Molecular and Cellular Endocrinology 331 (2011) 79–88

Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Isolation and characterization of the cDNA encoding DH_{31} in the kissing bug, *Rhodnius prolixus*

Meet Zandawala*, Jean-Paul Paluzzi, Ian Orchard

Department of Biology, University of Toronto Mississauga, Mississauga, Ontario, Canada L5L 1C6

ARTICLE INFO

Article history:
Received 5 May 2010
Received in revised form 24 July 2010
Accepted 17 August 2010

Keywords:
Calcitonin-like diuretic hormone (CLDH)
Diuretic hormone 31 (DH_{31})
Diuresis
Neuropeptide
In situ
Insect

ABSTRACT

Rhodnius prolixus undergoes a period of rapid diuresis after ingesting large blood meals. Neurohormones with either diuretic or anti-diuretic activity control diuresis by acting on several tissues including the Malpighian tubules. One of the neurohormones that potentially plays a role in diuresis is diuretic hormone 31 (DH_{31}) which belongs to the insect calcitonin-like family of diuretic hormones. Here we determine the complete cDNA sequences of three *Rhopr-DH_{31}* splice variants (*Rhopr-DH_{31}-A, Rhopr-DH_{31}-B and Rhopr-DH_{31}-C*) and characterize their expression in unfed fifth-instar *R. prolixus*. Reverse transcriptase-PCR demonstrates that *Rhopr-DH_{31}* is predominantly expressed in the central nervous system (CNS) of unfed fifth-instars. However, the expression of the three splice variants differs with *Rhopr-DH_{31}-B* expression being the highest followed by *Rhopr-DH_{31}-A* and *Rhopr-DH_{31}-C*, as determined using semi-quantitative Southern blot analysis. Fluorescent *in situ* hybridization reveals that *Rhopr-DH_{31}* is expressed in a variety of cells in the CNS, including some neurosecretory cells.

1. Introduction

Rhodnius prolixus is a blood-feeding hemipteran which is confined to drier savannah areas of Central and South America (Dujardin et al., 1998; Feliciangeli et al., 2004; Monroy et al., 2003; Ramsey et al., 2000). This insect is a principal vector of Chagas' disease, an incurable illness damaging the human heart and nervous system, caused by the parasite *Trypanosoma cruzi* (Koberle, 1968). *Rhodnius prolixus* ingests large blood meals and then undergoes a period of rapid diuresis. Infection occurs when *R. prolixus* releases protozoans in urine that it deposits near the site of feeding (see Prata, 2001). Diuresis, the process of urine production, involves a variety of processes and tissues. These include ion and water movement across the epithelium of the anterior midgut and the Malpighian tubules, and muscle contractions of the anterior midgut, hindgut and dorsal vessel which facilitate mixing of the blood meal, mixing of the haemolymph, and the expulsion of waste (Coast et al., 2001, 2005; Donini et al., 2008; Te Brugge et al., 2005, 2008, 2009). Neurohormones with either diuretic or anti-diuretic activity control the function of Malpighian tubules (see Coast et al., 2002; see Schooley et al., 2005). Diuretic hormones (DHs) cause acceleration in primary urine production, whereas anti-diuretic hormones (ADHs) either stimulate fluid reabsorption in the hindgut or reduce Malpighian tubules secretion (see Coast et al., 2002). Insect DHs include serotonin (5-hydroxytryptamine) and various families of neuropeptides such as the calcitonin-like DHs (CLDH), kinin-like DHs, corticotropin-releasing factor (CRF)-like DHs and CAPA-like DHs (see Coast et al., 2002; Furuya et al., 2000; Kean et al., 2002; Maddrell et al., 1991; Te Brugge et al., 1999, 2002, 2005, 2009).

One of the neurohormones that may play a role in diuresis in *R. prolixus* is referred to as diuretic hormone 31 (DH_{31}), and this belongs to the CLDH family of insect peptides (Te Brugge et al., 2005). The first CLDH was identified in the Pacific beetle cockroach, *Diploptera punctata* (Furuya et al., 2000). Since then, CLDHs have been identified in several insects, as well as crustaceans and chelicerates (Table 1) (Christie, 2008; Christie et al., 2010; Coast et al., 2001, 2005; Gard et al., 2009; Li et al., 2008; Schooley et al., 2005). Here, we annotated CLDHs in the southern house mosquito, *Culex quinquefasciatus* and the honey bee varroa mite, *Varroa destructor* following a BLAST search of their genome databases. CLDHs are 31 amino acids in length in arthropods, with the exception of *Ixodes scapularis* and *V. destructor* (predicted) where they contain 34 amino acids. These peptides show a high degree of amino acid identity and are amidated at their carboxyl termini. Invertebrate CLDHs are less similar to vertebrate calcitonin but share the C-terminal Gly-X-Pro-NH_2 (Furuya et al., 2000). CLDHs have been shown to stimulate fluid secretion by Malpighian tubules (Coast et al., 2001; Furuya et al., 2000; Maddrell et al., 1991; Te Brugge et al., 2002, 2005; Te Brugge and Orchard, 2008), to have potent

* Corresponding author at: Department of Biology, University of Toronto Mississauga, Room 3016A/B, South Building, 3359 Mississauga Road North, Mississauga, Ontario, Canada L5L 1C6. Tel.: +1 905 828 5333; fax: +1 905 828 3792.
E-mail address: meet.zandawala@utoronto.ca (M. Zandawala).

0303-7207/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mce.2010.08.012

Table 1
Structures of mature DH_{31} peptides (deduced or sequenced) from 16 species.

Species	Peptide structure	Reference
Insects		
Rhodnius prolixus[1]	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	This study, Te Brugge *et al.*, 2008
Diploptera punctata[1]	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	Furuya et al., 2000
Apis mellifera[1]	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	Schooley et al., 2005
Tribolium castaneum[2]	GLDLGLGRGFSGSQAAKHLMGLAAANFAGGP-NH_2	Li et al., 2008
Bombyx mori	AFDLGLGRGYSGALQAKHLMGLAAANFAGGP-NH_2	Schooley et al., 2005
Drosophila melanogaster	TVDFGLARGYSGTQEAKHRMGLAAANFAGGP-NH_2	Coast et al., 2001
Anopheles gambiae[3]	TVDFGLSRGYSGAQEAKHRMAMAVANFAGGP-NH_2	Coast et al., 2005
Aedes aegypti[3]	TVDFGLSRGYSGAQEAKHRMAMAVANFAGGP-NH_2	Schooley et al., 2005
Nasonia vitripennis	GLDLGLNRGFSGSQAAKHLMGLAAANYAGGP-NH_2	Predicted
Acyrthosiphon pisum	GLDLGLSRGYSGTQAAKHLMGMAAANFAGGP-NH_2	Predicted
Nilaparvata lugens	GLDLGLSRGFSGSQAAKHLMGLAAANYAAGP-NH_2	Predicted
Culex quinquefasciatus[3]	TVDFGLSRGYSGAQEAKHRMAMAVANFAGGP-NH_2	This study, Predicted
Crustaceans		
Daphnia pulex	GVDFGLGRGYSGSQAAKHLMGLAAANYAIGP-NH_2	Gard et al., 2009
Homarus americanus[2]	GLDLGLGRGFSGSQAAKHLMGLAAANFAGGP-NH_2	Christie et al., 2010
Chelicerates		
Ixodes scapularis	AGGLLDFGLSRGASGAEAAKARLGLKLANDPYGP-NH_2	Christie et al., 2008
Varroa destructor	SNGMLDFGLARGMSGVDAAKARLGLKYANDPYGP-NH_2	This study, Predicted

Amino acids that are shared by all sequences are highlighted in gray.
[a]Identical sequences.
[b]Identical sequences.
[c]Identical sequences.

natriuretic activity (Coast et al., 2005) and to increase dorsal vessel and hindgut contractility (Te Brugge et al., 2008). Moreover, Rhopr-DH_{31} causes an increase in cAMP concentration and contraction frequency in the anterior midgut (Te Brugge et al., 2009). Since these processes are associated with diuresis, CLDHs might play an important role in post-feeding diuresis in *R. prolixus*.

In the present study, complete cDNA sequences of three *Rhopr-DH_{31}* splice variants (*Rhopr-DH_{31}-A*, *Rhopr-DH_{31}-B* and *Rhopr-DH_{31}-C*) were obtained, which encode a mature peptide that is 100% identical to that determined previously by MALDI-TOF mass spectrometry for Rhopr-DH_{31} (Te Brugge et al., 2008). Northern blot hybridization was performed to confirm the approximate size of the DH_{31} gene (*DH_{31}*) transcripts. Reverse transcriptase-PCR (RT-PCR) analysis also confirmed *Rhopr-DH_{31}* expression in the central nervous system (CNS). Semi-quantitative analysis using Southern blot hybridization to determine the relative expression of the three splice variants within the CNS revealed that *Rhopr-DH_{31}-B* expression was the highest followed by *Rhopr-DH_{31}-A* and *Rhopr-DH_{31}-C* in fifth-instar *R. prolixus*. Using *Rhopr-DH_{31}-A* partial cDNA sequence to design a probe, fluorescent *in situ* hybridization (FISH) was performed to localize the cell-specific expression of *Rhopr-DH_{31}*, revealing a number of cells distributed throughout various regions of the CNS.

2. Methods

2.1. Animals

Fifth-instar *R. prolixus* of both sexes were taken from a long standing colony at the University of Toronto Mississauga. Insects were maintained at high relative humidity in incubators at 25 °C and were fed on rabbits' blood. All the tissues used were dissected from 6 weeks post-fed (as fourth-instars) insects in nuclease-free phosphate-buffered saline (PBS) (Sigma–Aldrich, Oakville, ON, Canada) and were either used immediately or stored at −20 °C in RNAlater™ RNA stabilization reagent (Qiagen Inc., Mississauga, ON, Canada).

2.2. Screening of fifth-instar CNS cDNA library

Degenerate forward primers (dh31forward1, dh31forward2, dh31forward3 and dh31forward4) (Supplementary Table S1) were designed using the Rhopr-DH_{31} peptide sequence (Te Brugge et al., 2008). These primers were used along with plasmid reverse primers (DNR-LIB REV110 and DNR-LIB REV77) (Supplementary Table S1) to obtain a partial *Rhopr-DH_{31}* cDNA sequence. Next, modified 5′ and 3′ rapid amplification of cDNA ends (RACE) PCR reactions were performed. The partial *Rhopr-DH_{31}* cDNA sequence was used to design gene-specific forward and reverse primers for 3′ and 5′ RACE, respectively. For 5′ RACE, semi-nested PCRs were performed using two plasmid forward primers (DNR-LIB FOR1 and DNR-LIB FOR2) and four gene-specific reverse primers (DH31 5race1, DH31 5race2, DH31 5race3 and DH31 5race4) (Supplementary Table S2). One plasmid reverse primer (pDNR-LIB 3–25 REV) and four gene-specific forward primers (FOR1DH31, FOR2DH31, FOR3DH31 and FOR4DH31) were used to perform the semi-nested PCRs for 3′ RACE (Supplementary Table S3). The forward and reverse primers were used successively in order to selectively amplify the specific products. PCR product of each reaction was either column purified using EZ-10 Spin Column PCR Products Purification Kit (Bio Basic Inc., Markham, ON, Canada) or gel extracted using EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc., Markham, ON, Canada) and used as a template for the subsequent reaction. Plasmid DNA isolated from a *R. prolixus* CNS cDNA library was used as the template for the above reactions (Paluzzi et al., 2008). All PCR reactions were performed using GeneAmp® PCR System 9700 (Applied Biosystems) Thermocylcer. Temperature-cycling profiles remained constant and were based on the following profile: initial denaturation at 95 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 50–60 °C (depending on the primers used) for 30 s and 72 °C for 1 min. A final 10 min extension at 72 °C was also included. Products generated after the final PCR were gel extracted, cloned using the pGEM®-T Easy vector (Promega Corporation, Madison, WI, USA), and the plasmid DNAs isolated from the overnight cell cultures and sequenced at the SickKids DNA Sequencing Facility (The Centre for Applied Genomics, Hospital for Sick Children, Toronto, ON, Canada).

Northern blot analysis revealed that the size of transcripts was comparatively larger (see below) than that obtained following the 5′ and 3′ RACE PCR reactions. *In silico* analyses confirmed that a considerable region at the 3′ end could not be cloned using the 3′ RACE reactions and thus the following approach was adopted to obtain the complete *Rhopr-DH_{31}* cDNA sequence for all three splice variants. *Rhopr-DH_{31}* 3′ UTR sequence was predicted using the *R. prolixus* preliminary genome assembly and used to design forward and reverse gene-specific primers for PCR (Supplementary Table S4). CNS single-stranded cDNA was used as a template for these PCR reactions (for cDNA synthesis, see Section 2.5).

2.3. cDNA sequence analysis

The deduced Rhopr-DH_{31} prepropeptide sequences were examined for potential signal peptides using SignalP 3.0 (Bendtsen et al., 2004) and potential ubiquitination sites using UbPred (Radivojac et al., 2010). The intron–exon boundaries of *Rhopr-DH_{31}* were determined using BLAST (Altschul et al., 1990) search results obtained from the *R. prolixus* preliminary genome assembly (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=insects – last accessed on April 12, 2010) and confirmed using Genie, online software for splice site prediction (Reese et al., 1997). The three Rhopr-DH_{31} variant prepropeptide sequences (products of transcript variant A – HM030716, transcript variant B – HM030715 and transcript variant C – HM030714) and their homologous sequences from other insect species, including *Drosophila melanogaster transcript variant A* (NM_078790.3), *D. melanogaster transcript variant C* (NM_164825.2), *Bombyx mori* (NP_001124379.1), *Anopheles gambiae* (EAA01397.3/XP_321755.3),

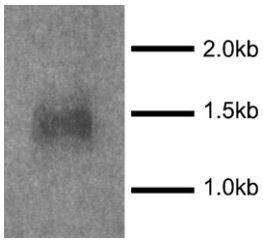

Fig. 1. Northern blot analysis of *Rhopr-DH$_{31}$* in *R. prolixus* fifth-instar CNS. Hybridization was performed with a probe synthesized using *Rhopr-DH$_{31}$-A* variant. The position of the RNA molecular weight markers (MW) is indicated.

Acyrthosiphon pisum (XP_001945901.1), *Aedes aegypti* (EAT40182.1), *Nilaparvata lugens* (DB826761) and *Nasonia vitripennis* (XP_001599948.1), were aligned using ClustalW (http://www.ch.embnet.org/software/ClustalW.html – last accessed on April 12, 2010) and the figure presented was obtained using the BOXSHADE 3.21 server (http://www.ch.embnet.org/software/BOX_form.html – last accessed on April 12, 2010). DH$_{31}$ sequence from *Homarus americanus* (GQ290461.1) was also included in the alignment and utilized as an outgroup for the phylogeny construction. *D. melanogaster transcript variant D* (NM_001169442.1) and *Tribolium castanuem* (EEZ99367.1) sequences were not included in the alignment as the prepropeptide sequence did not contain the mature peptide and the sequence was truncated at the 5′ end, respectively. The alignment obtained upon performing the multiple sequence alignment was exported and used to produce a phylogenetic tree using MEGA 4.02 (Tamura et al., 2007). The rooted phylogram was obtained using neighbour-joining analysis and presented along with bootstrap values based on 1000 replicates.

2.4. Northern blot hybridization

Total RNA was purified from CNS using Trizol® reagent (Life Technologies Corporation, Carlsbad, CA, USA) followed by mRNA extraction using PolyATtract® mRNA isolation systems III and IV (Promega Corporation, Madison, WI, USA). Approximately 300 ng CNS mRNA was denatured at 75 °C for 5 min followed by immediate incubation on ice. To enable size determination, a 250 ng sample of RiboRuler™ High Range RNA Ladder (Fermentas Canada Inc., Burlington, ON, Canada) was prepared in a similar manner. Samples of CNS mRNA and RNA ladder were then separated on a 1% agarose gel containing formaldehyde at 70 V for 140 min. In order to ensure that RNA samples were separated sufficiently, the gel was quickly observed under a UV transilluminator. Next, excess formaldehyde was rinsed off the gel by incubating the gel in diethyl pyrocarbonate-treated double-distilled water for 1 h, changing the water every 20 min. The RNA was then transferred overnight in 20× Saline-Sodium Citrate (Fisher Scientific Ltd., Ottawa, ON, Canada) to a positively charged nylon membrane (Roche, Mannheim, Germany) through downward capillary transfer and immobilized using a UV crosslinker at a setting of 30,000 μJ/cm^2. Hybridization was then performed using the DIG Northern Starter Kit (Roche, Mannheim, Germany). The protocol supplied by the manufacturer was followed with some minor modifications. Specifically, hybridization was performed using digoxigenin (DIG)-labeled RNA probe. Plasmid DNA containing a 445 bp partial *Rhopr-DH$_{31}$-A* cDNA fragment was linearized using *NcoI* and then purified using Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI, USA). The linearized recombinant plasmid was used as template to synthesize the DIG-labeled RNA probe via *in vitro* transcription using the DIG RNA labeling kit SP6/T7 (Roche Applied Science, Mannheim, Germany). SP6 RNA polymerase was used to synthesize this anti-sense probe. Following RNA probe synthesis, the template DNA was removed by incubating the probe with deoxyribonuclease I for 15 min at 37 °C. During prehybridization, the membrane was incubated in pre-warmed hybridization solution at 68 °C for 30 min. The prehybridization solution was then replaced with hybridization solution containing probe (50 ng/mL) and the membrane incubated at 68 °C for 16–18 h. The blot was then washed under stringency conditions and used for immunological detection. Exposures were made at room temperature for various times ranging from 1 min to 20 min using Bioflex Scientific Imaging Films (Clonex Corporation, Markham, ON, Canada). The blot was stripped off the probe and reprobed using DIG-labeled probe for the RNA ladder to determine the size of the observed fragment.

2.5. RT-PCR tissue profiling

Insect tissues were dissected and pooled into four different groups: (1) CNS, (2) salivary glands, (3) oesophagus, anterior midgut, posterior midgut, hindgut and Malpighian tubules, (4) dorsal vessel, fat bodies, trachea, abdominal nerves, immature testes, immature ovaries, dorsal diaphragm and ventral diaphragm. Following the manufacturer supplied protocols, total RNA was isolated from these tissues using SV Total RNA isolation system (Promega Corporation, Madison, WI, USA) and used for cDNA synthesis using an oligo(dT) primer supplied with iScript™ Select cDNA Synthesis Kit (Bio-Rad Laboratories Ltd., Mississauga, ON, Canada). An aliquot of this single-stranded cDNA was used as a template to perform a subsequent PCR. The following temperature-cycling profile was used for PCR to amplify a *Rhopr-DH$_{31}$* fragment: initial denaturation at 95 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 1 min and a final 10 min extension at 72 °C. The forward primer (FOR1DH31) had the 5′ CACTATATCGCCGGCAGTC 3′ sequence and the reverse primer (DH315race4) had the 5′ CAAGATCCAAACCACGCTTC 3′ sequence. The predicted size of the PCR product would be between 317 bp and 428 bp depending on the type of variant expressed. For a positive control to test the quality of the template cDNA, a 229 bp *ribosomal protein 49* (*RP49*) fragment was amplified. The forward primer (rp49-qPCR-for) had the 5′ GTGAAACTCAGGAGAAATTGGC 3′ sequence and the reverse primer (rp49-rev2) had the 5′ AGGACACACCATGCGCTATC 3′ sequence. This procedure was repeated using three independent biological replicates.

2.6. Southern blot hybridization

RNA was isolated from CNS tissue and subsequently used to synthesize cDNA as mentioned earlier (see Section 2.5). An aliquot of this single-stranded cDNA was used as a template for PCR, identical to the one described earlier for *Rhopr-DH$_{31}$* RT-PCR (see Section 2.5). It was assumed that the target amplification was equally efficient for the three transcript variants since the primers used were designed over regions common to all three variants (i.e. the transcript variant region was located within the amplified region and not over the primer annealing sites). Several different PCR reactions were carried out with the number of cycles ranging from 26 to 35 since no amplification products were observed in PCR reactions of less than 26 cycles. This was carried out to ensure that the relative abundance of each transcript was consistent over the exponential amplification phase of the PCR reaction and not an end-point inhibition artifact. The PCR product was purified as described earlier (see Section 2.2) and 100 ng of this purified PCR product was separated on a 2% agarose gel for 2 h at 170 V. The DNA was denatured, neutralized and transferred to a positively charged nylon membrane (Roche, Mannheim, Germany) through downward capillary transfer. The DNA was fixed onto the membrane using a UV crosslinker at a setting of 7500 μJ/cm^2. Hybridization was performed using the Amersham Gene Images AlkPhos Direct Labelling and Detection System (GE Healthcare, Piscataway, NJ, USA) as described previously (Paluzzi et al., 2008). The manufacturer protocol was followed with some minor modifications. In particular, hybridization was performed using an alkaline phosphatase-labeled probe, which was a 317 bp long fragment amplified from *Rhopr-DH$_{31}$-A* clone using the same primers as above for PCR. The probe was stored at −20 °C in 50% glycerol until used for hybridization. The membrane was incubated in hybridization solution containing probe (10 ng/mL) at 60 °C. After overnight incubation for 22–24 h, the membrane was washed in the primary wash buffer, in which the blocking solution was replaced by 0.1% (w/v) BSA and 0.1% (w/v) milk powder. The blot was subsequently treated following manufacturer recommendations. Lastly, the image was captured using a STORM 840 Phosphorimager (GE Healthcare, Piscataway, NJ, USA) at several time points after adding the ECF substrate to generate signal, and was analyzed using ImageQuant TL software (GE Healthcare, Piscataway, NJ, USA). The Southern blot analysis of the RT-PCR products was repeated using three independent biological replicates to obtain average relative expression of the three transcripts. To confirm that the average relative expression of the three transcripts differed significantly, two-tailed *t*-test ($P < 0.05$) was performed on arcsine transformed data.

2.7. Fluorescent in situ hybridization (FISH)

Rhopr-DH$_{31}$ expression localization in fifth-instar *R. prolixus* CNS was determined using the protocols described earlier (Paluzzi et al., 2008; Paluzzi and Orchard, 2010), with the following modifications. DIG-labeled anti-sense RNA was synthesized as mentioned earlier for northern blot hybridization (see Section 2.4). Similarly, to obtain the sense probe for the control experiment, the plasmid DNA was linearized using *NdeI* and T7 polymerase was used instead to label the RNA probe via *in vitro* transcription. The probe was treated with deoxyribonuclease I for 15 min at 37 °C to remove the template DNA and then stored at −20 °C for up to several weeks. Following the endogenous peroxidase quenching step, the tissues were incubated in 4% Triton X-100 (Sigma–Aldrich, Oakville, ON, Canada) in PBS with 0.1% Tween-20 (PBT) and the tissues incubated for 45 min at room temperature. Prehybridization was performed at 56 °C for 1 h. Approximately 0.5–1 ng probe (amount estimated using gel electrophoresis) per 1 μL hybridization solution was used to perform hybridization. Following the primary antibody incubation, tissues were washed six times for 10 min each with PBT containing 1% (w/v) blocking reagent (PBTB) supplied with TSA™ kit #24 (Molecular Probes Inc., Eugene, OR, USA). Following the incubation with streptavidin horseradish peroxidase, the tissues were washed six times for 10 min each with PBTB. This was followed by one wash in PBT and two washes in PBS for 5 min each. Following the incubation with tyramide substrate, the tissues were washed five times for 10 min each with PBS and then left overnight (12–14 h) in PBS at 4 °C before mounting.

(A)

```
                                                                5′ – CACTATATCGCC 12
GGCAGTCGCCCAAAGCAGCTCAGCCTCCGAGCCCTACTAACTGCATCCAGAGAGAGCGCATTCATATACATCCTACAAA 91

ATG GTA ACA AAT ATT GCT GTA GTA GGA GTG AGT CTG ATG CTG GGT ACA CTG ATA GTT CTG 151
MET Val Thr Asn Ile Ala Val Val Gly Val Ser Leu Met Leu Gly Thr Leu Ile Val Leu 20

TCT GCC GCC TCT GAA AAC ATC CCT TAC ATT GGA CAT AGG GCA TCC TAT TTT GGA GAC ATG 211
Ser Ala Ala Ser Glu Asn Ile Pro Tyr Ile Gly His Arg Ala Ser Tyr Phe Gly Asp Met 40
           ▲
GAC AAT GAG CCT GAC AGC GAA GTT ATG TTG GAA ATC CTT GCC AAA TTA GGT CAG ACT ATC 271
Asp Asn Glu Pro Asp Ser Glu Val Met Leu Glu Ile Leu Ala Lys Leu Gly Gln Thr Ile 60

ATG AGA GCA AAC GAC CTT GAA AAG CCA ATG ATA TAC AGC AGG GAA GCC TCG AAC CCC TGG 331
Met Arg Ala Asn Asp Leu Glu Lys Pro Met Ile Tyr Ser Arg Glu Ala Ser Asn Pro Trp 80

ACT GCT GTC AAC AAG TTG AGA CCA TCA AAT TTG CCC TAC AAC ATT GAA CTT GCT GAA AAC 391
Thr Ala Val Asn Lys Leu Arg Pro Ser Asn Leu Pro Tyr Asn Ile Glu Leu Ala Glu Asn 100

CCA GAT TCC ATT TAC TCG AAG CGT GGT TTG GAT CTT GGA TTA AGC AGA GGT TTC TCT GGA 451
Pro Asp Ser Ile Tyr Ser Lys Arg Gly Leu Asp Leu Gly Leu Ser Arg Gly Phe Ser Gly 120

TCA CAG GCT GCC AAA CAT CTG ATG GGA CTA GCT GCA GCT AAC TAT GCA GGA GGC CCA GGC 511
Ser Gln Ala Ala Lys His Leu Met Gly Leu Ala Ala Ala Asn Tyr Ala Gly Gly Pro Gly 140

AGA AGG AGG CGC CAA GCG TAG ACGCCTAAAATTCCTTTCTCTTCGAAATCAAAACTAATAATTAAATATATC 583
Arg Arg Arg Arg Gln Ala ---                                                     146

CCCCGAATAACGAAAATATATACATATATATATTTTTTTTTGTCAACAATAAACCCTCTCCTAAATATTATAAATAACA 662
ATTCAATGATTCTCAATCAAAACAAATAAAATTTAAGAATCACTTGTAGCTAACAATAACTTCTTCTATCTGTCGAGAT 741
GATCAGCTATTGTATTAGTGCTCACTTAGATGATCTGTCTATTTCTTACCATAAATGTACCTAAGTGAACTAAACATCC 820
TTTCTAGAATAAATGATAAATAAAAATTATTGCATTTATTTATTTAAATAATATAAATATATATATATATTTAAATA 899
AATCAAACTAATTTATTTAGGAGAAATAAGACATAAAGTATGTAAATGTTTTATTATTCATTTTGCATATACATAACTC 978
CCTTAGCTTGCTTTTACTTTCTGTTCATGTTCTAAGAAAAAGAATTATTGTTTAAAGATATACAACAAAACAGAACAAA 1057
CAGTTTCTACGAATAACAATAGCGTCTCATTTATTGTTTTTACATGATTTTAATTAAGTTATTGTATGCAGTGCAATTC 1136
CCTCAGGACAAAGAAAATTCTTATTATGTACGCTTTTCAAATACAGGTTTTGCTTTTACTCTTGATTTATGTTTGTTTG 1215
TTAAACCACTTTTTATTACTGAATTAAAAATTGAATTATTATTTACCCAGTTAAAGTTAGAATGTCCAAAAGTCAAATC 1294
TATATTTTTATTTAAAATGTATTGTATTTATTTCGATTGTATCATATATAAATAAACCTTTTTAGAATAAATAAAGGTG 1373
GACAAAAAAAAGCATGTTATCTTTGGGTCTATCTTTTGTTGAAAAAAAACCCTTTCTATGGAATTTATACCTAATAAAAT 1452
TTAGATGCAGCAAATTTAAAAAAAAAAAAAAAA – 3′                                          1485
```

(B)

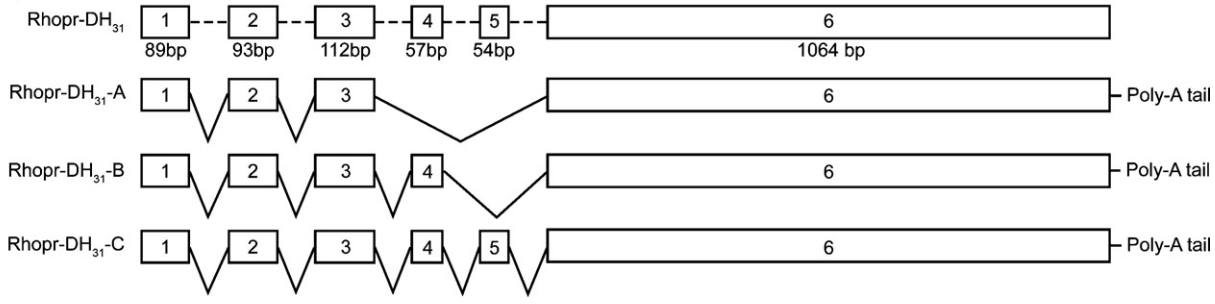

Fig. 2. *Rhopr-DH31* sequence and structure of *Rhopr-DH31* splice variants. (A) Nucleotide sequence of *Rhopr-DH31-C* cDNA and deduced amino acid sequence of the peptide precursor. The numbering for each sequence is shown at right. The initial methionine start codon has been capitalized. The amino acid sequence for the mature DH_{31} peptide is shown in bold. The dibasic amino acid pair at the N terminus and the arginine residues at the C terminus that are required for post-translational cleavage are shaded in gray. The glycine residue required for amidation is boxed while the lysine reside predicted to undergo ubiquitination is double underlined. The nucleotides highlighted in black indicate exon–exon boundaries. Potential polyadenylylation signals are underlined and the partial poly(A) tail is dash underlined. The most likely site for cleavage of a signal sequence is indicated by an arrow. (B) Molecular organization of *Rhopr-DH31* based on the BLAST analysis and *in silico* analysis using intron prediction. The boxes represent exons and the dashes represent introns.

3. Results

3.1. Three DH_{31} splice variants exist in R. prolixus

Three *Rhopr-DH31* splice variants are present in fifth-instar *R. prolixus*, referred to as *Rhopr-DH31-A*, *Rhopr-DH31-B* and *Rhopr-DH31-C*. The *Rhopr-DH31-A* cDNA sequence obtained after performing the initial 5′ and 3′ RACE was used to synthesize a labeled probe. Northern blot hybridization was then performed using this probe to determine the size of the *Rhopr-DH31* transcripts. The expected size of the transcripts was approximately 450–550 bp. However, the *Rhopr-DH31* transcripts were estimated to be just less than 1.5 kb long (Fig. 1), which was considerably larger than the cDNA sequences that had been determined. Hence, only partial cDNA sequences for the three variants had been obtained after the 5′ and 3′ PCR RACE reactions. Therefore a modified approach, as described earlier, was adopted to obtain the complete *Rhopr-DH31* cDNA sequence (Fig. 2A). *Rhopr-DH31-A* variant was found to

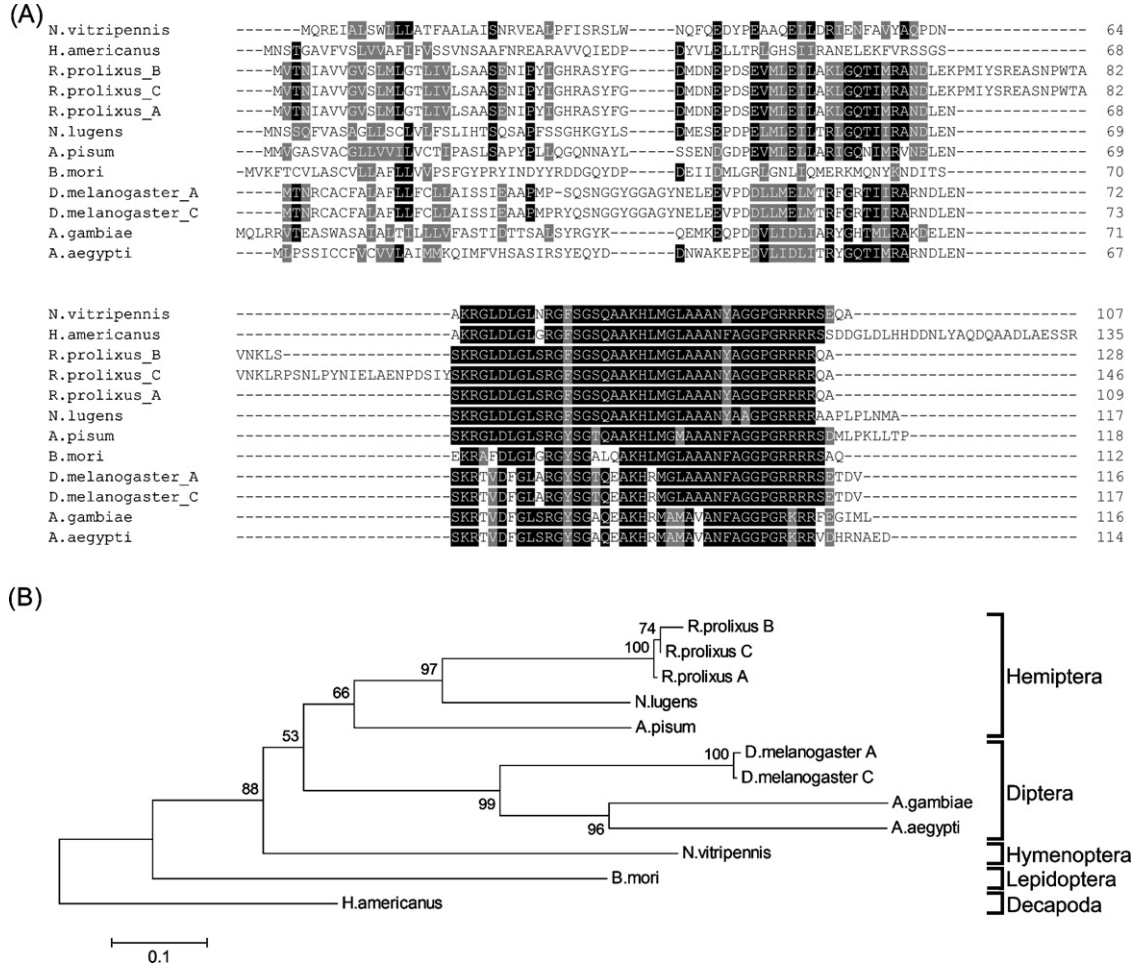

Fig. 3. Multiple sequence alignment of the DH_{31} prepropeptide from different insect orders and a crustacean and its corresponding phylogenetic tree. (A) ClustalX alignment of the DH_{31} prepropeptide from the parasitic wasp, *N. vitripennis* (XP_001599948.1), the kissing bug, *R. prolixus* (transcript variant A – HM030716, transcript variant B – HM030715 and transcript variant C – HM030714), the brown planthopper, *N. lugens* (DB826761), the pea aphid, *A. pisum* (XP_001945901.1), the silk worm, *B. mori* (NP_001124379.1), the fruit fly, *D. melanogaster* (transcript variant A – NM_078790.3 and transcript variant C – NM_164825.2), the malaria mosquito, *A. gambiae* (EAA01397.3/XP_321755.3) and the yellow fever mosquito, *A. aegypti* (EAT40182.1). The sequence from the American lobster, *H. americanus* (GQ290461.1) is also included in the alignment. Following the 50% majority rule, identical amino acids are highlighted in black and similar amino acids are highlighted in gray. (B) A phylogram showing the evolutionary relationships of the 12 taxa inferred using neighbour-joining analysis. *Homarus americanus* taxon is utilized as an outgroup. Boot strap values (only those above 50) obtained on performing the boot strap test with 1000 replicates are indicated where the branches split. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

be 1374 bp long while *Rhopr-DH_{31}-B* and *Rhopr-DH_{31}-C* were found to be 1431 bp and 1485 bp long, respectively. All three *Rhopr-DH_{31}* variants have the same 5′ and 3′ UTRs which are 91 bp and at least 953 bp long, respectively. Within the 3′ UTR of all the variants is a poly(A) tail which is at least 16 nucleotides long. All the variants contain a single open reading frame (ORF) but their lengths vary. *Rhopr-DH_{31}-A*, *Rhopr-DH_{31}-B* and *Rhopr-DH_{31}-C* have an ORF of 330 bp, 387 bp and 441 bp, respectively, which encode apparent prepropeptides of 109, 128 and 146 amino acids, respectively. A highly predicted signal peptide is present in the ORF with the most likely cleavage site occurring between alanine residue at position 23 and serine at position 24. Moreover, the lysine residue at position 68 is predicted to undergo ubiquitination. All three variants produce the same mature peptide consisting of 31 amino acids and located between amino acid positions 108 and 140 (Fig. 2A). Within the prepropeptide sequence, the sequence encoding the mature peptide is flanked by lysine and arginine dibasic amino acid residues at the N terminus and a glycine and four arginine residues at the C terminus. The dibasic amino acid residues at the N terminus and the arginine residues at the C terminus are required for post-translational proteolytic processing following signal peptide cleavage while the glycine residue is required for amidation by peptidyl-glycine α-amidating monooxygenase (see Coast et al., 2002). The mature peptide sequence predicted here is 100% identical to the sequence determined using MALDI-TOF mass spectrometry analysis of the CNS (Te Brugge et al., 2008). Polyadenylation signal consensus sequence (AATAAA) is usually present in the 3′ UTR of a gene. This sequence marks the binding site for proteins that cleave the 3′ end of the RNA after which the poly(A) tail is added (see Zhao et al., 1999). Interestingly, nine such consensus sequences are present in the 3′ UTR of *Rhopr-DH_{31}* sequence. After completing the PCRs to clone the *Rhopr-DH_{31}* 3′ UTR, two sequences of varying lengths were obtained – a larger sequence (presented in Fig. 2A) and a shorter sequence. The shorter sequence was a truncated version of the larger sequence and had a poly(A) tail after nucleotide position 1385 (Fig. 2A).

Molecular organization of *Rhopr-DH_{31}* was determined based on BLAST search and *in silico* analysis using intron prediction (Fig. 2B). The gene comprises six exons (sequentially referred to as 1–6) which are separated by five introns. The lengths of the exons are 89 bp, 93 bp, 112 bp, 57 bp, 54 bp and 1064 bp. *Rhopr-DH_{31}-C* is composed of exons 1–6, *Rhopr-DH_{31}-B* is comprised of exons 1–4 and 6, while *Rhopr-DH_{31}-A* is made up of exons 1–3 and 6. The size of the introns could not be determined since the *R. prolixus* genome

assembly is incomplete. The inclusion of exon 4 in *Rhopr-DH$_{31}$-B* and *Rhopr-DH$_{31}$-C* causes a change in the resulting amino acid sequence and results in a lysine residue at position 68 instead of an asparagine residue at the same position in *Rhopr-DH$_{31}$-A*. Similarly, the presence of exon 5 in *Rhopr-DH$_{31}$-C* results in an arginine residue at position 87 instead of a serine residue found in *Rhopr-DH$_{31}$-B*.

3.2. *DH$_{31}$* sequences in other insects

DH$_{31}$ is found in several insect species; cDNAs encoding DH$_{31}$ prepropeptides have been cloned and sequenced for several species and the sequence predicted for several others. We examined these prepropeptide sequences and aligned them to observe the conservation across various species (Fig. 3A). DH$_{31}$ sequence from the crustacean, *Homarus americanus* was also included in the alignment. 50% majority rule was implemented to highlight the identical and similar amino acids across sequences. The alignment shows that the mature peptide is very well conserved across all insects. Interestingly, prepropeptide sequences encoded by *Rhopr-DH$_{31}$-B* and *Rhopr-DH$_{31}$-C* are considerably longer than other sequences. These two sequences contain a prepropeptide sequence region that is absent in others. This sequence region is present just before the dibasic cleavage site for the mature peptide.

Phylogenetic relationship was inferred using a total of 12 DH$_{31}$ prepropeptide sequences from 8 insect species and 1 crustacean species (*H. americanus*), which was utilized as an outgroup (Fig. 3B). Hemimetabolous insect sequences, comprising hemipterans (*R. prolixus*, *A. pisum* and *N. lugens*), form a monophyletic group since it contains all the hemipteran sequences and their common ancestor. *DH$_{31}$* is present in all four major orders of holometabolous insects but only three of them were included in our analysis. Holometabolous insect sequences, comprising dipterans (*D. melanogaster*, *A. gambiae*, *A. aegypti*), hymenopterans (*N. vitripennis*), and lepidopterans (*B. mori*), form a paraphyletic group since it contains all the insect sequences except for the hemipteran sequences. This observation differs from that of insect phylogeny determined using nuclear gene sequences (Savard et al., 2006), where the holometabolous insects form one monophyletic group and hemimetabolous insects form another monophyletic group. Bootstrap support is fairly low within the clade comprising the hemipteran sequences (>66) but very high within the clade comprising only the dipteran sequences (>96). Moreover, the bootstrap support for the relationships between the different orders is variable, with values ranging from 88 to those less than 50.

3.3. Spatial expression of Rhopr-DH$_{31}$ in fifth-instars

Although Rhopr-DH$_{31}$-like immunoreactivity is present in the CNS, salivary glands, hindgut, anterior dorsal vessel and dorsal diaphragm of fifth-instar *R. prolixus*, immunoreactive cell bodies are present only in the CNS (Te Brugge et al., 2005, 2008). To determine the spatial expression pattern of *Rhopr-DH$_{31}$*, RT-PCR tissue profiling was performed ($n = 3$ independent trials). RT-PCR results demonstrate that *Rhopr-DH$_{31}$* is most likely expressed only within the CNS (Fig. 4) in unfed fifth-instars. The expected size of the PCR product for *Rhopr-DH$_{31}$* was between 317 bp and 428 bp depending on the type of variant being expressed in the tissue. All three transcript variants were observed in CNS, confirming that all are expressed in CNS of fifth-instar *R. prolixus*. In contrast, no PCR product was detected in other tissues in all three trials, except for salivary glands, where a very faint band corresponding to the expected product size for *Rhopr-DH$_{31}$-A* was observed in one of the three trials (not shown). A 229 bp *Rhopr-RP49* fragment was amplified as a positive control to assess the quality of the template cDNA.

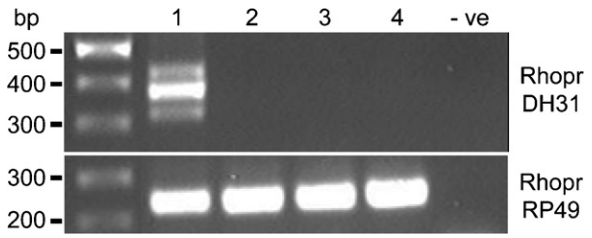

Fig. 4. RT-PCR of *Rhopr-DH$_{31}$* in fifth-instar *R. prolixus*. Tissues were dissected from at least 10 insects and pooled together into four categories: (1) CNS, (2) salivary glands, (3) oesophagus, midgut, hindgut and Malpighian tubules, and (4) remaining tissues (trachea, abdominal nerves, immature testes and ovaries, dorsal and ventral diaphragm and fat bodies). PCR product is strongly observed in the CNS. *Rhopr-RP49* was also amplified as a positive control to test the quality of the template cDNA ($n = 3$ biological replicates).

Bands of high intensity and of expected size were observed for all groups of tissues. Hence, the difference in RT-PCR result observed across the various groups of tissues is due to a difference in *Rhopr-DH$_{31}$* expression and not due to the template used for the reaction.

3.4. Relative expression of the three Rhopr-DH$_{31}$ variants

To determine the relative expression of the three variants, semi-quantitative Southern blot analysis was performed (Fig. 5A). Following a 35 cycle PCR amplification, *Rhopr-DH$_{31}$-A* variant had an average relative expression of 31.5%, *Rhopr-DH$_{31}$-B* variant of 52.6% and *Rhopr-DH$_{31}$-C* of 15.9% ($n = 3$ independent trials) (Fig. 5B). Two-tailed *t*-test shows that the relative expressions of the three transcript variants are significantly different from each other. In addition, a number of different cycles (26, 28, 30, 32 and 35 cycles) were chosen for analysis (not shown) to ensure that the results

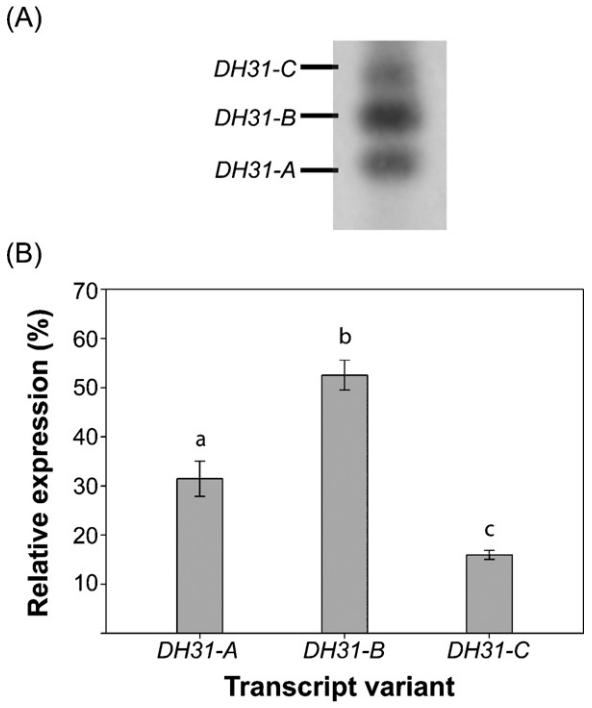

Fig. 5. Relative expression of *Rhopr-DH$_{31}$* variants in fifth-instar CNS. (A) Southern blot analysis of RT-PCR amplified *Rhopr-DH$_{31}$* variants. (B) Semi-quantitative Southern blot analysis of *Rhopr-DH$_{31}$* showing the average relative expression of the three variants. The average expression is based on three independent trials. Bars indicate the standard error. Two-tailed *t*-test ($P < 0.05$) comparison of arcsine transformed data indicate that the average relative expression of each variant is significantly different from others (denoted by letters a–c).

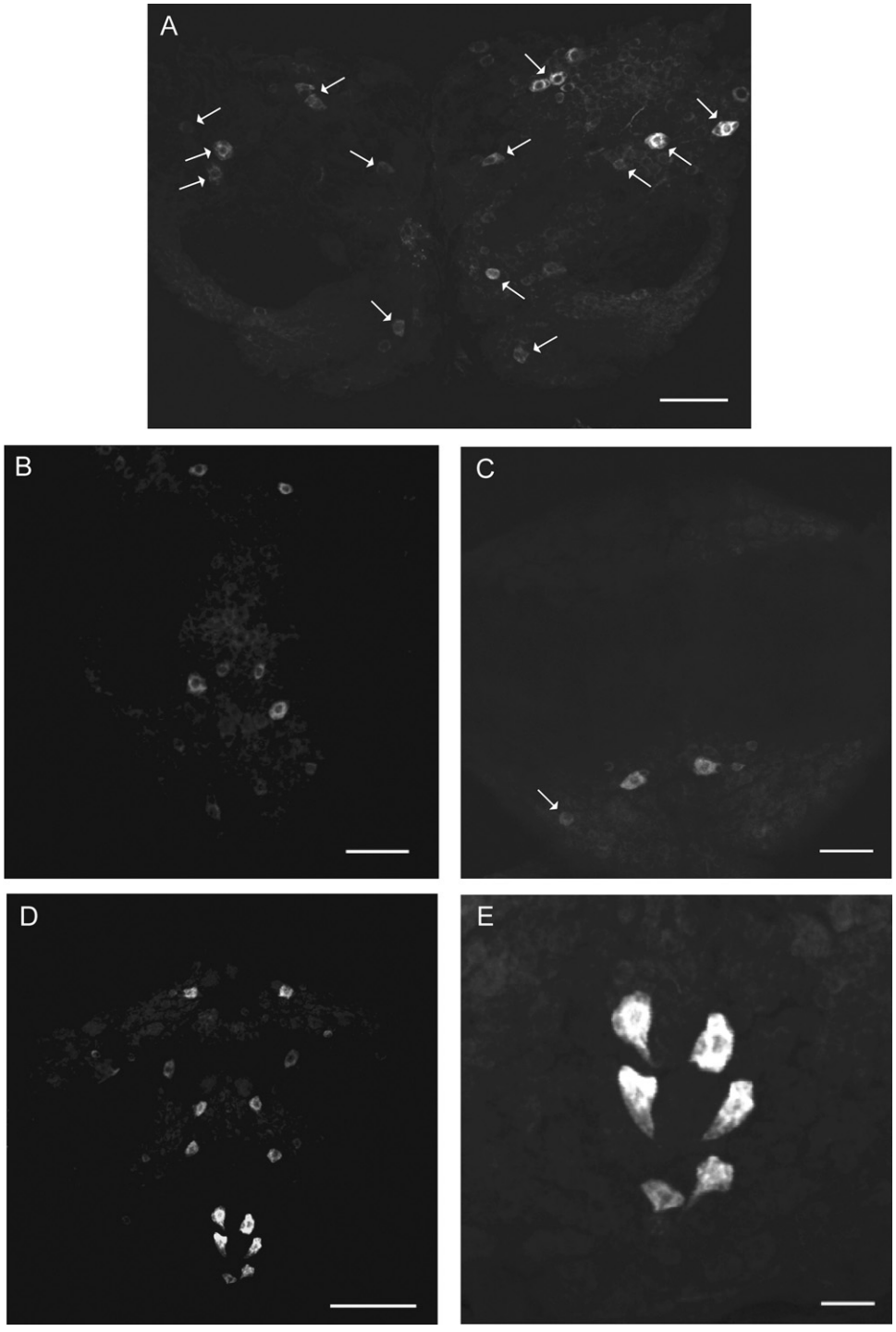

Fig. 6. Fluorescence *in situ* hybridization (FISH) showing *Rhopr-DH$_{31}$* expression in whole mount preparations of (A) dorsal brain, (B) ventral suboesophageal ganglion, (C) ventral prothoracic ganglion, (D) ventral mesothoracic ganglionic mass (MTGM) and (E) higher magnification of ventral MTGM of fifth-instar *R. prolixus*. Where it is unclear, cell-specific staining is indicated by arrows. Scale bars: (A)–(C) = 50 μm, (D) = 100 μm and (E) =20 μm.

were consistent over the exponential amplification phase of PCR and not an artifact of end-point product inhibition.

3.5. *Rhopr-DH$_{31}$ expression in fifth-instar CNS*

The cell-specific expression of *Rhopr-DH$_{31}$* in fifth-instar *R. prolixus* was localized using digoxigenin-labeled anti-sense RNA probe (Fig. 6A–E). *Rhopr-DH$_{31}$* expression is observed in approximately 46 cells of the CNS. No staining is observed in these cells in the control preparations with the digoxigenin-labeled sense RNA probe (not shown). On examining *Rhopr-DH$_{31}$* expression in all the components of the CNS which includes the brain, suboesophageal ganglion (SOG), prothoracic ganglion (PRO) and mesothoracic ganglionic mass (MTGM), it can be seen that the gene expression localization is basically organized into bilaterally symmetrical groups or pairs of cells. Within the dorsal brain, 18 cells containing the *Rhopr-DH$_{31}$* transcript(s) are faintly to moderately stained (Fig. 7). *Rhopr-DH$_{31}$* expression is also observed in six moderately-stained cells in the ventral SOG, and in four bilaterally-paired cells in the ventral PRO (Fig. 7). *Rhopr-DH$_{31}$* expression is observed in 16 cells within the

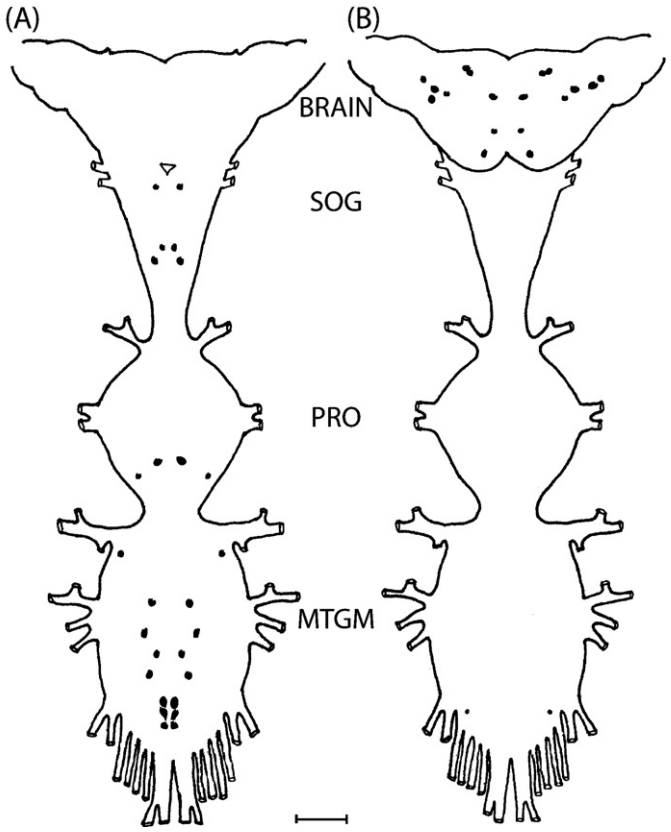

Fig. 7. Diagrammatic representation of *Rhopr-DH$_{31}$* transcripts expression (A) ventral aspect of the CNS (brain), suboesophageal ganglion (SOG), prothoracic ganglion (PRO), and mesothoracic ganglionic mass (MTGM). (B) Dorsal aspect of the CNS. Map based on 20 preparations. Scale bar: 200 μm.

ventral MTGM (Fig. 7), with a group of 6 strongly stained cells (Figs. 6D and E and 7) localized along the midline in the posterior MTGM; these cells stain the strongest and are observed in almost all the preparations. However, these cells were variable in number, with few a preparations staining either 5 or 7 cells. Lastly, *Rhopr-DH$_{31}$* expression is observed in a pair of cells in the dorsal MTGM (Fig. 7); these cells were very faintly stained and were observed in a few preparations only.

4. Discussion

We have isolated and characterized a calcitonin-like *DH$_{31}$* from *R. prolixus*. This is the first study to identify and isolate three *DH$_{31}$* splice variants in an arthropod species. A modified approach was adopted to obtain the complete cDNA sequences for the three variants. Surprisingly, the 3′ UTR contains nine polyadenylation signal consensus sequences. It is fairly unlikely that all these nine consensus sequences are functional and are most likely present due to the 3′ UTR being highly AT-rich. However, PCRs to clone the *Rhopr-DH$_{31}$* 3′ UTR resulted in two sequences of varying lengths – a larger sequence (presented in Fig. 2A) and a shorter sequence. Cleavage of the transcript occurs approximately 10–30 nucleotides downstream of the polyadenylation signal consensus sequence (see Zhao et al., 1999). There is a polyadenylation consensus sequence 19 nucleotides upstream of the larger sequence and there are two consensus sequences between 16 and 27 nucleotides upstream of the shorter sequence. Also, the sequence preceding the poly(A) tail of this shorter sequence consists of six adenosine nucleotides. Therefore, the shorter version of the sequence could be another result of unspecific binding of the oligo(dT) primer or the two sequences of varying lengths could be a result of alternative polyadenylation, similar to the mammalian calcitonin and calcitonin gene-related peptide (CGRP) gene processing (Amara et al., 1982; Sabate et al., 1985). The mammalian calcitonin gene undergoes alternative splicing to generate two products – calcitonin and CGRP, both of which have a different poly(A) site. It is unclear if the three *Rhopr-DH$_{31}$* transcripts have different poly(A) sites or if any possible effects result from the alternate use of poly(A) site in *R. prolixus*.

Recently, Ons et al. submitted *R. prolixus* CLDH ORF sequences of two isoforms (isoform A – GQ856316 and isoform B – GQ856317) to the GenBank database. Isoform A is identical to *Rhopr-DH$_{31}$-C* and isoform B is identical to *Rhopr-DH$_{31}$-A*. However, the *Rhopr-DH$_{31}$-A* and *Rhopr-DH$_{31}$-C* sequences presented here also include the additional 3′ UTR and 5′ UTR sequences. These isoforms are indeed splice variants as confirmed by our analyses. Moreover, BLAST search results indicate that *Rhopr-DH$_{31}$* may have undergone partial gene duplication (Hu and Worton, 1992). Specifically, *Rhopr-DH$_{31}$-C* cDNA sequence was used as a query to perform a BLAST search of the *R. prolixus* preliminary genome assembly. Interestingly, two unique highly similar hits, each located on a different supercontig, were returned for exons 5 and 6, suggesting that at least a portion of the gene is duplicated. Closer examination of the three transcript variants illustrates that residues present in the *Rhopr-DH$_{31}$-C* variant are missing in *Rhopr-DH$_{31}$-A* and *Rhopr-DH$_{31}$-B*. A gap is also present in the multiple sequence alignment of the DH$_{31}$ prepropeptide sequences from *R. prolixus* and other invertebrate species. The gap is present just before the sequence coding for the mature peptide, and is due to *Rhopr-DH$_{31}$-B* and *Rhopr-DH$_{31}$-C* variants being longer than other sequences. Comparing this result with the molecular organization of *Rhopr-DH$_{31}$* shows that exons 4 and 5 are responsible for the gap observed in the alignment. Phylogenetic analysis of DH$_{31}$ prepropeptide sequences indicates that *Rhopr-DH$_{31}$-A* is more closely related to the *N. lugens* sequence compared to the other two variants. This suggests that *Rhopr-DH$_{31}$-A* variant, which lacks exons 4 and 5, is the ancestral version of the gene, and the *Rhopr-DH$_{31}$-B* and *Rhopr-DH$_{31}$-C* variants possibly result from DNA insertion caused by an event such as horizontal gene transfer or through DNA transposons (McClintock, 1950; Syvanen, 1985).

Tissue-specific expression of the three *Rhopr-DH$_{31}$* splice variants could explain their presence in *R. prolixus*, especially since each variant is predicted to generate the same mature Rhopr-DH$_{31}$ peptide through post-translational modification and processing (see Veenstra, 2000). Although *Rhopr-DH$_{31}$* variants are most likely expressed only within the CNS, it does not eliminate the possibility of non-neuronal expression of this gene as is found in *D. melanogaster* midgut (Veenstra, 2009). The presence of a very faint product corresponding to *Rhopr-DH$_{31}$-A* in salivary glands for only one of the three RT-PCR trials (not shown) suggests that *Rhopr-DH$_{31}$-A* expression in salivary glands is very low compared to that of CNS. Likewise, one or more *Rhopr-DH$_{31}$* splice variants could be expressed in other tissues but their expression is too low to be detected. Furthermore, it is important to keep note of the fact that we are only examining one particular stage of *R. prolixus* development (6 weeks post-fed fifth-instars). Therefore, *Rhopr-DH$_{31}$* spatial expression pattern could be different when examined at other life stages.

The three *Rhopr-DH$_{31}$* variants significantly differ in their relative expression within the CNS. Since the 5′ end of all the three variants is similar, they probably share the same promoter. Thus, the difference in relative expression of the three splice variants could be due to the variants being expressed in a cell-specific manner within the CNS. This could be beneficial in regulating the amount of mature peptide synthesized in different neurons since increased amounts of mRNA would most likely result in increased amounts of protein. Moreover, the difference in relative expression

could be due to the difference in transcripts generated from alternative splicing. As mentioned earlier, exon 4, present in *Rhopr-DH$_{31}$-B* and exons 4 and 5, present in *Rhopr-DH$_{31}$-C* result in these transcripts being longer than their homologs from other insect species. The propeptide region coded by exons 4 and 5 is not similar to any known biologically active peptides (as determined by BLAST search). However, specific lysine residues within these peptide sequences could undergo various post-translational modifications such as ubiquitination or SUMOylation which would affect peptide stability (see Haglund and Dikic, 2005; Hay, 2005; Hochstrasser, 1995). Specifically, the presence of exon 4 in *Rhopr-DH$_{31}$-B* and *Rhopr-DH$_{31}$-C* results in a lysine residue at position 68 instead of asparagine residue at the same position in *Rhopr-DH$_{31}$-A*, which could be critical since this lysine is predicted to be ubiquitinated (Radivojac et al., 2010).

FISH was performed to localize *Rhopr-DH$_{31}$* expression within the CNS of fifth-instar *R. prolixus*. Comparing these results with the immunohistochemical localization of Rhopr-DH$_{31}$-like peptide (Te Brugge et al., 2005), additional cells stain with the immunohistochemical procedure than are observed with FISH. Specifically, Rhopr-DH$_{31}$-like immunoreactivity is observed in cells in the ventral brain, dorsal SOG and dorsal PRO where *Rhopr-DH$_{31}$* expression is not observed. *Rhopr-DH$_{31}$* expression was also not observed in dorsal unpaired median (DUM) neurons in the MTGM, although it must be noted that these stain quite weakly with immunohistochemistry. *Rhopr-DH$_{31}$* could yet still be expressed in DUM neurons; however, its expression appears to be too low to be detected using the approach and stage used here in this study. In terms of the similarities between the immunohistochemical localization and FISH, most of the cell-specific staining observed with FISH is also observed using immunohistochemistry. In particular, on the posterior ventral surface of MTGM, six strongly immunoreactive bilaterally-paired cells also demonstrate *Rhopr-DH$_{31}$* expression. These cells are very strongly stained and observed in almost all the preparations. However, the numbers of these cells varied as some preparations contained either five or seven cells instead of six. In addition, consistent with the immunohistochemical localization, *Rhopr-DH$_{31}$* expression is also localized to lateral and medial neurosecretory cells in the dorsal brain. This provides evidence that DH$_{31}$ produced by these cells might be released into the haemolymph via the corpus cardiacum for action as a neurohormone.

DH$_{31}$ is present in several arthropod species. Phylogenetic analysis of DH$_{31}$ prepropeptide sequences indicates that the *R. prolixus* sequence is most similar to that of *N. lugens*. It is interesting that holometabolous insect CLDH sequences form a paraphyletic group rather than a monophyletic group as inferred using nuclear gene sequences (Savard et al., 2006). All the dipteran sequences and hemipteran sequences form distinct clades, each with strong bootstrap support. Moreover, the predicted Rhopr-DH$_{31}$ mature peptide is 100% identical to Dippu-DH$_{31}$ and Apime-DH$_{31}$ peptides. Within the DH$_{31}$ sequence proper, 13 out of the 31 (or 34) amino acid residues are identical in sequences from all the species observed. Specifically, the C-terminal domain of the peptide sequence shows high identity across various species, which follow varying developmental patterns and utilize different feeding strategies. The high conservation observed for this peptide across various species suggests that apart from being involved in processes associated with diuresis, it could play other crucial roles in arthropods which are still unclear.

In conclusion, the complete cDNA sequences of three *Rhopr-DH$_{31}$* splice variants have been obtained in *R. prolixus*. Although it is evident from *in situ* hybridization that *Rhopr-DH$_{31}$* is expressed in the CNS, it is still unclear which cells within the CNS express the different transcripts. It is evident that CLDHs are multi-functional peptides and have a wide array of functions in arthropods including increasing Malpighian tubules secretion, and hindgut and dorsal vessel contractility.

Acknowledgements

We would like to thank Victoria Te Brugge for sharing the DH$_{31}$ peptide sequence prior to its publication. This research was funded through an NSERC Discovery Grant to I.O.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.mce.2010.08.012.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., Evans, R.M., 1982. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature 298, 240–244.
Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.
Christie, A.E., 2008. Neuropeptide discovery in Ixodoidea: an *in silico* investigation using publicly accessible expressed sequence tags. Gen. Comp. Endocrinol. 157, 174–185.
Christie, A.E., Stevens, J.S., Bowers, M.R., Chapline, M.C., Jensen, D.A., Schegg, K.M., Goldwaser, J., Kwiatkowski, M.A., Pleasant Jr., T.K., Shoenfeld, L., Tempest, L.K., Williams, C.R., Wiwatpanit, T., Smith, C.M., Beale, K.M., Towle, D.W., Schooley, D.A., Dickinson, P.S., 2010. Identification of a calcitonin-like diuretic hormone that functions as an intrinsic modulator of the American lobster, *Homarus americanus*, cardiac neuromuscular system. J. Exp. Biol. 213, 118–127.
Coast, G.M., Garside, C.S., Webster, S.G., Schegg, K.M., Schooley, D.A., 2005. Mosquito natriuretic peptide identified as a calcitonin-like diuretic hormone in *Anopheles gambiae* (Giles). J. Exp. Biol. 208, 3281–3291.
Coast, G.M., Orchard, I., Phillips, J.E., Schooley, D.A., 2002. Insect diuretic and antidiuretic hormones. In: Evans, P.D. (Ed.), Advances in Insect Physiology. Academic Press Ltd., London, pp. 279–409.
Coast, G.M., Webster, S.G., Schegg, K.M., Tobe, S.S., Schooley, D.A., 2001. The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian tubules. J. Exp. Biol. 204, 1795–1804.
Donini, A., O'Donnell, M.J., Orchard, I., 2008. Differential actions of diuretic factors on the Malpighian tubules of *Rhodnius prolixus*. J. Exp. Biol. 211, 42–48.
Dujardin, J.P., Munoz, M., Chavez, T., Ponce, C., Moreno, J., Schofield, C.J., 1998. The origin of *Rhodnius prolixus* in Central America. Med. Vet. Entomol. 12, 113–115.
Feliciangeli, M.D., Carrasco, H., Patterson, J.S., Suarez, B., Martinez, C., Medina, M., 2004. Mixed domestic infestation by *Rhodnius prolixus* Stal, 1859 and *Panstrongylus geniculatus* Latreille, 1811, vector incrimination, and seroprevalence for *Trypanosoma cruzi* among inhabitants in El Guamito, Lara State, Venezuela. Am. J. Trop. Med. Hyg. 71, 501–505.
Furuya, K., Milchak, R.J., Schegg, K.M., Zhang, J., Tobe, S.S., Coast, G.M., Schooley, D.A., 2000. Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects. Proc. Natl. Acad. Sci. U.S.A. 97, 6469–6474.
Gard, A.L., Lenz, P.H., Shaw, J.R., Christie, A.E., 2009. Identification of putative peptide paracrines/hormones in the water flea *Daphnia pulex* (Crustacea; Branchiopoda; Cladocera) using transcriptomics and immunohistochemistry. Gen. Comp. Endocrinol. 160, 271–287.
Haglund, K., Dikic, I., 2005. Ubiquitylation and cell signaling. EMBO J. 24, 3353–3359.
Hay, R.T., 2005. SUMO: a history of modification. Mol. Cell 18, 1–12.
Hochstrasser, M., 1995. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. Curr. Opin. Cell Biol. 7, 215–223.
Hu, X., Worton, R.G., 1992. Partial gene duplication as a cause of human disease. Hum. Mutat. 1, 3–12.
Kean, L., Cazenave, W., Costes, L., Broderick, K.E., Graham, S., Pollock, V.P., Davies, S.A., Veenstra, J.A., Dow, J.A., 2002. Two nitridergic peptides are encoded by the gene capability in *Drosophila melanogaster*. Am. J. Physiol. Regul. Integr. Comp. Physiol. 282, R1297–1307.
Koberle, F., 1968. Chagas' disease and Chagas' syndromes: the pathology of American trypanosomiasis. Adv. Parasitol. 6, 63–116.
Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., Williamson, M., Arakane, Y., Verleyen, P., Schoofs, L., Schachtner, J., Grimmelikhuijzen, C.J., Park, Y., 2008. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res. 18, 113–122.
Maddrell, S.H., Herman, W.S., Mooney, R.L., Overton, J.A., 1991. 5-Hydroxytryptamine: a second diuretic hormone in *Rhodnius prolixus*. J. Exp. Biol. 156, 557–566.
McClintock, B., 1950. The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. U.S.A. 36, 344–355.

Monroy, C., Rodas, A., Mejia, M., Rosales, R., Tabaru, Y., 2003. Epidemiology of Chagas disease in Guatemala: infection rate of *Triatoma dimidiata*, *Triatoma nitida* and *Rhodnius prolixus* (Hemiptera, Reduviidae) with *Trypanosoma cruzi* and *Trypanosoma rangeli* (Kinetoplastida, Trypanosomatidae). Mem. Inst. Oswaldo Cruz 98, 305–310.

Paluzzi, J.P., Orchard, I., 2010. A second gene encodes the anti-diuretic hormone in the insect, *Rhodnius prolixus*. Mol. Cell. Endocrinol. 317, 53–63.

Paluzzi, J.P., Russell, W.K., Nachman, R.J., Orchard, I., 2008. Isolation, cloning, and expression mapping of a gene encoding an antidiuretic hormone and other CAPA-related peptides in the disease vector, *Rhodnius prolixus*. Endocrinology 149, 4638–4646.

Prata, A., 2001. Clinical and epidemiological aspects of Chagas disease. Lancet Infect. Dis. 1, 92–100.

Radivojac, P., Vacic, V., Haynes, C., Cocklin, R.R., Mohan, A., Heyen, J.W., Goebl, M.G., Iakoucheva, L.M., 2010. Identification, analysis, and prediction of protein ubiquitination sites. Proteins 78, 365–380.

Ramsey, J.M., Ordonez, R., Cruz-Celis, A., Alvear, A.L., Chavez, V., Lopez, R., Pintor, J.R., Gama, F., Carrillo, S., 2000. Distribution of domestic triatominae and stratification of Chagas Disease transmission in Oaxaca, Mexico. Med. Vet. Entomol. 14, 19–30.

Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in Genie. J. Comput. Biol. 4, 311–323.

Sabate, M.I., Stolarsky, L.S., Polak, J.M., Bloom, S.R., Varndell, I.M., Ghatei, M.A., Evans, R.M., Rosenfeld, M.G., 1985. Regulation of neuroendocrine gene expression by alternative RNA processing. Colocalization of calcitonin and calcitonin gene-related peptide in thyroid C-cells. J. Biol. Chem. 260, 2589–2592.

Savard, J., Tautz, D., Richards, S., Weinstock, G.M., Gibbs, R.A., Werren, J.H., Tettelin, H., Lercher, M.J., 2006. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. Genome Res. 16, 1334–1338.

Schooley, D.A., Horodyski, F.M., Coast, G.M., 2005. Hormones controlling homeostasis in insects: endocrinology. In: Gilbert, L.I., et al. (Eds.), Comprehensive Molecular Insect Science. Elsevier, pp. 493–550.

Syvanen, M., 1985. Cross-species gene transfer; implications for a new theory of evolution. J. Theor. Biol. 112, 333–343.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Te Brugge, V.A., Ianowski, J.P., Orchard, I., 2009. Biological activity of diuretic factors on the anterior midgut of the blood-feeding bug, *Rhodnius prolixus*. Gen. Comp. Endocrinol. 162, 105–112.

Te Brugge, V.A., Lombardi, V.C., Schooley, D.A., Orchard, I., 2005. Presence and activity of a Dippu-DH_{31}-like peptide in the blood-feeding bug, *Rhodnius prolixus*. Peptides 26, 29–42.

Te Brugge, V.A., Miksys, S.M., Coast, G.M., Schooley, D.A., Orchard, I., 1999. The distribution of a CRF-like diuretic peptide in the blood-feeding bug *Rhodnius prolixus*. J. Exp. Biol. 202, 2017–2027.

Te Brugge, V.A., Orchard, I., 2008. Distribution and activity of a Dippu DH_{31}-like peptide in the large milkweed bug *Oncopeltus fasciatus*. Peptides 29, 206–213.

Te Brugge, V.A., Schooley, D.A., Orchard, I., 2002. The biological activity of diuretic factors in *Rhodnius prolixus*. Peptides 23, 671–681.

Te Brugge, V.A., Schooley, D.A., Orchard, I., 2008. Amino acid sequence and biological activity of a calcitonin-like diuretic hormone (DH_{31}) from *Rhodnius prolixus*. J. Exp. Biol. 211, 382–390.

Veenstra, J.A., 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch. Insect Biochem. Physiol. 43, 49–63.

Veenstra, J.A., 2009. Peptidergic paracrine and endocrine cells in the midgut of the fruit fly maggot. Cell Tissue Res. 336, 309–323.

Zhao, J., Hyman, L., Moore, C., 1999. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol. Mol. Biol. Rev. 63, 405–445.