RIDAA
Repositorio Institucional
Digital de Acceso Abierto de la
Universidad Nacional de Quilmes

Universidad
Nacional
de Quilmes

Sterkel, Marcos

Genómica y proteómica de neuropéptidos en Rhodnius prolixus

Esta obra está bajo una Licencia Creative Commons Argentina.
Atribución - No Comercial - Sin Obra Derivada 2.5
https://creativecommons.org/licenses/by-nc-nd/2.5/ar/

Documento descargado de RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes de la Universidad Nacional de Quilmes

Cita recomendada:
Sterkel, M. (2011). Genómica y proteómica de neuropéptidos en Rhodnius prolixus. (Tesis de doctorado). Universidad Nacional de Quilmes, Bernal, Argentina. Disponible en RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes
http://ridaa.unq.edu.ar/handle/20.500.11807/123

Puede encontrar éste y otros documentos en: https://ridaa.unq.edu.ar

Universidad
Nacional
de Quilmes

Roque Sáenz Peña 352 // Bernal
Buenos Aires // Argentina
t.: (+41 11) 4365 7100
f.: (+54 11) 4365 7101
info@unq.edu.ar

Sterkel, Marcos, Repositorio Institucional Digital de Acceso Abierto,
diciembre de 2011, pp. 112,
http://ridaa.demo.unq.edu.ar,
Universidad Nacional de Quilmes, Secretaría de Posgrado,
Doctorado en Ciencias Básicas y Aplicadas.

Genómica y proteómica de neuropéptidos en Rhodnius prolixus

TESIS DOCTORAL

Marcos Sterkel

msterkel@gmail.com

Resumen

La enfermedad de Chagas afecta a 10 millones de personas en el continente americano, mientras que 80 millones se encuentran en riesgo de contraerla. El agente causante de la enfermedad es el protozoo *Tripanosoma cruzi*, transmitido por insectos hematófagos del orden *Hemiptera,* siendo las especies más importantes en la transmisión R*hodinus prolixus* y *Triatoma infestans*.

Debido a la falta de tratamiento en la etapa crónica de la enfermedad, el control de las poblaciones del vector mediante la fumigación con insecticidas piretroides ha sido, y sigue siendo, la principal estrategia de prevención. Sin embargo, en los últimos años se han detectado poblaciones con una altísima resistencia a este tipo de insecticidas (Picollo *et al*, 2005). Dado que los neuropéptidos se encuentran involucrados en la regulación de todos los procesos fisiológicos de vida de un insecto (ver Nässel, 2002), se ha propuesto la manipulación de los sistemas de señalización peptídicos como una posible estrategia para el diseño de nuevos insecticidas.

En el presente trabajo hemos realizado una caracterización exhaustiva de los neuropéptidos presentes en *R. prolixus* a nivel de secuencia y estructura, utilizando técnicas bioinformáticas, proteómicas y de biología molecular. Por otra parte, hemos aplicado técnicas de peptidómica cuantitativa para caracterizar la respuesta post-ingesta sanguínea, a fin de comenzar la caracterización fisiológica de los neuropéptidos identificados. Nuestros resultados indican que el neuropeptidoma de *R. prolixus* es apenas diferente de los demás insectos con el genoma secuenciado y ensamblado, todos ellos holometábolos, presentando particularidades únicas a nivel de secuencia. Este trabajo es el punto de partida para posteriores estudios a nivel fisiológico, y sienta las bases necesarias para posteriores estudios funcionales, así como también para estudios de interacción con los receptores y para estudios comparativos y/o evolutivos con otros artrópodos. Se trata del primer trabajo de este tipo realizado en un insecto vector de enfermedades

Introducción general

1) La enfermedad de Chagas

La enfermedad de Chagas afecta a 10 millones de personas en el continente americano, mientras que 80 millones se encuentran en riesgo de contraerla. Su distribución abarca un área que va desde el sur de los Estados Unidos hasta el sur de nuestro país (Fig. 1) (http://www.who.int/mediacentre/factsheets/fs340/en/). El agente causante de la enfermedad es el protozoo *Tripanosoma cruzi*, transmitido a través de las heces por insectos hematófagos del orden *Hemiptera* pertenecientes a la familia *Reduviidae* (Subfamilia *Triatominae*), conocidos popularmente en Argentina como "vinchucas".

Figura 1: Distribución de la enfermedad de Chagas en América.

Las distintas especies de la familia *Reduviidae* (Figura 1.2) se distribuyen entre las latitudes 42° N y 46° S abarcando un área comprendida entre los grandes lagos de América del Norte y el sur de la Argentina (Gorla et al., 1997). Las especies más importantes en cuanto a la transmisión del parásito son R*hodinus prolixus* en el sur de América del Norte y América Central, mientras que *Triatoma infestans* y *Pastrongylus megistus* son los principales vectores en América del Sur. En Argentina se estima que existen aproximadamente 2.300.000 enfermos de Chagas, lo que representa el 7,2% de la población. (http://www.fac.org.ar/fec/chagas/fatala/enelpais.htm).

Debido a la falta de un tratamiento efectivo en la etapa crónica de la enfermedad, el control de las poblaciones del insecto vector mediante la fumigación con insecticidas piretroides ha sido, y sigue siendo, la principal estrategia de prevención (Zerba, 1999). Sin embargo, en los últimos años se han detectado poblaciones con una altísima resistencia a este tipo de insecticidas en las provincias argentinas de Mendoza, San Luis, Catamarca, Salta y Chaco, así como en el sur de Bolivia (Gonzalez Audino *et al*, 2004; Picollo *et al*,

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

2005). La aparición de poblaciones resistentes a insecticidas constituye una amenaza para los programas de control de Chagas existentes hasta la fecha.

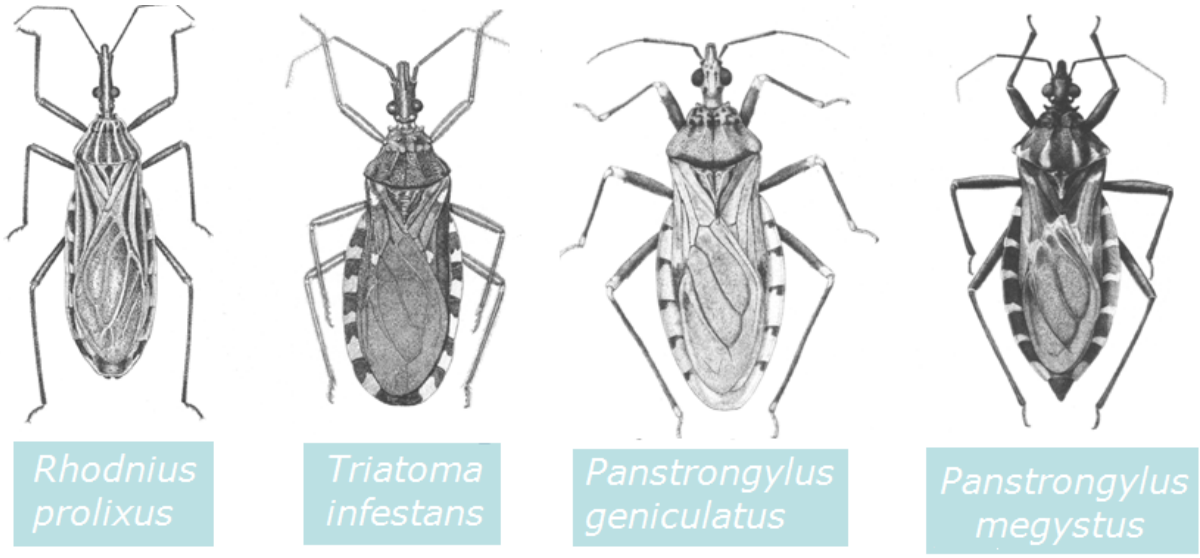


Figura 2: Especies de la familia *Reduviidae* con importancia vectorial.

La plasticidad genética en insectos que permite la rápida aparición de poblaciones resistentes a insecticidas es bien conocida (Read *et al*, 2009), y plantea la necesidad de dirigir esfuerzos hacia la obtención de nuevas estrategias de control. Dado que los neuropéptidos se encuentran involucrados en la regulación de todos los procesos fisiológicos durante el ciclo de vida de un insecto (ver Nässel, 2002), se ha propuesto la manipulación de los sistemas de señalización peptídicos como una posible estrategia para el diseño y desarrollo de insecticidas de nueva generación. Sin embargo, debido a sus características farmacocinéticas y la corta vida media en el ambiente, estos péptidos no pueden ser utilizados por sí mismos su como insecticidas. Por esta razón, el diseño de pseudopéptidos o péptido-miméticos, capaces de interferir con aspectos críticos de la fisiología de los insectos, se ha propuesto como una vía novedosa hacia la obtención de estos nuevos insecticidas (Scherkenbeck *et al*, 2009). De esta manera, el desarrollo de resistencia ligado al incremento en los mecanismos detoxificativos, que complica el éxito de los insecticidas tradicionales, podría ser superado con nuevos compuestos basados en moléculas endógenas. Con dicha estrategia además, se podría aumentar la especificidad por parte del insecticida hacia la especie a combatir, resultando en un mejor cuidado del medio ambiente y la salud humana. Para esto se requiere la identificación y caracterización, tanto molecular como fisiológica, de los neuropéptidos y sus receptores (ver Nässel, 2002).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

2) Comunicación celular en insectos: mensajeros químicos.

La comunicación celular resulta fundamental para coordinar las funciones entre células a nivel de poblaciones, tejidos y sistemas. Los sistemas de señalización aparecieron tempranamente durante el proceso de evolución; se encuentran desde grupos ancestrales como protistas (Le Roith *et al* 1980), cnidarios (ver Grimmelikhuijzen *et al* 1996) y turbellarios (Platyhelminthes; Adami *et al*, 2010) hasta vertebrados. En organismos multicelulares, las células se comunican a través de señales químicas de diferente naturaleza, incluyendo iones minerales, lípidos, aminoácidos y péptidos.

La comunicación celular puede darse entre células que se encuentran próximas (nanómetros), o bien a distancia (ver Hartenstein, 2006). En el primer caso las señales químicas, neurotransmisores, son liberadas al espacio extracelular en sitios particulares, llamados sinapsis. Entre los mecanismos de comunicación a distancia se encuentran los sistemas autócrinos, parácrinos y endocrinos. En los sistemas autócrinos y parácrinos las hormonas son liberadas al espacio extracelular y actúan a nivel local, regulando la actividad de las mismas células productoras de la señal (autócrinos) o de células vecinas (parácrinos). En el sistema endocrino las señales químicas son liberadas al espacio extracelular, desde donde difunden hasta la hemolinfa para actuar a nivel sistémico.

En insectos, la mayoría de las hormonas y neurotransmisores son moléculas peptídicas (neuropéptidos). Los neuropéptidos maduros, con actividad biológica, se generan a partir del clivaje proteolítico de precursores de mayor tamaño. Los precursores de neuropéptidos codifican en su secuencia aminoacídica un péptido señal, y un número variable de péptidos, separados entre sí por sitios de clivaje proteolítico (ver Nässel, 2002). El clivaje se produce generalmente en sitios dibásicos o monobásicos (ver Veenstra, 2000). En muchos casos, los péptidos son modificados post-traduccionalmente por diferentes enzimas para alcanzar el estado de maduración final (amidación C-terminal, oxidación, acetilación, etc). Las modificaciones post-traduccionales contribuyen a la estabilidad del péptido, y resultan importantes para su interacción con el receptor. En cuanto a los receptores de neuropéptidos, todos los conocidos pertenecen a la familia de receptores de membrana acoplados a proteína G (GPCRs).

Es común que un precursor dé origen a múltiples péptidos que presentan un motivo en común que los distingue como familia (ver Nässel, 2002). Algunas veces, los péptidos originados en un mismo precursor tienen la misma función, mientras que otras presentan funciones diferentes e incluso opuestas (Zhang *et al*, 2005). Es por ello que resulta importante conocer todos los péptidos originados a partir de un precursor, ya que cada uno de ellos puede tener una relevancia fisiológica diferente. Los precursores peptídicos son sintetizados y transportados al Retículo Endoplasmático Rugoso (RER), se procesan en el aparato de Golgi y luego los péptidos maduros se almacenan en vesículas hasta el momento de la exositosis (Pow y Golding, 1987). Además del Sistema nervioso (SN),

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

células del digestivo medio (Brown *et al*, 1999; Lee, 2002; Sarkar *et al*, 2003, Kwok *et al*, 2005) y testículos (Paluzzi *et al*, 2008) han demostrado ser capacer de producir péptidos regulatorios. Recientemente se ha demostrado que en células de los Túbulos de Malphigi de *T. infestans* se expresa un péptido simil-Allatotropina (Santini and Ronderos, 2009a).

Los sistemas de señalización peptídicos participan en la regulación de todas las actividades fisiológicas de un organismo: alimentación, reproducción, crecimiento, metamorfosis, etc. También participan en la regulación de las rutas metabólicas: síntesis, almacenamiento y liberación de carbohidratos y lípidos, regulación del balance osmótico, hídrico y mineral, formación y maduración de las gametas, crecimiento y regeneración tisular, y regulación de ritmos circadianos (ver más adelante). Varios de estos péptidos muestran un alto grado de conservación a nivel de secuencia entre distintas especies, por lo que se ha sugerido que han evolucionado a partir de un ancestro común. Como ejemplos se pueden citar el péptido simil-Arginina-Vasopresina (AVPL), los péptidos simil-insulina (ILPs) y el neuropéptido F largo (LNF o LNY en vertebrados), conservados tanto en vertebrados como en invertebrados.

Además de las hormonas peptídicas, derivados de lípidos y aminoácidos pueden actuar como hormonas. Como ejemplo en insectos se pueden citar la hormona de la muda (ecdisona), que deriva de esteroles, y las hormonas juveniles que son sintetizadas a partir de de ácidos grasos poly-insaturados. Estas hormonas son sintetizadas en el Retículo Endoplasmático Liso (REL) y no son almacenadas en vesículas. Sus receptores, llamados receptores nucleares, pertenecen a una clase de factores de trascripción que se localizan en el citoplasma de las células blanco en estado inactivo. Cambian al estado activo cuando se les une el ligando e ingresan al núcleo donde regulan la expresión de determinados genes (Schulster *et al*, 1976)

3) El sistema nervioso de triatominos

El sistema nerviosos (SN) de los insectos se encuentra conformado por tres regiones: el Sistema Nervioso Central (SNC), compuesto por el cerebro y el cordón nervioso ventral; el Sistema Nervioso Periférico (SNP), que se encuentra constituido por fibras nerviosas tanto sensoriales como motoras; y el Sistema Nervioso Estomatogástrico (SNS; Insausti T. C, 1994).

El SNC se compone del cerebro y la cadena gangliolar ventral que en insectos triatominos esta formada por el ganglio subesofágico (SOG), el ganglio prototoráxico (PG) y el ganglio mesometatoraraxico (MMTG) o posterior (PoG), siendo este último el resultado de la fusión de los ganglios meso y metatoraxicos, y los ganglios abdominales. A partir del MMTG se generan cinco pares de nervios que inervan los distintos segmentos del abdomen. El par central es el más largo e inerva el segmento genital (AbN V). Desde los nervios abdominales se liberan neuropéptidos a la hemolinfa. Del SOG se generan

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

proyecciones nerviosas que inervan la *Corpora Cardíaca* (NCC III), y las glándulas salivales (Figura 3). Este ganglio es también responsable del movimiento de las partes orales (Griss, 1990).

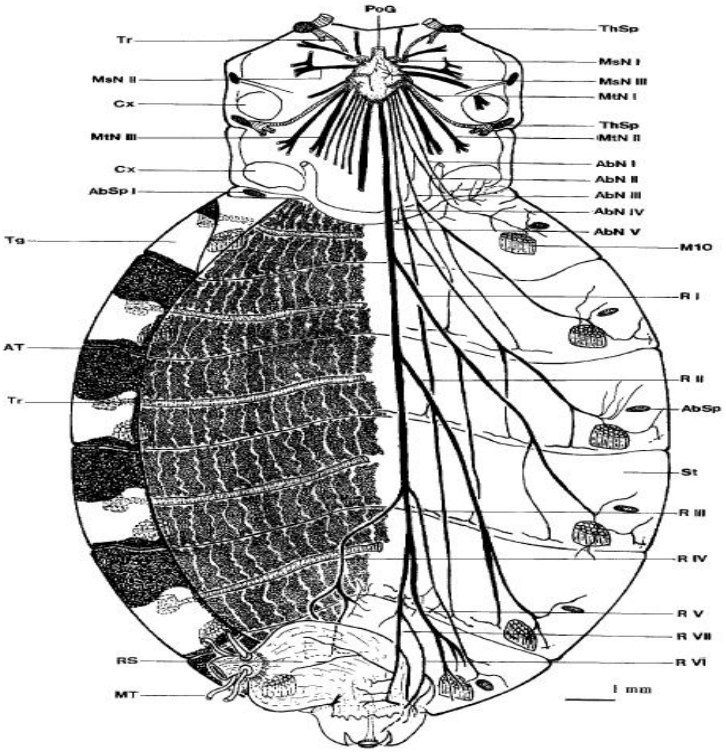


Figura 3 : Vista dorsal del MMTG. Se muestra la inervación sobre los distintos segmentos abdominales y el recto. Imagen tomada de (Insausti T. C, 1994). Tr: tráquea, MsN: nervios metatoráxicos, AbN: Nervios Abdominales, Cx: Coxa, M: Músculo, R: Ramificaciones de los nervios abdominales. ThSp: Espiráculo toráxico. AbSp: Espiráculo Abdominal. AT: tejido Adiposo. PoG: Ganglio Posterior.

El cerebro está formado por el protocerebro, el tritocerebro, el deuterocerebro y los lobulos ópticos (Figura 4). El tritocerebro y el deuterocerebro son más pequeños que el protocerebro y se encuentran en posición ventral respecto a éste. A partir de cada uno de los lóbulos del deuterocerebro se generan los nervios que inervan las antenas (AN). En el del protocerebro se generan los nervios ocelares (OcN) que inervan los ocelos, así como un nervio tegumentario (TN) y dos nervios delgados que inervan la *Corpora Cardíaca* (NCC I, NCC II). El protocebro, junto con el complejo *Corpora Cardíaca- Corpora Allata* (CC-CA) son los principales órganos neuroendócrinos. Las neuronas productoras de hormonas en el protocerebro se localizan principalmente en las regiones llamadas "*pars intercerebralis*" (PI) y "*pars lateralis*" (PL). Es a partir de estas regiones desde donde se proyectan NCC I y NCC II respectivamente (Pipa, 1978; ver Veelaert *et al*, 1998). En las neuronas de la región PI se expresan hormonas diuréticas (DH31 y DH44) péptidos simil-insulina (ILP), FMRFamida, la Hormona dispersora de pigmentos (PDF), Corazonina (CZ), SIFamida, Proctolina (Prt) y Allatostatina A (ALS-A), mientras que en la región PL se ha encontrado expresión de FMRFamida, PDF, CZ y ALS-A (Siga, 2003; ver Nässel, 2002). El *Par Lateralis* es el responsable de la regulación de la producción de las hormonas

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

juveniles por la *Corpora Allata* (Stay *et al*, 1996; 2000*)*. Además de estas dos regiones del protocerebro, se han descripto neuronas con actividad secretoria en el tritocerebro y en los distintos ganglios que conforman el cordón nervioso ventral (Kim, 1998; ver Nässel, 2002).

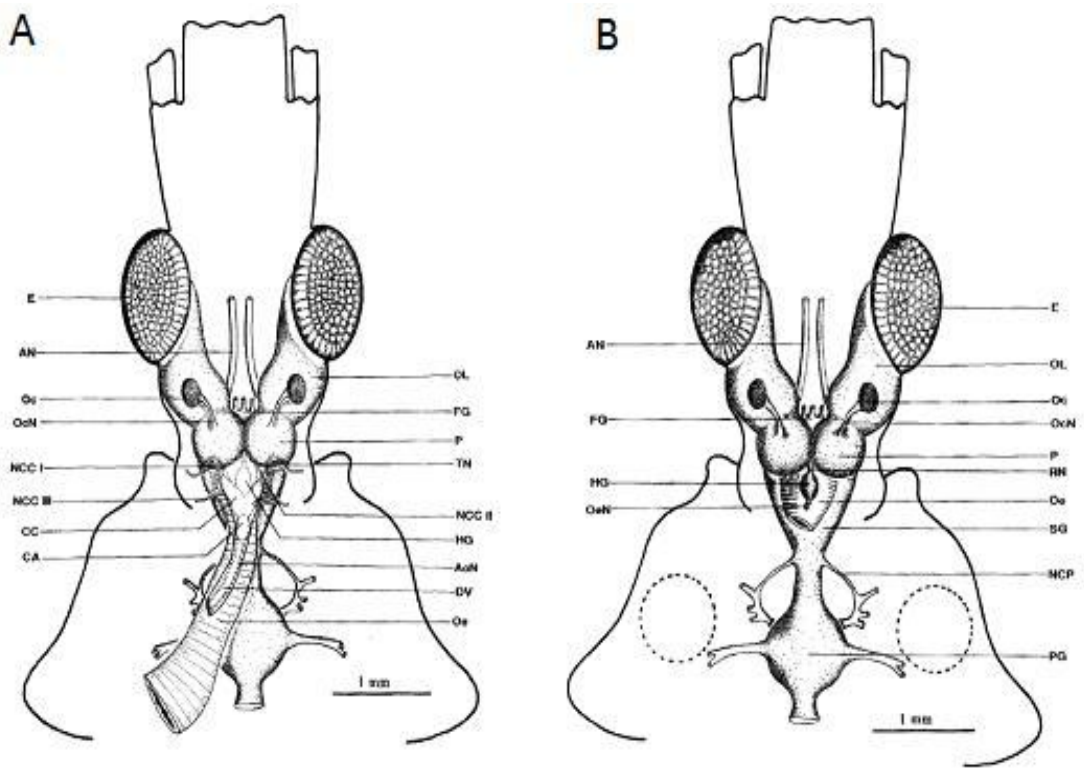


Figura 4: Vista dorsal (A) y ventral (B) del SNC. Imagen tomada de (Insausti T. C, 1994). E: Ojo. AN: nervio que inerva la antena, OL: lóbulo óptico, FG: ganglio frontal, Oc: ocelo, OcN: nervio que inerva el ocelo, P: Protocerebro, RN: Nervio Recurrente, HG: ganglio hipocerebral, OeN: Nervio esofágico, Oe: esófago, SG: Ganglio Subesofágico, PG: ganglio Prototaráxico.

El SNS consta del ganglio frontal (FG) el complejo retrocerebral, formado por la *Corpora cardíaca (CC)* y la *Corpora allata (CA),* y el ganglio hipocerebral (HCG). El FG se encuentra ubicado en posición anterior respecto al tritocerebro y se conecta con éste a través del nervio frontal. De la región antero dorsal emerge el nervio frontal medio (MFrN) que inerva la musculatura de la faringe. El nervio recurrente (RN) se extiende desde la región posterior del FG. De aquí también se extienden nervios que inervan las glándulas salivales. Por lo tanto estos órganos son inervados tanto por el SNC como por el SNS. La CA consta de un par de glándulas endócrinas fusionadas formando una estrucutra que se ubica ventralmente a la aorta, cuya principal función conocida es la producción de hormonas juveniles (JH). La CA se encuentra conectada con la CC a través del nervio NCA I. A partir del complejo CC-CA se liberan neuropéptidos a la hemolinfa, por lo que se

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

trata de un sitio de liberación neurohemal. El SNS tiene un papel central en la alimentación en distintos insectos. En particular el FG, el cual es responsable de la actividad motora del buche y el digestivo medio, y por lo tanto, del pasaje de alimento a través del sistema digestivo. Muchos de los neuropéptidos que muestran actividad sobre la musculatura del sistema digestivo *in vitro*, han sido detectados en el SNS de *Apis mellifera* (Boerjan *et al*, 2010).

4) Genómica y proteómica: Nuevas tecnologías en endocrinología de insectos.

Tradicionalmente, la identificación de péptidos bioactivos se realizaba mediante laboriosos procesos de purificación. Las muestras eran fraccionadas mediante técnicas cromatográficas, y posteriormente cada fracción era utilizada en un bioensayo con el fin de determinar la presencia de un compuesto activo. Con el advenimiento de las nuevas técnicas de "alto rendimiento" (genómica y proteómica) se hizo posible la identificación de nuevos neuropéptidos de manera más rápida y simple, a partir de una cantidad de tejido mucho menor. Esto ha llevado a una verdadera revolución en la neuroendocrinología de insectos en los últimos años, a partir de su primera utilización para la identificación de neuropéptidos en *Drosphila melanogaster* (Baggerman *et al,* 2002). Hoy es posible obtener información de secuencia y expresión de un alto número de moléculas en forma simultánea, acelerando los tiempos necesarios para la detección de nuevas moléculas. Si bien estas técnicas permiten una alta eficiencia en la identificación y descripción de neuropéptidos, presentan la desventaja de no asociar a las moléculas identificadas a una función determinada. Es por ello que, luego de una etapa descriptiva, resulta necesario llevar a cabo ensayos fisiológicos y de biología molecular para determinar la función de las mismas.

En el presente trabajo hemos realizado una caracterización exhaustiva de los neuropéptidos presentes en *R. prolixus* a nivel de secuencia y estructura, utilizando técnicas bioinformáticas, proteómicas y de biología molecular. Por otra parte, hemos aplicado técnicas de peptidómica cuantitativa para caracterizar la respuesta post-ingesta sanguínea, a fin de comenzar la caracterización fisiológica de los neuropéptidos identificados. Se trata del primer trabajo de este tipo realizado en un insecto vector de enfermedades.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Objetivos

a) Identificar, caracterizar y anotar en el genoma de *R. prolixus* los genes que codifican neuropéptidos (GPNs)
b) Realizar estudios de peptidómica cuantitativa en el SNC de *R. prolixus*, a fin de determinar variaciones en la concentración de neuropéptidos a distintos tiempos post ingesta sanguinea.
c) Determinar la expresión en SNC y demás tejidos de los GPNs identificados.
d) Sentar las bases para posteriores enyasos fisiológicos y moleculares a fin de determinar la/s función/es de los neuroéptidos caracterizados.

Capítulo 1

Identificación y caracterización de genes precursores de neuropéptidos en *R. prolixus*

1.1) Introducción

En el presente estudio hemos combinado racionalmente herramientas bioinformáticas, genómicas, proteómicas y de biología molecular con el objetivo de avanzar en la caracterización del sistema neuroendócrino de *R. prolixus* ¿Por qué *R. prolixus*? Estudios similares podrían haber sido llevados a cabo en *T. infestans*, el vector de Chagas más extendido en Argentina, sin embargo en el corto plazo hubiese resultado improbable acceder a la información genómica necesaria de esta especie. En cambio, gracias al esfuerzo de un consorcio internacional que nuestro laboratorio integra, el proceso de secuenciación del genoma de *R. prolixus* se encuentra en su fase final de ensamblado. Esto hace que se cuente con información genómica útil para llevar a cabo con éxito los estudios propuestos, que a su vez servirán de marco de referencia para avanzar en el conocimiento del neuropeptidoma de *T. infestans*. Por otro lado, *R. prolixus* ha sido utilizado como modelo de estudio del sistema neuroendócrino por casi 100 años. Ha sido el modelo elegido por Sir Vincent Wigglesworth para estudiar la regulación del crecimiento y el desarrollo (ver Edwards, 1994; 1998) y por S.H. Maddrell para el estudio de la regulación del balance hídrico-mineral (Maddrell, 1964; ver Beyenbach, 2003; ver Coast, 2009). Los insectos no alimentados permanecen en un estado de arresto del desarrollo, y solo son capaces de mudar al siguiente estado ninfal o adulto con posterioridad a la ingesta sanguínea. El proceso de alimentación actúa como un estímulo para el sistema neuroendócrino, desencadenando procesos como la diuresis, la muda y la reproducción. A pesar de su importancia sanitaria y de su extendida utilización como modelo para el estudio del sistema neuroendócrino, hasta la realización de nuestro trabajo sólo se habían descripto dos genes que codifican neuropéptidos en *R. prolixus* (dos genes parálogos del neuropéptido CAPA; Paluzzi, 2008; 2010).

En insectos, hasta la fecha se han secuenciado y publicado los genomas de *D. melanogaster* (2000), *Anopheles gambiae* (2002), *Aedes aegypti* (2007), *A. mellifera* (2006), *Tribolium castaneum* (2008) y *Bombyx mori* (2008). El estudio en estas especies

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

de los neuropéptidos, así como de sus receptores, ha aportado información crucial sobre su desarrollo, fisiología, comportamiento y relación evolutiva (Hewes and Taghert, 2001; Hummon *et al*, 2006; Riehle *et al*, 2002; Li *et al*, 2008). Nuestro trabajo es el primero de este tipo en llevarse a cabo en un insecto hemimetábolo. Dentro de este grupo se encuentran varias especies de interés sanitario, y otras que se comportan como plagas de cultivos, por lo que son causantes de grandes pérdidas económicas.

1.2) Metodología de trabajo

1.2.1) Estrategia experimental y flujo de trabajo

La estrategia experimental utilizada con el fin de caracterizar GPNs en el genoma de *R. prolixus* ha combinado herramientas bioinformáticas, proteómicas, transcriptómicas y técnicas de biología molecular (Figura 1.1)

En primer lugar, se realizó una búsqueda bioinformática en trazas y *contigs* de la secuencia genómica pre-ensamblada, comparándola con secuencias proteicas de neuropéptidos de *R. prolixus* identificadas previamente en cerebro por espectrometría de masas (MS) (Ons *et al*, 2009), o bien secuencias ortólogas conocidas en otros insectos. A partir de las regiones cortas detectadas en el genoma, se diseñaron *primers* específicos para cada gen. Los *primers* fueron utilizados para caracterizar el mRNA completo mediante RACE-PCR. Esta técnica permitió confirmar y/o corregir las predicciones bioinformáticas, y corroborar la transcripción de los genes en el SN. Por otro lado, hemos podido identificar distintas isoformas de algunos genes, generadas por *splicing* alternativo. También hemos analizado extractos del SN mediante MS, con el fin de confirmar la expresión de los péptidos predichos a partir del clivaje de los precursores. Esta técnica permite conocer además las modificaciones post-traduccionales (PTM) que presenta cada péptido maduro.

En el marco de secuenciación del genoma de *R. prolixus*, el Dr. Pedro Oliveira (Universidad Federal de Río de Janeiro, Brasil), ha obtenido bases de datos de EST a partir de mRNA extraído de diferentes órganos del insecto (recto, digestivo medio posterior, digestivo medio anterior y testículos) mediante secuenciación al azar con secuenciadores T454. Estas bases de datos fueron utilizadas para determinar la transcripción de los GPNs en los distintos órganos.

Una vez caracterizado el neuropeptidoma en SN, se realizó un *screening* de neuropéptidos mediante MS, a fin de determinar la expresión de los mismos en distintos tejidos. Los tejidos analizados fueron gándulas salivales, digestivo medio anterior, digestivo medio posterior, recto, túbulos de Malphigi (TM), cuerpo graso, testículos, ovarios y hemolinfa.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

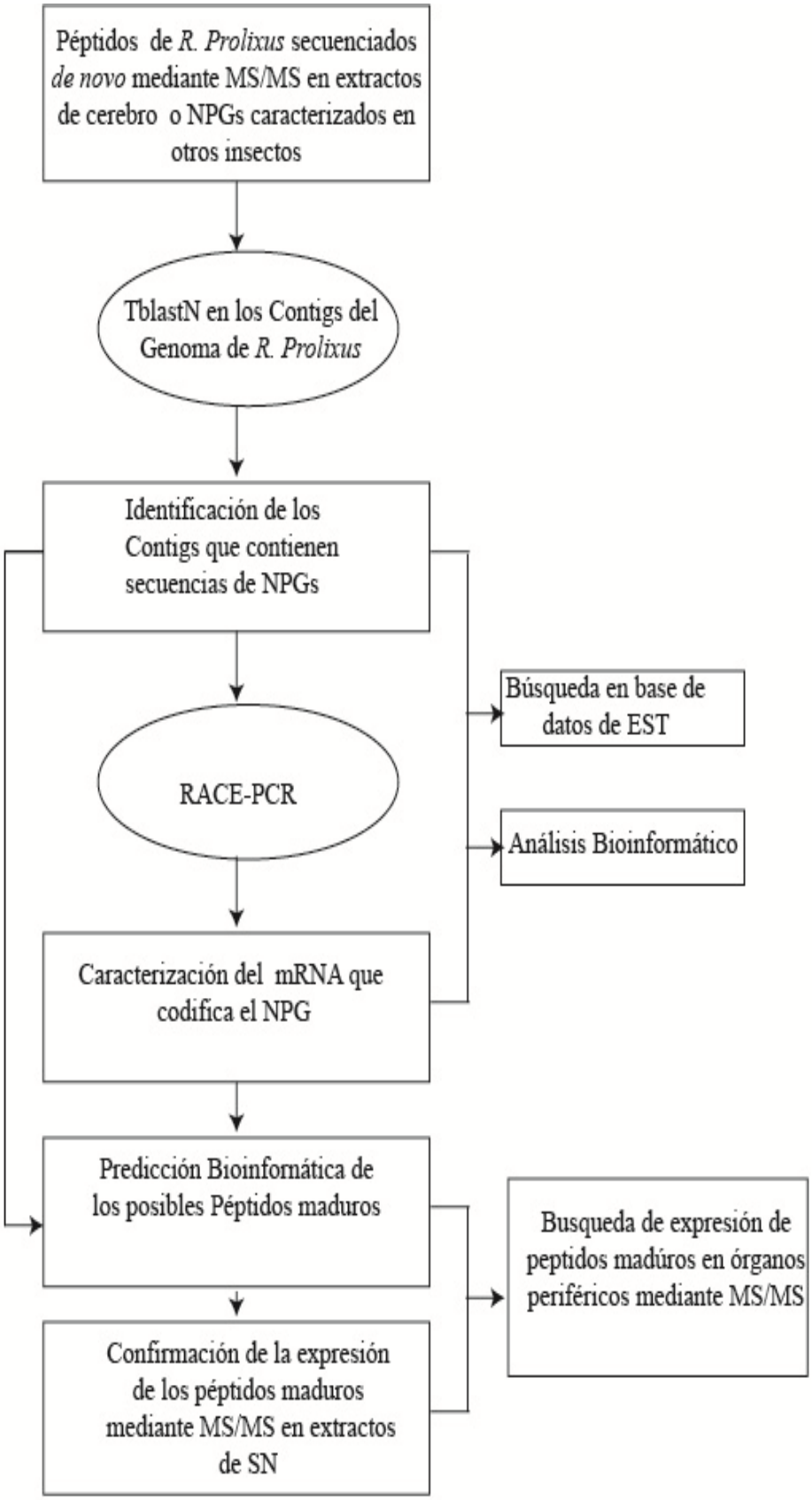


Figura 1.1: : Flujo de trabajo utilizado para la determinación de los GPNs y neuropéptidos maduros. MS/MS: espectrometría de masas en tándem. RACE-PCR: "*Rapid amplification of cDNA ends*"

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

1.2.2) Identificación y análisis bioinformático de genes

Las bases de datos del genoma de *R. prolixus* fueron utilizadas para la búsqueda bioinformática de GPNs por similitud de secuencia. Las búsquedas se realizaron mediante el algoritmo TBLASTN y la matriz BLOSUM62. Dicha base de datos ha sido elaborada en "Washinton University School of Medicine, St. Louis", y puede ser consultada en http://genome.wustl.edu/genomes/view/R.prolixus_prolixus/. El mismo procedimiento fue llevado a cabo para identificar los posibles receptores de neuropéptidos en el genoma.

Para las predicciones bioinformáticas de estructura de genes y marcos de lectura abiertos (ORF, según sus siglas en ingés). se utilizaron los programas FGENESH,(www.softberry.com); Genscan (http://genes.mit.edu/GENSCAN.html) y Augustus (http://augustus.gobics.de/). Los resultados obtenidos fueron corregidos manualmente.

La presencia de péptido señal fue corroborada mediante SignalP *server* (Bendtsen *et al*, 2004; http://www.cbs.dtu.dk/services/SignalP/). Para predecir los posibles sitios de clivaje proteolítico del precursor, y posibles PTM de los péptidos maduros, se utilizó un algoritmo estadístico (http://www.neuroproteomics.scs.uiuc.edu/neuropred.html).

Los alineamientos de secuencias se realizaron con el programa ClustalX2. El contenido informativo de cada posición se graficó mediante Sequence logo (http://weblogo.berkeley.edu/logo.cgi).

1.2.3) Confirmación de secuencia y transcripción en Sistema Nervioso

Insectos

R. prolixus adultos fueron obtenidos de una colonia establecida en el Centro Regional de Estudios Genómicos (CREG-UNLP), mantenida a una temperatura de 28 +/- 2 ºC y 45% de humedad relativa, con un ciclo de 12:12 horas de luz: oscuridad. Para extraer el RNA se utilizaron 40 individuos de ambos sexos con más de 14 días post-ecdisis, 20 de los cuales fueron alimentados con sangre de gallina 24 hs antes de la toma de muestra, mientras que los restantes 20 no fueron alimentados durante el estado adulto.

Disección del Sistema Nervioso

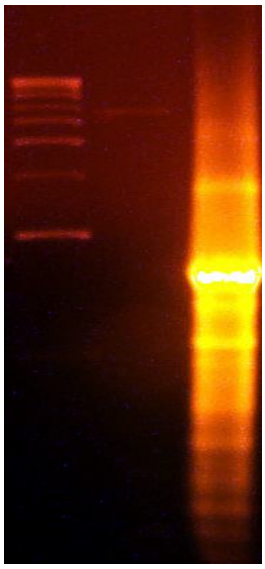
Se extrajo el SNC y SNS de los insectos mediante microdisección con pinzas quirúrgicas bajo lupa, en solución salina de *R. prolixus*. (NaCl, 129 mM; KCl, 8.6 mM; $CaCl_2$ 2.0 mM; $MgCl_2$ 8.5 mM; $NaHCO_3$ 10.2 mM; NaH_2PO_4 4.3 mM; Hepes 8.6 mM, PH 7).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Síntesis de cDNA.

Se extrajo RNA total a partir de SN utilizando el método descripto por Chomczynski y Sacchi (1987), con la diferencia de que todos los pasos se realizaron a PH neutro en lugar de ácido, y finalmente el RNA se precipitó con LiCl (Figura 1.2).

Figura 1.2: Gel de agarosa 1% tenido con bromuro de etidio, donde se muestra el RNA total extraído del SN. Pueden observarse bandas de mayor intensidad que corresponden a los RNA ribosomales y de transferencia.



La síntesis de cDNA se realizó utilizando el *kit* "GENERACER KIT" (INVITROGEN), siguiendo las instrucciones del fabricante hasta la obtención de la primera cadena. Luego se modificó el protocolo para obtener cDNA doble cadena (dscDNA; Figura 1.3).

Para sintetizar dscDNA se utilizó la sigueinte mezcla de PCR:

	Concentración final
Buffer 1 *long Span*	**1x**
dNTPs	**0.35 mM**
5´Generacer primer	**0.3 uM**
3´Generacer primer	**0.3 uM**
Long span	**2.5 U**
1ª cadena cDNA	**2 µl**
H_2O	**Hasta alcanzar volumen**
Volumen final	**20 µl**

El programa de PCR utilizado fue el siguiente:

Temperatura (ºC)	Tiempo (min.)	Número de ciclos
68 ºC	10	1 ciclo
95 ºC	1	
95 ºC	0.5	5 ciclos
68 ºC	9	

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

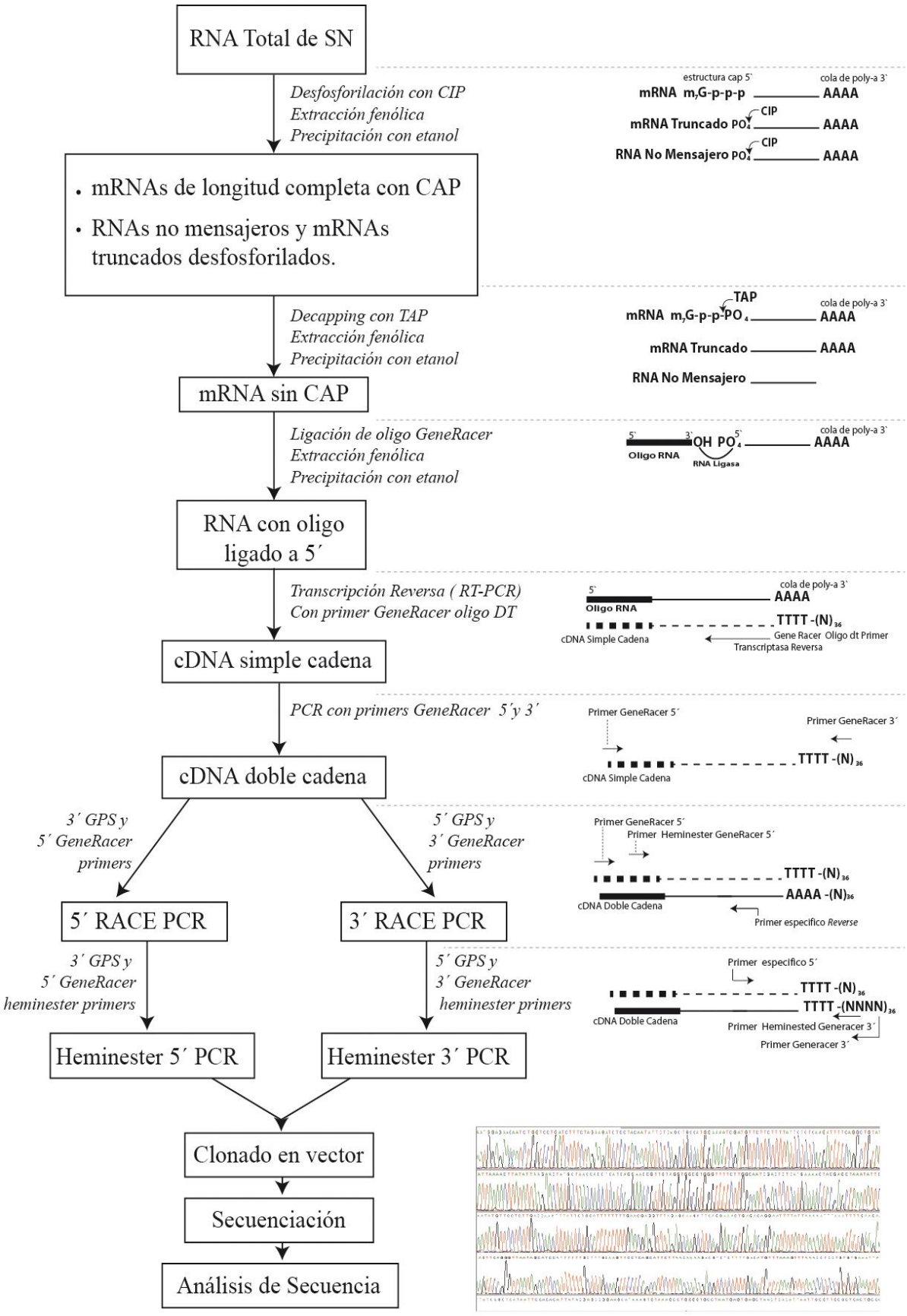


Figura1.3 Estrategia utilizada para caracterizar los mRNA de interés mediante RACE-PCR. CIP: *Calf intestinal phosphatase* , TAP: Tabacco Acid Pyrophosphatasa. Imagen modificada a partir del manual del kit Generacer (Invitrogen)

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

RACE-PCR

Con el fin de obtener la secuencia completa de cada uno de los mensajeros de interés, se realizaron RACE-PCRs en los extremos 5´y 3´. Para las RACE 5´se utilizaron los primers "*reverse*" de cada gen y el "*GeneRacer 5´primer*". Para obtener el extremo 3´ se utilizó el primer "*foward*" diseñado para cada gen y "*GeneRacer 3´primer*" (Tabla 1.1).

Nombre del GPN	**Primer *Forward***	**Primer *Reverse***
Hormona Adipokinética (AKH)	TTGGTACTGCTGACATTTCAT	AGTATCAGGATTTGGTGTACA
Allatostatina A (ATS-A)	TCCCATTTGCAGACCAGC	AATTGTGTAAGTTATTGGCTGG
Allatostatina B (ATS-B)	CTCTTGGGGGAAACGT	GCCCCATAATCCTTTCAAGT
Allatotropina (AT)	TTGGTGTTGGTGGCCTTGG	CTGGATAGGTTCTCTTGCCG
Pro-hormona 1 (PH1)	GGTTCAATTGACCGAGCTGT	GCAGGAGACCGCATTGAA
Hormona simil-Calcitonina (DH31 A y B)	AGCCTCCGAGCCCTACTAACT	CGCCTCCTTCTGCCTGGG
CCHamida (CCHa)	TGTTCAGCATTTGGGCATTCAT	ATCGTTTGCCATGGCCTCCAA
Péptido Cardioactivo de Crustaceos (CCAP)	GGATGAATCTATGGCCACATT	AGGGGAATTCGGTCTGACT
Corazonina (CZ)	TTCGATCAAGTTGTCTACTCATC	GAACATACCTGTGGTATAGTC
Hormona de Eclosión (EH)	GGAACTGTGCCCAGTGTAAGA	ATGGTCCAATTGAGGCAATG
Péptido simil-Insulina 2 (ILP-2)	GCCATCAAAAGGGGAGTACA	TATTTCTCCAAGTCGCCGAA
Péptido transportador de iones (ITP)	ATGCATCAGGAGCGGAGAGC	ACGTGCCAAGTGAATGTAGC
Neuropeptido F Largo (LNF)	CAATGGCACGTCCAGCTA	CTGCCGGCAACAGCATA
Miosupresina (MS)	CAGGACCAGATTGTTCTCCAT	ACGTCGACGTCCAAAGCG
Precursor de Neuropéptidos 1 (NPLP-1)	TCTCGCAAGCGGGGAGAATA	ATTTCTTCGCCTTCTTCATCA
Orcokinina (OK)	CTCAGCGTATGCCGCAGAC	AATGATGTGGCTCACTTT
Pirokinina (PK)	CGTTTAGGACGGGATGAAGA	CTTGGCCTGAAGACGATACG
Neuropéptido F Corto(SNF)	CAGATAGTGCTACTGATGGC	CTTCATAGCTGTGGCCCTCATGT
SIFa	ATGTCTCGCACTCTGTTCGT	TTAATTGTTGTCTTGAACGGG
Sulfakinina (SK)	TGTGTTTGGCTGAGCCAGC	CTAATGGTCTGGACCTTCC
Taquikinina (TK)	CCTGTGGGAAGTCTGCTGG	ATGGTTCAAGTCATCAGGGC

Tabla 1.1: Primers utilizados en las reacciones de RACE-PCR.

La mezcla de RACE-PCR utilizada fue la siguiente:

	Concentración Final
Buffer	**1 X**
$MgCl_2$	**2 mM**
dNTPs	**0.2 mM**
***Primer Fw* o *Rv* específico**	**0.2 µM**
GeneRacer 5´o 3´ *primer*	**0.6 µM**
Taq polimerasa	**1 U**
cDNA doble cadena	**1 µl**
H_2O	**Hasta alcanzar volumen**
Volumen final	**50 µl**

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Programa de ciclado

Temperatura (ºC)	Tiempo (min.)	Número de ciclos
94 ºC	5 min.	1 ciclo
94 ºC	0.75 min	35 ciclos
53 ºC	0.75 min.	
74 ºC	2 min.	
74 ºC	10 min.	1 ciclo

Con el fin de aumentar la especificidad, se realizó una segunda reacción de PCR (*heminested* PCR), utilizando como molde 1 µl del producto de la primer reacción. Para esto se utilizó el *primer forward.* de cada gen y el "*Generacer 3´ nested primer*" para los extremos 3´, y el *primer reverse.* de cada gen con el "*Generacer 5´ nested primer*" para obtener los extremos 5´ de cada gen. El programa de ciclado fue el mismo que para la primer reacción. La mezcla de PCR utilizada fue la siguiente:

	Concentración Final
Buffer	**1 X**
$MgCl_2$	**2 mM**
dNTPs	**0.2 mM**
***Primer Fw* o *Rv* específico**	**0.2 µM**
***GeneRacer nested primer* 5´o 3´**	**0.2 µM**
Taq polimerasa	**1 U**
PCR 1 (molde)	**1 µl**
H_2O	**Hasta alcanzar volumen**
Volumen final	**50 µl**

Los productos de PCR fueron visualizados en un gel 1% de agarosa conteniendo 5 µg/ml de bromuro de etidio. Las bandas de interés obtenidas en la *heminested* PCR fueron eluídas del gel utilizando el un *kit* comercial (QUIAGEN). Los productos de PCR fueron clonados en los vectores Clonjet (Fermentas), Topo Clonning (Invitrogen) o pGEM-T (Promega), según el caso. Los vectores portando el inserto fueron incorporados en bacterias *E. coli* XL1-Blue competentes, amplificados y purificados por *miniprep*. Se verificó la presencia de inserto mediante digestión con enzimas de restricción. Para verificar que el inserto clonado correspondía al gen de interés se realizaron PCRs con *primers forward* y *reverse* específicos para cada gen. Los insertos de interés fueron secuenciados en el servicio de secuenciación Macrogen (Seúl, Corea del Sur).

Una vez obtenidas las secuencias 3´ y 5´ de cada mRNA, estas fueron ensambladas, comparadas con el genoma, analizadas y corregidas manualmente.

1.2.4) Búsqueda de transcriptos de GPNs en tejidos

Mediante la secuenciación al azar con secuenciadores T454, el Dr. Pedro Oliveira (Universidad Federal de Río de Janeiro, Brazil, resultados no publicados) ha obtenido información del transcriptoma del recto, digestivo medio posterior, digestivo medio anterior

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

y testículos de *R. prolixus*. Una vez identificados los GPNs, estas bases de datos fueron utilizadas para determinar su posible transcripción en cada órgano. Las búsquedas bioinformáticas se realizaron mediante el algoritmo BLASTN y la matriz BLOSUM62.

1.2.5) Confirmación de expresión de GPN y determinación de modificaciones post-traduccionales mediante MS/MS

Con el fin de obtener un panorama completo de los péptidos expresados en el sistema neuroendócrino del insecto, se analizaron extractos de las siguientes estructuras: SN, gándulas salivales, digestivo medio anterior, digestivo medio posterior, recto, TM, cuerpo graso, testículos, ovarios y hemolinfa mediante *on line* nanoLC-ESI-Quadrupole-TOF MS, Nano-LC-MALDI-TOF/TOF MS y Nano-LC-ESI-Orbi-Trap MS.

Insectos y disección de los tejidos

Se realizaron microdisecciones de los órganos mencionados a partir de insectos adultos, en las mismas condiciones que las descriptas más arriba. Las muestras fueron congeladas a -80 °C y transportadas en hielo seco para su análisis en el Laboratorio de Espectrometría de Masas Bioanalítica, Instituto Max Plank de Biofísica Química. Göettingen, Alemania.

Extracción de péptidos

Cada muestra conteniendo un tipo de tejido de 10 individuos fue colocada en 100 µl de metanol/agua/ácido acético (90/9/1 v/v/v), sonicada por 5 min. y centrifugada a 7500g por 10 min. El sobrenadante fue recuperado y el precipitado fue re extraído dos veces más. Finalmente el solvente fue evaporado en vacío (*speedvacuum*) y los péptidos se resuspendieron en 20 µl de 0,1% ácido trifluoroacético (TFA)

Desalado de las muestras

Las muestras fueron desaladas usando discos de extracción C18 (Varian, Darmstadt, Germany), de acuerdo al protocolo descripto por Rappsilber *et al*, 2007, que permite concentrar y purificar parcialmente los péptidos de bajo peso molecular. Los discos de extracción C18 fueron activados con 80% acetonitrilo (ACN)/0,1% TFA y lavados con 0,1% TFA. Posteriormente se cargó la muestra, la que fue lavada 3 veces con 20 µl de 0,1% TFA. Finalmente los péptidos fueron eluídos con 70%ACN/0.1%TFA. El solvente orgánico fue evaporado en vacío y las muestras fueron re-diluídas en 12 µl 5% CH_3CN/1% ácido fórmico (FA) para ser analizadas mediante MS/MS.

Nano-LC-MALDI-TOF/TOF MS.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Se utilizó un sistema Dionex nano-LC (LC Packings). Las muestras se cargaron en la pre-columna durante 10 min con *buffer* A (0.1% v/v TFA). Los péptidos fueron eluídos en un gradiente lineal de 10-60% de *buffer* B (80% ACN, 0,1 % TFA) en 180 min. El gradiente fue finalmente elevado de 60 a 100% de solvente B en 10 min. A medida que eluían del sistema, las fracciones fueron mezcladas con ácido α-cyano-4-hydroxicinamico como matriz de ionización de MALDI, y sembradas en una placa LC-MALDI (Applied Biosystems/MDS Sciex), mediante un colector de microfracciones (LC Packing). Posteriormente las muestras fueron analizadas mediante MS tipo MALDI-TOF/TOF con un equipo ABI 4800 (Applied Biosystems). Se analizaron los espectros de masas entre 600 y 4000 m/*z*. La energía de colisión usada fue de 1.10-6 Torr. Las listas de picos generadas fueron utilizadas para la búsqueda en la base de datos usando los programas MASCOT 2.1.0 (Matrix Science,) y PEAKS STUDIO 5.2 (Bioinformatics Solutions,) con los siguientes parámetros: enzima: ninguna; modificaciones variables: Pyro-glutamico de Glutamina, amidación N-terminal; acetilación C-terminal, oxidación de metionina; Tolerancia de error de masa del ion parental: 200 mu; Tolerancia de error de masa del fragmento: 0.5 Da. Los resultados fueron validados manualmente. La base de datos utilizada para las búsquedas contenía todas las secuencias de *R. prolixus* presentes en *Genbank*, incluyendo los neuropéptidos maduros y precursores previamente identificados.

NanoLC- ESI-Quadrupole-TOF MS

Una fracción de las muestras fue analizada mediante *on line* nanoLC-ESI-Quadrupole-TOF MS (Thermo Fisher Scientific, USA). Se analizaron 4 mustras de cada órgano. Las muestras fueron cargadas en la pre-columna del HPLC a un flujo de 10 µl/min. *buffer* A (0,1% v/v TFA). Los péptidos fueron eluídos en un gradiente lineal de 10-60% v/v *de buffer* B (80% v/v ACN, 0.1% v/v TFA en agua) durante 38 minutos. Finalmente los péptidos restantes fueron eluídos con 100% *buffer* B durante 7 min. Las listas de picos fueron utilizadas para la búsqueda en la base de datos de todos los genes de *R.prolixus* registrados en Genebank (incluídos los GPNs). Para ello se utilizaron MASCOT 2.1.0 (Matrix Science,) y PEAKS STUDIO 5.2 (Bioinformatics Solutions) con los siguientes parámetros: enzima: ninguna; modificaciones variables: Pyro-glutamico de Glutamina, amidación N-terminal; acetilación C-terminal, oxidación de metionina; Tolerancia de error de masa del ion parental: 100 ppm; Tolerancia de error de masa del fragmento: 0.25 Da.

NanoLC-ESI-Orby-Trap MS

Muestras correspondientes a diferentes órganos fueron analizadas mediante un espectrómetro de masas *on line* nanoLC-(ESI)-Orby-Trap (Thermo Fisher Scientific). El protocolo de HPLC utilizado fue el mismo que el detallado en el punto anterior.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Los espectros fueron procesados y las listas de picos fueron utilizadas para la búsqueda de la base de datos usando los programas MASCOT 2.1.0 (Matrix Science) y PEAKS STUDIO 5.2 (Bioinformatics Solutions) con los siguientes parámetros: enzima: ninguna; modificaciones variables: Pyro-glutamico de Glutamina, amidación N-terminal; acetilación C-terminal, oxidación de metionina; Toleracia de error de masa del ion parental: 10 ppm; Tolerancia de error de masa del fragmento: 0.5 Da.

1.2.6) Inmunohistoquímica

La cutícula dorsal de *R. prolixus* de quinto estadío fue removida en solución salina con el fin de exponer el SNC. El tejido fue fijado *in situ* con paraformaldehído 2% (ph 7) por 12 hs. a 4°C. Posteriormente, el SNC fue disectado, lavado con solución salina y transferído a 4% Tritón X-100 con 2% BSA y 10% de suero normal de cabra (NGS) en *buffer* fosfato PH 7,4 (PBS) durante 1 h. a temperatura ambiente. El antisuero primario desarrollado contra el péptido en estudio fue diluído 1:2000 en 0.4% tritón X-100 en PBS con 2% BSA y 2% NGS, e incubado a 4°C por 12 hs. antes de ser usado. El tejido fue incubado con el antisuero primario por 48 hs. y posteriormente lavado durante 12 hs. con PBS. Finalmente, el tejido fue incubado por 12 hs con el anticuerpo secundario conjugado con el fluoróforo Cy3 diluido 1:500 con 10% NGS en PBS . Después de varios lavados con PBS a temperatura ambiente los SN fueron montados en un portaobjetos con Mowiol . Las imágenes fueron obtenidas con un microscopio confocal de barrido laser *LSM-510-META* (Carl Zeiss, Germany) y procesadas mediante el software *Zeiss LSM Image Browser*. En los grupos control de cada experimento se omitió el uso de antisuero primario.

El anticuerpo primario anti Orcokinina (OK) fue cedido gentilmente por el Dr. Heinrich Dircksen (Universidad de Estocolmo, Suecia). El anticuerpo primario anti SIFamida fue cedido gentilmente por el Dr. Jan A. Veenstra (Universidad de Bordeaux, Francia).

1.3) Resultados y Discusión

1.3.1) El neuropeptidoma de *R. prolixus*

Mediante técnicas bioinformáticas hemos podido identificar *in silico* 32 GPN en el genoma de *R. prolixus*. Por medio de técnicas transcriptómicas y/o proteómicas hemos confirmado la expresión de 28 de estos genes en SN (Tabla 1.2), así como formas de *splicing* alternativas para 4 de ellos. Teniendo en cuenta los dos únicos GPN caracterizados con anterioridad a nuestro trabajo (dos variantes del gen CAPA; Paluzzi *et al* 2008; 2010), los GPN identificados en *R. prolixus*, suman 34. Este número es similar al descripto en especies de insectos cuyos genomas se encuentran totalmente secuenciados y ensamblados: 40 en *B. mori (*Roller *et al*, 2008*),* 41 en *T. castaneum (*Li *et al*, 2008*),* 36

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

en *D. melanogaster (*Riehle *et al,* 2002)*,* 40 en *A. gambie* (Hewes and Taghert, 2001) y 34 en *A. mellifera* (Hummon *et al*, 2006) (Tabla 1.3).

Mediante MS hemos podido confirmar la expresión en el SN y verificar PTM de 107 péptidos codificados por los GPN presentes en el genoma. Resulta interesante observar que varios de los péptidos identificados presentan particularidades a nivel de secuencia, no descriptas anteriormente en insectos. La identificación de estas secuencias particulares, distintas a las conocidas para el resto de las especies, aporta las bases para futuros estudios farmacológicos orientados al diseño de una nueva generación de mecanismos de control que sean más específicos para la especie a combatir.

En algunos casos, las técnicas de MS revelaron la presencia de formas truncadas de neuropéptidos maduros. Dado que estos péptidos truncados se han detectado utilizando distintos espectrómetros de masas y a partir de diferentes muestras biológicas es poco probable que se trate de una fragmentación de los péptidos ocurrida durante el proceso de ionización. No puede descartarse que la fragmentación suceda durante el proceso de extracción, sin embargo se trata de una observación frecuente en los trabajos peptidómicos en insectos, en los que se han utilizado distintos métodos de extracción (Hummon *et al*, 2006, Neupert *et al*, 2010; Li *et al*, 2008, Boerjan *et al*, 2010). Por otro lado, los péptidos truncados que se observan en las diferentes muestras son siempre los mismos. Esto sugiere que las formas truncadas ocurren naturalmente en el SN de insectos, y probablemente sean capaces de llevar a cabo una función biológica , aunque hasta ahora no se han realizado estudios para corroborar esta hipótesis. Las formas truncadas podrían ser el producto de la metabolización parcial del péptido maduro, un proceso regulado en el que participan varias enzimas, como se ha demostrado para la degradación de Corazonina (Vandersmissen, 2006)

Además de los neuropéptidos maduros y variantes truncadas, se han detectado varios péptidos “espaciadores” o “acompañantes”. Se conoce con este nombre a los péptidos que se generan a partir del clivaje de un precursor, pero no presentan el dominio característico de una determinada familia, conservado entre distintas especies. Algunos autores sugieren que únicamente los péptidos con el dominio característico son capaces de interaccionar con el/los receptor/es, mientras que los péptidos espaciadores no tendrían función en la señalización celular, o llevarían a cabo una función diferente (Boerjan *et al*, 2010).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Tabla 1.2: Neuropéptidos en *R. prolixus*. Los signos + y – indican detección por RACE-PCR/EST. Los números indican la cantidad de péptidos maduros identificados para un determinado precursor mediante MS. En rojo se muestran los GPN detectados unicamente *in silico*

Nombre GPN *In silico*	**GeneBank AN**	**mRNA**[a] (RACE/EST)	**MS**[b]	**Secuencia con mayor homología en Swissprot**
Allatostatina A (ATS-A)	GQ856315	+	8	Allatostatin *Diploptera punctata* (E: 4 e-26).
Allatostatina B (ATS-B)	ID: 1273333	+	9	Prothoracicostatic hormone *Bombyx mori* (E: 5 e-29).
Allatotropina (AT)	GQ162783	+	1	Allatotropin precursor *Manduca sexta* (E: 4 e-09).
CCHamida	ID: 1273254	+	1	No significant match[b]
Corazonina (CZ)	ID: 1273265	+	1	Corazonin precursor *Bombix mori* (E: 6 e-10).
FMRFamida	ID: 1273272	-	4	FMRFa- related peptides. *Drosophila virilis.* (E: 2 e-12).
Factor Dispersor de Pigmentos (PDF)		-	-	Pigment-dispersing factor.*Romalea microptera* (E: 2e-07)
Hormona Adipokinetica (AKH)	GQ888667	+	-	Hypertrehalosaemic peptide *Blaberus discoidalis* (E: 3e-05).
Hormona Diurética simil-Calcitonina A (DH31-A)	GQ856316	+	1	CT-like DH *Drosophila melanogaster* (E: 4 e-12).
Hormona Diurética simil-Calcitonina B (DH31-B)	GQ856317	+	1	CT-like DH *D. melanogaster* (E: 4 e-19).
Hormona Diurética simil-Calcitonina C (DH31-C)	ID: 1273251	+	-	No significant match
Hormona Diurética simil-CRF (DH44)		-	1	DH44 *Bombyx mori (E:4e-07)*
Hormona Disparadora de la Ecdisis (ETH)	BK006926	+	-	No significant match[d]
Hormona de Eclosión (EH)	ID: 1273270	-	1	Eclosion hormone *Manduca sexta* (E: 3e-17).
kinina	ID: 1273279	-	6	No significant match[a]
Péptido Cardioactivo de Crustaceos isof. A (CCAP)	GQ888668	+	1	Cardioactive peptide *Periplaneta americana* (E: 4e-28).
Péptido Cardioactivo de Crustaceos isof. B (CCAP)	HQ264139	+	1	Cardioactive peptide *Periplaneta americana* (E: 4e-28).
Miosupresina (MS)	GQ344501	+	4	Miosupressin *Apis mellifera* (E: 2 e-09).
Neuroparsina A (NPA)	ID: 1273288	+	-	Neuroparsin A *Locusta migratoria* (E: 2e-08).
Neuropeptido F Corto (SNF)	ID: 1273338	+	4	Short Neuropeptide F *A. aegypti* (E: 6 e-08).
Neuropeptido F Largo (LNF)	ID: 1273282	+	4	Neuropeptide Y *Aplysia californica* (E: 4e-08).
Orcokinina Isof. A (OK-A)	ID: 1273300	+	7	Orcokinin precursor *Procambarus clarkii* (E: 4 e-14).
Orcokinina Isof. B (OK-B)		+	-	No significant match
Peptido simil-AKH/CZ (ACP)		-	-	No significant match [c]
Péptido simil-insulina-1 (ILP-1)		-	-	Bombyxin E1 *Bombyx mori* (E:5e-04)
Péptido simil-insulina-2 (ILP-2)	GU230850	+	-	Insµlin-like growth factor 2 *Mus muscµlus* (E:1e-07)
Péptido simil-insulina-3 (ILP-3)		-	-	Insµlin-like peptide 3 *Locusta migratoria* (E:4e-04)
Péptido simil-insulina-6 (ILP-6)		-	-	Insµlin-like peptide 6 *Drosophila melanogaster* (E:1.5)
Péptido Transportador de Iones Isof. A (ITP-A)	ID: 1273339	+	-	Ion transport peptide *Schistocerca gregaria* (E: 1e-35).
Péptido Transportador de Iones Isof. B (ITP-B)	ID: 1273340	+	-	Ion transport peptide *Schistocerca gregaria* (E: 4e-42).
Pirokinina (PK)/PBAN	ID: 1273305	-	1	PBAN neuropeptides *Aedes aegypti* (E: 0.006).
Precursor de Neuropéptidos 1 (NPLP1)	ID: 1273296	+	13	Neuropeptide like precursor 1 *D. melanogaster* (E:0.058)
Prohormona 1 (PH1)	ID: 1273236	+	4	Prohormone 1 *Apis mellifera* (E: 2e-21).
Prohormona 2 (PH2)	BK006932	-	3	Prohormone 2 *Apis mellifera* (E: 0.004).
Prohormona 3 (PH3)		-	1	Prohormone 3 *Apis mellifera* (E: 4e-10).
SIFamida	GQ253922	+	4	SIFa *Procambarus clarkii* (E: 4 e-12).
Sulfakinina (SK)	GQ162784	+	5	Sµlfakinins. *Calliphora vomitoria* (E: 1 e-07).
Tachikinina (TK)	GQ162785	+	20	Tachikinins precursor *A. mellifera* (E: 4 e-11).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

[a] Presenta similitud con LK de *Bombyx mori* (E:1.6 in non-redundant protein database)
[b] Presenta similitud con CCH-amide de *Bombyx mori* (E:0.014 in non-redundant protein database)
[c] Presenta similitud con ACP preprohomone Tribolium castaneum (E:7e-05 in non-redundant protein database)
[d] Presenta similitud con ETH de *Aedes aegypti* (E:1.4 in non-redundant protein database)
AN: Número de acceso asignado en Genbank

.**Tabla 1.3:** GPNs encontrados en los insectos cuyo genoma ha sido secuenciado (+) o que se encuentran ausentes (-). En los casos donde se han descripto formas alternativas de splicing estas se indican con la cantidad de (+) correspondiente

Nombre	***R. prolixus*** **(Hemiptera)**	*T. castaneum* (Coleoptera)	*D. melanogaster* (Diptera)	*A. gambie* (Diptera)	*A. mellifera* (Hymenoptera)	*B. mori* (Lepidoptera)
Hormona Adipokinetica I	+	+	+	+	+	+
Hormona Adipokinetica II	-	+	-	-	+	+
Peptido simil-AKH/CZ	+	+	-	+	-	+
Allatostatinas A	+	-	+	+	+	+
Allatostatinas B	+	+	+	+	-	+
Allatostatinas C (ALS-C)	-	+	+	+	+	+
Allatotropina	+	+	-	+	-	+
Factor Antidiurético b-1 (ADF-1)	-	+	-	-	-	-
Factor Antidiurético b-2 (ADF-2)	-	+	-	-	-	-
Factor Antidiurético b-3 (ADF-3)	-	+	-	-	-	-
Factor Antidiurético b-4 (ADF-4)	-	+	-	-	-	-
Factor Antidiurético b-5 (ADF-5)	-	+	-	-	-	-
Factor Dispersor de Pigmentos	+	-	+	+	+	+
Hormona Diurética simil-Calcitonina	+++	+	-	+	+	+
CAPA I	+ [a]	+	+	+	+	++
CAPA II	+ [b]	-	-	-	-	-
CCHamide	+	+	+	+	+	+
Corazonina	+	-	+	+	+	+
Péptido Cardioactivo de Crustaceos	++	+	+	+	+	+
Hormona Diurética simil-CRF	+	++	+	+	+	+++
Hormona Disparadora de la Ecdisis	+	+	+	+	+	+
Hormona de Eclosión	+	+	+	+	+	+
FMRFamide	+	+	+	+	+	+
kininas	+	-	+	+	-	+
Orcokinin	++	-	+	+	+	+
Péptido simil-Insulina-1	+	+	+	+	+	+
Péptido simil-Insulina-2	+	+	+	+	+	+
Péptido simil-Insulina-3	+	+	+	+	-	+
Péptido simil-Insulina-4	-	+	+	+	-	+
Péptido simil-Insulina-5	-	-	+	+	+	+
Péptido simil-Insulina-6	+	-	+	+	+	+
IMFa	-	-	-	-	-	+

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Péptido trasnportador de Iones	++	+	+	+	++	+
Precursor de Neuropéptidos 1 (NPLP1)	+	+	+++	+	+	+
Precursor de Neuropéptidos 2 (NPLP2)	-	-	+	-	+	-
Neuropeptido F Largo 1	+	-	+	+	-	++
Neuropeptido F Largo 2	-	-	-	-	+	+
Miosupresina	+	+	+	+	+	+
Neuroparsina A	+	+	-	+	+	+
Prohormona 1	+	+	+	+	+	+
Prohormona 2	+	+	-	-	+	-
Prohormona 3	+	+	+	+	+	+
Péptido simil-arginina-vasopressina (AVPL)	-	+	-	-	-	-
Pyrokinina	+	+	+	+	+	+
Hormone Prothoracicotropica	-	+	+	+	-	+
Proctolina (PR)	-	+	+	+	-	-
Neuropeptide F Corto	+	+	+	+	+	+
SIFamide	+	+	+	+	+	+
Sulfakininas	+	+	+	+	+	+
Tachikininas	+	+	+	+	+	+
Hormonas (proteínas)						
Bursicon	+	+	+	+	+	+
Hormone Glycoproteica alpha-2 (GPH-2)	+	+	+	+	-	+
Hormone Glycoproteica beta-5 (GPH-5)	-	+	+	+	-	+
Total GPNs en el genoma	**36**	**41**	**36**	**40**	**36**	**40**

[a] (Paluzzi *et al*, 2008)

[b] (Paluzzi *et al*, 2010)

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

1.3.2) Análisis comparativo de los neuropéptidos de *R. prolixus*

1.3.2a) Superfamilia de péptidos RF-amida

El primer neuropéptido perteneciente a la superfamilia de los péptidos con el motivo -RFamida (FaRPs) en su extremo carboxilo terminal fue caracterizado en 1977 en *Macrocallista nimbosa* (Price y Greenburg, 1977). Posteriormente se identificaron distintos péptidos con este dominio, todos los cuales presentan actividad miotrópica (Nachman *et al,* 1986; Spittaels *et al*, 1996; Tsang and Orchard, 1991). En esta superfamilia se incluyen los precursores FMRF-amida, Sulfakininas (SK), Miosupresina (MPS), Neuropéptido F largo (LNF) y Neuropéptido F corto (SNF). Todos ellos fueron detectados en el SN de *R. prolixus* en nuestro trabajo.

FMRFamida

En *R. prolixus,* así como en otras especies de insectos pertenecientes a distintos Órdenes, se ha detectado inmunomarcación positiva para estos péptidos FMRFamida (Tsang y Orchard, 1991; Settembrini *et al*, 2005; Matsushima *et al*, 2007). En *R. prolixus* se ha observado inmunomarcación distribuída por todo el SNS, tanto en el ganglio frontal como en el el complejo retrocerebral. También en terminaciones nerviosas que se extienden desde el ganglio hipocerebral hasta la superficie del buche. Péptidos FMRFamida han demostrado ser capaces de aumentar la frecuencia de contracción de este órgano (Tsang y Orchard, 1991). También se ha encontrado marcación positiva en las terminaciones nerviosas que inervan las glándulas salivales (Tsang y Orchard, 1991), por lo que se ha asociado a estos péptidos con la regulación de la secreción de saliva. Posteriormente, varios péptidos de la superfamilia de las FaRPs han demostrado aumentar la frecuencia y la amplitud de las contracciones de las glándulas salivales, mientras que miosupresina las inhiben (Orchard y Te Brugge, 2002).

Del clivaje proteolítico del precursor de FMRFamida que hemos caracterizado en *R. prolixus (Rhopr-FMRFamida)* se generan trece péptidos, cuatro de los cuales presentan el motivo característico de la familia, mientras que otros cuatro presentan el motivo FIRFamida y cinco son péptidos espaciadores (Figura 1.5). Uno de estos cinco péptidos presenta un residuo Gly en su extremo C terminal, lo que indica que el péptido maduro es amidado (DKALNRL-NH_2). Esta es una modificación característica de los péptidos con actividad regulatoria. De los trece péptidos, cuatro fueron detectados en extractos de SN y cerebro mediante MS. Uno de ellos posee una acetilación en el extremo N-terminal (Acetil-AKDNFIRF-NH_2). Esta PTM es única entre los péptidos maduros descriptos de esta familia en insectos. Otro de los péptidos presenta el motivo FLRLamida, el cual no había sido descripto previamente (Tabla 1.4). Resulta llamativo el hecho de que los tres motivos

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

(FMRFamida, FIRFamida y FLRLamida) se encuentran precedidos por un residuo Asn. Para este gen hemos obtenido la predicción bioinformática del ORF, por lo que no se conocen las regiones no traducidas (UTR) (Figura 1.4).

A partir del análisis bioinformático del genoma se ha podido predecir un GPCR que presenta un alto grado de homología con los receptores de FMRFamida caracterizados en otros insectos (BLASTN *E value*: E%=1e-126).

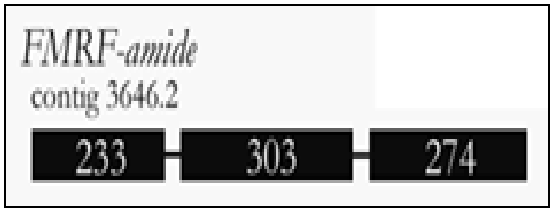


Figura 1.4: Estructura del gen de *Rhopr-FMRFamida*. El número en cada exón indica su longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Los exones 1 y 3 no están completos puesto que faltan los UTR 3´y 5´.

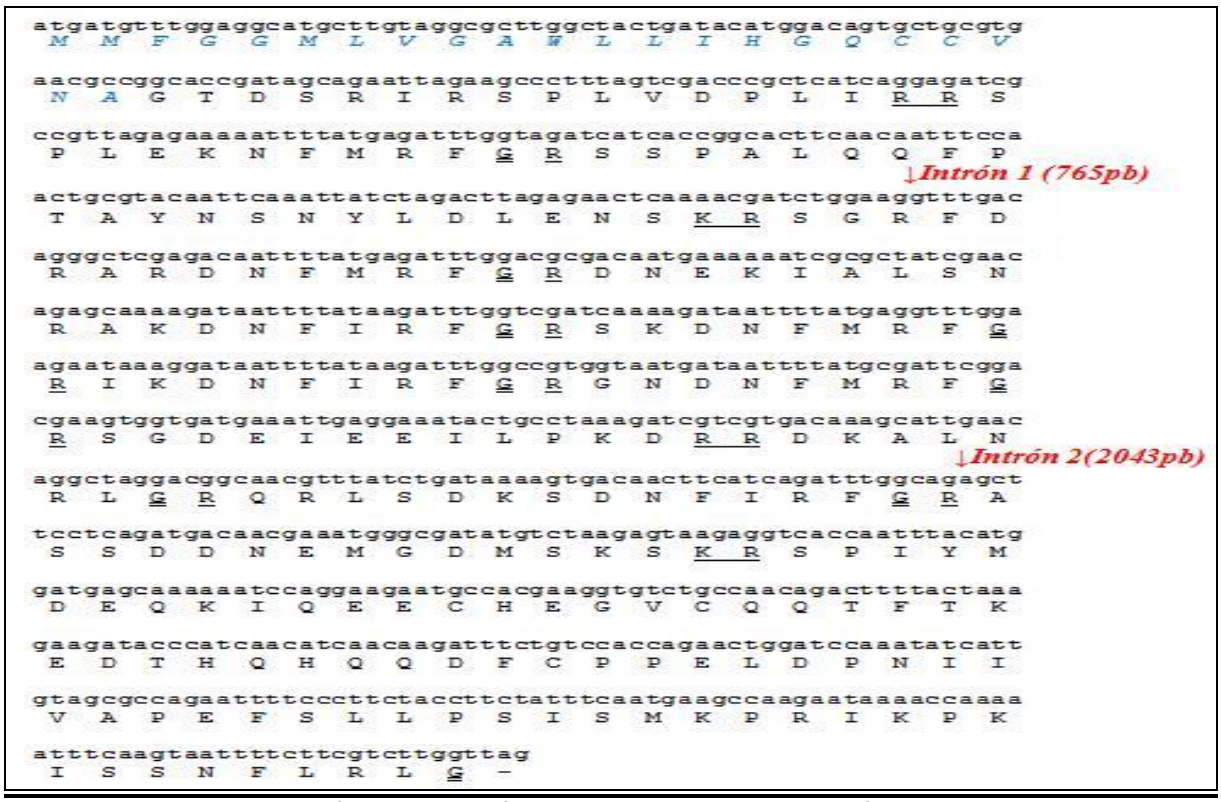


Figura 1.5: Secuencia nucleotídica y aminoacídica de *Rhopr-FMRFamida* . El péptido señal (probabilidad SignalP-HMM: 0.995) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Tabla 1.4: Péptidos predichos a partir del clivaje proteolítico del precursor *Rhopr-FMRFamida* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
GTDSRIRSPLVDPLI		1638,92	
SPLEKNFMRF-NH_2		1267,66	

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

SSPALQQFPTAYNSNYLDLENS		2459,14	
SGRFDRARDNFMRF-NH_2		1773,87	
DNEKIALSNRAKDNFIRF-NH_2		2150,15	
	AKDNFIRF-NH_2	1009,55	Cerebro
	(Acetyl-A)KDNFIRF-NH_2	1051,6	Cerebro
SKDNFMRF-NH_2		1043,51	
IKDNFIRF-NH_2	IKDNFIRF-NH_2	1051,61	SN
GNDNFMRF-NH_2		999,45	
SGDEIEEILPKD	SGDEIEEILPKD	1344,65	Cerebro
DKALNRL-NH_2		828,51	
QRLSDKSDNFIRF-NH_2		1624,86	
ASSDDNEMGDMSKS		1473,55	
SPIYMDEQKIQEECHEGVCQQTFTK EDTHQHQQDFCPPELDPNIIVAPEFS LLPSISMKPRIKPKISSNFLRL-NH_2		8477,2	

Miosupresina

El primer péptido de esta familia fue caracterizado en 1986 (pyro-EDVDHVFLRFamida) por su capacidad de inhibir las contracciones espontaneas del recto en la cucaracha *Leucophaea madera* (Nachman *et al,* 1986). Las Miosupresinas (MSP) pertenecen a la familia de neuropéptidos de insectos conocida como "FLRFamida", que presentan en común la presencia del heptapéptido DHVFLRFamida en su extremo C-terminal. Las MSP se caracterizan por ser fuertes inhibidores de la contracción muscular de las vísceras y la actividad cardíaca. Se ha demostrado que el heptapéptido es el fragmento más pequeño capaz de ejercer tal efecto (Nachman *et al*, 1993).

Se conocen compuestos no peptídicos que mimetizan la actividad de MSP. El compuesto *cloruro de benzetonio* (Bztc) es ser capaz de inhibir las contracciones espontaneas de los oviductos en *Schistocerca gragaria* con una EC_{50}= $6,4*10^{-7}$ M, que está en el rango del heptapéptido. Se ha demostrado que esta inhibición está dada por la interacción directa con el receptor de MSP (Lange *et al*, 1995).

El péptido *Rhopr-MS* que hemos caracterizado presenta en su sitio activo (región C-terminal) el motivo FMRF-amida, en lugar del motivo FLRF-amida, característico de todas las MSP conocidas hasta la fecha (ver Nässel, 2002; Audsley y Weaver, 2009). Al igual que los precursores descriptos en otros insectos, *Rhopr-MS* contiene un solo péptido con el dominio de la familia, situado en el extremo C-terminal del precursor (Figura 1.8).

Mediante MS hemos podido detectar 3 variantes para este péptido con distintas PTM, así como una forma truncada del péptido acompañante (Tabla 1.5). Resulta interesante el hecho de que se haya encontrado el péptido pyro-QDIDHVF(M-OH)RF-NH_2, ya que el residuo Met en esa posiciónes es único entre la MSP descriptas en insectos. La PTM en

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

este residuo (oxidación) la diferencia aún mas de los péptidos ortólogos. También se ha detectado el péptido QDIDHVFMRF-NH_2, que carece del grupo Pyro-Q en el extremo N terminal, característico del péptido maduro. Este péptido ha sido identificado en distintas muestras de SN y su frecuencia de aparición sugiere que no es un intermediario inmaduro de vida media breve, sino que podría tener función biológica. Esto mismo fue sugerido en *A. mellifera* (Boerjan *et al*, 2010).

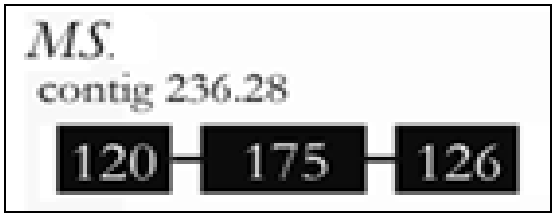


Figure 1.7: Estructura del gen *Rhopr-MSP*. El número en cada exón indica su longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones.

```
acacaattcaacgctatctcgcaaagtccacacttacctctactattttctctagataaa
cacgtcgatttggctgtggtctatctacccttatacgtgctctcctccactaattacaag
↓Intrón 1 (218.736pb)
atgattttggcgtggatgtgtgtgacactactcgctggagttttggctgctccaggacca
 M  I  L  A  W  M  C  V  T  L  L  A  G  V  L  A  A  P  G  P

gattgttctccatctgcactacaagtacagtcctctagagtgagaaacatgtgtgctttg
 D  C  S  P  S  A  L  Q  V  Q  S  S  R  V  R  N  M  C  A  L
                                        Intrón 2 (1.479pb) ↓
taccagatttcaagcgcactgcaggcttatctagatgaacaaaacaactatcagactgct
 Y  Q  I  S  S  A  L  Q  A  Y  L  D  E  Q  N  N  Y  Q  T  A

ctccgagacacgaacataccgtataacatacctgagaagagacaagatattgaccacgta
 L  R  D  T  N  I  P  Y  N  I  P  E  K  R  Q  D  I  D  H  V

ttcatgcgctttggacgtcgacgttaaaatacaagaaaatttaagtggcagagagtgaaa
 F  M  R  F  G  R  R  R  -
```

Figura 1.8: Secuencia nucleotídica y aminoacídica de *Rhopr-MSP* . El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y su longitud entre paréntesis .

Tabla 1.5: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-MS* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
APGPDCSPSALQVQSSRVRNMCALYQISSALQAYLDEQNNYQTALRDTNIPYNIPE		6213,97	
	LDEQNNYQTALRDTNIPYNIPE	2621,25	SN
QDIDHVFMRF-NH_2	QDIDHVFMRF-NH_2	1306,64	Cerebro
	pyro-QDIDHVFMRF-NH_2	1289,6	SN, Cerebro, Midgut, Hemolinfa
	pyro-QDIDHVF(M-OH)RF-NH_2	1305,6	SN, Cerebro

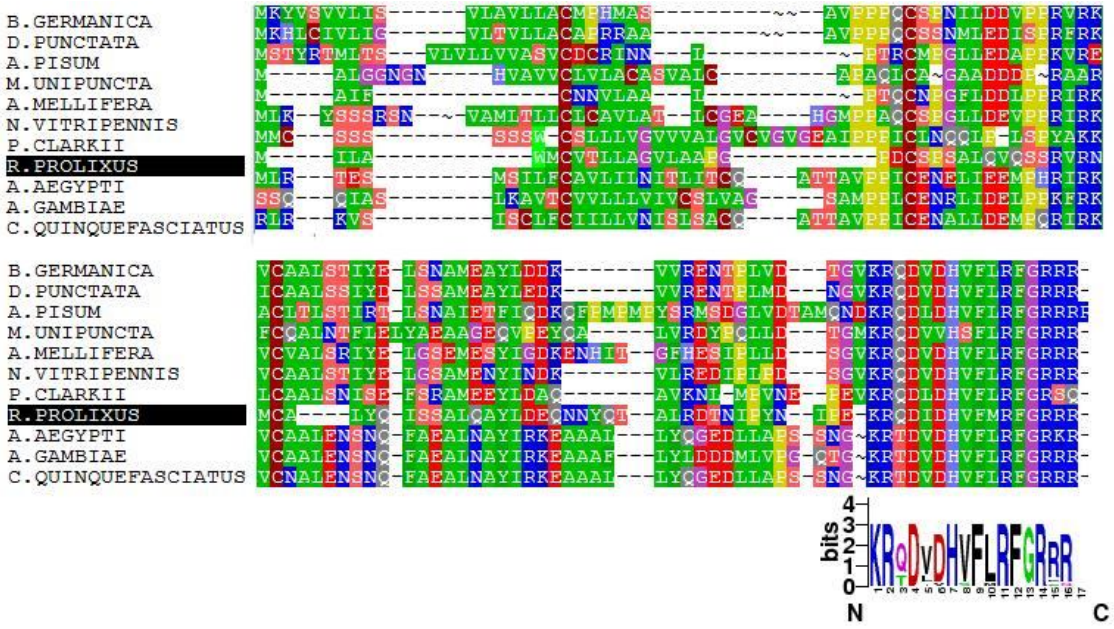
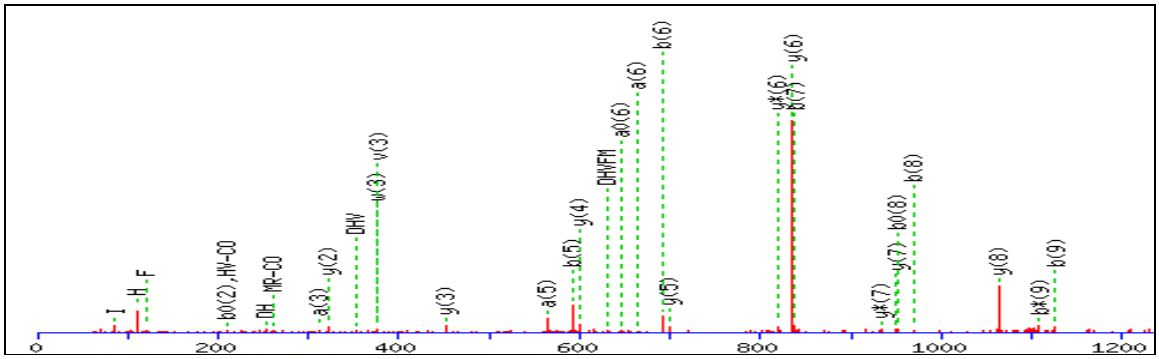


Figura 1.9: Alineamiento del precursor de *Rhopr-MS* con precursores ortólogos caracterizados en distintas especies de insectos. Se pueden observar 3 residuos de Cys conservados. El péptido activo con el motivo característico de esta familia se encuentra ampliamente conservado. En *R. prolixus*, sin embargo este motivo es FMRFa, a diferencia del resto que poseen el dominio FLRFa**.** Se puede observar el contenido informativo de cada posición expresado en bits para cada posición.

Mediante hibridización *in situ* se ha demostrado la transcripción de este GPN en células del digestivo medio posterior de *Pseudaletia unipuncta* y *M. sexta* (Lee, 2002; Lu, 2002). Posteriormente este péptido fue aislado y secuenciado a partir del digestivo medio en *Locusta migratoria* (Hill, 2007). En *Blattella germánica* las MSP inhiben la frecuencia y amplitud de las contracciones del digestivo medio anterior y posterior, reducen de manera dosis dependiente la cantidad de alimento ingerido, el cual se acumula sin pasar hacia el digestivo medio posterior (Aguilar *et al*, 2004). En nuestro trabajo, hemos podido detectar la expresión del péptido pyro-QDIDHVFMRF-NH_2 en el digestivo medio posterior de *R. prolixus*, tanto por MALDI-TOF-TOF MS/MS (Figura 1.10) como por ESI-Q-TOF MS/MS (Figura 1.11). Este hallazgo sugiere un papel regulador de MSP en la actividad del digestivo posterior de manera autócrina o parácrina. Por otra parte, hemos detectado el péptido MSP en hemolinfa (ver mas adelante), lo que sugiere una actividad a nivel sistémico. Previamente, se había detectado la presencia de péptidos simil-FMRFamida en hemolinfa luego de la ingesta sanguínea en *Rhodnius* (Elia *et al*; 1993). Tal vez lo que se detectó en realidad fue esta miosupresina

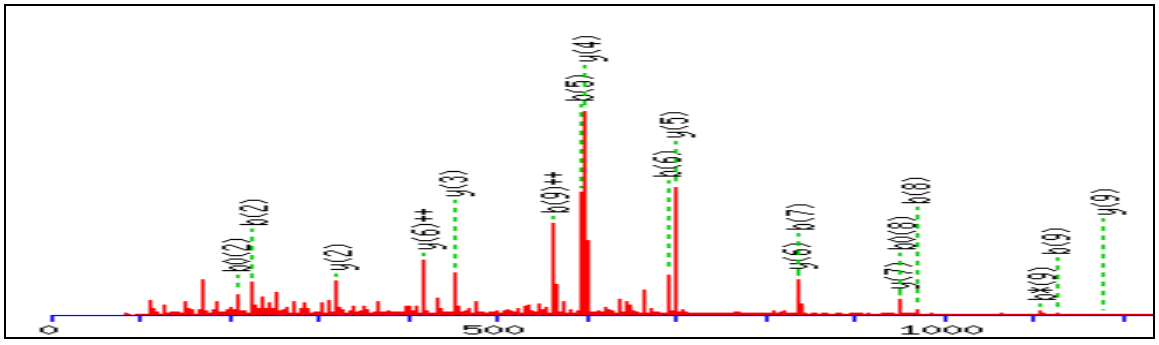
Mediante búsquedas bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de Miosupresinas caracterizados en otros insectos (BlastN *E value*: E%=3e-110).



Monoisotopic mass of neutral peptide Mr(calc): 1288.6023
Variable modifications:
N-term : Gln->pyro-Glu (N-term Q)
C-term : Amidated (C-term)
Ions Score: 42 Expect: 0.0069
Matches (Bold Red): 32/143 fragment ions using 58 most intense peaks

#	Immon.	a	a*	a^0	b	b*	b^0	d	Seq.	v	w	w'	y	y*	y^0	#
1	101.0709	84.0444	67.0179		112.0393	95.0128			Q							10
2	88.0393	199.0713	182.0448	181.0608	227.0663	210.0397	**209.0557**		D	1118.5564	1117.5611		1178.5775	1161.5510	1160.5670	9
3	**86.0964**	**312.1554**	295.1289	294.1448	340.1503	323.1238	322.1398		I	1005.4723	1018.4927	1032.5084	**1063.5506**	1046.5240	1045.5400	8
4	88.0393	427.1824	410.1558	409.1718	455.1773	438.1507	437.1667		D	890.4454	889.4501		**950.4665**	**933.4400**	932.4559	7
5	**110.0713**	**564.2413**	547.2147	546.2307	**592.2362**	575.2096	574.2256		H	753.3865			**835.4396**	**818.4130**		6
6	72.0808	**663.3097**	646.2831	**645.2991**	**691.3046**	674.2780	673.2940		V	654.3181	667.3385		**698.3807**	681.3541		5
7	**120.0808**	810.3781	793.3515	792.3675	**838.3730**	821.3465	820.3624		F	507.2496			**599.3122**	582.2857		4
8	104.0528	941.4186	924.3920	923.4080	**969.4135**	952.3869	**951.4029**		M	**376.2092**	**375.2139**		**452.2438**	435.2173		3
9	129.1135	1097.5197	1080.4931	1079.5091	**1125.5146**	**1108.4881**	1107.5040	1012.4557	R	220.1080	219.1128		**321.2033**	304.1768		2
10	**120.0808**								F				165.1022			1

Figura 1.10: Espectro MS/MS del péptido pyro-QDIDHVFMRF-NH_2 detectado en el digestivo medio posterior mediante Nano-HPLC - MALDI-TOF/TOF MS. En rojo se muestran los iones detectados. * indica la pérdida de un grupo amonio, ° la pérdida de una molécula de agua en un ión.



Monoisotopic mass of neutral peptide Mr(calc): 1288.6023
Variable modifications:
N-term : Gln->pyro-Glu (N-term Q)
C-term : Amidated (C-term)
Ions Score: 31 Expect: 0.038
Matches (Bold Red): 18/92 fragment ions using 46 most intense peaks

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	112.0393	56.5233	95.0128	48.0100			Q							10
2	**227.0663**	114.0368	210.0397	105.5235	**209.0557**	105.0315	D	**1178.5775**	589.7924	1161.5510	581.2791	1160.5670	580.7871	9
3	340.1503	170.5788	323.1238	162.0655	322.1398	161.5735	I	1063.5506	532.2789	1046.5240	523.7657	1045.5400	523.2736	8
4	455.1773	228.0923	438.1507	219.5790	437.1667	219.0870	D	**950.4665**	475.7369	933.4400	467.2236	932.4559	466.7316	7
5	**592.2362**	296.6217	575.2096	288.1085	574.2256	287.6164	H	**835.4396**	**418.2234**	818.4130	409.7101			6
6	**691.3046**	346.1559	674.2780	337.6427	673.2940	337.1507	V	**698.3807**	349.6940	681.3541	341.1807			5
7	**838.3730**	419.6901	821.3465	411.1769	820.3624	410.6849	F	**599.3122**	300.1598	582.2857	291.6465			4
8	**969.4135**	485.2104	952.3869	476.6971	**951.4029**	476.2051	M	**452.2438**	226.6256	435.2173	218.1123			3
9	**1125.5146**	**563.2609**	**1108.4881**	554.7477	1107.5040	554.2557	R	**321.2033**	161.1053	304.1768	152.5920			2
10							F	165.1022	83.0548					1

Figura 1.11: Espectro MS/MS del péptido pyro-QDIDHVFMRF-NH2 detectado en el digestivo medio posterior mediante nanoLC-ESI-Quadrupole-TOF MS. En rojo se muestran los iones detectados. * indica la pérdida de un grupo amonio, ° la pérdida de una molécula de agua en un ión.

Sulfakininas

Las Sulfakininas (SK) de insectos presentan homología a nivel de estructura primaria con la familia de colecistokininas (gastrinas) de vertebrados. Los péptidos pertenecientes a esta familia se caracterizan por la presencia de la secuencia DYGHMRFamida en el extremo C-terminal (ver *Nässel, 2002)*. En *L. migratoria*, SK reduce la ingesta de alimento, estimula la contracción muscular del recto y la liberación de enzimas digestivas en el intestino medio (Baggerman *et al*, 2000). En *S. gragaria*, inyecciones de este péptido inhiben hasta un 55% la ingesta de alimento de manera dosis dependiente. Si bien el mecanismo de acción de estos neuropéptidos es desconocido, se ha propuesto que activarían los receptores que censan la distensión del buche, reduciendo así la cantidad de alimento ingerido (Wei *et al*; 2000).

A partir de la secuencia del precursor de *Rhopr-SK* que hemos caracterizado se predicen cinco péptidos, dos con el motivo característico de esta familia (Figura 1.13; tabla 1.6). Uno de ellos, GGSDEKFDDYGYMRF-NH2, es distinto a las SK ortólogas, ya que presenta la sustitución de His^4 por Tyr^4 (en posición 4 a partir del extremo C terminal del péptido maduro; Figura 1.14)

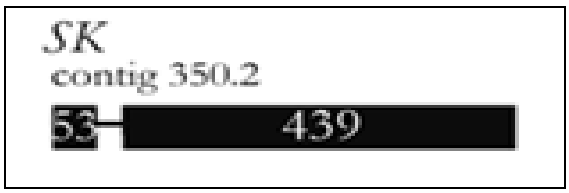


Figure 1.12: Estructura del gen de *Rhopr-SK*. El número en cada exón indica su longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones.

Figura 1.13: Secuencia nucleotídica y aminoacídica de *Rhopr-SK* . El péptido señal (probabilidad SignalP-HMM: 0.800) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición del intrón se muestra en rojo y entre paréntesis la longitud del mismo.

Tabla 1.6 Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-SK* y péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

Péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
SHFNCLAEPAE		1217,53	
SLIRIRPEPALFAAEDDPLDIVD	SLIRIRPEPALFAAEDDPLDIVD	2565,36	Cerebro
QFNEYGHMRF		1327,6	
	pyro-QFNEYGHMRF-NH_2	1310,6	Cerebro
	pyro-QFNEYGH(M-OH)RF-NH_2	1326,6	Cerebro
GGSDEKFDDYGYMRF-NH2	GGSDEKFDDYGYMRF-NH_2	1785,75	Cerebro
	GGSDEKFDDYGY(M-OH)RF-NH_2	1801,76	Cerebro
SRPLANSLPN		1068,58	

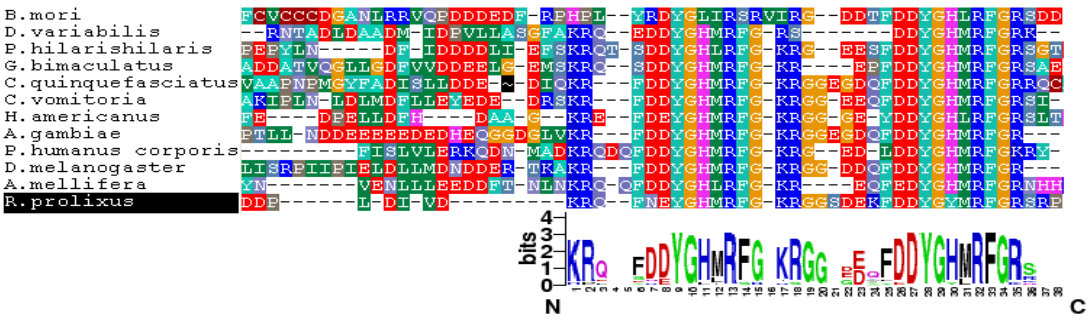


Figura 1.14: Alineamiento del precursor de *Rhopr-SK* con los precursores ortólogos caracterizados en distintas especies de insectos. El contenido informativo expresado en bits se muestra para cada posición. Se muestra el extremo C-terminal de los precursores.

Mediante búsquedas bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de SK caracterizados en otros insectos (BLASTN *E value*: E%=5e-76).

Neuropéptido F corto

Los primeros miembros de esta familia fueron aislados de *Locustra decemlineata* y designados con ese nombre debido a que presentan homología en el extremo C-terminal con el Neuropéptido Y de vertebrados (Spittaels *et al*, 1996). A partir de esa fecha, se han caracterizado péptidos ortólogos en varios Órdenes de insectos. Todos ellos comparten como característica común la presencia del motivo LRLRFamida en el extremo C-terminal. En *D. melanogaster* se ha sugerido que este péptido está involucrado en la regulación de la alimentación, ya que su sobre-expresión produce un aumento en la cantidad de alimento ingerido, mientras que los mutantes donde no se expresa no se alimentan (Lee *et al*, 2004). Posteriormente, se demostró que en *L. decemlineata* el Neuropéptido F corto (SNF) se expresa de manera diferencial en los insectos durante el estado activo, respecto a los que se encuentran en diapausa, por lo que se ha sujerido un papel de SNF durante la diapausa (Huybrechts *et al*, 2004). Por otra parte, se ha demostrado que SNF estimula el desarrollo ovárico en *L. migratoria* y *S. gregaria* (Cerstiaens *et al*, 1999; De Loof *et al*, 2001), y provoca un aumento en la concentración de vitelogenina en hemolinfa en esta última especie (Schoofs *et al*, 2001)

El precursor de SNF que hemos caracterizado en *R. prolixus* (*Rhopr-SNF*) codifica tres péptidos maduros, flanqueados por pares de residuos básicos (Figura 1.16). De éstos, solo uno (NNRSPQLRLRF-NH_2) presenta el motivo característico de la familia. La secuencia de este péptido se encuentra conservada entre distintas especies, exepto en las primeras dos posiciones del extremo N terminal. El residuo Gln^6 de *Rhopr-SNF* es atípico, ya que todas las secuencias parálogas poseen un residuo Ser en esa posición (Figura 1.15). Hemos detectado la presencia de este péptido en extractos de cerebro, SN total y hemolinfa, lo cual sugiere un papel humoral. De los péptidos acompañantes, uno de ellos y dos formas truncadas del otro han sido detectados en extractos del SN, (Tabla 1.7). Una particularidad que presenta *Rhopr-SNF* es su longitud, la cual es menor que los precursores de SNF caracetrizados en la mayoría de los artrópodos, con la excepción de *A. mellifera.*

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

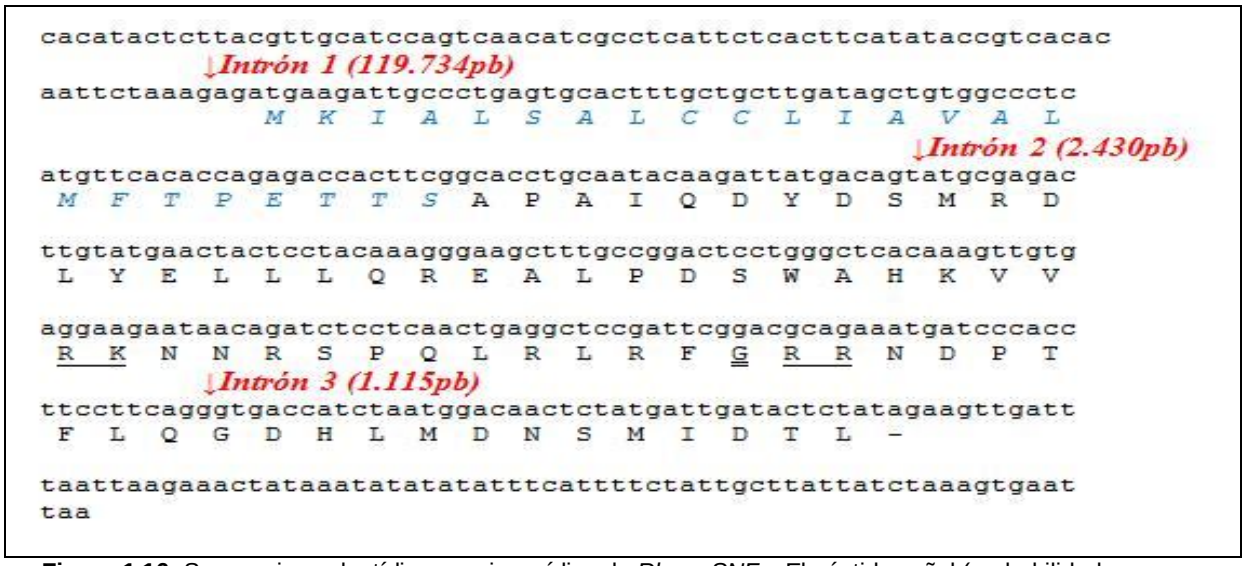
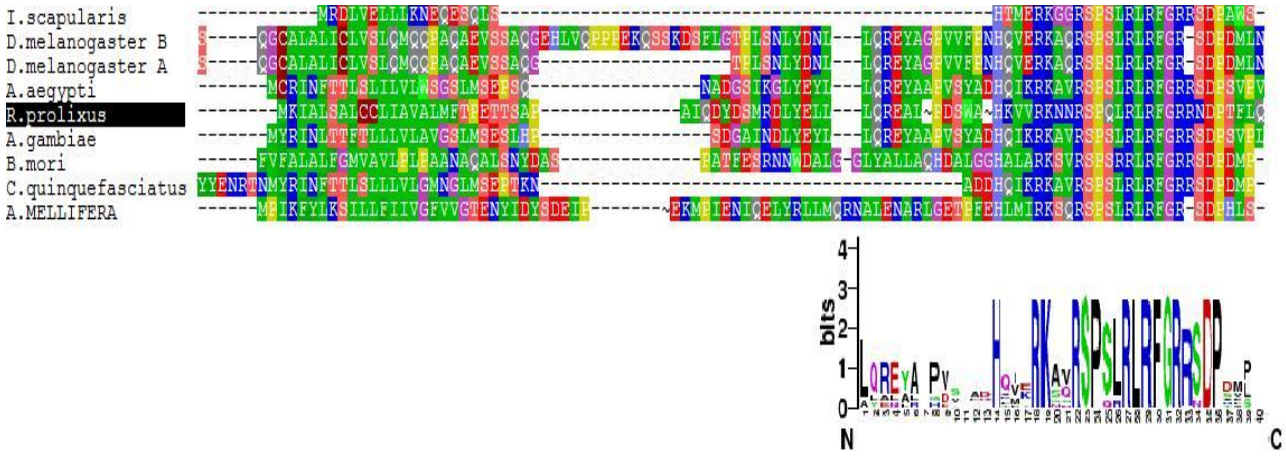


Figura 1.16: Secuencia nucleotídica y aminoacídica de *Rhopr-SNF* . El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Tabla 1.7: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-SNF* y péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
APAIQDYDSMRDLYELLLQREALPDSWAHKVV		3742,89	
	AIQDYDSMRDLYELL	1844,874	SN
	LQREALPDSWAHKVV	1749	SN
NNRSPQLRLRF-NH_2	NNRSPQLRLRF-NH_2	1399,79	SN, Cerebro, Hemolinfa



NDPTFLQGDHLMDNSMIDTL	NDPTFLQGDHLMDNSMIDTL	2276,01	SN, Cerebro

Figura 1.15: Alineamiento del precursor de *Rhopr-SNF* con los precursores ortólogos caracterizados en distintas especies de insectos. El contenido informativo expresado en bits se muestra para cada posición.

Mediante búsqueda bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de SNF caracterizados en otros insectos (BLASTN *E value*: E%=1e-122).

Neuropéptido F largo

El neuropéptido F largo (LNF) de insectos es homólogo al Neuropéptido Y de vertebrados. En *D. melanogaster y A. aegypti* LNF ha sido localizado en el cerebro, SOG y células endócrinas del digestivo medio posterior (Brown *et al*, 1999; Stanek *et al*, 2002). En *A. aegypti*, la concentración de LNF en hemolinfa es máxima 24 hs después de la alimentación (Stankes *et al*, 2002). En este insecto se ha demostrado también que reduce el transporte de iones y los movimientos peristálticos de buche (Onken *et al*, 2004). En el genoma de *B. mori* se han detectado 2 genes parálogos para este péptido, uno de los cuales presenta dos formas alternativas de *splicing*. Estos dos genes presentan un patrón de expresión diferencial en el SNC, y su expresión ha sido detectada también en el digestivo medio (Roller *et al*, 2008). En *R. prolixus* se ha observado inmunomarcación para este péptido en todo el SN, incluyendo el SNC, SNP y SNS (Gonzalez *et al*, 2008).

Mediante RACE-PCR hemos podido obtener el ORF completo del mRNA que codifica este neuropéptido en *R. prolixus,* así como el UTR 5´ (Figura 1.18). Mediante MS se han encontrado dos formas truncadas que presentan el dominio característico de esta familia en invertebrados, RX_1RFamida (sinedo X_1 un residuo de Pro abitualmente). También se han encontrado dos formas truncadas del péptido acompañante (Tabla 1.8).

Tanto *B. mory* como *A. mellifera* presentan el motivo RX_1RYamida, característico de vertebrados (Figura 1.19).

Mediante búsquedas bioinformática hemos podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de LNF caracterizados en otros insectos (BLASTN *E value*: E%=2e-42).

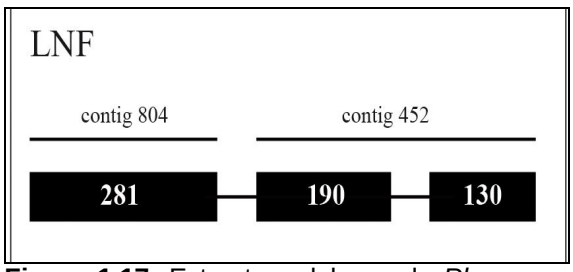


Figure 1.17: Estructura del gen de *Rhopr-LNF*. El número en cada exón indica su longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones. No se ha secuenciado el UTR 3´, por lo que el exón 3 se grafica de forma parcial.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 1.18: Secuencia nucleotídica y aminoacídica de *Rhopr-LNF*. El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Tabla 1.8: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-LNF* y péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
QPIPADAMARPARPKSFASPDDLRTYLDQLGQYYAVAGRPRF-NH_2		4705.42	
	AVAGRPRF-NH2	872,53	SN, Cerebro
	YAVAGRPRF-NH2	1035,58	SN
AGGINPRLHLAVDGVNYRYPLADASDLYDLLFQQQSTE		4220,1	
	AGGINPRLHLAVDGVN	1602,87	SN
	YRYPLADASDLYDLL	1787,88	SN

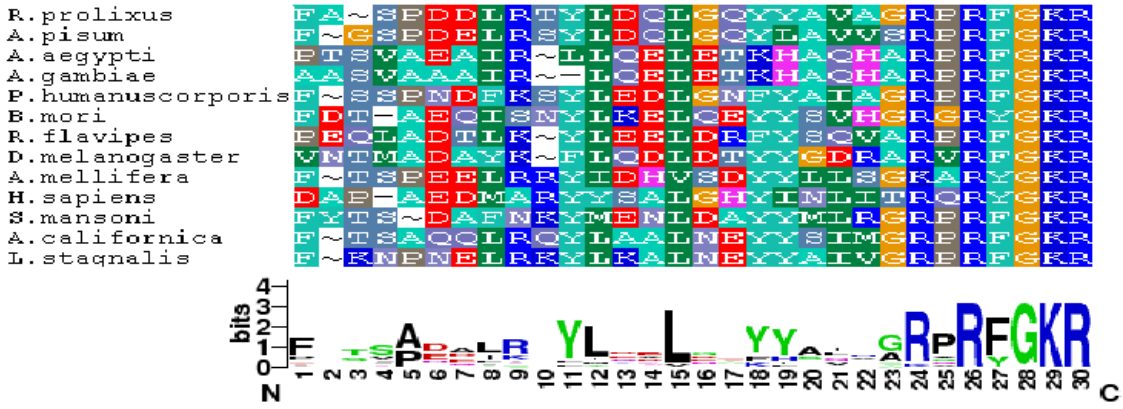


Figura 1.19: Alineamiento del precursor de Rhopr-LNF con los precursores ortólogos caracterizados en distintas especies de insectos. Las Pro y Tyr son importantes para la formación de la estructura "PP-fold" de los Neuropéptidos F (Larhammar, 1996). El contenido informativo expresado en bits se muestra para

cada posición. Se muestra el contenido informativo para cada posición expresado en bits. En el alineamiento se ha incluído el Neuropéptido Y de Homo sapiens.

1.3.2.b) Péptidos diuréticos y antidiuréticos

Los insectos hematófagos son capaces de ingerir grandes cantidades de sangre en cada comida, por lo que rápidamente se inicia el proceso de diuresis mediante el que se produce y elimina la orina. Durante la diuresis, la región distal de los TM genera la orina primaria, la cual es luego modificada en la región proximal mediante reabsorción selectiva de iones. La diuresis en insectos se encuentra regulada por serotonina y hormonas peptídicas (ver Orchard I, 2009).

En *R. prolixus* hemos encontrado tres hormonas caracetrizadas en distintos insectos como diuréticas: DH44, DH31 y Kinina; y dos antidiuréticas: Péptido Transportador de Iones (ITP) y Neuroparsina A (NPA, además de los genes CAPA descriptos previamente (Paluzzi et al, 2008; 2010). Las primeras estimulan la producción de orina por los TM, mientras que las antidiuréticas estimulan la reabsorción de fluído por el recto, o bien reducen la secreción por parte de los TM.

Hormona diurética simil- calcitonina

Uno de los procesos fisiológicos característicos en los insectos hematófagos es la rápida diuresis que se inicia durante la ingesta sanguínea y culmina 3-4 horas luego de su finalización. Este proceso involucra la acción coordinada de varios tejidos, regulados por el sistema neuro-hormonal. Una de las hormonas que regulan dicha función pertenece a la familia de “Hormonas Diuréticas simil-calcitonina” (CT-*like* DH o DH31).

Un péptido maduro producto del precursor de DH31 fue el primero en ser caracterizado en *R. prolixus* (Te Brugge y Orchard, 2008). El mismo fue aislado a partir del SNC de larvas de quinto estadío por cromatografía líquida de fase reversa (RPLC) y secuenciado *de novo* mediante MALDI-TOF MS/MS y degradación de Edman. La secuencia encontrada (GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2) resultó ser idéntica a la que había sido caracterizada previamente en la cucaracha *Diploptera punctata* (Furuya, 2000). En *R. prolixus,* ensayos de inmunohistoquímica (IHQ) habían demostrado previamente la presencia de un péptido simil-DH31 en el SNC (células neurosecretoras medias (PI) y laterales (PL) del protocerebro), MTGM, proyecciones nerviosas sobre las glándulas salivales, el recto y la arteria dorsal, y también en células endócrinas del digestivo medio (Donini *et al*, 2008). Sin embargo, DH31 demostró ser incapaz de aumentar la tasa de secreción de los TM ni de modificar la composición del fluido secretado. Tampoco fue capaz de estimular la recaptación de K^+ por parte de la región proximal (Donini *et al*, 2008). Por otro lado, el péptido demostró ser capaz de aumentar la frecuencia de contracción del recto, la arteria dorsal y del digestivo medio anterior, aunque no influyó en la tasa de

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

absorción de agua por este último epitelio, ni sobre la resistencia y el voltage trans-epitelial (Te Brugge y Orchard, 2008; Te brugge *et al*, 2009).

En el presente trabajo hemos logrado caracterizar el GPN que codifica el péptido DH31. Mediante RACE-PCR hemos logrado obtener 2 formas alternativas de *splicing*, mientras que una tercera isoforma fue encontrada en la base de datos de EST de testículo (Figuras 1.21, 1.22y 1.23). El gen consta de al menos 8 exones. Las isoformas A y B comparten los exones 1, 2, 4, y 8. El exón 2 codifica el péptido señal, mientras que el exón 8 codifica el péptido maduro con actividad diurética (GLDLGLSRGFSGSQAAKHLMGLAANYAGGP-NH2). Las isoformas A y C comparten los exones 4, 5, y 6. Los exones 3 y 7 se encuentran solo en la isoforma C (Figura 1.20). Por lo tanto, las isoformas A y B codifican péptidos de secreción, y al menos un péptido maduro con actividad diurética, mientras que la isoforma C codifica una proteína que no es de secreción, en la cual el péptido diurético se encuentra ausente. Es interesante remarcar que hemos observado un patrón de expresión tisular específico de las diferentes isoformas. Las isoformas A y B se encontraron presentes en el SNC, mientras que la isoforma C fue la única encontrada en la base de datos de EST generadas a partir de testículo. El neuropéptido CAPA, otra molécula asociada al control de la diuresis, también fue encontrado en testículos en *R. prolixus* (Palluzi *et al*, 2008). Nuestros resultados sugieren una nueva función para el producto proteico generado por la isoforma C del gen *Rhopr-DH31*, en la maduración sexual y/o la reproducción. Llamativamente, los exones 6 y 7 se encuentran tanto en el *contig* 15451 como en el *contig* 2671, mientras que el exón 8 se encuentra en los *contigs* 6076 y 5528 (Figura 1.20), lo que indica que parte de este gen se ha duplicado en el genoma. El hecho de que los exones parálogos sean iguales, sin que presenten mutaciones, indica que la duplicación ha sido reciente.

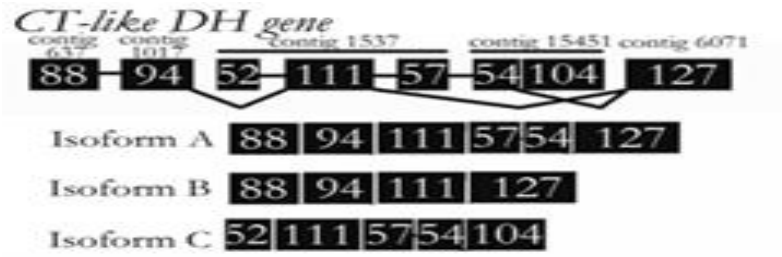


Figure 1.20 : Estructura del gen de DH31 y las tres formas de *splicing* encontradas. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones. El péptido señal está codificado en el exón 2 y el péptido con actividad diurética en el último exón. Las isoformas 1 y 2 se transcriben en SN, mientras que la isoforma 3 en Testículos. Para las isoformas A y B no se conocen los UTRs 3´, por lo que el exón 8 está incompleto.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

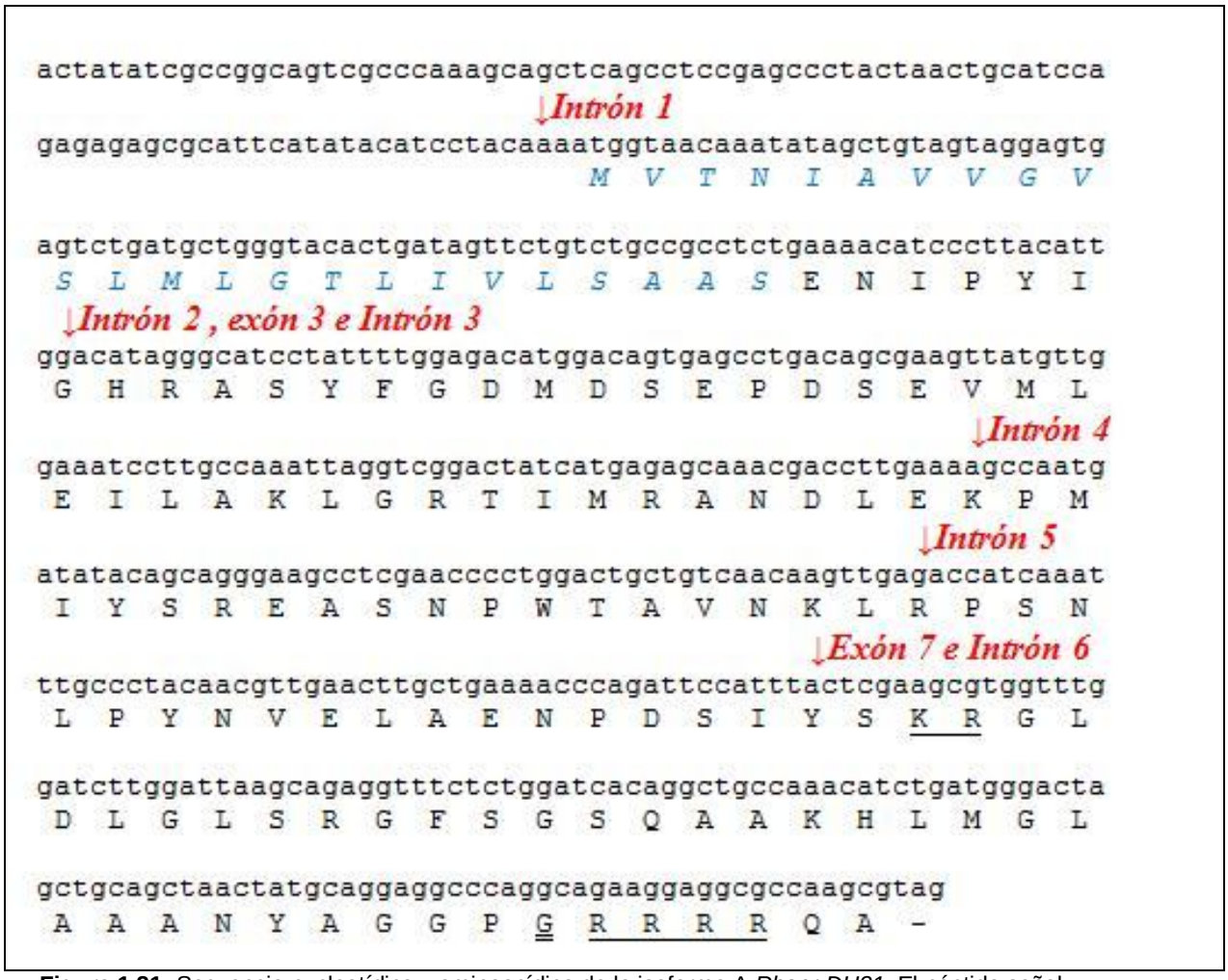


Figura 1.21: Secuencia nucleotídica y aminoacídica de la isoforma A *Rhopr-DH31*. El péptido señal (probabilidad SignalP-HMM: 0.966) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones y los exones ausentes en la isoforma se muestran en rojo.

Tabla 1.9: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-DH31* isoforma A y los péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
SENIPYIGHRASYFGDMDNEPDSEVMLEILAKLGRTIMRANDLEKPM-IYSREASNPWTAVNKLRPSNLPYNIELAENPDSIYS		9469,63	
GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	2986,5	SN, Cerebro

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

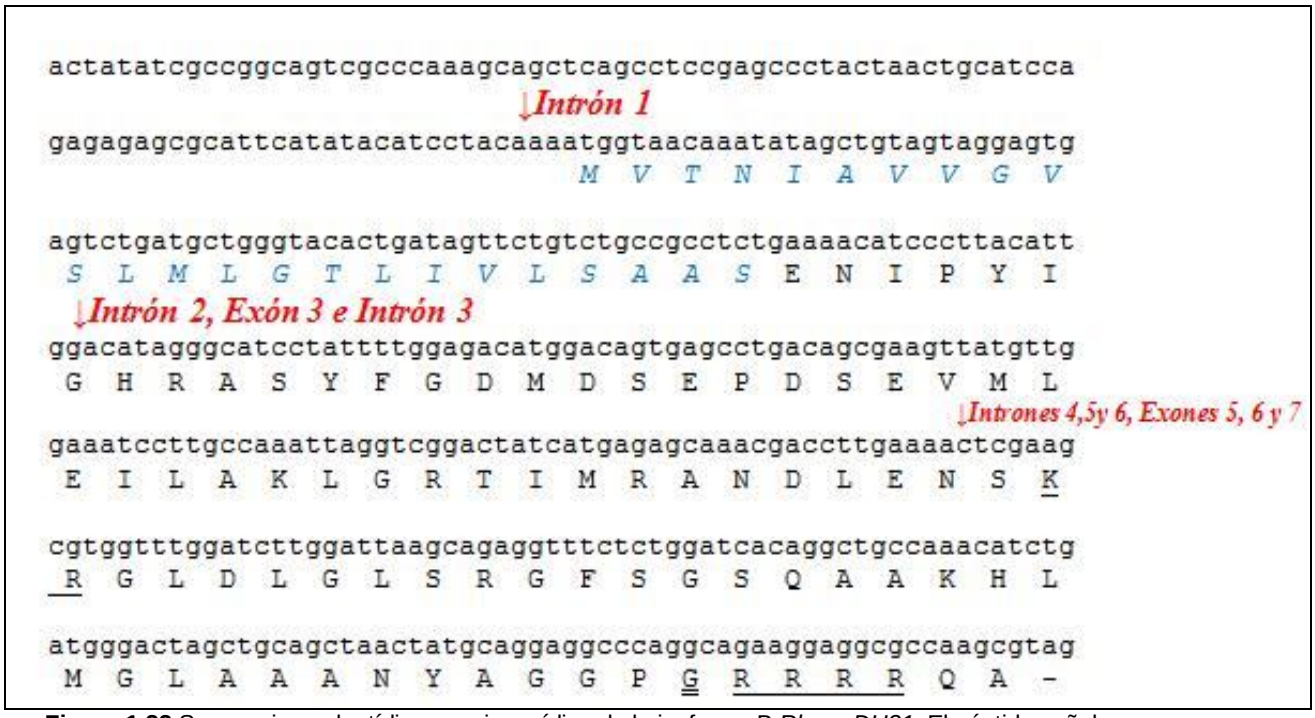


Figura 1.22 Secuencia nucleotídica y aminoacídica de la isoforma B *Rhopr-DH31*. El péptido señal (probabilidad SignalP-HMM: 0.966) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones y los exones ausentes se muestra en rojo.

Tabla 1.10: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-DH3*1 isoforma B y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
SENIPYIGHRASYFGDMDNEPDSEVMLEILAKLGRTIMRANDLENS		5211,48	
GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH_2	2986,5	SN, Cerebro

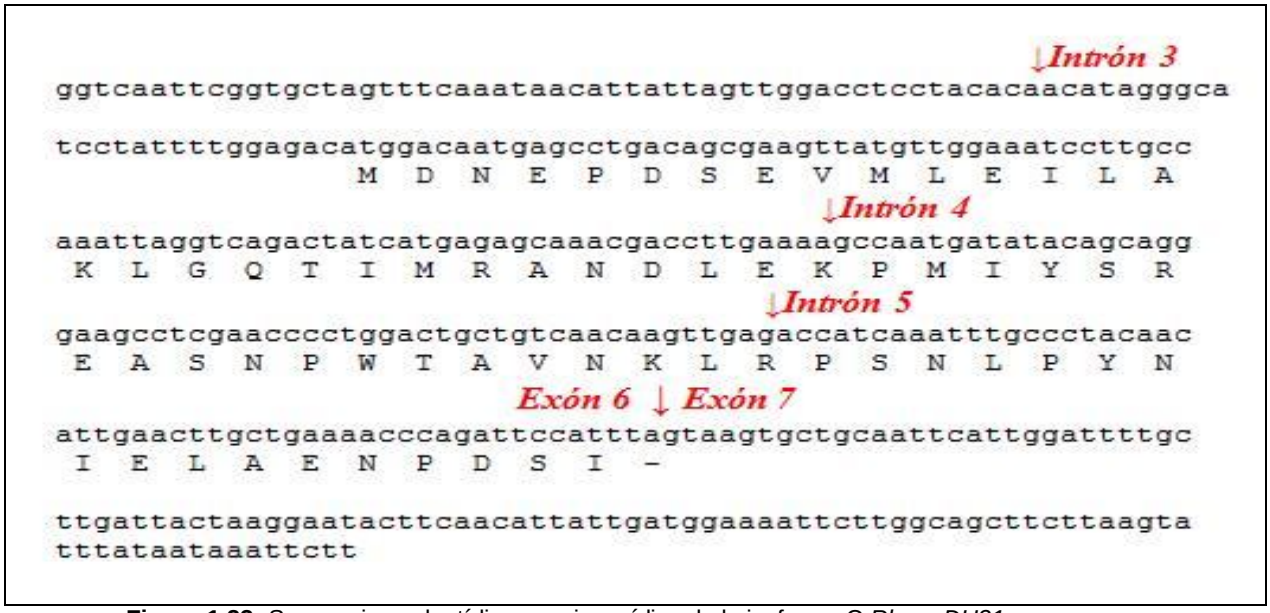
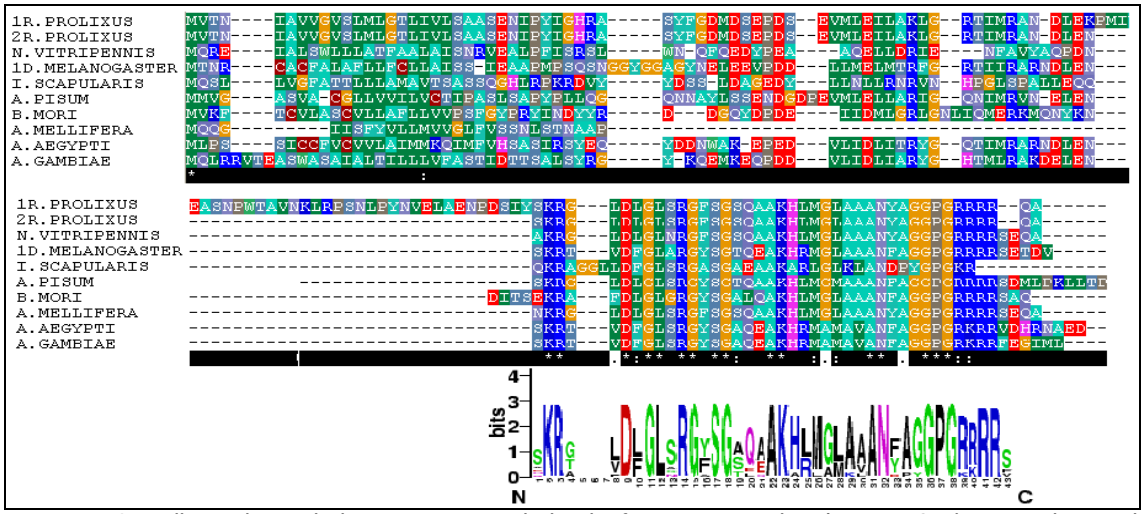


Figura 1.23: Secuencia nucleotídica y aminoacídica de la isoforma C *Rhopr-DH31*..

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Como puede observarse en el alineamiento de las secuencias ortólogas, la región codificada en la isoforma A que corresponde a los exones 5 y 6, no se encuentra en nigún otro insecto. Esto sugiere que la isoforma B es la ancestral y que la isoforma A ha surgido mas tarde durante la evolución de *R. prolixus* (Figura 1.24). Llamativamente, los exones 5 y 6 se encuentran también presentes en la isoforma C que se transcribe en testículos, no codifica el péptido diurético y no posee péptido señal (Figura 1.25). Dado que son formas alternativas de *splicing* del mismo gen, éstas se transcriben a partir del mismo promotor en ambos órganos (SN y testículo) y la expresión espacial está regulada por *splicing* alternativo del pre-mRNA.

In silico se ha encontrado un GPCR que presenta un alto grado de homología con los receptores de DH31 caracterizados en otros insectos (BLASTN *E value*: E%=9e-89).



Fugura 1.24: Alineamiento de los precursores de las isoformas A y B de *Rhopr-DH31* (que contienen el péptido diurético) con los precursores ortólogos de otras especies de insectos. Se puede apreciar que la isoforma A contiene una región que no se encuentra en precursores de otros insectos, precediendo el péptido con actividad diurética. Esta región se presenta también en la isoforma C, y es codificada en los exones 5 y 6. El contenido informativo expresado en bits se muestra para cada posición.

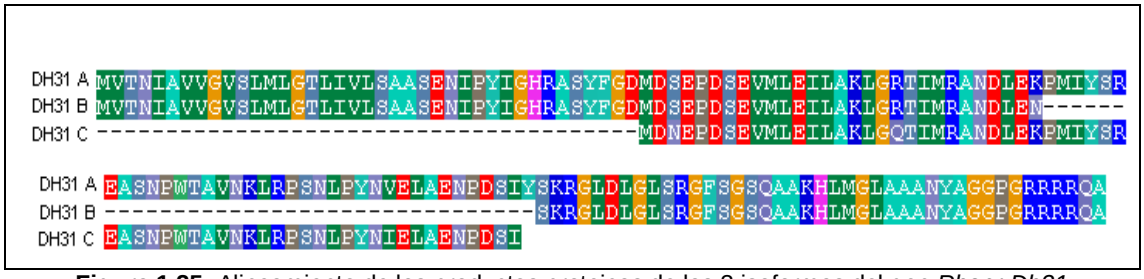


Figura 1.25: Alineamiento de los productos proteicos de las 3 isoformas del gen *Rhopr-Dh31*

Hormona diurética simil-CRF

En *R. prolixus*, se ha visto que el péptido ortólogo de *Zootermopsis nevadensis* (*Zoone-DH44*) es capaz de aumentar la absorción de agua a través del epitelio del digestivo medio anterior, así como de disminuir la resistencia y el voltaje trans-epiteliales de este órgano.

La cantidad de cAMP en digestivo medio anterior aumenta cuando se trata con *Zoone-DH44*, por lo que sería el segundo mensajero utilizado en la cascada se señalización de este péptido (Te Brugge *et al*, 2009). En TM el péptido *Zoone-DH44* ha demostrado ser capáz de aumentar la tasa de secreción, aunque fue incapaz de estimular la recaptación de K^+ en la región proximal (Donini *et al*, 2008). En estos órganos, al igual que en el digestivo medio anterior, Zoone-DH44 aumenta el contenido intracelular de cAMP. También mostró tener actividad sobre las contracciones peristálticas del recto (Te Brugge *et al*, 2002)

El dominio C-terminal de esta familia de péptidos (WRamida) es necesario para la interacción con el receptor, mientras que el grupo NH_2 aumenta la afinidad por el mismo. Por otro lado, el dominio N-terminal es necesario para la activación del receptor. En *M. sexta* adopta una estructura de hélice-giro-hélice, lo que produce que ambos extremos se encuentren próximos en el espacio (Schooley *et al*, 2005).

Para este GPN hemos podido identificar una secuencia parcial a partir del genoma de *R. prolixus*. Como puede observarse en la figura 1.26, además del péptido diurético con el motivo C terminal WRamida, se producen al menos otros dos péptidos a partir del cilvaje del precursor, uno de ellos presenta al extremo C-terminal amidado. Mediante MS hemos detectado la presencia de uno de estos péptidos maduros en el SN, lo que confirma la expresión de la secuencia parcial encontrada (Tabla 1.11).

Por otro lado, hemos podido predecir i*n silico* un GPCR que presenta un alto grado de homología con los receptores de DH44 caracterizados en otros insectos (BLASTN *E value*: E%=2e-136).

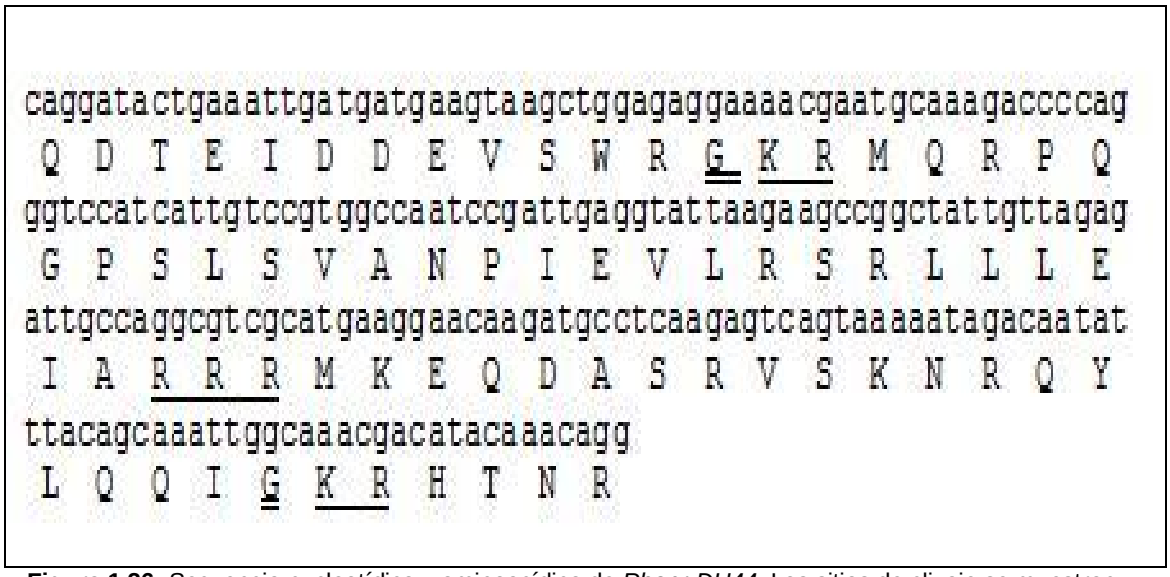


Figura 1.26: Secuencia nucleotídica y aminoacídica de *Rhopr-DH44*. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Tabla 1.11: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-DH44* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	[M+H]+	Tejido
…..QDTEIDDEVSWR-NH_2			
MQRPQGPSLSVANPIEVLRSRLLLEIA		2987,68	
	MQRPQGPSLSVANPIEVLRSRLLLE	2803,56	SN
MKEQDASRVSKNRQYLQQI-NH_2		2321,21	

Kininas

El péptidos de esta familia han sido caracterizados en especies pertenecientes a varios Órdenes de insectos: Díptera, Lepidóptera, Orthoptera y Dictyoptera. Estos péptidos han demostrado también ser agentes diuréticos potentes, capaces de aumentar la tasa de secreción de los TM (Nachman *et al*, 2002; 2009). También se los ha implicado en la regulación de la liberación de enzimas digestivas. (Harshini *et al*, 2003).

La presencia de péptidos simil-Kininas ha sido demostrada en ninfas de quinto estadío de *R. prolixus* mediante IHQ y radioinmunoensayo (RIA), tanto en el SN como en el digestivo medio posterior, así como también su co-localización con péptidos simil-CRF en células del MMTG y en los nervios abdominales I y II (Te brugge *et al*, 2001). En *R. prolixus*, el péptido Leucokinina 1 (LK-1) de *L. mederae* ha demostrado ser capaz de aumentar la frecuencia de contracción del recto y aumentar el tono basal de las glándulas salivales (Orchard y Brugge, 2002). En el buche, fue capaz de aumentar la frecuencia de contracción, disminuir la resistencia y el voltaje transepiteliales. (Te Brugge *et al*, 2009). Sin embargo, ha demostrado no tener actividad sobre la tasa de secreción y en la composición del fluido secretado por los TM (Donini *et al*, 2008).

Las kininas de insectos comparten la presencia del pentapéptido FX_1X_2WG-NH_2 (X_1: F, H, N, S o Y; X_2: A, P o S). Ésta es la secuencia mínima necesaria para mantener actividad biológica (miotrópica y diurética), la cual adopta una estructura de giro-β tipo 4 en la estructura secundaria. Dentro de este sitio activo, G^1, W^2 y en particular F^5 son los aminoácidos más importantes para la actividad (Nachman *et al*, 2003; 1991; 1993). El dipéptido amidado formado por G^1 y W^2, junto con un extremo amino terminal que mimetiza el anillo de F^5, han demostrado ser el requerimiento mínimo necesario para una unión eficiente con el receptor. F^5 y W^2 se asocian para formar un anillo aromático lipofílico que interacciona con el sitio activo del receptor, mientras que G^1 amidado refuerza la interacción al formar un puente de hidrógeno mediante el grupo amino (ver Scherkenbeck, 2009; figura 1.27).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

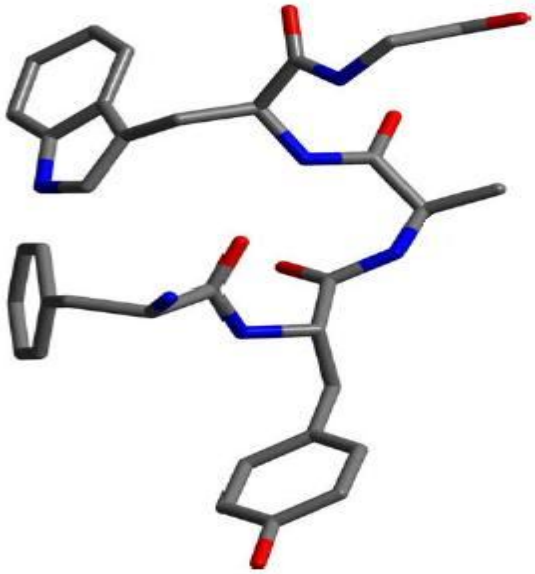


Figura 1.27: Modelo de estructura de péptido FYAWG-NH_2 al interaccionar con el receptor. Imagen tomada de Scherkenbeck *et al*, 2009

El conocimiento detallado de la interacción entre el neuropéptido y su receptor es un pre-requisito indispensable para el diseño racional de peptidomiméticos. Nosotros hemos encontrado un GPN en el genoma de *R. prolixus* que codifica el precursor de Kininas. De él, se deducen seis kininas con la secuencia C-terminal clásica FX_1X_2WG-NH_2 (X_1: S, N, Y; X_2: S, A), y cinco kininas con el pentapéptido C-terminal FSX_2WA-NH_2 (X_2: S, H). Este motivo es único a nivel se secuencia entre las Kininas descriptas para insectos. El precursor codifica además doce péptidos espaciadores (Figura 1.29). La transcripción de este gen no pudo ser confirmada mediante RACE-PCR, por lo que se desconocen los UTRs, pero la expresión de varios de los péptidos codificados en el precursor fue confirmada en SN por MS (Tabla 1.12).

El mayor desafío en la generación de peptidomiméticos con actividad insecticida es aumentar su estabilidad y capacidad de penetración de la cutícula. Se propone que esto puede lograrse reemplazando los aminoácidos naturales por moléculas no peptídicas que permitan mantener (o aumentar) la actividad biológica. Análogos miméticos para LK ya han sido identificados y han demostrado ser capaces de interferir los procesos que estos péptidos regulan (ver Nachman *et al*, 2009; Taneja-Bageshwar, 2009). Por lo tanto, las nuevas secuencias de las Kininas que hemos descripto en *R. prolixus* pueden ser un posible blanco para el desarrollo de nuevos insecticidas, más específicos y que generen menor daño ecológico, para el control de la población de este insecto vector.

Mediante búsquedas bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de Kininas caracterizados en otros insectos (BLASTN *E value*: E%=2e-112).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figure 1.28: Estructura del gen de Kininas. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Para este gen solo se conoce el ORF, no se conocen los UTRs, por lo que la longitud de los exones es mayor.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

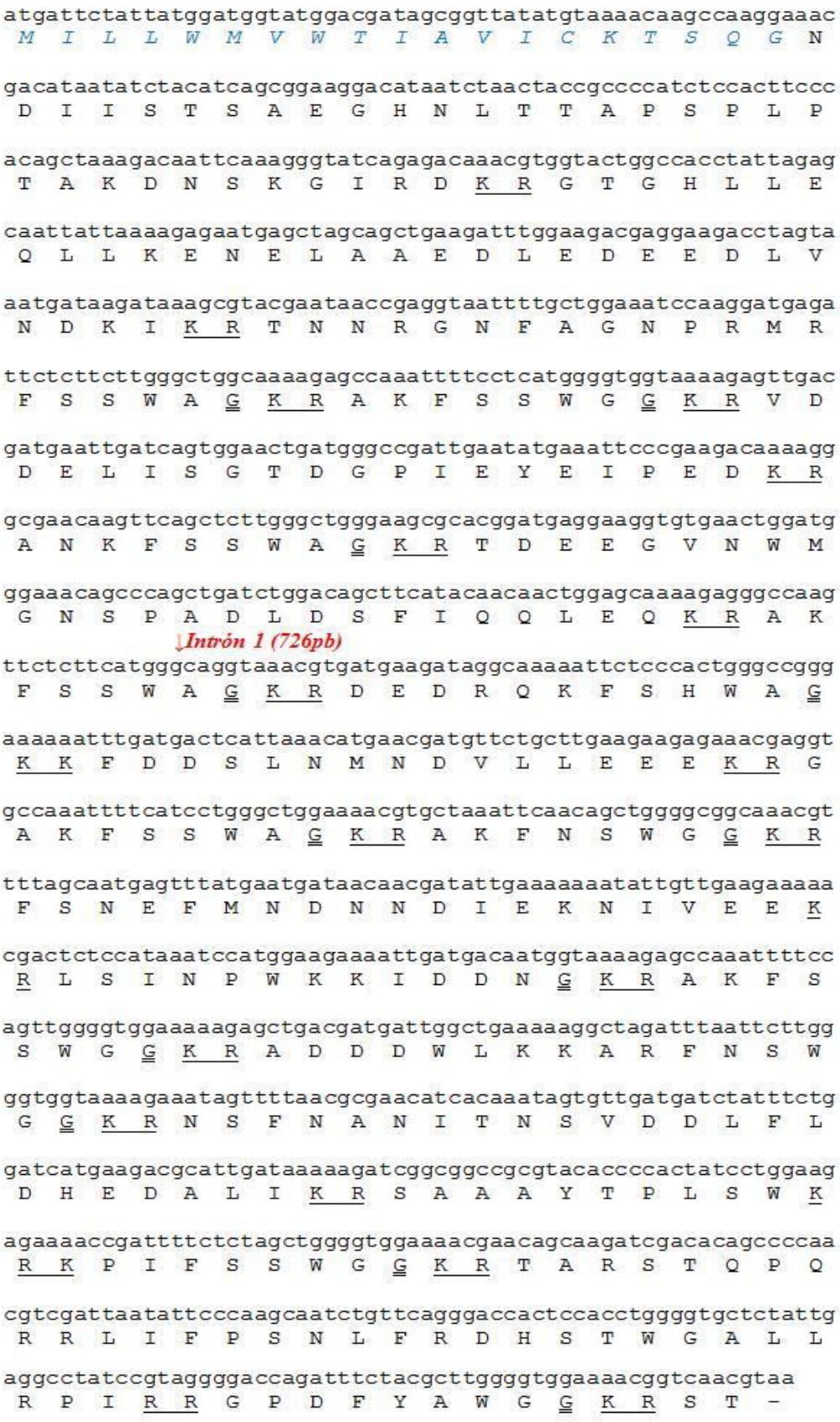


Figura 1.29: Secuencia nucleotídica y aminoacídica de *Rhopr-kininas*. El péptido señal (probabilidad SignalP-HMM: 0.983) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Tabla 1.12: Péptidos predichos a partir del clivaje proteolítico del precursor de *kininas* y péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
NDIISTSAEGHNLTTAPSPLPTAKDNSKGIRD		3320,67	
GTGHLLEQLLKENELAAEDLEDEEDLVNDKI	GTGHLLEQLLKENELAAEDLEDEEDLVNDKI	3492,72	Cerebro
TNNRGNFAGNPRMRFSSWA-NH_2		2182,05	
AKFSSWG-NH_2	AKFSSWG-NH_2	781,4	SN
VDDELISGTDGPIEYEIPED		2205,99	
ANKFSSWA-NH_2		909,46	
TDEEGVNWMGNSPADLDSFIQQLEQ		2823,24	
AKFSSWA-NH_2		795,41	
DEDRQKFSHWA-NH_2	DEDRQKFSHWA-NH_2	1417,66	Cerebro
FDDSLNMNDVLLEEE		1782,77	
GAKFSSWA-NH_2		852,44	
AKFNSWG-NH_2		808,41	
FSNEFMNDNNDIEKNIVEE	FSNEFMNDNNDIEKNIVEE	2300.10	Cerebro
LSINPWKK IDDN-NH_2		1441,78	
AKFSSWG-NH_2	AKFSSWG-NH_2	781,4	SN
ADDDWLKK ARFNSWG-NH_2		1807,88	
NSFNANITNSVDDLFLDHEDALI		2577,21	
	FLDHEDALI	1072,53	SN
SAAAYTPLSW		1066,52	
KPIFSSWG-NH_2		920,40	
TARSTQPQRR LIFPSNLFRDHSTWGALLRPI		3634,98	
GPDFYAWG-NH_2		911,40	

Péptido Transportador de Iones

El péptido transportador de iones (ITP) fue aislado por primetra vez a partir de la Corpora Cardíaca en base a su capacidad de estimular el transporte activo de Cl_2 en el recto de S. gregaria (Audsley et al, 1992)

En nuestro trabajo hemos identificado dos formas alternativas de splicing para ITP en R. prolixus, al igual que en T. castaneum y S. gregaria. La isoforma A fue obtenida mediante RACE-PCR 3´, mientras que la isoforma B fue encontrada en la base de datos de EST total. Llamativamente, los exones 4 y 5 detectados por transcriptómica no se han encontrado en la base de datos genómica. Probablemente la región que contiene estas secuencias se encuentre en el porcentaje del genoma que no ha sido secuenciado (Figura 1.30).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

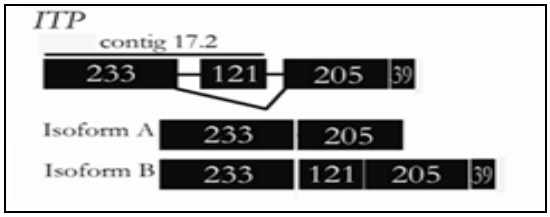
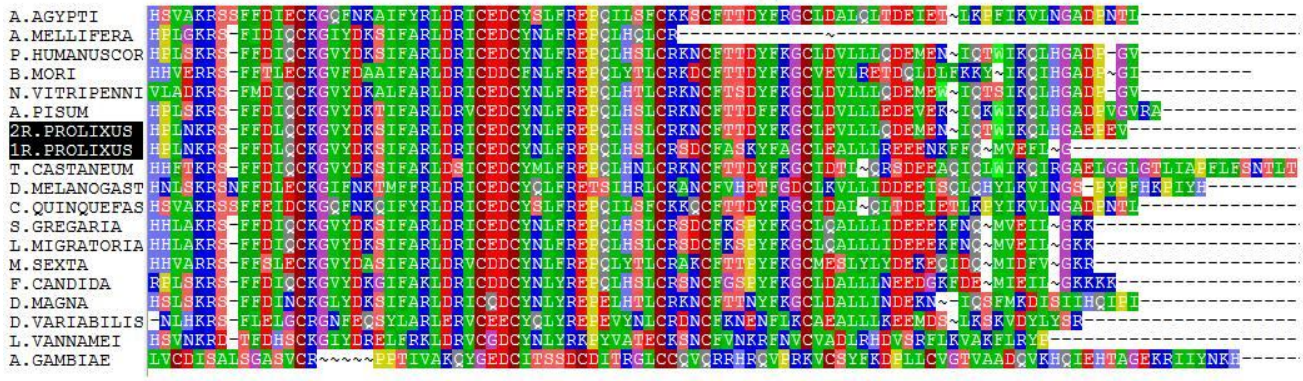


Figure 1.30: Estructura del gen de *Rhopr-ITP* y las dos isoformas encontradas. El número en cada exón indica su longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones. Para este gen no se conoce el UTR 5´, por lo que el exón 1 es en realidad de mayor longitud.

En *S. gregaria* se ha demostrado que los residuos Fen en en posiciones 2 y 3 del extremo N-terminal son esenciales para la activación de su receptor, así como la Leu en el extremo C-terminal (Zhao *et al,* 2005). En el alineamiento se pueden apreciar seis residuos Cys altamente conservados (Figura 1.31). Los puentes disulfuro que se forman entre ellos resultan críticos para el correcto plegamiento de la molécula, y por lo tanto para su correcta interacción con el receptor



Fugura 1.31: Alineamiento de los precursores de las isoformas A y B de *Rhopr-ITP* con los precursores ortólogos de distintas especies de insectos.

Figura 1.32: : Secuencia nucleotídica y aminoacídica de la isoforma A *Rhopr-ITP.* El péptido señal (probabilidad SignalP-HMM: 0.989) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones y los exones ausentes en esta isoforma se muestra en rojo

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Tabla 1.13: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ITP A* y los péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
GPSSRLVLSHPLN		1376,76	
SFFDLQCKGVYDKSIFARLDRICEDCYNLFREPQ-LHSLCRSDCFASKYFAGCLEALLLREEENKFFQMVEFL-NH_2		8586,12	

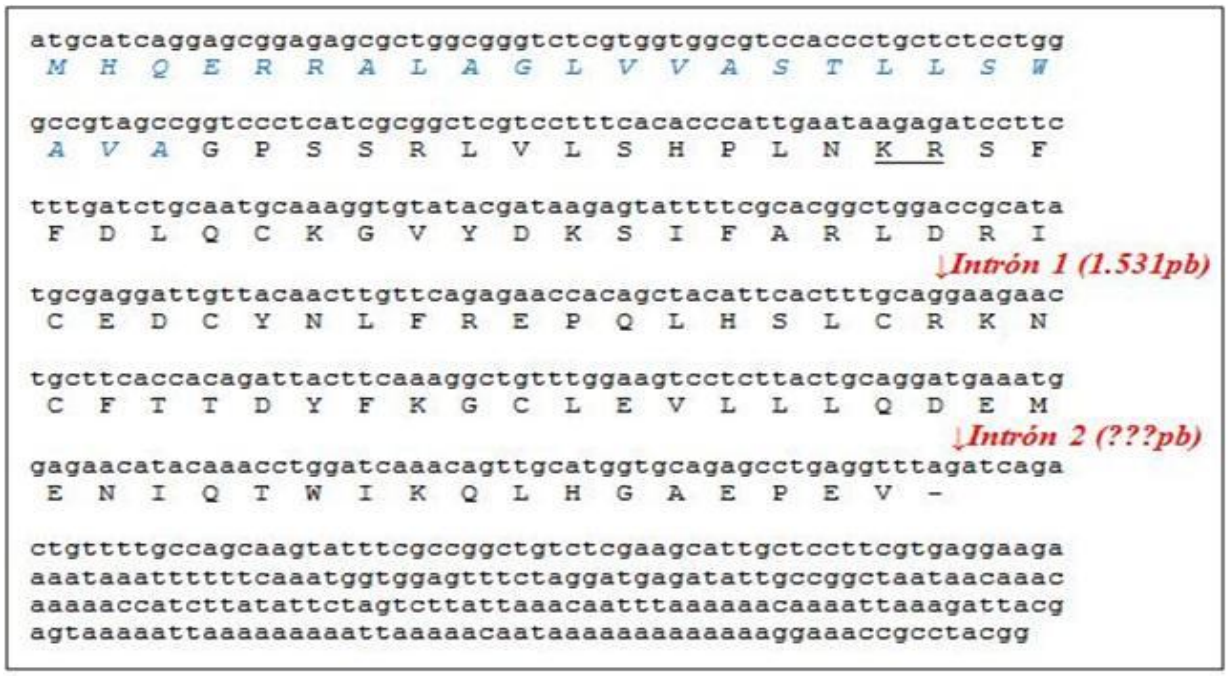


Figura 1.33: Secuencia nucleotídica y aminoacídica de la isoforma B *Rhopr-ITP*. El péptido señal (probabilidad SignalP-HMM: 0.989) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones y los exones ausentes en esta isoforma se muestra en rojo.

Tabla 1.14: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ITP B* y los péptidos maduros encontrados por MS, $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
GPSSRLVLSHPLN		1376,76	
SFFDLQCKGVYDKSIFARLDRICEDCYNLFREPQLHSLC - RKNCFTTDYFKGCLEVLLLQDEMENIQTWIKQLHGAEPEV		9392,53	

Neuroparsina-A

Se ha propuesto un papel antidiurético de Neuroparsina-A en *L migratoria* y *S gregaria* (Fournier *et al*, 1994; Girardie *et al*, 1998). En nuestro trabajo hemos logrado caracterizar el mRNA para este gen en toda su longitud mediante búsqueda en la base de datos de ESTs de digestivo medio anterior, confirmando la predicción bioinformática a partir del genoma. La expresión en digestivo medio anterior sugiere que en *R. prolixus*

Neuroparsina también estaría asociada a la regulación de la diuresis, actuando de manera autócrina o parácrina sobre la actividad de éste órgano.

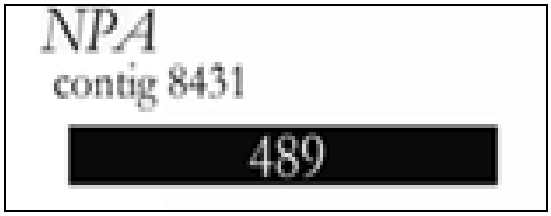


Figure 1.34: Estructura del gen de *Rhopr-NPA*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Para este gen no se conoce el UTR 5´ completo, por lo que el exón es en realidad de mayor longitud.

```
atcatacaatctcgagatgtcgtcccaatcgtccaaaacagcaactactgctttagctgtt
                  M  S  S  Q  S  S  K  T  A  T  T  A  L  A  V

ttaaccattttctgtatggtggcgttagtctcgggagtattctacggctgtgtaccatgt
 L  T  I  F  C  M  V  A  L  V  S  G  V  F  Y  G  C  V  P  C

atcggcgatgaatgtaatttgaatccaggaaattgtccgtacggtattgttcgcgatccg
 I  G  D  E  C  N  L  N  P  G  N  C  P  Y  G  I  V  R  D  P

tgtggaagattggtttgcgctgctggacctggtgaacgttgcggcggtcgagacttccat
 C  G  R  L  V  C  A  A  G  P  G  E  R  C  G  G  R  D  F  H

ctgggcaaatgtggtgaaggactgtcatgtaagtgcggcaaatgcagaggatgttcaatt
 L  G  K  C  G  E  G  L  S  C  K  C  G  K  C  R  G  C  S  I

aaacaaattatgaacggacgtatagattgtgacactacaaatcctatgtgccaataatat
 K  Q  I  M  N  G  R  I  D  C  D  T  T  N  P  M  C  Q  -

tctttcacgaagtaaatgtaacaaataaaatctttccccaaaatgtagattaaaaaacaa
aattatagacaaaaattgaaataaatttatttatta
```

Figura1.35: Secuencia nucleotídica y aminoacídica de *Rhopr-NPA*. El péptido señal (probabilidad SignalP-HMM: 0.809) se muestra en azul y letra cursiva.

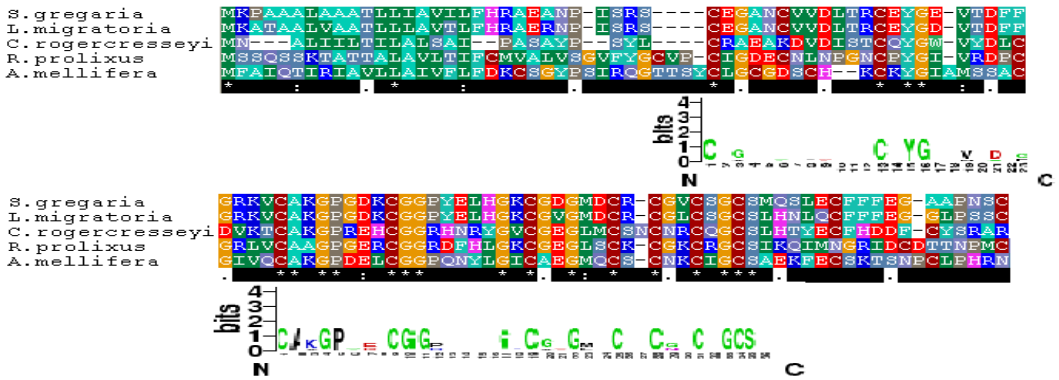


Figura 1.36: Alineamiento del precursor de *Rhopr-NPA* con los precursores ortólogos caracterizados en distintas especies de insectos. Se pueden observar varios residuos Cys conservados. El contenido informativo expresado en bits se muestra para cada posición

1.3.2.c) Hormonas asociadas al desarrollo y ecdisis

Durante el proceso de ecdisis, los artrópodos cambian el exoesqueleto al final de un período de crecimiento. Para esto se forma un nuevo exoesqueleto por debajo de la capa a ser cambiada, seguido por la separación de las capas. Este proceso se puede dividir en tres etapas: pre-ecdisis, ecdisis y post-ecdisis. La hormona esteroide 20-hidroxyecdisona es el principal factor iniciador de la ecdisis (Truman , 2005; ver Zitnan *et al*, 2007), mientras que el proceso es regulado por una compleja red de sistemas de señalización peptídicos. Estudios en organismos como *D. melanogaster* (Truman, 2005), *M. sexta* (Ewer *et al*, 2005a) y *T. castaneum* (Arakane *et al* , 2008) mostraron que, si bien las hormonas involucradas son las mismas, Corazonina (CZ) (exepto en coleópteros), la Hormona Disparadora de la Ecdisis (ETH), la Hormona de Eclosión (EH),el Péptido Cardioacelerador de Crustáceos (CCAP) y Bursicon, la función de las distintas moléculas no es igual entre los diferentes Órdenes estudiados, todos ellos holometábolos. Los fenotipos obtenidos mediante RNAi en *T. castaneum* (Arakane *et al*, 2008) difieren significativamente de los obtenidos en *D. melanogaster* (Truman *et al*, 2005), lo que sugiere que la función de estos sistemas de señalización ha sufrido cambios durante la evolución de los artrópodos.

Nosotros hemos encontrado cuatro GPNs asociados al desarrollo y la ecdisis en el genoma de *R. prolixus*: CZ, ETH, EH y CCAP**.** También se ha encontrado parte del ORF que codifica Bursicon, aunque este no es un péptido, sino una proteína conformada pos dos subunidades. El estudio de estos genes en *R. prolixus*, aportará valiosa información acerca del proceso de evolución que ha dado lugar al amplio rango funcional de sistemas de señalización que regulan la ecdisis.

Corazonina

Corazonina es un neuropéptido ampliamente distribuido en insectos con excepción del Órden Coleoptera. Los análisis filogenéticos han demostrado que CZ y su receptor presentan homología con la Hormona Adipokinética (AKH) y su receptor, por lo que se ha sugerido que estos sistemas evolucionaron a partir de la duplicación y mutación de un sistema de señalización ancestral. Posteriormente, mediante el mismo proceso, se habría originado el sistema de señalización "*AKH-CZ Related Peptide"* (ACP), reportado recientemente por primera vez (Hansen *et al*, 2010).

Se ha demostrado que CZ presenta actividad cardioaceleradora en *L. migratoria.* (Veenstra *et al*, 1989), mientras que en *S. gragaria* muestra efectos pigmentotrópicos e influye sobre el cambio de la fase solitaria a la fase de agregación (Tawfik *et al,* 1999) y sobre la muda (Zitnan *et al*, 2007). En *D. mallanogaster y M. Sexta* este péptido desencadena una cascada de señalización que da como resultado la liberación de ETH y el comienzo del proceso de ecdisis (Kim y Rulifson, 2004). Recientemente, se ha determinado el patrón de expreción de CZ en el SN de *T. infestans* y se ha secuenciado el

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

péptido maduro mediante MS, el cual resultó ser igual al que hemos encontrado en *R. prolixus* (Settembrini, 2010).

Mediante RACE-PCR hemos obtenido el ORF completa del mRNA que codifica el precursor de este neuropéptido. En todos los precursores de CZ descriptos hasta la fecha, la estructura es la misma, con el péptido biológicamente activo inmediatamente después del péptido señal, seguido por un péptido espaciador en el extremo 3´. El péptido biológicamente activo y el acompañante se encuentran separados por un par de residuos básicos donde se produce el clivaje proteolítico. El péptido maduro se encuentra altamente conservado en todas las especies donde se ha descripto, excepto en *A. mellifera*, (Figura 1.39). Hemos podido confirmar la presencia de CZ, tanto en SN como en hemolinfa de *R. prolixus*, mediante técnicas de MS. La presencia en hemolinfa sugiere una función hormonal para *Rhopr-CZ*.

Resulta interesante destacar la presencia del mRNA que codifica CZ en la base de datos de EST de testículos de *R. prolixus.* (Figura 1.38). El receptor de este neuropéptido presenta similitud de secuencia con el receptor de la familia de la hormona liberadora de gonadotropinas (GRH), de esto ha surgido la hipótesis de una función ancestral de CZ relacionada al control de la reproducción (Kim and Rulifson, 2004). La transcripción de CZ en testículo refuerza esta hipótesis, y sienta las bases para posteriores ensayos con el fin de determinar su función presisa.

In silico se ha podido predecir un GPCR que presenta un alto grado de homología con los receptores de CZ caracterizados en otros insectos (BLASTN *E value*: E%=1e-117).

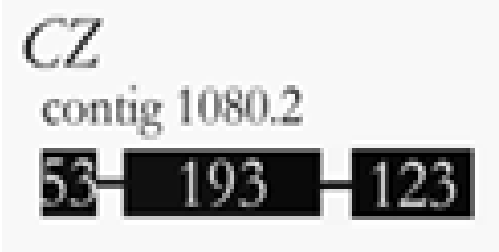


Figure 1.37: Estructura del gen de *Rhopr-CZ*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Dado que se desconoce el UTR 3´, el exón 3 está imcompleto, por lo que es mas largo de lo que se muestra en la figura.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 1.38: Secuencia nucleotídica y aminoacídica de *Rhopr-CZ.* El péptido señal (probabilidad SignalP-HMM: 0.809) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.15: Péptidos predichos a partir del clivaje proteolítico del precursor de Rhopr-CZ y los péptidos maduros encontrados por MS. [M+H]+masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
QTFQYSRGWTN-NH_2	**pyro-QTFQYSRGWTN-NH_2**	1369,65	Cerebro, Hemolinfa
GIPSKEVTACQLQRIKSLLEGKTIPQLYWPCEWSPFM-EAALSRQMKTSELTSLPVVAPLTPEIEEK		7425,865	

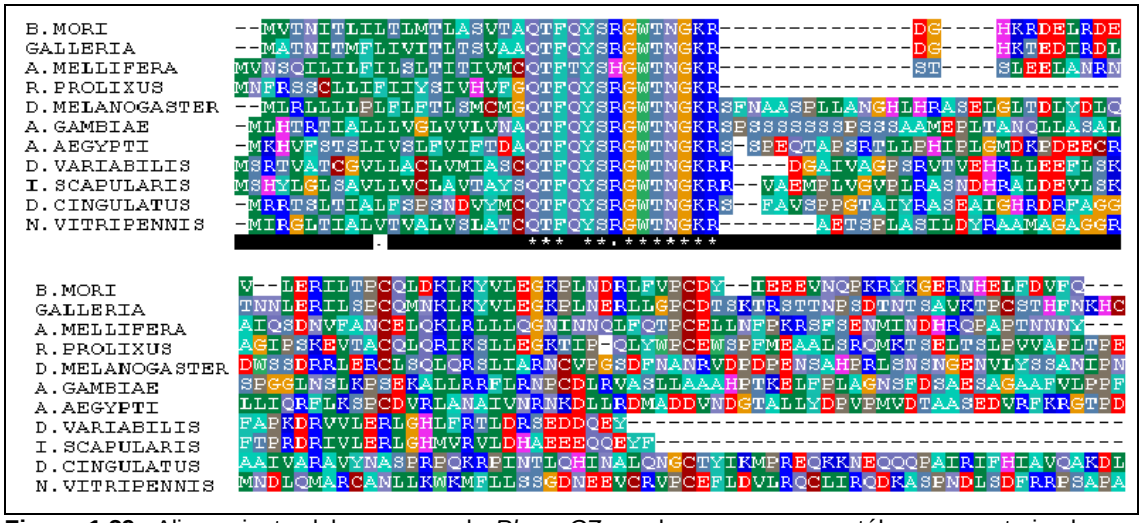


Figura 1.39: Alineamiento del precursor de *Rhopr-CZ* con los precursores ortólogos caracterizados en distintas especies de insectos. Se pueden observar el péptido activo es igual en todos los insectos donde se ha descripto el precursor excepto en *A. mellifera*.

Hormona Disparadora de la Ecdisis

Esta hormona es uno de los factores que desencadena el proceso de ecdisis tanto en *D. mallanogaster* como en *T. castaneum* , y ha demostrado ser necesario para que se

lleven a cabo los procesos de pre-ecdisis y ecdisis (Park *et al*, 2002; Arakane *et al* , 2008). La mayoría de los genomas de insectos conocidos poseen un solo gen que codifica esta familia de péptidos, mientras que *T. castaneum* tiene dos. En este insecto, inyecciones de dsRNA con el fin de bloquear la expresión tanto de ETH como de su receptor, presentaron un fenotipo letal, en el cual el proceso de ecdisis se detenía en la etapa pre-ecdisial. Un fenotipo similar se obtuvo al bloquear la expresión de EH (Arakane *et al* , 2008)

En *R. prolixus* hemos encontrado un EST que codifica este péptido (Figura 1.40). Aunque el ORF completo no está presente en la base de datos, la región identificada es suficiente para demostrar su transcirpción.

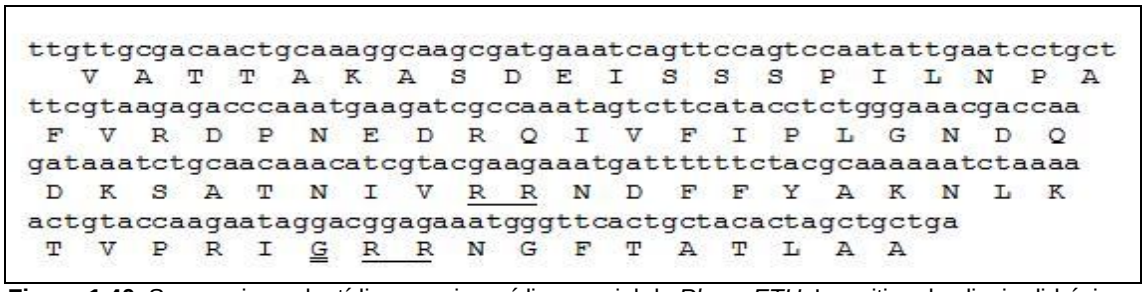


Figura 1.40: Secuencia nucleotídica y aminoacídica parcial de *Rhopr-ETH*. Los sitios de clivaje di-básicos se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado.

Tabla 1.16: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ETH* y los péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

peptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
...VLQSGNQPRPGCKLFISFNQFSECFLTIVATTAKASDEISSSPIL-NPAFVRDPNEDRQIVFIPLGNDQDKSATNIV			
NDFFYAKNLKTVPRI-NH_2		1825,01	
NGFTATLAADGRSGARVD		1778,89	

Hormona de Eclosión

EH ha demostrado ser necesario para que se lleve a cabo el proceso de ecdisis en *T. castaneum* (Arakane *et al*, 2008).

Mediante herramientas bioinformáticas hemos podido caracterizar el ORF completo de *Rhopr-EH* a partir del genoma de *R. prolixus* (Figura 1.41). La secuencia del precursor se encuentra altamente conservada en todos los insectos donde se ha descripto (Figura 1.43). En particular se puede observar la presencia de seis residuos Cys altamente conservados que estabilizan la estructura del precursor mediante puentes disulfuro. Dos de estos residuos se encuentran en el péptido biológicamente activo, y son importantes para su correcto plegamiento y estabilidad (Figura 1.43). Mediante MS hemos detectado la presencia del péptido maduro con actividad biológica codificado por este precursor en extractos de cerebro, confirmando se expresión (tabla 1.17).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figure 1.42: Estructura del gen *Rhopr-EH*. El número indica la longitud en nucleótidos. El *contig* en que se el gen se indica en la parte superior. Para este gen no se conocen los UTRs, por lo que su longitud real es mayor.

```
atgaagaagctgctgttggtgattctattgacgtccttcctagctgaaatcagtggccgc
 M  K  K  L  L  L  V  I  L  L  T  S  F  L  A  E  I  S  G  R
cagattggcgtgtgcatcaggaactgtgcccagtgtaagaaaatgtttggtgtctatttt
 Q  I  G  V  C  I  R  N  C  A  Q  C  K  K  M  F  G  V  Y  F
gaagggcagatgtgcgctgacacttgtctcaaatataagggcaaattgattccagactgt
 E  G  Q  M  C  A  D  T  C  L  K  Y  K  G  K  L  I  P  D  C
gaggacattgcctcaattggaccatttctaaataagttaatatga
 E  D  I  A  S  I  G  P  F  L  N  K  L  I  -
```

Figura 1.41: Secuencia nucleotídica y aminoacídica de *Rhopr-EH*. El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.17: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ETH* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
RQIGVCIRNCAQC	RQIGVCIRNCAQC	1463,70	Cerebro
MFGVYFEGQMCADTCLKYKGKLIPDCEDIASIGPFLNKL		4358,11	

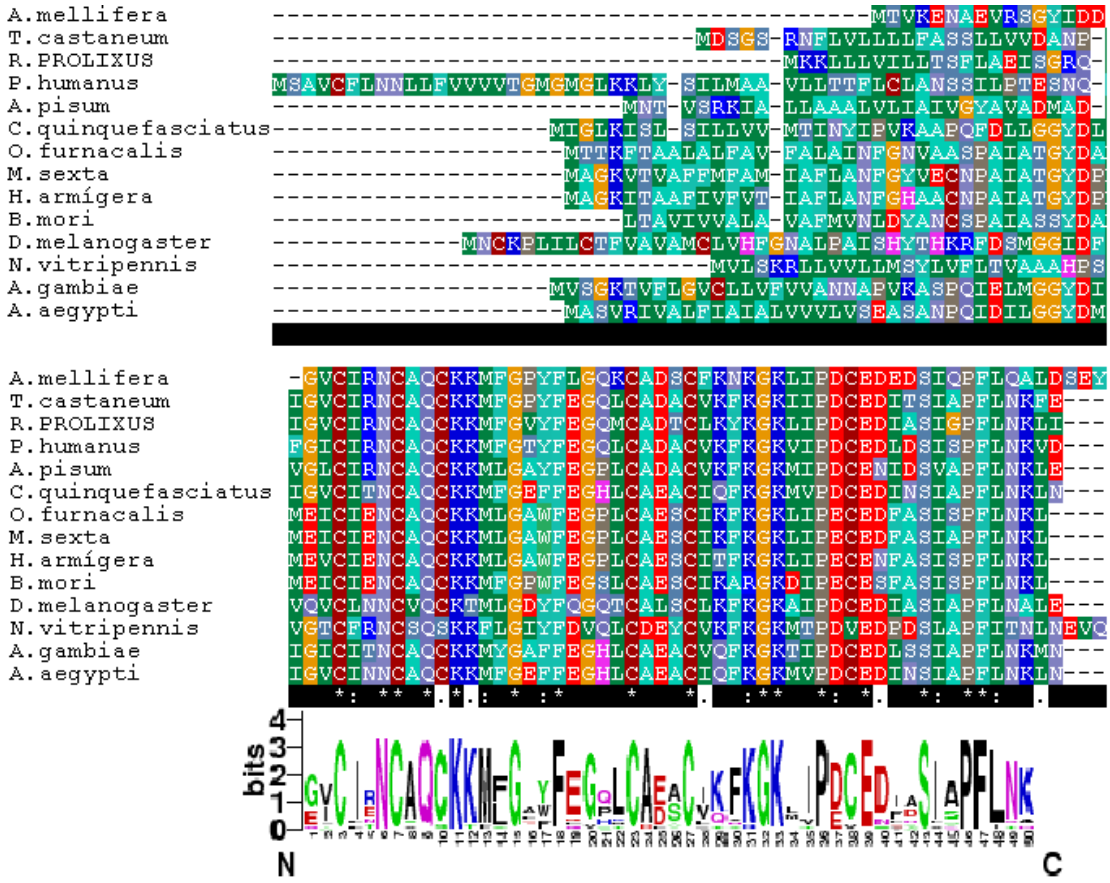


Figura 1.43: Alineamiento del precursor de *Rhopr-EH* con los precursores ortólogos caracterizados en distintas especies de insectos. Se pueden observar seis residuos Cys conservados. El contenido informativo expresado en bits se muestra para cada posición

Péptido Cardioactivo de Crustáceos

En *T. castaneum,* CCAP ha demostrado ser crítico para el proceso de ecdisis, ya que cuando su expresión fue bloqueada mediante RNAi, los insectos no fueron capaces de mudar normalmente y presentaron un fenotipo letal. Durante el proceso de pre-ecdisis mientras tanto, no se observaron diferencias con el grupo control. (Arakane *et al*, 2008). También se ha involucrado a CCAP en la cascada de señalización que regula la ecdisis en *M. sexta* (Kim *et al*, 2006) y *D. melanogaster* (ver Ewer, 2005b). Al igual que muchos otros neuropéptidos, CCAP ha resultado ser mutifuncional; es cardioestimulador en *M. sexta* (Tublitz, 1985) y capaz de aumentar la frecuencia de contracción de los ovarios en *L. migratoria* (Donini, 2001).

En *R. prolixus* hemos encontrado dos formas alternativas de *splicing* para CCAP. La isoforma A fue obtenida mediante RACE-PCR, mientras que la isoforma B se encontró en la base de datos total de ESTs. Llamativamente, la secuencia proteica codificada por estas dos isoformas es igual, aunque difieren en el UTR 3´. Mientras que la isoforma A posee el exón 6, la isoforma B posee el exón 7 (Figura 1.44). Esto podría estar asociado a la regulación de la expresión de cada isoforma (Figuras 1.46 y 1.47). El péptido que presenta actividad biológica conocida se encuentra totalmente conservado en todas las

especies donde se ha descripto (Figura 1.45), lo que implica que está sometido a una gran presión de selección. Las demás regiones del precursor presentan mayor variabilidad.

Mediante búsquedas bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de CCAP caracterizados en otros insectos (BLASTN *E value*: E%=1e-125).

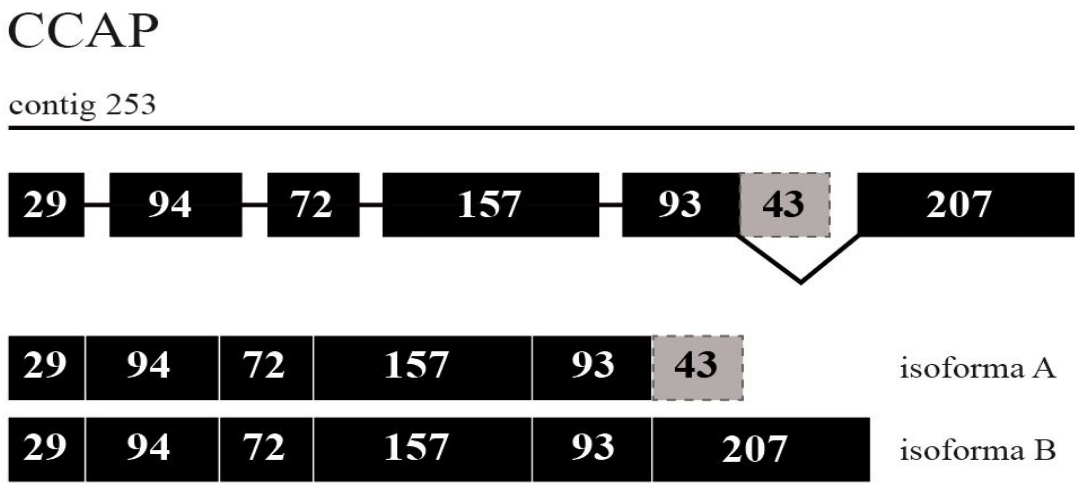


Figure 1.44: Estructura del gen de *Rhopr-CCAP* y las dos formas de splicing encontradas. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones

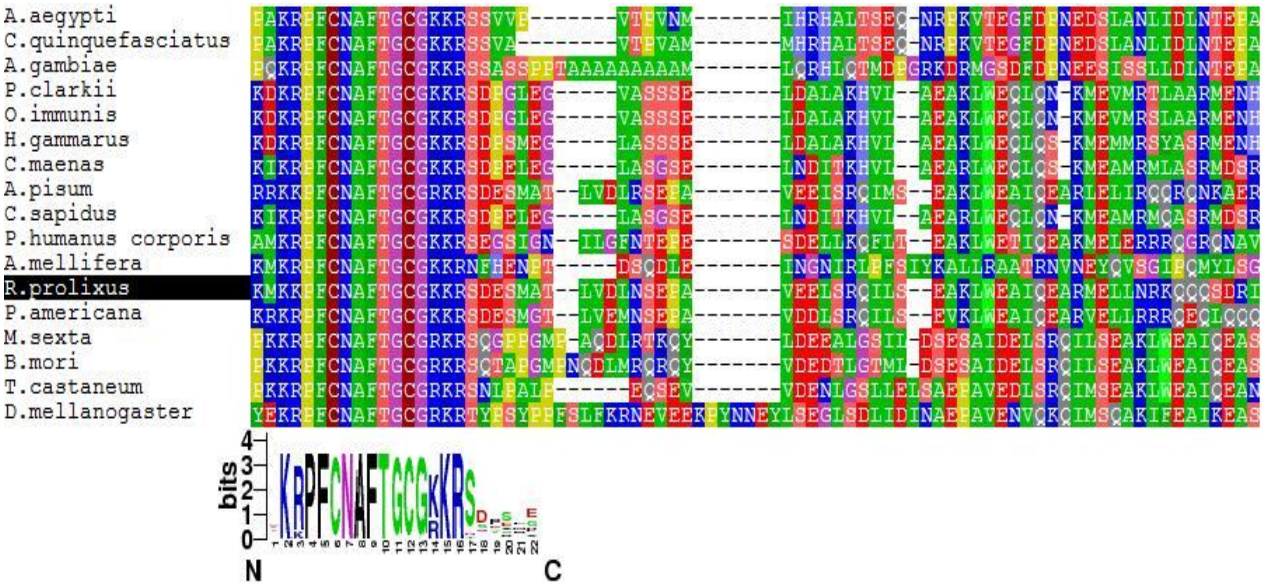


Figura 1.45: Alineamiento del precursor de *Rhopr-CCAP* con los precursores ortólogos caracterizados en distintas especies de insectos. Se pueden observar dos residuos Cys conservados en el péptido activo. El contenido informativo expresado en bits se muestra para cada posición..

Figura 1.46: Secuencia nucleotídica y aminoacídica de la isoforma A *Rhopr-CCAP*. El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

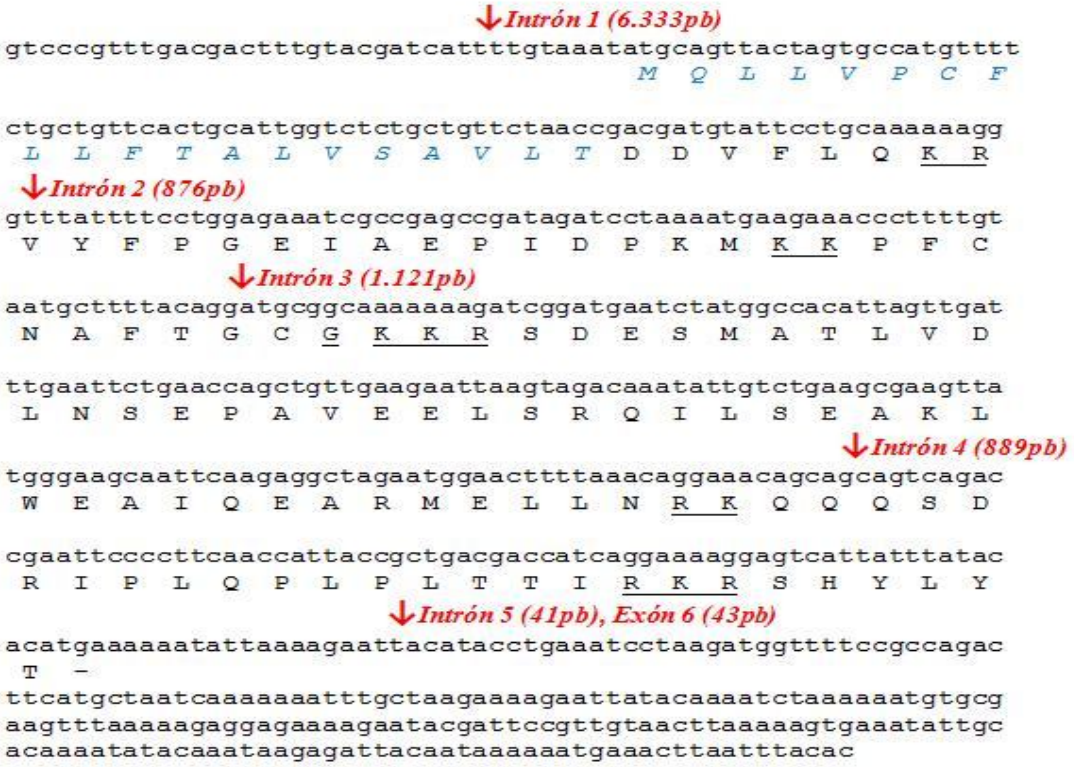


Figura 1.47: Secuencia nucleotídica y aminoacídica de la isoforma A *Rhopr-CCAP*. El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido

maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.18: Péptidos predichos a partir del clivaje proteolítico de los precursores A y B de *Rhopr-CCAP* y los péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
DDVFLQ		1049,55	
VYFPGEIAEPIDPKM		1705,85	
PFCNAFTGC-NH_2	PFCNAFTGC-NH_2	958,39	Cerebro
SDESMATLVDLNSEPAVEELSRQILSEAKLWEAIQEARMELLN		4858,41	
QQQSDRIPLQPLPLTTI		1948,09	
SHYLYT		783,37	

1.3.2.d) Péptidos Allatoregulatorios

Los péptidos allatoregulatorios fueron identificados inicialmente por su capacidad de estimular (Allatotropina; AT) o inhibir (Allatostatinas; ATS) la síntesis de Hormonas Juveniles (JH) por parte de la *Corpora Allata*. Se conocen tres familias de Allatostatinas: A, B y C, todas capaces de inhibir la síntesis de JH (ver Stay y Tobe, 2007; Weaver y Audsley, 2009). Además, estos péptidos han demostrado tener actividad sobre la musculatura del tracto digestivo de distintos insectos, tanto estimulando (AT) como inhibiendo (ATS) la contracción muscular (Fusé et al, 1999; Matthews et al, 2007).

Allatostatina-A

El primer péptido de la familia de las ATS fue descubierto en 1989 en *D. punctata* (Woodhead *et al*, 1989). Posteriormente, varios péptidos de este tipo fueron caracterizados en distintos órdenes de insectos. Todos ellos comparten el dominio característico X_1XFGLamida (X_1: F o Y) en el extremo C terminal, secuencia mínima requerida para que el péptido sea capaz de ejercer su función biológica (Matthews *et al*, 2007). Se ha demostrado la capacidad por parte de ATS-A de inhibir la producción de JH en la *Corpora Allata* de varios insectos (Edwards *et al*, 2001; Maestro *et al*, 2010), con la excepción de *A. mellifera* (Rachinsky *et al*, 2000), donde tampoco fue detectado ningún péptido de la familia de las ATS en el complejo retrocerebral (Audsley and Weaver, 2006; Boerjan *et al*, 2010).

Utilizando un suero policlonal desarrollado contra Allatostatina 7 de *D. punctata* se observó inmunomarcación positiva en SNC, complejo retrocerebral y células endócrinas del digestivo medio de *R. prolixus*. En este insecto, Dippu-ATS 7 mostró además ser capaz de inhibir las contracciones espontaneas del recto y la frecuencia de contracción de la arteria dorsal (Sarkar et al, 2003).

En nuestro trabajo hemos podido caracterizar el GPN de Allatostatinas-A de *R. prolixus (Rhopr-ATS-A)*. A partir de este GPN se deducen trece péptidos maduros, seis

de los cuales presentan el motivo característico de la familia. También se deduce un péptido amidado (LTELGL-NH_2) que comparte el motivo GL-amida con el resto de las ATS-A (Figura 1.48). Los otros seis péptidos que se generan del clivaje proteolítico del precursor son péptidos espaciadores. De los trece péptidos que pueden predecirse, 8 fueron detectados mediante MS, tanto en muestras de cerebro como en muestras del SN completo (Tabla 1.19).

Se ha detectado *in silico* un GPCR que presenta un alto grado de homología con los receptores de ATS-A caracterizados en otros insectos (BLASTN *E value*: E%=4e-128).

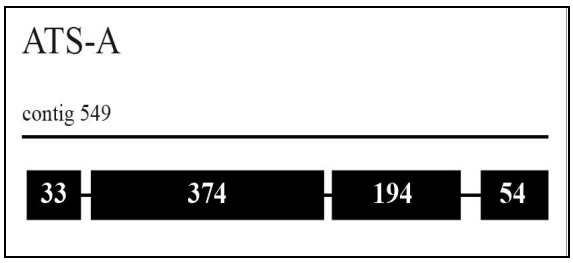


Figure 1.49: Estructura del gen de *Rhopr-ATS A*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Dado que solo se obtuvo el RACE 5´, no se ha podido caracterizar el UTR 3´, por lo que el exón 4 está incompleto, siendo su longitud mayor que la indicada.

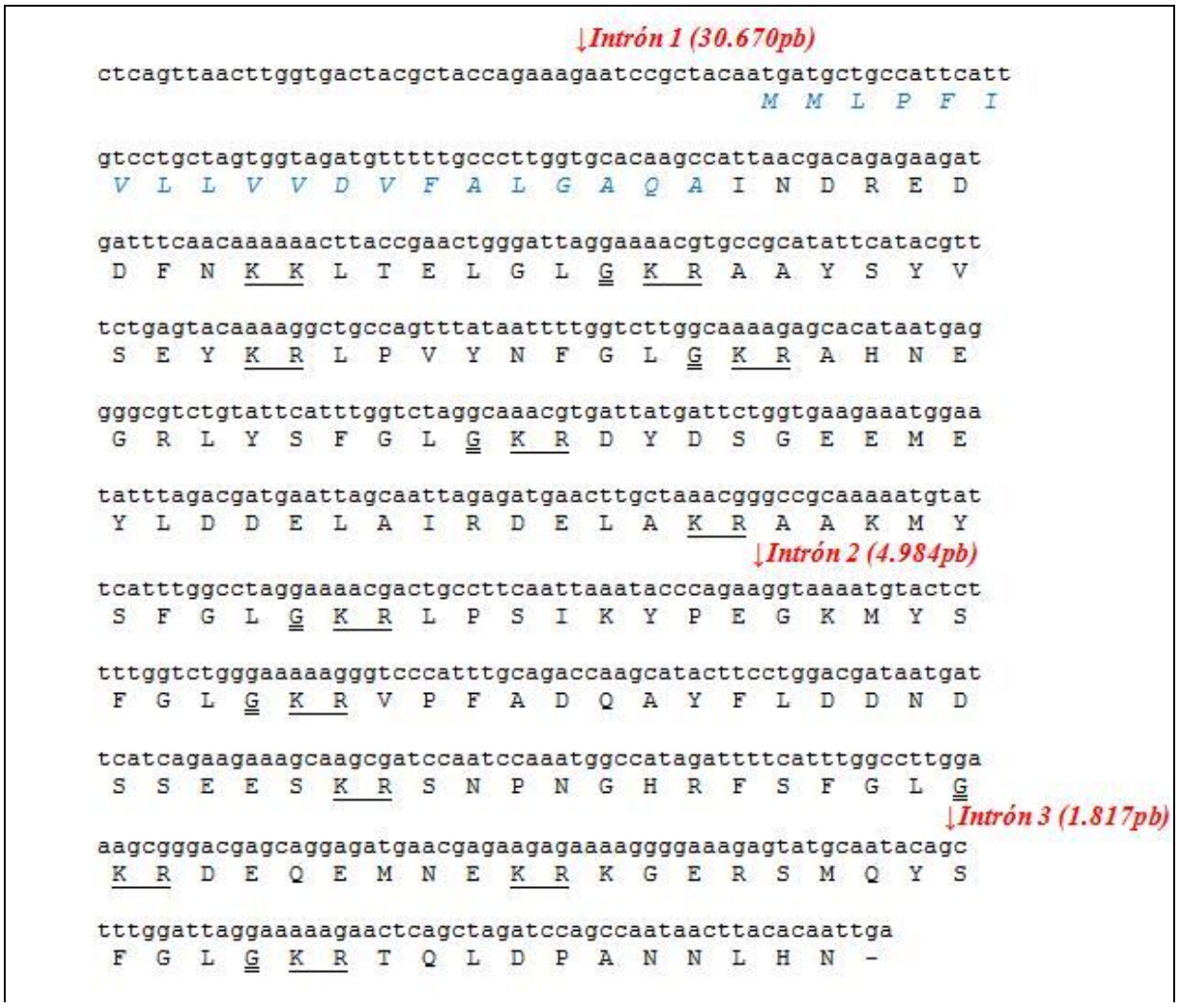


Figura 1.48: Secuencia nucleotídica y aminoacídica de *Rhopr-ATS A* . El péptido señal (probabilidad SignalP-HMM: 0.997) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran

subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos. Dado que sólo se obtuvo el extremo 5´ por RACE-PCR no se conoce el UTR 3´.

Tabla 1.19: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ATS A* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
INDREDDFN		1137,48	
LTELGL-NH_2		644,4	
AAYSYVSEY	AAYSYVSEY	1052,46	Cerebro
LPVYNFGL-NH_2	LPVYNFGL-NH_2	921,51	Cerebro
AHNEGRLYSFGL-NH_2	AHNEGRLYSFGL-NH_2	1362,69	SN, Cerebro
DYDSGEEMEYLDDELAIRDELA		2591,1	
AAKMYSFGL-NH_2	AAKMYSFGL-NH_2	986,52	Cerebro
LPSIKYPEGKMYSFGL-NH_2	LPSIKYPEGKMYSFGL-NH_2	1828,95	Cerebro
VPFADQAYFLDDNDSSEES	VPFADQAYFLDDNDSSEES	2148,89	Cerebro
SNPNGHRFSFGL-NH_2	SNPNGHRFSFGL-NH_2	1331,66	SN, Cerebro
DEQEMNE		894,32	
GERSMQYSFGL-NH_2		1273,6	
TQLDPANNLHN	TQLDPANNLHN	1236,6	SN, Cerebro

Allatostatina-B

Las Allatostatinas B de insectos tienen en común la secuencia $W(X)_6$GWamida en el extremo C terminal. El gen que codifica este precursor no ha sido descripto en *A. mellifera* así como tampoco se han detectado péptidos maduros mediante MS, por lo que este sistema de señalización estaría ausente en este insecto (Hummon *et al*, 2006; Boerjan *et al*, 2010).

En nuestro trabajo, hemos podido detectar un GPN que codifica para ATS-B en el genoma de *R. prolixus (Rhopr-ATS-B).* Del precursor se deducen once péptidos, ocho de los cuales presentan el dominio $W(X)_6$GWamida. Mediante MS se detectaros los ocho péptidos maduros con el dominio característico, y un péptido espaciador. (Tabla 1.20).

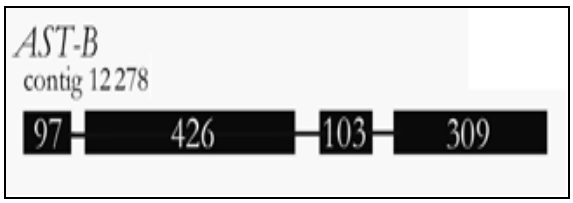


Figure 1.50: Estructura del gen de *Rhopr-ATS-B*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. La longitud del intrón 1 es mayor que la indicada, ya que se desconoce el UTR 5´.

Tabla 1.20 Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ATS B* y los péptidos maduros encontrados por MS, $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
QGPNTVTPMDNVIAEDEYLIPSNPALIDD		3141,49	
SWKDLQSSGW-NH_2	SWKDLQSSGW-NH_2	1192,58	SN, Cerebro

GWKDMQTVGW-NH_2	GWKDMQTVGW-NH_2	1206,57	SN, Cerebro
AWTDLPSSGW-NH_2	AWTDLPSSGW-NH_2	1118,53	Cerebro
DWKDMQSSGW-NH_2	DWKDMQSSGW-NH_2	1238,53	Cerebro
AWSDLQSSGW-NH_2	AWSDLQSSGW-NH_2	1135,52	SN, Cerebro
SKGDDQDEDIEENLAEED	SKGDDQDEDIEENLAEED	2050,82	Cerebro
AWNSLHGGW-NH_2	AWNSLHGGW-NH_2	1026,49	SN, Cerebro
TADWGSFRGSW-NH_2	TADWGSFRGSW-NH_2	1268,58	SN, Cerebro
EPAWQNLKGLWG-NH_2	EPAWQNLKGLWG-NH_2	1397,73	SN
SPVNLLQDNFQTGYVEPILN		2261,15	

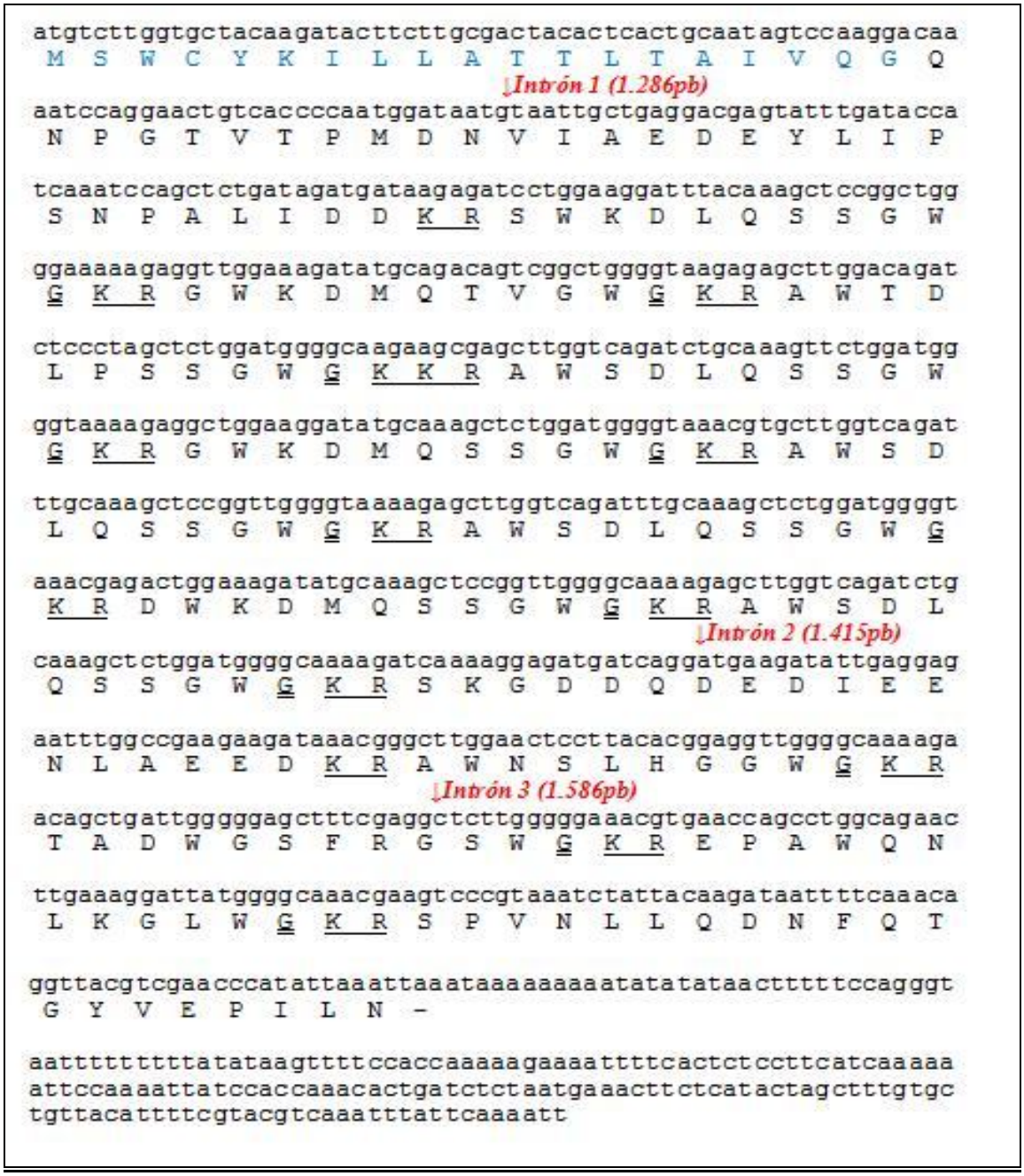


Figura 1.51: Secuencia nucleotídica y aminoacídica de *Rhopr-ATS B*. El péptido señal (probabilidad SignalP-HMM: 0.996) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran

subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Precursor de Allatotropina

Allatotropina es un neuropéptido que fue aislado del cerebro del lepidóptero *M sexta* (Kataoka *et al*, 1989), y posteriormente caracterizado en varias especies de insectos (Abdel-Latief *et al*, 2003; Park *et al*, 2002b, Truesdell *et al*, 2000; Veenstra y Costes, 1999). Sin embargo, no se ha encontrado en *D. melanogaster* (Riehle *et al,* 2002) ni *A. mellifera* (Hummon *et al*, 2006), lo que indica que este sistema de señalización se ha perdido varias veces a lo largo de la evolución de artrópodos. A pesar de esto, el péptido de *M. sexta* ha demostrado ser capaz de estimular la síntesis de JH en la CA de *A. mellifera* (Rachinsky y Feldlaufer, 2000).

Si bien ha sido caracterizada sobre la base de su capacidad de estimular la síntesis de JH por la *Corpora Allata* en varias especies de insectos (Kataoka et al, 1989; Rachinsky et al, 2003, ver Weaver y Audsley, 2009), se constataron otras funciones para este péptido , como cardioacelerador (Koladich *et al*, 2002; Veenstra et al, 1994), y estimulador de la contracción muscular en diferentes especies (Duve et al, 1999, 2000; Matthews et al, 2007; Rudwall et al, 2000;. Santini y Ronderos, 2007). Además, se propuso su implicación en el intercambio de iones (Lee *et al*, 1998) y, más recientemente, en la regulación de la secreción de enzimas digestivas en el digestivo medio (Lwalaba et al., 2009). A pesar de que AT se aisló en insectos, los miembros de esta familia de péptidos también se han encontrado en otros *Phyla* de invertebrados (Elekonich y Horodyski, 2003).

A partir del precursor de *Rhopr-AT* se duducen tres péptidos. El péptido con actividad allatoregulatoria presenta el motivo conservado en el extremo C terminal TARGFamida. Además, el precursor presenta otra región con un alto grado de homología entre las diferentes secuencias ortólogas en su extremo C terminal, (Figura 1.54). El grado de conservación indica que esta región se encuentra bajo presión evolutiva y podría tener función biológica.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

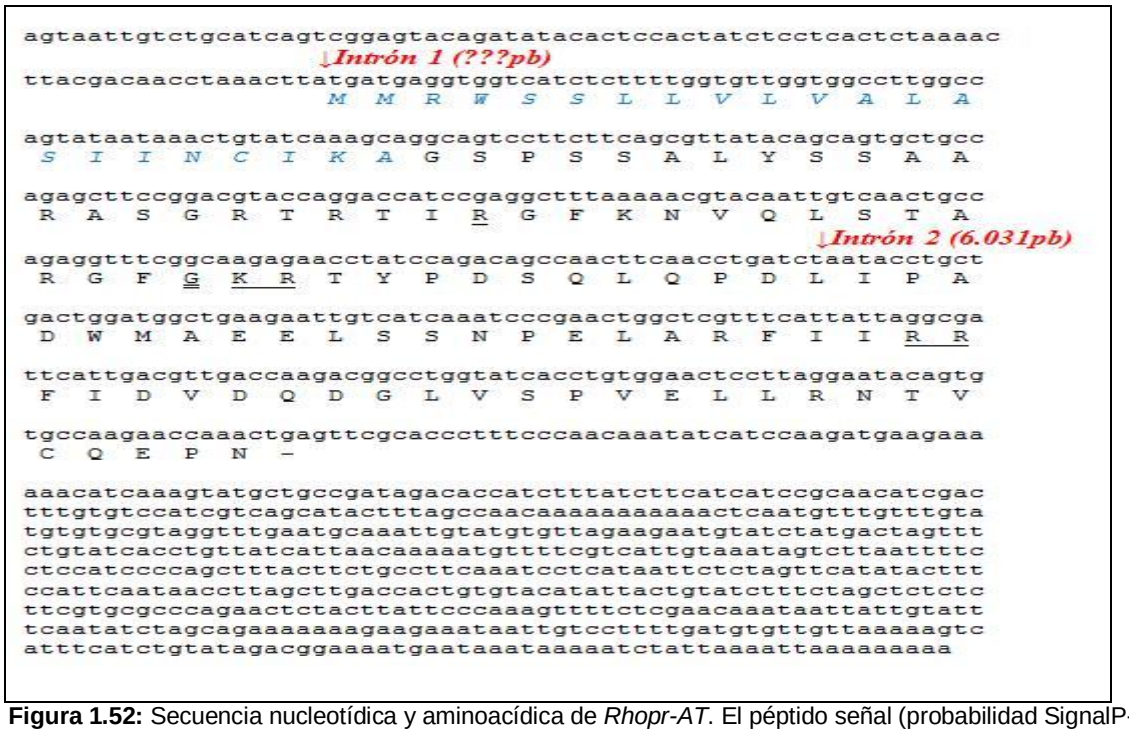


Figura 1.52: Secuencia nucleotídica y aminoacídica de *Rhopr-AT*. El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

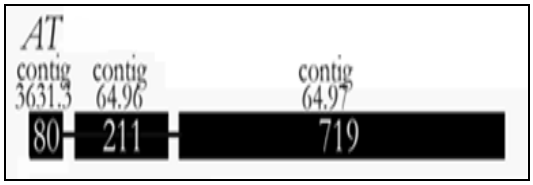


Figure 1.53: Estructura del gen de *Rhopr-AT*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón se indica en la parte superior. Las líneas representan los intrones.

Tabla 1.21: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-AT* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

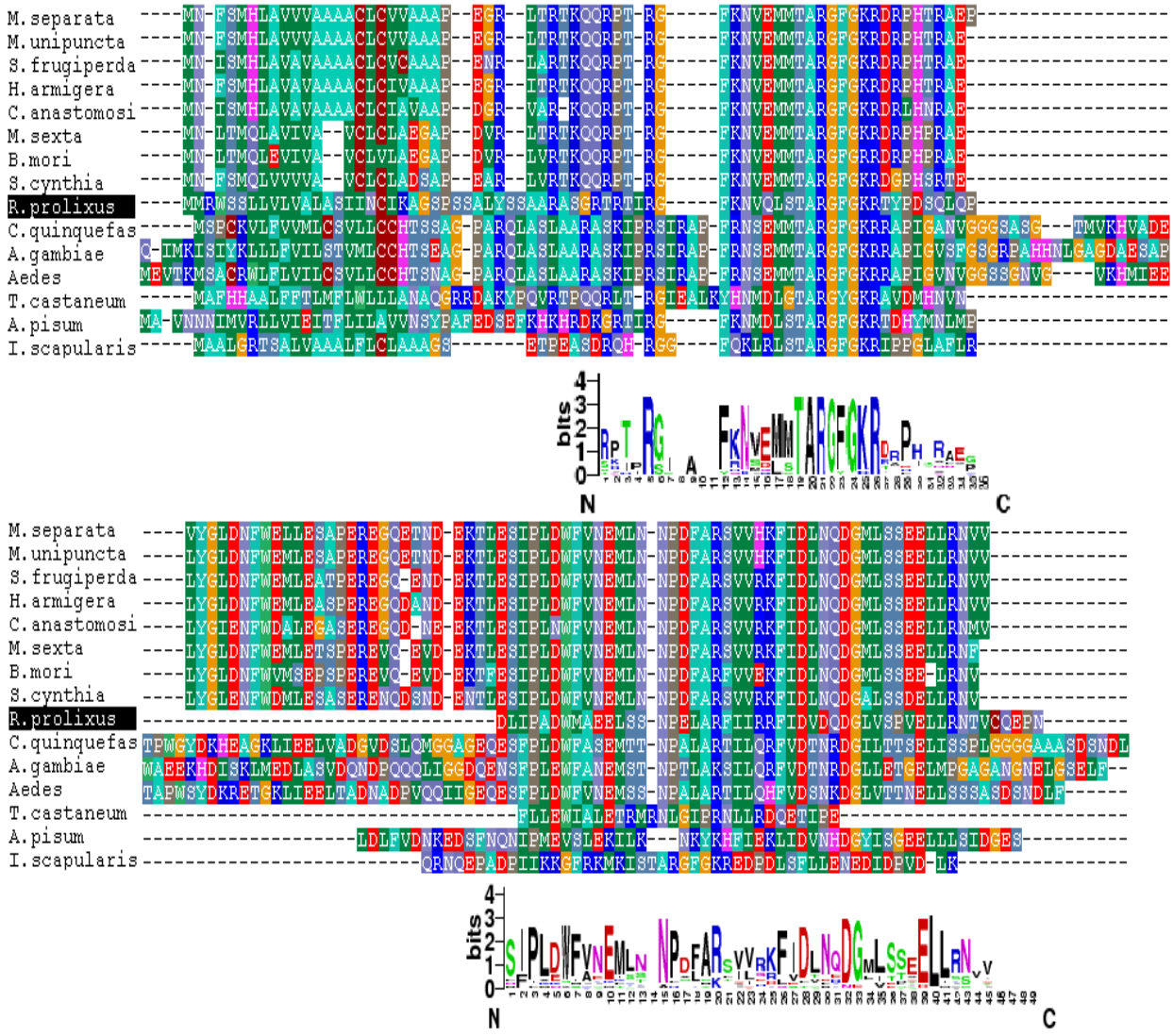


Figura 1.54: Alineamiento del precursor de *Rhopr-AT* con los precursores ortólogos caracterizados en distintas especies de insectos. Para las regiones que se encuentran más conservadas entre especies se indica el contenido informativo para cada posición. Además del péptido biológicamente activo, se observa otra región con alto grado de conservación en el precursor, lo que sugiere que se encuentra bajo presión evolutiva y podría tener función biológica.

1.3.2.e) Péptidos asociados al metabolismo energético

Las hormonas Adipokinéticas (AKH) se encuentran involucradas en la homeostasis de carbohidratos en hemolinfa mediante la estimulación de su liberación en momentos en que los insectos no se encuentran en intensa actividad física. En *D. melanogaster,* este proceso es contrarrestado por péptidos simil-insulina (ILP), los que disminuyen la concentración de carbohidratos en hemolinfa. Por lo tanto, AKHs e ILPs en insectos actúan de manera similar al glucagon e insulina en mamíferos durante la homeostasis de azúcar en sangre (Kim *et al,* 2004b).

Hormona Adipokinetica

AKH es producido en la *Corpora Cardíaca* de insectos. En *D. melanogaster*, esta hormona regula la liberación de energía durante el vuelo o locomoción intensa (Schaffer *et al*; 1990). Durante este período es liberada a la hemolinfa y trasportada hasta el cuerpo graso, donde estimula la liberación de carbohidratos.

Se han encontrado péptidos de la familia AKH en todos los insectos analizados así como en crustáceos, de lo que se deduce que este sistema de señalizacón celular se ha originado tempranamente durante la evolución de los artrópodos.

Mediante RACE- PCR hemos podido confirmar la transcripción de este gen en el SN de *R. prolixus* (Figura 1.55). A partir del clivaje del precursor se generan dos péptidos además del péptido señal; el péptido característico de la familia AKH y un péptido acompañante (Tabla 1.22). Llamativamente, el péptido AKH no ha sido detectado por MS ni en nuestro trabajo, ni en ningún trabajo peptidómico en insectos realizado hasta la fecha. Esto puede deberse a su alta hidrofobicidad, que lo dejaría retenido en la columna de RP-HPLC.

```
atggccaccaatctgttcataacaagtgtattggtactgctgacatttcattacaccctc
 M  A  T  N  L  F  I  T  S  V  L  V  L  L  T  F  H  Y  T  L

gcacagctaacattttcaaccgactggggtaagagatcagttcgccacaatgcacctgat
 A  Q  L  T  F  S  T  D  W  G  K  R  S  V  R  H  N  A  P  D
                                                ↓Intrón 1 (1.775pb)
tgtacaccaaatcctgatactgtaatttttctgtacaagtatttgcagaatgagttttac
 C  T  P  N  P  D  T  V  I  F  L  Y  K  Y  L  Q  N  E  F  Y

aaaatgatagagtgtggaaaaacggaaggtctttgacaccaagcttaatttagctgacac
 K  M  I  E  C  G  K  T  E  G  L  -

aacgaaatgaataatgtatttaattgataattatcattgtgaaggcaaaaaaaaaaatc
aatttttaagactaa
```

Figura 1.55: Secuencia nucleotídica y aminoacídica de *Rhopr-AKH*. El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

.

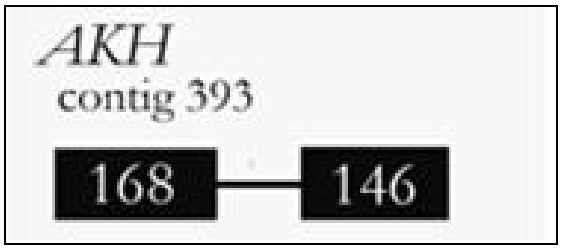


Figure 1.56: Estructura del gen de *Rhopr-AKH*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. El exón 1 está incompleto ya que se desconoce el UTR 5´para este mRNA.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Tabla 1.22: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-AKH* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos RACE SNC	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
QLTFSTDW-NH_2		996,48	
SVRHNAPDCTPNPDTVIFLYKYLQNEFYKMIECGKTGGL		4462,17	

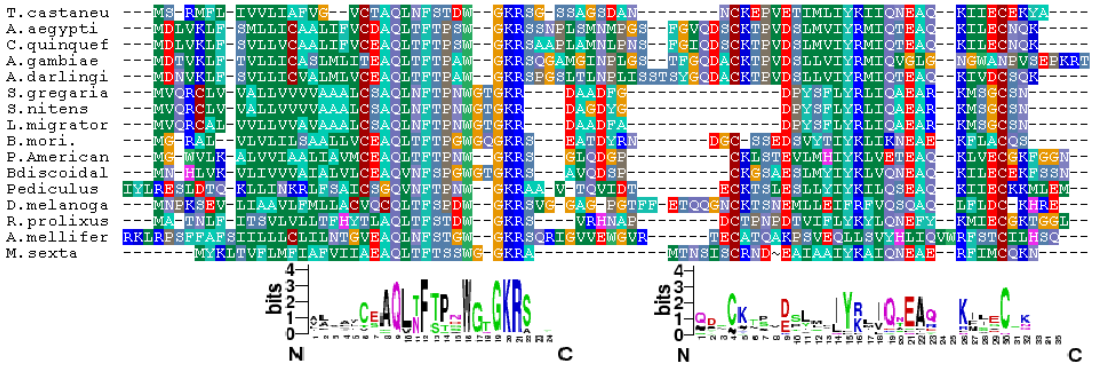


Figura 1.57: Alineamiento de los precursores de *Rhopr-AKH* de distintas especies de artrópodos. En la parte inferior se muestra el contenido informativo de las distintas posiciones expresado en bits.

Péptido símil-Insulina-2

Los péptidos simil insulina (ILPs) se encuentran tanto en vertebrados como en invertebrados, de lo que se deduce que han aparecido tempranamente durante el proceso de evolución. En vertebrados se dividen en dos grupos que presentan similitudes a nivel de secuencia, pero difieren a nivel de estructura cuanternaria y llevan a cabo funciones distintas. La insulina es un heterodímero compuesto por las cadenas A y B que regula la homeóstasis de azúcares en sangre, mientras que los péptidos simil-insulina están compuestos por una sola subunidad y regulan el crecimiento tisular (Froesch y Zapf, 1985; Clemmons, 1989).

Desde su primer descubrimiento en *B. mori* a partir de extractos de 2.800.000 cerebros, por su capacidad de estimular el desarrollo en larvas sin cerebro (Nagasawa et al, 1980), se han caracterizado péptidos simil-insulina en varias especies de insectos (ver Nässel, 2002; Wu & Brown, 2006). En *D. melanogaster* los ILPs han demostrado participar en la regulación de diversas funciones como el metabolismo de azúcares en hemolinfa y el crecimiento (Rulifson *et al*, 2002; Ikeya *et al*, 2002; Geminard *et al*, 2006; Kim *et al*, 2004b). Los genomas de los insectos que han sido secuenciados hasta la fecha poseen múltiples genes que codifican este tipo de péptidos (Iwami, 2000; Brogiolo *et al*, 2001; Riehle *et al*, 2006; Li *et al*, 2008). Los ILP de tipo B, al igual que las insulinas de vertebrados, se caracterizan por la presencia del motivo CCXXXC.

En nuestro trabajo hemos detectado cuatro genes que codifican péptidos simil-insulina en el genoma de *R. prolixus*, aunque sólo se ha podido determinar la transcripción de uno

de ellos: ILP-2. Sin embargo, los análisis bioinformáticos y transcriptómicos no han permitido conocer el ORF completo para este gen, por lo que solo se conoce parte del mismo (Figura1.60).

Mediante búsquedas bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de ILP caracterizados en otros insectos (BLASTN *E value*: E%=1e-129).

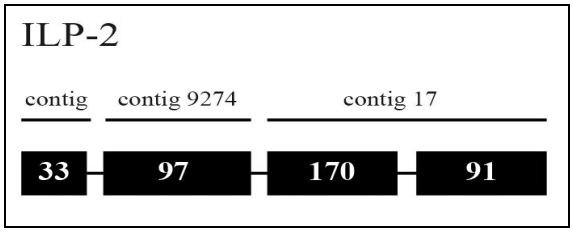


Figure 1.59: Estructura del gen *Rhopr-ILP-2*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones. El exón 4 está incompleto, ya que se desconoce parte del ORF y el UTR 3´para este mRNA.

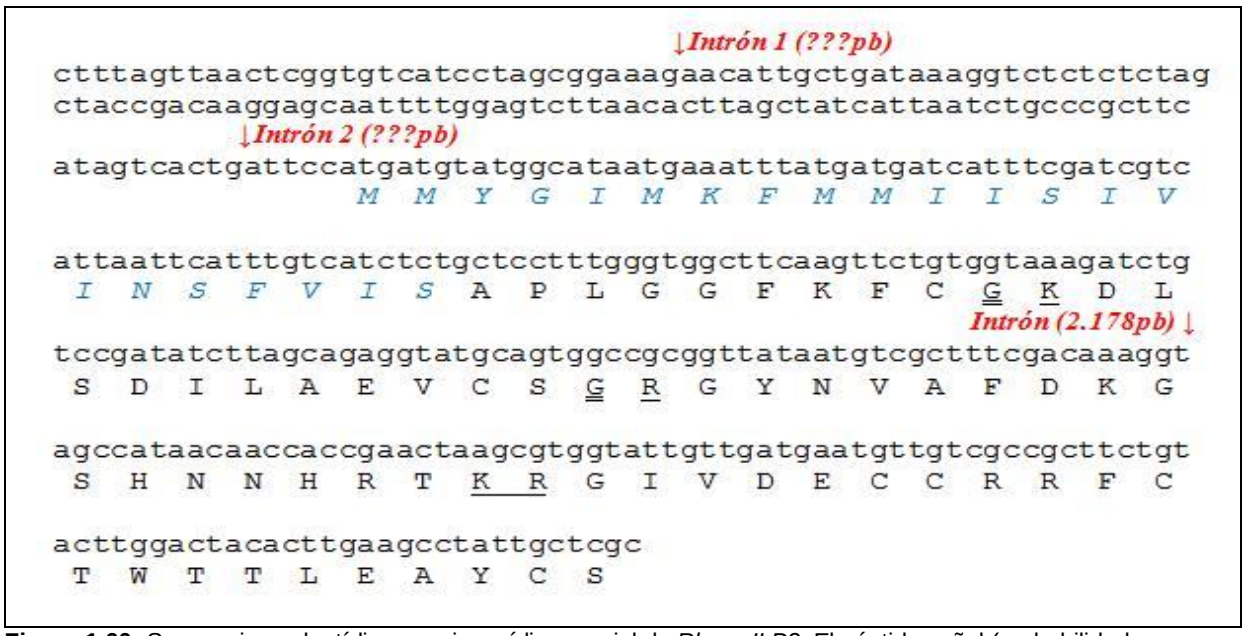


Figura 1.60: Secuencia nucleotídica y aminoacídica parcial de *Rhopr-ILP2*. El péptido señal (probabilidad SignalP-HMM: 0.404) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

3.2.f) Orcokininas

Las orcokininas (OK) son una familia de neuropéptidos que ha sido caracterizada por primera vez en crustáceos (Stangier J, 1992) y más recientemente en los insectos *L maderae* (Hofer and Homberg, 2006) y *A. mellifera* (Hummon et al, 2006). También han sido identificados mediante búsqueda bioinformática en el genoma de *A. gambiae* (Pascual et al, 2004), *D. melanogaster* (Liu et al, 2006) y *B. mory* (Roller *et al*, 2008)

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

posibles ortólogos de OK, pero su expresión aún no ha sido demostrada experimentalmente. Este sistema de señalización no ha sido descripto en *T. castaneum.*

Si bien en crustáceos se ha establecido un papel miotrópico para OK (Stangier J, 1992), su función en insectos no parece ser miotrópica (Pascual et al, 2004), y no ha sido aún dilucidada, aunque hay evidencia que las asocia a la regulación de los ritmos circadianos (Hofer and Homberg, 2006).

En el genoma de *R. prolixus* hemos podido identificar un GPN que codifica OK (*Rhopr-OK*), y determinar su patrón de expresión en SN por IHQ (Figura 1.61). Hemos detectado la expresión de OK en el cerebro, los diferentes ganglios del SNC central y en regiónnes neurohemales (CC-CA y Nervios Abdominales. Figura 1.61). Por otro lado, se ha podido detectar la presencia de OK en hemolinfa mediante MALDI TOF/TOF MS. Su presencia en hemolinfa demuestra por primera vez una función hormonal de OK en insectos (ver mas adelante), lo que concuerda con la inmunomarcación en las regiones neurohemales.

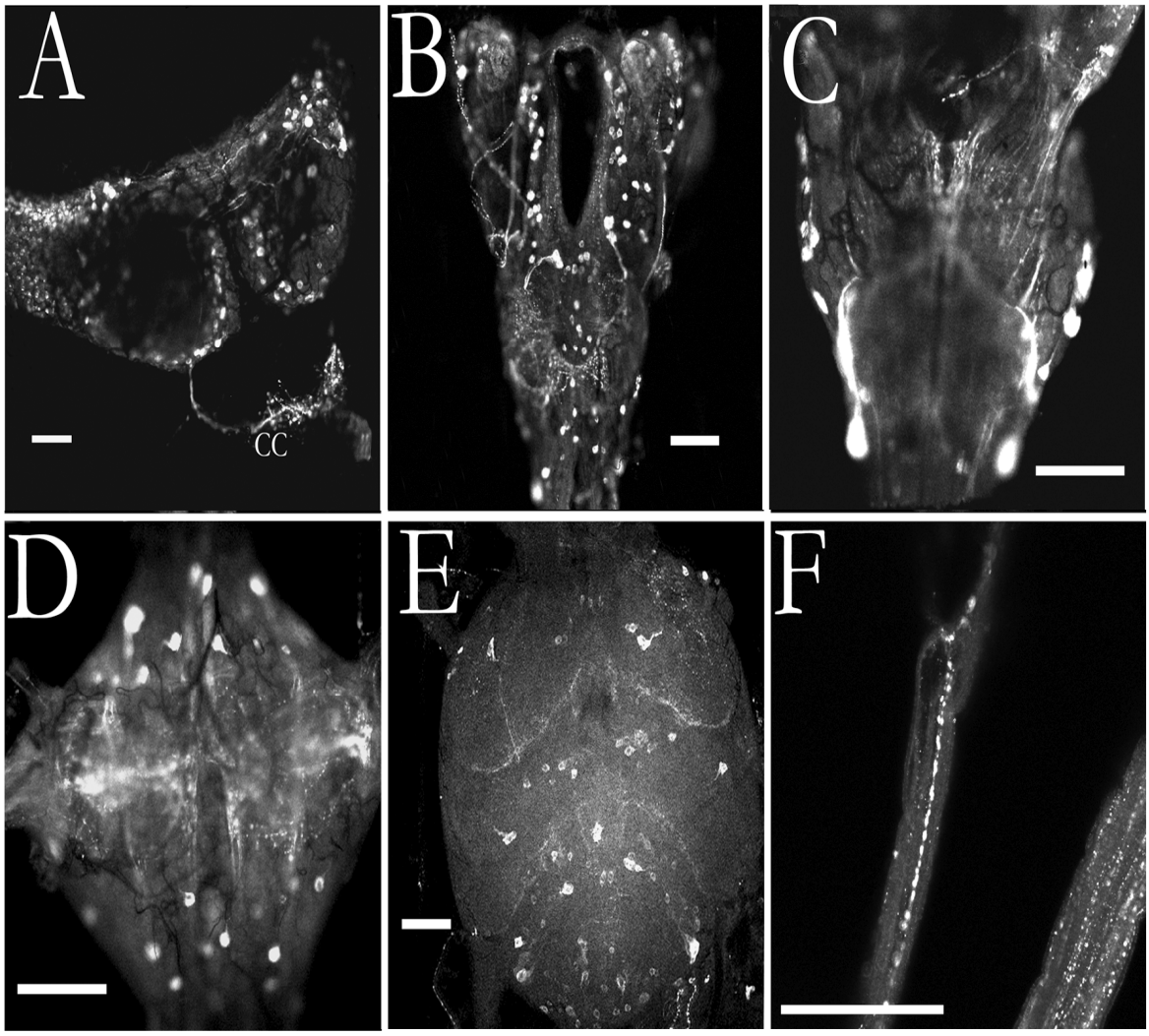


Figura 1.61: Inmunoreactividad (IR) de Orcoquininas en el Sistema Nervioso Central de *R. prolixus* (A) Vista dorsal del cerebro. Nótese la IR en Corpora Cardiaca (CC) sitio de liberación neurohemal. (B) Vista ventral del cerebro y ganglio subesofágico (SOG). (C) Vista dorsal del SOG. (D) Ganglio protoraxico (E) Masa ganglionar meso-toraxica. (F) Detalle de nervios abdominales. Barra=100

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Para este GPN observamos dos formas alternativas de *splicing* en *R. prolixus*, una de ellas fue caracterizada mediante RACE-PCR en SN. Esta isoforma (A) consta de 4 exones (1, 2, 4 y 5). En exón 1 codifica el péptido señal, mientras que el exón 5 codifica los péptidos que contienen en dominio característico de la familia (NFDEIDR en el extremo N-terminal). De este precursor se producen además cuatro péptidos que no corresponden a la familia de OK. La segunda isoforma (B) de este gen ha sido encontrada en la base de EST de digestivo medio anterior. Si bien el ORF no está completo en la base de datos, hemos podido determinar que las isoformas A y B comparten los exones 1 y 2, mientras que la isoforma B tiene un exón (3) ausente en la isoforma A (Figura 1.62). La isoforma B sólo está representada en su extremo 5' en la base de datos de EST, por lo que no hemos podido determinar la presencia del exón 5, que codifica los péptidos de la familia de las OK. Sin embargo, ambas comparten la presencia de péptido señal, que está codificado en el exón 1, por lo que son productos proteicos de secreción.

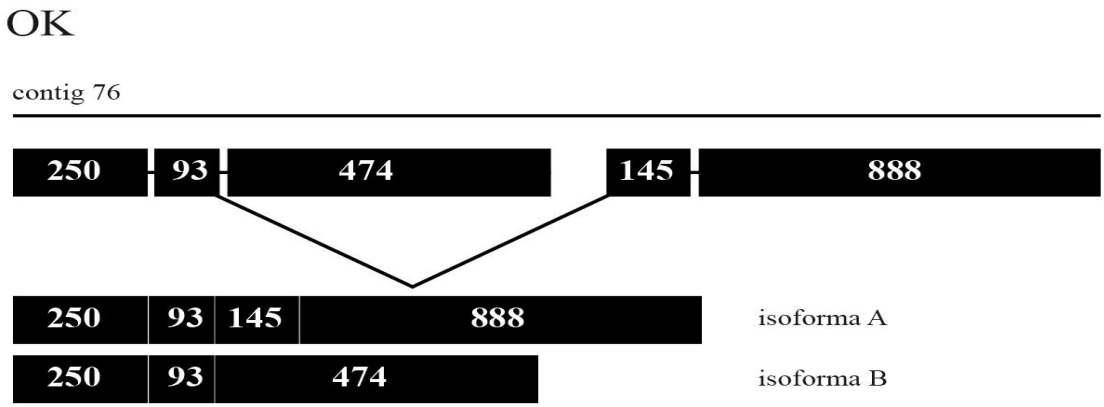


Figure 1.62: Estructura del gen *Rhopr-OK* y las 2 formas de splicing encontradas. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. Tanto la isoforma B como el exón 3 están incompletos, ya que solo se conoce una secuencia parcial.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 1.63: Secuencia nucleotídica y aminoacídica de la isoforma A *Rhopr-OK*. La misma fue obtenida mediante RACE en el SNC. El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Tabla 1.23: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-OK* y los péptidos maduros encontrados por MS. $[M+H]^+$ masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	[M+H]+	Tejido
FPRGELGVEEGNLYPGLYRDQTMEDKE-NH_2		3141,5	
NLDTLGSGNLLRDLEAVLAHPNLFYGPA		2980,55	
NHDTLDSLSGITPGSQ	NHDTLDSLSGITPGSQ	1691,79	SN
FDPLSSAYAAD		1156,516	
NFDEIDRSGFNSFI	NFDEIDRSGFNSFI	1660,77	SN, Cerebro, Hemolinfa
	NFDEIDRSGFN	1313,58	Cerebro
NFDEIDRSGFDGFV	NFDEIDRSGFDGFV	1617,73	SN, Cerebro
	NFDEIDRSGFDG	1371,59	Cerebro
NFDEIDRVGFGSFI	NFDEIDRVGFGSFI	1615,76	SN, Cerebro, Hemolinfa
	NFDEIDRVGF	1211,57	Cerebro

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

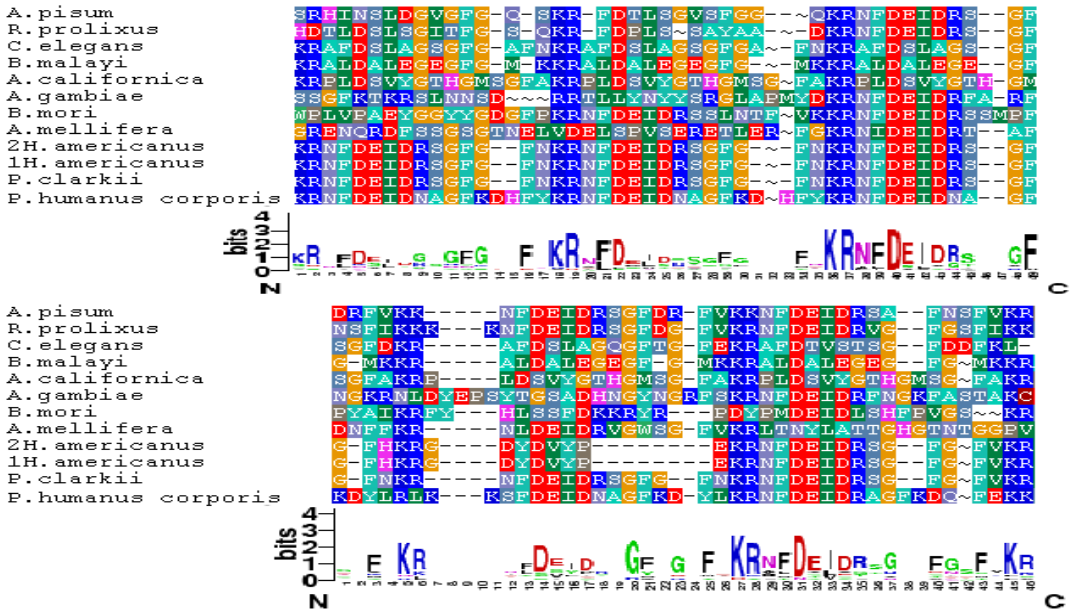


Figura 1.64: Alineamiento del precursor de *Rhopr-OK* con los precursores ortólogos caracterizados en distintas especies de insectos. El contenido informativo expresado en bits se muestra para cada posición.

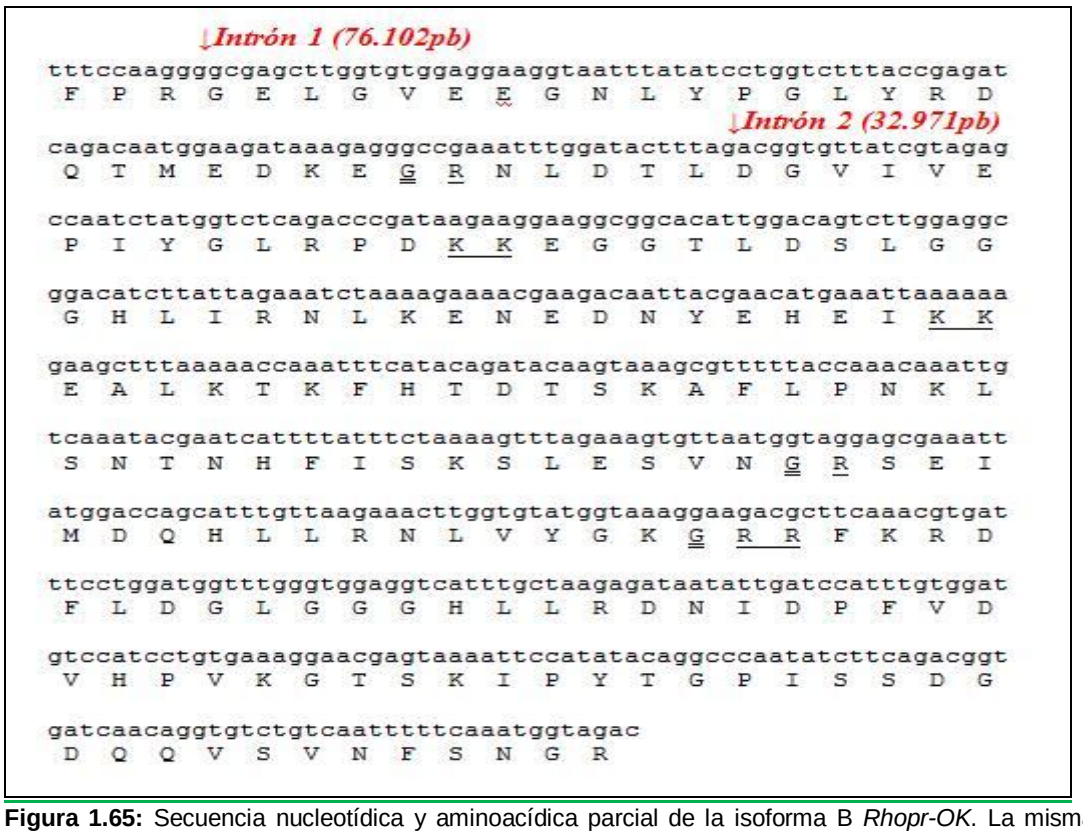


Figura 1.65: Secuencia nucleotídica y aminoacídica parcial de la isoforma B *Rhopr-OK*. La misma fue obtenida mediante búsqueda en base de datos de ESTs de digestivo medio anterior. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

1.3.2.g) SIFamida

SIFamida es uno de los neuropéptidos más conservados a nivel de secuencia en artrópodos (ver Verleyen *et al*, 2009b. Figura 1.66). A pesar de que especies del Órden *Orthoptera* como *L. migratoria y S. gregaria* fueron extensamente analizadas mediante técnicas peptidómicas (Clynen *et al*, 2001; 2003), SIFamida nunca había sido encontrado en un insecto hemimetábolo, por lo que se había sugerido una relación entre su ocurrencia y la metamorfosis completa. Nuestro trabajo ha sido el primero en reportar la presencia de SIF-amida en un insecto de este tipo, *R. prolixus*.

Si bien existe una conservación a nivel de secuencia en SIF-amida entre crustáceos e insectos, el patrón de expresión del péptido es distinto. En crustáceos se observa una amplia distribución en el SNC y células del digestivo medio (Yasuda, 2004). En los insectos donde se ha estudiado, la expresión de este gen se restringe a cuatro neuronas de la región media del protocerebro, cuyos axones se extienden a lo largo del cordón ventral, con abundantes ramificaciones (Roller *et al*, 2008; Terhzaz *et al*, 2007; Verleyen *et al*, 2004).

A partir del análisis de la secuencia del precursor de SIFamida en *R. prolixus,* se puede deducir un péptido maduro perteneciente a esta familia: TYKKPPFNGSIF-NH_2 (Figura 1.66) . Mediante MS hemos podido confirmar la expresión de éste péptido, así como tres formas truncadas del mismo (Tabla 1.24). Un hecho similar fue observado en la descripción del peptidoma de *A. mellifera* (Hummon *et al*, 2006). Cabe destacar que la secuencia dibásica KK no es un sitio de clivaje en este precursor, de acuerdo a lo establecido por Veenstra (2000). La secuencia del ORF completo del mRNA de este gen pudo ser obtenida mediante RACE-PCR, aunque el UTR 3´no ha sido secuenciado.

Figure 1.68: Estructura del gen de *Rhopr-SIFamida*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones. El exón 3 no está completo ya que no se conoce el UTR 3´.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

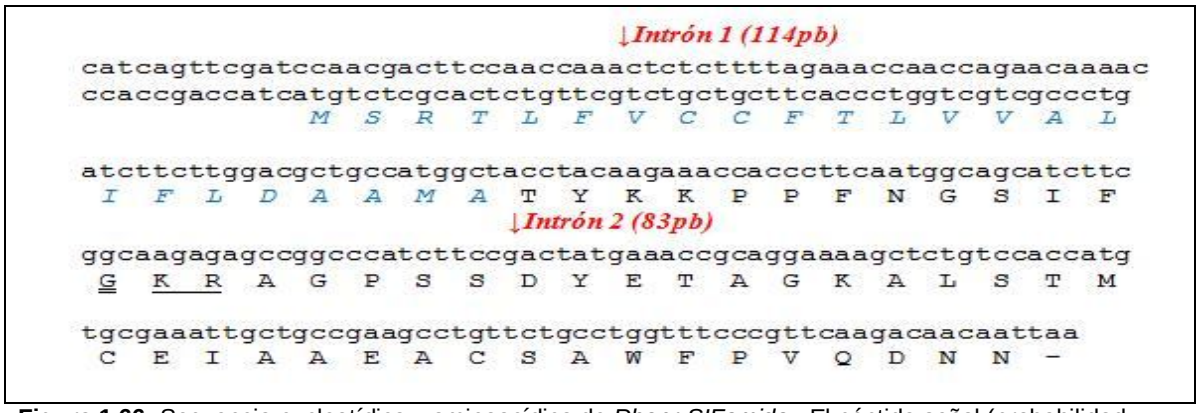


Figura 1.66: Secuencia nucleotídica y aminoacídica de *Rhopr-SIFamida* . El péptido señal (probabilidad SignalP-HMM: 0.988) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.24: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-SIFamida* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
TYKKPPFNGSIF-NH_2	TYKKPPFNGSIF-NH_2	1397,74	Cerebro
	YKKPPFNGSIF-NH_2	1296,7	Cerebro
	KKPPFNGSIF-NH_2	1133,66	Cerebro
	KPPFNGSIF-NH_2	1005,56	Cerebro
AGPSSDYETAGKALSTMCEIAAEACSAWFPVQDNN		3634,58	

La estructura del precursor es igual en todos los artrópodos en los que ha sido descripta: el péptidos activo con el motivo SIFamida se encuentra precedido por el péptido señal, y está separado de un péptido espaciador por un sitio de clivaje di básico. La misma estructura se observa para los precursores de AKH, ACP, CZ y CCAP. Además, pueden observarse dos Cys conservadas en todos los precursores, separadas por seis aminoácidos (Figura 1.67).

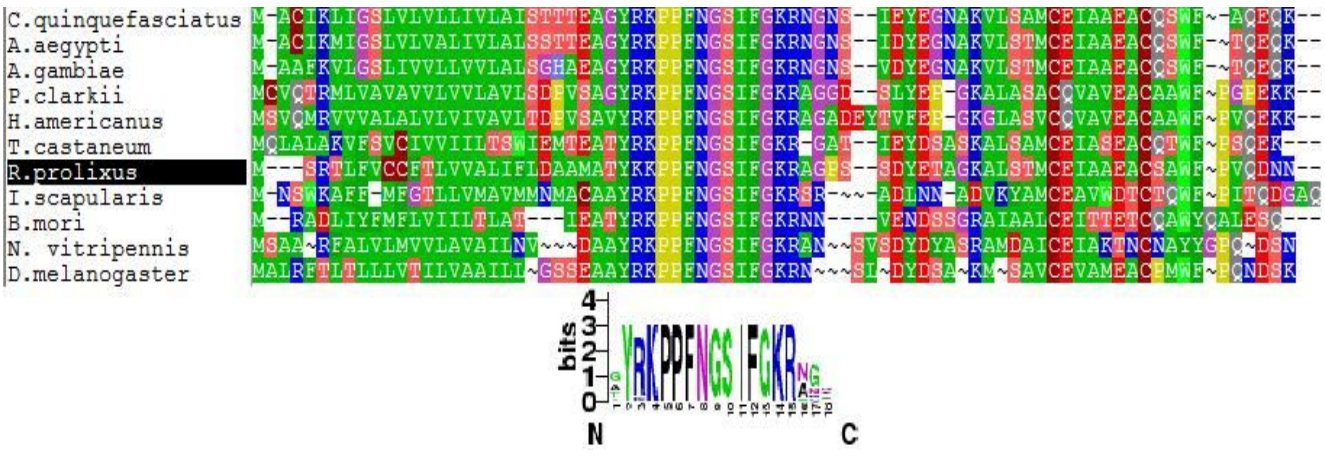


Figura 1.67: Alineamiento del precursor de *Rhopr-SIFa* con los precursores ortólogos caracterizados en distintas especies de insectos. El contenido informativo expresado en bits se muestra para cada posición. Se pueden observar dos residuos de Cys separados exactamente por seis aminoácidos.

Si bien no se cuenta hoy con suficiente información acerca de la función biológica de SIF-amida, estudios moleculares y fisiológicos mostraron que juega un papel importante en el comportamiento sexual de *D. melanogaster*. El silenciamiento de este gen mediante RNAi produce un comportamiento sexual promiscuo en individuos de ambos sexos. Las hembras con el gen silenciado son significativamente más receptivas que las salvajes, mientras que los machos cortejan indistintamente a hembras y otros machos. Además se observó que la expresión de este péptido aumenta durante la fase adulta (Terhzaz *et al*, 2007). La expresión de su receptor fue localizada dentro del SNC, por lo que este péptido actuaría como un neurotransmisor, procesando o transmitiendo información sensorial (Jorgensen *et al*, 2006).

La confirmación de la presencia del mRNA y el péptido de SIF-amida en el SN de *R. prolixus* indica que este neuropéptido no es una característica distintiva de insectos con metamorfosis completa. Sin embargo, el patrón de expresión podría ser diferente, y llevar a cabo una función distinta. Para estudiar el patrón de expresión de SIF-amida, hemos realizado ensayos de IHQ en SN de *R. prolixus*. Los resultados muestran que no solo la secuencia del precursor se encuentra conservada, sino también su patrón de expresión (Figura 1.69), lo cual fortalece la hipótesis de una conservación de función. La expresión restringida a cuatro neuronas en del protocerebro, así como la ausencia de marcación en regiones neurohemales, sustenta la hipótesis del papel de SIFamida como neurotransmisor, involucrado en el procesamiento de la información sensorial.

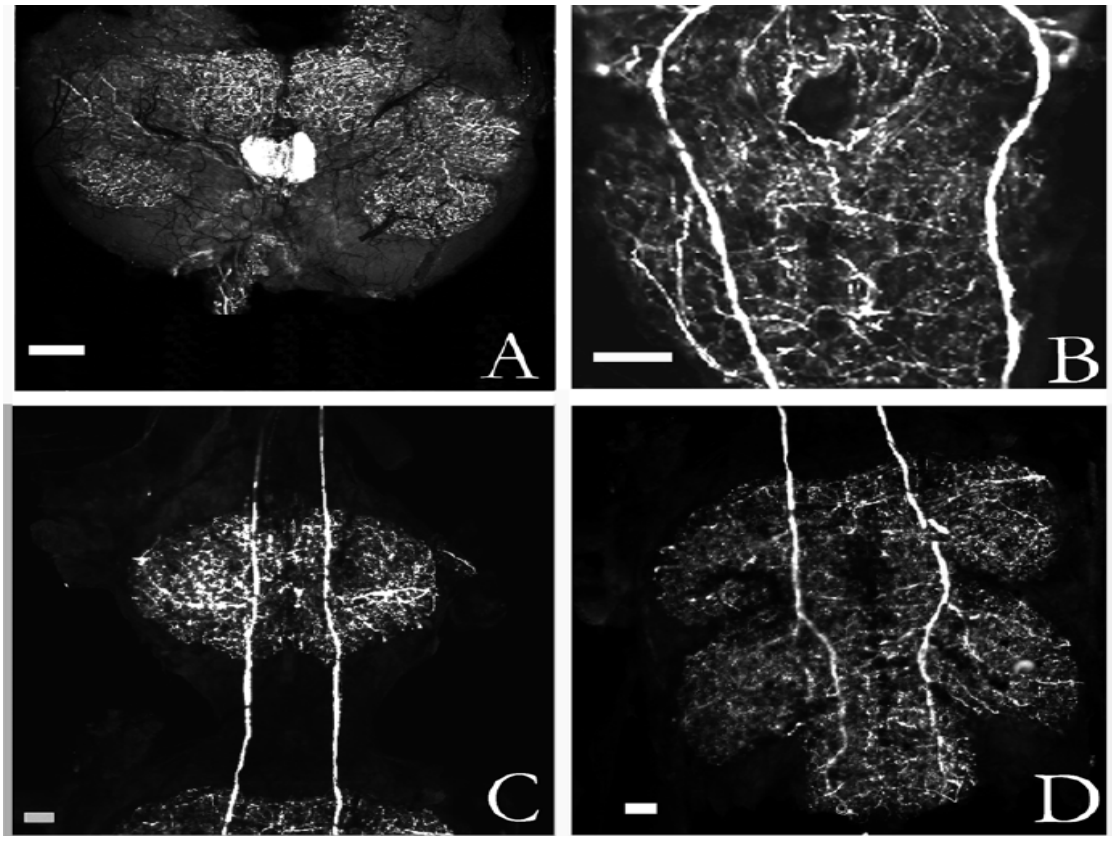


Figure 1.69: Inmunomarcación positiva para un péptido simil-SIFamida en el Sistema Nervioso Central de R. prolixus. (A) Vista dorsal del Cerebro. Se puede observar marcación positiva en cuatro neuronas y las proyecciones nerviosas de éstas. (B) Vista del ganglio subesofágico donde se muestra marcación positiva para las proyecciones nerviosas que descienden del cerebro originadas en las 4 neuronas mostradas en A. (C) Se muestran las mismas proyecciones nerviosas mostradas en B a nivel del Ganglio Protoráxico.
(D) Vista de Ganglio Mesometatoráxico donde se muestran las proyecciones nerviosas antes mencionadas. Bar=100 µm

Mediante búsquedas bioinformática hemos podido predecir *in silico* un receptor que presenta un alto grado de homología con los receptores de SIFamida caracterizados en otros insectos (BLASTN *E value*: E%=8e-157).

1.3.2.h) Taquikininas

Las Taquikininas (TK) pertenecen a una familia conservada de neuropéptidos, se ha determinado su expresión tanto en el SN como en intestino en vertebrados e invertebrados (ver Satake *et al*, 2003). Las TK de vertebrados se caracterizan por tener la secuencia consenso –FXGLMamida (X representa cualquier aminoácido) en el extremo C-terminal, mientras que las de invertebrados contienen la secuencia –FXGXRamida. Ensayos de expresión heteróloga del receptor demostraron que el pentapéptido C-terminal es suficiente para activarlo, y determina la especificidad entre los receptores de vertebrados e invertebrados. El reemplazo de Fen por Ala provoca pérdida de actividad (Torfs *et al*, 2002).

La primer TK descripta en insectos fue aislada de extractos del SNC de *L. migratoria* a partir de sus propiedades miotrópicas en la musculatura del recto (Schoofs *et al*, 1990). Posteriormente, en varios grupos de insectos fueron descriptos péptidos de esta familia, así como los genes que los codifican. En todas las especies de invertebrados estudiadas, un solo precursor da origen a todos los péptidos de la familia de TK. Los péptidos de esta familia han demostrado ser multifuncionales, y se encuentran asociados al procesamiento de información sensorial, la actividad motora y la contracción de la musculatura viceral (ver Satake *et al*, 2003).

En *R. prolixus*, se encontró inmunomarcación positiva para TK, tanto en el SNC como en el intestino medio, utilizando un suero policlonal generado contra TK1 de *L. migratoria*. Se demostró también que TK-1 y TK-2 de *L. migratoria* son capaces de causar un incremento en la frecuencia de contracción del recto en *R. prolixus.* (Kwok *et al*, 2005).

En nuestro trabajo hemos logrado caracterizar el mRNA de *Rhopr-TK* (Figura 1.71). Varios neuropéptidos maduros codificados en este GPN fueron detectados por MS en extractos de cerebro y SN en nuestro análisis peptidómico (Tabla 1.25). Una PTM que se observa con frecuencia en estos péptidos es la oxidación de los residuos de Met. Se han encontrado péptidos maduros con esta modificación y sin ella. También se han encontrado varios péptidos espaciadores generados a partir del precursor TK. Una particularidad de

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Rhopr-TK es que da origen a dos péptidos con igual secuencia: APASGFFGMR-NH_2. Este GPN codifica por otra parte un péptido que solo consta de los cinco residuos requeridos para la interacción con el receptor (FMGVR-NH_2) (Tabla 1.25).

Mediante búsqueda bioinformática se ha podido predecir *in silico* un GPCR que presenta un alto grado de homología con los receptores de TK caracterizados en otros insectos (BLASTN *E value*: E%=1e-102).

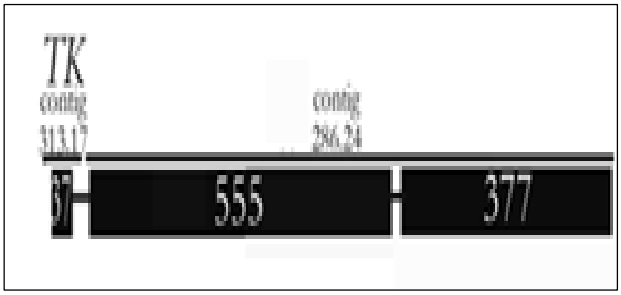


Figure 1.70: Estructura del gen de *Rhopr-TK*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones.

```
                                                          ↓Intrón 1 (???pb)
attcgctcacaacccgtccagcctgagccactccaggatgcctgtgggaagtctgctggta
                                                    M  P  V  G  S  L  L  V

atgagctgtgttctggccgcctgtttagcccaggagagaagagctatgggctttgtggga
 M  S  C  V  L  A  A  C  L  A  Q  E  R  R  A  M  G  F  V  G

atgcgaggcaagaaggacacgccagacatggaggagtacaagagggcaccctcaacaatg
 M  R  G  K  K  D  T  P  D  M  E  E  Y  K  R  A  P  S  T  M

gggtttcaaggagtccgaggcaagaaggatgacctcatcggcgagccggatgatactttc
 G  F  Q  G  V  R  G  K  K  D  D  L  I  G  E  P  D  D  T  F

ctggaagaattcaaaagagcaccggctgcaatgggctttcaaggaatgagaggcaaaaaa
 L  E  E  F  K  R  A  P  A  A  M  G  F  Q  G  M  R  G  K  K

acgcctgcaatgggattcatgggcatgcgcggtaaaaaggacagtgattatggttggtgg
 T  P  A  M  G  F  M  G  M  R  G  K  K  D  S  D  Y  G  W  W

gaggaggacaagagggcgcccgcttcgggggttcttcggcatgcggggcaaaaaagccccg
 E  E  D  K  R  A  P  A  S  G  F  F  G  M  R  G  K  K  A  P

gcctcaggcttctttgggatgcgaggcaagaagggaccatccagctcggcattcttcggc
 A  S  G  F  F  G  M  R  G  K  K  G  P  S  S  S  A  F  F  G

atgcgaggcaagaaaggaccttccggcttcatgggcgtcagaggcaaaaaagacagccct
 M  R  G  K  K  G  P  S  G  F  M  G  V  R  G  K  K  D  S  P
                                                          ↓Intrón 2 (1.009pb)
gatgacttgaaccatctgctgcaattgcttcgtgaatcagctcttaaacaggagatggag
 D  D  L  N  H  L  L  Q  L  L  R  E  S  A  L  K  Q  E  M  E

gagatgttggaagatgggcggggactgaagaggttcgctggcctgagcgactccttcgag
 E  M  L  E  D  G  R  G  L  K  R  F  A  G  L  S  D  S  F  E

gactacccggccgaactgctctaagccagccgcaagcctccacctcctgccccaaactgt
 D  Y  P  A  E  L  L  –

tgatttgtgcccttgcgtccctaccccggtccgagtgagggtgccaaccccgctgtttcc
tgtgatgtcaaccttcaagagatctcttcccgttaattttctagattgtgtagaccttgt
acctttattatataagcaactcgttacatatatctacatacataagtagatatatgtaag
atacatctaagaagacaatgtatgataaatacacagatataaatatactatacgatatat
aaaccgaaaaaa
```

Figura 1.71: Secuencia nucleotídica y aminoacídica de *Rhopr-TK*. El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azul y letra cursiva. Los sitios de clivaje proteolítico se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.25: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-TK* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	$[M+H]^+$	Tejido
QERRAMGFVGMR-NH_2	QERRAMGFVGMR-NH_2	1436,74	Cerebro
	pyro-QERRAMGFVGMR- NH_2	1419,72	Cerebro
DTPDMEEY		999,36	
STMGFQGVR-NH_2	STMGFQGVR-NH_2	981,49	Cerebro
APSTMGFQGVR-NH_2	APSTMGFQGVR-NH_2	1149,53	SN, Cerebro
	APST(M-OH)GFQGVR-NH_2	1165,53	SN
DDLIGEPDDTFLEEF	DDLIGEPDDTFLEEF	1754,76	SN, Cerebro
APAAMGFQGMR-NH_2	APAAMGFQGMR-NH_2	1135,51	SN, Cerebro
	APAA(M-OH)GFQG(M-OH)R-NH_2	1150,51	SN
TPAMGFMGMR-NH_2	TPAMGFMGMR-NH_2	1097,48	SN, Cerebro
DSDYGWWEED		1301,46	
APASGFFGMR-NH_2	APASGFFGMR-NH_2	1039,52	SN, Cerebro
	APASGFFG(OH-M)R-NH_2	1055,52	SN
APASGFFGMR-NH_2	APASGFFGMR-NH_2	1039,52	SN
	APASGFFG(-OH)MR-NH_2	1055,52	SN
GPSSSAFFGMR-NH_2	GPSSSAFFGMR-NH_2	1142,53	SN
	GPSSSAFFG(M-OH)R-NH2	1158,53	SN
GPSGFMGVR-NH_2	GPSGFMGVR-NH_2	906,46	SN
	GPSGF(M-OH)GVR-NH_2	922,46	SN
FMGVR-NH_2	FMGVR-NH_2	608,33	Cerebro
DSPDDLNHLLQLLRESALKQEMEEMLEDGRGL		3694,8	
	SPDDLNHLLQL	1264,65	SN
FAGLSDSFEDYPAELL	FAGLSDSFEDYPAELL	1773,82	

1.3.2.i) Pirokinina/PBAN/Diapause Hormone

El primer péptido de esta familia fue aislado en *L. maderae* por su capacidad de estimular la actividad muscular en el tracto digestivo (Holman *et al*, 1986). Las pirokininas (PK) son una familia de péptidos ampliamente distribuida, que comparte la presencia del motivo $FX_1PRLamida$ (X_1: S, G, V o T) en el extremo C terminal (Nachman *et al*, 1993). En *B. mori* el pentapéptido FTPRLamida ha demostrado ser necesario y suficiente para estimular la síntesis de feromonas (Kuniyoshi *et al*, 1992; Davis *et al*, 1992). En *H. virescence* la aplicación tópica del pentapéptido FTPRLamida conjugado con 7-bromofluoro, induce la producción de feromona. La inyección de 500 nmol de este compuesto produjo una mortandad del 100% (Nachman et al, 2002).

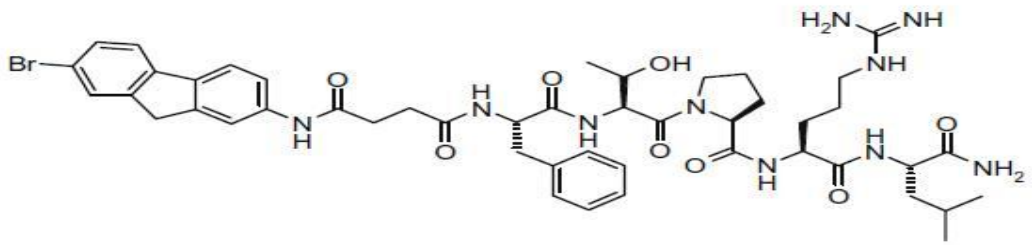


Figura 1.85: Estructura del compuesto 7-bromofluorene-FTPRLamida. Imagen tomada de Scherkenbeck, 2009

Tanto en pentapéptido como el 7-bromofluoro demostraron ser necesarios para lograr este efecto, lo que sugiere que el mismo está dado por un mecanismo de interacción específica ligando-receptor (Nachman *et al*, 2002). Estos resultados son importantes pues demuestran que pseudopéptidos lipofílicos son capaces de atravesar la cutícula e inducir una respuesta. También se han reportado cuatro pseudopéptidos con propiedades antagónicas a PK, que mostraron inhibir totalmente la producción de feromona en *H. peltigera* (Altstein *et al*, 1999; Zeltser *et al* ,2001).

En nuestro trabajo hemos podido identificar el precursor *Rhopr-PK* en el genoma de *R. prolixus* (Figura 1.86). Del análisis de este precursor se deducen dos péptidos con el dominio característico FX_1PRL-amida. Sin embargo, a diferencia de los péptidos ortólogos descriptos en otras especies, X_1 puede ser Ser, Ala o Arg. Estos dos últimos aminoácidos no fueron descriptos previamente en esta posición en moléculas ortólogas. El precursor da origen por otra parte, al péptido IVFRPRF-NH_2, que presenta en el extremo C-terminal el motivo caractrístico de LNF (RPRFamida).

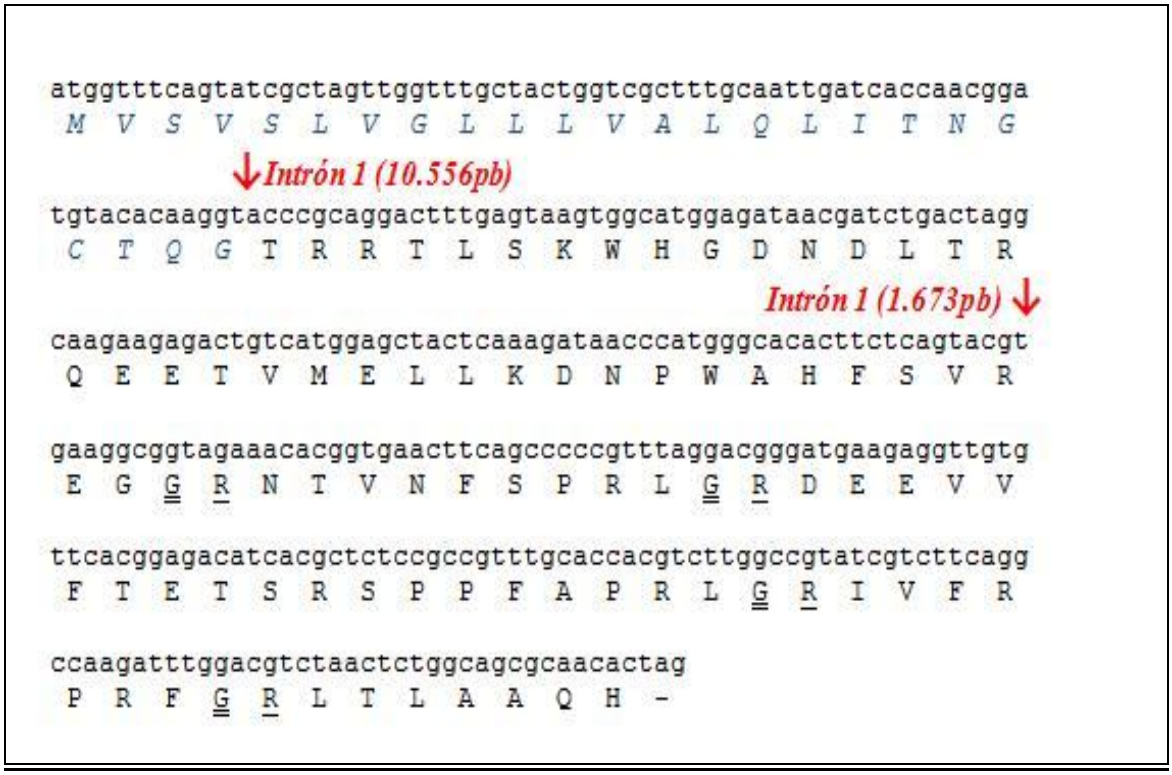


Figura 1.86: Secuencia nucleotídica y aminoacídica de *Rhopr-PK*. El péptido señal (probabilidad SignalP-HMM: 0.999) se muestra en azµl y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

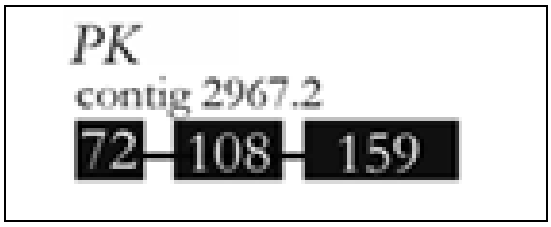


Figure 1.87: Estructura del gen *Rhopr-PK.* El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones.

Tabla1.30: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-TK* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
TRRTLSKWHGDNDLTRQEETVMELLKDNPWAHFSVREG-NH_2		4551,3	
NTVNFSPRL-NH_2		1046,6	
DEEVVFTETSRSPPFAPRL-NH_2		2176,10	
	SPPFAPRL-NH2	883,51	Cerebro
IVFRPRF-NH_2		933,58	
LTLAAQH		753,43	

Mediante búsquedas bioinformática se ha podido predecir *in silico* GPCR que presenta un alto grado de homología con los receptores de PK caracterizados en otros insectos (BLASTN E value: E%=7e-116).

1.3.2.j) Péptidos con función no establecida en insectos

Precursor de Neuropéptidos 1

El precursor del Neuropéptido 1 (NPLP1) fue identificado por primera vez en *D. melanogaster* mediante análisis peptidómicos (Baggerman et al, 2002) y posteriormente en *A. mellifera* (Hummon et al, 2006), *B. mori* (Roller et al, 2008) y *T. castaneum* (Li et al, 2008). Si bien la función de los péptidos codificados por este precursor es desconocida, se ha sugerido que podrían estar asociados a la regulación de la metamorfosis en larvas de insectos holometábolos. Esta hipótesis se basó en datos obtenidos a partir de microarreglos que muestran que la expresión del mRNA aumenta durante la pupación en *D. melanogaster* (Baggerman et al, 2002). Sin embargo, ni la inyección de tres péptidos sintéticos codificados por este precursor, ni anticuerpos desarrollados contra los mismos, afectaron la velocidad de retracción o contracción en los ensayos de pupación en *Neobellieria bullata*. El patrón de expresión en este insecto se restringe a 28 neuronas en el SNC (Verleyen et al, 2009a).

Una característica particular de NPLP1 es la escasa similitud de secuencia que presentan entre si los péptidos que codifica, así como la escasa homología entre los genes ortólogos de distintas especies. Los péptidos que se generan a partir del precursor de

NPLP1 no presentan ningún motivo común, ni entre distintas especies, ni dentro de una misma especie. Lo anterior hace suponer la existencia de diferentes receptores para los péptidos codificados por NPLP1, aunque hasta la fecha no se ha descripto ninguno de ellos.

En nuestro trabajo, hemos podido caracterizar *Rhopr-NPLP1* en toda su longitud mediante RACE-PCR (Figura 1.73). De su traducción y clivaje proteolítico se generan 20 posibles péptidos, 13 de los cuales pudieron ser identificados por MS en el SN de *R. prolixus* (Tabla 1.26). Hasta el momento, NPLP1 nunca había sido descripto en un insecto hemimetábolo.

Figure 1.72: Estructura del gen de *Rhopr-NPLP1*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

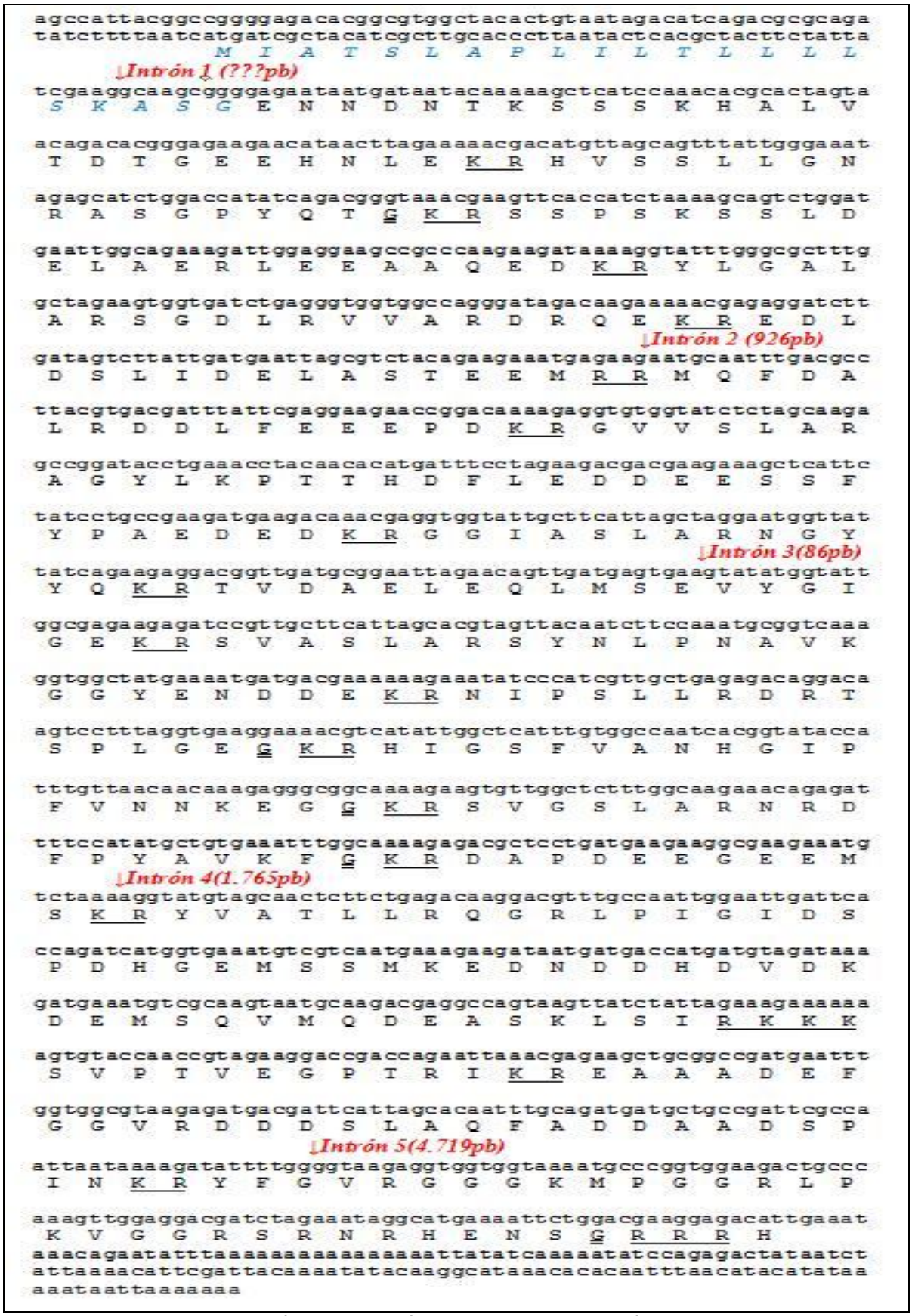


Figura 1.73: Secuencia nucleotídica y aminoacídica de *Rhopr-NPLP1*. El péptido señal (probabilidad SignalP-HMM: 1.000) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

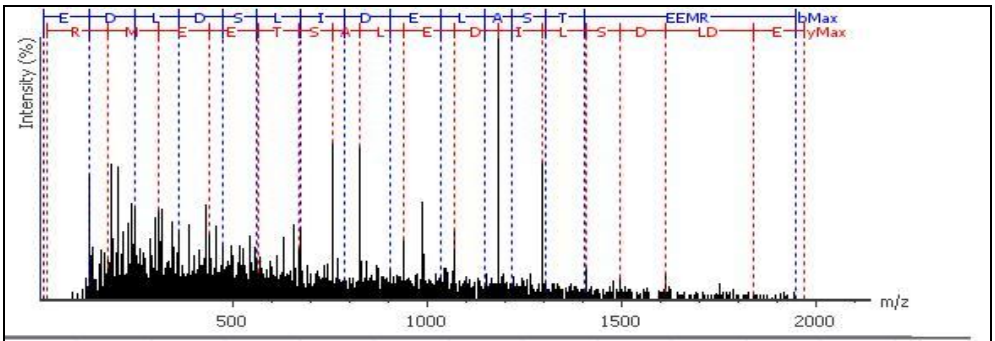
doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.26: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-NPLP1* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada. . Entre paéntesis se muestran los tejidos donde los péptidos han sido encontrados sólo por *Mass match.*

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
NNDNTKSSSKHALVTDTGEEHNLE		2640,21	
HVSSLLGNRASGPYQT-NH_2	HVSSLLGNRASGPYQT-NH_2	1685,87	SN, Cerebro, GS
SSPSKSSLDELAERLEEAAQED	SSPSKSSLDELAERLEEAAQED	2391,12	SN, Cerebro, GS
YLGALARSGDLRVVARDRQE		2245,22	
EDLDSLIDELASTEEM	EDLDSLIDELASTEEM	1809,8	SN
MQFDALRDDLFEEEPD		1969,85	
	FDALRDDLFEEEPD	1710,75	SN
GVVSLARAGYLKPTTHDFLEDDEESSFYPAEDED	GVVSLARAGYLKPTTHDFLEDDEESSFYPAEDED	3802,72	(GS)
	GVVSLARAGY	992,55	SN
GGIASLARNGYYQ	GGIASLARNGYYQ	1369,69	SN, GS
TVDAELEQLMSEVYGIGE	TVDAELEQLMSEVYGIGE	1982,93	SN
SVASLARSYNLPNAVKGGYENDDE	SVASLARSYNLPNAVKGGYENDDE	2569,28	SN,(GS)
NIPSLLRDRTSPLGE-NH_2	NIPSLLRDRTSPLGE-NH_2	1666,92	Cerebro
HIGSFVANHGIPFVNNKEG-NH_2	HIGSFVANHGIPFVNNKEG-NH_2	2036,05	Cerebro
SVGSLARNRDFPYAVKF-NH_2	SVGSLARNRDFPYAVKF-NH_2	1926,03	SN
DAPDEEGEEMS		1208,43	
YVATLLRQ-NH_2	YVATLLRQ-NH_2	962,58	SN, Cerebro, (GS)
LPIGIDSPDHGEMSSMKEDNDDHDVDKDEMSQVMQDEASKLSI		4788,09	
SVPTVEGPTRI		1155,64	
EAAADEFGGVRDDDSLAQFADDAADSPIN	EAAADEFGGVRDDDSLAQFADDAADSPIN	2982,29	SN, Cerebro, (GS)
YFGVRGGGKMPGGRLPKVG-NH_2		1932,08	
SRNRHENS-NH_2		998,49	

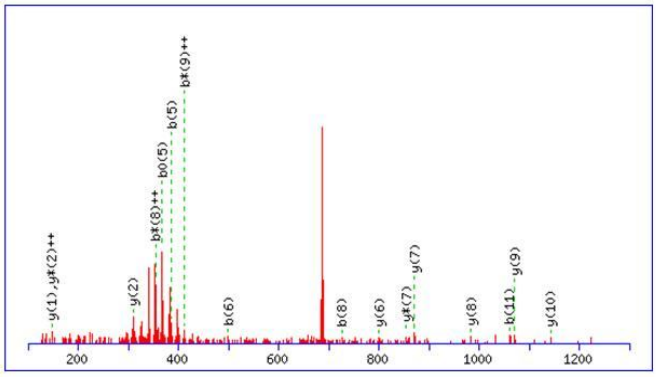
En nuestro análisis peptidómico de tejidos hemos podido detectar 7 péptidos maduros codificados por el precursor NPLP1 en glándulas salivales (Tabla 1.26, figura 1.76, 1.77, 1.78). Resulta interesante el hecho de que se haya encontrado el péptido EDLDSLIDELASTEEMR (Figura 1.75), el cual es un intermediario del proceso de maduración, que aún conserva el residuo Arg en el extremo C-terminal antes de su remoción por la enzima carboxyl-peptidasa. El haber encontrado un intermediario inmaduro indica que NPLP-1 está siendo sintetizado en las glándulas salivales, ya que a las terminaciones nerviosas que las inervan llegan solo los péptidos maduros, los cuales son procesados en el soma neuronal. Hasta la fecha, este es el primer indicio de que las glándulas salivales son capaces de producir péptidos con actividad regulatoria, y sugiere la implicancia de NPLP-1 en la regulación de la producción y/o secreción de saliva

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes



#	Imm...	b	b-H2O	a	c	Seq	y	y-H2O	z	z'	y (2+)	#
1	102.06	130.12	112.04	102.06	147.09	E						17
2	88.04	245.09	227.16	217.09	262.14	D	1837.05	1818.84	1819.83	1820.83	918.93	16
3	86.10	358.18	340.17	330.18	375.26	L	1721.83	1703.82	1704.80	1705.80	861.41	15
4	88.04	473.22	455.19	445.26	490.20	D	1608.92	1590.73	1591.84	1592.72	804.87	14
5	60.04	560.23	542.27	532.29	577.27	S	1493.78	1475.77	1476.69	1477.69	747.36	13
6	86.10	673.32	655.33	645.33	690.42	L	1406.65	1388.67	1389.66	1390.78	703.84	12
7	86.10	786.43	768.41	758.41	803.53	I	1293.65	1275.59	1276.69	1277.57	647.30	11
8	88.04	901.44	883.40	873.51	918.44	D	1180.54	1162.49	1163.49	1164.48	590.76	10
9	102.06	1030.60	1012.54	1002.60	1047.50	E	1065.49	1047.50	1048.50	1049.53	533.24	9
10	86.10	1143.68	1125.53	1115.55	1160.60	L	936.52	918.43	919.52	920.42	468.72	8
11	44.05	1214.60	1196.69	1186.72	1231.61	A	823.40	805.36	806.37	807.48	412.18	7
12	60.04	1301.73	1283.60	1273.62	1318.64	S	752.35	734.41	735.41	736.36	376.66	6
13	74.06	1402.70	1384.70	1374.66	1419.93	T	665.35	647.32	648.30	649.42	333.15	5
14	102.06	1531.70	1513.69	1503.71	1548.73	E	564.26	546.26	547.26	548.26	282.62	4
15	102.06	1660.74	1642.93	1632.75	1677.77	E	435.23	417.20	418.24	419.20	218.10	3
16	104.05	1791.78	1773.77	1763.79	1808.81	M	306.17	288.15	289.14	290.10	153.58	2
17	129.11					R	175.12	157.12	158.10	159.10	88.06	1

Figura 1.75: Espectro MS/MS del péptido EDLDSLIDELASTEEMR en glándulas salivales. La presencia de este intermediario del proceso de maduración es un indicio que NPLP-1 es expresado y procesado en estos órganos. En el espectro se muestran solo los iones ***b*** (azµl) e ***y*** (rojo), necesarios para la secuenciación del péptido (PEAKS).

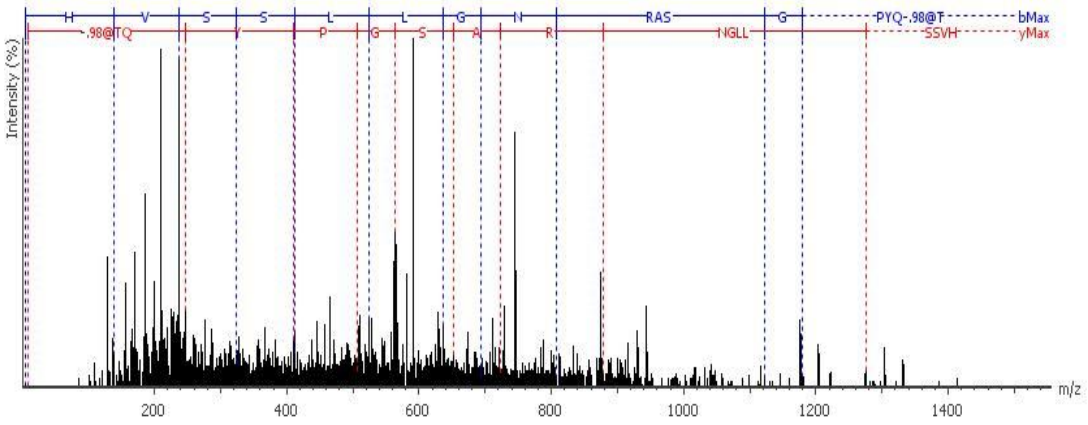


Monoisotopic mass of neutral peptide Mr(calc): 1368.6786
Ions Score: 44 Expect: 0.00073
Matches (Bold Red): 16/106 fragment ions using 45 most intense peaks

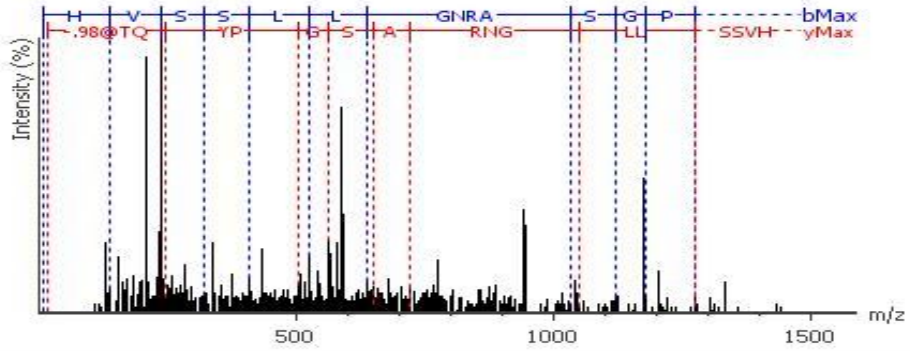
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	58.0287	29.5180					G							13
2	115.0502	58.0287					G	1312.6644	656.8359	1295.6379	648.3226	1294.6539	647.8306	12
3	228.1343	114.5708					I	1255.6430	628.3251	1238.6164	619.8118	1237.6324	619.3198	11
4	299.1714	150.0893					A	**1142.5589**	571.7831	1125.5324	563.2698	1124.5483	562.7778	10
5	**386.2034**	193.6053			**368.1928**	184.6001	S	**1071.5218**	536.2645	1054.4952	527.7513	1053.5112	527.2592	9
6	**499.2875**	250.1474			481.2769	241.1421	L	**984.4898**	492.7485	967.4632	484.2352			8
7	570.3246	285.6659			552.3140	276.6606	A	**871.4057**	436.2065	**854.3791**	427.6932			7
8	**726.4257**	363.7165	709.3992	**355.2032**	708.4151	354.7112	R	**800.3686**	400.6879	783.3420	392.1747			6
9	840.4686	420.7380	823.4421	**412.2247**	822.4581	411.7327	N	644.2675	322.6374	627.2409	314.1241			5
10	897.4901	449.2487	880.4635	440.7354	879.4795	440.2434	G	530.2245	265.6159	513.1980	257.1026			4
11	**1060.5534**	530.7803	1043.5269	522.2671	1042.5429	521.7751	Y	473.2031	237.1052	456.1765	228.5919			3
12	1223.6167	612.3120	1206.5902	603.7987	1205.6062	603.3067	Y	**310.1397**	155.5735	293.1132	**147.0602**			2
13							Q	**147.0764**	74.0418	130.0499	65.5286			1

Figura 1.76: espectro MS/MS del péptido GGIASLARNGYYQ en glándulas salivales. En el espectro se muestran solo los iones ***b*** e ***y*** necesarios para la secuenciación del péptido (MASCOT).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes



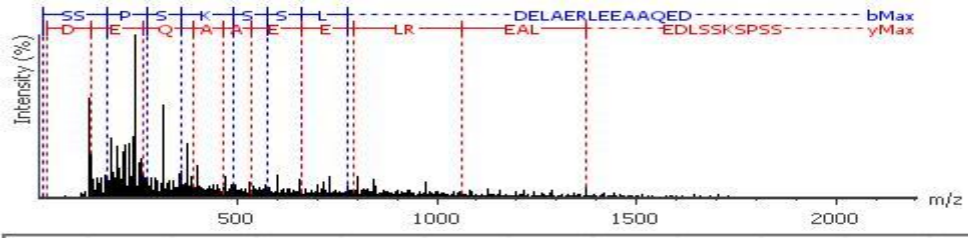
#	Immonium	b	b-H2O	a	c	Seq	y	y-H2O	z	z'	y (2+)	#
1	110.07	138.06	120.07	110.07	155.08	H						16
2	72.08	237.13	219.14	209.14	254.14	V	1548.81	1530.80	1531.79	1532.79	774.91	15
3	60.04	324.17	306.16	296.12	341.17	S	1449.74	1431.73	1432.72	1433.72	725.38	14
4	60.04	411.23	393.22	383.21	428.26	S	1362.71	1344.70	1345.69	1346.69	681.86	13
5	86.10	524.28	506.31	496.28	541.30	L	1275.70	1257.67	1258.65	1259.65	638.33	12
6	86.10	637.37	619.33	609.36	654.39	L	1162.60	1144.59	1145.59	1146.57	581.79	11
7	30.03	694.37	676.31	666.31	711.37	G	1049.51	1031.50	1032.51	1033.49	525.26	10
8	87.06	808.40	790.39	780.44	825.43	N	992.49	974.48	975.46	976.43	496.75	9
9	129.11	964.53	946.52	936.41	981.56	R	878.41	860.44	861.46	862.42	439.74	8
10	44.05	1035.57	1017.49	1007.57	1052.60	A	722.38	704.34	705.33	706.40	361.67	7
11	60.04	1122.62	1104.59	1094.61	1139.63	S	651.34	633.31	634.31	635.35	326.17	6
12	30.03	1179.68	1161.61	1151.63	1196.65	G	564.30	546.28	547.32	548.29	282.64	5
13	70.07	1276.68	1258.67	1248.68	1293.70	P	507.28	489.29	490.23	491.28	254.14	4
14	136.08	1439.74	1421.73	1411.74	1456.77	Y	410.25	392.22	393.22	394.24	205.60	3
15	101.07	1567.80	1549.79	1539.80	1584.83	Q	247.13	229.13	230.11	231.11	124.07	2
16	73.08					-.98@T	119.08	101.07	102.05	103.05	60.04	1



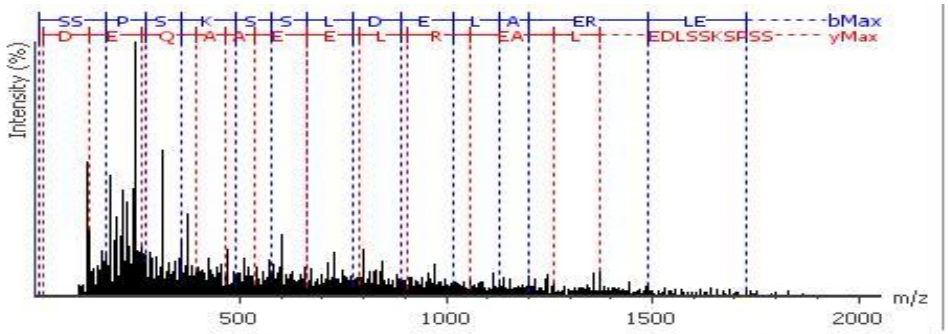
#	Imm...	b	b-H2O	a	c	Seq	y	y-H2O	z	z'	y (2+)	#
1	110.07	138.07	120.07	110.07	155.09	H						16
2	72.08	237.14	219.15	209.15	254.20	V	1548.81	1530.80	1531.79	1532.79	774.91	15
3	60.04	324.17	306.16	296.13	341.25	S	1449.74	1431.73	1432.84	1433.72	725.37	14
4	60.04	411.21	393.22	383.24	428.02	S	1362.71	1344.70	1345.69	1346.69	681.86	13
5	86.10	524.30	506.26	496.27	541.31	L	1275.73	1257.67	1258.65	1259.65	638.37	12
6	86.10	637.35	619.37	609.30	654.34	L	1162.60	1144.59	1145.61	1146.57	581.81	11
7	30.03	694.39	676.38	666.39	711.42	G	1049.45	1031.53	1032.55	1033.49	525.24	10
8	87.06	808.43	790.39	780.49	825.46	N	992.49	974.48	975.46	976.46	496.75	9
9	129.11	964.53	946.45	936.54	981.56	R	878.45	860.44	861.48	862.42	439.72	8
10	44.05	1035.65	1017.67	1007.57	1052.60	A	722.47	704.34	705.38	706.40	361.67	7
11	60.04	1122.64	1104.60	1094.61	1139.63	S	651.39	633.28	634.41	635.28	326.20	6
12	30.03	1179.63	1161.61	1151.63	1196.65	G	564.32	546.27	547.32	548.25	282.64	5
13	70.07	1276.73	1258.67	1248.68	1293.70	P	507.29	489.25	490.23	491.32	254.20	4
14	136.08	1439.74	1421.73	1411.74	1456.77	Y	410.20	392.17	393.22	394.26	205.60	3
15	101.07	1567.80	1549.79	1539.80	1584.83	Q	247.13	229.14	230.12	231.13	124.05	2
16	73.08					-.98@T	119.08	101.07	102.05	103.05	60.04	1

Figura 1.77: Espectros MS/MS del péptido HVSSLLGNRASGPYQamida en glándulas salivales obtenidos a partir de dos diferentes muestras de estos órganos. En el espectro se muestran solo los iones ***b*** (azul) e ***y*** (rojo), necesarios para la secuenciación del péptido (PEAKS) .

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes



#	Imm...	b	b-H2O	a	c	Seq	y	y-H2O	z	z'	y (2+)	#
1	60.04	88.04	70.03	60.04	105.07	S						22
2	60.04	175.08	157.13	147.08	192.10	S	2304.08	2286.07	2287.06	2288.06	1152.54	21
3	70.07	272.13	254.12	244.14	289.15	P	2217.05	2199.04	2200.02	2201.02	1109.03	20
4	60.04	359.17	341.17	331.17	376.17	S	2120.00	2101.99	2102.97	2103.97	1060.53	19
5	101.11	487.25	469.23	459.27	504.28	K	2032.97	2014.96	2015.94	2016.94	1016.98	18
6	60.04	574.28	556.27	546.28	591.31	S	1904.87	1886.86	1887.84	1888.84	952.94	17
7	60.04	661.30	643.37	633.35	678.35	S	1817.84	1799.83	1800.81	1801.81	909.42	16
8	86.10	774.43	756.38	746.39	791.40	L	1730.81	1712.80	1713.78	1714.78	865.90	15
9	88.04	889.43	871.45	861.43	906.48	D	1617.72	1599.71	1600.70	1601.70	809.40	14
10	102.06	1018.47	1000.46	990.47	1035.50	E	1502.70	1484.69	1485.67	1486.67	751.85	13
11	86.10	1131.55	1113.54	1103.56	1148.58	L	1373.69	1355.70	1356.63	1357.63	687.35	12
12	44.05	1202.59	1184.58	1174.60	1219.62	A	1260.57	1242.56	1243.64	1244.64	630.79	11
13	102.06	1331.63	1313.62	1303.64	1348.66	E	1189.53	1171.52	1172.51	1173.51	595.27	10
14	129.11	1487.73	1469.72	1459.74	1504.76	R	1060.53	1042.52	1043.46	1044.46	530.75	9
15	86.10	1600.82	1582.81	1572.82	1617.85	L	904.39	886.38	887.48	888.36	452.69	8
16	102.06	1729.86	1711.85	1701.87	1746.89	E	791.40	773.38	774.43	775.42	396.22	7
17	102.06	1858.90	1840.89	1830.91	1875.93	E	662.33	644.34	645.32	646.32	331.63	6
18	44.05	1929.94	1911.93	1901.95	1946.97	A	533.25	515.24	516.22	517.30	267.13	5
19	44.05	2000.98	1982.97	1972.98	2018.00	A	462.27	444.17	445.16	446.23	231.59	4
20	101.07	2129.04	2111.03	2101.04	2146.06	Q	391.15	373.14	374.13	375.19	196.10	3
21	102.06	2258.08	2240.07	2230.08	2275.11	E	263.09	245.08	246.08	247.13	132.04	2
22	88.04					D	134.05	116.03	117.02	118.02	67.52	1



#	Im...	b	b-H2O	a	c	Seq	y	y-H2O	z	z'	y (2+)	#
1	60.04	88.04	70.03	60.04	105.07	S						22
2	60.04	175.08	157.11	147.08	192.11	S	2304.08	2286.07	2287.06	2288.06	1152.58	21
3	70.07	272.13	254.13	244.13	289.15	P	2217.05	2199.04	2200.02	2201.02	1109.03	20
4	60.04	359.20	341.16	331.17	376.20	S	2120.00	2101.99	2102.97	2103.97	1060.53	19
5	101.11	487.24	469.25	459.18	504.26	K	2032.97	2014.96	2015.94	2016.94	1016.98	18
6	60.04	574.31	556.29	546.28	591.37	S	1904.87	1886.86	1887.84	1888.84	952.94	17
7	60.04	661.34	643.37	633.33	678.35	S	1817.84	1799.83	1800.81	1801.81	909.55	16
8	86.10	774.42	756.42	746.42	791.48	L	1730.81	1712.80	1713.78	1714.78	865.90	15
9	88.04	889.46	871.36	861.55	906.51	D	1617.72	1599.71	1600.70	1601.70	809.36	14
10	102.06	1018.60	1000.57	990.47	1035.62	E	1502.70	1484.69	1485.67	1486.76	751.87	13
11	86.10	1131.63	1113.65	1103.58	1148.59	L	1373.66	1355.72	1356.71	1357.74	687.33	12
12	44.05	1202.57	1184.58	1174.62	1219.62	A	1260.63	1242.69	1243.64	1244.71	630.79	11
13	102.06	1331.63	1313.62	1303.64	1348.66	E	1189.53	1171.65	1172.58	1173.67	595.33	10
14	129.11	1487.74	1469.80	1459.74	1504.76	R	1060.53	1042.62	1043.53	1044.54	530.75	9
15	86.10	1600.82	1582.81	1572.82	1617.85	L	904.52	886.53	887.52	888.50	452.69	8
16	102.06	1729.97	1711.85	1701.87	1746.89	E	791.48	773.36	774.42	775.40	396.18	7
17	102.06	1858.90	1840.89	1830.98	1875.93	E	662.35	644.37	645.27	646.31	331.63	6
18	44.05	1929.94	1911.93	1901.95	1946.97	A	533.26	515.27	516.25	517.33	267.16	5
19	44.05	2000.98	1982.97	1972.98	2018.00	A	462.22	444.28	445.21	446.23	231.59	4
20	101.07	2129.04	2111.03	2101.04	2146.06	Q	391.17	373.15	374.15	375.20	196.09	3
21	102.06	2258.08	2240.07	2230.08	2275.11	E	263.10	245.09	246.10	247.14	132.04	2
22	88.04					D	134.05	116.03	117.02	118.02	67.52	1

Figura 1.78: Espectros MS/MS del péptido SSPSKSSLDELAERLEEAAQED en glándulas salivales obtenidos a partir de dos diferentes muestras de estos órganos. En el espectro se muestran solo los iones b (azul) e y (rojo), necesarios para la secuenciación del péptido (PEAKS).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Pro-hormona 1

Las prohormonas (PH) fueron identificadas experimentalmente por primera vez en un estudio peptidómico del cerebro de *A. mellifera* (Hummon *et al*, 2006). Posteriormente los genes que los codifican fueron encontrados *in silico* en el genoma de varios insectos (ver Tabla 1.3). La función biológica de estos péptidos es desconocida.

En nuestro trabajo, hemos podido secuenciar competamente el mRNA que codifica *Rhopr-PH1* mediante RACE-PCR (Figura 1.79) En el análisis peptidómico hemos logrado identificar péptidos maduros y formas truncadas de ellos, generados a partir de éste precursor (Tabla 1.24).

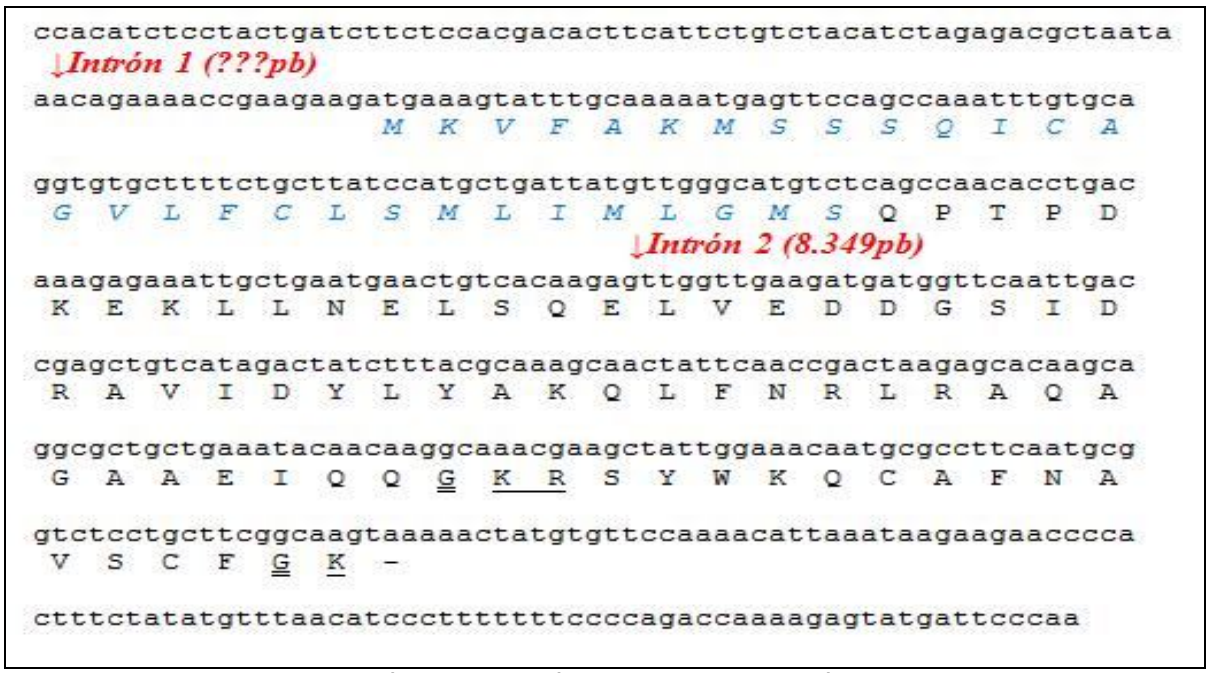


Figura 1.79: Secuencia nucleotídica y aminoacídica de *Rhopr-PH1*. El péptido señal (probabilidad SignalP-HMM: 0.982) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

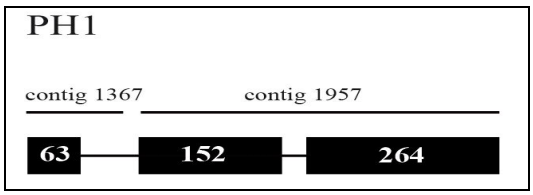


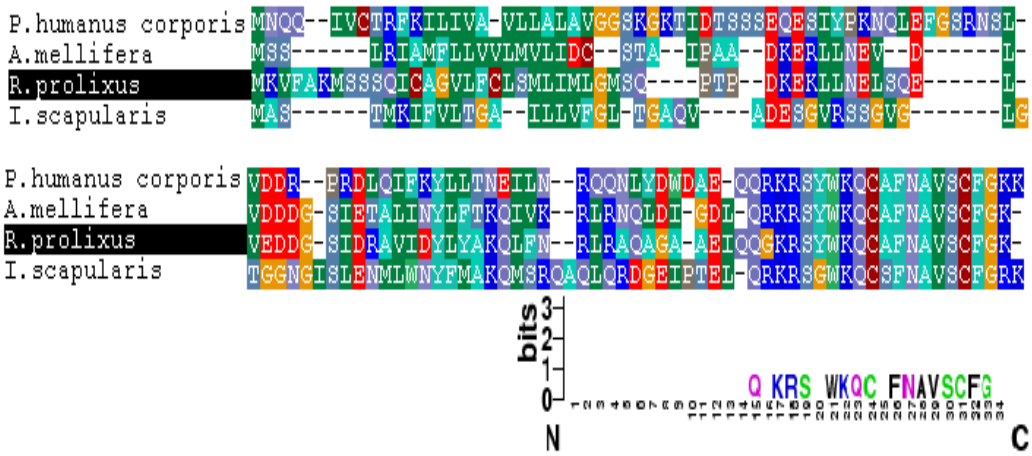
Figure 1.80: Estructura del gen de pro-hormona 1. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra cada exón o grupo de exones se indica en la parte superior. Las líneas representan los intrones.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

péptidos predichos	**péptidos encontrados**	**[M+H]$^+$**	**Tejido**
QPTPDKEKLLNELSQELVEDDGSIDRAVIDYLYAKQLFNRLRAQAGAAEIQQ-NH_2	(pyro-Q)PTPDKEKLLNELS		SN
	(pyro-Q)PTPDKEKLL		SN
	AQAGAAEIQQ-NH2	985,51	SN
	RAQAGAAEIQQ-NH2	1141,61	SN
SYWKQCAFNAVSCF-NH_2		1652,74	

Tabla 1.27: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-PH1* y los péptidos maduros encontrados por MS. [M+H]$^+$masa monoisotópica protonada.

Figura1.81: Alineamiento del precursor de *Rhopr-PH1* con los precursores ortólogos caracterizados en distintas especies de insectos. El contenido informativo expresado en bits se muestra para cada posición. Al igual que en el precursor de SIFa se pueden observar dos Cys separadas porv seis residuos.



Pro-hormona 2

Se han detectado péptidos codificados en el precursor de Prohormona 2 en el complejo CC-CA de *A. mellifera,* de lo que se deduce una probable función hormonal (Boerjan *et al*, 2010).

En nuestro trabajo hemos detectado un gen que codifica *Rhopr-PH2* en la secuencia genómica (Figura 1.83). La expresión del mismo fue confirmada por la presencia de su transcripto en la base de datos de EST total (*Whole Body*), y de péptidos maduros detectados por MS en SN (Tabla 1.28). Los exones 1, 2, 3, 4 y 5 de este gen se encuentran duplicados en el genoma. Los mismos se encuentran en los contigs 6856 y

165, mientras que los exones 6 y 7 se encuentran solo en el contig 165, y el exón 8 en el contig 2683 (Figura 1.82). Los 5 exones que se han duplicado presentan varias sustituciones de nucleótidos.

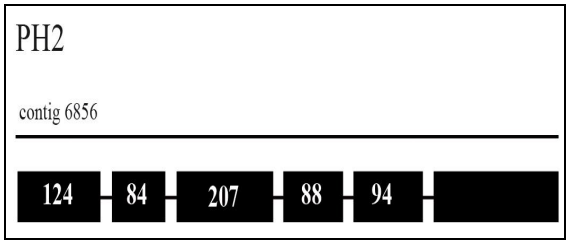


Figure1.82: Estructura del gen *Rhopr-PH2*. El número en cada exón indica la longitud en nucleótidos. Los *contigs* en los que se encuentran los exones se indican en la parte superior. Las líneas representan los intrones

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

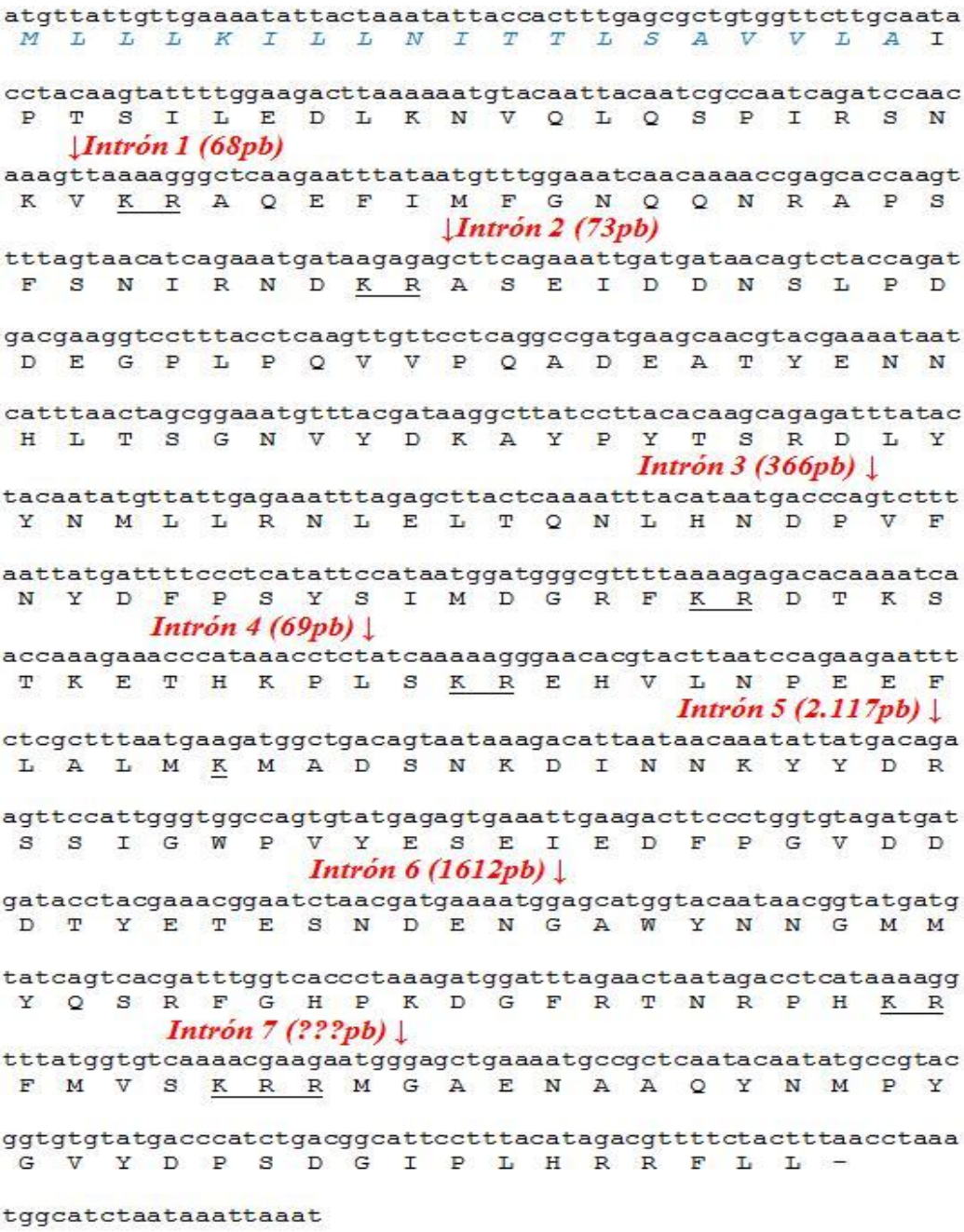


Figura 1.83: Secuencia nucleotídica y aminoacídica de *Rhopr-PH2*. El péptido señal (probabilidad SignalP-HMM: 0.990) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.28: En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-PH2* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
FMVSKRRMGAENAAQYNMPYGVYDPSDGIPLHRRFLL		4300,14	
AQEFIMFGNQQNRAPSFSNIRND		2684,26	
ASEIDDNSLPDDEGPLPQVVPQADEATYENNHLTSGNVYDKAYPYTS-RDLYYNMLLRNLELTQNLHNDPVFNYDFPSYSIMDGRF		9804,54	
DTKSTKETHKPLS		1471,78	
EHVLNPEEFLALM	EHVLNPEEFLALM	1541,77	SN

	EHVLNPEEFLAL(M-OH)	1557,77	SN
	EHVLNPEEFLAL	1393,72	SN
MADSNKDINNKYYDRSSIGWPVYESEIEDFPGVDDDTYETESNDEN-GAWYNNGMMYQSRFGHPKDGFRTNRPH		8554,73	
FMVSKRRMGAENAAQYNMPYGVYDPSDGIPLHRRFLL		4283,13	

Pro-hormona 3

Hemos detectado la secuencia parcial de un gen que codifica *Rhopr-PH3* en la base de datos genómica de *R. prolixus* (Figura 1.84). Su expresión ha sido confirmada por MS.

Secuencia parcial del mRNA de PH3 in silico

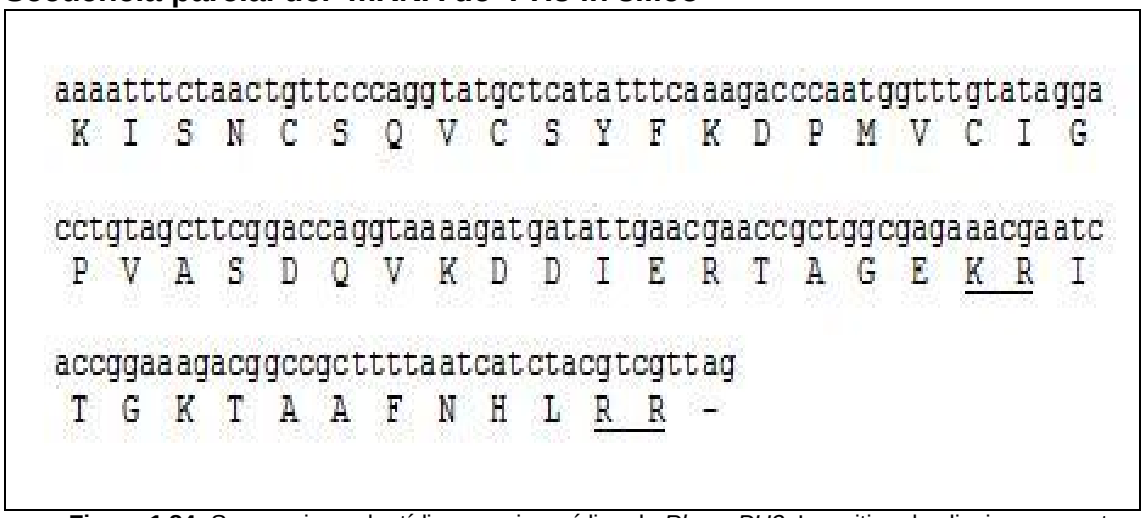


Figura 1.84: Secuencia nucleotídica y aminoacídica de *Rhopr-PH3*. Los sitios de clivaje se muestran subrayados.

Tabla 1.29 En la tabla se muestra los péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-PH3* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	péptidos encontrados	[M+H]+	Tejido
…..QVCSYFKDPMVCIGPVASDQVKDDIERTA GE			
ITGKTAAFNHL	ITGKTAAFNHL	1172,64	SN

CCHamida

Este neuropéptido fue caracterizado por primera vez en 2008 en B. mori (Roller et al, 2008). Fue denominado CCHamida por presentar dos residuos Cys conservadas y una His amidada en el extremo C terminal (Figura 1.90). En todos los insectos cuyo genoma ha sido secuenciado se ha encontrado este neuropéptido mediante búsquedas bioinformáticas, pero su expresión solo fue demostrada en B. mori, tanto en el SNC como en las células endócrinas del digestivo medio. Su función permanece desconocida.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

En nuestro trabajo hemos determinado la expresión de Rhopr-CCHamida en el SN de R. prolixus. El ORF completo del mensajero que codifica este precursor, así como su UTR 5´, fue caracterizado mediante RACE-PCR (Figura 1.88). Por otra parte, hemos confirmado su expresión al detectar la presencia de uno de los péptidos maduros generados a partir del clivaje del precursor en extractos de cerebro (Tabla 1.31)

Figura 1.88: Secuencia nucleotídica y aminoacídica de *Rhopr-CCHamida*. El péptido señal (probabilidad SignalP-HMM: 0.986) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

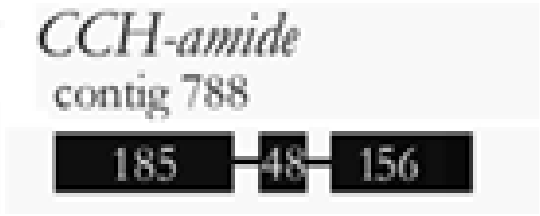


Figure 1.89: estructura del gen *Rhopr-CCHamida*. El número en cada exón indica la longitud en nucleótidos. El *contig* en que se encuentra el gen se indica en la parte superior. Las líneas representan los intrones.

Tabla 1.31: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-CCHamida* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
AAFKGARDGDASF	AAFKGARDGDASF	1312,63	Cerebro
GGCSAFGHSCFGGH-NH_2		1322,52	
SDDYMAQIQSRQLQRLPPADIVRQWVSVFCISNLL		3014,53	

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 1.90: Alineamiento del precursor de *Rhopr-CCHamida* con los precursores ortólogos caracterizados en distintas especies de insectos. Se pueden apreciar 3 Cys conservadas en el precursor, y dos en el péptido activa, separadas por 6 residuos. El precursor de *Ixodes scapularius* es más corto que los demás. Es probable que este precursor no esté completo, sino que el ORF comience en otro residuo de metionina previo, ya que no posee péptido señal.

1.3.2.k) GPN identificados solo *in silico*

Para algunos de los GPN detectados en el genoma, su expresión no ha sido confirmada por técnicas transcriptómicas ni peptidómicas.

Péptido similar a AKH/Corazonina

El gen que codifica este péptido en *T. castaneum* fue descripto inicialmente como AKH 3 (Li, 2008). Péptidos maduros de esta familia habían sido encontrados mediante MS en *L. migratoria* (Siegert, 1999) y *A. gambiae* (Kaufmann *et al*, 2006), y llamados con el mismo nombre. Recientemente, el sistema de señalización celular conformado por este péptido y su receptor fue el último en ser descripto en insectos como un sistema de señalización independiente (Hansen *et al*, 2010). Tanto el ligando como el receptor tienen una estructura que es intermedia entre los sistemas de AKH y CZ , por eso se lo ha llamado ACP : “AKH/CZ-like peptide”. Este péptido demostró no activar los receptores de AKH y CZ en *T. castaneum*, mientras que AKH y CZ no activan el receptor de ACP (Hansen *et al*, 2010). Estos resultados prueban que se trata de un sistema de señalización independiente, generado posiblemente por duplicación génica seguida de mutaciones y co-evolución del ligando y el receptor. De acuerdo al patrón de expresión temporal en *T. castaneum,* los autores sugirieron un papel de este sistema en el desarrollo larval, aunque los experimentos de RNAi no mostraron efectos sobre este proceso, ni ningún fenotipo distintivo (Hansen *et al*, 2010).

Luego de su identificación, el sistema ACP fue encontrado en varias especies de insectos pertenecientes a distintos órdenes: *Díptera, Lapidóptera, Coleóptera, Hymenóptera,* y *Hemíptera*, lo que significa que es un sistema que se ha originado tempranamente durante la evolución de los artrópodos. Sin embargo, no se ha encontrado en ninguna de las 12 especies de *Drosophila (Díptera)* ni *A. mellifera (Hymenóptera)*, por lo que este sistema se ha perdido varias veces durante el proceso de evolución (Hansen *et al*, 2010).

En *R. prolixus* hemos podido detectar el ORF completo para este GPN. El precursor presenta la misma estructura que CZ, AKH, CCAP y SIFamida, con el péptido biológicamente activo precedido inmediatamente por el péptido señal, y separado por un sitio de clivaje de otro péptido espaciador o acompañante en el extremo C-terminal de precursor.

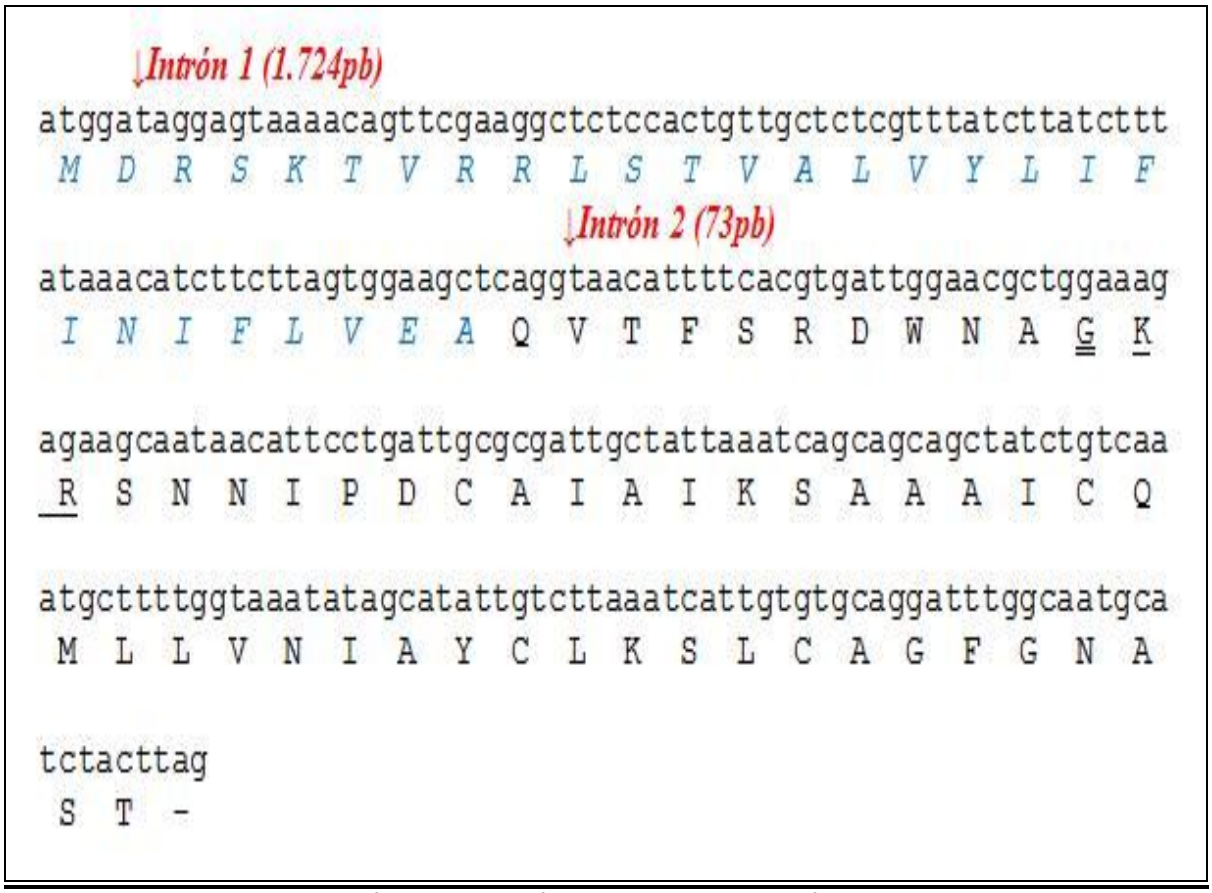


Figura 1.91: Secuencia nucleotídica y aminoacídica de *Rhopr-ACP*. El péptido señal (probabilidad SignalP-HMM: 0.851) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.32: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ACP* y los péptidos maduros encontrados por MS. $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
QVTFSRDWNA-NH_2		1222,6	
SNNIPDCAIAIKSAAAICQMLLVNIAYCLKSLCAGFGNAST		4173,07	

Factor Dispersor de Pigmentos

Este neuropéptido (PDF) es homologo a la hormona Pigment-Dispersing Hormone de crustaceos. En insectos, su expresión ha sido detectada tanto en el protocerebro como en los

lóbulos ópticos, y se lo ha asociado a una función efectora dentro del reloj circadiano en L. maderae y D. melanogaster (Petri and Stengl, 1997; Park and Hall, 1998)

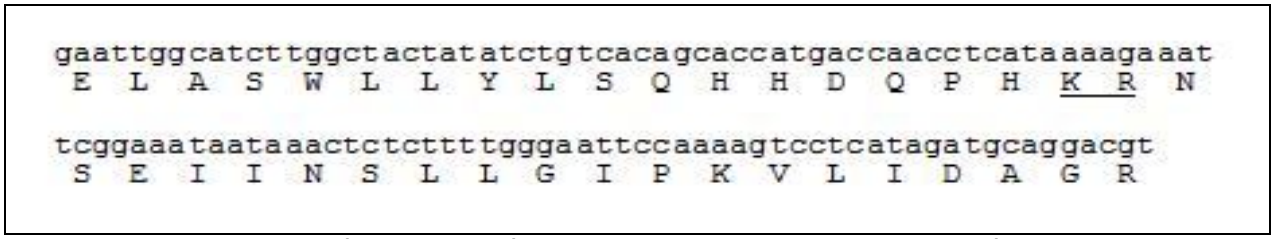


Figura 1.92: Secuencia nucleotídica y aminoacídica de *Rhopr-PDF*. Los sitios de clivaje di-básicos se muestran subrayados.

Péptidos simil-insulina 1 y 3

Al igual que ILP-2, ILP-1 e ILP-3 pertenecen a la familia de las Insulinas tipo B, las que se caracterizan por el motivo CCXXXC.

Secuencia parcial del mRNA de ILP 1 in silico

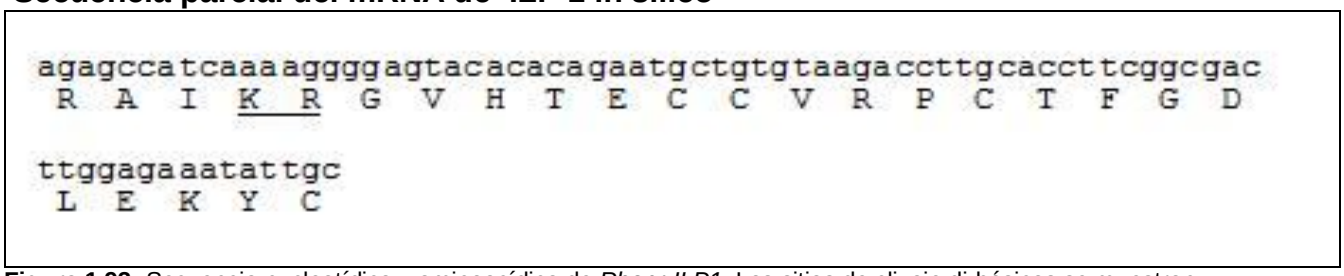


Figura 1.93: Secuencia nucleotídica y aminoacídica de *Rhopr-ILP1*. Los sitios de clivaje di-básicos se muestran subrayados.

Péptido simil-insulina 3

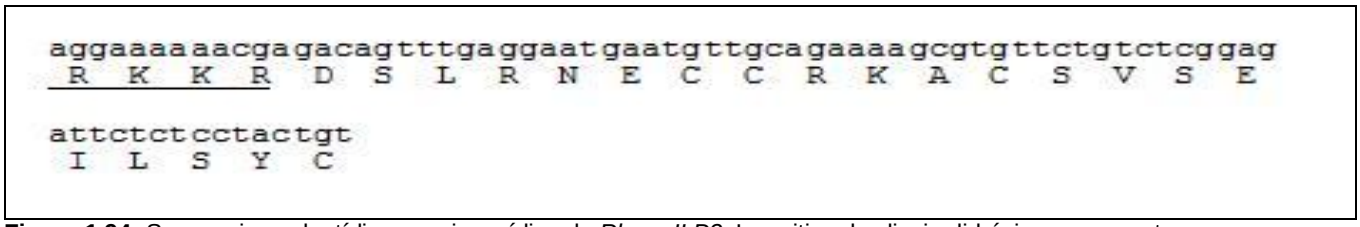


Figura 1.94: Secuencia nucleotídica y aminoacídica de *Rhopr-ILP3.* Los sitios de clivaje di-básicos se muestran subrayados.

Péptido simil-insulina 6

A diferencia de las insulinas del grupo B (ILP-B), las ILP-A se caracterizan por el motivo CCXXXXC. En *R. prolixus* hemos podido podido predecir un gen que codifica un péptido de esta familia (Figura 1.95).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 1.95: Secuencia nucleotídica y aminoacídica de *Rhopr-ILP6.* El péptido señal (probabilidad SignalP-HMM: 0.986) se muestra en azul y letra cursiva. Los sitios de clivaje se muestran subrayados. El residuo de Gly necesario para la amidación en el extremo C-terminal del péptido maduro se muestra con doble subrayado. La posición de los intrones se muestra en rojo y entre paréntesis la longitud de los mismos.

Tabla 1.33: Péptidos predichos a partir del clivaje proteolítico del precursor de *Rhopr-ILP6* y los péptidos maduros encontrados por MS, $[M+H]^+$masa monoisotópica protonada.

péptidos predichos	**péptidos encontrados**	**$[M+H]^+$**	**Tejido**
QRVTLCL-NH_2		831,49	
NQITLTAGLFTG-NH_2		1234,68	
VGSGNCV		635,28	
GLSPVVRLCCKSSGCAMKPLELWDQ		2720,34	

1.3.2.h) GPN aparentemente ausentes en el genoma

Algunos de los GPN conservados en insectos se encuentran aparentemente ausentesen el genoma de *R. prolixus*. Entre ellos se pueden mencionar AKH II, Allatostatina-C (ALS-C), Proctolina (Prt), simil-Arginina-Vasopresina(AVPL), Neuropeptide-like precursor-2 (NPLP2) y la hormona Protoracicotrópica (PTTH). Sin embargo, hemos podido detectar la presencia de un posible GPCR para ALS-C en el genoma de *R. prolixus* por aproximaciones bioinformáticas (E= 7e-148). En el caso de Prt y PTTH, dos péptidos con receptores conocidos, no se ha detectado ninguna secuencia homóloga a sus receptores, lo que refuerza la hipótesis de que estos sistemas de señalización se encuentran ausente en *R. prolixus.*

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

1.3.2.i) Neuropéptidos en hemolinfa

Una vez identificado en el SNC, la función hormonal de un neuropéptido se comprueba mediante la detección del mismo en la hemolinfa. Este último paso, sin embargo, es omitido en la mayoría de los trabajos debido a las complicaciones técnicas asociadas al análisis de hemolinfa. La presencia de moléculas en muy baja concentración, como los neuropéptidos, es enmascarada por la complejidad de la composición de la muestra y las concentraciones relativamente altas de otros péptidos y proteínas presentes, como los péptidos antimicrobianos.

Sin embargo, en el presente trabajo hemos realizado un intento para determinar la función humoral de los GPN caracterizados. Utilizando un protocolo que incluye separación por nano-HPLC y espectrometría de masas (MALDI-TOF MS/MS) pudimos encontrar dos OK, CZ, pyro-MS y SNF en la hemolinfa de *R. prolixus.* Para el caso de OK, en nuestrotrabajo pudimos apreciar inmuno-marcación positiva en el CC-CA y nervios abdominales de *R. prolixus,* ambos sidios de liberación neuro-hemal. La presencia de OK en hemolinfa demuestra por primera vez su papel hormonal en insectos. Es importante destacar que muchos de los GPN identificados en el SNC podrían estar presentes también en la hemolinfa, sin que hayan sido identificados por encontrarse en concentraciones por debajo del límite de detección de la técnica, o enmascarados por otras moléculas.

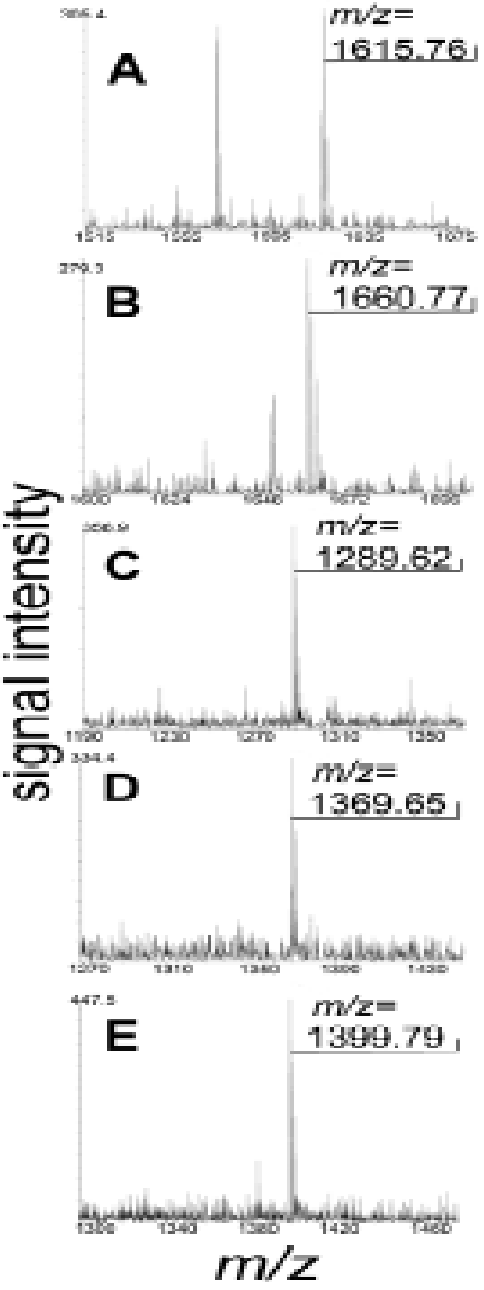


Figura 1.96: Péptidos identificados en hemolinfa mediante MALDI -TOF-TOF-MS (A) Orcokinina-1 **NFDEIDRVGFGSFI;** (B) Orcokinina-2 **NFDEIDRSGFNSFI;** (C) pyro-miosuppressina **pyro-QDIDHVFMRFa;** (D) pyro-corazonina **pyro-QTFQYSRGWTNa;** (E) Neuropéptido F Corto **NNRSPQLRLRFa;**

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

1.4) Conclusiones

En el presente trabajo se combinaron técnicas de bioinformática, genómica y peptidómica con el fin de realizar un análisis exhaustivo de los neuropéptidos y sus precursores en una de las especies transmisoras de la enfermedad de Chagas: *R. prolixus*. Sorprendentemente, este tipo de estudio nunca antes había sido aplicado para el estudio de insectos de interes sanitario.

Hemos logrado identificar *in silico* 32 GPN, y por medio de técnicas transcriptómicas y/o proteómicas hemos confirmado la expresión de 28 de estos genes en SN, así como formas de *splicing* alternativas para 4 de ellos. Se detectaron 107 péptidos maduros mediante MS. Este trabajo contribuye a la anotación de genes en el genoma de *R. prolixus*. Además se presenta evidencia experimental de la expresión de GPNs en diferentes tejidos.

Mediante MS, además de los péptidos maduros predichos a partir del clivaje del precursor, se han encontrado formas truncadas en el extremo C y N terminal para muchos de ellos. De la misma manera, se han podido identificar péptidos inmaduros o en proceso de maduración.

Para los GPNs descriptos en *R. prolixus* se pueden encontrar ortólogos con motivos ampliamente conservados en otros insectos. La conservación de dominios entre especies filogeneticamente lejanas, es prueba de que se trata de secuencias bajo presión de selección. Por otro lado, la estructura de los precursores se encuentra conservada. Como ejemplos se pueden mencionar AKH, ACP, CZ y SIFamida, donde el péptido biológicamente activo se encuentra localizado inmediatamente después del péptido señal, y es seguido por un péptido asociado más largo localizado en el extremo C-terminal. Precursores como TK, ALS, NPLP1 por su parte, codifican múltiples péptidos cortos separados por sitios de clivaje di-básicos.

Para algunos de los GPNs cuya transcripción ha sido probada por RACE-PCR en SN, no se han detectado los péptidos maduros mediante MS. Esto puede deberse al mayor tamaño de los mismos, lo que impide una extracción eficiente, por ejemplo AKH, ILPs, ITP, NPA, EH y ETH. Por otro lado estos péptidos presentan puentes disulfuro dentro de su estructura, lo que dificulta su extracción.

La evolución de los sistemas de señalización celular está asociada a la pérdida o duplicación de los genes que codifican neuropéptidos y sus receptores. La pérdida “frecuente” de sistemas peptídicos de señalización celular es una característica única de artrópodos (en comparación con vertebrados). Para que se forme un nuevo sistema de señalización celular es necesario que tanto el ligando como el receptor se hayan duplicado, mutado y haya actuado la presión de selección sobre ellos, lo que da como resultado un sistema de señalización independiente que regula un proceso diferente. Nuestros resultados indican que en *R. prolixus* parte de los genes de DH31 y PH2 se han duplicado, mientras que otros sistemas de señalización parecen haberse perdido, como PTTH y Prt. Además de la duplicación de genes, el *splicing* alternativo, las diferentes modificaciones post-traduccionales y la expresión diferencial de genes en distintos tejidos, son mecanismos efectivos para controlar diferentes funciones biológicas mediante moléculas estructuralmente relacionadas.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Se ha podido determinar la expresión de péptidos con actividad regulatoria en el digestivo medio anterior, digestivo medio posterior, testículos y glándulas salivales. La expresión de GPN en estos órganos brinda indicios acerca de la posible función de los péptidos, ya que posiblemente actúen de manera autócrina o parácrina regulando la actividad del órgano en que se expresan.

En conjunto, nuestros resultados constituyen un importante aporte al conocimiento del sistema neuroendócrino de triatominos, y el punto de partida para estudios fisiológicos, farmacológicos y moleculares orientados al diseño de insecticidas de nueva generación. Los datos presentados, incluyendo la secuencia de los GPN y la identificación de PTM en péptidos maduros, aportan herramientas para la investigación en endocrinología de insectos. A partir de nuestros resultados es posible realizar estudios funcionales tales como RNAi, peptidómica cuantitativa, estudios de hibridización *in situ*, etc, a fin de lograr un conocimiento profundo de la fisiología de *R. prolixus*.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Capítulo 2

Dinámica del neuropeptidoma de *Rhodnius prolixus* en respuesta a la ingesta sanguínea

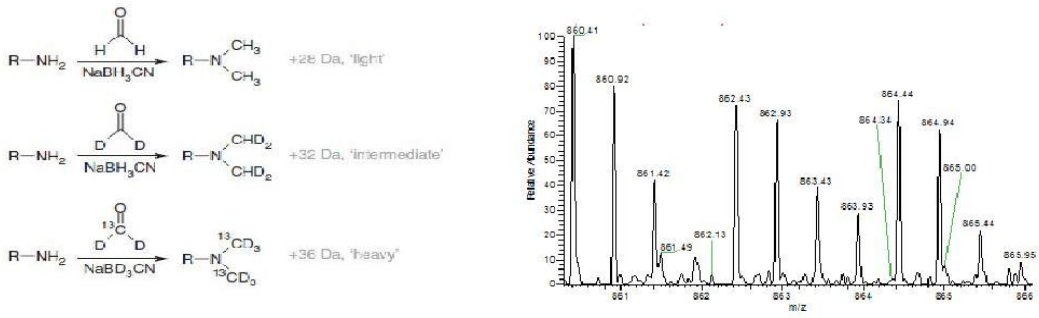
2.1) Introducción

Una vez finalizada la etapa descriptiva, donde caracterizamos el neuropeptidoma de *R. prolixus*, realizamos ensayos de peptidómica cuantitativa a fin de comenzar a caracterizar el posible papel fisiológico de las moléculas descriptas en los procesos desencadenados por la ingesta sanguínea.

En insectos hematófagos, el sistema neurosecretor se activa rápidamente con la alimentación. Este proceso causa varios cambios a nivel fisiológico y desencadena procesos inmediatos como la diuresis, y otros no inmediatos como la digestión, la muda en ninfas y reproducción en adultos. Todos estos procesos se encuentran regulados por neuropéptidos presentes en el SN. Por eso resulta importante estudiar los cambios en el sistema neurosecretor a lo largo del tiempo en respuesta a la alimentación. El proceso de diuresis es crítico en insectos hematófagos, e involucra la actividad coordinada de diferentes órganos, entre los que se incluyen el digestivo medio anterior, los TM y el recto. Este proceso es regulado por neuropéptidos, tanto diuréticos como anti diuréticos, así como también por Serotonina (ver Orchard, 2006). Es durante la alimentación cuando se produce la transmición de *T. cruzi* a través de las heces que los insectos triatominos depositan sobre el huésped. En *R. prolixus*, el proceso de diuresis dura entre 3 y 4 hs. Durante este tiempo, aproximadamente el 50% del volumen adquirido es eliminado (Maddrell, 1964). Por otro lado, los componentes nutritivos de la sangre, glóbulos rojos y proteínas presentes en el plasma, son almacenados en el digestivo medio anterior, y luego transferidos hacia el digestivo medio posterior, donde se realiza la digestión y asimilación de los nutrientes. Estos procesos tienen lugar en los días posteriores a la alimentación. Durante los períodos de ayuno entre cada ingesta sanguínea, es necesario para los insectos conservar agua y por lo tanto no hay producción de orina.

Tradicionalmente, la cuantificación de péptidos en diferentes condiciones fisiológicas se realizaba mediante ELISA. En la actualidad, los adelantos de las técnicas basadas en peptidómica cuantitativa permiten un análisis más rápido y global de la expresión de los neuropéptidos. Recientemente, se ha desarrollado el etiquetado de péptidos en el extremo N-terminal y los grupos amino *epsilon* de los residuos Lys con isótopos de dimetilo (Boersema *et al*, 2009). El uso de ^{2}H y ^{13}C combinados diferencialmente permite la comparación de tres condiciones diferentes; la dimetilación de las aminas primarias genera un incremento de masas de 28, 32 o 36 Da (figura 2.1), que puede ser detectado por espectrometría de masas, y permite la cuantificación relativa de los péptidos etiquetados con cada isótopo mediante comparación de las intensidades de picos generados por los iones precursores. La cuantificación con este método es llevada a cabo a nivel del espectro de masa, mientras que el espectro de masa-masa (Tadem MS) se utiliza solo para la identificación de los péptidos.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes



Label	Light	Intermediate	Heavy
Mass increase per label	+28.0313 Da	+32.0564 Da	+36.0757 Da
Formaldehyde isotope	CH_2O	CD_2O	$^{13}CD_2O$
Cyanoborohydride isotope	$NaBH_3CN$	$NaBH_3CN$	$NaBD_3CN$

Figura 2.1: Esquema del proceso de etiquetado con isotopos de dimetilo (tomada de Boersema *et al*, 2009). Se muestra a modo de ejemplo el espectro de masas del péptido NHDTLDSLSGITPGSQ perteneciente a la familia de las Orkokininas, marcado diferencialmente con los tres isótopos

Dado que las propiedades químicas de los diferentes isótopos son las mismas, los péptidos iguales presentes en las muestras marcadas con diferentes isótopos eluyen con los mismos tiempos de retención de la columna de HPLC. Este método de etiquetado presenta varias ventajas respecto a otras aproximaciones: es económico, proporciona una eficacia de 100% de etiquetado en casi todos los casos y es aplicable a virtualmente cualquier muestra.

2.2) Metodología de trabajo

Con el fin de identificar los neuropétidos involucrados en la regulación de los procesos desencadenados por la ingesta sanguínea (diuresis, anti-diuresis, reproducción) hemos llevado a cabo un análisis peptidómico cuantitativo del SN de *R. prolixus* en respuesta a la ingesta de sangre. Este análisis nos ha permitido obtener información acerca de los cambios en la concentración de neuropéptidos en respuesta al proceso de alimentación. Para esto se utilizó un protocolo que involucra *on line* nano-LC-ESI-Orbi-TRAP MS y marcación isotópica estable de los extractos de SN con distintos isótopos de dimetilo. Se analizaron tres grupos experimentales: insectos sin alimentar (SA), 4 horas post alimentación (4H) y 24 horas post alimentación (24H; Figura 2.2).

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

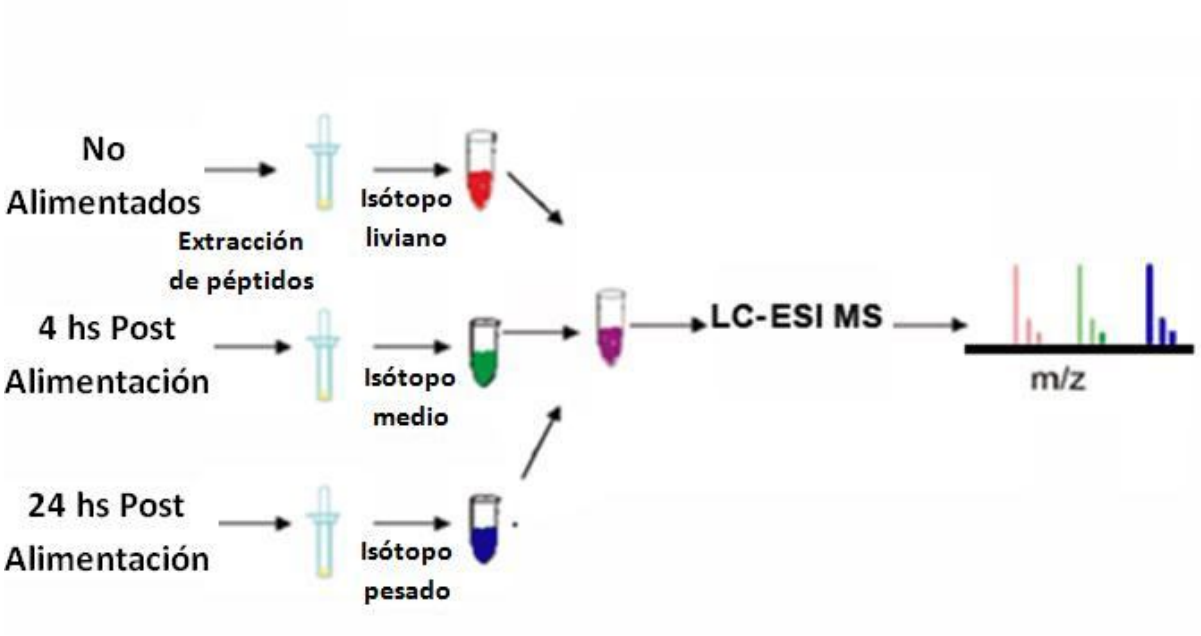


Figura 2.2: : Diseño experimental empleado para la determinación de los cambios en el nivel de neuropéptidos en el SN en respuesta a la ingesta sanguínea

2.2.1) **Insectos y diseño experimental**

R. prolixus adultos fueron obtenidos de la colonia de nuestro laboratorio, mantenida bajo las condiciones descriptas en el Capítulo 1. Todos los insectos utilizados en cada réplica experimental mudaron a la etapa adulta en la misma fecha. Éstos fueron alimentados con sangre de gallina y sacrificados a las 4 horas (4H) o 24 horas (24H) después de la ingesta. Los insectos del grupo SA no fueron alimentados durante el estado adulto (Fugura 2.2). Se analizaron cinco replicas experimentales, cada una conteniendo SNs provenientes de 10 individuos en cada condición alimentaria (SA, 4H y 24H). Las disecciones de los SN se realizaron en solución salina de *R. prolixus*. Las muestras fueron almacenadas a -80°C hasta su utilización.

2.2.2) Extracción de péptidos

Los péptidos fueron extraídos en metanol ácido (metanol/agua/el ácido acético 90/9/1, v/v/v), sonicados por 5 minutos y centrifugados por 10 minutos a 7500g. El sobrenadante fue recogido y el precipitado fue re-extraído dos veces mas. Los solventes orgánicos fueron evaporados por vacío en speedvaccum a 5000g.

2.2.3) Etiquetado de la muestra con isótopos de demetilo

Las muestras pertenecientes a insectos que no habían sido alimentados (SA) fueron etiquetadas con el isótopo liviano (CH_2O), las muestras de insectos disecados 4 horas (4H)

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

después de la ingesta sanguínea fueron etiquetadas con el isótopo mediano (CD_2O), y las muestras provenientes de insectos disecados 24 horas (24H) después de la alimentación, fueron etiquetadas con el isótopo pesado ($^{13}CD_2O$, Figura 2.2). La reacción de etiquetado fue realizada según el protocolo "*In Solution*" descripto por Boersema PJ (2009). Brevemente, las muestras fueron re diluidas en 100 µl TEAB 100mM (Sigma, St. Louis, MO). 4 µl del isótopo dimethyl 4% v/v correspondiente (CH_2O, CD_2O o $^{13}CD_2O$) fueron agregados a cada muestra. Luego se añadieron 4 µl de $NaBH_3CN$ 0,6 M a las muestras SA y 4H, y 4 µl de $NaBD_3CN$ 0,6 M a la muestra 24H. Las muestras fueron incubadas 1,5h a 20°C y posteriormente transferidas a hielo. Se agregaron 16 µl de NH_3-H_2O y 8 µl de ácido fórmico (FA) para detener la reacción de etiquetado. Las muestras SA, 4H y 24H fueron mezcladas juntas en relación 1:1:1. El solvente fue evaporado bajo vacío y la muestra re-diluída en 20 µL 0,1% TFA.

2.2.4) Desalado de la muestra

La muestra conteniendo 10 SNC de insectos en cada estado alimenticio (SA, 4H y 24H) fue desalada según lo descripto en el Capítulo 1. Luego de evaporado el solvente las muestras fueron re-suspendidas en 12 µl 5% CH_3CN/1% FA para ser analizadas mediante MS.

2.2.5) Cromatografía líquida y espectrometría de masas (LC-MS)

Las muestras correspondientes a cada réplica experimental fueron divididas en dos y analizadas por duplicado mediante *on line* nanoLC-Electrospry (ESI)-Orby-Trap Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Como resultado se obtuvieron dos réplicas de cada una de las cinco muestras biológicas. Las muestras fueron cargadas en la pre-columna del HPLC a un flujo de 10 µl/min durante 5 min. en buffer A (0,1% v/v TFA). Posteriormente los péptidos fueron eluídos en un gradiente lineal de 10-60% v/v de solvente B (80% v/v ACN, 0.1% v/v TFA en agua) durante 38 min. Finalmente los péptidos restantes fueron eluídos con 100% buffer B durante 7 min. En la figura 2.3 se muestra una de las corridas de HPLC.

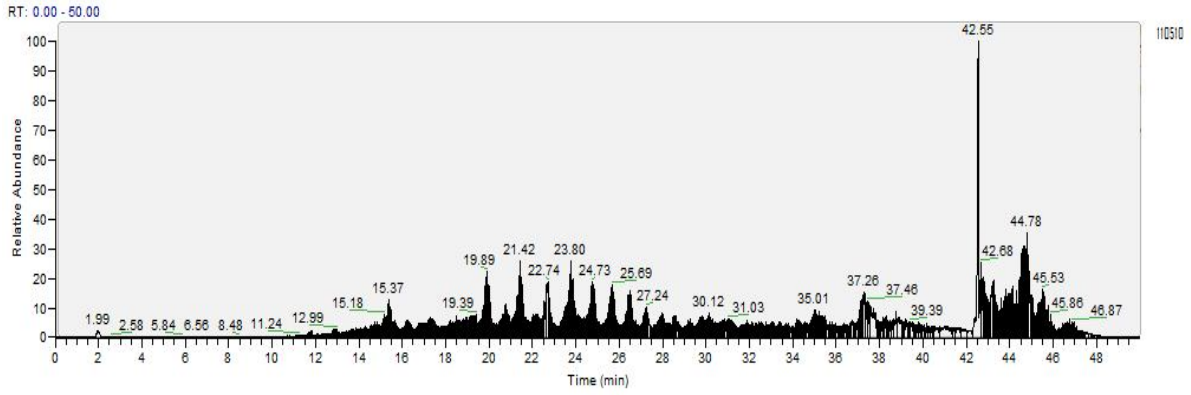


Figura 2.3: Cromatograma de HPLC de una muestra biológica utilizada para la cuantificación relativa de los péptidos

2.2.6) Identificación y cuantificación de péptidos

Los espectros fueron procesados y las listas de picos fueron utilizadas para la búsqueda en la base de datos usando los programas MASCOT 2.1.0 (Matrix Science,) y PEAKS STUDIO 5.2 (Bioinformatics Solutions) con los siguientes parámetros: enzima: ninguna; modificaciones variables: Pyro-glutamico de Glutamina, amidación N-terminal; acetilación C-terminal, oxidación de metionina; Tolerancia de error de masa del ion parental: 10 ppm; Tolerancia de error de masa del fragmento: 0.5 Da. La base de datos usada para la búsqueda contenía todas las secuencias de *R. prolixus* presentes en en Genbank, incluídos los precursores generados a partir de los GPN identificados previamente.

La cuantificación de los péptidos se realizó con la ayuda del programa MSQuant (Mortensen *et al*, 2010). Los resultados fueron validados o corregidos manualmente utilizando el programa *XCalibur* (Thermo Fisher Scientific, Waltham, MA, USA). Para los análisis estadísticos, las intensidades de los picos de los péptidos marcados con el isótopo liviano, medio o pesado fueron normalizadas dividiéndolas por la suma de las intensidades de los péptido marcados con los isótopos liviano y pesado. De esta manera, se eliminan las diferencias en las intensidades de los picos debidas a la eficiencia de ionización y a la sensibilidad del detector en las distintas corridas de HPLC-MS. Los resultados obtenidos fueron analizados mediante ANOVA de un factor. Cuando fue necesario se realizaron transformaciones logarítmicas a fin de obtener homogeneidad en las varianzas.

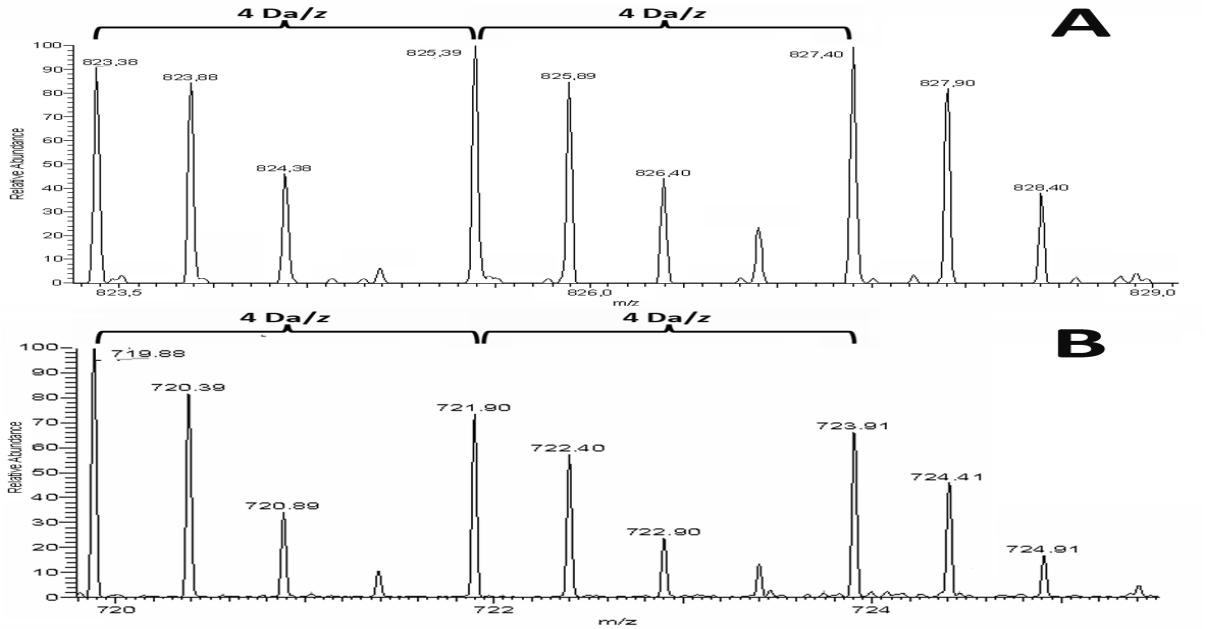


Figura 2.4: Espectros de masa donde se muestran dos péptidos marcados con distintos isótopos de dimetilo. Las intensidades de los picos correspondiente s a cada estado alimenticio fueron utilizadas para cuantificar las cantidades relativas del péptido. Los extractos de SN de insectos no alimentados fueron etiquetados con el isótopo liviano (CH_2O), los extractos provenientes de insectos sacrificados 4 horas post alimentación (4H) fueron marcados con el isótopo medio (CD_2H), mientras que las muestras correspondiente a insectos disecados 24 horas post alimentación fueron etiquetados con el isótopo pesado ($^{13}CD_2O$). Se puede observar una diferencia de 4 Da/z entre los péptidos marcados con los diferentes isótopos. (A) Espectro de masa de péptido NFDEIDRSGFDGFV perteneciente a la familia de las Orcokininas. (B) Espectro de masa de péptido EHVLNPEEFLAL generadoa a partir del precursor de Pro Hormona 2.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

2.3) Resultados y discusión

Mediante el análisis on line nanoLC-Electrospry (ESI)-Orby-Trap MS combinado con la marcación con isótopos de dimetilo, se han podido examinar estadisticamente los cambios cuantitativos de 18 neuropéptidos maduros provenientes de 9 precursores distintos en el SN de R. prolixus.

2.3.1) Péptidos que muestran cambios significativos en el SN luego de la ingesta sanguínea

Pro-hormonas 2 y 3

Hemos podido determinar que la concentración relativa en el SN del péptido EHVLNPEEFLAL, codificado por el GPN Rhopr-PH2, se reduce significativamente a las 4 ($p<0.01$) y 24 ($p<0.001$) horas post alimentación en relación a insectos sin alimentar. A su vez, este péptido se encuentra en menor concentración a las 24 hs que a las 4 hs post ingesta ($p<0.05$; Figura 2.5). La disminución observada indica que el péptido fue liberado en respuesta a la ingesta sanguínea en el período durante el cual la diuresis tiene lugar, y los niveles descienden aún más luego de transcurridas 24 hs, lo cual indica que durante el período entre 4 y 24 horas el péptido continúa siendo liberado. Los resultados sugieren la implicancia de Rhopr-PH2 en la regulación de los procesos inmediatos y no inmediatos desencadenados por la ingesta sanguínea.

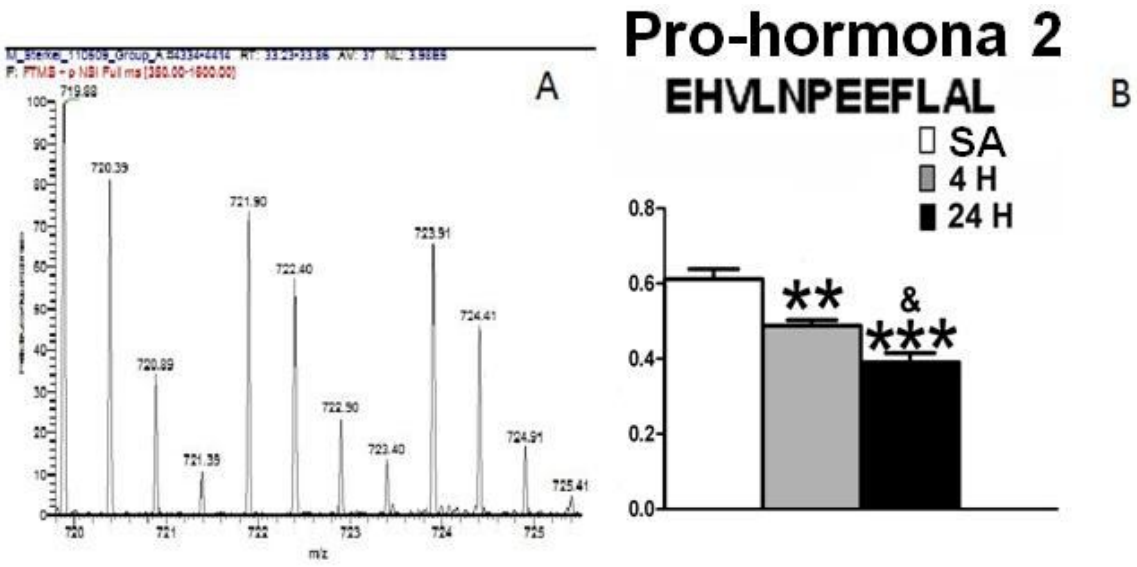
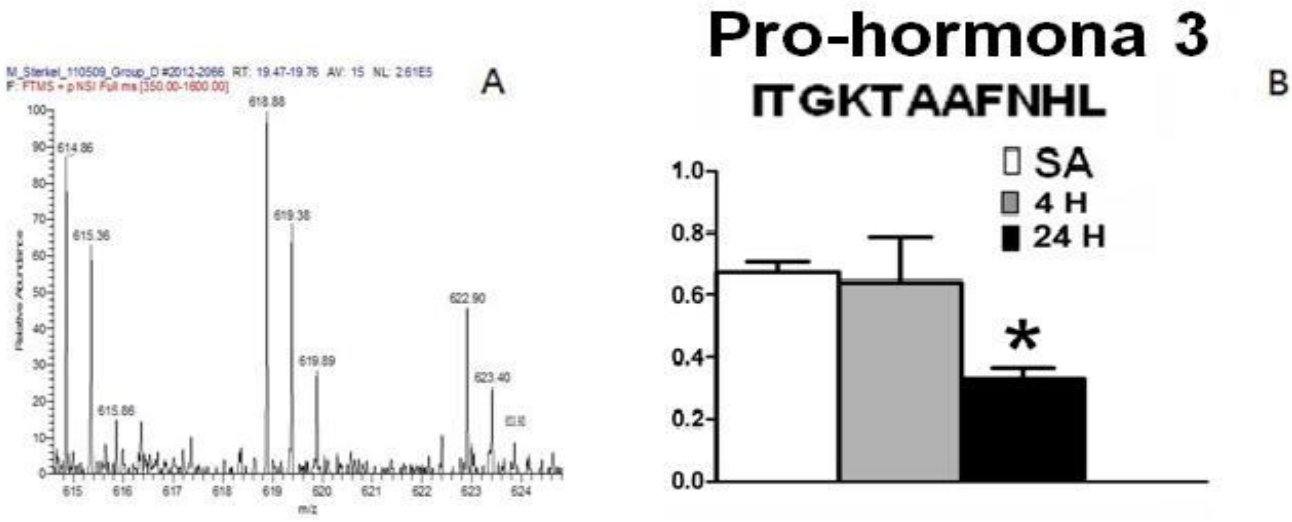


Figura 2.5: Cambios en la concentración del péptido EHVLNPEEFLAL codificado por el precursor Rhopr-PH2 en el SN de *R. prolixus* después de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representadas (medias ± SEM; N= 4-5). **=P<0.01 *vs*. el grupo sin alimentar, ***= P<0.001 vs el grupo sin alimentar, &= P<0.05 versus el grupo 4H. 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea. En la figura A se muestra un espectro MS representativo para este péptido.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Hemos encontrado cambios significativos en los niveles del péptido ITGKTAAFNHL codificado por el GPN *Rhopr-PH3* a las 24 hs. post-alimentación ($p<0.05$) mientras que los niveles entre los insectos no alimentados y 4 hs. no varían (Figura 2.6). Este hecho muestra que hay una liberación tardía del péptido desde el SN en respuesta a la alimentación, lo cual sugiere un papel en la regulación de los procesos que dan lugar a la reproducción.

Nuestros resultados sientan las bases para posteriores estudios fisiológicos orientados a determinar la función exacta de estos neuropéptidos, cuya función es totalmente desconocida en



insectos.

Figura 2.6: Cambios en la concentración del péptido ITGKTAAFNHL codificado por el precursor Rhopr-PH3 en el SN de *R. prolixus* después de la ingesta sanguínea. Las concentraciones del péptido en las diferentes condiciones se muestran representadas (medias ± SEM; N= 4-5). *P<0.05 *vs*. el grupo sin alimentar y 4H. 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea. La figura A muestra un espectro MS representativo para este péptido.

Kinina

Nuestros resultados muestran un incremento significativo en la concentración del péptido FLDHEDALI, codificado por el precursor de Kininas en *R. prolixus,* a la 4 hs. después de la alimentación ($p<0.01$), mientas que a las 24 hs. los niveles se asemejan a los de los insectos SA (Figura 2.7). Cabe destacar que, si bien este péptido es codificado por el GPN de Kinina, no presenta el dominio característico de la familia ($R\text{-}FX_1X_2WG\text{-}NH_2$), sino que es un péptido espaciador. El aumento de este péptido 4 hs. post-alimentación indica que la síntesis y/o maduración aumenta durante este período, sin que sea liberado. Por otro lado, a las 24 hs. post alimentación los niveles se asemejan a los de los insectos que no fueron alimentados, lo que indica que hay una liberación tardía de este péptido o bien que es degradado durante el período entre las 4 y 24 horas post alimentación. Estos resultados sugieren que este péptido, codificado por el precursor de kininas, no actua como un factor diurético en *R. prolixus*.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

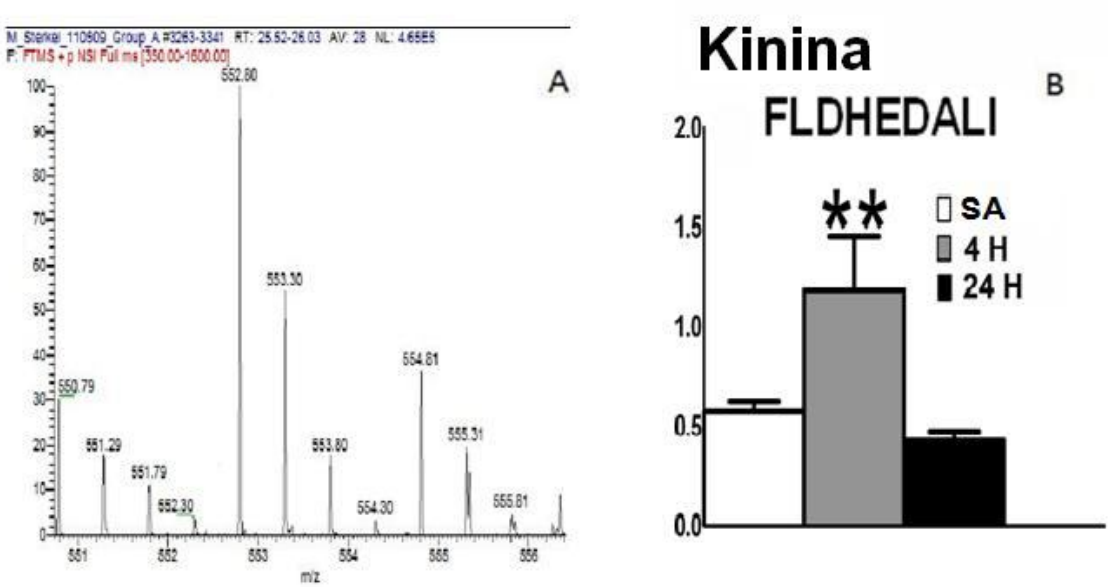


Figura 2.7: Cambios en la concentración del péptido FLDHEDALI (codificado por el precursor de Kininas) en el SN de *R. prolixus* después de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representadas (medias ± SEM; N= 4-5). **=P<0.01 *vs.* el grupo sin alimentar y 24H. 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea. se muestra un espectro representativo para este péptido.

NPLP-1

Nuestros resultados aportan evidencia que relaciona péptidos codificados por *Rhopr-NPLP1* con la regulación de los procesos fisiológicos desencadenados por la ingesta sanguínea. Resulta interesante el hecho de que estos péptidos presentan una regulación diferencial en respuesta a la alimentación. El péptido TVDAELEQLMSEVYGIGE muestra un aumento a las 4 hs. post-alimentación ($p<0.01$), mientras que la concentración a las 24 hs. es menor que en los insectos sin alimentar ($p<0.05$). Por otro lado el péptido SVASLARSYNLPNAVKGGYENDDE muestra una disminución significativa a las 24 hs ($p<0.01$), pero no a las 4 hs. respecto a los insectos sin alimentar (Figura 2.8). De los resultados se deduce que los péptidos generados por NPLP1 son regulados por mecanismos post-traduccionales, y ejercerían una función biológica diferente, lo que concuerda con la hipótesis de que interaccionan con diferentes receptores (ver Capítulo 1). Los resultados muestran una asociación de NPLP1 con los procesos fisiológicos que se desencadenan como consecuencia de la alimentación. Si bien es necesario continuar investigando para delinear la función exacta de estos péptidos, este es el primer reporte de la implicancia fisiológica de NPLP1, un neuropéptido ampliamente distribuído en insectos.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

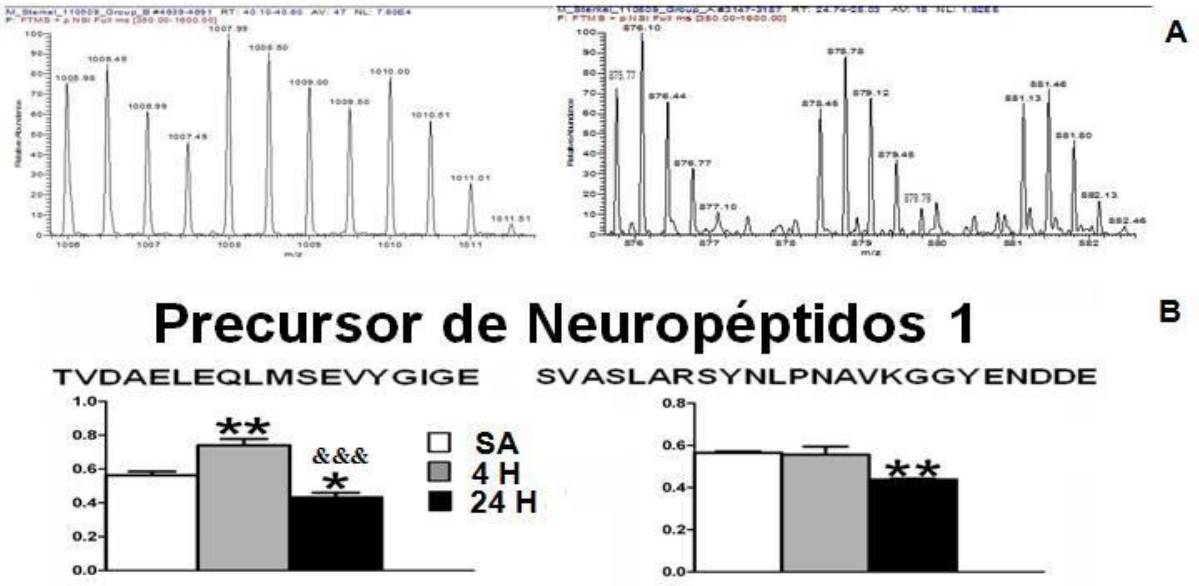


Figura 2.8: Cambios en la concentración de los péptidos TVDAELEQLMSEVYGIGE y SVASLARSSYNLPNAVKGGYENDDE (codificados por *Rhopr-NPLP1*) en el SN de *R. prolixus* luego de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representados (medias ± SEM; N= 4-5).*= P<0.05 vs. el grupo sin alimentar ,* *P<0.01 vs los demás grupos, &&&= P<0.001vs el grupo 4H. 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea. La figura A muestra un espectro representativo para cada péptido.

2.3.2) Péptidos que no muestran cambios significativos en el SN luego de la ingesta sanguínea.

Prohormona 1

Los resultados obtenidos de la cuantificación de tres péptidos generados a partir de este precursor no muestran variaciones significativas a las 4 ni a las 24 hs post alimentación (Figura 2.9), por lo que no estaría involucrado en la regulación de los procesos asociados a la ingesta sanguínea. Sin embargo, dos de los péptidos presentan gran variabilidad en el grupo de 4 hs. Esta variabilidad no se observa en los demás tiempos estudiados, por lo que no parece ser producto del protocolo experimental aplicado sino a una variabilidad propia de la mustra.

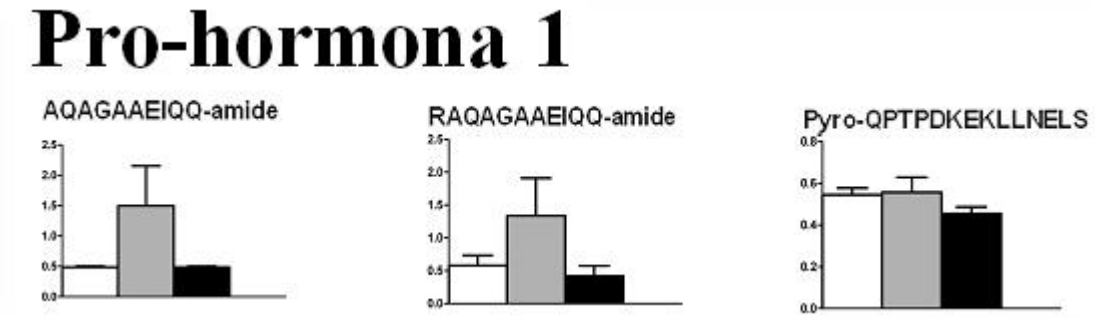


Figura 2.9: Cambios en la concentración de los péptidos AQAGAAEIQQamida, RAQAGAAEIQQamida y Pyro-QPTPDKEKLLNELS (codificados por *Rhopr-PH1*) en el SN de *R. prolixus* luego de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representados (medias ± SEM; N= 4-5). 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Neuropéptido F largo

Nuestros resultados no revelan cambios significativos en la concentración del péptido YAVAGRPRFamida, codificado por el precursor *Rhopr-LNF* en los tiempos estudiados (Figura 2.10).

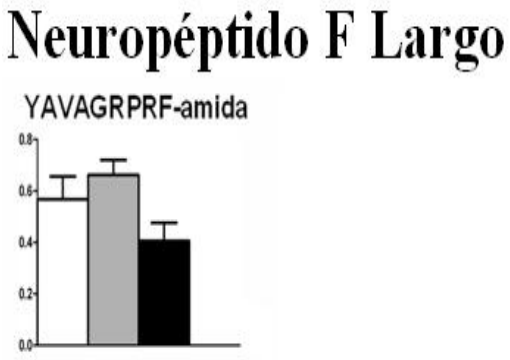
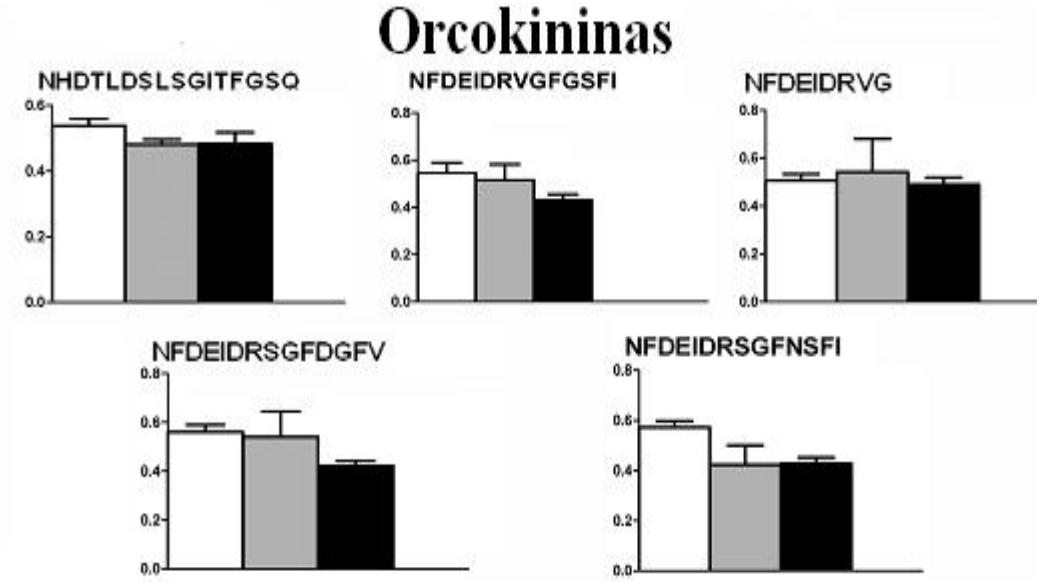


Figura 2.10: Cambios en la concentración del péptido YAVAGRPRFamida (codificado por *Rhopr-LNF*) en el SN de *R. prolixus* luego de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representados (medias ± SEM; N= 4-5). 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea.

Orcokininas

Varios péptidos codificados por el GPN de Orkokininas han sido identificados por MS en cerebro y SN, e incluso dos de ellos fueron encontrados en hemolinfa, confirmando una función humoral (ver Capítulo 1). Cuatro péptidos de este precursor pudieron ser cuantificados, y todos ellos mostraron concentraciones similares en los tres grupos experimentales estudiados, por lo que no estarían involucrados en los procesos desencadenados por la alimentación en insectos adultos. La función de los péptidos de esta familia, así como su receptor, es desconocida en insectos. Estudios cuantitativos de los cambios del neuropeptidoma en respuesta al proceso de alimentación en el cangrejo *Cancer borealis* arrojaron resultados similares para esta familia de péptidos (Chen *et al*, 2009).



Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Figura 2.11: Cambios en la concentración de los péptidos NHDTLDSLSGITPGSQ, NFDEIDRSGFDGFV, NFDEIDRSGFNSFI, NFDEIDRVGFGSFI y NFDEIDRVG (codificados por el precursor de Orcokininas) en el SN de *R. prolixus* luego de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representados (medias ± SEM; N= 4-5). 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea.

Taquikininas

No se han detectado cambios significativos en ninguno de los cuatro péptidos codificados por el precursor *Rhopr-TK* que fueron cuantificados. Sin embargo, es importante resaltar el hecho de que éstos exhiben gran variabilidad entre las distintas réplicas experimentales. Estos resultados sugieren la existencia de diferencias individuales en cuanto a la respuesta a la alimentación. Puesto que este aumento en la variabilidad se observa en todos los péptidos TK y no en el resto de los péptidos estudiados, no es atribuíble al procedimiento experimental, sino a una variabilidad propia de las muestras biológicas. En concordancia con nuestros resultados, en *A. mellifera* se ha visto que la cantidad de mRNA que codifica el precursor de TK varía considerablemente con el sexo y el tipo de labor llevado a cabo (Takeuchi *et al*, 2003; 2004; Brockmann, 2009).

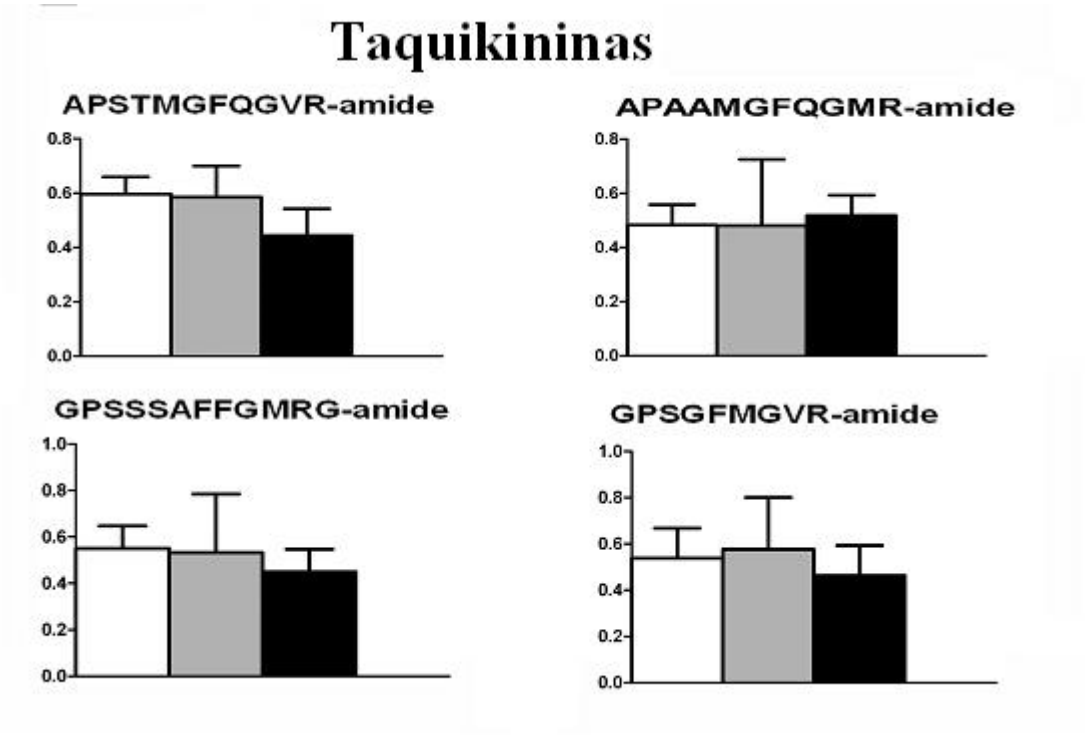


Figura 2.12: Cambios en la concentración de los péptidos APSTMGFQGVRamida, APSAAMGFQGMRamida, GPSSSAFFGMRGamida y GPSGFMGVRamida (codificados por el precursor *Rhopr-TK*) en el SN de *R. prolixus* luego de la ingesta sanguínea. Las cantidades del péptido en las diferentes condiciones se muestran representados (medias ± SEM; N= 4-5). 4H: insectos sacrificados 4 horas luego de la ingesta sanguínea. 24H: insectos sacrificados 24 horas luego de la ingesta sanguínea.

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

Conclusiones

Aquí presentamos el primer trabajo peptidómico cuantitativo en un insecto vector de enfermedades. Hemos podido determinar la implicancia de cinco neuropéptidos en la respuesta del sistema neuroendócrino a la ingesta sanguínea. De ellos, péptidos codificados por los precursores PH2, PH3 y NPLP-1 han sido descriptos recientemente por primera vez y su función fisiológica es totalmente desconocida. Nuestros resultados son un indicio de la asociación de estos neuropéptidos en la regulación de los procesos desencadenados por la ingesta sanguínea. Es interesante destacar el hecho de que los péptidos codificados por NPLP-1 que muestran variaciones en la cantidad en el SN luego de la alimentación se encuentran presentes en las glándulas salivales, órganos que tienen un papel fundamental durante este proceso.

Nuestro trabajo aporta indicios importantes para la caracterización funcional del sistema neuroendócrino, y su activación en respuesta a la ingesta sanguínea, y sienta las bases para posteriores ensayos fisiológicos a fin de determinar el proceso biológico particular que cada neuropéptido regula.

Discusión y Conclusiones Generales

En el presente trabajo se combinaron técnicas bioinformática con técnicas proteómicas y de biología molecular para caracterizar el conjunto de neuropéptidos que regulan la fisiología de *R. prolixus*. Las principales conclusiones a remarcar de este trabajo son:

- En el genoma se encontraron 32 nuevos genes que codifican precursores de neuropéptidos (GPNs), por lo que en total 34 GPNs han sido descriptos en este insecto. La expresión de 28 de estos genes pudo ser confirmada, y 107 péptidos maduros fueron detectados en el Sistema Nervioso mediante MS.
- Un péptido de la familia de las Sulfakininas, el péptido Miosupresina y péptidos generados a partir del precursor de Kininas y FMRFamida presentan particularidades a nivel de secuencia únicas en insectos. Este hecho es importante en el diseño y desarrollo de insecticidas de nueva generación, específicos hacia la especie a combatir, más seguros para el medio ambiente y la salud humana, y menos pasibles de generar poblaciones resistentes.
- Mediante RACE-PCR se pudieron identificar formas alternativas de *splicing* para los genes DH31, ITP, OK y CCAP. Las isoformas encontradas de OK y DH31 presentan un patrón de expresión diferencial en distintos tejidos.
- Por técnicas transcriptómicas y proteómicas se pudo determinar la expresión de NPA y la isoforma B de OK en el digestivo medio anterior, miosupresina en el digestivo medio posterior, la isoforma C de DH31 y CZ en testículos, y NPLP-1 en glándulas salivales. Resulta importante destacar el hecho que hasta la fecha no se sabía que las glándulas salivales eran capaces de producir péptidos con función regulatoria en ningún artrópodo. La presencia de estos péptidos en tejidos da indicios de su posible función, ya que probablemente los mismos actuen de manera parácrina o autócrina regulando las funciones de los órganos que los producen.
- Se han descripto varios neuropéptidos por primera vez en un insecto hemimetábolo, y su expresión ha sido confirmada. Entre ellos se pueden mencionar: SIFamida, NPLP1, PH1, PH2, PH3 y CCHamida.
- Utilizando técnicas de MS se ha determinado la función hormonal de SNF, CZ, MSP y Orcokininas por su presencia en hemolinfa. Esta es la primera vez que se demuestra que los péptidos de la familia de las Orcokininas actúan a nivel sistémico. La detección

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes

de neuropéptidos en hemolinfa es en si misma un desafío técnico, por la baja concentración en que se encuentran en relación a otros productos proteicos.

- Péptidos generados a partir de los precursores de PH2, PH3, Kininas y NPLP1 mostraron cambios significativos luego de la ingesta sanguínea al menos en alguno de los tiempos examinados. Cabe resaltar el hecho que la función de estos péptidos es desconocida. Los resultados obtenidos en este trabajo los vinculan a la regulación de los procesos desencadenados tras la ingesta sanguínea. Los dos péptidos generados a partir de *Rhopr-NPLP1* que pudieron ser cuantificados presentan una respuesta diferencial al proceso de alimentación, por lo que su expresión estaría regulada por mecanismos post traduccionales. Por otro lado, péptidos generados a partir del precursor de PH1, LNF, cinco péptidos de la familia de las Orcokininas y cuatro de la familia de las Taquikininas no mostraron cambios significativos en respuesta a este estímulo.

En resumen, nuestros resultados indican que el neuropeptidoma de *R. prolixus* es apenas diferente de los demás insectos con el genoma secuenciado y ensamblado, todos ellos holometábolos. Estas pequeñas diferencias podrían reflejar la metamorfosis incompleta de este insecto. Sin embargo, se requerán futuros estudios fisiológicos para determinar la función de cada uno de ellos.

Por último, la identificación de las particularidades únicas de los neuropéptidos a nivel de secuencia es el punto de partida para posteriores estudios a nivel fisiológico y farmacológico, que puedan llevar al diseño de una nueva generación de insecticidas como los peptidomiméticos.

Nuestro trabajo sienta las bases necesarias para posteriores estudios funcionales, así como también para estudios de interacción con los receptores y para estudios comparativos y/o evolutivos con otros artrópodos.

El aporte sustancial que el trabajo presentado constituye en la comprensión de la fisiología de triatominos es el punto de partida necesario, a la vez que un dinamizador, de futuros proyectos moleculares, proteómicos y fisiológicos, orientados a profundizar nuestro conocimiento en este campo, para poder aplicarlo en la lucha contra la enfermedad de Chagas.

Para citar este documento

Sterkel, Marcos. (2015) Genómica y proteómica de neuropéptidos en Rhodnius prolixus (Tesis de posgrado). Universidad Nacional de Quilmes, Bernal, Argentina: Repositorio Institucional Digital de Acceso Abierto. Disponible en: http://ridaa.demo.unq.edu.ar

Repositorio Institucional Digital de Acceso Abierto, Universidad Nacional de Quilmes