Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/euprot

New insights into the FLPergic complements of parasitic nematodes: Informing deorphanisation approaches

Ciaran J. McCoy[a], *Louise E. Atkinson*[a], *Mostafa Zamanian*[b], *Paul McVeigh*[a], *Tim A. Day*[b], *Michael J. Kimber*[b], *Nikki J. Marks*[a], *Aaron G. Maule*[a], *Angela Mousley*[a,*]

[a] *Molecular Biosciences-Parasitology, Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, UK*

[b] *Department of Biomedical Sciences, Iowa State University, Ames, IA, USA*

ARTICLE INFO

Article history:
Received 24 January 2014
Received in revised form
8 April 2014
Accepted 9 April 2014
Available online 19 April 2014

Keywords:
FMRFamide-like peptide
FLP
FLP-receptor
Neuropeptide GPCR
Nematode

ABSTRACT

FMRFamide-like peptide (FLP) receptors are appealing as putative anthelmintic targets. To date, 31 *flp*-encoding genes have been identified in *Caenorhabditis elegans* and thirteen FLP-activated G-protein coupled receptors (FLP-GPCRs) have been reported. The lack of knowledge on FLPs and FLP-GPCRs in parasites impedes their functional characterisation and chemotherapeutic exploitation. Using homology-based BLAST searches and phylogenetic analyses this study describes the identification of putative *flp* and *flp*-GPCR gene homologues in 17 nematode parasites providing the first pan-phylum genome-based overview of the FLPergic complement. These data will facilitate FLP-receptor deorphanisation efforts in the quest for novel control targets for nematode parasites.

1. Introduction

The discovery and development of novel anti-worm control strategies has been recognised as a priority in the human and veterinary health sectors and by the horticultural industry [1–3]. Despite the recent introduction of several new drugs to the veterinary market [4,5] and mass drug administration programmes for prioritised human helminthiases [6], roundworm infections remain widespread with significant socio-economic impacts [7]. In addition, the negative impacts of plant pathogenic nematodes on global food security are underscored by the current deficiencies in chemical control options (see [8] for review).

There has been a long-standing interest in the neuropeptidergic system as a source of novel targets for anthelmintic drugs (see [9] for review), with the FMRFamide like peptide (FLP) signalling system emerging as a leading candidate [10]. The primary drivers for this interest include: (i) the importance of FLPs to parasite behaviour (and survival) and their role in modulating neuromuscular function (a proven drug target for nematode control), (ii) the lack of drugs targeting the FLPergic

* *Corresponding author.*
E-mail address: a.mousley@qub.ac.uk (A. Mousley).
http://dx.doi.org/10.1016/j.euprot.2014.04.002

system such that resistance would not be a pressing concern, and (iii) the fact that most FLPs activate G-protein coupled receptors (GPCRs), proteins which are readily exploitable for drug discovery.

Whilst many facets of the FLP signalling system provide appeal as drug targets, FLP-GPCRs emerge as the most attractive. A major impediment to the exploitation of FLP GPCRs is the lack of data on the expression and function of FLP receptors in nematode parasites. Our current understanding of FLP receptor biology has been derived primarily from the model nematode *C. elegans*; 13 FLP-GPCRs, encoded on 10 genes, have been matched with their associated FLP ligands as determined by receptor activation potencies in heterologous expression systems [11–20].

Whilst focus on the identification of putative FLP-GPCRs in parasitic nematodes is of primary importance, much can be accomplished by re-mining for FLP ligands, especially in therapeutically-important species where they have not previously been reported. Indeed, better understanding of FLP complementarity across parasitic nematodes could expedite deorphanisation or, at least, influence deorphanisation approaches by providing more comprehensive species-specific peptide libraries to feed into screening platforms. The FLP-ligand data that we have for the phylum Nematoda are outdated (see [10,21]). The availability of ten parasitic nematode draft genomes and transcriptome data for over 60 nematode species [22] warrant re-interrogation of these datasets.

Here we report a pan-phylum homology-based BLAST interrogation of *flp* and *flp*-GPCR complements in 17 parasitic nematode species and perform phylogenetic analyses to identify additional putative *flp*-GPCRs. These data: (i) represent the most up to date, comprehensive insight into the *flp* and *flp*-GPCR complement of parasitic nematodes, (ii) support the re-designation of selected *flp*-encoding genes, (iii) expose the most highly conserved *flp*-GPCRs in key pathogenic species and, (iv) reveal putative novel *flp*-GPCRs in *C. elegans* and parasitic nematodes.

2. Methodology

2.1. Bioinformatics

A reciprocal BLAST (Basic Local Alignment Search Tool) based approach was implemented to identify *flp* gene sequelogues and *flp*-GPCR gene homologues within genomic and transcriptomic datasets of 17 key pathogenic nematodes representing five clades [23] and including plant, animal and human parasites. The parasite species selected for the BLAST analysis were primarily those with a published draft genome (seven species; see Supplementary Table 1). Secondary species-selection criteria were employed to include parasitic nematodes for which genomic and transcriptomic datasets were available and filtered based on their importance to human, animal or plant health or their status as a model parasite (ten species; see Supplementary Table 1). The BLAST analysis was completed between May 2012 and August 2013; the servers employed and databases queried are outlined in Supplementary Table 1. The draft genomes of two *Haemonchus contortus* strains [24,25] were published following the completion of the BLAST analysis in this study. To facilitate accuracy of the data presented, those genes that were not identified within the *H. contortus* databases outlined in Supplementary Table 1 were employed as search strings to query the whole-genome shotgun contigs (wgs) database found on the National Centre for Biotechnology Information (NCBI) BLAST server. Any datasets updated between May 2012 and August 2013 were similarly reinvestigated.

Prepropeptide and protein sequences for previously identified *flp*- [11], *flp*-GPCR [13–17,19,20,26,27], and selected orphan GPCR [28] -encoding genes in *Caenorhabditis elegans* (see Supplementary Tables 2 and 3) were retrieved from the NCBI protein database (www.ncbi.nlm.nih.gov/protein/) and used as search strings in translated nucleotide (tBLASTn) and protein (BLASTp) BLAST analysis of available datasets (see above). Only the largest of the splice variants encoded by any given *C. elegans flp*- or *flp*-GPCR gene were selected as query sequences. Additionally, prepropeptide sequences derived from *flp*-encoding genes not found in *C. elegans* (see [21]) were also used as BLAST search queries; these included FLP-29 (derived from *Ascaris suum* EST data; [21]), FLP-30 and FLP-31 (derived from *Meloidogyne incognita* EST data; [21]). The prepropeptide search string based methodology used in this study deviates from previously published methods (based on concatenated peptide search strings) employed to identify *flp* gene sequelogues within nematode genomes [29,30]. In this study, the prepropeptide approach has been shown to be as sensitive in identifying *flp* gene sequelogues as those methods previously published.

BLAST-generated alignment outputs (high scoring pairs) of the initial BLAST hits, with an expect value $\leq$1000 (or $\leq$100, where this was the maximum expect value threshold available), were manually inspected. In efforts to identify putative FLP-encoding genes in the selected nematode species, hits containing conserved FLP motifs [10,11] flanked by mono/dibasic cleavage sites were selected for further analysis. The motifs conserved within parasitic *flp* genes were used to designate initial hits as *C. elegans* gene sequelogues [21].

For *flp*-GPCR primary BLAST analysis, high-scoring return sequences (typically the hits with the smallest expect value and largest bit score) were concatenated into a single sequence to facilitate reciprocation (see [31]). Manually curated predicted protein (FLP-GPCR) sequences were used as search strings in reciprocal protein BLAST (BLASTp) queries against the *C. elegans* non-redundant protein sequence (nr) database on the NCBI BLAST server, using default settings. The top reciprocal BLAST hit was used to designate parasitic *flp*-GPCR genes as predicted *C. elegans* gene homologues.

2.2. Post-BLAST sequence analysis

Predicted FLP prepropeptide sequelogues and FLP-GPCR homologues were aligned using the Vector NTI Advance™11 AlignX® multiple sequence alignment tool [32], using default settings. Prepropeptide cleavage sites were identified using a previously described prediction method [21]. Predicted inter-peptide regions from each prepropeptide alignment were removed to provide an unambiguous representation of FLP conservation within sequelogue alignments (see