Purification and characterization of an insulin-related peptide in the desert locust, *Schistocerca gregaria*: immunolocalization, cDNA cloning, transcript profiling and interaction with neuroparsin

Liesbeth Badisco*, Ilse Claeys*, Matthias Van Hiel, Elke Clynen, Jurgen Huybrechts, Tim Vandersmissen, Sofie Van Soest, Luc Vanden Bosch, Gert Simonet and Jozef Vanden Broeck

Department of Animal Physiology and Neurobiology, Zoological Institute K U Leuven, Naamsestraat 59, PO Box 02465, B-3000 Leuven, Belgium

*Correspondence should be addressed to J Vanden Broeck; Email: jozef.vandenbroeck@bio.kuleuven.be

Abstract

Members of the insulin superfamily are not restricted to vertebrates, but have also been identified in invertebrate species. In the current report, we present the characterization of Scg-insulin-related peptide (IRP), an insulin-related peptide in the desert locust, *Schistocerca gregaria*. This peptide was isolated from corpora cardiaca (CC) extracts by means of a high-performance liquid chromatography (HPLC)-based purification strategy. Subsequent cloning and sequencing of the corresponding cDNA revealed that the encoded Scg-IRP precursor displays the structural organization that is typical for members of the insulin superfamily. Moreover, immunocytochemistry on brain tissue sections demonstrated the presence of Scg-IRP in median neurosecretory cells of the pars intercerebralis and their projections towards the storage part of the CC. Quantitative real-time RT-PCR studies revealed the presence of Scg-IRP transcripts in a variety of tissues, including nervous tissue and fat body. Furthermore, these transcripts showed a tissue- and phase-dependent, temporal regulation during the reproductive cycle of adult males and females. Finally, we demonstrated that Scg-IRP interacts in vitro with a recombinant neuroparsin, a locust protein displaying sequence similarity with vertebrate IGF binding proteins.

Journal of Molecular Endocrinology (2008) 40, 137–150

Introduction

Members of the insulin superfamily are not restricted to vertebrates but have also been identified in a large variety of invertebrate species where they are usually designated as ‘insulin-like peptides’ (ILPs) or ‘insulin-related peptides’ (IRPs). An invertebrate substance with insulin-like activities was first demonstrated in the mollusk, *Mya arenaria* (Collip 1923), and subsequently various studies suggested the existence of insulin-like substances in a wide variety of invertebrate species. A small *Bombyx mori* hormone displaying prothoracicotrophic activity in the related lepidopteran species, *Samia cynthia*, showed similarity to vertebrate insulin. The determination of its amino acid sequence represented the first structural identification of an insect IRP that was termed ‘bombyxin’ (Nagasawa et al. 1986). The second insect IRP to be identified was *Lom*-IRP, which had been isolated as a 5 kDa peptide from *corpora cardiaca* extracts of the migratory locust, *L. migratoria* (Hetru et al. 1991). Subsequently, IRP/ILPs have also been identified in various other insects (Wu & Brown 2006). Genome sequence information was the basis for the identification of seven distinct ILP genes in the fruit fly, *Drosophila melanogaster* (dilp 1–7), as well as in the malaria mosquito, *Anopheles gambiae* (AgamILP 1–7; Vanden Broeck 2001, Krieger et al. 2004). Similarly, eight ILP genes were discovered in the yellow fever mosquito, *Aedes aegypti* (AaegILP 1–8; Riehle et al. 2006), and two *Apis mellifera* ILPs (AmILP 1–2) were predicted from honey bee genome data (Wheeler et al. 2006). Furthermore, multiple studies in different metazoans have shown that the insulin/IRP signaling pathway is evolutionarily conserved and plays a crucial role in a range of fundamental and interrelated physiological processes, such as metabolism, growth, reproduction and ageing (as reviewed by Claeys et al. 2002). Interestingly, these processes also appear to be affected during phase transition in the desert locust, the species studied in the present report.

Locusts exhibit an extreme form of phenotypic plasticity: they can develop into two distinct ‘phases’, the solitarious and gregarious phase (Pener & Yerushalmi 1998). Solitarious locusts tend to avoid each other whereas gregarious animals are likely to aggregate. Combined with a high reproductive capacity, this...
aggregation behavior can lead to the formation of huge locust swarms consisting of billions of individuals. Solitary and gregarious locusts not only differ in behavior, but also display remarkable morphological and physiological differences (Uvarov 1966, Pener & Yerushalmi 1998). In adults, the phase shift leads to important changes in reproductive physiology. The complete switch from one phase to the other (i.e. phase transition) usually requires several generations and is reversible. In addition to the insect developmental hormones, juvenile hormone and 20-hydroxyecdysone, brain-derived peptides, such as corazonin, adipokinetic hormone and several ‘parsins’, were already suggested to be linked to this process (Ayali et al. 1996a,b, Tawfik et al. 1999). Moreover, a phase-dependent regulation of locust neuroparsin (NPs) transcript levels was reported recently (Claeys et al. 2005). The ‘parsins’ were initially discovered as small neurosecretory proteins that are present in the pars intercerebralis–CC complex of the locust brain (Girardie et al. 1989, 1998, Lagueux et al. 1990, Hetru et al. 1991). Although NPs were first identified in locusts, genome and expressed sequence tags data have revealed the presence of various members of the NPs family in other arthropod species (Claeys et al. 2003). Furthermore, these small Cys-rich proteins display sequence similarity with the conserved N-terminal region of insulin-like growth factor binding proteins (IGFBPs), which possesses the hormone binding capacity (Claeys et al. 2003, Badisco et al. 2007). An IGFBP-like peptide was also identified in another invertebrate, namely the mollusk, *Haliotis laevigata* (Weiss et al. 2001). This molluskan IGFBP-like peptide, termed ‘perlustrin’, was indeed shown to interact in vitro with vertebrate ILPs, such as IGF and insulin. In adult female mosquito ovaries, a NPs-like factor, the ‘ovary ecdysteroidogenic hormone’ (OEH), displays ecdysteroidogenic activity (Brown et al. 1998) in a similar way to (vertebrate) insulin (Riehle & Brown 1999). Altogether, these previous studies suggest the possible existence of a functional relationship between both IRP and NPs-like proteins in insects. Therefore, based on our interest in studying the role of parsins in desert locust reproduction and phase transition, we were eager to identify the IRP of *Schistocerca gregaria*, to clone the corresponding cDNA and to analyze possible tissue-, gender- and phase-dependent differences in its transcript levels. In addition, we tried to provide experimental evidence for the hypothesis that locust NPs are capable of interacting with this endogenous locust IRP.

**Materials and methods**

**Rearing of the animals**

Gregarious desert locusts, *S. gregaria* (Forskál), were reared under crowded conditions according to the method previously described by Vanden Broeck et al. (1998). Breeding of solitarious desert locusts was performed under isolated conditions as described by Hoste et al. (2002). Temperature, photoperiod and food supply were similar for both phases.

**Purification of Scg-IRP**

**Preparation of CC extract**

A total of 1800 locust CC were micro-dissected, rinsed in a Ringer solution (1L: 8.766 g NaCl; 0.188 g CaCl₂; 0.746 g KCl; 0.407 g MgCl₂; 0.336 g NaHCO₃; 30.807 g sucrose; 1.892 g trehalose; pH 7.2) and centrifuged at 10 000 g for 30 min (Sorvall, Beckmann, Germany). The remaining pellet was extracted once again with acidified ethanol. Both supernatants were pooled and the organic solvent was evaporated ( Büchi Rotavapor, Büchi Laboratory Equipment, Flawil, Switzerland). The remaining aqueous solution was delipidated by ethylacetate and n-hexane extractions. Remnants of organic solvents were removed in the rotavapor and the resulting extract was used for HPLC analysis.

**High performance liquid chromatography**

The first separation step was performed on a Waters Delta 600 HPLC. The chromatography was carried out on a preparative Deltapack C₁₄ column (25 mm × 100 mm; particle size 15 μm, 100 Å) at ambient temperature and flow rate of 12 ml/min. The prepared sample was then transferred to the injector and immediately after injection a linear gradient, from 0% to 90% CH₃CN containing 0.1% trifluoroacetic (TFA), was initiated. The eluting peptides were detected with a variable wavelength u.v. detector, set at 214 nm (Waters 2487). Fractions of 12 ml were collected every minute, from 0 to 90 min. Aliquots (1/60 volume) of these fractions were tested in a dot blot assay, using an antibody against *Lom*-IRP (Riehle et al. 2006). The most intensely stained fraction was evaporated to dryness and subjected to the second run was carried out on a Waters semi-preparative Spherisorb C₁₃ column (10 mm × 250 mm, particle size 10 μm) at ambient temperature and at a flow rate of 2 ml/min. Immediately after injection, a linear gradient was started from 2% to 50% CH₃CN containing 0.1% TFA. The eluting peptides were detected with a variable wavelength u.v. detector, which was set at 214 nm (Waters 486 Tunable Absorbance Detector). Fractions
of 2 ml were collected every minute, from 0 to 60 min. Small aliquots (1/60 volume) were again sacrificed in a dot blot assay.

**Dot blot assay**

Aliquots of HPLC fractions were evaporated in a vacuum centrifuge and subsequently dissolved in 5 μl 10% CH₃CN/0.1% TFA. Two microliters of each sample was spotted onto a nitrocellulose membrane (Hybond-C, Amersham), which was then baked for 30 min at 120°C. Membranes were blocked with 3% skimmed milk in 50 mM Tris-buffered saline and rinsed with distilled water for 12 h, dehydrated in an ethanol series, cleared overnight in histosol/paraplast, and then transferred to Bouin Hollande’s solution, and then embedded in paraplast. Alternating sections of 4 μm were made with a LKB Historange glass microtome (LKB, Stockholm, Sweden). Sections were rehydrated and processed according to the peroxidase anti-peroxidase immunocytochemistry method (Vandesande & Dierickx 1976) using DAB as chromogenic substrate. The primary rabbit antibody, raised against the A-chain of Lom-IRP (Riehle et al. 2006), was applied in a dilution of 1/500.

**Preparation of RNA and cDNA**

Desert locust tissues were micro-dissected under a binocular microscope and immediately collected in RNAlater solution (Ambion) to prevent degradation. Until further processing, pooled tissue samples (each sample was derived from ten individuals) were stored at -20°C. Samples were added to reaction tubes containing Green beads and homogenized in the MagNA Lyser instrument (Roche). Subsequently, total RNA was extracted from the resulting homogenates utilizing the RNeasy Lipid Tissue Mini Kit (Qiagen). In combination with this extraction procedure, a DNase treatment (RNase-free DNase set, Qiagen) was performed to eliminate potential genomic DNA contamination. After spectrophotometric quantification and quality control with the Agilent 2100 Bioanalyzer (Agilent Technologies), 1 μg of the resulting total RNA was reverse transcribed (Superscript II, Invitrogen) utilizing random hexamers according to the company’s protocol. Afterwards, the resulting cDNA was diluted tenfold.

**Cloning of the Scg-IRP cDNA**

**Partial cDNA cloning by PCR**

The PCR primer sets (I and II) were initially based on the amino acid sequence of the A- and B-chains of Scg-IRP (cf. Edman degradation) and on the nucleotide sequence of Lom-IRP. These primers had the following sequences:

- → 5'-CGGCCGAGAGCTCTCCAA-3' (I)
- ← 5'GCCTTCTTGAACATGGTGTTGATGTTG-3' (I)
- → 5'GGAGGATGTTGCGCCGCAAGA-3' (II)
- ← 5'GACGAGTCGTCGCCGCAGTAG-3' (II)

Since sequence information of the resulting PCR fragments revealed that the coding sequences of Scg-IRP and Lom-IRP were nearly identical, a new set of primers spanning the open reading frame (ORF) was designed based on the Lom-IRP coding sequence:

- → 5'-ATGTGGAAGCTGTGCCTCCGACTGCTCG-3'
- ← 5'GGGCCGAGAGCTCTCCAA-3'

The PCR was performed with Pwo SuperYield DNA Polymerase (Roche), because the amplification efficiency appeared to be low when Taq polymerase was
used, probably due to the presence of GC-rich regions in the Scg-IRP cDNA. The total reaction volume was 50 μl, including 5 μl cDNA template, 5 μl 10× PCR buffer (Roche), 1 μl dNTP mix (2.5 mM for each dNTP), 10 μl GC-rich resolution solution (Roche), 5 μl each primer (10 μM) and 0.5 μl Pwo SuperYield DNA polymerase (5 U/μl, Roche). Hot-start PCR was performed in a Thermocycler (Biometra, Göttingen, Germany). After an initial incubation at 95 °C for 2 min, thermal cycling (45 cycles) consisted of a denaturation step at 95 °C, an annealing step at 70 °C and an extension step at 68 °C, for 1 min each. A final extension step was applied at 72 °C for 7 min.

Rapid amplification of cDNA Ends (RAcE)

Based on the sequence of the ORF fragment, the following primers were designed for RAcE.
- Primer for 5′-RAcE: 5′-GGTGGAGAGCTTCTCGGGCAGGCA-3′
- Primer for 3′-RAcE: 5′-GGTTCCTCCAAAGATGTGTCGGACGGG-3′.

The RAcE reactions were performed with the BD SMART RAcE cDNA amplification kit (BD Biosciences, Clontech, San Jose, CA, USA) according to the manufacturer’s protocol. However, due to amplification difficulties, Pwo SuperYield DNA polymerase had to be employed for the 3′-RAcE reaction, instead of Taq polymerase. The following temperature profile was applied: 5 cycles with a denaturation step at 94 °C for 30 s and annealing/extension at 72 °C for 3 min, 5 cycles with denaturation at 94 °C for 30 s, annealing at 70 °C for 30 s and extension at 72 °C for 3 min, and 40 cycles with denaturation at 94 °C for 30 s, annealing at 68 °C for 30 s and extension at 72 °C for 3 min.

Analysis of PCR and RAcE fragments

Amplification products were analyzed by horizontal agarose gel electrophoresis and purified using the GenElute gel extraction kit (Sigma–Aldrich). The DNA fragments were subcloned into the pCRII vector via the TOPO TA Cloning Kit (Invitrogen). The DNA sequences were determined using the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems) following the protocols outlined in the ABI PRISM BigDye Terminator Ready Reaction Cycle Sequencing Kit (Applied Biosystems).

Quantitative real-time RT-PCR

The RT-PCRs were performed in 25 μl volume, containing 5 μl diluted cDNA sample, according to the Power SYBR Green PCR Master Mix protocol (Applied Biosystems). The final concentration of the primers was 300 nM. In order to compensate for possible variations due to pipetting errors and differences in reverse transcriptase efficiency, a S. gregaria β-actin transcript was analyzed as an endogenous control. Our previous studies indicated that the levels of this mRNA remain quite constant in locust tissues, regardless of developmental or physiological conditions (Vanden Broeck et al., 1998, Janssen et al. 2001, Claes et al. 2003). Primers for the endogenous control were described previously (Claes et al. 2005). Those for the Scg-IRP target sequence were designed by means of the Primer Express software package (Applied Biosystems):

\[
\text{Scg-IRP} \rightarrow 5′-\text{CCGTGGCAACTACAACACCAT-3′} \\
\leftarrow 5′-\text{TCCCGTGCTCGACACATCT-3′}.
\]

The reactions were run in duplicate on an ABI PRISM 7000 Sequence Detection System (Applied Biosystems) applying the following thermal cycling profile: 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 60 s at 60 °C. Data were analyzed according to the comparative CT method by means of the ABI PRISM 7000 SDS software (Applied Biosystems, version 1.2.3). The specificity of the PCRs was double-checked. Analysis of the dissociation curves of Scg-IRP and β-actin amplification products revealed a single melting peak. In addition, PCR products were analyzed via agarose gel electrophoresis, showing the presence of a single band of the expected size for each transcript. Furthermore, sequencing of the PCR products ultimately confirmed the identity of the amplified DNA. For each sample, the relative amount of transcript was normalized to the endogenous control and transcript levels were calculated relative to a calibrator sample (day 4 female brains). All experiments were repeated three times with independent samples. Statistical analysis was performed by means of Statistica 7.1 (StatSoft, Tulsa, OK, USA) and consisted of the Mann–Whitney U test for comparing two independent groups. A level of \( P<0.05 \) was considered significant.

Binding Of Scg-NP4

Vector for Scg-NP4 expression

The Drosophila inducible/secreted expression system (Invitrogen) results in the biosynthesis of a recombinant gene product in Drosophila Schneider 2 (S2) cells. The vector pMT/BiP/V5-His codes for an N-terminal, secretory signal peptide, which directs the product towards the culture medium, as well as a C-terminal peptide containing a V5-epitope for antibody detection and a polyhistidine (6×His) tag for binding onto a nickel column and subsequent affinity purification of the protein of interest. Furthermore, the expression of recombinant product is controlled by a metallothionein gene promoter (pMT). Transcription of metallothionein genes is regulated in a heavy metal-dependent manner.
Hence, expression of pMT-controlled genes can be induced by adding copper ions (or other heavy metal ions) to the cell culture medium.

Based on the nucleotide sequence of the Scg-NP4 precursor (Claeys et al. 2003), a set of primers was designed for selective amplification of the cDNA fragment that codes for Scg-NP4. The upstream primer contained a BglII restriction site whereas the downstream primer included a recognition site for XhoI. Since both sites are also present in the multiple cloning site of the pMT/BiP/V5-His vector, this allowed for directional insertion of the fragment. Digestion of pMT/BiP/V5-His and the amplified fragment was performed with BglII and XhoI (Roche). After ligation (Rapid DNA Ligation Kit, Roche) and transformation in Escherichia coli DH5α cells (Invitrogen), the expression construct was verified by DNA sequencing. Subsequently, recombinant colonies were subjected to an endotoxin-free Maxiprep (Qiagen) procedure.

**Biosynthesis and purification of Scg-NP4**

The Drosophila S2 cells were cultured at 23 °C in Schneider’s Drosophila medium (Serva Electrophoresis GmbH, Heidelberg, Germany) supplemented with 5·45 mM CaCl2, 4·44 mM NaHCO3, 50 U/ml penicillin, and 50 μg/ml streptomycin (Invitrogen) and 10% heat-inactivated fetal calf serum (Invitrogen). Transfections

---

**Figure 1** Chromatograms of the HPLC runs that resulted in the purification of Scg-IRP, a desert locust insulin-related peptide. (A) In the first run, an extract of 1800 corpora cardiaca was fractionated on a preparative Waters DeltaPack C4 column with a linear gradient from 100% A (0·1% TFA) to 100% B (0·1% TFA/90% acetonitrile) within 90 min at a flow rate of 12 ml/min. (B) Fraction 31 was further fractionated in the second run on a Waters semi-preparative Spherisorb C1 column with a linear gradient from 100% A (0·1% TFA) to 100% B (0·1% TFA/50% acetonitrile) within 60 min at a flow rate of 2 ml/min. The eluted peptides were monitored with a variable wavelength u.v. detector, set at 214 nm. Fractions were collected every minute, their numbers corresponding to the elution time. Arrows indicate the range of fractions that were stained positive in a dot blot assay with anti-Lom-IRP antiserum.
were carried out in serum-free Schneider’s *Drosophila* medium with 6 μl Cellfectin (Invitrogen) and 1 μg DNA per 10^6 cells. Co-transfection was performed with a pCoHygro selection plasmid (carrying a hygromycin B resistance gene) in order to obtain a stable cell population. For peptide production, hygromycin B selected cells were grown in *Drosophila* serum-free medium (Invitrogen). After the cells had reached the exponential phase of their growth curve (2–4×10^6 cells/ml), recombinant NPs expression was induced.

**Figure 2** Immunohistochemical analysis of IRP-like immunoreactivity in desert locust (*S. gregaria*) brain. (A) Immunopositive staining is visualized in the pars intercerebralis (full arrow), nervus corporis cardiacl (broken arrow) and corpus cardiacum (dotted arrow). (B) Alternating tissue section incubated with the pre-immune serum as negative control. (C) Scheme of the desert locust brain showing the localization of IRP-like immunoreactivity. Immunoreactive fibers and terminals are present in the storage part of the CC (arrow); cell bodies are located in the PI (arrow). Abbreviations: CA, corpus allatum; CCg, corpus cardiacum glandular part; CCs, corpus cardiacum storage part; COC, circumesophageal connectives; DEUTO, deutocerebrum; FG, frontal ganglion; HC, hypocerebral ganglion; LA, lamina; LO, lobula; M, medulla; MN, median nerve; NCA I, II, nervi corporis allati; NCC I, II, III, nervi corporis cardiacl; OL, optic lobe; PI, pars intercerebralis; PL, pars lateralis; PROTO, protocerebrum; TRITO, tritocerebrum.
Figure 3 (A) Sequence of the Scg-IRP precursor cDNA (Scg-IRP T2) and the corresponding amino acid sequence. Consecutively, the coding sequences for the signal peptide (highlighted in gray, characters in bold black), the IRP copptide (characters in bold black, dotted underline), the B-chain (highlighted in black, characters in white), the C-chain (characters in bold gray) and the A-chain (highlighted in gray, characters in black) are represented. Start and stop codons, and the polyadenylation signal are underlined (single line). Predicted basic cleavage sites are denoted in bold italics. The sequence of an alternative transcript variant (Scg-IRP T1) is denoted in gray italics. The primers used for cloning the ORF are double underlined, the primer used for 5' RACE is underlined with dashes and points, and the primer used for 3' RACE is dashed underlined. (B) Pairwise sequence alignment of the newly identified Scg-IRP with that of Lom-IRP. The alignment was performed by ClustalW analysis (http://www.ebi.ac.uk/clustalw/), available from the European Bioinformatics Institute. All parameters were set at default values. Consecutively, the amino acid sequences of the signal peptide, the IRP copptide, the B-chain, the C-chain and the A-chain (representation as in A) are shown. Predicted basic cleavage sites are denoted in bold italics. Conserved residues are denoted with ‘*’ in the consensus line.
(48 h) by adding CuSO₄ to a final concentration of 500 µM. Recombinant Scg-NP4-V5-His in the culture medium (240 ml) was then concentrated by means of Centricon 80 cartridges (cut-off membrane of 5000 kDa; Amicon) to a final volume of 6 ml. A 10 ml column was filled with 3 ml His-Select Nickel Affinity Gel (Sigma–Aldrich). First, the column was washed with 6 ml H₂Odest and subsequently three times with 6 ml equilibration buffer (50 mM sodium phosphate, pH 8-0; 0.3 M sodium chloride; 10 mM imidazole). Then, the column was loaded with 6 ml concentrated Scg-NP4-V5-His solution. The gel and medium were incubated for 1 h at ambient temperature, keeping the gel suspended by attaching the column to a rotating wheel. Next, the column was washed eight times with 6 ml equilibration buffer. Finally, the recombinant protein was eluted with elution buffer (50 mM sodium phosphate, pH 8.0; 0.3 M sodium chloride; 200 mM imidazole) and 10 fractions of 1 ml were collected. Fractions were analyzed by means of SDS-PAGE. The eluted fractions containing the recombinant protein were desalted on PD-10 columns (Amersham-Pharmacia Biotech) and concentrated in a vacuum centrifuge to a final volume of 1 ml. The presence, purity and integrity of the protein were assessed by SDS-PAGE (Laemmli 1970) using a precast NuPAGE Novex 4–12% Bis-Tris gel, MES buffer and the protein electrophoresis system (all Invitrogen). A protein ladder (SeeBlue Prestained standard, Invitrogen) was run in parallel with an affinity purified protein sample. Following electrophoresis, the gel was stained in Coomassie Brilliant Blue solution (overnight) and proteins were visualized after destaining in a methanol acetic acid solution. As an additional control, purified material was partially sequenced by Edman degradation.

**Binding assay**

The Scg-NP4-V5-His (200 µl concentrated sample) was loaded onto a Ni²⁺-column (1 ml volume), which was pretreated as described above. After rinsing the column three times with 2 ml equilibration buffer, 200 CC equivalents of purified Scg-IRP, dissolved in 2 ml equilibration buffer, were added to the nickel resin already containing Scg-NP4-V5-His and both were incubated overnight at ambient temperature using a rotating wheel in order to keep the resin suspended. In parallel, the same amount of Scg-IRP was loaded onto an empty nickel column (which did not contain Scg-NP4-V5-His). Both resins were rinsed eight times with 2 ml equilibration buffer. Next, peptides were eluted from the resin with 2 ml elution buffer and this was repeated seven times. Wash and elution fractions were analyzed for the presence of Scg-NP4-V5-His and/or Scg-IRP by means of a dot blot assay, in which recombinant Scg-NP4-V5-His and Scg-IRP were detected by antibodies against the V5 epitope (Invitrogen) and the A-chain of Lom-IRP respectively.

**Figure 4** Relative quantity of the Scg-IRP mRNA in brain and fat body, represented per gender, in function of the number of days (4 or 10) after adult emergence. Dark gray graphs represent the results for gregarious desert locusts and light gray graphs for solitarious locusts. Results were obtained by analyzing three independent groups of ten individuals per condition and are represented as means ± s.d. Statistical analysis was performed by means of ‘Statistica 7.1’ and consisted of the Mann–Whitney U test for comparing two independent groups. Significant differences (P<0.05) are represented by connecting lines between the corresponding graphs. (A) Male brain, (B) male fat body, (C) female brain and (D) female fat body.
Results

HPLC purification of Scg-IRP

Fractions of the first HPLC run, which eluted between 29 and 36 min, resulted in positive staining (Fig. 1A) during the dot blot assay. Next, fraction 31 that was the most positively stained fraction was further analyzed by reversed-phase HPLC. From the subsequent dot blot assay, one positively immunostained fraction was obtained, which eluted at 51 minutes corresponding to 42.5% CH$_3$CN (Fig. 1B). Part of this fraction (1/60) was subsequently analyzed by MALDI-TOF MS, revealing an average mass of 5736 kDa, which was in accordance with the expected mass of a potential ILP. Since the IRP was not reduced, each Edman degradation cycle revealed two amino acids, one of the A-chain and one at the corresponding position in the B-chain. When all cycles were run, the following amino acid sequences were revealed:

**A-chain:** GVFDCECRKSCSIXLQITYCG

**B-chain:** SGAPQPVARYCGEKLSNALKIVCRGNYNTMF.

Edman degradation revealed the complete amino acid sequence of both the A- and B-chains of Scg-IRP, with the exception of the amino acid at position 15 of the A-chain (X). Comparison of the amino acid sequence with the only other identified IRP from Orthopteroidea, namely *Lom*-IRP, revealed that one amino acid was different: at position 21, the B-chain of Scg-IRP contains an isoleucine (I) instead of a leucine (L) in *Lom*-IRP. The A-chain appears to be identical in both locusts. Hence, the antibody against the complete A-chain of *Lom*-IRP can also be used for specific detection of Scg-IRP.

Immunolocalization in the brain

Alternating desert locust brain sections were incubated with rabbit polyclonal antibody raised against the A-chain of *Lom*-IRP and the rabbit pre-immune serum respectively. Intensely stained median neurosecretory cells were observed within the pars intercerebralis. Furthermore, immunopositive staining was also observed in the neurohemal storage part of the CC and in the connecting nerves between the pars intercerebralis and the CC, namely the nervi corporis cardiaci I (NCC-I; Fig. 2).

Cloning of the Scg-IRP cDNA

Sequencing of the obtained PCR fragments revealed the amino acid at position 15 as glutamic acid (E), identical to the A-chain of *Lom*-IRP (Fig. 3B). Next, primers for the RAcE reactions were derived from the partial cDNA sequence obtained by PCR. The RAcE-PCR strategy resulted in the identification of two transcript sequences that differed only in their 5'-untranslated region. These cDNA sequences and the deduced amino acid sequence of the encoded pre-pro-Scg-IRP precursor are represented in Fig. 3(A). When entering the Scg-IRP precursor amino acid sequence into the SignalP 3.0 algorithm (Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark) (Bendtsen et al. 2004), the most probable cleavage site for the signal peptide is between position 22 and 23, as was also observed for *Lom*-IRP (Clynen et al. 2003).

Distribution of the Scg-IRP transcripts

In both solitarious and gregarious locusts, the presence of Scg-IRP mRNA was detectable in virtually all tissues that were tested, i.e. nervous system, fat body, gonads, male accessory glands, flight muscles and salivary glands. The Scg-IRP transcript levels were most pronounced in...
nervous tissue and in fat body. Figure 4 shows the relative quantities of Scg-IRP mRNA that were determined in brain and fat body of adult locusts as a function of gender and of age (4 and 10 days after adult emergence). Significant temporal fluctuation was noticed in the fat body of both adult males and females, as well as in the brain of adult gregarious females. In the fat body of gregarious males, the Scg-IRP transcripts reached a significantly higher level (ca. threefold) at day 10, when compared with day 4. This is contrary to the situation in Aaeg peptide A1-3/B1-2 (GenPept: BAA03019, BAA03021, BAA03023, BAA03020, BAA03022); bombyxin A1-G1, product of a pair of duplicate genes, only differing in their C-peptides (GenPept: AAQ89694, AAQ89698); and of age (4 and 10 days after adult emergence).

Significant temporal fluctuation was noticed in the fat body of solitary males, where relative Scg-IRP mRNA quantities were lower (ca. twofold) at day 10 (Fig. 4B). As a result, significant differences in the significantly higher level (ca. threefold) at day 10, when compared with day 4. In solitary female fat bodies, a much smaller increase (ca. threefold) was observed (Fig. 4D).

**Binding to Scg-NP4**

Culture medium containing concentrated Scg-NP4-V5-His was subjected to a one-step nickel ion chelate affinity chromatography. All elution fractions were tested for the presence of Scg-NP4-V5-His protein by means of SDS-PAGE. Fractions containing the eluted protein were combined, desalted and concentrated. In the concentrated sample, one single protein band of ~12 kDa was observed by SDS-PAGE (Fig. 5A). This estimated molecular mass is in line with the calculated mass of Scg-NP4-V5-His (12 404-87 kDa). In addition, amino acid sequencing indeed identified the purified material as recombinant Scg-NP4-V5-His.

Figure 5B shows the results of the dot blot analysis in which wash and elution samples of the binding assay were tested for the presence of Scg-NP4-V5-His and/or Scg-IRP. No (or only very weak) immunoreactivity was detected in any of the wash fractions which were incubated with the anti-V5 antibody, indicating that no detectable amount of Scg-NP4-V5-His was recovered from the resin during the washing steps (as expected, since the capacity of the column for the poly-His containing protein is quite high and possible unbound recombinant NPs had already been removed during rinsing steps prior to the incubation of the column with Scg-IRP). In contrast, wash fractions 1–5 were stained with decreasing intensity after treatment with the antibody raised against the A-chain of Lom-IRP. Wash fractions 6–8 did not show any immunopositive staining upon incubation with anti-Lom-IRP (A-chain) antibody, indicating that the unbound Scg-IRP was eliminated after the subsequent wash steps. Immunopositive staining was clearly observed in the first two elution fractions after incubation with anti-V5 monoclonal antibody, as well as after treatment with the anti-Lom-IRP antiserum. The same experiment was repeated twice with reproducible results. In addition, both antibodies did not show any cross-reactivity for each others targets and Scg-IRP was not retained on the column in control experiments in which no Scg-NP4-V5-His was added.

**Discussion**

In the present study, we describe the characterization of a novel member of the insulin superfamily in *S. gregaria*. The Scg-IRP was isolated from CC extracts by means of an HPLC-based purification strategy (Fig. 1). The resulting HPLC fractions were screened for the presence of an insulin-like substance by means of an antibody raised against the A-chain of Lom-IRP. Immunocytochemical data obtained with the same antibody (Fig. 2) indicated that the localization of the immunoreactive material in locust brain is in accordance with that of several other insect IRPs. Localization studies in various insects (*Iwami et al. 1996a, Cao & Brown 2001, Riehle et al. 2006, Van de Velde et al. 2007*) demonstrated that many IRPs are produced in neurosecretory cells of the pars pars.

Figure 6 Multiple sequence alignment of the (A) A-chain and (B) B-chain of several insect insulin-related peptides. The alignment was performed by ClustalW analysis (http://www.ebi.ac.uk/clustalw/), available from the European Bioinformatics Institute. All parameters were set at default values. Highly conserved residues are highlighted in black. Residues that are similar to the column’s consensus are shaded. The consensus line includes the symbols: ‘*’ representing a particular position at which an amino acid is completely conserved, ‘.’ representing a position for which a consensus is defined, but at which not all amino acids are equal to the consensus. Abbreviations: Scg-IRP, *Schistocerca gregaria* insulin-related peptide; Lom-IRP, *Locusta migratoria* insulin-related peptide (GenPept: P15131); dlp1-7, *Drosophila melanogaster* insulin-like peptide 1-7 (GenPept: Q9VT50, Q9VT51, Q9VT52, Q9VT53, Q7KUD5, Q9W4Z4, Q9W4Q9); AaeGlp1-8, *Aedes aegypti* insulin-like peptide 1-8 (GenPept: ABI64116, ABI64118, ABI64117, ABI64119, ABI64124, ABI64122, ABI64123, ABI64120); AgamLP1-7, *Anopheles gambiae* insulin-like peptides 1 and 7 that are the product of a pair of duplicate genes, only differing in their C-peptides (GenPept: AAQ89693, AAQ89695, AAQ89697); Sac-bombA1-3/B1-2, *Samia cynthia* bombyxin-related peptide A1-3/B1-2 (GenPept: BAA03019, BAA03021, BAA03023, BAA03020, BAA03022); bombyxin A1-G1, *Bombyx mori* insulin-like peptides A1-G1 (GenPept: Q17192, P26733, P15410, P26736, P21808, P91896, O16271); SbIRP1-2, *Spodoptera littoralis* insulin-like peptide 1-2 (29), Agc-bombA-B, *Agrion convolvuli* bombyxin-related peptide A–B (GenPept: O09209, O09210).
intercerebralis, an important production site of neurohormones in insects. From there on, neurosecretory products are transported via the nervi corporis cardiaci (NCCI) to the storage part of the CC, where they reside until their release into the circulating hemolymph.

In combination with this immunological screening method, our research strategy has resulted in the purification of Scg-IRP, as well as in the determination of its amino acid sequence. The Scg-IRP possesses the typical structure of a metazoan insulin-like hormone (with the exception of vertebrate IGFs or IGFs and several Caenorhabditis elegans ILPs) that consists of A- and B-chains that are interconnected by Cys-bridges. Subsequently, the corresponding cDNA was cloned by means of a PCR- and RACE-based strategy and sequenced (Fig. 3). The encoded hormone precursor exhibits a very similar organization as observed for most other members of the insulin superfamily. The pre-pro-Scg-IRP polypeptide indeed contains a signal peptide, which directs the hormone to the secretory pathway, as well as a pro-insulin portion with a B-, C- and A-chain. Upon removal of the signal peptide and the C-chain, the A- and B-chains are covalently coupled by means of disulfide bridges and constitute the mature peptide. However, this desert locust IRP precursor still contains another peptide sequence that is situated between the signal sequence and the B-chain. This decapeptide has previously been termed ‘IRP-copeptide’. Its existence was first demonstrated in the migratory locust, Locusta migratoria, where it may be involved in the regulation of carbohydrate metabolism (Clynen et al. 2003). The IRP co-opeptide is identical in L. migratoria and S. gregaria. In contrast to the well-conserved sequences of Scg-IRP and IRP-copeptide, the C-chain is clearly less conserved between locust species (Fig. 3B). This observation is not unexpected, since, for all members of the insulin superfamily, the B- and A-chains are noticeably more conserved than the C-chain (Claeys et al. 2002). Multiple alignment of the newly identified Scg-IRP sequence with other known insect IRPs (Fig. 6) show that the cystein pattern in both the A- and B-chains is extremely well conserved. In addition, a tyrosine (Y) and leucine (L) residue in the A- and B-chains respectively, appear to occur in all known insect IRPs. Therefore, the observed results concerning the primary structure of the peptide and its cDNA indicate that pre-pro-Scg-IRP shares several general characteristics with other members of the insulin superfamily, while it also possesses some properties that may be more specific for locusts. In contrast to genome-derived nucleotide sequence data obtained from lepidopteran, dipteran and hymenopteran species, only one single IRP has so far been identified in each of the orthopteran species, L. migratoria and S. gregaria. For both Lom-IRP and Scg-IRP, two transcript (cDNA) variants have now been identified, but these only differ in their 5' untranslated region (Kromer-Metzger & Lagueux 1994). Since the genomes of these species have not (yet) been sequenced, it is still unclear whether other IRP genes remain to be discovered in these locusts.

Quantitative RT-PCR analysis revealed that the Scg-IRP transcripts occur in a wide variety of tissues, including nervous tissue and fat body, where expression levels were most pronounced. Our observation that IRP transcripts are not restricted to the brain is in line with several studies in other insects (Wu & Brown 2006). Although bombyxins in B. mori are mainly expressed in neurosecretory cells of the brain, bombyxin A and B transcripts have also been detected in other ganglia, as well as in peripheral tissues (Iwami et al. 1996). In D. melanogaster, the mRNAs for most dils occur in both nervous and peripheral tissues (Brogiolo et al. 2001, Ikeya et al. 2002, Rulifson et al. 2002, Broughton et al. 2005). Moreover, in the migratory locust, L. migratoria, Lom-IRP transcripts were also found in a wide variety of tissues (Kromer-Metzger & Lagueux 1994). The presence of IRP transcripts in various peripheral tissues may be an indication for a possible role of the encoded peptide(s) as a paracrine and/or endocrine signal substance, a situation that may resemble that of IGFs and many other growth factors. Our data also show the occurrence of a temporal fluctuation in relative Scg-IRP transcript quantities, especially in the fat body (Fig. 4). In adult, male and female, gregarious fat bodies, the Scg-IRP mRNA levels are significantly higher at day 10 after adult emergence, when compared with day 4 and with solitarious locusts. The fat body plays a key role in insect reproductive physiology. It is the major production site of vitellogenins (yolk protein precursors) that are released into the hemolymph and internalized in the growing oocytes. The entire process is believed to be under a strict hormonal control. Desert locust phases display remarkable differences in their reproductive physiology. Gregarious desert locusts show accelerated sexual maturation compared to their solitarious congeners (Uvarov 1966). In addition, the egg pods of gregarious females contain more eggs than those of solitarious females (90–160 eggs per gregarious egg pod versus a maximum of 80 eggs per solitarious egg pod; Uvarov 1966). On the other hand, solitarious females can produce more egg pods during their life time, since they tend to live longer (Claeys et al. 2005). In our gregarious locust lab cultures (the exact timing of the reproductive cycle may vary depending on culture conditions and food supply), the initial vitellogenic cycle in females started around day 6 after adult emergence, and resulted in well-developed ovaries around day 10. The first egg batches were deposited around day 14. In gregarious males, initial copulation behavior was also observed around day 10. Therefore, the observed increase of Scg-IRP transcript levels in fat body occurred in parallel with the sexual maturation of gregarious animals, suggesting a possible link between the hormonally and phase-dependently controlled reproductive cycle and Scg-IRP expression. Various
studies in other insects already indicated a possible role for IRP/ILP signaling in the control of insect reproduction. Fullbright et al. (1997) observed putative insulin receptors in ovarian cells of three different lepidopteran species, suggesting a function in ovarian development that may be similar to that of IGFs and relaxins in vertebrates. In *Aedes aegypti*, a mosquito insulin receptor (MIR) was also demonstrated in follicular cells, the primary source of ecdysteroids in adult females. In addition, MIR appeared to be expressed in function of the reproductive cycle (Riehle & Brown 2002) and treatment with bovine insulin caused *Ae. aegypti* follicular cells to produce ecdysteroids (Riehle & Brown 1999) that further stimulate vitellogenesis in this species. Female *D. melanogaster* require an intact insulin/IRP signaling pathway in the ovaries to regulate egg production in response to dietary changes (Drummond-Barbosa & Spradling 2001). Moreover, mutations of the insulin receptor, a receptor tyrosine kinase, produced fruit fly phenotypes displaying impaired ovarian ecdyseroidogenesis (Tu et al. 2002), reduced folliculogenesis and juvenile hormone deficiency (Tatar 2004). Although the situation in locusts is less clear, various studies also suggested the existence of a (temporal) relationship between ovarian ecdyseroidogenesis and vitellogenesis (Girardie & Girardie 1996, 1998, Tawfik et al. 1997, Girardie et al. 1998, Tawfik et al. 2002).

In addition to insulin (this study), locusts possess different NPs, which are structurally related to the OEH of mosquitoes and to vertebrate IGFBP (Badisco et al. 2007). In contrast to OEH, NP (i.e. *Scg*-NP1) was initially discovered as a factor that prevented oocyte growth and thus delayed the gonotrophic cycle (Girardie et al. 1998). At present, four NPs (*Scg*-NP1-4) have been identified from *S. gregaria* (Janssen et al. 2001, Claey s et al. 2003), and in the fat body of adult gregarious locusts the transcripts coding for *Scg*-NP3 and for *Scg*-NP4 are strongly upregulated during sexual maturation, similarly as *Scg*-IRP. In the present study, we demonstrate the *in vitro* binding of recombinant *Scg*-NP4 with the purified, endogenous *Scg*-IRP (Fig. 5). This result represents the first experimental evidence that insect NPs possess the capacity to interact with insulin-related peptides, as IGFBPs do in vertebrates and perlustrin in the mollusk, *H. laevisgata* (Weiss et al. 2001). If such interactions would also occur *in vivo*, NPs may exert regulatory effects situated upstream of IRP receptor signaling. These likely include effects that modulate molecular interactions, turnover, transport and tissue targeting of the corresponding insulin-like hormone(s) (Simonet et al. 2004). However, as exemplified by IGFBPs, this does not necessarily exclude the possible existence of additional modes of action (Mohan & Baylink 2002). In the future, it will be of interest to further examine the *in vivo* role of *Scg*-IRP and to analyze whether *Scg*-NPs may exert an inhibiting or stimulating effect on its activities. Furthermore, more insight in the mechanisms controlling reproduction and phase transition could offer novel opportunities to improve prevention and control of locust swarms.

Acknowledgements

The authors especially thank Mark Brown (Athens, GA, USA) for his kind gift of antiserum and for his suggestions, R Jonckers for taking care of the insect cultures and J Puttemans for providing assistance with the figures. The authors gratefully acknowledge the K U Leuven Research Foundation (GOA 2005/06), the Belgian Interuniversity Attraction Poles program (IUAP/PAI P6/14, Belgian Science Policy), the IWT and the FWO-Vlaanderen for financial support. EMBL accession numbers: AM889088 (*Scg*-IRP T1), AM889089 (*Scg*-IRP T2). The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work. The authors have nothing to declare.

References


www.endocrinology-journals.org

Collip JB 1923 The demonstration of an insulin-like substance in the tissues of the clam (Mya arenaria). **Journal of Biological Chemistry** 55 R39.

Drummond-Barbosa D & Spradling AC 2001 Stem cells and their progeny respond to nutritional changes during **Drosophila** oogenesis. **Developmental Biology** 231 265–278.

Fullbright G, Lacy ER & Bullesbach EE 1997 The prothoraciotropic hormone bombyxin has specific receptors on insect ovarian cells. **European Journal of Biochemistry** 245 774–780.


Tu MP, Yin CM & Tatar M 2002 Impaired ovarian ecdysone synthesis of **Drosophila melanogaster** insulin receptor mutants. **Aging Cell** 1 158–160.


Vandesande F & Dierickx 1976 Immunocytochemical demonstration of separate vasotocinergic and mesotocinergic neurons in amphibian hypothalamic magnocellular neurosecretory system. **Cell and Tissue Research** 175 289–296.


Weiss IM, Gohring W, Fritz M & Mann K 2001 Periluin, a *Holotis laevigata* (abalone) nacre protein, is homologous to the insulin-like growth factor binding protein N-terminal module of vertebrates. **Biochemical and Biophysical Research Communications** 285 244–249.
