Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular & Biochemical Parasitology 142 (2005) 116–120
MOLECULAR & BIOCHEMICAL PARASITOLOGY

Brief communication

Sst-tgh-1 from *Strongyloides stercoralis* encodes a proposed ortholog of *daf-7* in *Caenorhabditis elegans*☆

Holman C. Massey*, Michelle L. Castelletto, Veena M. Bhopale, Gerhard A. Schad, James B. Lok

Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, 3800 Spruce Street, Philadelphia, PA 19104, USA

Received 14 July 2004; received in revised form 15 March 2005; accepted 15 March 2005
Available online 7 April 2005

Keywords: Parasitic nematode; *Caenorhabditis elegans*; Dauer larva; Infective larva; DAF-7; TGF-β

Many species of parasitic nematode invade their definitive hosts as infective third-stage larvae (L3i). Infection may occur by various routes, but whatever their mode of invasion, L3i exhibit a state of programmed developmental arrest reversible only when the larvae receive physicochemical cues indicating that they have entered the definitive host. Little is known about the endogenous molecular mechanisms governing developmental arrest in parasitic L3i or their reactivation on encountering the host. Dauer arrest in *Caenorhabditis elegans*, which occurs facultatively in third stage larvae (L3) under conditions of starvation, overcrowding or elevated temperature, has been proposed as a model of how parasitic nematodes regulate development in the L3i [1–4]. The environmental signals that regulate switching between dauer and continuous development in *C. elegans*, namely temperature, food and the concentration of a pheromone mediating the population density effect, are transduced by at least three molecular signal transduction pathways in which key intermediates are encoded by *da*uer *f*ormation or *daf* genes [5]. A G protein-mediated odourant receptor pathway involving the guanylyl cyclase encoded by *C. elegans daf-11* is thought to be involved in transducing the pheromone signal [6], while an insulin-like signal pathway involving the insulin-like receptor tyrosine kinase DAF-2 affects dauer switching, reproduction and life span. The third major dauer regulatory pathway contains elements of a TGF-β-like signal transduction cascade with *Cel-daf-7* encoding the TGF-β-like ligand [7], *Cel-daf-1* and *Cel-daf-4* encoding Type 1 and Type 2 receptor serine–threonine kinases, respectively [8,9], and *Cel-daf-3* [10], *Cel-daf-8* [5], and *Cel-daf-14* [11] encoding downstream receptor- and co-Smad transcription factors. Mutations resulting in loss of functional DAF-7 or its receptors give conditional dauer constitutive phenotypes. These facts suggest that in *C. elegans*, DAF-7 signaling drives development towards the continuous reproductive cycle.

To test the hypothesis that dauer-like signaling is responsible for regulation of developmental arrest in infective larvae and reactivation of development in the definitive host, searches for orthologs of key *C. elegans* dauer pathway intermediates in parasitic nematodes have been undertaken in several laboratories. From these searches in *Strongyloides stercoralis*, we have reported orthologs of intermediates in the *C. elegans* G-protein-mediated [12] and insulin-like [13] dauer signal transduction pathways, and the *S. stercoralis* ortholog of *Cel*-DAF-12 has been described by others [14]. In this communication we report the sequence and expression patterns of *Sst-tgh-1* (*S. stercoralis t*ransforming *g*rowth factor-β *h*omolog-*one*), which encodes a TGF-β-like growth factor that we propose to be orthologous to *Cel*-DAF-7. A recent paper [15] reports similar findings for *Strongyloides ratti* and for *Parastrongyloides trichosuri*. A putative *Cel*-DAF-7 ortholog, *Bm* TGH-2, has also been described in the filaria *Brugia malayi* [16]. A putative ortholog of *Cel-daf-4*, Type 2 receptor for the *Cel*-DAF-7 ligand, has also been identified in the *S. stercoralis* EST database (GenBank BE029357,

Abbreviations: *Bma*, *Brugia malayi*; *Cbr*, *Caenorhabditis briggsae*; *Cel*, *C. elegans*; *Hsa*, *Homo sapiens*; *Ptr*, *Parastrongyloides trichosuri*; *Sra*, *Strongyloides ratti*; *Sst*, *S. stercoralis*; *Tsp*, *Trichinella spiralis*

☆ *Note:* Nucleotide sequence data reported in this paper for *Sst-tgh-1* is available in the EMBL, GenBank, and DDJB databases under the accession number AY662390.

* Corresponding author. Tel.: +1 215 573 4756; fax: +1 215 573 7023.
E-mail address: hmassey@vet.upenn.edu (H.C. Massey).

doi:10.1016/j.molbiopara.2005.03.004

AW496611), further substantiating a role for TGF-β-like signaling in the physiology of this parasite.

To obtain the sequence for the *S. stercoralis* ortholog of *Cel-daf-7*, RNA and DNA templates for PCR were prepared from cultures of specific life stages of the UPD strain of *S. stercoralis* as described [12,13]. The forward primer for initial screening, D7VBF2 (ATTGAATTTGAAAAAATTGGITGG), was designed around the amino acid sequence IEFEKIGW near the N-terminal end of the ligand domain of *C. elegans* DAF-7. The first of two reverse primers, D7nSCR (TCRTATTCWGTTGGATGICAIC), corresponded to the conserved peptide CCPTEYD in *C. elegans*. The second reverse primer, D7R3′20 (AIGARCAICCRCAYTTYTTIGC), corresponds to the terminal residues AKKCGCS* of the *C. elegans* sequence (with * representing the termination codon). Taq DNA polymerase PCR, performed as described previously [12] yielded products with the expected sizes of 179 and 268 bp, respectively. These were sequenced directly from PCR. Authentic primers SsD7-R6 (TGTGCTTGAAGAATTTTTGGTG), SsD7-R7 (TATCAGGATGATAACAATCACC), SsD7-F8 (GTAAATTTAAACATCTTCAGGAC) and SsD7-F9 (AAAGTTGCTGCTATCCAACAG) were designed for expansion of the known sequence by inverse PCR using circularised *Eco* RI, *Bgl* II and *Bcl* I digests of *S. stercoralis* genomic DNA as template as described [12]. The full *Sst-tgh-1* cDNA sequence was obtained by assembling overlapping sequences of 5′ and 3′ rapid amplification of cDNA ends (RACE) products amplified from first-strand preparations from total L3i RNA using gene specific primers. Sequence editing and analysis were performed using the SeqEd® program (Applied Biosystems) and various online BLAST search engines as detailed previously [12]. Introns were located by comparing genomic and cDNA sequences. Using a primer with the *Cel*-SL1 sequence as the 5′ end tag for RACE, we only obtained PCR products under non-stringent conditions at a site in the 5′ untranslated region of the message with sequence similarity to *Cel*-SL1, but lacking a good consensus splice acceptor site. The *Cel-daf-7* message (GenBank NM064864) reportedly contains no 5′ spliced leader [7]. Our conclusion is that the *Sst-tgh-1* message is not normally trans-spliced to an SL1-like sequence. The *Sst-tgh-1* gene (Fig. 1A) comprises 1052 bp including a single 59 bp intron not found in *Cel-daf-7*. Exon 1 encodes the transit signal peptide. The entire pro- and ligand domains of the *Sst*-TGH-1 precursor protein, including the intervening proteolytic cleavage site, are encoded by exon 2. By contrast, *Cel*-DAF-7 is encoded by five exons of a gene spanning 1338 bp.

Alignment of the predicted sequences of *Sst*-TGH-1, *Cel*-DAF-7, *Bm* TGH-2 with other known nematode and representative vertebrate and insect cysteine knot type growth factors was performed using the Clustal W program in the Vector NTI® software suite (InforMax) with manual correction. These alignments reveal highest levels of sequence conservation in the ligand domains with relatively low levels of similarity in the pro-domains of these growth factors. A partial alignment of nematode sequences is shown in Fig. 1B. *Sst*-TGH-1 has a transport signal sequence at its N-terminus, in common with other members of the cysteine knot superfamily of growth factors (not shown). In *Sst*-TGH-1, the N-terminal pro-domain is separated from the ligand domain by a four-residue sequence, RIKR, corresponding to a basic protease recognition site, RRKR, in *Cel*-DAF-7, which is required for cleavage of the mature growth factor from its inactive precursor [17]. Also present are seven cysteine residues that are invariant in all members of the TGF-β superfamily. Six of these form the intra-chain disulfide bonds necessary for formation of the rigid cysteine knot structure characteristic of the group. The seventh cysteine participates in formation of an interchain disulfide bond during dimerisation of ligand monomers. *Sst*-TGH-1 contains two additional cysteine residues found in *Cel*-DAF-7 and members of the activin, inhibin and TGF-β subfamilies of growth factors that participate in an additional intrachain disulfide bond [17]. *Sst*-TGH-1, *Cel*-DAF-7, *Bm* TGH-2, and orthologs from related species form a cluster in a phylogram generated by an alignment of nematode cysteine knot growth factors that distinguishes them as a clade distinct from other closely related sequences (Fig. 1C). This is consistent with phylogenetic analysis of the closely related DAF-7-like protein from *S. ratti* [15]. *C. elegans* and *B. malayi* represent the divergent clades V and III, respectively, in the contemporary small subunit ribosomal DNA sequence based nematode phylogeny [18], and *S. stercoralis*, *S. ratti* and *P. trichosuri* represent the distinct clade IVa. Strong conservation of sequence motifs in *Sst*-TGH-1 and its orthologs in such distantly related nematodes argues for conservation of conformation, processing and function of these molecules in these organisms. This does not rule out the possibility that the developmental role of these pathways have diverged significantly as the varying life histories of these organisms evolved. It remains to be determined whether *Sst*-TGH-1 is the native ligand for the serine/threonine receptor kinase represented by two *S. stercoralis* ESTs, identified as orthologs of *Cel-daf-4*, the serine/threonine kinase Type 2 subunit of the *Cel*-DAF-7 receptor. Nevertheless, identification of *Sst-tgh-1* as an ortholog of *Cel-daf-7* and the identification of ESTs orthologous to *Cel-daf-4*, together with the identification of orthologs of genes in the insulin-like [13] and G-protein dauer signaling pathways [12] strongly supports the hypothesis that all three of the major *C. elegans* dauer signaling pathways exist in *S. stercoralis*.

If *Sst*-TGH-1 has a developmental regulatory function analogous to *Cel*-DAF-7, we would expect it to be up regulated in populations of postparasitic L1, 90% of which are destined for free living development in the UPD strain of *S. stercoralis* [19]. Conversely, we would expect it to be down regulated in post-free-living L1, all of which develop to L3i. In fact, the *Sst-tgh-1* transcript is significantly up-regulated in the L3i relative to the other life stages assayed (Fig. 2; $p < 0.025$ in all pairwise comparisons). Low levels