TPR domain with functionally important aromatic residues (37). Site-directed mutagenesis was therefore applied to determine the effect of two amino acid differences between the *C. briggsae* and *C. elegans* PHY-2 TPR domains on the assembly properties of the full-length polypeptides. Double mutagenesis of the *C. elegans* and *C. briggsae* PHY-2 TPR motifs changed their assembly properties towards that of the other species but did not lead to a complete reversal, which was achieved only after the entire N-terminal halves of the polypeptides were swapped. These results indicate that although the identified TPR residues are important in determining the assembly properties of a PHY-2 polypeptide with PDI-2, other adjacent residues in the N-terminal half also play a vital role. Previous domain swap experiments have shown that residues Gln121-Ala271 and Asp1-Leu122 of *C. elegans* PHY-1 and PHY-2, respectively, are critical for the assembly of the mixed PHY-1/PHY-2/(PDI-2)₂ tetramer (20). The current study further emphasizes the importance of the N-terminal non-catalytic half of the PHY polypeptides in dictating their assembly properties.

It is interesting to note that the other available PHY-2 sequence, from the closely related *Caenorhabditis ramanai* (wormbase.org; and Winter and Page, unpublished results), is 97.5% and 94.5% identical to the *C. briggsae* and to *C. elegans* sequences, respectively. In addition, the aforementioned two crucial amino acids of the TPR domain shown to be involved in active (PHY-2)₂(PDI-2)₂ formation are completely conserved between *C. briggsae* and *C.*

ramanai, whereas these amino acids from the *Brugia malayi* (31) and *Onchocerca volvulus* (34) PHY sequences differ from all three *Caenorhabditis* species. It may be proposed that the presence of two *phy* genes (*phy-1* and *phy-2*) results in partial redundancy, thereby making them less-sensitive to mutation and ultimately allowing for the evolution of their respective roles and the wide range of active complexes that they form.

The results presented in this study reinforce the findings that cuticle collagen P4Hs are essential for the survival of nematode species, but also emphasize the fact that the assembly efficiency of active P4H forms and even their molecular composition can differ between two closely related species despite extremely high amino acid sequence identity of the enzyme subunits. Therefore, as shown here, mutation of genes that are highly conserved between two species may not necessarily have identical phenotypic consequences. The existence of unique collagen P4H forms in nematodes, PHY-1/PHY-2/(PDI-2)₂ and (PHY-2)₂/(PDI-2)₂ tetramers and PHY-1/PDI-2 and PHY-2/PDI-2 dimers in the *Caenorhabditis* species and even active (PHY-1)₄ tetramers lacking a PDI/ β subunit in *Brugia malayi* (31) will help to elucidate the assembly process of collagen P4Hs in general.

Acknowledgements

This work was supported the Medical Research Council of Great Britain through the award of a Senior Fellowship (G117/476) to APP; and by the Health Science Council (Grants 200471 and 202469), the Finnish Centre of Excellence Programme 2000-2006 (Grant 44843) of the Academy of Finland and from the S. Juselius Foundation to JM; and an EMBO Short Term Fellowship (ASTF 44.03) to ADW. We thank Raija Juntunen, Eeva Lehtimäki and Liisa Äijälä for expert technical assistance. V. Robert and J. -L. Bessereau are thanked for plasmids and their precious advice on the *Mos1* mutagenesis and A. Fire and his laboratory for plasmids.

References

1. Kivirikko, K. I., and Pihlajaniemi, T. (1998) *Adv. Enzymol. Rel. Mol. Biol.* **72**, 325-400
2. Myllyharju, J. (2003) *Matrix Biol.* **22**(1), 15-24
3. Myllyharju, J., and Kivirikko, K. I. (2004) *Trends Genet.* **20**(1), 33-43
4. Lamberg, A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) *J. Biol. Chem.* **270**(17), 9926-9931
5. Myllyharju, J., and Kivirikko, K. (1997) *EMBO J.* **16**, 1173-1180
6. Myllyharju, J., and Kivirikko, K. (1999) *EMBO J.* **18**, 306-312
7. Hieta, R., Kukkola, L., Permi, P., Pirila, P., Kivirikko, K. I., Kilpelainen, I., and Myllyharju, J. (2003) *J. Biol. Chem.* **278**(37), 34966-34974
8. Vuori, K., Pihlajaniemi, T., Myllyla, R., and Kivirikko, K. I. (1992) *EMBO J.* **11**(11), 4213-4217
9. Helaakoski, T., Annunen, P., Vuori, K., MacNeil, I. A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4427- 4431
10. Annunen, P., Helaakoski, T., Myllyharju, J., Veijola, J., Pihlajaniemi, T., and Kivirikko, K. I. (1997) *J. Biol. Chem.* **272**(28), 17342- 17348
11. Kukkola, L., Hieta, R., Kivirikko, K. I., and Myllyharju, J. (2003) *J. Biol. Chem.* **278**(48), 47685-47693
12. Van den Diepstraten, C., Papay, K., Bolender, Z., Brown, A., and Pickering, J. G. (2003) *Circulation* **108**(5), 508-511
13. Annunen, P., AutioHarmainen, E., and Kivirikko, K. I. (1998) *J. Biol. Chem.* **273**(11), 5989-5992
14. Nissi, R., Autio-Harmainen, H., Marttila, P., Sormunen, R., and Kivirikko, K. I. (2001) *J. Histochem. Cytochem.* **49**(9), 1143-1153
15. Annunen, P., Koivunen, P., and Kivirikko, K. I. (1999) *J. Biol. Chem.* **274**, 6790-6796
16. Abrams, E. W., and Andrew, D. J. (2002) *Mech. Dev.* **112**(1-2), 165-171
17. Winter, A. D., and Page, A. P. (2000) *Mol. Cell. Biol.* **20**, 4084-4093
18. Hill, K. L., Harfe, B. D., Dobbins, C. A., and Hernault, S. W. L. (2000) *Genetics* **155**, 1139-1148
19. Friedman, L., Higgin, J. J., Moulder, G., Barstead, R., Raines, R. T., and Kimble, J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 4736-4741
20. Myllyharju, J., Kukkola, L., Winter, A.D., and Page, A.P. (2002) *J. Biol. Chem.* **277**, 29187-29196
21. Veijola, J., Koivunen, P., Annunen, P., Pihlajaniemi, T., and Kivirikko, K. I. (1994) *J. Biol. Chem.* **269**(43), 26746-26753
22. Veijola, J., Annunen, P., Koivunen, P., Page, A. P., Pihlajaniemi, T., and Kivirikko, K. I. (1996) *Biochem. J.* **317**, 721-729
23. Page, A. P., and Winter, A. D. (2003) *Adv. Parasitol.* **53**, 85-148
24. Riihimaa, P., Nissi, R., Page, A.P., Winter, A.D., Keskiaho, K., Kivirikko, K. I., and Myllyharju, J. (2002) *J. Biol. Chem.* **277**, 18238-18243.
25. Stein, L. D., Bao, Z. R., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N. S., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H. A., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C. C., Willey, D., Wilson, R. K., Durbin, R., and Waterston, R. H. (2003) *Plos Biol.* **1**(2), 166-+
26. Bessereau, J.-L., Wright, A., Williams, D. C., Schuske, K., Davis, M. W., and Jorgensen, E. M. (2001) *Nature* **413**, 70-74
27. Williams, D. C., Boulin, T., Ruaud, A. F., Jorgensen, E. M., and Bessereau, J. L. (2005) *Genetics* **169**(3), 1779-1785
28. Crossen, R., and Gruenwald, S. (1998) *Pharmingen, San Diego, CA*
29. Kivirikko, K. I., and Myllylä, R. (1982) *Meth. Enzymol.* **82**, 245-304
30. Vuori, K., Pihlajaniemi, T., Marttila, M., and Kivirikko, K.I. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7467-7470
31. Winter, A. D., Myllyharju, J., and Page, A. P. (2003) *J. Biol. Chem.* **278**, 2554-2562
32. John, D. C. A., and Bulleid, N. J. (1994) *Biochem.* **33**(47), 14018-14025

33. Myllyharju, J., Lamberg, A., Notbohm, H., Fietzek, P. P., Pihlajaniemi, T., and Kivirikko, K. I. (1997) *J. Biol. Chem.* **272**, 21824-21830
34. Merriweather, A., Guenzler, V., Brenner, M., and Unnasch, T. R. (2001) *Mol. Biochem. Parasitol.* **116**(2), 185-197
35. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005) *Nuc. Acid. Res.* **33**, W116-W120
36. D'andrea, L. D., and Regan, L. (2003) *Trends Biochem. Sci.* **28**(12), 655-662
37. Pekkala, M., Hieta, R., Bergmann, U., Kivirikko, K. I., Wierenga, R. K., and Myllyharju, J. (2004) *J. Biol. Chem.* **279**(50), 52255-52261
38. Sulston, J., and Horvitz, H. (1977) *Dev. Biol.* **56**, 110-156
39. Thein, M. C., McCormack, G., Winter, A. D., Johnstone, I. L., Shoemaker, C. B., and Page, A. P. (2003) *Dev. Dyn.* **226**, 523-539

Footnote

The gene sequences described in this paper have been deposited in the public databases under accession numbers, AM260973 (*Cb-phy-1*), AM260974 (*Cb-phy-2*) and AM260975 (*Cb-pdi-2*).

Figure Legends

FIGURE 1. Amino acid alignment of the *C. briggsae* (*Cb* prefix) PHY-1 (GenBank no. AM260973) and PHY-2 (GenBank no. AM260974) with *C. elegans* (*Ce* prefix) PHY-1 (GenBank no. Z81134) and PHY-2 (GenBank no. Z69637). Numbering starts from the mature polypeptides from which the signal peptide sequences have been removed. Gaps (dashes) were introduced for maximal alignment. The cysteine residues required for intrachain disulfide binding and the histidine, aspartic acid and lysine residues required for binding of Fe²⁺ and 2-oxoglutarate at the catalytic site are indicated with an asterisk. The predicted TPR motifs are indicated by an *inverted L*. The sites of the point mutations introduced into the TPR motifs are indicated by a black downward arrow in the PHY-1 and a grey downward arrow in the PHY-2 sequences and the site of exchange to generate inter-species hybrid PHY-1 and PHY-2 polypeptides are indicated by an upward arrow.

FIGURE 2. The *C. briggsae phy-1 (dpy-18)* mutant and rescue of the *C. briggsae* and *C. elegans dpy-18* strains. *A-C* and *E-F*, Nomarski images of adult hermaphrodites. *A*, Wild type *C. briggsae* (length approximately 1050 μm). *B*, *Cb-dpy-18 (mf104)* (length approximately 670 μm). *C*, Partial rescue of *Cb-dpy-18 (mf104)* by injection of a wild type copy of *Cb-phy-1*. The two uppermost nematodes are rescued (lengths approximately 920 and 930 μm) relative to the non-rescued nematode below (length approximately 720 μm). *D*, Epifluorescence image of the nematodes shown in *C*, depicting expression of GFP from the co-injected selectable marker *dpy-7::gfp* in the rescued nematodes, but missing it in the non-rescued one. *E*, Wild type *C. elegans* (length approximately 1000 μm). *F*, *Ce-dpy-18 (e1096)* (length approximately 700 μm). *G*, Rescue of *Ce-dpy-18 (e1096)* by injection of a wild type copy of *Cb-phy-1* (length approximately 1050 μm). *H*, Epifluorescence image of the rescued nematode shown in *G*, depicting expression of GFP from the co-injected selectable marker *dpy-7::gfp*. Scale bars, 100 μm. *I* and *J*, RT-PCR confirmation of transgene expression. *I*, Transgene expression in transformed (*lanes 1, 3* and *5*) and non-transformed (*lanes 2, 4* and *6*) *Ce-dpy-18*. *J*, Transgene expression in transformed (*lanes 1, 3, 5* and *7*) and non-transformed (*lanes 2, 4, 6* and *8*) *Cb-dpy-18*. The transgene expressed is denoted above individual *lanes*. The sizes of the products, approximately 1.6 kb, confirm that the amplification was derived exclusively from the spliced transgenic messages.

FIGURE 3. The effect of RNAi of *phy-2* in *C. briggsae dpy-18 (mf104)* mutant background. Double-stranded RNA corresponding to the *C. briggsae phy-2* coding sequence was injected to the *C. briggsae dpy-18* nematodes and the phenotypes of the progeny were analyzed. *A*, Newly hatched first stage larva (L1) displaying severe Dpy and coiling phenotypes. *B*, Typical severe Dpy L1 larvae displaying a larval arrest phenotype. Scale bar, 50 μm.

FIGURE 4. Analysis of P4H forms assembled from recombinant *C. briggsae* P4H subunits. Recombinant *C. briggsae* PHY-1 and PHY-2 were coexpressed in insect cells together with *C. briggsae* or *C. elegans* PDI-2 polypeptides. The corresponding *C. elegans* PHY polypeptides were expressed in control experiments. The cells were harvested 72 h after infection, homogenized in a

Triton X-100-containing buffer, and centrifuged. The Triton X-100-soluble fractions were analyzed by nondenaturing PAGE followed by Western blotting using antibodies against PHY-1 (top panel) and PHY-2 (bottom panel). The positions of the P4H tetramers (T) and dimers (D) are indicated by arrows.

FIGURE 5. Gel filtration analysis of a Triton X-100-soluble fraction from insect cells expressing recombinant *C. briggsae* PHY-2 and PDI-2. A Triton X-100-soluble fraction from insect cells expressing recombinant *C. briggsae* PHY-2 and PDI-2 was applied to a Superdex 200 column, and the eluted fractions were assayed for P4H activity (— ■ —). In control experiments, a Triton X-100-soluble fraction from insect cells expressing recombinant *C. elegans* PHY-1 and human PDI-2 and purified recombinant human type I collagen P4H were applied to the column, and the fractions were assayed for P4H activity (— ▲ —) and absorbance at 280 nm (--- ▲ ---), respectively.

FIGURE 6. Analysis of P4H forms in wild type and *dpy-18* *C. briggsae* in vivo. Nematodes from the wild type (*lanes 1* and *5*) and *dpy-18* (*mf104*) (*lanes 2* and *6*) *C. briggsae* strains were homogenized in a buffer containing Triton X-100, and the soluble fractions were analyzed by nondenaturing PAGE followed by Western blotting using antibodies against PHY-1 (*A*) and PHY-2 (*B*). Triton X-100-soluble proteins from insect cells coexpressing *C. briggsae* PHY-1, PHY-2 and PDI-2 were used as a control in *A* and *B* (*lanes 3* and *7*) and PHY-1 and PDI-2 in panel *A* (*lane 4*). The positions of the PHY-1/PHY-2/(PDI-2)₂ and (PHY-2)₂/(PDI-2)₂ tetramers and PHY-1/PDI-2 dimers are indicated by arrows.

FIGURE 7. Schematic representation of the P4H forms present in wild type and mutant *C. briggsae* and *C. elegans* strains. The PHY-1/PHY-2/(PDI-2)₂ mixed tetramers and PHY-1/PDI-2 dimers are present in both wild type *C. briggsae* and *C. elegans*. In the absence of the PHY-1 polypeptide, a (PHY-2)₂/(PDI-2)₂ tetramer is assembled in *C. briggsae dpy-18*, while a PHY-2/PDI-2 dimer is found in the *C. elegans dpy-18*. When the PHY-2 polypeptide is lacking, a PHY-1/PDI-2 dimer is still present and the assembly of this dimer has been previously shown to be markedly increased in *C. elegans* relative to the wild type (20). When both PHY-1 and PHY-2 are lacking, no active collagen P4H is formed in either species.

FIGURE 8. Effect of point mutations in the TPR motifs of *C. briggsae* and *C. elegans* PHY-1 and PHY-2 on their assembly properties with PDI-2. Two point mutations were introduced into the TPR motifs of the *Caenorhabditis* PHY polypeptides by site directed mutagenesis to convert these residues to ones present in the other species. The double mutant (*DM*) Ce-PHY-1-T220A/Q147K, Cb-PHY-1-A220T/K174Q, Ce-PHY-2-W190S/V177N and Cb-PHY-2-S190W/N177V, and the wild type PHY-1 and PHY-2 from both species were expressed in insect cells together with PDI-2, and the cells were harvested and homogenized as described in the legend to Fig. 4. Triton X-100-soluble fractions were analyzed by non-denaturing PAGE followed by Western blotting using antibodies against PHY-1 (*A*) and PHY-2 (*B*). The positions of the (PHY-2)₂/(PDI-2)₂ tetramers (T) and PHY-1/PDI-2 dimers (D) are indicated by arrows.

FIGURE 9. Analysis of tetramer formation by inter-species hybrid *Caenorhabditis* PHY-2 polypeptides with PDI-2. Hybrid PHY-2 polypeptides in which the N-terminal 262 amino acids were changed between *C. elegans* and *C. briggsae*, named CeN/CbC-PHY-2 and CbN/CeC-PHY-2, and wild type PHY-2 from both species were expressed in insect cells together with PDI-2. The cells were harvested and homogenized as described in the legend to Fig. 4, and Triton X-100-soluble fractions were analyzed by non-denaturing PAGE followed by Western blotting using an antibody against PHY-2. The position of the (PHY-2)₂/(PDI-2)₂ tetramers (T) is indicated by an arrow.

TABLE 1

Rescue experiments of *C. elegans* and *C. briggsae* *dpy-18* phenotype by injection of wild type *phy-1* and *phy-2* genomic constructs from both species

Genomic rescue clone injected	Recipient Strain	Outcome
<i>Ce-phy-1</i>	<i>C. elegans dpy-18 (e364 & e1096)</i>	Rescue
<i>Ce-phy-2</i>	<i>C. elegans dpy-18 (e364)</i>	Partial Rescue
<i>Cb-phy-1</i>	<i>C. elegans dpy-18 (e1096)</i>	Rescue
<i>Cb-phy-2</i>	<i>C. elegans dpy-18 (e1096 & e364)</i>	No Rescue
<i>Cb-phy-1</i>	<i>C. briggsae dpy-18 (mf104)</i>	Partial Rescue
<i>Cb-phy-2</i>	<i>C. briggsae dpy-18 (mf104)</i>	No Rescue
<i>Ce-phy-1</i>	<i>C. briggsae dpy-18 (mf104)</i>	Partial Rescue
<i>Ce-phy-2</i>	<i>C. briggsae dpy-18 (mf104)</i>	No Rescue

TABLE 2**Effect of double-stranded RNA interference of *C. briggsae phy-1*, *phy-2* and *pdi-2* in wild type *C. elegans* and *C. briggsae* and *C. briggsae dpy-18***

RNAi construct injected	Effect on <i>Caenorhabditis</i> strains		
	Wild type <i>C. elegans</i>	Wild type <i>C. briggsae</i>	<i>C. briggsae dpy-18</i>
<i>Cb-phy-1</i>	Slight/medium Dpy	Slight Dpy/ Sma	Slight Dpy/ Sma
<i>Cb-phy-2</i>	Wild type	Wild type	Severe Dpy/ Lva
<i>Cb-pdi-2</i>	Severe Dpy/Emb	Severe Dpy/Emb	ND

Dpy, dumpy; Sma, small; Lva, larval lethal; Emb, embryonic lethal; ND, not determined.

TABLE 3

Collagen P4H activity in Triton X-100 extracts of insect cells expressing recombinant *C. elegans* (prefix Ce) and *C. briggsae* (prefix Cb) PHY-1, PHY-2 and PDI-2 polypeptides in various combinations.

Values are given in dpm/100 μ g of extractable cell protein, mean \pm S.D. from 4-6 independent experiments.

Polypeptides expressed	Collagen P4H activity
	<i>dpm/100 μg</i>
CePHY-1 + CePDI-2	780 \pm 130
CePHY-1 + CbPDI-2	1230 \pm 140
CePHY-2 + CePDI-2	190 \pm 130
CePHY-2 + CbPDI-2	250 \pm 90
CbPHY-1 + CePDI-2	5570 \pm 1990
CbPHY-1 + CbPDI-2	20,340 \pm 2420
CbPHY-2 + CePDI-2	31,900 \pm 9150
CbPHY-2 + CbPDI-2	30,570 \pm 4280
CePHY-1 + CePHY-2 + CePDI-2	35,550 \pm 1260
CePHY-1 + CePHY-2 + CbPDI-2	21,920 \pm 1910
CbPHY-1 + CbPHY-2 + CePDI-2	92,700 \pm 18,850
CbPHY-1 + CbPHY-2 + CbPDI-2	80,720 \pm 5120

Table 4

Collagen P4H activity in Triton X-100 extracts of insect cells expressing recombinant *C. elegans* (prefix Ce) and *C. briggsae* (prefix Cb) PHY-1 and PHY-2 polypeptides with point mutations in the TPR motifs, and inter-species hybrid PHY-2 polypeptides with PDI-2.

Values are given in dpm/100 μ g of extractable cell protein, mean \pm S.D. from 3-6 independent experiments.

Polypeptides expressed	Activity
	<i>dpm/100 μg</i>
CePHY-1 + CePDI-2	1750 \pm 370
CePHY-1-T220A + CePDI-2	2100 \pm 230
CePHY-1- Q147K/T220A + CePDI-2	5640 \pm 300
CePHY-2 + CePDI-2	520 \pm 160
CePHY-2-W190S + CePDI-2	2130 \pm 60
CePHY -2- V177N/W190S + CePDI-2	4040 \pm 260
CbPHY-1 + CbPDI-2	19,830 \pm 1820
CbPHY-1-A220T + CbPDI-2	32,520 \pm 3940
CbPHY-1-K174Q/A220T + CbPDI-2	9370 \pm 900
CbPHY-2 + CbPDI-2	35,850 \pm 3410
CbPHY-2-S190W + CbPDI-2	26,070 \pm 3900
CbPHY-2-N177V/S190W + CbPDI-2	30,990 \pm 5030
CeN/CbCPHY-2+ CePDI-2 or CbPDI-2	780 \pm 130
CbN/CeCPHY-2+ CePDI-2 or CbPDI-2	38,935 \pm 8482
CePHY-1 + CePHY-2 + CePDI-2	30,840 \pm 2690
CbPHY-1 + CbPHY-2 + CbPDI-2	101,930 \pm 5620

Figure 1

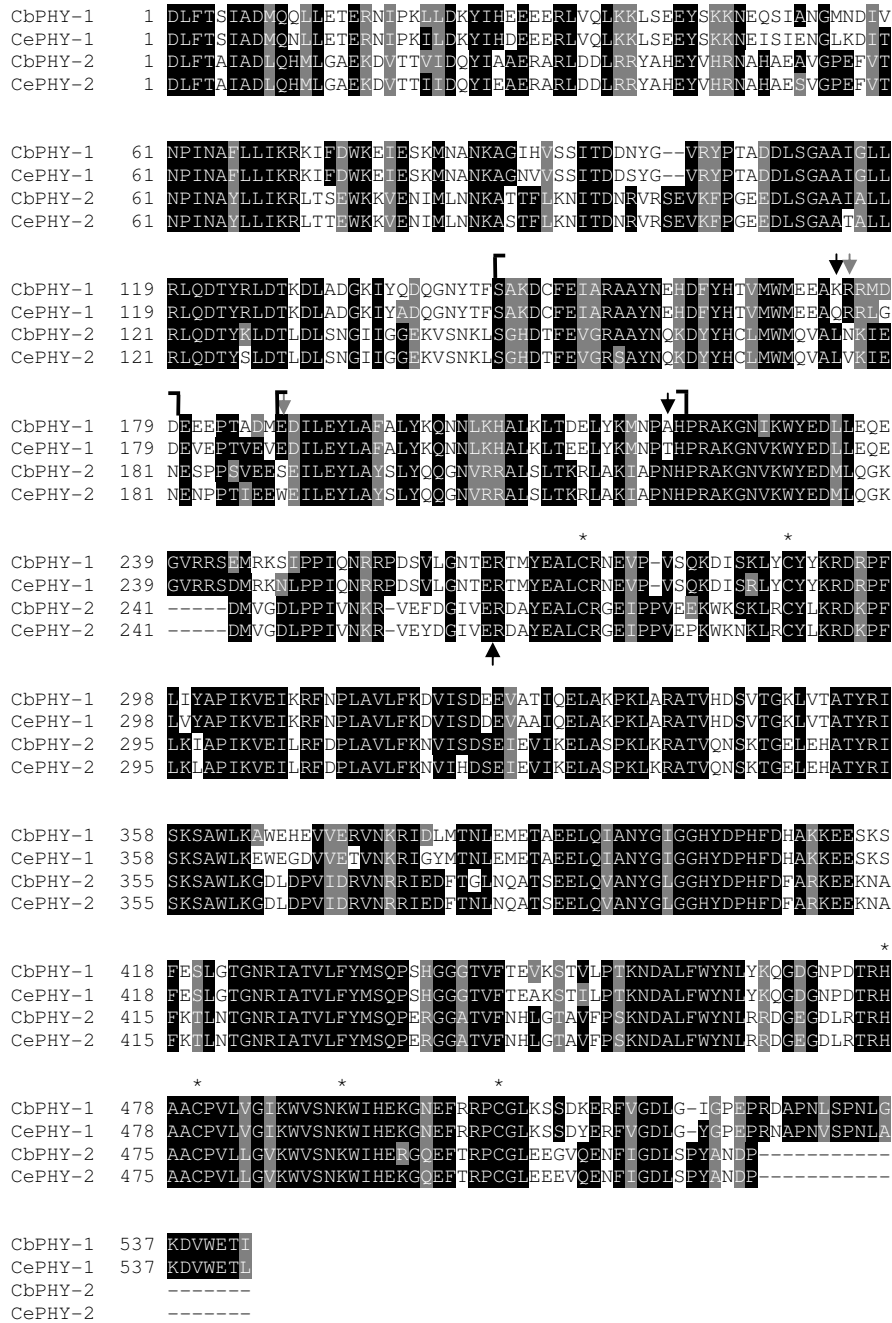


Figure 2

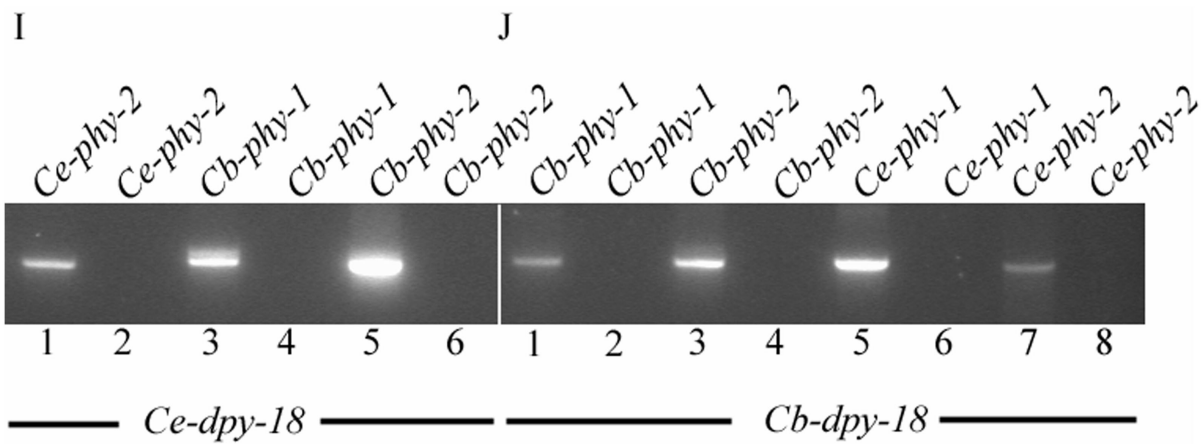
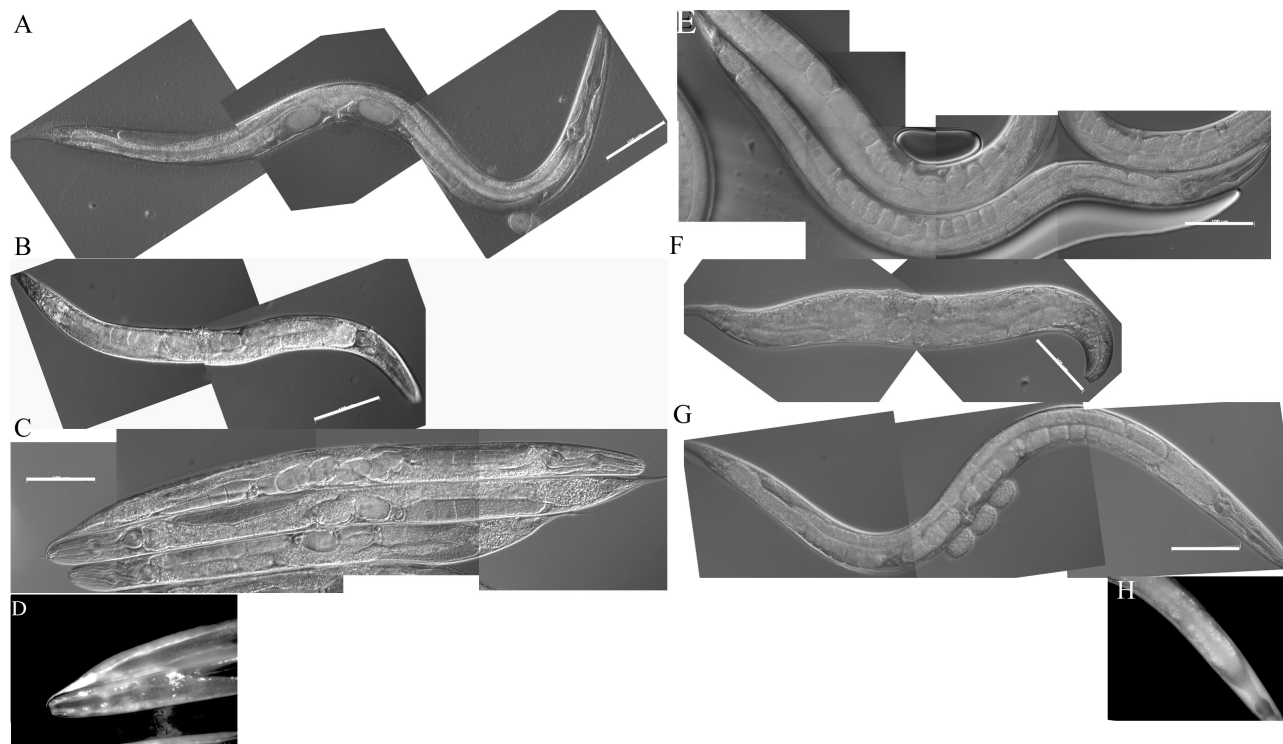


Figure 3

A



B

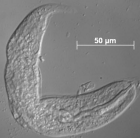


Figure 4

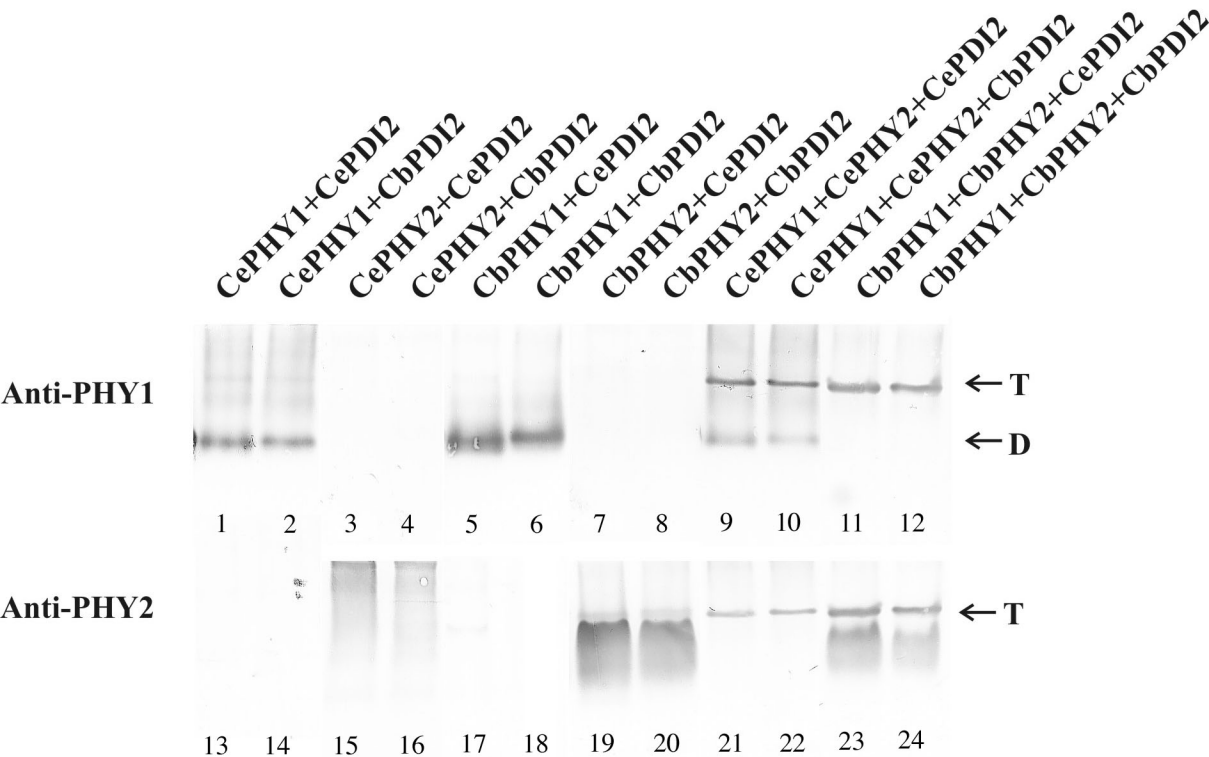


Figure 5

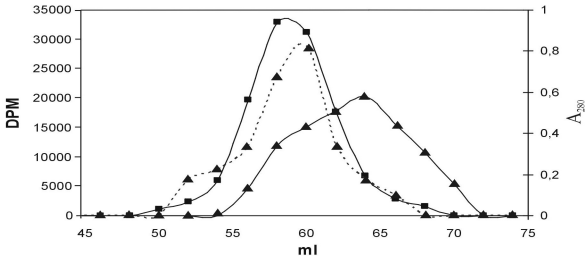


Figure 6

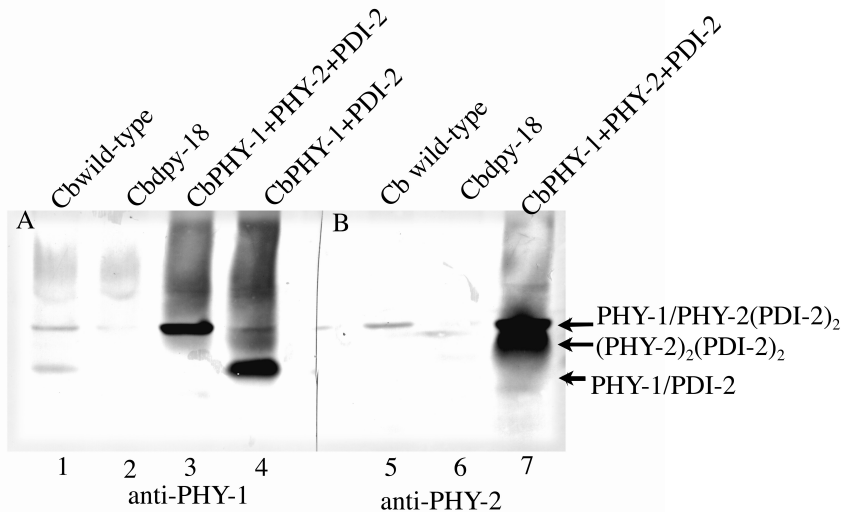


Figure 7









<u><i>C. briggsae</i> strain</u>	<u>P4H forms present</u>			<u>Phenotype</u>
Wild type (AF16)			—	Wild type
<i>Cb-dpy-18</i> (<i>mf104</i>)	—	—		Small
<i>Cb-phy-2</i> RNAi in <i>Cb</i> wild type	—		—	Wild type
<i>Cb-phy-2</i> RNAi in <i>Cb-dpy-18</i>	—	—	—	Embryonic lethal
<u><i>C. elegans</i> strain</u>				
Wild type (N2)			—	Wild type
<i>Ce-dpy-18</i> (<i>e364</i>)	—	—		Dumpy
<i>Ce-phy-2Δ</i> (<i>ok199</i>)	—		—	Wild type
<i>Ce-dpy-18</i> / <i>Ce-phy-2Δ</i>	—	—	—	Embryonic lethal

Figure 8

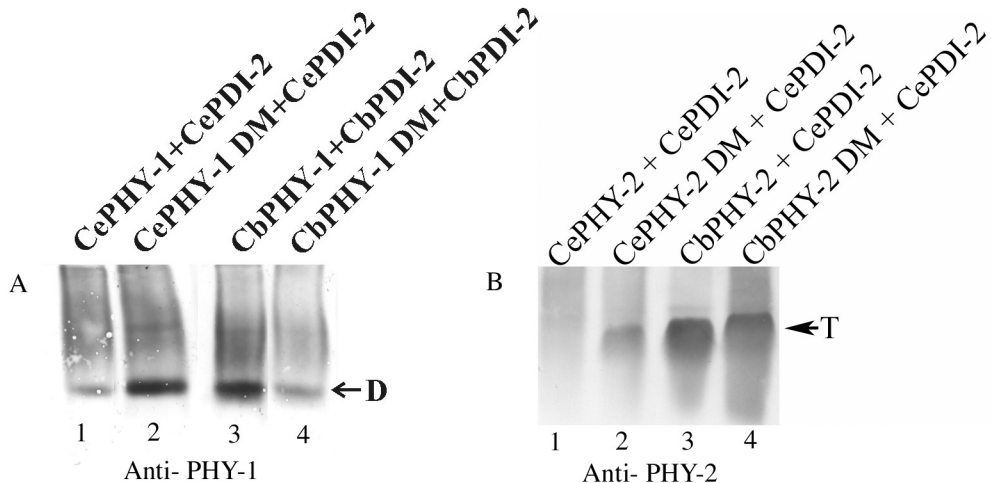
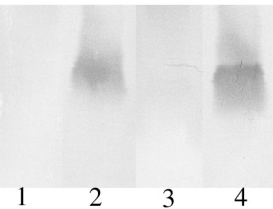


Figure 9

CePHY-2+CePDI-2
CbPHY-2+CePDI-2
CeN/CbCPHY-2+CePDI-2
CbN/CeCPHY-2+CePDI-2



} ← T