Molecular and Cellular Probes
(~5300 words)
080110bec

Characterization of a *Caenorhabditis elegans* Glc seven-like phosphatase (*gsp*)
orthologue from *Haemonchus contortus* (Nematoda)

Bronwyn E. Campbell^a, Elida M. Rabelo^{a,b}, Andreas Hofmann^c, Min Hu^a, Robin B. Gasser^{a,*}

^a*Department of Veterinary Science, The University of Melbourne, Werribee, Victoria 3030, Australia*

^b*Departamento de Parasitologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil*

^c*Eskitis Institute for Cell & Molecular Therapies, Griffith University, Brisbane, Queensland 4111, Australia*

^d*Department of Agricultural Sciences, La Trobe University, Bundoora, Victoria 3086, Australia*

*Corresponding author. Tel.: +61-3-97312000; fax: +61-3-97312366.

E-mail address: robinbg@unimelb.edu.au (R.B. Gasser).

Note: Nucleotide sequence data reported in this paper are available in the GenBank database under accession numbers GQ280009 and GQ280010.

Abstract

Full-length complementary DNA (cDNA) sequence (designated *Hc-stp-1*) encoding a serine/threonine phosphatase (*Hc*-STP-1) was isolated from *Haemonchus contortus*, a strongylid nematode parasite of small ruminants. *Hc-stp-1* was shown to be transcribed in males of the adult and fourth larval stage, but not in females, early larval stages or eggs. The full-length gene (2854 bp) contained ten exons and nine introns, and encoded a cDNA of 951 bp. Comparisons of the conceptually translated protein (317 amino acids, estimated at ~35 kDa) with serine/threonine phosphatases (STPs) from other organisms revealed the presence of the conserved motif LRGNHE. Structural analysis, by comparative modelling, confirmed strict conservation of residues and features involved in catalytic activity, and variation in the ligand-binding interface. Phylogenetic analysis of amino acid sequence data revealed that *Hc*-STP-1 clustered with STPs from other nematodes (including *Caenorhabditis elegans*, *Trichostrongylus vitrinus*, *Oesophagostomum dentatum*, *Ascaris suum* and *Brugia malayi*) to the exclusion of STPs from other organisms. The protein was inferred to be most closely related to the PP1 class of STPs of *C. elegans*, within a group containing STPs encoded, amongst others, by the genes *gsp-3* and *gsp-4* in this free-living nematode. The functions of proteins GSP-3 and GSP-4 are known to be central to spermatogenesis and other male-specific processes in *C. elegans*. The findings from the present and previous studies support the proposal that *Hc-stp-1* and its product play a significant role in reproductive and/or developmental processes in maturing or adult male *H. contortus*.

Keywords: *Haemonchus contortus*, serine/threonine phosphatase (class PP1), transcription, genetic relationships

1. Introduction

Serine/threonine phosphatases (STPs) are involved in essential biological processes, such as cell division, apoptosis and exocytosis, through the dephosphorylation of proteins [1-3]. Dephosphorylation by STPs involves the catalytic removal of phosphate groups from serine and threonine residues of proteins [4,5]. STPs can be divided into four classes of proteins (PP1, PP2A, PP2B and PP2C) based on their specificity for various substrates and inhibitors [4,6]. The two classes PP1 and PP2A require the protein to be bound to other regulatory proteins to regulate their catalytic activity (i.e. they are holo-enzymes) [4]. Previous studies [7,8] characterized *stp* genes which are transcribed in a male-specific manner in *Trichostrongylus vitrinus* (black scour worm of sheep) and *Oesophagostomum dentatum* (nodule worm of pigs) (Strongylida), suggesting that these molecules play essential roles in reproductive and/or developmental processes in these parasitic nematodes.

In a recent study, Campbell et al. [9] identified, using oligonucleotide microarray analysis, molecules (accession nos. BM138885 and CB020068) exhibiting male-specific transcription in adult *Haemonchus contortus*, the economically important, barber's pole worm of small ruminants. Following this study, an unpublished investigation showed that the proteins inferred from these two transcripts were related to PP1s encoded by a number of *Caenorhabditis elegans* genes, including *gsp-3*. This latter gene, for instance, is functionally tightly linked to reproduction, because gene silencing leads to maternal sterility (Ste) in *C. elegans* hermaphrodites [7,10]. The protein GSP-3 has been shown to be abundantly expressed in sperm, predominantly around meiotic and mature sperm DNA [11], supporting previous reports describing that the majority of protein phosphatases are highly expressed in germline tissues [12] and/or involved in spermatogenesis and/or the regulations of sperm motility [13,14]. Taken together, this information indicated that *gsp-3* orthologues and their products in

parasitic nematodes might play similar roles. In the present study, we isolated and characterized the full-length cDNA and gene (designated *Hc-stp-1*), explored transcription in different developmental stages and both sexes of *H. contortus* and studied the relationship of the inferred protein with GSP-3 and other homologues from a range of species of invertebrates and vertebrates.

2. Materials and methods

2.1. Parasite material

H. contortus was produced in Merino lambs (~3 months of age) which were raised and maintained under helminth-free conditions. Sheep were infected intra-uminally with 10,000 infective third-stage larvae (L3s) of the Haecon 5 strain of *H. contortus*. After ~ 1 month, eggs were isolated from the faeces from infected sheep using an established sucrose flotation procedure [15]. First-, second- and third-stage larvae (L1s, L2s and L3s, respectively) were produced in coproculture [16] at 28 °C. L1s, L2s and L3s (identified according to Veglia [17]) were collected 1, 4 and 8 days, respectively, and washed extensively in H₂O. L3s were exsheathed/activated in CO₂ using an established method [18], washed and then suspended in Earle's balanced salt solution (EBSS). An exsheathment rate of >95% was consistently achieved. Fourth-stage larvae (L4s) and adults of *H. contortus* were collected from the abomasa of infected lambs (euthanased by intravenous injection with an overdose of pentobarbitone sodium, Virbac; Animal Ethics approval number 0707528, The University of Melbourne) 11 days and 1 month following intra-uminal inoculation with L3s, respectively. These latter two developmental stages were washed extensively in physiological saline, and

males and females separated prior to freezing. All developmental stages of *H. contortus* collected were snap frozen in liquid nitrogen and then stored at -70 °C until use.

2.2. Isolation, synthesis, purification, treatment and storage of nucleic acids

Genomic DNA was extracted from 50 mg of pooled worms using a small-scale sodium dodecyl-sulphate (SDS)/proteinase K extraction procedure [19], followed by purification over a mini-column (Wizard Clean-Up, Promega). The specific identity and mono-specificity of the parasite material was verified by PCR-coupled, automated sequencing of the second internal transcribed spacer (ITS-2) of nuclear ribosomal DNA from genomic DNA as described previously [20].

Total RNA was extracted separately from different developmental stages (eggs, L1s, L2s, L3s and L4s) and sexes (males and females) of *H. contortus* (following homogenization under liquid nitrogen), employing the TriPure isolation reagent (Roche Molecular Biochemicals). RNA yields were estimated spectro-photometrically (NanoDrop 1000), and the integrity of RNA was confirmed by detecting discrete 18S and 28S ribosomal RNA bands on ethidium bromide-stained agarose gels. Each RNA sample (~10 µg) was treated with 2 U of *DNase* I (Promega) and incubated at 37 °C for 30 min prior to heat denaturation of this enzyme (75 °C for 5 min). Single-stranded (ss) cDNA was synthesized from *DNase* I-treated total RNA (500 ng) from each developmental stage and each sex of *H. contortus* by oligo d(T)-priming using SuperScript III reverse transcriptase (cat. no. 18064, Invitrogen), as recommended by the manufacturer. Nucleic acids were stored at -70 °C.

2.3. Isolation of the full-length cDNA/gene and reverse transcription PCR (RT-PCR)

The full-length cDNA of *Hc-stp-1* was amplified by PCR from ss cDNA derived from RNA from adult males (SuperScript III cDNA Synthesis Kit, Invitrogen; cf. section 2.2) using primers STALLF3 (5'-ATGGACACTACTCAATTGATCACTAAC-3') and STALLR3 (5'-TCATTGACAAGGAGGCGCTG-3') designed to the *Tv-stp-1* gene of *T. vitrinus* (see [8]). High fidelity PCR was conducted, using 100 pmol of each primer, 200 μ M of each dNTP and Phusion DNA polymerase (1 U, Finnzyme) under the following cycling conditions: 98 °C/30 s, followed by 30 cycles of 98 °C/10 s, 63 °C/30 s and 72 °C/90 s, followed by a final extension of 72 °C/10 min. Following agarose gel electrophoresis, the amplicon produced (representing a single band on agarose) was purified over a mini-column (Wizard PCR Preps DNA Purification System, Promega) and cloned into the pGEM-T-Easy vector (Promega). *Escherichia coli* (strain JM109) (10^8 colony forming units/ μ g, Promega) was transformed with recombinant plasmids via heat shock and grown overnight at 37 °C on Luria Bertani (LB) plates containing 10 mg/ml ampicillin, 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and 80 μ g/ml X-gal (5-bromo-4-chloro-3-indolyl- β -galactosidase). Plasmid DNA was isolated from recombinant clones and column-purified (Wizard *Plus* SV Minipreps DNA Purification Systems, Promega) from cultures and the inserts sequenced in both directions using the oligonucleotide primers T7 and SP6 (Promega), employing Big Dye Terminator v. 3.1 chemistry in an automated ABI-PRISM sequencer (Applied Biosystems). Primers STPALLF3 and HcSTPR (5'-TTATTGACAAGGTGGAGCACCAG-3'; designed to the 3' region of *Hc-stp-1* cDNA; accession no. GQ280009) were used for the amplification of the gene by long-PCR (Advantage 2 PCR Enzyme System, Clontech) from ~100 ng of genomic DNA from adult (mixed-sex) *H. contortus*. The cycling conditions in a 2400 thermocycler (ABI) were 92 °C/2 min, followed by 35 cycles of 94 °C/30 s, 60 °C/30 s and 68 °C/3 min, followed by a final extension of 70 °C/10 min. The amplicon was purified, cloned and sequenced conventionally using a primer walking strategy.

RT-PCR was performed using an approach similar to that described by Hu et al. [8]. The ss cDNA (~ 200 ng), synthesized as described in section 2.2, was subjected to PCR (50 µl volume) under the following cycling conditions: initial denaturation at 95 °C/2 min, followed by 30 cycles of 95 °C/30 s, 55 °C/30 s and 72 °C/1 min, with a final extension of 72 °C/5 min. The transcript representing *Hc-stp-1* was amplified using primers HcSTPF (5'-ATGGGTACTCGTGTTGATGTGG-3') and HcSTPR. The primers BTUB9F (5'-TTCCCTGGACARCTRAATGCT-3') and BTBU10R (comment ah: Is it BTUB or BTBU?) (5'-CGCATGCTCATYCGACGACGG-3') were used in PCR to amplify a short region (248 bp) from the beta-tubulin gene (accession no. FJ981641.1) as a reference control. Following PCR, an aliquot of each amplicon (10%) was loaded and resolved in a 2% agarose gel, stained with ethidium bromide and photographed upon transillumination. The specificity and identity of individual amplicons were confirmed by sequencing using the same primers (individually) as employed for their amplification.

2.4. Bioinformatic analyses

The exon/intron boundaries in full-length genes were inferred based on the alignment with the cDNA, using ClustalX [21] and manual adjustment, following the AG-GT rule [22]. The genomic sequence of *Hc-stp-1* was compared with those of selected *C. elegans* serine/threonine phosphatase genes (accession nos. CAB62794, NP505733, NP001022616, NP491429, NP491237 and CE35227), *T. vitrinus stp-1* (accession no. AM691045-AM691049) and of *Oesophagostomum dentatum Od-mpp-1* (accession no. AF496634) obtained from the National Center for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov>) or WormBase (<http://www.wormbase.org/>) [23] (see Table 1). Sequences were compared with those in non-redundant databases (GenBank

<http://www.ncbi.nlm.nih.gov/Genbank/>, EMBL-EBI <http://www.ebi.ac.uk/>) using the BLAST 2.0-suite of programs from NCBI (<http://www.ncbi.nlm.nih.gov/BLAST>), the Sanger Centre (www.sanger.ac.uk/Projects/Celegans/) and the Parasite Genome database (www.ebi.ac.uk/parasites/parasite_blast_server.html). The conceptual translation of cDNA into amino acid sequences was performed using the selection “translate” available at <http://bioinformatics.org/sms/>. Protein motifs were identified by interrogating the databases PROSITE (www.expasy.ch/tools/scnpsit1.html) [24] and Pfam (www.sanger.ac.uk/Software/Pfam/) [25]. Full-length amino acid sequences were chosen from the BLASTx analysis, aligned using ClustalW (<ftp://ftp.ebi.ac.uk/pub/software/clustalw2>) and the alignment adjusted manually. Orthologous sequences from a range of different parasites, other invertebrates and vertebrates were included (Table 1). Phylogenetic analysis was conducted using Bayesian Inference (BI) (software package MrBayes 3.1.2; <http://mrbayes.csit.fsu.edu/index.php>) [26,27]. Homologous sequences from the fungi *Aspergillus fumigatus* and *A. clavatus* were used as outgroups. Posterior probabilities (pp) were calculated using 200,000 generations (ngen = 200,000 burnin = 20) employing four simultaneous tree-building chains (nchains = 4) and saving every 100th tree (samplefreq = 100).

In addition to phylogenetic analysis, genetic interactions with *C. elegans* orthologues (*gsp-3*, *gsp-4* and C47A4.3) of *Hc*-STP-1 were predicted using probabilistic functional gene networking in Gene Orienteer (<http://www.geneorienteer.org/>) [28] and WormNet (<http://www.functionalnet.org/wormnet/>) [29]. Gene Orienteer includes ~20% of the genes in *C. elegans* and contains key information on cellular localization, gene ontology and expression. WormNet includes ~82% of predicted *C. elegans* genes and also includes relevant data for *Drosophila melanogaster* (vinegar fly), *Saccharomyces cerevisiae* (yeast) and *Homo sapiens* (humans). These computational networks make predictions regarding essentiality for

diverse cellular and developmental processes. In the present study, the threshold values for the algorithms in WormNet and Gene Orienteer were set at 0.89 and 4.6, respectively.

Structural elements in the proteins *Hc*-STP-1 from *H. contortus*, *Tv*-STP-1s (accession nos. CAM84505-CAM84509) from *T. vitrinus*, as well as proteins CAB62794, GSP-1, GSP-2, GSP-3 and GSP-4 from *C. elegans* were predicted using PSIPRED [30], and structure-based sequence alignments were generated using the catalytic subunit of human PP1-alpha. Independent homology models for each of the ten STP-1 proteins were computed with MODELLER [31], and those with the lowest energy were selected. The overall geometry of each structure was scrutinized using the program PROCHECK [32], and models were inspected visually with O [33].

3. Results

The *Hc-stp-1* cDNA isolated (accession no. GQ280009) was 951 bp and was differentially transcribed between the two sexes and among some developmental stages of *H. contortus*. Specifically, *Hc-stp-1* was transcribed exclusively in males of the adult and L4 stages, but not in females (adult and L4s) or any other developmental stage (Fig. 1). The cDNA of *Hc-stp-1* encoded the predicted protein (*Hc*-STP-1) of 317 amino acids (~35 kDa) (Fig. 2). *Hc*-STP-1 had significant amino acid sequence identity to proteins from a range of organisms (see Table 1). The levels of sequence identity among 62 STPs from organisms representing different taxonomic groups (amoebae, amphibians, arthropods, choanoflagellate, chordates, echinoderms, fish, fungi, mammals, nematodes, plants, platyhelminths, protozoa and yeast) ranged from 50-90%. *Hc*-STP-1 had the highest identity (90%) to *Tv*-STP-1 of *T. vitrinus* (accession nos. CAM84505-CAM84509) [8]. *C. elegans* GSP-3 encoded by *gsp-3* (accession no. W09C3.6) had the highest sequence identity (56%; e-value: 2e-103) to *Hc*-

STP-1. The alignment of *Hc*-STP-1 with STP sequences from other organisms revealed the presence of a conserved motif (LRGNHE, positions 115-120) and nine amino acids (Asp58, His60, Asp86, Asp89, Asn118, His168, His243 and Tyr267) involved in metal ion-binding and proton donation (His 119) for catalytic activity (Fig. 2) [4]. Although the amino acids involved in the catalytic activity were conserved, the N- and C-termini were more variable and are proposed to be involved in binding to proteins that target and regulate their activity [34].

The *Hc-stp-1* gene (accession no. GQ280010) was 2854 bp in length. A sequence alignment of this gene with the *Hc-stp-1* cDNA revealed ten exons and nine introns (Fig. 3). Comparisons with the structures of *stp* genes of other nematodes revealed that *Tv-stp-1* (5950-6476 bp) from *T. vitrinus* had the same number of exons and introns as *Hc-stp-1*, whereas *Od-mpp-1* (6448 bp) from *O. dentatum* had eight and seven, and orthologous *C. elegans* genes (1043-3321 bp) had 3-7 and 2-6 exons and introns, respectively. The *Hc-stp-1* gene was approximately half of the length of *Tv-stp-1* and *Od-mpp-1*, similar to *C. elegans gsp-1*, but approximately twice as long as *C. elegans* genes CA47A4.3, *gsp-2*, *gsp-3* and *gsp-4*.

Phylogenetic analysis by Bayesian inference of full-length amino acid sequence data (cf. Table 1) was conducted to establish the relationship between *Hc*-STP-1 and other STPs from a range of invertebrates and vertebrates (Table 1; Fig. 4). *Hc*-STP-1 grouped together with key homologues from four parasitic and two free-living nematodes with strong support (pp = 0.9). Specifically, *Hc*-STP-1 grouped most closely with *Tv*-STP-1 from *T. vitrinus*, which clustered with proteins from other nematodes including *C. elegans* (GSP-3, GSP-4, XP_002642146 and CAB62794), *C. briggsae* (XP_002632333, XP_002631681, *O. dentatum* (*Od*-MPP-1), *Ascaris suum* (accession no. CAJ98743) and *Brugia malayi* (XP_001902172, XP_001897890) (pp = 0.96). Grouping externally to these 15 STPs from nematodes were

other STPs from invertebrates, including the nematodes *C. elegans* (i.e. GSP-1 and GSP-2), *C. briggsae* (CAE57617) and *B. malayi* (XP_001894042) as well as from vertebrates (pp 0.8-0.9) (Fig. 4).

Genetic interactions and gene ontologies (GOs) of *C. elegans* orthologues (*gsp-3*, *gsp-4* and C47A4.3) of *Hc-stp-1* were predicted using probabilistic functional gene networking for *C. elegans* (Table 2). Sixty-five and fifty-two interactions were predicted for *gsp-3* and *gsp-4*, respectively, whilst no interactions were identified for the gene encoding protein CAB62794. The GO terms of interactors with *gsp-3* and *gsp-4* were very similar. Of the 117 interactors, 61 (52%) had GO classifications for at least one of the following terms: biological process, cellular component and/or molecular function. The main biological processes for the interactors included reproduction (such as gametogenesis and genital development in the hermaphrodite) (37.5%), embryonic and larval development (25%) as well as protein amino acid phosphorylation/dephosphorylation. Interactors with defined GO terms for cellular components included membrane (43%), nucleus (29%), cytoplasm (29%), and intracellular and extracellular molecules (14% each). Molecular functions inferred were linked mainly to structural molecule activities (41%), binding of proteins and nucleotides (27%), hydrolase activity (10%), ATP binding (8%) and protein kinase activity (5%). Less common molecular functions of the interactors included those associated with the activities of, ligases, G-protein coupled receptors (G-PCRs), nucleoside triphosphatase, protein phosphatases and metal ion binding (2-3% each). One of the interactors identified, *akt-2*, was linked to lifespan determination, the insulin receptor-signaling pathway and dauer entry (see www.wormbase.org).

Structurally, STPs share a common catalytic subunit of ~ 280 amino acid residues in length [4]. The fold of this subunit (see Fig. 5) is characterized by a central β -sandwich consisting of two mixed β -sheets, comprising five and six β -strands, respectively. The central

β -structure packs against an α -helical moiety, with seven α -helices on one side and an α/β moiety (three α -helices and three β -strands) on the other. At the top of the central β -sandwich, the interface between the three β -sheets forms an elongated cavity harbouring the active site. The conserved fold of the catalytic subunit is reflected by a high level of identity (up to 90%) among the amino acid sequences of the ten different nematode STPs. Major structural differences appear only at the level of the holo-enzymes and are generated by the diversity of the associated regulatory subunits. Given that the three-dimensional structures of some enzymes of all four classes of STPs are known (<http://www.rcsb.org/pdb>), homology models can be constructed for the STPs of nematodes for which sequence data are available. When aligning amino acid sequences of PP1, PP2A and PP2B molecules of known structures with the amino acid sequence of *Hc*-STP-1, the highest significance score (p-value of $2e-17$) was obtained for human PP1-alpha (PDB accession code 3e7a) [35]. Using this PP1-alpha structure (residues 7-299) as a template, homology models for *Hc*-STP-1, *Tv*-STP-1 from *T. vitrinus* (accession nos. CAM84505-CAM84509) and the proteins CAB62794, GSP-1, GSP-2, GSP-3 and GSP-4 of *C. elegans* were generated for the catalytic subunit (Fig. 5). For all ten nematode proteins, the features of the active site and the catalytic residues were conserved, which infers an enzymatic activity consistent with that of previously characterized serine/threonine phosphatases [34]. In the periphery of the prolate ligand-binding site, Ser129 (PP-1 numbering), which could potentially interact with substrates or drug inhibitors, is replaced by asparagine in the proteins *Hc*-STP-1 and *Tv*-STP-1. Interestingly, the *C. elegans* homologue (CAB62794) has a phenylalanine at this position, whereas this was not the case for the other *C. elegans* proteins examined (GSP-1, GSP-2, GSP-3 and GSP-4). Ser129 is conserved in GSP-1 and GSP-2, but replaced by alanine in GSP-3 and GSP-4. Adjacent to Ser129, and also lining the substrate channel, is Asp197 (PP-1 numbering), which is replaced by glycine in the proteins *Hc*-STP-1 and *Tv*-STP-1, but which relates to a lysine residue in the

C. elegans protein (CAB62794). GSP-1 and GSP-2 show conservation of this aspartate, whilst GSP-3 and GSP-4 possess a semi-conservative asparagine residue instead. Closer to the catalytic site, Asp220 (PP-1 numbering) lines the substrate channel. This position is conserved in GSP-1 and GSP-2, but shows a semi-conservative mutation to glutamine in the *C. elegans* (CAB62794) molecule. In GSP-3 and GSP-4, Asp220 is replaced by a threonine residue, and hydrophobic residues are found at this position in the molecules *Hc*-STP-1 (isoleucine) and *Tv*-STP-1 (methionine).

The *C. elegans* serine/threonine phosphatases GSP-1 and GSP-2 are phylogenetically quite distant from *Hc*-STP-1 and related molecules (Fig. 4), which is reflected by a significantly higher amino acid identity (90-95%) to the modelling template, human PP1-alpha. GSP-3 shared ~ 53% amino acid identity with this PP1-alpha. Assessing the homology models, GSP-1 and GSP-2 differed structurally only by several amino acid substitutions in the substrate/regulator binding region of the ten STPs. Several hydrophobic residues of *Hc*-STP-1 are substituted by charged residues in GSP-1 and GSP-2. For example, residues Ile206, Ala213, Ile215 in *Hc*-STP-1 are substituted by Lys205, Glu212 and Asp214 in each of these two *C. elegans* proteins. Additionally, some polar residues (Gln145, Gln247 and Gln270) in *Hc*-STP-1 are represented by charged residues (Lys144, Glu246 and Glu269) in both GSP-1 and GSP-2. In summary, the homology models obtained for *Hc*-STP-1 and a number of other STPs from nematodes reflect conserved molecular mechanisms representative of classes PP1, PP2A and PP2B, which is enabled by high amino acid sequence conservation in the core areas of these proteins.

4. Discussion

The male-enriched transcript *Hc-stp-1* is inferred to encode a PP1-type serine/threonine phosphatase with identity/similarity (comment ah: 'Identity' or 'similarity'? In the structural papers people now only tend to use 'identity') (38-90%) to a number of molecules from a range of nematodes, including *T. vitrinus*, *O. dentatum*, *B. malayi*, *A. suum* (parasitic), *C. briggsae* and *C. elegans* (free-living). In spite of relative conservation among *Hc*-STP-1 and related molecules (e.g., *Tv*-STP-1, *Od*-MPP-1 and those encoded by *gsp-3*, *gsp-4*; Fig. 1), there were significant differences among them in genomic structure as well as number and size of exons/introns (Fig. 2), suggesting significant differences in gene regulation and transcription. For instance, the *Hc-stp-1* gene (2854 bp) was substantially shorter than those of both *Tv-stp-1* (5041-5362 bp) and *Od-mpp-1* (10271 bp), being most similar in length to *gsp-1*, for the genes examined. Nonetheless, the numbers of exons (10) and introns (nine) for *Hc-stp-1* was the same as for *Tv-stp-1* (see Fig. 2).

The results of comparative modeling indicate that catalytic residues and structural features around the catalytic site are highly conserved between *Hc*-STP-1 and nine other nematode STPs. Structural differences were predicted to appear only in the ligand binding interface, which is affirmative of the concept that a specific regulatory subunit confers the unique characteristics of an individual STP holo-enzyme. *Hc*-STP-1 had 38-56% amino acid similarity/identity (comment ah: 'Identity' or 'similarity'? In the structural papers people now only tend to use 'identity')(Fig. 1) to *C. elegans* proteins GSP-3 and GSP-4, for which some biochemical and functional information is available (see WormBase). These two molecules belong to a group of at least eight PP1s in *C. elegans*, some of which have been found to be involved in spermatogenesis in this nematode [10,36]. The presence of the *Hc-stp-1* transcript in L4 and adult males (Fig. 1) of *H. contortus* also suggests a key role for this protein in male-

specific biological processes, such as spermatogenesis. Although no functional data are presently available for *H. contortus*, in *C. elegans*, proteins GSP-3 and GSP-4 are expressed abundantly in sperm, predominantly around meiotic and mature DNA [11]. In particular, these two PP1s have been specifically associated with chromatin in the pachytene-diakinesis transition, meioses I and II, anaphase I and in mature sperm cells [11]. An evaluation of functionality using gene silencing (RNAi) has revealed aberrant chromosomal segregation, resulting in male sterility and in embryos surrounded by soft egg shells [10,11].

Probabilistic functional genetic interaction data further supported an involvement in male-specific biological processes (Table 2). The *gsp-3* gene was predicted to interact with *gsp-4*, a serine/threonine kinase (*akt-2*) gene and also with various genes encoding major sperm and sperm-specific proteins, a glutathione S-transferase and a glutathione synthetase (Table 2). The *akt-2* gene and its product have been well characterized in *C. elegans* and are integral in the regulation of DNA damage-induced apoptosis in the male germline [37]. The serine/threonine kinases AKT-1 and AKT-2 (encoded by *akt-1* and *akt-2*, respectively) are major factors in the insulin-like growth factor 1 signalling (IIS) pathway, which appears to be conserved for a wide range of organisms [38,39]. The IIS pathway regulates essential processes, such as growth, development, reproduction and longevity [38,39]. Given that *akt-2* and its interactor *akt-1* are both linked to anti-apoptotic activity, GSP-3 may have a role in dephosphorylating the proteins that the two kinases phosphorylate. GSP-3 may also be involved in regulating apoptosis in the male germline. AKT-2 is known to be expressed in somatic muscles, vulva muscles and spermatheca [38]. RNA interference (RNAi) and co-immunoprecipitation experiments indicate that a regulatory subunit of a PP2A holo-enzyme (designated PPTR-1) interacts directly with *akt-1* and, in turn, modulates processes, including dauer formation and longevity in *C. elegans* [38,39]. The molecules AKT-1 and AKT-2 are also implicated in the regulation of lifespan and dauer formation in nematodes *via* the

repression of the forkhead transcription factor DAF-16 [39,40], inducing adult development and a short lifespan, whilst inhibiting dauer formation and a long lifespan. This information indicates that GSP-3 might also play an indirect, albeit significant role in these crucial biological processes. Similarly, the *gsp-4* gene is predicted to interact with major sperm protein, sperm-specific family protein (groups S, P and Q) and *gsp-3* genes, indicating that the two STPs are essential for male germline development.

Independent support for a male-specific role of some PP1 genes, such as *Hc-stp-1*, can be suggested based on previous microarray studies of *C. elegans* [12,36]. For example, Reinke et al. [12] reported that ~50% of the phosphatases and ~30% of the kinases identified or predicted to be encoded are linked to sperm production in both hermaphrodites and males of this nematode. These studies have suggested that the abundance of these molecules could explain the observation that the terminal differentiation of spermatids into motile spermatozoa occurs without gene expression, because ribosomes are discarded during spermatid development [41]. This information would suggest that all of the proteins required for this differentiation are produced before the ribosomes are lost. Thus, it is possible that the functions/activities of these proteins are regulated through dephosphorylation and phosphorylation via the involvement of phosphatases and kinases, respectively. Another possibility is that some of these enzymes relate specifically to the fertilization of oocytes, perhaps being involved in signal transduction cascades and/or in the modification of oocyte proteins. A comparison, by Hanazawa et al. [10], of the gene expression profiles of two mutant strains of *C. elegans* hermaphrodites, one (*fem-3*) producing sperm but not oocytes, and another strain (*fem-1*) producing oocytes but not sperm, showed that the expression level of two PP1 genes (designated T03F1.5 and W09C3.6) was significantly higher in sperm-producing hermaphrodites than in those producing oocytes only. In addition, a number of investigations of various, related STP genes [7,8,10,12,36] suggest that they are associated

with ‘male’ germline tissues/processes and with either sperm maturation, motility and/or the capacity of sperm to fertilize oocytes. Previous studies [42-44] have also demonstrated that the negative regulation of three genes, *fem-1*, *fem-2* and *fem-3*, required for male development in *C. elegans* (via the product of the *tra-2* gene) results in the normal development of female worms. The *fem-2* gene encodes a type 2C serine/threonine phosphatase, indicating that this STP plays a key role in sex-determination in *C. elegans*. Moreover, this phosphatase has been shown to be important in spermatogenesis and normal somatic development of the male worm [45,46]. The *tra-2* gene is predicted to encode a membrane protein, whereas *fem-1* codes for a novel protein containing ANK (amino acid code) repeats. The *fem-3* gene product is also a novel protein (cf. [42-44]) with no homology to any known protein in current databases.

Hc-stp-1 was shown to be transcribed in developing and reproductively active males of *H. contortus*, which reflects the ‘expression’ profiles for STPs in *C. elegans* and *O. dentatum* (which are also restricted to reproductively-active stages) (cf. [7]). This information suggests that *Hc-stp-1* and orthologous genes in *C. elegans* may have similar biological functions in males or sperm-producing hermaphrodites. The apparent conservation of two closely linked GATA transcription factor-binding motifs in the promoters of selected STP genes in *C. elegans* and in some other male-specific genes (such as *msps* in nematodes; [47,48], suggests that there is a common element regulating components of germline expression in male nematodes. Hence, detailed molecular and biochemical analyses in *C. elegans* could have important implications for understanding gene expression and reproductive processes in *H. contortus*, although, clearly, caution is required in the interpretation of findings, given the biological differences between the free-living and this parasitic nematode [49].

Since there is currently no effective approach for investigating the effects of gene expression *in vivo* in *H. contortus*, comparative functional genomic analyses in *C. elegans* of *gsp-3*, *gsp-4* and *Hc-stp-1* (considered an orthologue of *gsp-3*) might assist in inferring precise

function(s) for *Hc*-STP-1 (cf. [50,51]). The results obtained from previous investigations of the stronglyloid nematode *O. dentatum* [7] and *C. elegans* [7,10,12,36] have indicated that a number of STPs are functionally involved in reproductive processes, including spermatogenesis, and appear to reduce the capacity of the sperm to fertilize oocytes, although the precise mechanism by which the latter occurs is not yet understood. On the other hand, some STP genes (e.g., C47A4.3) subjected to high throughput RNAi analysis have been shown previously to possess a ‘wild-type’ phenotype [11]. The reason for this difference in findings is not yet clear, but it is possible that there is some functional redundancy in STPs or that the RNAi phenotype was not sufficiently evident to be detected microscopically in previous large-scale screens, or that the *C. elegans* strains and/or different RNAi approaches used produced phenotypes of differing penetrance. In spite of distinct differences in genomic organization compared with orthologues in *C. elegans*, *H. contortus* and other strongylids, genetic complementation studies should be attempted to restore *gsp-3* and *gsp-4* functions in a *C. elegans* strain carrying a null mutation at these loci, as has been conducted recently for *daf-16* [52].

In an experimental investigation of transgenic *C. elegans* (promoter::reporter gene constructs), Boag et al. [7] showed that the expression of *C. elegans* PP1s (encoded by genes T03F1.5 and W09C3.6) related to *Hc*-STP-1 was not restricted to the germline. Indeed, expression in neuronal or neuron-associated cells and in musculature was inferred. The promoters directed green fluorescent protein (GFP) expression of the translational fusion-constructs throughout the cell bodies, indicating that the proteins were not localized to any particular sub-cellular organelle, which indirectly suggested that they were localized in the cytoplasm. The W09C3.6 promoter directed reporter-gene expression to the tail regions of adult males *C. elegans*, inferred to be linked to either the HOA or HOB neurons. Interestingly, a large serine/threonine-rich protein, LOV-1, has been identified in *C. elegans* [53]. LOV-1 is

expressed in various male neurons, including the HOA and HOB, and associated with aspects of male mating behaviour, such as locating the vulva [53]. Although the function of this protein has not been unequivocally determined, it might be involved in the signal transduction of chemo-sensation and/or mechano-sensation during the mating process. If the W09C3.6 and homologous genes are expressed in these or similar neurons, it is possible that the serine/threonine-rich domain present in the LOV-1 protein represents a substrate for the phosphatase activity linked to this *C. elegans* gene. This information suggests that, in addition to their involvement in spermatogenesis, some male-enriched PP1s could function in neural signalling linked to mating behaviour of the male nematode. Although the neural anatomy of *H. contortus* and related nematodes is not yet known, it is possible that there is similarity to that of *C. elegans* [49,54]. The inferred localization to and/or function of key STPs in the germline, neuronal tissue, neuron-associated cells and musculature indicate prospects for the development of specific inhibitors of spermatogenesis and/or mating as anthelmintic drugs.

In conclusion, the present study of *Hc-stp-1* of *H. contortus* adds to a growing body of knowledge about gender-enriched or -specific genes and their products in parasitic nematodes [7,8,55-60]. This study has shown that some STPs are relatively conserved between parasitic and free-living nematodes, whereas others are distinctly different, perhaps reflecting their roles in distinct pathways required for the growth, development, survival and/or reproduction in nematodes. Further work should focus on the functional aspects of *Hc-stp-1* and related STPs in parasitic nematodes, such as gene knockout/complementation in *C. elegans* as the surrogate system. However, although some progress in transgenesis has been made for parasitic nematodes [61], no practical gene knockout/complementation method is yet available for any strongylid nematode. Therefore, developing a technique for functional analysis of *Hc-STP-1* and related PP1s would have major fundamental significance and could assist in the discovery of a novel anthelmintic target, amenable to a specific inhibitor of *Hc-stp-1* function, thus

disrupting essential processes. This would be of major practical value, given the current problems with anthelmintic resistance in *H. contortus* and related strongylid nematodes.

Acknowledgements

Funding from the Australian Research Council is gratefully acknowledged.

References

- [1] Klumpp S, Kriegelstein J. Serine/threonine protein phosphatases in apoptosis. *Curr Opin Pharmacol* 2002;2:458-62.
- [2] Berndt N. Roles and regulation of serine/threonine-specific protein phosphatases in the cell cycle. *Prog Cell Cycle Res* 2003;5:497-510.
- [3] Sim AT, Baldwin ML, Rostas JA, Holst J, Ludowyke RI. The role of serine/threonine protein phosphatases in exocytosis. *Biochem J* 2003;373:641-59.
- [4] Barford D, Das AK, Egloff, MP. The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Ann Rev Biophys Biomol Struct* 1998;27:133-64.
- [5] Gallego M, Virshup DM. Protein serine/threonine phosphatases: life, death, and sleeping. *Curr Opin Cell Biol* 2005;17:197-202.
- [6] Barford D. Molecular mechanisms of the protein serine/threonine phosphatases. *Trends Biochem Sci* 1996;21:407-12.
- [7] Boag PR, Ren P, Newton SE, Gasser RB. Molecular characterisation of a male-specific serine/threonine phosphatase from *Oesophagostomum dentatum* (Nematoda: Strongylida), and functional analysis of homologues in *Caenorhabditis elegans*. *Int J Parasitol* 2003;33:313-25.
- [8] Hu M, Abs EL-Osta YG, Campbell BE, Boag PR, Nisbet AJ, Beveridge I, et al. *Trichostrongylus vitrinus* (Nematoda: Strongylida): molecular characterization and transcriptional analysis of *Tv-stp-1*, a serine/threonine phosphatase gene. *Exp Parasitol* 2007;117:22-34.
- [9] [Campbell BE, Nagaraj SH, Hu M, Zhong W, Sternberg PW, Ong EK, et al.](#) Gender-enriched transcripts in *Haemonchus contortus* - predicted functions and genetic interactions based on comparative analyses with *Caenorhabditis elegans*. *Int J Parasitol* 2008;38:65-83.
- [10] Hanazawa M, Mochii M, Ueno N, Kohara Y, Iino Y. Use of cDNA subtraction and RNA interference screens in combination reveals genes required for germ-line development in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 2001;98:8686-91.
- [11] Chu DS, Liu H, Nix P, Wu TF, Ralston EJ, Yates III JR, et al. Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. *Nature* 2006;443:101-05.
- [12] Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, et al. A global profile of germline gene expression in *C. elegans*. *Mol Cell* 2000;6:605-16.
- [13] Smith GD, Wolf DP, Trautman KC, da Cruze Silva EF, Greengard P, Vijayaraghavan S. Primate sperm contain protein phosphatase 1, a biochemical mediator of motility. *Biol Reprod* 1996;54:719-27.
- [14] Varmuza S, Jurisicova A, Okano K, Hudson J, Boekelheide K, Shipp EB. Spermiogenesis is impaired in mice bearing a targeted mutation in the protein phosphatase 1c gamma gene. *Dev Biol* 1999;205:98-110.
- [15] Cox DD, Todd AC. Survey of gastrointestinal parasitism in Wisconsin dairy cattle. *J Am Vet Med Assoc* 1962;141:706-09.
- [16] Nikolaou S, Hartman D, Presidente PJA, Newton SE, Gasser RB. HcSTK, a *Caenorhabditis elegans* PAR-1 homologue from the parasitic nematode, *Haemonchus contortus*. *Int J Parasitol* 2002;32:749-58.
- [17] Veglia F. The anatomy and life-history of the *Haemonchus contortus* (Rud.), *Rep Dir Vet Res* 1915;3-4:347-500.
- [18] Conder GA, Johnson SS. Viability of infective larvae of *Haemonchus contortus*, *Ostertagia ostertagi*, and *Trichostrongylus colubriformis* following exsheathment by various techniques. *J Parasitol* 1998;82:100-2.
- [19] Gasser RB, Chilton NB, Hoste H, Beveridge I. Rapid sequencing of rDNA from single

- worms and eggs of parasitic helminths. *Nucleic Acids Res* 1993;21:2525-6.
- [20] Bott NJ, Campbell BE, Beveridge I, Chilton NB, Rees D, Hunt PW, et al. A combined microscopic-molecular method for the diagnosis of strongylid infections in sheep. *Int J Parasitol* 2009;39:1277-87.
 - [21] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23:2947-8.
 - [22] Breathnach R, Chambon P. Organization and expression of eucaryotic split genes coding for proteins. *Ann Rev Biochem* 1981;50:349-83.
 - [23] Stein L, Sternberg PW, Durbin R, Thierry-Mieg J, Spieth J. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res* 2001;29:82-6.
 - [24] Bairoch A. The PROSITE dictionary of sites and patterns in proteins, its current status. *Nucleic Acids Res* 1993;21:3097-103.
 - [25] Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res* 2000;28:263-6.
 - [26] Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 2001;17:754-5.
 - [27] Ronquist F, Huelsenbeck JP. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003;19:1572-4.
 - [28] Zhong W, Sternberg PW. Genome-wide prediction of *C. elegans* genetic interactions. *Science* 2006;311:1481-4.
 - [29] Lee I, Leher B, Crombie C, Wong W, Fraser AG, Marcotte EM. A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nature Genet* 2008;40:181-8.
 - [30] Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT. Protein structure prediction servers at University College London. *Nucleic Acids Res* 2005;33:W36-8.
 - [31] Sali A, Blundell T. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993;234:779-815.
 - [32] Laskowski R, MacArthur M, Moss D, Thornton J. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993;26:283-91.
 - [33] Jones TA, Zou JY, Cowan S, Kjeldgaard M. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr A* 1991;47:110-19.
 - [34] Egloff MP, Johnson DF, Moorhead G, Cohen PT, Cohen P, Barford D. Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatases 1. *EMBO J* 1997;16:1876-87.
 - [35] Kelker MS, Page R, Peti W. Crystal structures of protein phosphatase-1 bound to nodularin-R and tautomycin: a novel scaffold for structure based drug design of serine/threonine phosphatase inhibitors. *J Mol Biol* 2009;385:11-21.
 - [36] Jiang M, Ryu J, Kiraly M, Duke K, Reinke V, Kim SK. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 2001;98:218-23.
 - [37] Quevedo C, Kaplan DR, Derry WB. AKT-1 Regulates DNA-damage-induced germline apoptosis in *C. elegans*. *Curr Biol* 2007;17:286-92.
 - [38] Padmanabhan S, Mukhopadhyay A, Narasimhan SD, Tesz G, Czech MP, Tissenbaum HA. A PP2A regulatory subunit regulates *C. elegans* insulin/IGF-1 signaling by modulating AKT-1 phosphorylation. *Cell* 2009;136:939-51.
 - [39] Shaw R, Dillin A. PPTR-1 counteracts insulin signalling. *Cell* 2009;136:816-8.
 - [40] Vanfleteren JR, Braeckman BP. Mechanisms of life span determination in *Caenorhabditis elegans*. *Neurobiol Aging* 1999;20:487-502.

- [41] Nelson GA, Robert TM, Ward S. *Caenorhabditis elegans* spermatozoan locomotion: amoeboid movement with almost no actin. *J Cell Biol* 1982;92:121-31.
- [42] Pilgrim D, McGregor A, Jäckle P, Johnson T, Hansen D. The *C. elegans* sex-determining Gene *fem-2* encodes a putative protein phosphatase. *Mol Biol Cell* 1995;6:1159-71.
- [43] Chin-Sang ID, Spence AM. *Caenorhabditis elegans* sex-determining protein FEM-2 is a protein phosphatase that promotes male development and interacts directly with FEM-3. *Genes Dev* 1996;10:2314-25.
- [44] Mehra A, Gaudet J, Heck L, Kuwabara PE, Spence AM. Negative regulation of male development in *Caenorhabditis elegans* by a protein-protein interaction between TRA-2A and FEM-3. *Genes Dev* 2006;13:1453-63.
- [45] Kimble J, Edgar L, Hirsch D. Specification of male development in *Caenorhabditis elegans*: the *fem* genes. *Dev Biol* 1984;105:234-9.
- [46] Hansen D, Pilgrim D. Molecular evolution of a sex determination protein: FEM-2 (PP2C) in *Caenorhabditis*. *Genetics* 1998;149:1353-62.
- [47] Klass M, Ammons D, Ward S. Conservation in the 5' flanking sequences of transcribed members of the *Caenorhabditis elegans* major sperm protein gene family. *J Mol Biol* 1988;199:15-22.
- [48] Scott AL, Dinman J, Sussman DJ, Yenbutr P, Ward S. Major sperm protein genes from *Onchocerca volvulus*. *Mol Biochem Parasitol* 1989;36:119-26.
- [49] Nikolaou S, Gasser RB. Prospects for exploring molecular developmental processes in *Haemonchus contortus*. *Int J Parasitol* 2006;36:859-68.
- [50] Geldhof P, Murray L, Couthier A, Gilleard JS, McLauchlan G, Knox DP, et al. Testing the efficacy of RNA interference in *Haemonchus contortus*. *Int J Parasitol* 2006;36:801-10.
- [51] Kotze AC, Bagnall NH. RNA interference in *Haemonchus contortus*: suppression of beta-tubulin gene expression in L3, L4 and adult worms in vitro. *Mol Biochem Parasitol* 2006;145:101-10.
- [52] Hu M, Lok JB, Ranjit N, Massey Jr HC, Sternberg, PW, Gasser RB. Structural and functional characterisation of the fork head transcription factor-encoding gene, *Hc-daf-16*, from the parasitic nematode *Haemonchus contortus* (Strongylida). *Int J Parasitol* 2009;Sep 29. [Epub ahead of print].
- [53] Barr MM, Sternberg PW. A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* 1999;401:386-89.
- [54] Brownlee DJA, Fairweather I. Exploring the neurotransmitter labyrinth in nematodes. *Trends Neurosci* 1999;22:16-24.
- [55] Boag PR, Newton SE, Gasser RB. Molecular aspects of sexual development and reproduction in nematodes and schistosomes. *Adv Parasitol* 2001;50:153-98.
- [56] Boag PR, Ranganathan S, Newton SE, Gasser RB. A male-specific (cysteine-rich) protein of *Oesophagostomum dentatum* (Strongylida) with structural characteristics of a serine protease inhibitor containing two trypsin inhibitor-like domains. *Parasitology* 2002;125:445-55.
- [57] Nisbet AJ, Gasser RB. Profiling of gender-specific gene expression for *Trichostrongylus vitrinus* (Nematoda: Strongylida) by microarray analysis of expressed sequence tag libraries constructed by suppressive-subtractive hybridisation. *Int J Parasitol* 2004;34:633-43.
- [58] Nisbet AJ, Cottee P, Gasser RB. Molecular biology of reproduction and development in parasitic nematodes – progress and opportunities. *Int J Parasitol* 2004;34:125-38.

- [59] Cottee PA, Nisbet AJ, Boag PR, Larsen M, Gasser RB. Characterization of major sperm protein genes and their expression in *Oesophagostomum dentatum* (Nematoda: Strongylida). *Parasitology* 2004;129:479-90.
- [60] Cottee PA, Nisbet AJ, Abs EL-Osta YG, Webster TL, Gasser RB. Construction of gender-enriched cDNA archives for adult *Oesophagostomum dentatum* by suppressive-subtractive hybridization and a microarray analysis of expressed sequence tags. *Parasitology* 2006;132:691-708.
- [61] Grant WN, Skinner SJ, Newton-Howes J, Grant K, Shuttleworth G, Heath DD, et al. Heritable transgenesis of *Parastrongyloides trichosuri*: a nematode parasite of mammals. *Int J Parasitol* 2006;36:475-83.
- [62] Hurley TD, Yang J, Zhang L, Goodwin KD, Zou Q, Cortese M, et al. Structural basis for regulation of protein phosphatase 1 by inhibitor-2. *J Biol Chem* 2007;282:28874-83.
- [63] DeLano WL. The PyMOL Molecular Graphics System. 2002;<http://www.pymol.org>.

LEGENDS FOR FIGURES

Fig. 1. Analysis of transcription of serine threonine phosphatase (*Hc-stp-1*) in eggs (E), first-stage (L1), second-stage (L2), ensheathed third-stage (L3), exsheathed third-stage (xL3), male (L4M) and female fourth-stage (L4F) larvae as well as adult male (M) and female (F) *Haemonchus contortus* by reverse-transcription PCR. An internal control PCR using primers to a region (248 bp) of the beta-tubulin gene was included (βt). The transcription of *Hc-stp-1* (951 bp) was greater in adult and L4 males compared with other developmental stages and females.

Fig 2. Alignment of the inferred amino acid sequence for the *Haemonchus contortus* serine/threonine phosphatase (*Hc-STP-1*) with those predicted from *Trichostrongylus vitrinus* (*Tv-STP-1*, accession no. CAM84506), *Caenorhabditis elegans* genes C47A4.3, T03F1.5, W09C3.6 (*gsp-3*), F56C9.1 (*gsp-2*) and F29F11.6 (*gsp-1*) (*Ce*; accession nos. CAB62794, NP_491237, NP_491429, NP_001022616 and NP_505733 respectively), *Oesophagostomum dentatum* (*Od-MPP-1*, accession no. AF496634) and *Homo sapiens* (*Hs*; accession no. AAV38549). Amino acids predicted to be involved in the catalytic pocket (metal binding) of the enzyme are indicated by asterisks, and the histidine residue predicted to be involved in proton donation is indicated with an arrow. Bold, black and blue letters define amino acids that are conserved, semi-conserved substitutions and non-conserved residues respectively. The Prosite motif PS00125 is indicated with a horizontal line.

Fig. 3. Diagrammatic representation of the genomic organizations of the *Hc-stp-1* gene from *Haemonchus contortus* and orthologues from other nematodes. The organization of the *Hc-stp-1* gene was determined by aligning the cDNA and genomic DNA sequences, with intron-exon boundaries being defined using the AG-GT rule [22]. Black boxes represent exons, whilst lines represent introns. Numbers above the boxes indicate the size (bp) of exons, whereas numbers below the lines indicate the intron sizes (bp).

Fig 4. A tree constructed following phylogenetic analysis by Bayesian inference of amino acid sequence data for serine-threonine phosphatases from *Haemonchus contortus* (*Hc-STP-1*) and other organisms for which full-length peptide sequences are available in current databases (www.ncbi.nlm.nih.gov and www.ebi.ac.uk/parasites/parasite-genome; see Table 1). Serine-threonine phosphatase sequences from fungi of the genus *Aspergillus* were used as outgroups. Nodal support values ≥ 0.80 are indicated above each branch.

Fig. 5. Homology model of *Hc-STP-1* (residues 1-294) and predicted assembly with regulating subunits. The fold of the catalytic subunit is characterized by a central β -sandwich (pale red) packing against an α -helical moiety with seven helices (blue) on one side, and an α/β moiety (green) on the other. At the top of the central β -sandwich, the interface between the three β -sheets forms an elongated cavity harbouring the active site with two metal ions (orange). Amino acid residues 129(S $\rightarrow$ D), 197(D $\rightarrow$ G) and 220(D $\rightarrow$ I) which are mutated when compared with human PP-1 (see text) are indicated in red. Protein phosphatase inhibitor 2 from mouse PDB accession code 2o8a) [62] is superimposed in dark blue. This image was prepared using the program PyMol [63].

Table 1

Selected serine/threonine phosphatase amino acid sequences (and relevant information from BLASTx analysis) used for the analysis of phylogenetic relationship to *Hc-stp-1* from *Haemonchus contortus*.

Group	Species	Protein accession no.	Gene accession no.	Gene name/locus tag	Length (bases)	% identity
Amoeba	<i>Dictyostelium discoideum</i>	XP 643639	DOBO185058	PPPB	321	53
Amphibian	<i>Xenopus laevis</i>	AAH72730.1	BC072730	MGC79074	327	53
Arthropod	<i>Apis mellifera</i>	XP 393296.2	XM 393296	LOC409804	328	52
Choanoflagellate	<i>Monosiga brevicollis</i>	XP 0017485	XM 001748529.1	-	307	54
Chordate	<i>Oikopleura dioica</i>	AAS21337.1	OO2-15	<i>gsp-1</i>	322	53
Echinoderm	<i>Strongylocentrotus purpuratus</i>	XP 001175801	LOC752338	-	324	51
Fish	<i>Carassius auratus</i>	ABQ18261	ABQ18261	-	327	53
	<i>Danio rerio</i>	CAD61270.1	AL929207	-	327	52
	<i>Scophthalmus maximus</i>	ABC94584	DQ364569.2	-	327	53
Fungus	<i>Ashbya gossypii</i>	NP985713	4621968	AGOS AFR166C	314	51
	<i>Aspergillus clavatus</i>	XP001269665	4700735	ACLA 029690	324	53
	<i>Aspergillus fumigatus</i>	XP750244	3507427	AFUA 1G04950	323	52
	<i>Aspergillus nidulans</i>	XP658014	2876189	AN0410.2	323	53
	<i>Aspergillus terreus</i>	XP001211774	4317084	ATEG 02596	324	53
	<i>Cryptococcus neoformans</i>	XP569133	3255713	CNB02030	328	51
	<i>C. neoformans</i>	XP569131	3255713	CNB02030	314	52
	<i>Phaeosphaeria nodorum</i>	XP0017908	XM 001790816.1	-	321	51
	<i>Ustilago maydis</i>	XP759227	3631163	UM03080.1	331	52
Mammals	<i>Homo sapiens</i>	BAA82664	AB030255.2	-	294	56
	<i>Macaca mulatta</i>	XP001108364	709935	LOC709935	326	52
	<i>Mus musculus</i>	BAC40733.1	AK089067.1	-	327	53
	<i>M. musculus</i>	AAC53385.1	AAC53385	PP1cgamma	337	52
	<i>Pan troglodytes</i>	XP001144460	462795	PPP1CC	304	54
Nematode	<i>Ascaris suum</i>	CAJ98743	AM261207.1	-	308	60
	<i>Brugia malayi</i>	XP001902172	6105592	Bm1 53525	317	61
	<i>B. malayi</i>	XP001897890	6101334	Bm1 32220	304	62
	<i>B. malayi</i>	XP001894042	6097494	Bm1 12875	334	54
	<i>Caenorhabditis briggsae</i>	XP001666795	5632004	CBG00341	304	55
	<i>C. briggsae</i>	CAE57617	CAE57617	-	333	52
	<i>C. briggsae</i>	CAE73431	CAE73431	-	305	56
	<i>C. briggsae</i>	CAE71230	CAE71230	-	308	55
	<i>Caenorhabditis elegans</i>	CAB62794	C47A4.3	-	333	55
	<i>C. elegans</i>	NP505733	F29F11.6	<i>gsp-1</i>	329	54
	<i>C. elegans</i>	NP001022616	F56C9.1	<i>gsp-2</i>	333	52
	<i>C. elegans</i>	NP491429	W09C3.6	<i>gsp-3</i>	305	56
	<i>C. elegans</i>	NP491237	T03F1.5	<i>gsp-4</i>	305	56
	<i>C. elegans</i>	CE35227	C09H5.7	-	333	55
	<i>C. elegans</i>	CE02116	C06A1.3	-	364	39
	<i>C. elegans</i>	CE35562	F58G1.3	-	364	38

	<i>Oesophagostomum dentatum</i>	AAO85519	AF496635.1	<i>Od-mpp1</i>	311	54
	<i>Trichostrongylus vitrinus</i>	CAM84504-9	AM691044-9.1	<i>Tv-stp1</i>	316	90
Plant	<i>Arabidopsis thaliana</i>	NP200724	836034	TOPP2	312	53
	<i>A. thaliana</i>	CAA78153	CAA78153	-	312	53
	<i>Malus pumila</i>	AAD56010	AF178530.1	PP1	316	54
	<i>Oryza sativa</i>	EAZ25104	CM000139.1	-	344	51
	<i>O. sativa</i>	EAY88048	CM000127.1	-	325	51
	<i>Vitis vinifera</i>	CAO65054	CAO65054	-	345	53
	<i>V. vinifera</i>	CAN71489	CAN71489	-	319	53
	<i>Zea mays</i>	NP001105341	542269	prh1	316	51
Platyhelminth	<i>Schistosoma japonicum</i>	AAW24965	AY813233.1	-	314	53
Protozoa	<i>Chlamydomonas reinhardtii</i>	AAD38856	AF156101.1	-	304	55
	<i>Cryptosporidium hominis</i>	XP667490	3415346	Chro.70303	320	52
	<i>Cryptosporidium parvum</i>	XP001388388	3371830	cgd7_2670	320	52
	<i>Leishmania braziliensis</i>	XP001566103	5416977	LbrM28_V2.0710	301	54
	<i>Leishmania infantum</i>	XP001470105	5070290	LinJ28.0700	301	54
	<i>Leishmania major</i>	XP001684347	5653275	LmjF28.0690	301	54
	<i>Toxoplasma gondii</i>	ABD96038	DQ437871.1	-	306	53
	<i>Trypanosoma brucei</i>	EAN79771	CH464491.1	-	303	54
	<i>Trypanosoma cruzi</i>	XP815599	3547356	Tc00.1047053508815.110	303	54
Yeast	<i>Candida albicans</i>	XP711142.1	3647257	glc-7	330	52
	<i>Saccharomyces cerevisiae</i>	NP0110	NC_001137.2	-	312	51
	<i>Schizosaccharomyces pombe</i>	NP596	2542632	cdc25	327	53

Table 2 Genetic interactions and gene ontology terms predicted for *Caenorhabditis elegans* orthologues (*gsp-3* and *gsp-4*) of *Hc-stp-1* from *Haemonchus contortus*. Molecules interacting with *gsp-3* and *gsp-4* include those associated with reproduction, protein phosphorylation, embryonic and larval development. Cellular localities include nucleus, cytoplasm, membrane and both intracellular and extracellular ?what?. Molecular functions include serine/threonine kinase, hydrolase and structural molecule activities.

<i>gsp-3</i>				
Locus ID	Symbol	Biological process	Cellular component	Molecular function
T21G5.4	-	Larval development (<i>sensu</i> Nematoda); gametogenesis; growth; hermaphrodite genitalia development	???	Protein binding FIX THROUGHOUT
ZK938.1	-	-	-	Hydrolase activity
Y69E1A.2	-	Embryonic development (<i>sensu</i> Metazoa);	-	-
F52H3.6	-	-	-	Hydrolase activity
R13H9.1	-	-	-	Binding
T04A6.3	-	metal ion transport;	-	Metal ion binding
C04G2.8	-	-	-	-
R10E9.2	-	-	-	-
Y69E1A.1	-	-	-	-
T28H11.7	-	-	-	-
Y57G11A.2	-	-	-	-
T27A3.4	-	-	-	-
ZK265.3	-	-	-	-
F32A11.3	-	reproduction;	-	-
K06A5.2	-	-	-	-
C10G11.9	-	-	-	-
F53B6.4	-	-	-	structural molecule activity;
T16A9.5	-	-	-	-
F36H12.11	-	-	-	binding;
F32B6.5	<i>sss-1</i>	-	-	-
T27A3.3	<i>ssp-16</i>	-	-	structural molecule activity;
F44D12.4	-	-	-	protein binding;
C10G11.8	-	protein catabolic process;	nucleus; cytoplasm;	nucleotide binding; ATP binding; hydrolase activity; nucleoside-triphosphatase activity;

Y38H8A.3	-	protein amino acid phosphorylation;	-	protein ki-se activity; protein serine/threonine ki-se activity; ATP binding;
F36D3.4	-	-	-	structural molecule activity;
C14C11.1	-	-	-	-
B0379.7	-	-	-	-
C25A8.1	-	-	-	-
F36H12.9	-	protein amino acid phosphorylation;	-	protein ki-se activity; ATP binding;
C05B5.2	-	-	-	-
ZK484.8	-	-	-	-
T13F2.9	-	-	-	-
T03F1.5	<i>gsp-4</i>	-	-	hydrolase activity;
F21H7.5	-	-	-	structural molecule activity;
C45G9.9	-	-	-	-
R09E10.6	-	reproduction;	-	-
T16H12.6	<i>kel-10</i>	-	-	protein binding;
ZK546.6	<i>msp-152</i>	embryonic development (sensu Metazoa);	-	structural molecule activity;
C35D10.2	-	-	-	protein binding;
C55C2.2	<i>ssp-19</i>	-	-	structural molecule activity;
C48E7.7	-	-	-	-
ZK1248.6	<i>msp-64</i>	-	-	structural molecule activity;
W03D8.10	-	-	-	-
T23B3.5	-	-	-	-
C06A8.6	-	antibacterial humoral response (sensu Protostomia);	extracellular region;	-
C04G2.9	-	-	-	-
F27C1.1	-	-	-	-
F26G1.7	<i>msp-3</i>	-	-	structural molecule activity;
K07F5.2	<i>msp-10</i>	-	-	structural molecule activity;
F47B8.11	<i>sss-2</i>	-	-	-
C16A11.7	-	-	-	-
T10E9.4	-	-	-	-
C24D10.7	-	-	-	-
K05F1.7	<i>msp-63</i>	-	-	structural molecule activity;
T08B2.12	-	-	-	-

C25G4.6	-	reproduction;	-	protein binding;
T23F11.2	-	-	-	-
ZK1248.5	-	-	-	-
F17E9.5	-	-	-	-
W03D8.9	-	-	-	-
ZK484.5	-	-	-	-
F44D12.7	-	-	-	structural molecule activity;
B0379.2	-	-	-	-
C14C10.1	-	transport; determi-tion of adult life span, insulin receptor sig-ling pathway, dauer entry,	membrane;	binding;
F28H6.1	<i>akt-2</i>	protein amino acid phosphorylation	intracellular	protein serine/threonine ki-se activity, ATP binding, protein tyrosine ki-se activity

<i>gsp-4</i>				
Locus ID	Symbol	Biological process	Cellular component	Molecular function
F36H12.11	-	-	-	binding;
ZK546.7	-	-	-	-
T28H11.7	-	-	-	-
Y57G11A.2	-	-	-	-
Y69E1A.2	-	embryonic development (sensu Metazoa);	-	-
C01G10.14	-	-	membrane;	G-protein coupled receptor activity; structural molecule activity;
C10G11.9	-	-	-	-
ZK1248.17	-	-	-	structural molecule activity;
F26G1.7	<i>msp-3</i>	-	-	structural molecule activity;
ZK546.3	-	-	-	structural molecule activity;
R13H9.1	-	-	-	binding;
ZK637.12	-	-	-	-
ZK484.5	-	-	-	-
K08C9.2	-	reproduction;	-	-
W09C3.6	<i>gsp-3</i>	reproduction;	-	hydrolase activity;
F58A6.9	-	-	-	structural molecule activity;
Y69E1A.1	-	-	-	-

C04G2.9	-	-	-	-
T08B2.12	-	-	-	-
K05F1.9	-	-	-	structural molecule activity;
C25G4.6	-	reproduction;	-	protein binding;
T27E7.1	-	-	-	-
ZK484.8	-	-	-	-
T27A3.4	-	-	-	-
F58A6.8	<i>msp-45</i>	-	-	structural molecule activity;
C55C2.2	<i>ssp-19</i>	-	-	structural molecule activity;
C45G9.4	-	-	-	-
F44D12.4	-	-	-	protein binding;
R10E9.2	-	-	-	-
T27A3.3	<i>ssp-16</i>	-	-	structural molecule activity;
F47B8.11	<i>sss-2</i>	-	-	-
C35D10.2	-	-	-	protein binding;
K07F5.2	<i>msp-10</i>	-	-	structural molecule activity;
T21G5.4	-	larval development (sensu Nematoda); gametogenesis; growth; hermaphrodite genitalia development;	-	protein binding;
T28H11.1	<i>ssq-4</i>	-	-	-
R13H9.4	<i>msp-53</i>	-	-	structural molecule activity;
F32A11.3	-	reproduction;	-	-
T23F11.2	-	-	-	-
F17E9.5	-	-	-	-
C14C10.1	-	transport; glutamine biosynthetic process; nitrogen compound metabolic process;	membrane;	binding;
K03H1.1	<i>gln-2</i>	-	-	glutamate-ammonia ligase activity;
W03D8.10	-	-	-	-
R05F9.13	<i>msp-31</i>	embryonic development (sensu Metazoa);	-	structural molecule activity;
T27A3.5	-	protein amino acid dephosphorylation;	-	phosphoprotein phosphatase activity; protein tyrosine phosphatase activity;
C10G11.8	-	protein catabolic process;	nucleus; cytoplasm;	nucleotide binding; ATP binding; hydrolase activity; nucleoside-triphosphatase activity;
ZK1251.6	<i>msp-76</i>	-	-	structural molecule activity;
F07A5.2	-	-	-	-

F53B6.4	-	-	-	structural molecule activity;
T16A9.5	-	-	-	-
C04G2.8	-	-	-	-
K06A5.2	-	-	-	-
C09B9.6	<i>msp-55</i>	-	-	structural molecule activity;

Campbell et al.

Fig. 1. Analysis of transcription of serine threonine phosphatase (*Hc-stp-1*) in eggs (E), first-stage (L1), second-stage (L2), ensheathed third-stage (L3), exsheathed third-stage (xL3), male (L4M) and female fourth-stage (L4F) larvae and adult male (M) and female (F) *Haemonchus contortus* by reverse-transcriptase PCR. This included an internal control PCR using primers for the amplification of a region (~250 bp) from beta-tubulin (*β-tubulin*). The transcription of *Hc-stp-1* (951 bp) was detected in adult males and fourth-stage larval males, but not in any other lifestage.

Campbell et al.

Hc-STP-1	M-----DPTQLITNLLNVGLPDKGLTKTVSENDIMEVLGKAREMFLSQPPMVELD
Tv-CAM84506	M-----DTTQLITNLLSVGLPDKGLTKTVSENDIMEVLGKAREMFLSQPPAMVELD
Ce-CAB62794	M-----QNNVVDVSIIDVLSASTHEKPLCKVITEERVVKLLDLALGVFKAKQPMVENV
Ce-AAB65386	MASGYRSLDGMHRLLTMTISFFLLLENFAARLDIDELILKVLVSGHPEKSLTKTVKEHELQCLQMAKKSFLDQPMVEVD
Ce-GSP-4	M-----TATIDVDNLMRSRLNVGMSGGRLTTSVNEQELQTCCAVAKSVFASQASLLEVE
Ce-GSP-3	M-----TAPMDVDNLMRSRLNVGMSGGRLTTSVNEQELQTCCAVAKSVFASQASLLEVE
Ce-GSP-2	M-----DVEKLNLDNIISRLLSVRGSKPGKNVQLTSEIKGLCQKSREIFLSQPILLELE
Ce-GSP-1	M-----SNDGDLNLDNITRLLEVRCRPGKPVMTSEAEIRALCHKRSREIFLSQPILLELE
Od-AA085519	M-----AQLDVDSLMSRLNVGMAGGRLTTSVSEQELQQCCFVARQVFISQSSLIECE
Hs-AAV38549	M-----ADGELNVDSLITRLLEVRCRPGKIVQMTAEVRGLCLIKRSREIFLSQPIPLELE

	* *	* *
Hc-STP1	SPVKICGDTHGQYI-----DLLRLEFNKGGFPPLSNYLFLGDYVDRGKQNL	EVILLMIAYKLRFPKNFFLL
Tv-CAM84506	SPVKICGDTHGQYP-----DLLRLEFNKGGFPPLSNYLFLGDYVDRGKQNL	EVILLVIAKLRFPKNFFLL
Ce-CAB62794	APIKVCGDIHGQFP-----DLLRLEFHRGGWPPPTANYLFLGDYVDRGRFSIETIV	LLLAYKVKFPCNFFLL
Ce-AAB65386	APIKVCGDVHGQYSGKCSFTLTGKFIIIIIDVIRMFSAIRFPPHHSNYIFLGDYVDRGRQNL	ELITFLCYKIKFYDRFYML
Ce-GSP-4	PPIIVCGDIHGQYS-----DLLRIFDKNGFPPDINFLFLGDYVDRGRQNIETIC	LMFCFKIKYPENFFML
Ce-GSP-3	PPIIVCGDIHGQYS-----DLLRIFDKNGFPPDVNFLFLGDYVDRGRQNIETIC	LMFCFKIKYPENFFML
Ce-GSP-2	APLKICGDIHGQYY-----DLLRLEFYGGFPPESNYLFLGDYVDRGKQSL	ETICLLLAYKIKYPENFFLL
Ce-GSP-1	APLKICGDIHGQYN-----DLLRLEFYGGFPPEANYLFLGDYVDRGKQSL	ETICLLLAYKIKYPENFFLL
Od-AA085519	PPLVVCGDIHGQYS-----DLLRIFDKNGFPPETNYLFLGDYVDRGRQNIETIC	LMFCYRIKYPESFFML
Hs-AAV38549	APLKICGDIHGQYT-----DLLRLEFYGGFPPEANYLFLGDYVDRGKQSL	ETICLLLAYKIKYPENFFLL

	*↓	*
Hc-STP1	RGNHECANVNRAYGFYEENRRYQSQRMWQAFQDVLVCMPLTALVSDKILCMHGGLSPHLQ---	SLDQLRNITRPTDALG
Tv-CAM84506	RGNHECANVNRAYGFYDECMRRYQSQRMWQLFQDVFVCMPLTALVGEKILCMHGGLSPHLE---	SLDQLRNIPRPTATG
Ce-CAB62794	RGNHECEFNKTYGFYECCQKRYQSVRMVAAFQDVFNWLP	LTGLIATKILCMHGGLSPLMTKEFTLDRKIERPTEGK-
Ce-AAB65386	RGNHECPAVNRVYGFYEENKRYASTRLWLAFQEAFAAMPFTGLISGRILCMHGGLSPKLT---	NLDVLRDLTRPMDPPS
Ce-GSP-4	RGNHECPAINRVYGFYEENRRYKSTRLWSIFQDTFNWMP	LCGLIGSRILCMHGGLSPHLQ---
Ce-GSP-3	RGNHECPAINRVYGFYEENRRYKSTRLWSIFQDTFNWMP	LCGLIGSRILCMHGGLSPHLQ---
Ce-GSP-2	RGNHECASINRIYGFYDECKRRYN- IKLWKTTFTDCFNCLP	VAAIIDEKIFCHGGLSPDLQ---
Ce-GSP-1	RGNHECASINRIYGFYDECKRRFS- IKLWKTTFTDCFNCLP	AALIDEKIFCHGGLSPDLQ---
Od-AA085519	RGNHECPAINRIYGFYEENRRYHSSRLWSSSQDTFNWMP	LCGFIAGRILCMHGGLSPQLT---
Hs-AAV38549	RGNHECASINRIYGFYDECKRRFN- IKLWKTTFTDCFNCLP	AAIVDEKIFCHGGLSPDLQ---

	*	*
Hc-STP1	ATLEMDLLWADPVIGLNGFQANIRGASYGFGPDILAKYCQLLNIDLV	VARAHQVVQDGYEFFGGRKLVTIFSAPHYCGQFD
Tv-CAM84506	ATLEMDLLWADPVIGLNGFQANMRGASYGFGPDILAKYCQALNIDLV	VARAHQVVQDGYEFFGGRKLVTIFSAPHYCGQFD
Ce-CAB62794	EGLVADLLWADPISGLSGFMNQAGCGFGRDSVLNLCSEFQLDLV	CRAHQVVQDGYEFFAGRKLVTIFSAPHYCGQFD
Ce-GSP-4	PSIGIDLLWADPDQVVGWQANTRGVSVYVFGQDVVADVCSRLD	IDLVARAHQVVQDGYEFFASKKMVTIFSAPHYCGQFD
Ce-GSP-3	PSIGIDLLWADPDQVVGWQANTRGVSVYVFGQDVVADVCSRLD	IDLVARAHQVVQDGYEFFASKKMVTIFSAPHYCGQFD
Ce-GSP-2	QGLLCDLLWSDPKDVTGWGENDRGVSF	TFGPEVVAKFLHKHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFD
Ce-GSP-1	TGLLCDLLWSDPKDVTGWGENDRGVSF	TFGPDVVAKFLNRHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFD
Od-AA085519	PSMGIDLLWADPDQVVGWQANTRGVSVYVFGQDVVFETCQKLN	IDLVARAHQVVQDGYEFFASKKMVTIFSAPHYCGQFD
Hs-AAV38549	TGLLCDLLWSDPKDVQGWGENDRGVSF	TFGADVSKFLNRHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFD

Hc-STP1	NAAAMMTVDENLQCSFDAFRPSCAKPQPKIVATSMGSPGAPPCQ-----
Tv-CAM84506	NAAAMMTVDENLQCSFEILRPSVGKQPKIIPPTIGSPAAPPCQ-----
Ce-CAB62794	NCAAFMSCDEKLQCSFEILRPTTGRLEIREKPLLKDETN-----
Ce-AAB65386	NSAAMNVDDNLI CSFHVLRPSNRKAKQAN-----
Ce-GSP-4	NSAATMKVDENMVCTFVMYKPTPKSLRKG-----
Ce-GSP-3	NSAATMKVDENMVCTFVMYKPTPKSMRRG-----
Ce-GSP-2	NAGSMMTVDETLMCSFQILKPADKKKYPYGAGGVGSNRPVTPPRNAPAAQP--
Ce-GSP-1	NAGGMMSVDETLMCSFQILKPSKKAK-YQYQGMNSGRPAVGGGRPGTTAGKK----
Od-AA085519	NFAATMKVSEDLV CNFAMYKPTAKALRMAAGVSRAS-----
Hs-AAV38549	NAGGMMSVDETLMCSFQILKPSKKAK-YQYGGGLNSGRPVTPPRNANPPKKR-----

Fig. 2. Alignment of the inferred amino acid sequence (*Hc*-STP1) for the *Haemonchus contortus* serine/threonine phosphatase with those predicted from *Trichostrongylus vitrinus* (CAM84506), *Caenorhabditis elegans* genes C47A4.3, C09H5.7, T03F1.5 (*gsp-4*), W09C3.6 (*gsp-3*), F56C9.1 (*gsp-2*) and F29F11.6 (*gsp-1*) (Ce; accession nos. CAB62794, AAB65386, NP_491237, AAC24414, NP_001022616 and CAA98273, respectively) and from *Oesophagostomum dentatum* (accession no. AF496634) and *Homo sapiens* (Hs; accession no. AAV38549). Amino acids predicted to be involved in the catalytic pocket (metal binding) of the enzyme are indicated by asterisks, and the histidine residue predicted to be involved in proton donation is indicated with an arrow. Identical amino acids are shown in bold, similar amino acids in black and non-conserved ones are in blue. Dashes indicate gaps in the sequence, included for alignment purposes. The Prosite motif PS00125 is indicated with a horizontal line.

Campbell et al. **GENE NAMES FOR PARASITIC NEMATODES.**

H. contortus

C. elegans gsp-1

C. elegans gsp-2

C. elegans gsp-3

C. elegans gsp-4

C. elegans C47A4.3

T. vitrinus

O. dentatum

Fig. 3. Diagrammatic representations of the genomic organization of the *Hc-stp-1* gene from *Haemonchus contortus* and orthologues from other nematodes. The genomic organization of the *Hc-stp-1* gene was determined by alignment of the cDNA and genomic DNA sequences, with intron-exon boundaries defined by following the AG-GT rule (Breathnach and Chambon, 1981). Black boxes represent exons, whilst lines represent introns. Numbers above the boxes indicate the size (bp) of exons, whilst numbers below the lines indicate the size (bp) of introns.

Fig. 5. Homology model of Hc-stp-1 (residues 1-294) and potential assembly with regulating subunits.

The fold of the catalytic subunit characterised by a central β -sandwich (pale red) packing against an α -helical moiety with seven helices (blue) on one side, and an α/β moiety (green) on the other. At the top of the central β -sandwich, the interface between the three β -sheets forms an elongated cavity harbouring the active site with two metal ions (orange).

Amino acid residues 123(S->D), 192(D->G) and 215(D->I) which are mutated when compared to human PP-1 (see text) are indicated in red. Protein phosphatase inhibitor 2 from mouse (Hurley *et al.*, 2007; PDB accession code 2o8a) is superimposed in dark blue.

Figure prepared with PyMol (DeLano, 2002).