A Crayfish Insulin-like-binding Protein

ANOTHER PIECE IN THE ANDROGENIC GLAND INSULIN-LIKE HORMONE PUZZLE IS REVEALED*§

Received for publication, May 16, 2013, and in revised form, June 10, 2013 Published, JBC Papers in Press, June 17, 2013, DOI 10.1074/jbc.M113.484279

Ohad Rosen†§, Simy Weil†§, Rivka Manor†§, Ziv Roth*, Isam Khalaila*, and Amir Sagi†§

From the †Department of Life Sciences, §National Institute for Biotechnology in the Negev, and *Avram and Stella Goldstein-Goren Department of Biotechnology Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

Background: In Crustacea, a male-specific insulin-like androgenic hormone governs sexual differentiation. Results: A crayfish insulin-like growth factor-binding protein (Cq-IGFBP) was expressed and found to bind the androgenic hormone. Conclusion: Cq-IGFBP is part of the scheme modulating the crustacean androgenic hormone. Significance: Cq-IGFBP is the first IGFBP found that interacts with an insulin-like gender-specific ligand linking early evolution of insulin pathways with gender.

Across the animal kingdom, the involvement of insulin-like peptide (ILP) signaling in sex-related differentiation processes is attracting increasing attention. Recently, a gender-specific ILP was identified as the androgenic sex hormone in Crustacea. However, moieties modulating the actions of this androgenic insulin-like growth factor were yet to be revealed. Through molecular screening of an androgenic gland (AG) cDNA library prepared from the crayfish Cherax quadricarinatus, we have identified a novel insulin-like growth factor-binding protein (IGFBP) termed Cq-IGFBP. Based on bioinformatics analyses, the deduced Cq-IGFBP was shown to share high sequence homology with IGFBP family members from both invertebrates and vertebrates. The protein also includes a sequence determinant crucial for ligand binding, which according to three-dimensional modeling is assigned to the exposed outer surface of the protein. Recombinant Cq-IGFBP (rCq-IGFBP) protein was produced and, using a “pulldown” methodology, was shown to specifically interact with the insulin-like AG hormone of the crayfish (Cq-IAG). Particularly, using both mass spectral analysis and an immunological tool, rCq-IGFBP was shown to bind the Cq-IAG prohormone. Furthermore, a peptide corresponding to residues 23–38 of the Cq-IAG A-chain was found sufficient for in vitro recognition by rCq-IGFBP. Cq-IGFBP is the first IGFBP family member shown to specifically interact with a gender-specific ILP. Unlike their ILP ligands, IGFBPs are highly conserved across evolution, from ancient arthropods, like crustaceans, to humans. Such conservation places ILP signaling at the center of sex-related phenomena in early animal development.

The insulin-like peptide (ILP)² superfamily, found in both invertebrates and vertebrates, comprises some of the best studied peptides to date, such as insulin, relaxin, and bombyxin (1). Although each of these hormones may be involved in its own well defined pathway and thus in a distinctive physiological response, they all share fundamental structural traits that allow their classification as insulin-like peptides. Indeed, all ILP family members are thought to have evolved from an ancestral insulin-like gene (1). Generally, an ILP is encoded as a preprohormone corresponding to a single-chained polypeptide containing a signal peptide followed by a B-chain, a connecting peptide (C-peptide), and an A-chain. Insulin-like growth factors (IGFs) are known to contain additional D- and E-domains at their C terminus (1). Usually, a proteolytic cleavage event in which the C-peptide is removed occurs before secretion (2), except in the cases of IGFs, which remain in a functional intact chain form. In all cases the tertiary structure of the mature and functional hormone is acquired via the formation of a pair of disulfide bonds linking the A and B chains.

As proteinaceous hormones that exert their biological effects via dedicated membrane-spanning receptors, ILPs are released either to the local environment (in an autocrine or paracrine manner) or to distal target cells and tissues (i.e. at the endocrine level) via secretion into the circulatory system. Although insulin and ILPs are generally associated with the systemic regulation of metabolism, IGFs are specifically involved in potentiating growth and are considered the safe-keepers of proper cell proliferation and survival processes, such as during cell differentiation, proliferation, and migration of cells and in the case of apoptosis modulation (3). Imbalanced IGF signaling can result in developmental abnormalities, such as exaggerated or reduced growth and also cancer (4–6). Thus, tight regulation of

---

* This work was supported by Research Grant QB-9308-06 from BARD of the United States-Israel Binational Agricultural Research and Development Fund.

† This article contains supplemental Figs. 1 and 2.

± To whom correspondence should be addressed: P. O. Box 653, Beer-Sheva 84105, Israel. Tel.: 972(8)6461364; Fax: 972(8)6479062; E-mail: sagia@bgu.ac.il.

² The abbreviations used are: ILP, insulin-like peptide; IGF, insulin-like growth factor; IGFBP, IGF-binding protein; Cq-IGFBP, C. quadricarinatus IGF binding protein; rCq-IGFBP, recombinant Cq-IGFBP; brCq-IGFBP, biotinylated rCq-IGFBP; Cq-IAG, C. quadricarinatus insulin-like AG hormone; AG, androgenic gland; RACE, rapid amplification of cDNA ends; contig, group of overlapping clones; IB, insulin-like-binding; KI, kazal-type serine protease inhibitor; rIBP, recombinant Cq-IGFBP; Bis-Tris, 2-(bis[2-hydroxyethyl]amino)-2-(hydroxymethyl)-propane-1,3-diol.
IGFBP Binds an Insulin-like Androgenic Hormone in Decapods

these highly potent hormones so as to render them freely available for receptor binding is essential. In contrast to insulin, which is known to freely circulate the blood system, IGFs move in the blood stream as part of complexes with specific high affinity IGF-binding proteins (IGFBP) (7). In humans, for example, the vast majority of IGF-I is transported in association with IGFBPs. These binding proteins subsequently regulate IGF-I signaling by acting as competitors to the designated receptors (8). Although all IGFBPs were found to be inhibitors of growth signaling via a sequestering of IGFs, some were also found to serve as positive regulators of IGF actions (9). IGFBPs bind IGFs with affinities that are equal to or greater than those of the various receptors (8) to dictate the availability of these hormones and, accordingly, to inhibit their abilities to potentiate growth. Mechanistically, IGF-based signaling is enabled due to a reduction of this affinity, leading to IGF release at the target site, either a result of IGFBP binding to the extracellular matrix or to an IGFBP-receptor or due to IGFBP-specific proteolysis (7).

In addition to serving as regulators of metabolism and growth, insulin-like peptides have in recent years been assigned possible involvement in various aspects of sexual differentiation in a diverse set of animal species. Briefly, ILP signaling was shown to correlate with various sexual-related manifestations ranging from primary gonad development to secondary phenotypes (e.g. dimorphic growth, weaponry, attractiveness, etc.) (10-13). With respect to ILP sex dependence, a gender-specific insulin-like hormone has been recently recorded. In the red claw crayfish, Cherax quadricarinatus, a male-unique gland termed the androgenic gland (AG) (14). IAG was shown to specifically bind the crayfish insulin-like AG hormone (Cq-IAG) (14). Since this pioneering study numerous IAG-encoding transcripts have been identified in other decapod crustaceans (e.g crab, crayfish, and shrimp) (15). IAG was shown to be the major sex hormone, inducing both masculine differentiation and maintenance (16), and was found to be crucial for the development and sustainability of the male gonad (17) as well as for secondary sex characteristics, such as male-superior growth, aggression, courtship, and others (18-20). With respect to sexual dimorphism, implantation of the AG into females resulted in higher growth rates (similar to those of males) as compared with control females (20). This AG-related effect was later specifically linked to the crustacean IAG via an RNAi-based study, with such treatment resulting in a significant reduction of growth rates and whole body weights (21).

At present the signaling schemes employed by IAG as an insulin-like growth-promoting factor remain unknown. Indeed, only a limited number of the conserved components from ILP growth pathways have been reported in crustaceans. These include two IGFBP homologs, although neither is linked to sexuality or growth regulation of the IGF-like AG hormone, instead being found to be associated with immunological responses (22, 23). In the current study we report the identification of the first IGFBP homolog to be fully sequenced in the crayfish C. quadricarinatus, namely the organism in which the first decapod IAG transcript was discovered (14). The newly identified C. quadricarinatus IGFBP (Cq-IGFBP) is predicted to contain three key domains, namely an insulin-like binding domain, a kazal-type serine protease inhibitor domain, and an immunoglobulin C-2-type domain and, based on the organization of these domains, was found to be one of many homologs found in invertebrate and vertebrate species. A recombinant version of the Cq-IGFBP protein was generated and purified and shown to specifically bind the crayfish insulin-like AG hormone.

EXPERIMENTAL PROCEDURES

Tissue Dissection and Extraction—Male crayfish were anesthetized in ice-cold water for about 10 min. Subsequently, the AG, green gland, hepatopancreas, muscle, sperm duct, and testes were sampled. After dissection, RNA was extracted using the EZ-RNA Total RNA Isolation kit (Biological Industries, Beit Ha’emek, Israel). AG homogenates were also produced. Approximately 1000 glands were dissected from eyestalk-ablated crayfish (24, 25). The glands were manually homogenized in Tris-buffered saline (TBS) (150 mM NaCl, 10 mM Tris-HCl, pH 7.5) supplemented with 2 mM 4-(2-aminoethyl)benzenesulfonyl fluoride-hydrochloride protease inhibitor (AppliChem), yielding a total protein concentration of 10 μg/gland.

RT-PCR and Rapid Amplification of cDNA Ends—cDNA was generated by a reverse transcriptase-catalyzed reaction containing 1 μg of total RNA using Expand reverse transcriptase (Roche Applied Science) according to the manufacturer’s instructions. The cDNA was then amplified by PCR. Amplification of Cq-IGFBP (no. KC952011) was performed using the forward primer (5'-AGAATGTGATAGGTCGAAATGTG-3') and the reverse primer (5'-TGGTCTTATCTTATCTTCTGGCTTG-3'). PCR products were separated by electrophoresis through a 1.3% agarose gel. Fragments were excised, purified (HiYield Gel/PCR DNA Fragments Extraction Kit, RBC Biosciences, Taiwan), and cloned (pGEM-Teasy vector, Promega). Selected clones were isolated, and plasmid DNA was purified (HiYield Plasmid Mini Kit, RBC Biosciences) and sequenced. The sequences of the 5’ and 3’ ends of Cq-IGFBP were obtained by 5’ and 3’ rapid amplification of cDNA ends (RACE) using the Clontech SMARTer RACE kit and following the manufacturer’s protocol. PCR amplification of the 5’ region was achieved using the gene-specific reverse primer Cq-IGFBP RACE-r (5’-GCATAAGACACAGTCTCGCCATGA-3’) and the Universal Primers Mix provided with the kit. PCR amplification of the 3’ region was performed with the Universal Primers Mix as a reverse primer and the gene-specific forward primer Cq-IGFBP RACE-f (5’-TAGTGCTGATGCTATGTGGATGCTGC-3’). The PCR products were cloned and sequenced as described above.

Expression and Purification of Recombinant Cq-IGFBP—Escherichia coli BL-21 strain cells were transformed with the pET-28a expression vector encoding an N-terminal polyhistidine-tagged version of Cq-IGFBP lacking its signal peptide (amino acids 23–248). The cells were grown overnight in liquid LB medium supplemented with 50 μg/ml kanamycin at 37 °C under constant shaking (200 rpm). The cell culture was then diluted 1:100 into fresh LB with kanamycin and grown under the same conditions until reaching an A600 of 0.6. An overnight incubation with 100 μM isopropyl-β-D-thiogalactopyranoside at 30 °C ensued. The cells were pelleted by centrifugation at
IGFBP Binds an Insulin-like Androgenic Hormone in Decapods

12,000 \times g for 5 min and resuspended in 10 ml of lysis buffer (50 mM Tris-HCl, pH 8, 2 mM EDTA, 100 \mu g/ml lysozyme). The suspension was frozen (−80 °C) and thawed (37 °C) 3 times followed by DNase treatment (20 \mu g/ml) in the presence of 20 mM MgCl₂, for 30 min at 37 °C. Inclusion bodies were separated from the cytosolic protein pool by centrifugation at 20,000 \times g for 30 min, and the inclusion bodies-containing pellet was solubilized with 4 ml of solubilization buffer (6 M urea, 200 mM NaCl, 100 mM Tris-HCl, pH 8.3) supplemented with 10 \mu M β-mercaptoethanol and incubated overnight at 4 °C under constant rotation. Thereafter, the sample was spun at 25,000 \times g for 30 min to remove any insoluble material, and the His-tagged protein was purified using HisPurTM Ni-NTA Resin (Thermo Fisher Scientific) according to the manufacturer’s protocol.

**Binding Assay**—Biotinylation of recombinant Cq-IGFBP (rCq-IGFBP) was performed using an EZ-Link HPDP-biotin kit (Thermo Fisher Scientific), according to the manufacturer’s protocol. Briefly, 2 mg of rCq-IGFBP were dissolved in phosphate-buffered saline (PBS) (150 mM NaCl, 10 mM Na₂HPO₄, pH 7.4) and incubated with 400 \mu M HPDP-biotin for 2 h at room temperature. Unbound biotin was removed using Vivaspin concentrating devices (Sartorius) with PBS. To examine whether Cq-IGFBP binds insulin-like peptides, three different avidin-agarose precipitations were performed. The biotinylated rCq-IGFBP (25 \mu g) was incubated with a total AG protein homogenate (125 \mu g), human recombinant insulin (10 \mu g, Biological Industries), or human recombinant IGF-I (10 \mu g, PEP-ROTECH) for 2 h at room temperature in TBS (supplemented with 2 mM 4-(2-aminoethyl)benzenesulfonyl fluoride) under constant agitation.Thirty microliters of agarose-bound avidin (high capacity neutravidin agarose resin; Thermo Fisher Scientific) was added to each of the above binding assays for an additional hour of incubation at room temperature. The reactions were then centrifuged at 8000 \times g for 2 min, unbound supernatants were collected, and the pellets were washed 3 times with TBS. All fractions (unbound and avidin-bound proteins) were separated on SDS-PAGE.

**SDS-PAGE and Immunoblotting**—Proteins were separated on 4–12% NuPAGE Bis-Tris mini gels (Invitrogen) using MES running buffer according to the manufacturer’s instructions. Bands were visualized by either Coomassie Brilliant Blue or silver staining (26). For immunoblot analyses, separated protein bands were transferred to a nitrocellulose membrane. After blocking with 3% skim milk in TBS containing 0.1% Tween 20 (TBS-T), the membrane was incubated with the relevant primary antibodies, namely either His-probe (1:500, Santa Cruz Biotechnology) or α-rec-Cq-IAG (1:1000) antibodies for the blotting of rCq-IGFBP or Cq-IAG, respectively. After washing with TBS-T, the membrane was incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (1:15,000, Santa Cruz Biotechnology). Antibody binding was visualized using an EZ-ECL chemiluminescence detection kit (Biological Industries). Biotinylated rCq-IGFBP was directly blotted with HRP-conjugated streptavidin (1:1000, Jackson ImmunoResearch Laboratories) with binding detected as described above.

**Ligand Dot Blot**—The membrane was initially dot-blotted with 2.5 \mu g of rCq-IGFBP, 3 \mu g of recombinant gastrinostilin protein 35 (rGAP35), and 30 ng of recombinant peptide A, corresponding to residues 23–38 of the Cq-IAG A-chain (synthesized by Adar Biotech). Upon blocking with 3% skim milk in TBS-T, the membrane was incubated with ligand solution (100 \mu g of peptide A in TBS-T) overnight at 4 °C. The membrane was washed with TBS-T and incubated with a rabbit anti peptide A IgG (1:500, synthesized by Adar Biotech) primary antibody. Antibody binding was visualized using an EZ-ECL chemiluminescence detection kit. For staining of the spotted proteins, rCq-IGFBP was blotted with HRP-conjugated anti-His antibodies (1:6,000, Abcam), whereas rGAP35 was visualized using Ponceau stain (Sigma).

**Mass Spectrometry**—Protein bands were excised, reduced, alkylated, and trypsinized according to Roth et al. (27). The tryptic peptides were purified on a C-18 column and dissolved in 0.1% formic acid. Nano-liquid chromatography and mass spectrometry (MS) analysis were performed as described previously (28) using a 75-μm internal diameter fused silica column packed with C-18 resin (New Objective) connected to an Eksigent nano-LC system (Eksigent). Mass spectra were acquired using an LTQ-ORBITRAP XL (Thermo Fisher Scientific). Full MS and MS/MS fragmentation were performed in the data-dependent mode. After full MS acquisition, the multiply-charged six most abundant masses where chosen for collision-induced dissociation MS/MS fragmentation in the LTQ spectrometer performed at 35° collision energy and 30 ms activation time. Protein identification and validation were performed using Sequest and Mascot algorithms operated under Proteome Discoverer 2.0 software (Thermo Fisher Scientific) using an unpublished database containing the Cq-IAG and Cq-IGFBP sequences, allowing for 0.5 Da and 10 ppm tolerance.

**Bioinformatics**—The full-length Cq-IGFBP cDNA sequence was computationally translated using the ExpASy Proteomics Server. The deduced amino acid sequence was subsequently analyzed for the presence of conserved domains using the SMART program. Evolutionary analysis was conducted using MEGA, Version 5.0 (29). Briefly, evolutionary history was inferred using the Neighbor-Joining method (30). The bootstrap consensus tree inferred from 5000 replicates (31) was reproduced in the evolutionary history of the selected mature (i.e. signal peptide-lacking) IGFBPs among the taxa analyzed (31) (Table 1). Branches corresponding to partitions reproduced in <50% of the bootstrap replicates were collapsed. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test are shown (31). Evolutionary distances were computed using the Poisson correction method. All positions containing gaps and missing data were eliminated, resulting in 139 positions being included in the final dataset.

Three-dimensional modeling of Cq-IGFBP was performed using the ESyPred3D web server (32) and processing by Swiss-PdbViewer software (33). The reference structures, which were chosen based on SWISS-MODEL, were human HtrA1 (UniProt ID Q92743 and PDB 3TJQ) and the fifth immunoglobulin-like domain from human Roundabout homolog 2 (UniProt ID 3 R. Zarivach and A. Sagi, unpublished data.)
TABLE 1
Sequences used for phylogenetic analysis

| Organism                      | Common name                       | Accession no. |
|-------------------------------|-----------------------------------|---------------|
| A. americanum                 | Lone star tick                    | ADE06668.1    |
| Xenopus laevis                | African clawed frog               | ABF71729.1    |
| Branchiostoma floridae        | Florida lancelet                  | XP_0026086.41 |
| Rattus norvegicus             | Common rat                        | NP_00102442.1 |
| Xenopus (Silurana) tropicalis | Western clawed frog              | XP_00294300.1 |
| Takifugu rubripess            | Puffer fish                       | XP_00397942.7 |
| Homo sapiens                  | Human                             | NP_00100756.4 |
| Anolis carolinensis           | Green anole                       | XP_00322380.1 |
| Salmo salar                   | Atlantic salmon                   | ACN11329.1    |
| Sus scrofa                    | Wild boar                         | XP_00192384.7 |
| Bos taurus                    | Bovine                            | NP_00109263.1 |
| Rana catesbeiana              | American bullfrog                 | ACO51872.1    |
| Branchiostoma belcheri        | Lancelet                          | BAB97382.1    |
| Musculus                      | House mouse                       | NP_001211.1   |
| Tetraodon nigroviridis        | Spotted green puffer fish         | CAP9909.1     |
| Oncorhynchus mykiss           | Rainbow trout                     | NP_00118120.0 |
| C. rogerceseyi                | Sea louse                         | ACO10479.1    |
| C. quadricarinatus            | Red claw crayfish                 | KCO92101      |
| Gorilla gorilla gorilla       | Western lowland gorilla           | XP_00404102.1 |
| Papio anubis                  | Olive baboon                      | XP_00391152.5 |
| Meleagris gallopavo           | Wild turkey                       | XP_00320560.1 |
| Gallus gallus                 | Chicken                           | XP_420577.3   |
| Biophila glabrata             | Freshwater snail                  | ARU4453.1     |

Q9HC4 and PDB 2EDJ). Structures were superimposed based on their backbones, improved fit was applied, and root mean square deviations were obtained in Angströms (Å) using the Swiss-Pdb Viewer. The final structural figures were generated using the PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA).

RESULTS

In screening a previously constructed AG suppression subtractive hybridization cDNA library prepared from male C. quadricarinatus crayfish (14), a contig encompassing a putative insulin-like-binding domain was identified. In this study, the complete sequence of this putative transcript was obtained and found to be just under 1200 bp in length (Fig. 1A). The cDNA sequence includes 5′ and 3′-UTRs of 33 bases and ~400 bases, respectively, with the 3′-UTR end terminating with a poly(A) tail (Fig. 1A). The open reading frame is 747 bases long and encodes a deduced protein of 248 amino acids. According to the SMART bioinformatics tool, the newly identified transcript encodes a deduced protein comprising an insulin-like-binding domain (IB), a kazal-type serine protease inhibitor domain (KI), and an immunoglobulin-c-2 domain (IgC2) in that order (Fig. 1B). Based on the presence of these domains, the gene product is classified as an insulin-like growth factor-binding protein homolog of the crayfish and was termed C. quadricarinatus IGBP (Cq-IGFBP).

Assigned to the IGBP family of peptides, Cq-IGFBP was considered in a multiple alignment with other IGBPs sharing a similar domain organization (Fig. 1B) using the ClustalW algorithm (supplemental Fig. S1). A phylogram, calculated from the obtained alignment (Fig. 2), emphasized the fairly limited number of invertebrate IGBPs that have been fully sequenced to date and the relatively low sequence conservation level that they share. It should be stressed that although these sequences had seemed to divert quite substantially one from another, they are nonetheless similarly organized, sharing the same domain-based architecture (Fig. 1B). By contrast, vertebrate IGBP sequences share greater protein sequence similarity. Within vertebrates, fish, and birds sequences were grouped together, whereas amphibian and reptile sequences were grouped together and found to be more similar to the mammalian group than to others. Cq-IGFBP showed the greatest sequence homology to that sequence from yet another crustacean, the copepod Caligus rogerceseyi. The sequence from a fellow arthropod, the lone star tick Amblyomma americanum, was found in the closest branch to that of the crustaceans, although the two branches are quite remote from one another.

As the linear organization of the three Cq-IGFBP signature domains was shared with selected IGBPs (Fig. 2 and supplemental Fig. S1), conservation at the level of three-dimensional structure was next considered. Accordingly, two comparative homology models, one of the IB and KI domains (residues 28–133) (Fig. 3A) and the second of the IG domain (residues 143–248) (Fig. 3B), were built using SWISS-MODEL for identifying suitable templates. The template used for building the comparative model of the Cq-IGFBP IB and KI domains was the N-terminal domain of human HtrA1 (Fig. 3C) (sequence identity 31%; coverage ~95%, supplemental Fig. S2), a protein previously shown to comprise an IgF binding domain that was structured in detail by x-ray crystallography (34). To generate a homology of the IG domain, the fifth IG-like domain from human Roundabout homolog 2 (Fig. 3D) (sequence identity 27%; coverage ~95%, supplemental Fig. S2) was employed as template. The latter has been shown to be conserved in other invertebrate IGBP homologs (35, 36) although its function with this respect is yet to be revealed. The final models of the IB and KI domains and of the IG domain were both found to resemble their reference structures with high accuracy and minimal structural variations, as indicated by 0.95 and 0.64 Å root mean square deviation values, respectively. Other indicators of the accuracy of these models exist. Previously, the (R/L)XXXX determinant was found to be crucial for insulin-like binding by human IGBP-3 and IGBP-5 (37, 38). A highly similar determinant was identified within the IB domain of Cq-IGFBP (XXRLXXLXX). These residues are predicted to form a part of a β-sheet and a small loop, which are located in an exposed position (Fig. 3A, gold ribbon).

The tissue expression pattern of Cq-IGFBP was demonstrated by RT-PCR (Fig. 4). Although the primers used were devised for cDNA originally isolated from an AG suppression subtractive hybridization cDNA library, the transcript was shown to be amplified in all of the tissues considered, namely the AG, hepatopancreas, testes, muscle, sperm duct, and green glands in mature male crayfish.

As a prerequisite for assessing Cq-IGFBP ability to bind the Cq-IAG hormone, a recombinant version of the early (rCq-IGFBP) was produced using the pET28 expression system in E. coli cells. The His₆-tagged rCq-IGFBP protein had to be retrieved from inclusion bodies via solubilization before its purification via immobilized metal affinity chromatography. A Coomassie Blue-stained SDS-PAGE gel showed that most of the proteins present in the solubilized inclusion bodies (Fig. 5A, BC) flowed through the nickel column, with minimal amounts of protein being removed from the resin by repetitive washing (Fig. 5A, FT W₃). A single protein band of ~30 kDa was greatly enriched in the column eluate (Fig. 5A, E lanes). In a subsequent
immunoblot performed with anti-His antibodies, the purified protein was specifically labeled (Fig. 5B, arrowhead). Bovine serum albumin, which served as a negative control for the immunoblot, was shown to be present in a replica gel stained with Coomassie Blue (Fig. 5B, CB). The purified 30-kDa protein band (Fig. 5A, E3) was excised from the gel and subjected to peptide analysis via tandem MS. A Sequest search performed against a database of the rCq-IGFBP protein that includes a His tag and an adjacent thrombin cleavage site confirmed that the isolated band indeed corresponds to rCq-IGFBP. In these experiments, the generated tryptic fragments were shown to cover ~59% of the mature Cq-IGFBP sequence (i.e. lacking the signal peptide) (Fig. 5C).

Co-precipitation experiments were performed to test whether Cq-IGFBP specifically binds the insulin-like AG hormone (Cq-IAG). Biotinylated rCq-IGFBP (brCq-IGFBP) was incubated with AG homogenate, coprecipitated with streptavidin-conjugated agarose beads, and resolved by SDS-PAGE. In the streptavidin-precipitated fraction, a band visualized using Coomassie stain (Fig. 6A, arrowhead) was excised from the gel and analyzed by MS. Such analysis of the tryptic digest of this band revealed two peptides corresponding to residues 48–59 of the Cq-IAG B-chain and residues 96–102 of the C-peptide (Fig. 6B1). The scores calculated for these two peptides when double-charged were 1.04 and 2.26, respectively. For the second peptide (FTVPDAR, corresponding to residues 96–102 of the C-peptide), the possible ‘b’ and ‘y’ ions are denoted (Fig. 6B2), and an intensity map is provided (Fig. 6B3).

The streptavidin-precipitated proteins were also resolved by SDS-PAGE and transferred onto nitrocellulose membranes for immunoblot analysis. For this purpose, a specific antibody against recombinant proCq-IAG was previously raised (16). In the pulldown fraction of the immunoblot, the antibodies specifically reacted with two peptide bands (Fig. 6C, arrowheads). The same peptides were also recognized in the AG homogenate, serving as a control (Fig. 6C, rightmost lane). The presence of the brCq-IGFBP in the pulldown fraction was verified by blotting using HRP-conjugated avidin (Fig. 6C, bottom).
IGFBP Binds an Insulin-like Androgenic Hormone in Decapods

FIGURE 2. Evolutionary relationship of proteins across taxa having the same order of domains as Cq-IGFBP. Based on a ClustalW algorithm, insulin-like-binding proteins having the same domains organization as Cq-IGFBP, from a wide array of organisms, are mapped on a phylogenetic tree (constructed using MEGA, Version 5.0 (29)). This evolutionary history was inferred using the Neighbor-Joining method and a bootstrap consensus tree of 5000 replicates. The branch representing Cq-IGFBP is indicated (arrow), and the bar represents the number of amino acid substitutions per site.

FIGURE 3. Three-dimensional modeling of the three domains of Cq-IGFBP. Models of the Cq-IGFBP IB, KI, and immunoglobulin c-2 type domains were designed based on their relative similarities to human references and are presented in a ribbon diagram. A, shown are the IB (blue, residues 28–102) and KI (purple, residues 103–133) domains. The 5’RXLXXLX’ (gold) sequence is the putative equivalent of the conserved determinant (IR/K)XXLXXL found to be crucial for insulin-like peptide-binding in human IGFBPs. B, shown is the IG domain (residues 143–248). C, shown is the solved x-ray structure of human HtrA1 (34). D, shown is the solved solution NMR structure of the fifth IG-like domain from human Roundabout homolog 2.

ligand dot blot approach, the resolution of the IAG-IGFBP binding association was partially enhanced. It was shown that a portion of the Cq-IAG A-chain (residues 23–38) is recognized by recombinant IGFBP in vitro (Fig. 6D). The ability of brCq-IGFBP to bind other members of the insulin-like superfamily (e.g. human recombinant insulin and IGF-I) was also examined using streptavidin-based precipitation and SDS-PAGE separation, as above. The specificity of Cq-IGFBP for the IAG hormone was further established as it did not bind human insulin completely, as it was present in the not bound (NB) fraction and absent from the pulled-down (PD) fraction even though the gel was visualized by silver staining (Fig. 6E). A fairly low amount of IGF-I might have in fact been bound to Cq-IGFBP, as indicated by a faint band in the PD lane (Fig. 6E).

DISCUSSION

The findings reported here solve another piece in the puzzle of what seems to be a complicated scheme involving a unique insulin-like peptide with clear androgenic activities. Indeed, these results contribute to an emerging body of evidence revealing the involvement of ILP signaling in primary and/or secondary sexual differentiation phenomena in the animal kingdom from the relatively early appearance of invertebrates to the later evolution of vertebrates (10, 11, 13, 15, 39, 40).

In recent years, thanks in part to the giant leaps made in sequencing technology, many new invertebrate ILP homologs have been identified (41–45), including the crustacean insulin-like AG hormone (15). In the past, growth-related insulin-like peptides (namely IGFs) were thought to be exclusive to vertebrates. However, it is becoming ever clearer that this insulin-like growth-related family of peptides originated far earlier in the evolution of animals. Based on our knowledge and comparative studies showing that IGFs are secreted into the circulatory system where they interact with highly specific IGFBPs, it is less surprising that such binding proteins have been thus far identified in several invertebrate species, including crustaceans. Thus, our findings present exciting evidence adding new perspective to studies on the early evolution of the IGFBP family, which currently focus mainly on vertebrates (46). It should be emphasized that previous studies in decapods have linked these binding proteins to immune responses (22, 23) rather than to reproduction and growth. With the present findings, Cq-IGFBP joins the list of other arthropod IGFBP homologs, namely Imp-L2 and Sf-IBP, previously reported in the fruit fly Drosophila melanogaster and the moth Spodoptera frugiperda (Sf), respectively (35, 47). The insight gained here thus further supports the common incidence of insulin-like signaling, specifically in invertebrates and, generally, in Animals. Nevertheless, in contrast to Cq-IGFBP, Imp-L2 was shown to bind various Drosophila ILPs with no gender spec-
FIGURE 5. Purification of recombinant Cq-IGFBP. A, shown is mature recombinant Cq-IGFBP (rIBP) was retrieved from solubilized inclusion bodies subjected to a nickel-based chromatography. Fractions were separated on SDS-PAGE and stained with Coomassie Blue (CB). B, purified rIBP and BSA separated by SDS-PAGE were either stained with CB or blotted with a rabbit anti-His IgG (IB). C, the identity of rIBP was validated using mass spectrometry. On the full sequence of rIBP (positions are indicated) validated residues are highlighted in dark gray (high identification score) and light gray (medium score). The His tag and thrombin sequences derived from the expression vector are underlined.

MW, molecular weight standard; Bf, before column; FT, flow-through; W, wash (1, 3 M urea; 2, 1.5 M urea; 3, no urea); E, elution.

FIGURE 6. Cq-IGFBP specifically binds the Cq-IAG hormone. A, brCq-IGFBP was incubated with AG homogenate followed by protein precipitation. Fractions were separated on SDS-PAGE and stained with Coomassie Blue. A band corresponding to a suspected Cq-IGFBP ligand is indicated (arrowhead). AP, streptavidin precipitation; Ho, untreated homogenate. B, shown is identification of Cq-IAG by mass spectrometry. The suspected band (arrowhead in Fig. 6A) was shown to include two peptides (underlined) shown to be within the Cq-IAG prohormone sequence (B1). One of the two peptides is shown as an example, with its relevant b and y ions indicated (B2) as is an intensity map (B3) of the peptide A chain. E, the ability of rCq-IGFBP to bind other insulin-like peptides was also examined. Ten g of either human recombinant insulin or IGF-I were incubated with brCq-IGFBP, streptavidin-precipitated, and resolved on SDS-PAGE. Before precipitation (Bf), not bound (NB), and pulled down (PD) are shown. The gel was silver-stained.
IGFBP Binds an Insulin-like Androgenic Hormone in Decapods

ificity (35, 48, 49), whereas Sf-IBP was shown to bind various exogenous human ILPs (47). Thus, crustacean IAG signaling presents a novel twist on these interactions in that an IGFBP is suggested to serve as a pivotal component able to discriminatively bind primarily an endogenous male ILP sex-hormone. Cq-IAG has been suggested to have pleiotropic effects (16, 20, 21, 50) and, therefore, must induce signaling in cells of various origin (e.g., muscle, testis, and epithelium). It is noteworthy that the tissue expression pattern of Cq-IGFBP differs from that of Cq-IAG as previously described (14). The spatial profiling of Cq-IGFBP, revealing the wide expression of this transcript, supports the hypothesis that IAG signaling is modulated by the expression/production of Cq-IGFBP in multiple tissues.

To date, a native proteinaceous IAG hormone was isolated in isopod but not in decapod crustaceans (51, 52). In the latter, where IAG mRNA sequences were obtained, bioinformatics predicts sites for C-peptide cleavage, although this has never been demonstrated. In addition, the deduced IAG peptide does not fall under the classic definition of an IGF given that it lacks D and E domains and is suggested to undergo C-peptide removal and is instead considered an ILP. As previously discussed, IGFs are secreted as a single-chained peptide (1) and transported in complex with IGFBP family members in the blood. We have shown that regardless of whether proteolysis occurs, the endogenous proCq-IAG hormone, which has been identified in this study for the first time, is specifically recognized by Cq-IGFBP based on MS and immunological studies. These results thus join the observation made in S. frugiperda (47) and show that arthropod-IGFBPs are capable of binding proinsulin-like hormones. Moreover, our findings support those of Yamanaka et al. (53) arguing the common dogma suggesting that only IGFs interact with high affinity soluble binding proteins (8). As such, the findings reported here shed new light on the early development and co-evolution of the IGFBP superfamily and its diverse ligands.

As for the modeling of the IGF-IGFBP interaction, it was previously reported that such binding is primarily mediated via the B and A chains, with conflicting evidence regarding the involvement of the C-peptide having been shown (54). Our usage of a recombinant Cq-IAG-based peptide has provided new details on the specific sites used for hormone recognition by crayfish IGFBP. It is evident that residues 23–38 of the IAG A-chain are sufficient for binding by Cq-IGFBP in vitro. As for the sites of interaction in IGFBP, numerous studies have shown the involvement of the N-terminal region in ligand binding, as recently reviewed (55). Specifically, site-directed mutagenesis efforts have shown the importance of the (R/L)LXXXLL motif for such binding (37, 38). Cq-IGFBP provides another example of a family member bearing this sequence (excluding the final leucine). Based on homology modeling, the N-terminal domain of Cq-IGFBP encompassing the IB domain shows the highest structural homology to human HtrAI. Although the structure of HtrAI was solved in detail and found to present an IB domain, it was shown not to bind either human IGF-1 or IGF-II (34). Thus, although both proteins possess structurally similar IB domains, only Cq-IGFBP is functional, i.e. capable of binding ILP. This can be rationalized by the presence of the conserved determinant (R/L)LXXXLL (37, 38) that is absent in HtrAI. Furthermore, previous studies have clearly shown the necessity of the IGFBP C-terminal domain for ligand binding (55–57), which may also hold true for Cq-IGFBP.

In conclusion, ILP signaling is found at the core of sexual differentiation in various organisms. Cq-IGFBP, however, is the first family member involved in this process shown to specifically interact with a gender-specific androgenic ILP. It is striking that Crustacea, reaching back to the Cambrian (~500 million years ago) (58), present IGFBPs that share high sequence similarity and physiological roles with homologous proteins across evolution (despite the low sequence conservation of ILP ligands). This suggests that parallel situations yet to be identified exist in other fundamental life forms.

Acknowledgments—We thank Omri Lapidot for technical assistance with crayfish maintenance. We also thank Dr. Raz Zarivach and Natalie Zeytuni for assistance with protein modeling.

REFERENCES

1. Chan, S. J., and Steiner, D. F. (2000) Insulin through the ages. Phylogeny of a growth promoting and metabolic regulatory hormone. Am. Zool. 40, 213–222
2. Steiner, D. F. (1978) On the role of the proinsulin C-peptide. Diabetes 27, 145–148
3. Paveli, J., Matijevi, T., and Knezevi, J. (2007) Biological and physiological aspects of action of insulin-like growth factor peptide family. Indian J. Med. Res. 125, 511–522
4. Capoluongo, E. (2011) Insulin-like growth factor system and sporadic malignant melanoma. Am. J. Pathol. 178, 26–31
5. Gallagher, E. J., and LeRoith, D. (2011) Minireview. IGF, insulin, and cancer. Endocrinology 152, 2546–2551
6. Laron, Z. (2001) Insulin-like growth factor 1 (IGF-1). A growth hormone. Mol. Pathol. 54, 311–316
7. Hwa, V., Oh, Y., and Rosenfeld, R. G. (1999) The insulin-like growth factor-binding protein (IGFBP) superfamily. Endocr. Rev. 20, 761–787
8. Kelley, K. M., Schmidt, K. E., Berg, L., Sak, K., Galima, M. M., Gillespie, C., Balogh, L., Hawayek, A., Reyes, J. A., and Jamison, M. (2002) Comparative endocrinology of the insulin-like growth factor-binding protein. J. Endocrinol. 175, 3–18
9. Baxter, R. C. (2000) Insulin-like growth factor (IGF)-binding proteins. Interactions with IGFs and intrinsic bioactivities. Am. J. Physiol. Endocrinol. Metab. 278, E967–E976
10. Emlen, D. J., Warren, I. A., Johns, A., Dworkin, I., and Lavine, L. C. (2012) A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. Science 337, 860–864
11. Fedina, T. Y., Kuo, T. H., Dreisewerd, K., Dierick, H. A., Yew, J. Y., and Pletcher, S. D. (2012) Dietary effects on cuticular hydrocarbons and sexual attractiveness in Drosophila. PLoS ONE 7, e9799
12. Kawamura, K., Kumagai, J., Sudo, S., Chun, S. Y., Pisarska, M., Morita, H., Toppuri, I., Fu, P., Wade, J. D., Bathgate, R. A., and Hsueh, A. J. (2004) Paracrine regulation of mammalian oocyte maturation and male germ cell survival. Proc. Natl. Acad. Sci. U.S.A. 101, 7323–7328
13. Net, S., Verma-Kurvari, S., Menenmies, J., Vassalli, D. J., Efstratiadis, A., Accili, D., and Parada, L. F. (2003) Testis determination requires insulin receptor family function in mice. Nature 426, 291–295
14. Manor, R., Weil, S., Oren, S., Glazer, L., Aflalo, E. D., Ventura, T., Chalifa-Caspi, V., Lapidot, M., and Sagiv, A. (2007) Insulin and gender. An insulin-like gene expressed exclusively in the androgenic gland of the male crayfish. Gen. Comp. Endocrinol. 150, 326–336
of insulin receptor activation by insulin-like growth factor binding proteins. J. Biol. Chem. 272, 30729—30734
54. Denley, A., Cosgrove, L. J., Booker, G. W., Wallace, J. C., and Forbes, B. E. (2005) Molecular interactions of the IGF system. Cytokine Growth Factor Rev. 16, 421–439
55. Forbes, B. E., McCarthy, P., and Norton, R. S. (2012) Insulin-like growth factor binding proteins. A structural perspective. Front. Endocrinol. 3, 1–13
56. Bach, L. A. (2004) The insulin-like growth factor system. Towards clinical applications. Clin. Biochem. Rev. 25, 155–164
57. Firth, S. M., and Baxter, R. C. (2002) Cellular actions of the insulin-like growth factor binding proteins. Endocr. Rev. 23, 824–854
58. Walossek, D. (1998) in Crustaceans and the Biodiversity Crisis (Schram, F. R., and von Vaupel Klein, J. C., eds), pp. 3–27, Brill, Leiden, Holland