The expression and regulation of lipid transport proteins in the desert locust, *Schistocerca gregaria*

by

Ahmed Sayed Ahmed

M.Sc., Zagazig University, Egypt, 2006

B.Sc., Zagazig University, Egypt, 2001

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the

Department of Biological Sciences
Faculty of Science

© Ahmed Sayed Ahmed 2018

SIMON FRASER UNIVERSITY

Fall 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

Approval

Name:	**Ahmed Sayed Ahmed**
Degree:	**Doctor of Philosophy**
Title:	**The expression and regulation of lipid transport proteins in the desert locust, *Schistocerca gregaria***
Examining Committee:	**Chair: Michael Silverman** Professor
	Norbert Haunerland Senior Supervisor Professor
	Carl Lowenberger Supervisor Professor
	Inigo Novales Flamarique Supervisor Professor
	Jim Mattsson Internal Examiner Associate Professor
	Erwin Huebner External Examiner Professor Department of Biological Sciences University of Manitoba
Date Defended/Approved:	November 19, 2018

Abstract

Lipids play a central role in insects, both for storage of nutrients and as an energy source during development and dispersal. Due to their low water solubility, special transport mechanisms are required for their efficient mobilization and utilization. In this thesis, I studied intra- and extracellular proteins involved in lipid transport in the desert locust, *Schistocerca gregaria*. Vitellogenins are very high-density lipoproteins produced by adult females and deposited into the developing eggs. Two different vitellogenins, named VG-A and VG-B, are expressed in locust fat body. Their complete cDNA transcripts of ~5.6 kb each have been sequenced, coding for two proteins of ~200 kDa each. VG-A and VG-B are co-expressed in similar amounts by mature females, commencing 11 days after adult eclosion, and continuing at high levels throughout the entire adult life. The expression of both proteins is dependent on the nuclear transcription factors Met or RXR, and knockdown of each of these proteins almost completely eliminates VG expression. A similar expression profile was observed in adult muscle for the cytosolic fatty acid binding protein FABP, albeit in both sexes. The direct knockdown of the strongly expressed FABP by RNA inference reduced its levels to less than 2% of what is normally found 3 weeks after adult eclosion. In a series of flight experiments, it was demonstrated that in the absence of FABP, insects are incapable of engaging in flight longer than 30 min; at this time, most carbohydrate resources have been depleted, and locusts normally switch to lipids as the sole fuel for muscle energy production. Short-term flight performance of FABP knockdown locusts was identical to control insects, suggesting that the lack of FABP does not interfere with carbohydrate metabolism. Moreover, the mobilization of lipids in the fat body and their transport by the major hemolymph lipoprotein lipophorin was indistinguishable from control animals. In contrast, knockdown of apolipophorin III, which is essential for lipid transport during flight, completely eliminated flight capability, even for short duration flights. Taken together, this thesis highlights the essential role of lipid transport proteins for locust reproduction and dispersal and identifies potential targets for insect control strategies.

Keywords: vitellogenin; FABP; apoLp-III; Met; RXR; *Schistocerca gregaria*

Dedication

To my parents, my wife, and my dear sons Abdel Rahman and Omar.

Acknowledgments

I would like to express my gratitude and sincere thanks to everyone who has helped me over the past few years to finish this work. I may start by thanking Allah who has chosen me for this path and empowered me to complete this thesis.

To my senior supervisor, Dr. Norbert Haunerland, thank you for every moment you have been there for me with your support, knowledge, patience, and encouragement; not only in the lab to overcome the scientific problems and challenges but with a significant impact in my personal life. Without you being there for me, I would not be able to come to this moment and finish this work. A Special thank goes to Dr. Jutta Haunerland for her daily support that extends beyond the vast knowledge of molecular techniques that I have learned from her to her soothing words which have enlightened my way during the hard times.

I would like to thank my supervisory committee members Drs. Carl Lowenberger, Inigo Flamarique, and the former committee member; Allison Kermode. Dr. Carl has shown a great interest in our work. His suggestions and contributions have added a lot to this work. Dr. Carl has offered me the opportunity to attend his lab meetings as he treated me as one of his students. During these meetings, I have learned a lot about other molecular techniques and his amazing projects attempting to prevent the transmittance of different viruses by mosquitoes. His editing on the first manuscript of this thesis has solidified and enriched this work. Dr. Inigo, thank you for accepting the role to serve as a committee member on short notice. I have worked with Dr. Inigo on different projects where I have enjoyed learning about salmon and their vision mechanisms. During these projects, he has been following up with my thesis progress with valuable suggestions and support. Dr. Allison Kermode had served on the supervisory committee for a few years before she decided to resign, and then to step down from the supervision. We have sadly had to say thank you and goodbye and wishing her the best in the future out of the academia. Her valuable discussion and support during the annual committee meetings have helped us to finish the work.

Many thanks to my former colleagues in Haunerland lab; Cecily Bernales, Katrina Salvante and Xin Jiang for their assistance in learning all the molecular techniques that I have used in my research and for creating that welcoming, friendly environment that I

have enjoyed since my first day in the lab. I would like to thank David Qu and Sanjeeva Rajapakse for their valuable contribution in chapter four of this thesis by designing the FABP-dsRNA and knockdown the expression of the FABP in the flight muscles.

My sincere acknowledgment to the Egyptian government, Zagazig University in Egypt, and the Bureau of Cultural and Educational Affairs of Egypt in Canada for their support during the first two years of my research and choosing me for the joint-supervision scholarship which was entirely sponsored and financially supported by the Egyptian government.

Last but not least, I would like to thank and acknowledge the great support of my family; my mother, brothers, and my wife “Nora” and my kids for being there with their unimaginable understanding, support, and love.

Thank you, everyone.

Table of Contents

Approval ... ii
Abstract ... iii
Dedication ... iv
Acknowledgments ... v
Table of Contents ... vii
List of Tables ... x
List of Figures ... xi
List of Acronyms ... xii

Chapter 1. Introduction ... 1
1.1. Lipid storage and functions ... 2
1.2. Lipid Transport ... 5
1.3. Cytosolic lipid binding proteins ... 6
1.4. Vertebrate Lipoproteins ... 6
1.5. Lipid Transport in Insects ... 8
1.5.1. Hemolymph lipophorins ... 9
1.5.2. Storage lipoproteins ... 10
1.5.3. Vitellogenin (VG) ... 11
1.6. Endocrine regulation in insects ... 12
1.6.1. Ecdysone hormone ... 12
1.6.2. Juvenile hormone (JH) ... 13
1.7. Nuclear receptors as transcription factors ... 14
1.8. The desert locust, *Schistocerca gregaria* ... 14
1.9. Dissertation objectives ... 16
Connecting statement 1: Structure and expression of vitellogenin ... 17

Chapter 2. Primary structure and temporal expression of two vitellogenins from the desert locust, *Schistocerca gregaria* ... 18
Abstract ... 19
2.1. Introduction ... 20
2.2. Materials and methods ... 22
2.2.1. Insects and samples collection ... 22
2.2.2. RNA collection and preparation ... 22
2.2.3. cDNA synthesis and PCR reactions ... 23
2.2.4. 3` and 5` RACE library construction ... 23
2.2.5. Real-time qPCR ... 24
2.2.6. SDS PAGE gel-electrophoresis ... 25
2.3. Results ... 26
2.3.1. cDNA sequences of VG-A and VG-B ... 26
2.3.2. Nucleotide and amino acid sequence analysis: ... 27
2.3.3. Sequence similarity and phylogenetic analysis: ... 28
2.3.4. Vitellogenin expression in adult females: ... 31

2.3.5. Vitellogenin protein analysis in the hemolymph....33
2.4. Discussion:....34
Connecting statement 2: Nuclear receptors involved in vitellogenin expression....39

Chapter 3. The nuclear receptors Met and RXR promote vitellogenin expression in the desert locust....40
Abstract....41
3.1. Introduction....42
3.2. Material and Methods....44
3.2.1. Insects....44
3.2.2. RNA extraction and cDNA synthesis....44
3.2.3. Polymerase chain reaction (PCR)....44
3.2.4. Double-stranded RNA interference....45
3.2.5. Treatment and sample collection....46
3.2.6. Quantitative real-time PCR....46
3.2.7. SDS-gel protein electrophoresis....47
3.3. Results....47
3.3.1. Search for the Met and RXR Sequences....47
3.3.2. The effect of Met dsRNAi on vitellogenin expression....49
3.3.3. The effect of RXR dsRNAi on vitellogenin expression....51
3.4. Discussion....53
Connecting statement 3: Lipids are needed for reproduction and dispersal....57

Chapter 4. FABP silencing impedes extended flight in the desert locust, *Schistocerca gregaria*....58
Abstract....59
4.1. Introduction....60
4.2. Material and Methods....62
4.2.1. Insects....62
4.2.2. Flight experiments....63
4.2.3. Protein extraction and analysis....63
4.2.4. Lipid analysis....63
4.2.5. Total RNA preparation....64
4.2.6. cDNA synthesis....64
4.2.7. Synthesis of dsRNA....64
4.2.8. Injection of dsRNA....65
4.2.9. Quantitative real-time PCR (qPCR)....65
4.3. Results....66
4.3.1. FABP real-time PCR....66
4.3.2. Flight experiments....68
4.3.3. Lipids content analysis....70
4.4. Discussion....72
Connecting Statement 4: apolipophorin-III and lipid transport in the hemolymph....75

Chapter 5. Apolipophorin III: sequence, expression pattern, and role in sustained flight…………………………………………………………………………………...76
Abstract...77
5.1. Introduction...78
5.2. Materials and Methods 80
5.2.1. Insect treatment and sample collection 80
5.2.2. RACE libraries 80
5.2.3. ApoLp-III knockdown 80
5.2.4. Flight experiments 81
5.2.5. Gene expression 81
5.3. Results 82
5.3.1. cDNA nucleotide sequence of apoLp-III..82
5.3.2. Deduced amino acid sequence...83
5.3.3. Temporal expression analysis 85
5.3.4. Knockdown of apoLp-III expression in the fat body 86
5.3.5. Flight experiments 88
5.4. Discussion 89

Chapter 6. General Conclusions 94

References..97

Appendix A. Supplemental data for Chapter 2...115

Appendix B. Supplemental data for Chapter 3...128

Appendix C. Supplemental data for Chapter 5...129

List of Tables

Table 2-1 Specific real-time primers used to measure the expression of VG genes 25
Table 2-2 Sequence identity and similarity between different vitellogenins....30
Table 3-1 Primers used to generate the Met and RXR dsRNAi 44
Table 3-2 Primers used to measure the gene expression of Met, RXR, and EF1α 47
Table 3-3 Ct values of the housekeeping genes of the control and the treated insects....50
Table 3-4 Transcript changes after treatment with Met dsRNAi....51
Table 3-5 Transcript changes after treatment with RXR dsRNAi 52
Table 4-1 Expression of FABP mRNA in adult locusts. 67
Table 4-2 Hemolymph lipid content of resting and flying locusts. 70
Table 5-1 Primers used to amplify apoLp-III cDNA and synthesize the dsRNAi. 81
Table 5-2 Real-time primers used to assess apoLp-III expression level 81
Table 5-3 Ct values of the housekeeping genes GAPDH and EF1α in control and treated insects....87
Table 5-4 The relative expression of apoLp-III in the treated insects with apoLp-III dsRNA. 88
Table 5-5 The relative expression of apoLp-III after flight experiment. 89

List of Figures

Figure 1-1 Major lipid classes 3
Figure 1-2 The nymph and the adult of the desert locust, *Schistocerca gregaria* 15
Figure 2-1 Vitellogenin mRNA detection in fat body of adult females. 26
Figure 2-2 Conserved sequence motifs and domains in vitellogenin A and B 28
Figure 2-3 Predicted gene structure for VG-A and VG-B from *Locusta migratoria* ... 29
Figure 2-4 Phylogenetic tree of the vitellogenin gene family. 31
Figure 2-5 Relative temporal expression of VG A and B in the fat body of the desert locust. 32
Figure 2-6 Identification of different VG protein bands in the hemolymph of the adult locust female. 34
Figure 3-1 PCR results of the amplified potential Met nuclear receptor. 48
Figure 3-2 Sequence of the Met fragment from *S. gregaria* 48
Figure 3-3 Schematic representation of the two RXR isoforms from *S. gregaria* 49
Figure 3-4 Amplified PCR products of RXR 49
Figure 3-5 SDS PAGE of hemolymph proteins from14 day old adult females 51
Figure 3-6 SDS PAGE of hemolymph proteins from RXR RNAi treated 14-day old adult females 53
Figure 4-1 FABP expression in adult locusts. 67
Figure 4-2 SDS PAGE of soluble muscle proteins. 68
Figure 4-3 Flight performance of adult locusts. 69
Figure 4-4 Tethered flight of locusts. 69
Figure 4-5 Behavioral changes of exhausted FABP knockdown locusts. 70
Figure 4-6 TLC separation of hemolymph lipids. 71
Figure 5-1 Amplification of apoLp-III using primers from the migratory locust. 82
Figure 5-2 An illustration of the complete cDNA sequence of *S. gregaria* apoLp-III. 83
Figure 5-3 Schematic graph of the deduced amino acids sequence of apoLp-III. 83
Figure 5-4 Alignment of locust apoLp-III sequences 84
Figure 5-5 Graphical illustration of the *L. migratoria* apoLp-III gene. 85
Figure 5-6 Alignment of the nucleotide sequences and their deduced amino acid sequences between apoLp-III a & b from *L. migratoria*. 85
Figure 5-7 The expression of ApoLp-III measured using real-time qPCR in the total RNA of the fat body. 86
Figure 5-8 3-D structures of insect apoLp-III 91

List of Acronyms

apoLp-III	Apolipophorin-III
DAG	Diacylglycerol
dsRNAi	Double-strand RNA interference
EF1α	Elongation factor 1-alpha
EST	Expressed sequence tag
FABP	Fatty acid binding protein
FFA	Free fatty acid
GAPDH	Glyceraldehyde 3-Phosphate dehydrogenase
HDLp	High-density lipoprotein
iLBP	Intracellular lipid binding protein
JH	Juvenile hormone
LDLp	Low-density lipophorin
Met	Methoprene tolerant nuclear receptor
PBS	Phosphate buffer
PCR	Polymerase chain reaction
RACE	Rapid amplification of cDNA ends
RXR	Retinoid-X-Receptor
TAG	Triacylglycerol
VG	Vitellogenin

Chapter 1. Introduction

1.1. Lipid storage and functions

Insects represent the most successful and diverse class among living organisms. Over millions of years, they have evolved unique morphological, anatomical and physiological adaptations. Although these adaptations have allowed them to succeed in living in almost every habitat on earth, they make it difficult to cast a general physiological theme of any of these evolutionary mechanisms based on limited observations of a few insect species. For decades, lipids as other biological molecules in insects have received considerable attention in diverse species in an attempt to highlight their variable roles in this success.

The success of insects is profoundly linked to the significant roles of lipids in various physiological and developmental schemes such as in embryogenesis, reproduction, metamorphosis, migration, diapause, and flying (reviewed in Downer and Matthews, 1976). Lipids (Figure 1-1) comprise a broad class of biological macromolecules that play a variety of indispensable roles in living cells, for example as an energy storage resource, signaling molecules, and structural components. Many of these functions are linked to the hydrophobic nature of these hydrocarbon molecules, which include fatty acids and their derivatives (waxes, acyl-glycerols, phospholipids), as well as steroids and other fat-soluble vitamins.

Fatty acids are carboxylic acids with a long aliphatic chain (>14 carbons); these can be saturated or contain between one and four double bonds. The chain length and degree of desaturation influence their biological properties. For example, the longer a chain is, the more hydrophobic the molecule will be; saturated fatty acids are relatively rigid molecules, while double bonds lower the melting point and provide fluidity to fatty acid-containing structures. Fatty acids are mostly present as esters with glycerol, which has three hydroxyl groups that can react with fatty acids, resulting in mono-, di-, or tri-acyl glycerol molecules (MAG, DAG, TAG, often called mono-, di-, and triglycerides); TAG is the common form of lipid storage, and is often referred to as a “fat”. In phospholipids, which are the main components of membranes, one hydroxyl-group of glycerol forms an ester with phosphoric acid; as phosphoric acid has three acid groups, the resulting phospholipid is still negatively charged under physiological conditions. Therefore, phospholipids form bi-layers in which the hydrophobic fatty acid tails stick together, and the hydrophilic phosphate head group faces outwards to the environment. In the

membranes that embrace living cells or the intracellular compartments like the Golgi apparatus, the endoplasmic reticulum or the nucleus, lipids appear to be the main structural component. The properties of these membranes are defined by the fatty acid composition of the phospholipids, the presence of other lipid molecules including cholesterol, and numerous membrane proteins that can act as receptors, enzymes, or transport proteins. As such, membrane lipids are also involved in various cellular signal transduction pathways (see Fernandis and Wenk, 2007).

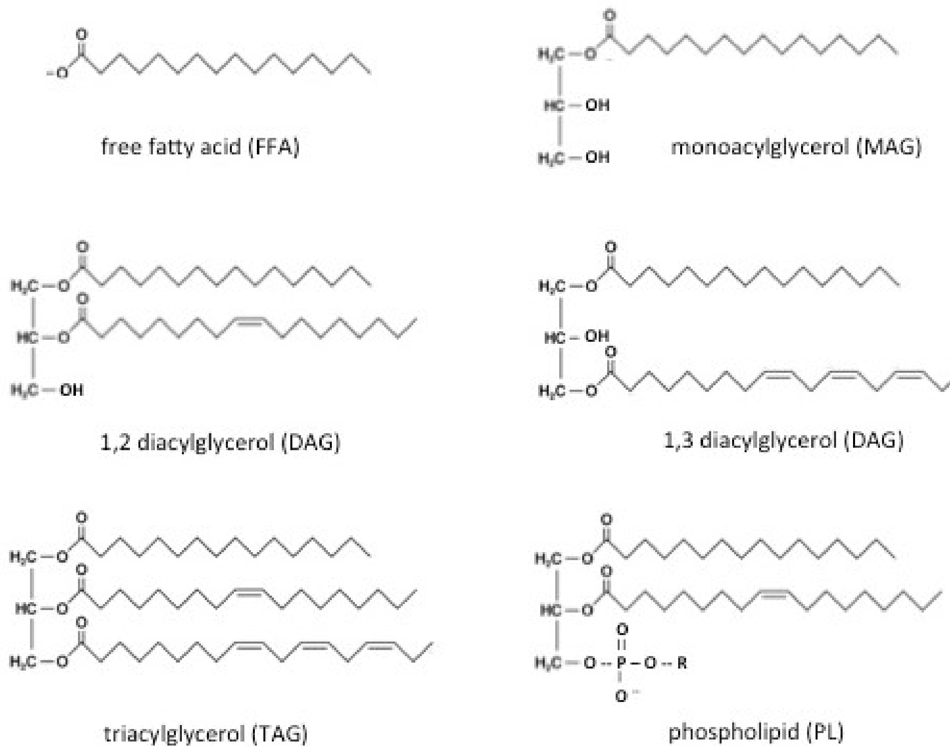


Figure 1-1 Major lipid classes

Most naturally occurring lipids contain fatty acids with an even number (between 14 and 20) of carbon atoms, with up to 4 double bonds. Their non-polar hydrocarbon tails make lipids hydrophobic, but the polar head group in phospholipids normally turns to the aqueous environment.

Another essential use of lipids in insects is protection from dehydration; considering the relatively high surface area to the volume ratio in insects and the open circulatory system which make them more prone to drought. Cuticular lipids create a barrier against water loss (Downer and Matthews, 1976). Further, upon oxidation lipids

yield more metabolic water than other fuel sources that makes lipids the favorable fuel source during the insects' flight and migration.

Lipids are an essential part of the insect's food, and they are digested and taken up in the midgut (review: Canavoso *et al*., 2001). Lipids are transported as DAG through the hemolymph and eventually stored as a TAG in the fat body. The various lipid components can be synthesized through complex biosynthetic pathways, with the exception of a few essential lipid molecules which have to be part of the animal's diet. For insects, these include the polyunsaturated fatty acids; linoleic acid and linolenic acid, as well as cholesterol (Canavoso *et al*., 2001).

As energy source reserves, lipids accumulate as triglycerides (TAGs) mainly in the adipose tissue. The stored TAGs must be mobilized and transported to their target tissue, to provide the cells with the required energy through β-oxidation of their constituent fatty acids (Athenstaedt and Daum, 2006). In addition to triglycerides, animal cells also store glycogen, a storage polysaccharide, which can be readily metabolized to release soluble sugar molecules as a quick energy source. The biochemical and biophysical properties of each fuel source determine the suitable usage strategy in different animal tissues depending on the metabolic needs of each tissue. Triglycerides are stored in a dense anhydrous form that allows the storage of large amounts within the hydrophilic environment of the cell. Though the mobilization of lipids is slower than glycogen which quickly associates with water, they generate more energy per unit of weight and yield more metabolic water (Arrese and Soulages, 2010; Weber, 2011). These characteristics have made lipids the most appropriate fuel source to sustain prolonged tasks despite their lower rate of ATP production and the low degree of solubility in water. The physical properties of lipids determine most of their functions *in vivo*, and how they interact with other lipids or other types of biological molecules like proteins and carbohydrates.

As water represents the central molecule of life and forms the substrate environment for most metabolic reactions, the low solubility of lipids in polar compounds such as water dramatically influences the interactions of these molecules in complex organisms and hinders their transport, and subsequently their bioavailability in different tissues. Consequently, living cells express numerous proteins to support the transport of lipids. Lipids noncovalently bind intra- and extracellularly with various types of lipid-binding proteins known as lipoproteins. Extracellularly, the lipoproteins contain a variety of fats

(acyl-glycerides and phospholipids) which form large complexes with apolipoprotein components, while within cells, small binding proteins bind to individual fatty acids or other monomeric lipids.

1.2. Lipid Transport

Lipids are stored in adipose tissue in specialized intracellular cytoplasmic compartments known as lipid droplets. The lipid droplet consists of a monolayer of phospholipids and cholesterol surrounding nonpolar lipids (triglycerides and cholesterol esters). During high energy demand activities, lipolysis takes place to liberate the fatty acids into the circulatory system as acyl-glycerides or as free fatty acids (FFA), where they associate with proteins that facilitate the transport to their target tissues. In mammals, serum albumin binds to several molecules of FFAs and distributes them through the body fluids. While the endogenous triglycerides can be carried in the bloodstream in the core of the very low-density lipoprotein (VLDL) (Van der Vusse *et al*., 1998), exogenously absorbed dietary lipids are transported as triglycerides to the target tissue in the form of chylomicrons. At the target cells, lipoprotein lipases release free fatty acids and glycerol from VLDL and chylomicrons.

The translocation of the free fatty acids through the plasma membrane of target cells is not well understood. While fatty acid may diffuse through the hydrophobic bilayer of the membrane, other evidence suggests the involvement of different membrane proteins like plasmalemmal fatty acid binding protein ($FABP_{pm}$), fatty acid translocase (FAT/CD36), and fatty acid transport protein (FATP). These transmembrane proteins are thought to increase the efficiency of fatty acid intake through the cell membrane (Baillie *et al*., 1996; Jia *et al*., 2007). Within the target cell, cytosolic fatty acid binding proteins (FABPs) are responsible for the intracellular transport of sequestered fatty acids, for example to the mitochondria for β-oxidation. Thus, the inner coupling with FABP maintains the fatty acid gradient and allows the continuing influx of free fatty acids into the cells. It may also prevent damage to intracellular membranes and their associated proteins due to the detergent-like activity of free fatty acid molecules (Glatz *et al*., 1998).

1.3. Cytosolic lipid binding proteins

FABPs are members of a highly conserved multiprotein superfamily, the intracellular lipid binding proteins (iLBPs). The members of this family have been recognized in the animal kingdom in both vertebrates and invertebrates, but have not been found in other organisms. The tertiary structure of these proteins is highly conserved as a "barrel" of antiparallel beta-strands, and a "lid" of two connected alpha helices. The core of this "beta-barrel" in most of these proteins has the ability to accommodate a single long fatty acid chain (Banaszak *et al*., 1994), with only one exception: liver FABP, which has been found to bind to two fatty acid molecules (Haunerland *et al*., 1984; Thompson *et al*., 1997).

The nomenclature of these proteins mostly has followed a conventional way, naming FABPs after the tissues in which they were initially discovered, even though most of these proteins were later found to be expressed in multiple tissues. Based on phylogenetic analyses, four subfamilies of iLBPs have been described in vertebrates, while in insects two different iLBPs have been identified; of these, one appears to be limited to inside the midgut while the expression of the other one seems to be spread between many tissues such as the brain, muscles, and fat body (Haunerland and Spener, 2004).

The cytosolic FABPs are mainly engaged in carrying free fatty acids intracellularly to the mitochondria for β-oxidation. In both insects and mammals, it was found that induced extensive exercise up-regulates the expression of these genes in muscle tissues (Haunerland, 1994; Clavel *et al*., 2002), which promotes the hypothesis that these proteins are needed for efficient fatty acid translocation.

1.4. Vertebrate Lipoproteins

As it has been mentioned, animals recruit various lipoproteins to overcome the low solubility of lipids in the internal aqueous environments. Animals exploit the vast arrangements flexibility and the resulting conformational variations of polypeptide chains to create shuttle systems to transport lipids between different compartments. These lipoproteins facilitate the mobilization of water-insoluble fats from various cellular compartments and for the extracellular translocation between the different tissues.

Besides, lipids and other essential nutrients may associate with protein components to form lipoproteins that can enter target tissues by endocytosis and stored there for later use as an energy source or as biosynthetic building blocks during different developmental processes.

Lipoproteins can be classified in a variety of ways, based on their physical, functional, or structural characteristics. For instance, they can be categorized based on their density into different classes: high density lipoprotein (HDL, >1.063 g/ml), intermediate-density lipoprotein (IDL, 1.006–1.019 g/ml), Low-density lipoprotein (LDL, 1.019–1.063 g/ml), very low-density lipoprotein (VLDL, 0.95–1.006 g/ml) and chylomicrons (<0.95 g/ml) (Cox and García-Palmieri, 1990; German *et al.*, 2006). These different classes of lipoproteins all contain triglycerides, cholesterol, cholesterol esters, and phospholipids, in various amounts and proportions. The transport of lipids in the circulatory system takes place through two pathways: the endogenous path to carry the lipids between different tissues, and the exogenous way in which the absorbed dietary lipids are brought from the intestinal epithelial cells to internal target tissues.

Before the excretion into the bloodstream, the plasma very low-density lipoproteins (VLDLs) are synthesized in the liver through complex processes, with sizes ranging from 30 to 70 nm (Griffin, 1999). With respect to size and density, differences of the secreted VLDLs depend on various genetic and external conditions such as the diet. Their prime function in the endogenous pathway of lipids transportation is to carry triglycerides and other lipids to different tissues inside the body. The interaction between VLDL and lipoprotein lipases lead to the hydrolysis of triglycerides with the release of free fatty acids and the formation of Intermediate-density lipoprotein (IDL). Low-density lipoproteins (LDL) are formed through subsequent conversions of IDL by substituting the remaining triglycerides with cholesterol esters and removing all apoproteins except apoprotein-B. The fundamental role of LDL in the blood is to transfer cholesterol to cells with LDL-receptors, via receptor-mediated endocytosis. The reverse translocation of cholesterol from different tissues to the liver for excretion is the responsibility of highly dynamic and composition-diverse particles called high-density lipoproteins (HDLs).

In the exogenous lipid dietary pathway, absorbed short and medium chain fatty acids bind to serum albumin for direct transport to the liver. In contrast, long chain fatty acids form triglycerides inside the intestinal epithelium before their integration with

phospholipids, cholesterol, and other apolipoproteins into large particles called chylomicrons, which enter the bloodstream through lymphatic vessels. The hydrolysis of the triglycerides by lipoprotein lipase liberates free fatty acids at the target site, while the remnant chylomicron particles are removed from the blood by binding to specific receptors and uptake into hepatic cells.

1.5. Lipid Transport in Insects

In insects, the adipose tissue represents the main compartment for lipid storage and metabolism. The amount of stored lipid in the fat body is a crucial element during metamorphosis of the holometabolous insects and has a substantial impact on the fertility of adult insects. Thus, throughout the feeding periods in various stages of the insects' life, the adipose tissues keep collecting lipids to serve as a reservoir of fuel to nourish the tissues with the needed energy during different physiological and behavioral activities. The transformations among the various insect stages require the insects to reach a particular size with a minimal amount of stored nutrients. For instance, the pupal stage of holometabolous insects is a nonfeeding phase between the larvae and adult stage, during which the larval components break down, and the adult structures are build up. This transition requires prior storage of a lot of nutrient reserves including lipids and proteins, to serve as building blocks and energy source. In other insects species, non-feeding adults rely on the stored nutrient reserves to sustain their life. Reproduction and egg maturation in adult females utilize significant amounts of these reserves as well. Also, extended flight in many adult insects involves the mobilization of triglycerides from the fat body and the subsequent translocation into muscle tissues. In addition to these well-known roles, the fat body also plays various roles as an endocrine organ in the insect's immune system (Choi and Hyun, 2012; Azeez *et al.*, 2014). Hence, it is not surprising of the fat body to be the site of synthesis for most of the hemolymph proteins and other storage proteins in insects.

In insects, lipoproteins play various functional roles throughout the different developmental stages. The hemolymph lipoproteins create an efficient delivery system to relocate lipids among the various tissues. Intracellularly, members of the cytosolic ILBP family are responsible for the cytoplasmic transport inside the cells. Furthermore, different insect lipoproteins belong to another protein family with a storage function. Nutritious amino acids stored inside the core of these lipoproteins simultaneously with lipids and

other important attached molecules represent a pool of building blocks that are needed for future development. Most of these proteins are synthesized in the fat body during different stages of the insect's life cycles and secreted into the hemolymph. In the aqueous environment of the hemolymph, they form vehicles to carry lipids inside their hydrophobic cores and deliver them to their utilization sites. Other lipoproteins get incorporated into the target tissues with their lipid content and stored there for later use.

As in vertebrates, lipids in insects are stored in the form of triglycerides in the adipose tissue, but distinct transport mechanisms have evolved. During energy-demanding activities such as flying in a process described in various reviews (Haunerland, 1997; Van der Horst *et al*., 2001; 2002), adipokinetic hormone (AKH) is released from the corpora cardiaca to stimulate the conversion of the TAG into diacylglycerol (DAG) in the fat body. Once released into the hemolymph, the cycle of loading and unloading of the hydrolyzed lipids starts as the DAG binds to the surface of the hemolymph high-density lipophorin (HDLp), creating large and dense particles. To stabilize these molecules, various molecules of a small hemolymph protein, apolipophorin-III (apoLp-III), assemble at the outside, thus minimizing the exposure of the hydrophobic lipids to the aqueous environment, and develop a lower density form of lipophorin known as low-density lipophorin (LDLp). At the muscle tissue, lipophorin lipase hydrolyzes the DAG into free fatty acids, which in turn are taken up into the cytosol of the muscle cells, and free HDLp and apoLp-III to get ready for another cycle. Eventually, the free fatty acids bind to FABP within the muscle cell and get transported to the mitochondria to release energy.

1.5.1. Hemolymph lipophorins

The main hemolymph protein lipophorin was first discovered in insect hemolymph as diacyl-glycerides carrying protein (Chino *et al*., 1969). These lipoproteins seem to exist in all insect species in addition to other arthropods, which indicates that they evolved before the separation of these taxa from a common ancestor (Haunerland and Bowers, 1989). In insects, it consists of three apoproteins, apolipophorins I, II, and III, with approximate molecular weights of 250, 80, and 18 kDa, and ratios 1:1:2, respectively (Kawooya *et al*., 1984; Kawooya *et al.,* 1986). During the resting phase of the adult insect, most of the lipophorin is present in the form of high-density lipoprotein (HDLp) which is composed of apoLp-I and II (Ryan *et al*., 1984; Shapiro *et al*., 1988). The increasing demand of energy during flight is associated with the loading of diglycerides and

apolipophorin-III (apoLp-III) onto the HDLp particle, resulting in the formation of LDLp (Chino *et al.,* 1986; Haunerland *et al*., 1986; Wells *et al*., 1987). At the muscle tissues, apoLp-III and the diglycerides dissociate from the LDLp which in turn restores the high-density form (HDLp) (Kawooya *et al*., 1988). Through cycles of loading and unloading, the apolipoproteins in the insects' hemolymph serve as a shuttle system similar to the mammalian lipoproteins to transport lipid without the internalization of any of these proteins in the origin or the target site (see; Kawooya *et al*., 1988).

1.5.2. Storage lipoproteins

During their postembryonic development, many insect species undergo metamorphosis, with profound changes in physical and physiological characteristics. Hemimetabolous insects, such as locusts (Orthoptera), experience incomplete (or partial) metamorphosis, when adults that are capable of flight and reproduction emerge from wingless juvenile nymphs. In the holometabolous insects which undergo complete metamorphosis, the transformation to the reproductive adult stage passes through two immature forms. The first is the physically active feeding period known as the larval phase. The second is the pupal stadium that represents the junction between the two other phases, in which the insect is solitary, enclosed, and nonfeeding. During the pupal stadium, many of the larval components are broken down, and the adult tissues are built up. To accomplish this, larvae synthesize large amounts of specific proteins for the purpose of storing nutrients needed during the pupal stage; these proteins are gathered and stored inside the fat body as dense protein granules prior to pupation (reviewed by Levenbook, 1985; Haunerland, 1996; Burmester and Scheller, 1999). The amino acids of these storage proteins form the precursor of the adult proteins as they get hydrolyzed during the pupal stage and used to synthesize the adult proteins.

In general, insect storage proteins are large proteins of ~ 500 kDa that are glycosylated and contain varying amounts of lipids. Many belong to the gene family of the hexamerins that encode hexameric proteins of ~72-83 kDa subunits (Kanost *et al*., 1990), such as arylphorin, which appears to be expressed on most insect species. Arylphorins are proteins rich in aromatic amino acids that are glycosylated with ~2-4% of their total weight being oligosaccharide moieties (Haunerland and Bowers, 1986b). Their lipid content can count for 1-2% of the total mass of these proteins (Levenbook, 1985; Kanost *et al*., 1990).

Some holometabolous insects contain another class of storage proteins that contain larger amounts of lipids. First discovered in the hemolymph of the corn earworm, *Helicoverpa zea* (Haunerland and Bowers, 1986b), these proteins are very high-density lipoproteins (VHDL). Unlike the arylphorin, VHDL from *H. zea* is a tetramer of four identical subunits (~150 kDa each) with a lipid content of ~8-10% and a density of 1.26 g/ml. In *H. zea*, VHDL is first detected in the hemolymph of 2nd larval instar (Haunerland and Bowers, 1986a), and its concentration increases until the last larval instar. VHDL is a larval specific protein, and it is not expressed in the adult stage of the corn earworm (Sum and Haunerland, 2007). Just like the yolk protein precursor vitellogenin, VHDL belongs to the large lipid transfer protein (LLTP) gene family, and it appears to fulfill a similar function during metamorphosis as vitellogenin during reproduction and embryonic development.

A particular form of lipoproteins not present in mammals is the yolk protein precursor vitellogenin (see Robinson, 2008). This protein contains a smaller proportion of lipids (<10%), resulting in very high-density lipoprotein (VHDL, density >1.20 g/ml). Vitellogenin is a highly conserved protein present in the circulatory system of all oviparous species, including birds, fish, amphibians, as well as invertebrates. Following its uptake into oocytes, the protein serves as a storage form for energy and biosynthetic precursors needed for the development of the embryo.

1.5.3. Vitellogenin (VG)

Vitellogenin is the precursor of the yolk protein in most oviparous animals. Similar to other lipoproteins in insects, the fat body is the central compartment in the production of VGs before the secretion into the hemolymph. Subsequently, VGs are captured by specific receptors on the membrane of the developing oocytes (Roehrkasten *et al*., 1989; Hafer *et al*., 1992; Schneider, 1996; Richard *et al*., 2001), and internalized through receptor-mediated endocytosis (Telfer *et al*., 1982). After the uptake of these proteins, the oocytes send the cytoplasmic coated vesicles that contain the VG proteins to specialized lysosomes (known as yolk bodies) to be processed for storage for later use during embryogenesis.

VGs are large lipoglycoproteins (~200-700 kDa) that are composed of one to four subunits. The VGs are often proteolytically cleaved and undergo co- and posttranslational modifications. VGs carry lipids, carbohydrates, and other nutrients to the oocytes

(reviewed in Tufail *et al*., 2014). In the oocytes, VGs are stored as vitellin (Vn), which is the precursor reservoir of the developing embryo.

1.6. Endocrine regulation in insects

The transformations that happen during metamorphosis are associated with changes in the expression of lipid binding proteins and related to their specific roles throughout the insects' life cycles. The expression profile of these proteins is not well understood at the molecular level, but it may involve the direct action of the endocrine system.

Juvenile hormone (JH) and 20-hydroxyecdysone are two major hormones in insects involved in development and reproduction. The cross-talk between these two hormones is responsible for the transition between different developmental stages, and the initiation of metamorphosis.

1.6.1. Ecdysone hormone

Ecdysone controls metamorphic and molting events in the insect life cycle. It is a steroid hormone that plays various essential functions such as regulating metamorphosis, the transitions between different insect phases, and oogenesis. Ecdysone is synthesized in the prothoracic gland of the immature insects (Terashima *et al*., 2005). The prothoracic gland is activated to release ecdysone by prothoracicotropic hormone (PTTH), which is produced by the corpora allata (Gilbert and Goodman, 1981). In some adult insects where the prothoracic gland is no longer present, the egg chambers produce ecdysone. The conversion of ecdysone into its active form 20-hydroxyecdysone (20E) takes place in the fat body (Chapman, 1998).

The molecular mechanism of the ecdysone is well understood, as the active form 20E binds to a nuclear receptor (ecdysone receptor, EcR), which in turn forms a heterodimer with another nuclear receptor, (ultraspiracle, USP) (Koelle *et al.,* 1991; Thomas *et al*., 1993). This heterodimer complex influences the expression of target genes through a coordinated hormonal response by binding to specific regulatory sequences in the promoter region of these genes (Uryu *et al*., 2015). In adult mosquitoes, it was found that the ovarian ecdysteroid regulates the expression of vitellogenin in the fat body (Belles

and Piulachs, 2015). Furthermore, it was concluded that in *Drosophila,* the ecdysteroid concentration during the egg developmental stages is crucial especially around stage 8, which is a critical checkpoint when each egg chamber decides if it will continue the growth or it will go under apoptosis (reviewed by Uryu *et al*., 2015).

1.6.2. Juvenile hormone (JH)

Since the first discovery of the juvenile hormone in 1934 by Wigglesworth, its molecular mechanism has been challenging to be fully understood, due to its unique chemical and physical characteristics and the many different roles it plays in various aspects during insects life cycles, in metamorphosis, reproduction, adult diapause, and different behavioral activities. It is a hormonal secretion of the corpora allata (CA), a pair of glands attached to the base of the insect's brain. Wigglesworth defined it as a status quo hormone, or juvenile hormone, as it was found to prevent metamorphosis in last instar nymphs of the blood-sucking bug, *Rhodnius prolixus*. Subsequently, JH was shown to be necessary for egg production and maturation in adult females (Wigglesworth, 1936).

The uniqueness of JH comes from its chemical structure. JH is found in at least six different forms (JH 0, I, II, III, 4-methyl JH I, and the bis-epoxide of JH; JH B3). While JH III seems to be the most dominant form in insects, the first three JH forms (JH 0, I, and II) have only been identified in lepidopteran insect species. The primary function of JH, i.e., to prevent metamorphosis during the immature stages, has been exploited extensively in the field of pest control. For decades after its discovery, many natural analogs (juvenoids) and synthetic JH analogs have appeared in the insecticide market as potent natural pesticides. However, it is not clear if the natural JHs or their synthetic analogs have the same mode of action and exhibit the same functions (Steiner *et al*., 1999).

The JHs share some general characteristics with the steroid hormones, such as lipid solubility, biosynthetic pathways, and the specificity induced by the side chains. The hydrophobic steroid hormones exert their actions through direct binding with intracellular cytoplasmic or nuclear receptors, which in turn acts as a transcriptional regulator that either up- or down-regulate the expression of specific genes. There is much evidence that JH uses similar mechanisms of action. In addition, it can act through the binding to specific membrane receptors, as in its role in vitellogenin transport into the developing oocytes in a process called "patency." In this instance, JH acts on the follicle cells of the ovary

causing them to shrink, allowing for big spaces in between the follicle cells which permit vitellogenin to enter the ovary to access the oocytes and eventually to enter these cells (Davey, 1996).

Understanding the molecular mechanism of JH, the crosstalk with the other hormones, and the search for the potential nuclear receptors have been long and challenging. With the advent of advanced molecular techniques and large amounts of genome data from insects, it is now worthwhile to revisit some of the older work, in order to investigate the involvement of candidate nuclear receptors in the regulation of processes under the control of JH.

1.7. Nuclear receptors as transcription factors

Nuclear receptors are mediator proteins that reside inside the eukaryotic cells. Upon binding to specific ligands, these nuclear receptors mediate the expression of particular genes vis interacting with the regulatory response elements of these genes (Owen and Zelent, 2000). Nuclear receptor for which the ligand has not been identified, or which act in a ligand-independent manner, is called an orphan receptor (Giguere, 1999).

The protein structure of nuclear receptors is highly conserved throughout the animal kingdom, as the members of the nuclear receptor superfamily are all composed of five major domains (N-terminus (A/B-C-D-E-F) C-terminus) (Brelivet *et al*., 2004). At the N-terminal end, the transactivation domain AF-1 (A/B) is responsible for the transcription activation. Next, C is the DNA binding domain (DBD) which is the most conserved region in the nuclear receptor family, with two conserved zinc finger motifs. The D region is the hinge that connects the DBD to the following area E which usually contains two domains: the ligand binding domain (LBD) and a second transcription activation domain (AF-2). At the C terminus, the F region with a coactivator binding role is the least conserved region, and even some nuclear receptors lack this domain.

1.8. The desert locust, *Schistocerca gregaria*

In our study, the desert locust, *Schistocerca gregaria* (Figure 1.1) was used as the model organism to examine the expression profile of some lipoproteins, and the involvement of the main candidate nuclear receptors in the regulation of the expression of

these genes. Locusts are hemimetabolous insects; i.e., their life cycle starts with a small nymphal wingless stage that looks similar to the adults, and after a few moltings, it transforms into the fertile adult that has wings and is capable of flying. Locusts belong to the order Orthoptera which also includes grasshoppers and crickets.

Figure 1-2 The nymph and the adult of the desert locust, *Schistocerca gregaria*

The desert locust can be found in two different forms, a solitary and gregarious form. The initiation of the transition from solitary to gregarious appears to take place in a short time – perhaps as short as thirty minutes – after the exposure of young nymphs to other locusts (Geva *et al*., 2010). This phenotypic plasticity in locusts has a great impact on their lives and involves morphological, physiological, and behavioral differences between the two forms (see Guershon and Ayali, 2012 and the references therein). While the solitarious insects favor to be inactive and avoid contact with other members, the changes to the gregarious form create a more active conspecific species that tend to cluster in swarms with other individuals in huge numbers, causing these insects to become potentially ravaging pests. In our research, insects are always kept in under the gregarious form in crowded cages.

In nature, as it was reported by Symmons and Cressman, 2001, the adult females in the gregarious phase lay around 140 eggs/female in pods in wet soil. The incubation period of the egg stage is primarily based on the temperature and moisture of the sand and takes around 10-65 days, followed by five nymphal instars called hoppers with an

average 35 days before the transformation into adults. The maturation of the adult stage differs according to the environmental conditions such as temperature, moisture, and the food kind and availability. The adult female can start laying eggs in around three weeks after the last molt if it has access to fully grown foliages (Schoonhoven *et al*., 2005).

Desert locust distribute over vast desert areas that extend from North Africa to southwest Asia, and in South America. Despite the existence of early warning systems in these countries, sometimes the upsurge of this pest is considered economically dangerous, as it becomes difficult to contain their proliferation, as it happened, for example, in the upsurge in West Africa between 2003-2005 (Ceccato *et al*., 2007). Different insect control agents may be needed to control this insect in various pest management systems, both in their breeding grounds and during the adult migration. A better understanding of the physiological mechanisms that stand behind the ability to reproduce and to migrate vast distances might be helpful to find novel targets for the control of these insects.

1.9. Dissertation objectives

Using the desert locust as a model system, the goal of this dissertation was to explore the function and regulation of various lipid transport proteins. Specifically, the following research projects were carried out:

- Obtain VG sequences of the desert locust. As vitellogenins have not been cloned in this species and their sequences are not known, a pre-requisite for this work was the cloning and sequencing of VG cDNA, and study their temporal expressions through the adult female lifetime.

- Investigate the role of the two candidate nuclear receptors, Methoprene-tolerant (Met) and retinoid X receptor (RXR) in the regulation of the VG genes.

- Carry out functional studies through knockdown experiments on the role of FABP in the intracellular lipid transport during flight.

- Obtain the sequence of apolp-III to study its temporal expression, and investigate the effects of the protein on the flight capability.

Connecting statement 1: Structure and expression of vitellogenin

In the introduction, various roles of different lipoproteins in transporting and storing lipids in insects and mammals have been discussed. One of these lipoproteins is the yolk-precursor protein, vitellogenin (VG). This protein is mainly synthesized in the fat body of adult females before its secretion into the hemolymph and deposition of its protein and lipid resources into the developing oocytes. It is well established in many insect species that vitellogenin synthesis is under hormonal control, and directly linked to fecundity and reproductive success. In the migratory locust, two vitellogenin molecules were discovered more than 30 years ago, and several studies have focussed on the role of juvenile hormone in their expression. However, at the onset of this thesis, the sequence of vitellogenins and their expression patterns had not been elucidated for any locust species. The next chapter describes the primary structure of the two vitellogenins expressed in the desert locust, as well as their expression pattern in the fat body throughout the adult female lifespan.

.

Chapter 2. Primary structure and temporal expression of two vitellogenins from the desert locust, *Schistocerca gregaria*

Abstract

The large storage protein vitellogenin (VG) is the precursor of the major yolk protein vitellin (Vn). In most insect species, it is produced in the fat body and secreted into the hemolymph prior to selective sequestration and deposition into the ovaries. Previous findings in the closely related species, *Locusta migratoria*, have shown the expression of two different vitellogenins. In the desert locust, the complete cDNA of both vitellogenin A and B were amplified and sequenced using the 5` and 3` Race techniques. They are 5570 bp and 5591 bp in size, with putative translated proteins of ~203 kDa and ~202 kDa; respectively. They share several of the conserved motifs found in vitellogenins from other species, like GL/ICG motif at the C-terminus, which is conserved among most of the known insect vitellogenins. The expression study reveals that VGs are not expressed in nymphs and early adults. VG mRNA is first detected around 11 days after adult ecdysis and reaches its first peak in the next 72 hours; mRNA levels remain high throughout the adult lifespan. In this study, we have elucidated for the first time full-length vitellogenin sequences from orthopteran species. The comparison of the primary structures and expression patterns of VG-A and VG-B enhance the knowledge of the vitellogenin gene family and provide insights into the characteristics and the regulation of these genes on a molecular scale.

2.1. Introduction

Like other oviparous species, insects provide their eggs with all nutrients required for the development of their offspring. Lipid and amino acid resources are stored in the egg by the major yolk protein vitellin (VN). Its precursor vitellogenin (VG) is synthesized in the fat body of mature adult females in a temporally and spatially specific manner and secreted into the hemolymph. Subsequently, the protein is taken up by the developing oocytes via receptor-mediated endocytosis. It was first described by Telfer (1954) in the silkmoth, *Hyalophora cecropia* (L.), and given the name vitellogenin (from Latin: *vitellus* – yolk, and Greek: *genesis* – creation) as the precursor of the egg yolk protein by Pan *et al.* (1969).

In the biosynthetic pathway, the emergent vitellogenins are large monomeric proteins that undergo co- and post-translational modifications and are cleaved into two or more subunits that form the mature oligomeric proteins. These subunits vary widely between different species and range in size from ~50 kDa to ~180 kDa (see reviews: Sappington and Raikhel, 1998, Tufail *et al.*, 2014). Vitellogenins are expressed in the fat body as pre-pro-vitellogenins that are proteolytically cleaved in the endoplasmic reticulum by a subtilisin-like vitellogenin convertase (VC) that recognizes a conserved cleavage site, as it was found in mosquito pro-VG (Chen and Raikhel, 1996). This cleavage and other post-translational modifications like glycosylation, phosphorylation, and sulfation enable the proteins to carry lipids, carbohydrates and other essential elements along with their amino acids so that they can serve as a reservoir for the various nutrients needed by the developing embryos inside their protective egg structure. In addition, some of these alterations aid in the secretion of these proteins from the fat body and in the deposition into the oocytes. Despite the considerable sequence variations between insect species, the amino acids composition of VG proteins has shown a remarkable degree of conservation between different species. Thus, it appears that the selection pressure on these genes reflected the nutritional needs of the developing embryo more than any specific structural requirements (Hughes, 2010).

In most insects, the female fat body is the primary site of VG expression; from there, the protein is released in large amounts into the insect's hemolymph before it is incorporated into the oocytes. However, in various insects, VG has been found to be expressed in a less sex- and tissue-dependent manner. For instances, in most higher

dipteran insects, including *D. melanogaster*, and some coleopteran species, vitellogenin is expressed not just in the fat body but in the ovarian follicle as well (Belles, 1998). In some dipteran insects, the expression takes mainly or even exclusively in the ovarian tissues (reviewed in Valle, 1993). Furthermore, while vitellogenins are generally considered to be specific to fertile adult females, recently these proteins were found in some male insects, namely in the honeybee drones, and also in immature stages such as pupae of the honey bee workers. The expression of these genes appears to start already before adult emergence, though the expression signal is weak but still detectable (Piulachs *et al.*, 2003).

In studies on the regulation of vitellogenesis by juvenile hormone in *Locusta migratoria*, Wyatt and coworkers (Wyatt *et al*.,1986; Locke *et al*.,1987) identified more than 30 years ago two separate mRNAs that encode two non-identical proteins called vitellogenin A and B. Only the approximate size and short amino-terminal sequences were obtained, and even today, little is known about the molecular characteristics of these proteins. In fact, at the onset of our work, no cDNA or genomic sequences for vitellogenin from any locust species had been published. The large size of the locust genome, with an abundance of repetitive sequences and very large introns, would have made cloning of these genes by traditional methods challenging. While the genome of *Locusta migratoria* has recently been sequenced (Wang *et al.*, 2014), vitellogenin sequences were not found in the automatic annotation process and were not identified at the outset of our study.

The size of the genome of the desert locust, *Schistocerca gregaria*, is estimated to be 8.55 Gb, and thus 50% larger than that of the migratory locust, making the sequencing of these genes from genomic sources even more challenging (Camacho *et al*., 2015). However, a transcriptome analysis of the central nervous systems of various developmental stages of the desert locust has been carried out by Badisco *et al*. (2011), and it is likely that the EST library will contain ESTs of proteins expressed in the ubiquitous fat body of the insect.

The current study aims to sequence the complete cDNA of both VG-A and B from the fat body of the desert locust and measure the expression levels in the fat body and the corresponding proteins in the hemolymph. Also, we seek to compare their sequences and determine their phylogenetic relationship to vitellogenins from different species and identify conserved motifs in these sequences.

2.2. Materials and methods

2.2.1. Insects and samples collection

Schistocerca gregaria (Orthoptera, Acrididae) were reared under controlled environmental conditions. The gregarious insects were kept at 30 ºC and 16 h daily illumination in crowded cages. To obtain and track the desired ages, newly emerged adults were collected and transferred to a new cage after marking their thoraxes by means of a permanent marking pen.

Fat body and hemolymph samples were collected from the last nymphal instar and at specific time points between the first day of adult eclosion and two months afterward, for gene expression and protein studies. For hemolymph collection, one hind leg was cut off, and the hemolymph was flushed out into a pre-cooled Eppendorf tube by injection of PBS buffer, pH 7.4, into the body cavity, as previously described by Chino *et al.* (1987). The tubes were centrifuged at 4 ºC for 10 min at 3500xg to remove the hemocytes before storing at -20 ºC freezer. Fat body tissue was dissected from the same insects in RNase-free tubes and used immediately for RNA isolation, or rapidly cooled in liquid nitrogen and stored at -80 ºC for later use.

2.2.2. RNA collection and preparation

Total RNA extraction was carried out using the RNeasy Plus Universal Mini Kit (Qiagen, Cat.# 73404), following the manufacturer's protocol. Fat body was homogenized in 900 µl QIAzol Lysis Reagent, and the homogenate was left at room temperature for 5 min to promote the complete dissociation of nucleic acid/protein complexes. To reduce genomic DNA contamination, 100 µl of gDNA Eliminator Solution was added to the homogenate, and the tube was shaken for 15 s. Chloroform (180 µl) was added, and the tube was shaken again for 15 s, followed by 3 min incubation at room temperature. The sample was spun at 16000xg for 15 min at 4 ºC. The upper aqueous phase was transferred to a new tube, mixed with an equal volume of chilled 70% ethanol, loaded in aliquots onto an RNeasy Mini spin column and centrifuged for 15 s at 9500xg at room temperature. Subsequently, the RNeasy column was rinsed with RWT buffer and twice with RPE buffer. The elution step was done in a new collection tube with 100 µl DEPC-

treated water, and total RNA concentration and quality were determined on a Nanodrop 2,000C UV-Vis spectrophotometer (Thermo Scientific).

2.2.3. cDNA synthesis and PCR reactions

The first complementary DNA strands were generated using the iScript cDNA Synthesis kit (BioRad, Cat#: 170-8890), following the manufacturer's protocol. Total RNA (1 µg) was mixed with 4 µl of 5X iScript reaction mix, and 1 µl of iScript reverse transcriptase and DEPC treated water was added to reach a total reaction volume of 20 µl. The mixture was incubated in a thermocycler for 5 min at 25ºC, followed by 30 min at 42 ºC and 5 min at 85 ºC. The resulting cDNA was diluted ten-fold with DEPC-treated water and stored at -20 ºC.

The amplification of VG was carried out using the Fermentas 2X PCR master mix (Cat#: K0172). The PCR products were analyzed on 1.0% agarose gel, purified with QIAquick® Extraction Kit from QIAGEN (CAT#:28704). The purified DNA was then sent for sequencing. The PCR was carried out at 98 ºC for 3 min, 40 cycles of 98 ºC for 30 s, the annealing temperature which depended on the designed primers for 30 s, 72 ºC for 1 min, and eventually with an elongation step for 15 min at 72 ºC. Gene-specific primers were designed from available sequences on databases and the newly sequenced VGs partial sequences, shown in appendix A (Table A1 & A2).

2.2.4. 3` and 5` RACE library construction

To obtain the complete cDNA for both VGs that contains the 3` and 5` untranslated regions (UTR), RACE libraries were constructed using the SMARTer™ RACE cDNA Amplification kit from Clontech (Cat#: 634860). The first-strand cDNAs to create the 5`and 3`RACE libraries were synthesized as described in the manufacturer's protocol as follows: For the preparation of 5` RACE cDNA, 1 µg of total RNA was mixed with 1 µl of 5`-CDS Primer A (12 µM) with the sequence (5`-$(T)_{25}$VN-3`), where N is any nucleotide while V is A, G or C, and sterile H_2O was added to reach 11 µl total volume. For the 3`Race library, 1 µg of total RNA was mixed with 3`-CDS (12 µM) Primer A, (5`-AAGCAGTGGTATCAACGCAGAGTAC$(T)_{30}$VN-3`) and sterile H_2O in a total volume of 12 µl. The contents were mixed and incubated in a thermocycler for 3 min at 72 ºC, followed by a cooling step at 42 ºC for 2 min then the tubes were spun for 10 s at maximum speed.

The SMARTER II A Oligo (1 µl) with a sequence (5`-AAGCAGTGGTATCAACGCAGAGTACXXXXX-3`), where X is an unknown nucleotide, was added to the 5`-RACE reaction. A master mix was prepared of: 5.5 µl buffer mix which was prepared of the following reagents (4 µl 5x first strand buffer, 0.5 µl DTT (100 mM) and 1 µl dNTPs), 0.5 µl RNase inhibitors (40 U/µl) and 2 µl SMARTScribe Reverse transcriptase (100U). The master mix was added to the 3` and 5`-RACE reactions to a total volume of 20 µl. Each reaction was incubated at 42 ºC for 90 min followed by 10 min at 70 ºC. Finally, the reactions were diluted with 90 µl of Tricine-EDTA buffer and stored at -20 ºC.

An optimized PCR kit SeqAmpTM DNA polymerase from Clontech (Cat#: 638504) was used to amplify the 3` and 5` ends of the VG transcripts. Different specific internal gene-specific primers were used along with Universal primer mix (UPM) to amplify and sequence the 3`and 5` ends which contain the untranslated regions (UTR), transcription start site (TSS) and the termination signals.

2.2.5. Real-time qPCR

Gene-specific qPCR primers were designed from the obtained sequences of VG-A and VG-B along with primers targeting GAPDH (Genbank accession # HQ851387.1) to be used as internal control genes (Table 2-1). In order to determine the efficiency of each pair of primers, a regular PCR was performed as described above. The PCR products were excised and eluted from the gel; then the DNA concentration was measured with the NanoDrop spectrophotometer. A series of 10-fold dilutions were prepared using sterile H_2O, to create at least seven dilutions ranging from 500 pg to 0.5 fg. From each dilution, 3 replicates were prepared by mixing 2 µl of DNA with 5 µmoles of the corresponding primers and 5 µl of BioRad's SsoFast^{TM} Evagreen Supermix (Cat#: 172-5200) in three separate tubes. Sterile nuclease-free water was added to bring the total reaction volume to 10 µl in each tube. A BioRad's MJ MiniTM thermocycler (Cat#: PTC-1148C) was used to run the qPCR program with the following parameters: 30 s initiation at 95 ºC, 45 cycles of 5 s at 95 ºC, 30 s at 52 ºC. Subsequently, a melting curve was generated by slowly increasing the temperature in increments of 0.5 ºC from 65 ºC to 95 ºC, followed by a plate read. The BioRad's CFX ManagerTM Software (Cat#: 184-5003) automatically generated the standard curve and calculated the corresponding efficiency. The efficiency was considered 100% for the different primers as they ranged between 90-100%.

To determine the expression level of each selected gene in each experimental subject, three PCR replicates with 10 ng of cDNA template were done for each sample for the target gene and the reference gene. The results of the temporal expression level of both VGs in the fat body were presented as ΔCt by normalizing the expression of the gene of interest to the internal reference gene for each sample, following the equation:

$$\Delta Ct = \text{average Ct of the gene of interest} - \text{average Ct of reference} \quad \text{(for each sample)}$$

Table 2-1 Specific real-time primers used to measure the expression of VG genes

Primers	Sequence 5`--------3`	Efficiency	Annealing temperature
VG-A forward	CGCTGAACATCACTGCAATAC	98.2%	52 ºC
VG-A reverse	CACCCAGCAAGTCTTCATAGT		
VG-B forward	ACTGCTCTTCTGGGATTTGG	99.5%	52 ºC
VG-B reverse	GGACTGTCTAACTTGCCGTATAA		
GAPDH forward	CCCAATGTATGTTGTTGGTGTAA	92%	52 ºC
GAPDH reverse	GTGCCAGGCAATTTGTAGTG		

2.2.6. SDS PAGE gel-electrophoresis

Analysis of hemolymph proteins was carried out by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were separated by 8% separating gel and 5% stacking gel. The electrophoresis was timed for 1 h at 200 mA and 130 V in a mini-slab electrophoresis unit. A total protein concentration of 50 mg of each sample was mixed with the double volume of the Laemmli Sample Buffer (BioRad, Cat#: 161-0737) and boiled for 90 s before loading on the gel. The gel was stained overnight with EZBlueTM staining reagent from Sigma-Aldrich (Cat#: G1041). Finally, it was washed with and kept in dist. water for destaining. The bands of the SDS-gel was visualized and quantified by a gel imager.

2.3. Results

2.3.1. cDNA sequences of VG-A and VG-B

The first degenerate primers (Appendix A - Table A1) used to amplify and sequence partial cDNA of VGs in *Schistocerca gregaria* were designed based on the available sequence using the Blast result of the known partial VG sequence from the Lubber grasshopper; *Romalea microptera* (Borst *et al*., 2000; GenBank: AF135420.1) against a *Schistocerca gregaria* EST database. The blast results revealed one partial sequence with accession number LC01027B2G10.f1. PCR reactions with these primers produced two strong bands ~ 300 and 600 bps (Figure 2-1). The PCR products were excised, eluted from the gel and sent for sequencing. Both bands gave together ~500 bps sequence which by blasting against other VG sequences was confirmed to be a member of the vitellogenin family.

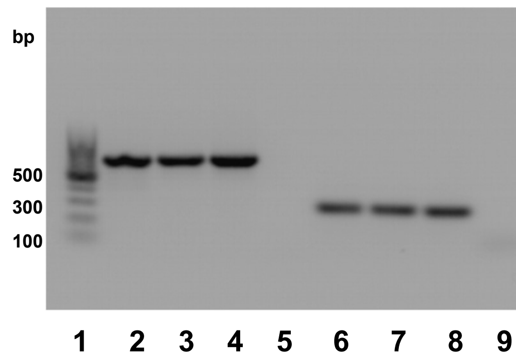


Figure 2-1 Vitellogenin mRNA detection in fat body of adult females.
Using primer sets designed from EST sequence (Accession # LC01027B2G10.f1). PCR template: cDNA from two weeks locust adult female. Lane 1: DNA ladder, Lane 2-4: PCR product ~ 600bp, Lane 6-8: PCR product ~ 300 bp, Lane 5,9: no template control.

Gene-specific primers obtained from this sequence were used to amplify the entire cDNA of VG-A using the 3` and 5` RACE techniques in multiple and overlapping sequencing runs. The 5' end of the mRNA is almost identical to the exon in the VG-A clone identified by Locke *et al.* (1987) (55 out of 56 nucleotides), and hence it was concluded that the cDNA encodes VG-A. The resulting sequence contains an open reading frame of 5455 nucleotides with a conceptually translated protein of 1806 amino acids (GenBank accession # MK206969 - Appendix A; sequence A1). The calculated molecular weight for VG-A is ~203 kDa, and the estimated isoelectric point is 6.73.

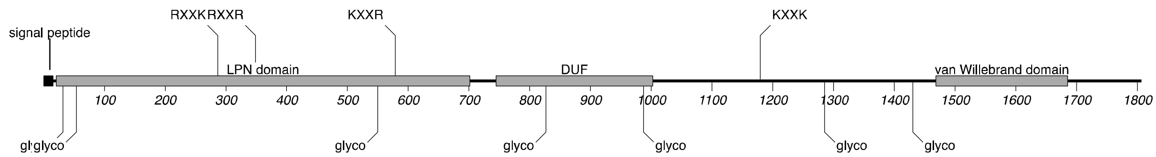
We could not identify any partial nucleotide fragments from the other vitellogenin, VG-B, until 2014 when the whole genomic sequence of *Locusta migratoria*, the migratory locust, was published (Wang *et al.*, 2014). The migratory locust and the desert locust are evolutionary very close, and they share a very high degree of sequence similarity in many other genes, as they belong to the same family (Acrididae). While automatic gene annotation did not identify any vitellogenin gene in the locust genome, we were able to detect the putative introns and exons for both VG genes by computer analysis with Genscan intron/exon predictions (http://genes.mit.edu/GENSCAN.html). Based on the assembled cDNA sequence of the putative *Locusta migratoria* VG-B, degenerate primers were designed to obtain partial overlapping sequences for *Schistocerca gregaria* VG-B. The 5' and 3' sequences were revealed with 5` and 3` RACE methods, using universal primers that anneal to the 3` and 5` RACE Libraries and internal specific primers from the obtained sequences. The cDNA of *Schistocerca gregaria* VG-B is 5591 bp in length, with an open reading frame of 5439 bp. This open reading frame translates to a protein with 1812 amino acid residues, a molecular weight of ~ 202 kDa, and an isoelectric point of 7.28 (GenBank accession # MK206970 – Appendix A; sequence A2).

2.3.2. Nucleotide and amino acid sequence analysis:

The analysis of the VG-A and VG-B cDNA sequences showed an untranslated 5` region of 33 and 35 bp, respectively, preceding the start codon ATG. The following 45 bp (VG-A) or 54 bp (VG-B) encode the N-terminal 15, or 18, amino acid residues that match the criteria for classical signal peptides of exported proteins (predicted by SignalP 4.1 Server, Petersen *et al.*, 2011). VG-A cDNA has the stop codon TGA followed by 87 bp that contains the polyadenylation signal AATAAA before its poly A tail, while in VG-B cDNA the same polyadenylation signal is located 67 bp after a TAA stop codon.

The deduced amino acid sequences were scanned with the PROSITE database (http://prosite.expasy.org/prosite.html) to identify potential conserved domains. Two domains conserved in all vitellogenins were identified in both sequences, the vitellogenin N–terminal domain (residues 22-770, VG-A; 23-790, VG-B) and a van Willebrand domain (residues 1469 – 1685, VG-A; 1498 – 1717, VG-B) (Figure 2-2).

VG-A



VG-B

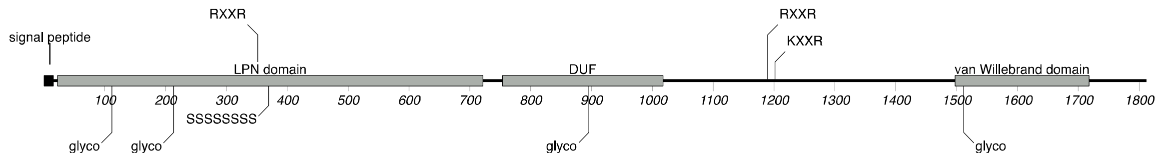


Figure 2-2 Conserved sequence motifs and domains in vitellogenin A and B

Both protein sequences contain several potential cleavage sites for di-basic protein convertases motifs which act on consensus motifs RXR/KR. One of these motifs, which is located at identical positions in both VG-A (349-352, RYRR) and VG-B (351-354, RPRR) has the consensus sequence for subtilisin-like convertases that has frequently found in insect vitellogenins which undergo post-translational proteolytic cleavage (Sappington and Raikhel, 1998).

Other motifs commonly found in vitellogenin proteins include a highly conserved GL/ICG site at the C-terminus in both VG-A (1613) and VG-B (1645), with a DGXR site close by (1595 for VG-A; 1616 VG-B). Noticeably, the poly-cysteine sequence which follows the GL/ICG motif in other insects is absent. Also, a poly-serine sequence that is common to many vitellogenins is found only in VG-B (369-376).

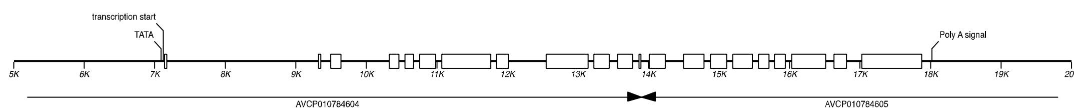
Potential glycosylation sites were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). There are 7 potential glycosylation sites in VG-A, and 4 in VG-B, as shown in Figure 2-2.

2.3.3. Sequence similarity and phylogenetic analysis:

With the knowledge of the full-length cDNA sequences for VG-A and VG-B from *S. gregaria*, we were able to search for the homologous sequences in the genome of *L. migratoria* (Appendix A; sequences A3 & A4). Recently, additional sources of sequence information about *L. migratoria* have become publicly available through the i5k workspace

at the USDA National Agricultural Library including both genome and transcript data. Through BLAST searches of *Schistocerca gregaria* VG-A and VG-B against these databases, we were able to assemble the likely gene structure for the homologous proteins in *L. migratoria* (Fig. 2-3). The gene for *L. migratoria* VG-A contains 21 exons, spanning over a 10.9 kb sequence found on assembly contigs AVCP010784604 and AVCP010784605; the modeled mRNA (gnl|Locusta_migratoria_JAMg-OGSv1_mRNA|JAMg_model_1297.1), while missing more than 400 bp in the center, codes for more 90% of the protein sequence. The 21 exons of the VG-B gene are spread over 11.4 kb on assembly contigs AVCP010784614, AVCP010784615, AVCP010784616, and AVCP010784617, and most of its cDNA is modeled in gnl|Locusta_migratoria_JAMg-OGSv1_mRNA|JAMg_model_1298.1.

VG-A



VG-B

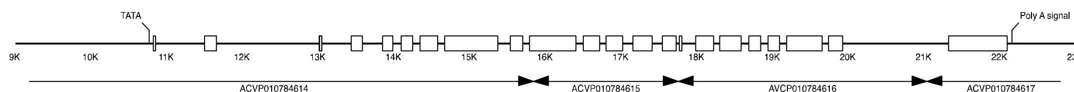


Figure 2-3 Predicted gene structure for VG-A and VG-B from *Locusta migratoria*

Translation of the resulting cDNA sequences revealed a large degree of sequence similarity between the respective proteins from *L. migratoria* and *S. gregaria* (VG-A: 70% identity, 83% similarity; VG-B: 70% identity, 83% similarity).

BLAST searches showed considerable sequence similarity with other insect vitellogenins and vitellogenin-like proteins. However, the alignment of the amino acid sequences with vitellogenins from each of the insect orders for which sequences have been reported showed, in general, relatively low similarity between them (Table 2-1). The number of VG transcripts identified varies between different insect species; most species seem to possess only one, but two or three genes have been found in some species (Tufail and Takeda, 2008). Although some researchers suggest other VG's roles than the nutritional role during embryogenesis, It is still not clear why some insects possess more

than one VG; perhaps VG not only provides nutrients during embryogenesis, but fulfill other, yet undefined roles as well (See Azevedo *et al*., 2016 and references therein). The topology of the constructed phylogenetic tree (Figure 2-4) is mostly consistent with the known insect systematic classification. The phylogenetic analysis shows that the VG genes might have evolved through multiple gene duplications at different times before and after the divergence of the various orders. The three VGs found in the stink bug seem to have evolved just in Hemiptera, but the two VGs in locusts and cockroaches may have existed prior to the origin of these orders, as these paralogs are not much more similar to each other than to VGs from other insect orders.

sequence identity [%]

	Tm VG	Pa VG	Ps VG3	Ps VG1	Ps VG2	Ha VG	Sg VG-B	Lm VG-B	Sg VG-A	Lm VG-A	Pa VG-2	Am VG	Cm VG
Tm VG	**100.0**	23.5	25.6	25.0	26.2	22.2	23.0	22.2	24.6	23.2	26.3	26.4	27.3
Pa VG	42.1	**100.0**	25.5	27.0	25.6	20.7	22.8	22.7	23.1	22.7	29.7	22.9	24.4
Ps VG3	45.3	43.9	**100.0**	44.0	43.7	21.9	23.3	23.3	22.6	22.3	26.1	24.5	26.1
Ps VG1	42.8	44.1	62.9	**100.0**	49.9	21.9	22.5	21.9	21.9	21.5	25.8	24.4	24.3
Ps VG2	44.9	43.2	62.4	67.1	**100.0**	22.1	23.0	23.0	21.9	22.1	25.8	25.2	25.4
Ha VG	41.7	39.8	41.1	40.0	41.0	**100.0**	20.1	19.7	19.7	19.2	20.2	22.2	23.3
Sg VG-B	42.4	40.2	41.9	40.2	41.7	38.7	**100.0**	69.9	26.5	26.2	23.3	22.0	23.3
Lm VG-B	42.0	39.3	41.4	39.4	41.2	37.6	82.7	**100.0**	25.0	25.7	23.7	21.3	22.7
Sg VG-A	41.4	39.2	40.8	37.9	38.3	36.7	43.2	42.3	**100.0**	70.2	24.0	22.4	22.2
Lm VG-A	40.9	39.6	40.9	38.3	39.2	37.4	44.2	43.8	82.8	**100.0**	24.5	22.0	21.6
Pa VG-2	44.3	49.9	44.7	43.2	43.6	38.8	39.7	40.1	40.3	40.6	**100.0**	24.4	25.4
Am VG	43.8	41.2	44.5	42.6	44.4	40.7	41.6	40.1	39.7	39.9	42.8	**100.0**	24.6
Cm VG	45.9	43.6	45.3	42.8	44.2	41.6	42.5	41.6	40.4	40.6	44.8	44.2	**100.0**

sequence similarity [%]

Table 2-2 Sequence identity and similarity between different vitellogenins

Pairwise sequence alignment was carried out with the MUSCLE algorithm (MacVector 16.0.9) for the following vitellogenins: *Tenebrio molitor* VG (GenBank accession AAU20328.2); *Periplaneta americana* VG (BAA86656.1); *Plautia stali* VG-3 (BAA88077.1), VG-1 (BAA88075.1), VG-2 (BAA88076.1); *Helicoverpa armigera* VG (AGL08685); *Schistocerca gregaria* VG-A (Appendix A; sequence A1, MK206969), VG-B (Appendix A; sequence A2, MK206970); *Locusta migratoria* VG-A (this thesis), VG-B (this thesis); *Apis mellifera* (CAD56944); *Periplaneta americana* VG-2 (BAB32673.1), *Apis mellifera* VG (NP_001011578.1); *Clunio marinus* VG *(* CRK93348.1).

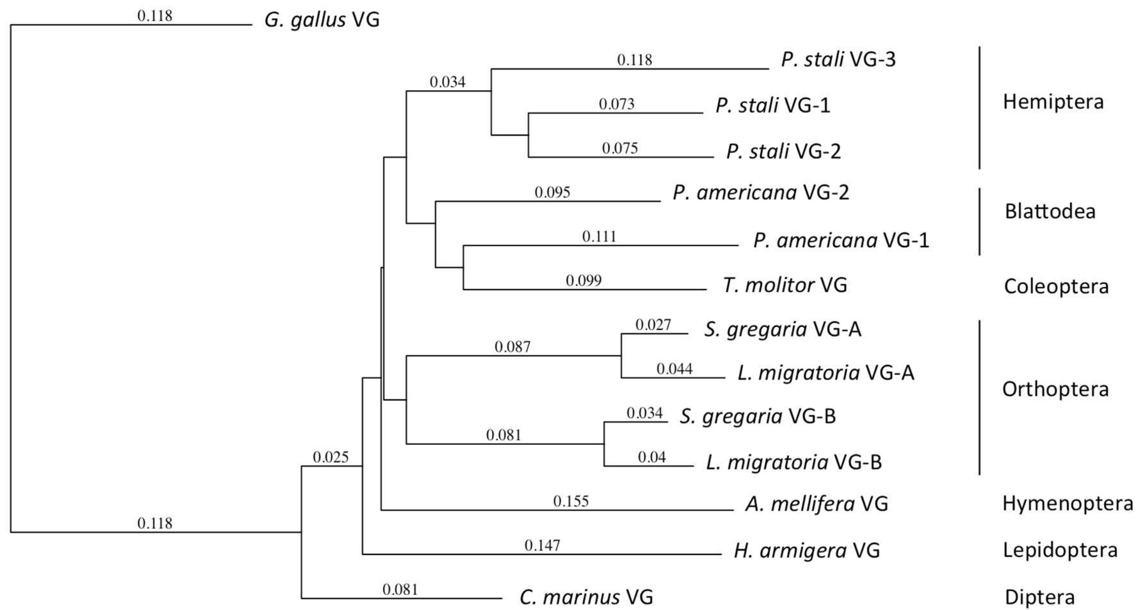
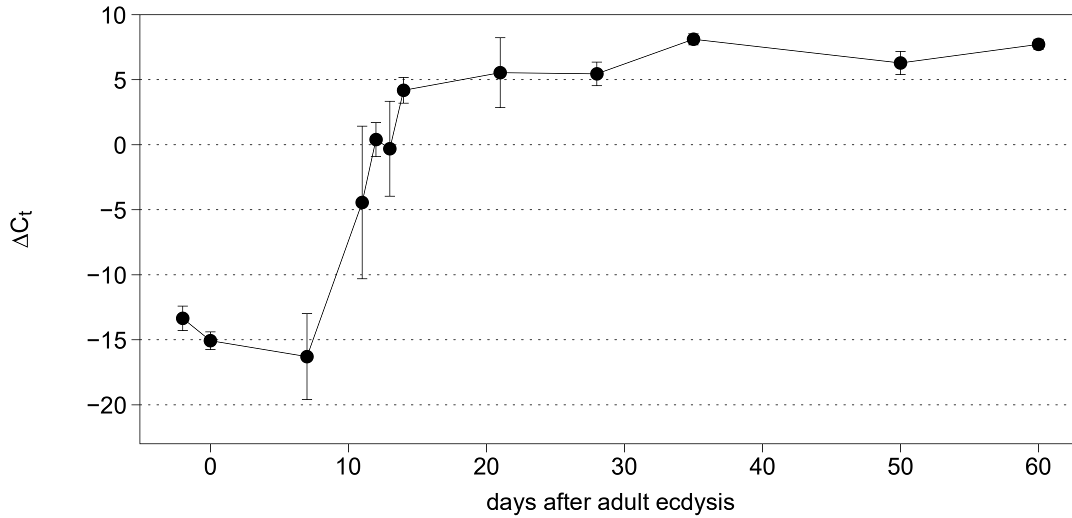


Figure 2-4 Phylogenetic tree of the vitellogenin gene family.
Evolutionary analysis was conducted in MacVector 16.0.9., using the Neighbor-Joining method (Best Tree). Chicken (*Gallus gallus*) VG was used as an outgroup. The best tree is drawn to scale, with branch lengths expressed as the number of amino acid substitutions per site.

2.3.4. Vitellogenin expression in adult females:

The quantitative real-time PCR method (qPCR) was used to measure the expression levels of these genes in fat body of adult female locusts. Specific real-time primers that anneal selectively with the VG-A or VG-B sequence and produce small amplicons that generate one precise amplification curve for each gene were designed using the IDT website tools (https://www.idtdna.com/site). GAPDH was used as an internal control gene to normalize the results. In order to calculate the efficiency of these primers, normal PCR reactions targeting VG-A, VG-B, and GAPDH were performed with a cDNA from fat body of a two-week-old female. For each of these genes, one clear band matching the expected amplicon size was apparent on the gel image. These bands were excised, and the DNA was eluted from the gel and used to prepare a series of dilutions for each amplicon. Standard real-time PCR curves were generated using these DNA dilutions, and the corresponding efficiencies were calculated with the software ($R^2 > 0.99$ for all genes, efficiencies for GDPH 92%, for VG-A and B > 98%). Primers specificity was confirmed by the melting curves of the products. As the amplification efficiencies approaching 100% for all primers, quantification by the ΔCt method was carried out assuming an efficiency of 100% for all targets.

VG-A



VG-B

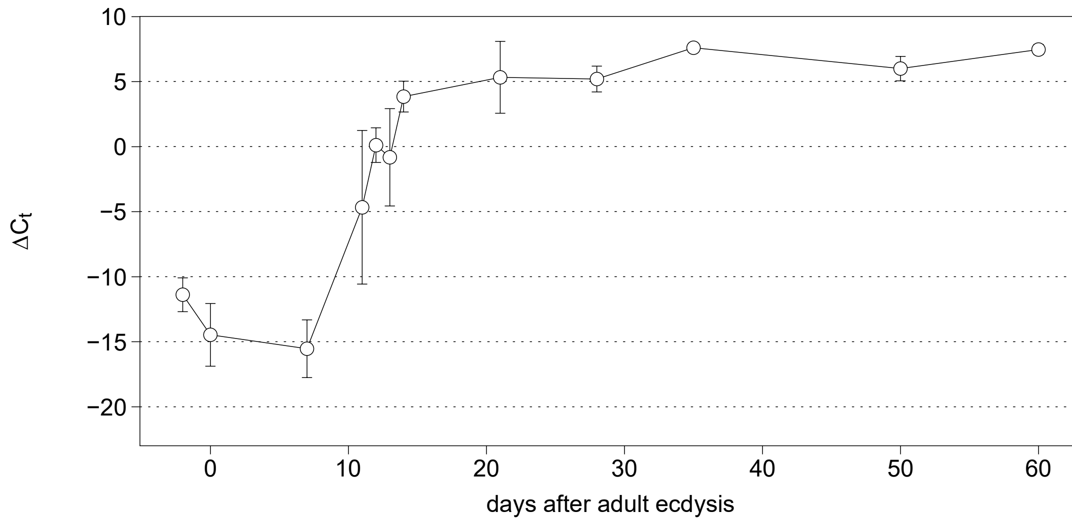


Figure 2-5 Relative temporal expression of VG A and B in the fat body of the desert locust.

Relative expression of VG-A and VG-B was determined by qRT-PCR at specific time-points starting at the last nymphal instar to two months old adult female. The fat body was extracted at each time point and used to extract the total RNA and to prepare the corresponding cDNA. The relative expression level presented as ΔCt value and determined by normalizing the expression of each sample to the internal housekeeping gene GAPDH. Mean ± S.D; n=3-5, Where n is the number of insects used at each age.

The first detectable signal of both VGs was in adult females 11 days after adult eclosion, followed by a gradual increase in the following 3 days. At 14 days, both VGs are strongly expressed, and their transcript levels remain high throughout their life (Figure 2-

5). The expression levels came around the same Ct value (14-16 cycle) in most individual adult females from age 14 days old to two months old for both genes. To exclude the possibility of the transcription of vitellogenins during the nymphal stages, total RNA was extracted from the fat body of the last nymphal instar. No vitellogenin mRNA could be detected at this juncture with the real-time qPCR.

As evident from Fig. 2-5, the ratio of these transcripts does not change significantly during the adult stage. The difference between their ΔCt values was generally less than one cycle, with VG-A normally slightly higher than VG-B.

2.3.5. Vitellogenin protein analysis in the hemolymph

The hemolymph was extracted and analyzed by SDS electrophoresis to reveal the presence of vitellogenin proteins. By comparing the protein bands from females with those from males, several strong bands were visible in the hemolymph of mature females only (Figure 2-6). The first band was ~200 kDa, which resembles the calculated molecular weight of the unprocessed nascent apoproteins. While the band of 200 kDa does not appear in the ovaries, there were two strong bands at around 100-120 kDa, as well as two weaker bands at around 50-60 kDa that appear both in hemolymph of older females and ovarian extracts, but not in nymphs, adult males, or females within the first 12 days after adult eclosion. As the appearance of these protein bands coincides with the timing for the VG mRNA, it is likely that these bands represent components of the mature vitellogenins, after their processing by endoproteases.

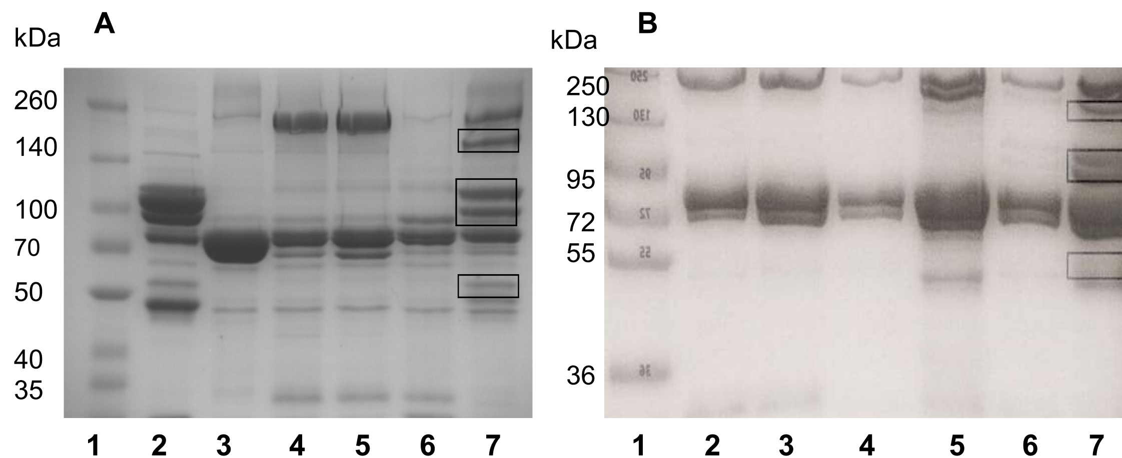


Figure 2-6 Identification of different VG protein bands in the hemolymph of the adult locust female.

SDS polyacrylamide gel electrophoresis separated the hemolymph proteins from the nymph, adult male, and adult females at various time points along with an egg extract. Samples were run on 8% separating gel for 1 hour. **(A)** Lane 1: Protein Ladder; Lane 2: egg extract; Lane 3: nymph hemolymph; Lane 4: one-week adult male hemolymph; Lane 5: one-week adult female hemolymph; Lane 6: 3 weeks old male hemolymph; Lane 7: three weeks old female hemolymph. **(B)** Lane 1: Protein ladder; Lane 2: one-week old female hemolymph; Lane 3: 10 days old female hemolymph; Lane 4: 11 days old female hemolymph; Lane 5: 12 days old female hemolymph; Lane 6: 13 days old female hemolymph; Lane 7: 14 days old female hemolymph. The potential VG protein bands in the hemolymph are shown in rectangles around 200 kDa, 100-120 kDa, 50-60 kDa.

2.4. Discussion

The complete cDNA sequences of the two vitellogenins of the desert locust, *Schistocerca gregaria,* have been determined using 3' and 5' RACE methods. The two transcripts were named VG-A and VG-B after their initial description in the migratory locust*, Locusta migratoria* (Wyatt *et al.*, 1986). The latter study had identified abundant vitellogenin mRNA of ~6500 bp in juvenile hormone-treated females and obtained from genomic libraries two clones of ~500 bp that coded for two distinct vitellogenin apoproteins. These authors described the native vitellogenin as a ~550 kDa lipoglycoprotein, expressed in the fat body as two large glycosylated subunits of ~220 kDa (~185 kDa in the presence of tunicamycin, which prevents glycosylation), and later cleaved into smaller fragments between 60 and 120 kDa.

While the primary structure of numerous insect vitellogenins has been described since then, it is surprising that until recently no additional sequence data on locust vitellogenins have been obtained. This may be a consequence of the very large size of the locusts genome, and the high degree of sequence variability between vitellogenins

from different insect orders. Our initial attempts to elucidate sequence data for desert locust VG similar to the short sequences obtained by Locke *et al.* (1987) and to other closely related species, namely the lubber grasshopper; *Romalea microptera,* relied on EST data from the central nervous system of *S. gregaria* at various developmental stages, assuming that contaminating fat body tissue of mature females would contribute ESTs for the strongly expressed vitellogenins. Indeed, starting with conserved sequences from these sources, we were able to assemble the full-length cDNA for vitellogenin A, but not vitellogenin B. The publication of the first draft of the complete genome of the migratory locust, *Locusta migratoria*, in 2014, provided another means for obtaining partial sequences of potential vitellogenins that could be used for RACE primer design. We were able to find two sequences with similarity to vitellogenin, of 3 kb and 5.5 kb, which helped us to obtain the full-length transcripts for VG-A and VG-B of the desert locust.

The automatic annotation of the locust genome did not identify vitellogenin, but Song *et al.*, (2013) identified partial sequence for the 3' end of vitellogenin A cDNA, coding for the amino-terminal part of VG-A (GenBank accession KF171066). While the conceptually translated protein misses the first 465 amino acid residues (that include the original sequences found by Locke *et al.*,1987), it is obvious that this protein is vitellogenin A. CLUSTAL W alignment between this sequence and the *S. gregaria* VG-A sequence reveals 69.3% sequence identity, and an additional 13.3% conservative substitutions, over the 1346 overlapping residues, confirming the high degree of sequence similarity between the two species. The sequence of vitellogenin B from *L. migratoria* was not identified by these authors, but very recently Wang and Zhou (Genbank accession # ASU09687) reported a partial sequence of 239 amino acids that is highly similar to *S. gregaria* VG-B (84% identity, 11% similarity over 239 bp overlapping sequence). In contrast, VG-A and VG-B within either species are far less similar (27% identity, 17% similarity), a fact that indicates that a gene duplication of the ancestor vitellogenin gene has occurred long before these species separated.

As can be expected from their close phylogenetic relationship, both VG-A and VG-B from *S. gregaria* are very similar to their the homologous proteins from *L. migratoria* (>80%) similarity, but far less to each other (~40%). A similar, relatively low degree of sequence similarity is observed with many other insect vitellogenins, as previously reported (Comas *et al.*, 2000 and Tufail *et al.*, 2001). Despite the low degree of sequence identity, multiple alignments of these proteins and phylogenetic tree construction by the

Neighbour-joining algorithm shows a tree that mostly agrees with the phylogenetic trees previously constructed by others (Pitulachs *et al.*, 2003, Hughes, 2010; Tufail *et al.*, 2014).

The complete mRNA nucleotide sequences of VG-A and VG-B in desert locust were almost identical in size with about 5.5 kb. Their amino acid sequences contain a number of conserved motifs common to members of the vitellogenin gene family: The vitellogenin N-terminal domain, also called lipoprotein N-terminal domain (LPN) because of its presence in most other members of the LLTP superfamily as well (Anderson *et al.*, 1998), spans the region from position 21-701 (VG-A) and 23-721 (VG-B); towards the C-terminus (VG-A, VG-B, a von Willebrand factor D domain is present that is believed to be involved in multimerization and optimal secretion of extracellular proteins, as well as the interactions with other ligands. Thus, it may play a role in the interactions between vitellogenin subunits in the multimer in holoprotein. Finally, a domain of unknown function (DUF1493) that can be found in many vertebrate and invertebrate vitellogenins is identified in VG-A (745-1002) and VG-B (754-1017). Other motifs previously linked to insect vitellogenins are also present: a high degree of conservation in the positions of 9 cysteine residues in the C-terminal region, following a short conserved motif GL/ICG (Romans *et al.*, 1995; Nose *et al.*, 1997 and Tufail *et al.*, 2000). In addition, the motif DGXR was found 18 residues upstream of the motif GL/ICG in VG-A, as it was identified in most insect VGs (review by Tufail *et al.*, 2014). In VG-B, however, this motif was found as well but three residues downstream of the GL/ICG motif. Like in vitellogenins from many, but not all insect species, VG-B (but not VG-A) harbors a small polyserine sequence in the vicinity to the potential RXRR cleavage site in the aminoterminal region of the protein (369-376). While the function of such polyserine domains is not well understood, they may be involved in the phosphorylation of these proteins (Sappington and Raikhel, 1998).

There are seven potential N-glycosylation sites in VG-A and four in VG-B. The vitellogenin proteins are known to be glycolipoproteins, as they are responsible for transferring lipids and carbohydrates along with their amino acids and other minerals to the ovaries. The glycosylation happens as a co- or post-translational modification in the fat body and prior to excretion of VG into the hemolymph (Dhadialla and Raikhel, 1990). The glycosylation of VGs may play a role in keeping these proteins highly soluble in the yolk fluid (Giorgi *et al.*, 1998). Also, the added N-glycan moieties on these proteins may be a source of carbohydrates during embryogenesis (Roth *et al.*, 2010).

SDS gel electrophoresis of hemolymph proteins from mature reproductive female showed five bands with distinct molecular weights which were absent in the hemolymph of males and young adult females: one at ~200 kDa, two at ~100 kDa, and two at ~60 kDa. The four smaller polypeptides correspond to the polypeptides from the egg extract, in which vitellin is the primary component. The highest band may represent vitellogenins without further cleavage other than the removal of the small signal peptide at the N-terminus. However, as this band is fairly weak, it appears that most of the protein has been cleaved by endoproteases. All insect vitellogenins have a cleavage site for proteases of the subtilisin family (Barr, 1991; Chen and Raikhel, 1996) in the N-terminal domain, and an additional cleavage site in the center or the C-terminal domain was identified in hemimetabolous insects in which the vitellogenin precursor is cleaved into several fragments (Tufail and Takeda, 2008). The amino acid sequences of VG-A and VG-B both have the consensus cleavage site RXR/KR in the N-terminal domain (VG-A: RYRR, 349-352; VG-B: RPRR, 351-354), as well as a similar di-basic site closer to the C-terminal end (VG-A: RNRK, 1183-1186; VG-B: RSKR, 1190-1193). Cleavage in these positions would result in a smaller N-terminal fragment (<40 kDa), and a large and small subunit for each protein of ~94 kDa and 69 kDa. The exact size of the resulting fragments depends on glycosylation or other posttranslational modifications.

Our results are in general agreement with earlier findings on locust vitellogenin. Chen *et al.* (1978) reported 8 polypeptides for *L. migratoria* that fall into two groups of molecular weights, the first group is at 52 to 65 and the second one ranges from 105 to 140 KDa. Mahamat *et al.* (1997) found nine apoproteins of about 124, 120, 105, 60, 59, 58, 57, 53 and 34 kD in *S. gregaria*. Different experimental techniques by these authors, such as allowing the hemolymph to clot for 30 min past collection, may have lead to partial proteolysis from other hemolymph enzymes. A similar subunit pattern is also seen in other hemimetabolous insects, such as the American cockroach, *Periplaneta americana,* where four polypeptides were seen in SDS PAGE, at 170, 150, 100 and 50 kDa; the 150 band was a minor polypeptide band (Tufail *et al.*, 2001).

As pointed out above, both vitellogenins are specific for adult female locusts. Their mRNA can be detected with quantitative real-time PCR in the fat body around eleven days after adult eclosion. Their levels rise during the following three days to peak at day 14. VGs expression continues at a high level throughout the lifetime of the adult female, with potential periodic fluctuations.

In early 1978, both vitellin and vitellogenin were purified and identified from the yolk and the hemolymph of the migratory locust; both yielded various polypeptides ranging from 52 to 140 KDa on the SDS-gel electrophoresis (Chen *et al.*, 1978). In locusts and many other insect species, VG production in the fat body is strictly under hormonal control. Juvenile hormone (JH) and JH analogs have been shown to be involved in the transcriptional regulation of vitellogenin expression (Chen *et al.*, 1979; Chinzei and Wyatt, 1985; Glinka and Wyatt, 1996; Wyatt and Davey, 1996). These authors showed that vitellogenin is the most abundant transcript in adult females, as it is necessary to produce enormous amounts of this protein prior to oviposition. It is possible that the presence of two genes helps in meeting these demands. Indeed, we found that VG-A and VG-B are co-expressed during this period, It has been found that the expression of VG-B is always less than VG-A. Tufail *et al.,* (2001) pointed out that the existence of multiple VGs in other insect species may be advantageous in the provision of the required amounts of yolk protein to the developing oocytes; alternatively, duplicated genes may have survived as they acquired different functions. For example, vitellogenins are also believed to be involved in the hormonal regulation in the honey bees: the inhibition of VG expression in the honey bee workers induced a significant increase in JH titer and its putative receptor, which in turn should stimulate the expression of VG (Guidugli *et al.*, 2005). In social insects like honeybees, vitellogenin may play different coordinating roles on behavioral traits and social organization, and influence the lifespan of the honey bee worker (Nelson *et al.*, 2007).

With the availability of the full-length transcripts of VG-A and VG-B from *S. gregaria* and the corresponding sequences we have identified in the genome of *L. migratoria,* it is now possible to study the role of vitellogenin and its regulation at the molecular level. Upstream sequences in the promoter regions of these different genes would be helpful to identify regulatory elements involved the hormonal regulation of these genes. If one can selectively block the expression of either of these genes, at the transcriptional or post-transcriptional level, it may be possible to get insights into their specific physiological functions.

Connecting statement 2: Nuclear receptors involved in vitellogenin expression

Juvenile hormone regulates the expression of the vitellogenin genes in locusts. It has been proposed that JH-III acts by binding to different nuclear receptors that in turn bind to regulatory elements and influence the expression of these genes. The identification of these receptors has been challenging for a long time due to the various functions and processes that are controlled by the JH. Two possible receptors have been identified across multiple insect species, namely Met and RXR. In chapter 2, the two VG sequences in the desert locust have been sequenced and analyzed, and their expression pattern in the fat body measured throughout the adult female lifespan. The next chapter describes the use of the dsRNAi knockdown technique to test the involvement of these two candidate nuclear receptors in the regulation of the expression of the two VGs.

Chapter 3. The nuclear receptors Met and RXR promote vitellogenin expression in the desert locust

Abstract

In all oviparous species, the vitellogenesis is a highly regulated hormonal process. The juvenile hormone (JH-III) in the orthopteran species like locusts is the main hormone controlling this mechanism, in addition to many other physiological processes throughout the insects' life history. Upon binding with transcription factors, JH mediates the expression of the vitellogenesis-related genes. Methoprene-tolerant (Met) and retinoid X receptor (RXR) are believed to work as JH receptors in insects. We used RNA interference (RNAi) to test the involvement of these candidate nuclear receptors in the expression of the two vitellogenin (VG) genes present in the desert locust. Quantitative real-time PCR and SDS-gel electrophoresis confirmed that the depletion of these receptors eliminates or substantially reduce the transcription of both VG genes in the fat body and on the proteins present in the hemolymph. These results expand previous studies from other insect species that have shown the involvement of these nuclear receptors in the regulation of vitellogenesis.

3.1. Introduction

During oogenesis, adult females need to produce very large amounts of the yolk protein precursor; vitellogenin. The regulation of vitellogenin genes in insects takes place at the transcriptional level via nuclear receptors (Edwards *et al*., 1993; Zhang and Wyatt, 1996; Song *et al*., 2013). These receptors are mediator proteins that regulate the expression of their target genes by binding to the regulatory hormones to form a hormone-receptor complex which in-turn interacts with hormone response elements (HREs) usually located in the upstream promoter sequence of these genes. There is abundant evidence that juvenile hormone is the key regulating factor that plays a critical role in the transcription regulation pathway of VG genes. In some dipteran species, the mutual actions of both JH and ecdysone are required (Adams and Filipi, 1988; Agui *et al*., 1991). In addition to the Juvenile hormone and ecdysteroids in many insect species, numerous small gonadotropic neuropeptides have been identified to modulate and control the oocyte maturation and VGs transcription directly or indirectly (see review by De Loof *et al*., 2001; Kuczer *et al*., 2007; Tufail *et al*., 2014).

In locusts, VG synthesis can be completely abolished by preventing JH production by the *corpora allata*, either surgically or chemically (Chen *et al*., 1979); the administration of external juvenile hormone or its more stable analog methoprene fully restores VG expression. Comparative analysis of the promoter regions of the JH-induced genes has suggested the presence of potential HREs (Wyatt, 1988). This laboratory later identified in *Locusta migratoria* a 15 bp, partially palindromic sequence (GAGGTTCGAG$^{A}/_{T}$CCT$^{T}/_{C}$) upstream of the JH-induced gene *jhp21* as a putative juvenile hormone response element (JHRE) which binds to a nuclear factor in the fat body in a JH-dependent manner (Zhang and Wyatt, 1996). These nuclear factors may mediate the molecular action of JH in regulating the transcription of various genes that control distinct physiological processes in insects. While there is no conclusive evidence for a single JH-receptor, studies with a mutant fruit fly, *Drosophila melanogaster* and other insect species revealed two potential candidates: ultraspiracle (USP) and methoprene-tolerant (Met) (Riddiford, 2008; Parthasarathy *et al.,* 2010; Charles *et al*., 2011; Elgendy *et al*., 2014; Hult *et al.,* 2015; Jindra *et al*., 2015).

Methoprene-tolerant (Met) protein is a member of the basic helix-loop-helix/Per-Arnt-Sim (PAS) nuclear receptor family that modulates the JH molecular actions (Ashok

et al., 1998). Met was found to sense and bind directly through its C-terminal PAS domain to JH-III and its hormonal analogs methoprene and pyriproxyfen, at physiological concentrations with the nanomolar range (Miura *et al*., 2005; Charles *et al*., 2011). Noteworthy, the members of this nuclear receptor family usually form heterodimers with other receptors to exert their actions (Kewley *et al*., 2004).

The other potential factor, ultraspiracle (USP), is believed to form a heterodimer with the ecdysone receptor (EcR) to modulate ecdysone-regulated genes in a positive or a negative pathway relying on the presence or absence of the EcR-ligand; 20-hydroxyecdysone (Riddiford *et al*., 2001; Riddiford, 2008). Also, it was found that USP as a homodimer can bind to JH-III (Jones and Sharp, 1997), and perhaps function as its nuclear receptor during the larval stadium in preventing metamorphosis. While the binding affinity of USP to JH-III is too low for the physiological concentration of JH (Jones *et al*., 2001), later studies from the same research group have shown that USP binds with higher affinity to the JH-III precursor methyl farnesoate (Jones *et al.,* 2006). USP is only found in dipteran and lepidopteran insects, and its ortholog in beetles and other hemimetabolous insects is called retinoid X receptor (RXR).

Most investigated insect species possess a single vitellogenin, but two different vitellogenins are found in locusts. We have determined the sequences of these proteins and their genes in the desert locust, *Schistocerca gregaria,* and found that VG-A and VG-B are co-expressed in the same temporal pattern, suggesting an identical hormonal regulatory mechanism. Both vitellogenins are adult specific proteins that are entirely absent in immature stages. Their expression rises after 10 days of the adult female ecdysis, and the levels of mRNA remain high throughout the life of the adult female.

In this study, the RNA interference technique was used to study the role of the nuclear receptors RXR and Met in the transcriptional regulation of the VG genes in the desert locust fat body.

3.2. Material and Methods

3.2.1. Insects

The insects were kept and reared as described in the second chapter under 16 h light / 8 h dark cycles. Freshly cut seedling grass and wheat bran were mainly used as the food source for the insects. Selected females were marked on the ventral side of the thorax with a permanent marker pen and reared in separate cages starting from the first day of the adult female eclosion.

3.2.2. RNA extraction and cDNA synthesis

Total RNA was extracted from the treated and the control insects fat body of the two weeks old adult females using the RNeasy Plus Universal Mini Kit (Qiagen, Cat.# 73404) following the manufacturer's protocol. The iScript cDNA Synthesis kit (BioRad, Cat#: 170-8890) was used to generate the complementary DNA library as illustrated in Chapter 2.

3.2.3. Polymerase chain reaction (PCR)

PCR reactions were conducted to generate cDNA specific amplicons for Met and RXR that could be used to produce double-stranded RNA molecules required to knock down the corresponding transcripts. Specific primers (Table 3-1) were designed based on the available sequences on the GenBank database (RXR: accession # KU139127; Met: Accession # KF471131). Fermentas 2X PCR Master Mix (Cat#: K0172) was used to amplify the PCR products from 200 ng total cDNA at annealing temperatures 51 °C (Met) or 60 °C (RXR), as described in the manufacturer's protocol.

Table 3-1 Primers used to generate the Met and RXR dsRNAi

Primers	Sequence 5\`-----------3\`
Met- Forward	CTCATTCGGGAACTACGGATAG
Met- Reverse	CTCTTCCTTCCTGGCATGTATC
RXR- Forward	CCCATCAAGCGGGAGTGGGCA
RXR- Reverse	ACAAATGCTCCAAACACTTAA

3.2.4. Double-stranded RNA interference

To induce gene silencing for the genes of interest (Met and RXR), we conducted PCR reactions to target the coding area of these genes. First, the targeted amplicons were amplified with simple PCR reactions with the specific primers (table 3-1). The PCR products were purified by agarose gel electrophoresis. The purified PCR products were used as a template in two separate PCR reactions at (95 °C for 3 min, followed by 10 cycles of 95 °C for 30 s, 53 °C for Met or 60 °C for RXR for 45 s and 72 °C for 45 s, then 35 cycles of 95 °C for 30 s, 65 °C for 45 s and 72 °C for 45 s) eventually 72 °C for 7 min, where the forward or reverse specific primers used with the respective modified reverse or forward primers that include the T7 promoter sequence (5`-GGATCCTAATACGACTCACTATAGG -3`) at the 5` end. The PCR products were excised and purified from the gel and sent for sequencing to confirm the specificity of the products and including the T7 promoter sequence. The generated amplicons were used to create the dsRNA interference as described in the manufacturer protocol using the T7 RiboMAX™ Express RNAi System, Promega (Catalogue # P1700).

Briefly, 1 µg DNA of each amplicon was mixed with 10 µl T7 2X Buffer, 2 µl T7 Express Enzyme Mix and DEPC water added to bring the total volume to 20 µl. The reaction was mixed and incubated in the thermocycler for 2 h at 37 °C. To anneal the complementary RNA of each gene, the products were mixed in one tube and incubated at 70 °C for 10 min in the thermocycler followed by a gradual decreasing by 2 °C and incubation for 1 min at each degree until room temperature. Then the product was mixed and incubated for 30 min at 37 °C with the provided RNase (2 µl freshly diluted at 1:200) and DNase (2 µl RQ1 RNase free DNase) solutions to remove the ssRNA and the DNA templates. To purify the dsRNAi, 4.4 µl 3 M sodium acetate and 110 µl 95% ethanol were added to the reaction and incubated for 5 minutes on ice, followed by 10 min spinning at 16000 xg. The pellet was removed, washed with ethanol and left to dry for 15 min. The dsRNA was resuspended in 100 µl DEPC-treated water. After measuring the dsRNA concentrations with a Nanodrop spectrophotometer, the products were diluted and aliquoted into 10 µg/10 µl portions, and stored at -20 °C.

3.2.5. Treatment and sample collection

For knockdown experiments, dsRNA samples were allowed to thaw on ice and spun for few seconds in a microcentrifuge. For each treatment, 10 µg of dsRNA was injected into the body cavity of the adult females. Treatment started within 16 h of adult eclosion, and the injection was repeated twice, at day 7 and day 13. Tissue samples were collected one day later, at day 14 after adult eclosion. For control insects, the same volume of DEPC treated water was injected at the same time intervals. The fat body and the hemolymph were collected as described in Chapter 2, and stored at -20 °C if not processed immediately.

3.2.6. Quantitative real-time PCR

Real-time PCR was carried out as described in Chapter 2, with primers specific for Met and RXR (table 3-2). To evaluate the expression of vitellogenin, VG-A and VG-B real-time primer sets shown in (Chapter 2) were used. To normalize the reactions, both GAPDH and EF1α were evaluated for their use as reference genes. For each data point, the average Ct value of three PCR replicates with 10 ng of cDNA template for each of the reference genes and the target genes were determined. The expression of each target gene was normalized against each of the reference genes, and the differences between different experimental groups were calculated using the $2^{-\Delta\Delta Ct}$ method described by Livak and Schmittgen (2001). The relative expression was measured using the Ct of each reaction as follow:

Control: $\Delta Ct\ (H_2O) = Ct\ (H_2O)$ of the gene of interest – $Ct\ (H_2O)$ of reference

Treatment x: $\Delta Ct\ (x) = Ct\ (x)$ of gene of interest – $Ct\ (x)$ of reference

Change in cycle number for treatment x: $\Delta\Delta Ct\ (x) = \Delta Ct\ (x) - \Delta Ct\ (H_2O)$

Relative transcript level after treatment: $100 / 2^{\Delta\Delta Ct}$

Table 3-2 Primers used to measure the gene expression of Met, RXR, and EF1α

Primers	Sequence 5`-----------3`	Efficiency %
Met- Forward	CATGCAGCGACAGGAAACAC	99.8%
Met- Reverse	TGTGTGAGGTTTGAGCGGAA	
RXR- Forward	ATCCAGAGGTGAGGGGTTTG	94.6%
RXR- Reverse	GAAGGCAGACGAAGCAGAAG	
EF1α - Forward	GATGCTCCAGGCCACAGAGA	92.5%
EF1α – Reverse	TGCACAGTCGGCCTGTGAT	

3.2.7. SDS-gel protein electrophoresis

SDS-gel electrophoresis was used to analyze the effectiveness of the knockdown of each nuclear receptor on the translated and released VG proteins in the hemolymph. SDS electrophoresis was carried out as described in Chapter 2, using 8% separating gel and 5% stacking gel.

3.3. Results

3.3.1. Search for the Met and RXR Sequences

Before designing dsRNA to knock down the Met nuclear receptor, it was necessary to identify at least a partial sequence of this nuclear receptor in the desert locust, *Schistocerca gregaria*. Based on the sequence from the migratory locust, *Locusta migratoria* (Accession # KF471131), Met-specific primers were designed and used to amplify the Met sequence from *S. gregaria* fat body. At the annealing temperature of 51 °C, PCR produced one strong band with the anticipated size of ~ 900 bp (Figure 3-1). The PCR product was excised from the gel and sent for Sanger's sequencing.

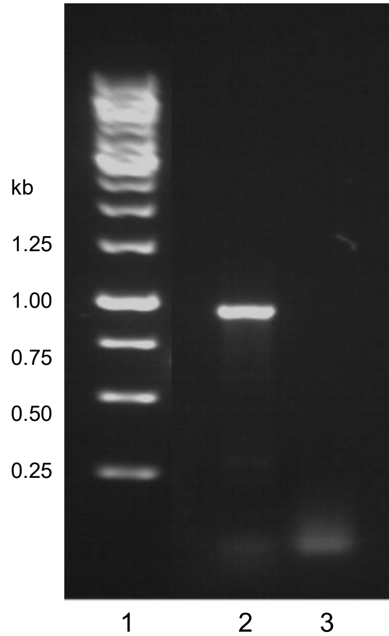


Figure 3-1 PCR results of the amplified potential Met nuclear receptor.
Primers were designed based on the sequence of Met from the migratory locust and used against cDNA from fat body of the desert locust. Lane 1: DNA ladder, Lane 2: PCR product band ~ 900bp, Lane 3: no template control.

The sequence of the PCR product was confirmed to be the Met nuclear receptor of the desert locust, as it was 88% identical to the *L. migratoria* sequence over the entire 893 bp. With additional PCR reactions, a slightly larger partial sequence of the Met fragment was obtained (Appendix B; sequence B1), which served as a template for primer design for RT-PCR and dsRNA production.

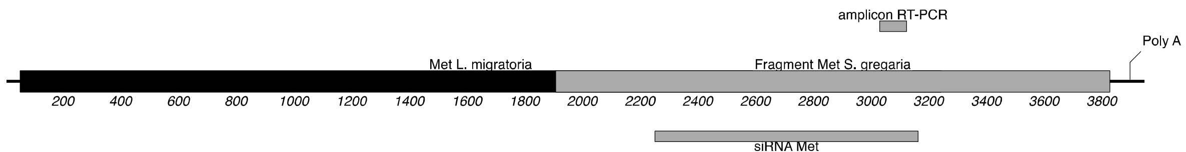


Figure 3-2 Sequence of the Met fragment from *S. gregaria*
The schematic graph shows the obtained Met sequence of *Schistocerca gregaria,* and the positions of the dsRNAi used to knock down the corresponding transcript, and the amplicon of the qPCR real-time primers.

RXR has been sequenced from both *Locusta migratoria* (Hayward *et al.,* 1999 & 2003) and *Schistorcerca gregaria* (Cui, 2008). In both species, two isoforms exist, a short form (Genbank accession KU139126) and a long form (Genbank accession KU139127),

distinguished only by a 66 bp insertion in the latter one (Figure 3-3). Primers specific for RXR were used to amplify the sequence from *S. gregaria* fat body, in order to design RXR dsRNAi based on the exact sequence for this species. These primers were designed to anneal downstream of the 66 bp insert, to assure the knockdown of both isoforms through the injections of the dsRNAi generated. The PCR product of ~ 300 bp (Figure 3-4) was excised from the gel and sent for sequencing to confirm the specificity of the primers and the sequence before using them to synthesize the dsRNA.

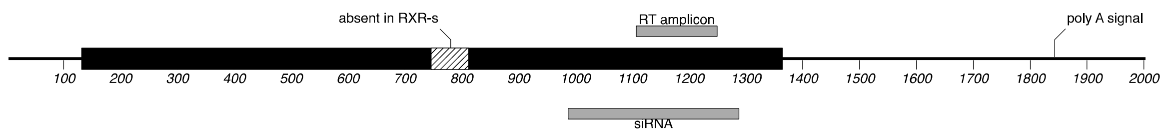


Figure 3-3 Schematic representation of the two RXR isoforms from *S. gregaria*
The graph represents the sequence of the RXR gene from *Schistocerca gregaria, the* position of the exon that is absent in the short form, and the areas targeted with dsRNAi and real-time qPCR.

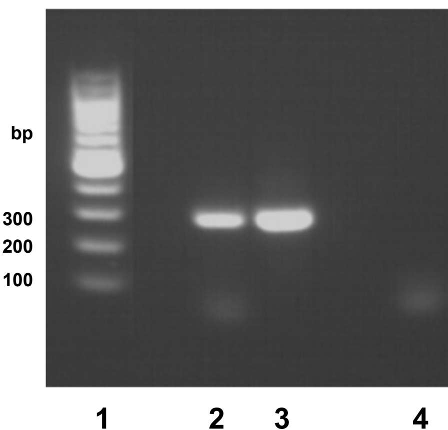


Figure 3-4 Amplified PCR products of RXR
cDNA from adult female locust was used to amplify the two isoforms as one product. Lane 1: DNA ladder, Lanes 2-3: ~300 bp of the amplified amplicon, Lane 4: no template control.

3.3.2. The effect of Met dsRNAi on vitellogenin expression

Under our rearing conditions, the expression onset of VGs in adult desert locust females starts around 11 days after adult eclosion; this represents the first gonadotrophic cycle. The expression increases in the following three days to reach its first peak around 14 days and continues at varying rates for the adult female lifespan. As our knockdown

attempts were started along with the VG expression studies and before knowing the exact temporal expression pattern, we decided to start administering the injections of Met dsRNAi at the first day of the adult female eclosion to eliminate any possibilities of early expression of the VGs. Another two boosting doses were applied to 7 and 13 days old adults, and samples were collected 24 h after the last injection.

We assessed the effect of Met dsRNAi on the corresponding transcript, using quantitative real-time PCR. Specific real-time primers were designed from the identified partial sequence of *S. gregaria* Met and used to target the expression of its gene in the fat body. Both GAPDH and EF1α were initially used as internal controls, to normalize the expression of genes of interest. The Ct values for GAPDH (Table 3-3) were consistently lower than the control for both experiments, averaging 1.4 cycles for animals treated with RXR RNAi, and 2.2 cycles for Met RNAi, indicating that the expression of this gene may be affected by these transcription factors. In contrast, the expression of EF1α was stable and varied by less than 0.5 cycles between the different treatment groups, and thus we used EF1α to normalize the expression of the other genes.

Table 3-3 Ct values of the housekeeping genes of the control and the treated insects

The values represent the Ct average of 11 control insects, 13 Met-dsRNAi treated, and 15 RXR-dsRNAi treated ± S.D. of the two internal reference genes in control and treated insects.

Experiment	GAPDH	EF1α
Control (H_2O)	20.49±0.54	19.76±0.73
Met dsRNAi	18.28±0.45	19.30±0.75
RXR dsRNAi	19.06±0.46	19.60±0.85

The injections with Met dsRNAi reduced the Met mRNA by more than 80% of the levels found in control animals injected with the same volume of DEPC-treated water. There were no noticeable morphological effects on the treated insects. Met dsRNA injections resulted in the complete absence of mRNA for both, VG-A and VG-B (Table 3-4).

Table 3-4 Transcript changes after treatment with Met dsRNAi
Relative expression of different genes in control and Met dsRNAi-treated insects on day 14 after adult eclosion. Expression levels were determined by the ΔΔCt method with normalization against EF1α and are expressed as a percentage relative to the control insects.

Experiment	Met	VG-A	VG-B	RXR
Ct Met	28.38±0.72	32.82±2.21	32.82±1.80	27.73±0.77
ΔCt Met	9.02±0.70	13.45±1.89	13.26±1.53	8.38±0.53
Ct H_2O	26.21±0.93	18.55±3.22	18.95±3.32	28.84±1.18
ΔCt H_2O	6.45±0.91	-1.21±3.03	-0.81±3.12	9.08±1.03
ΔΔCt	2.57	14.68	14.07	-0.70
$2^{\Delta\Delta Ct}$	5.95	26179.57	17171.07	0.61
% of control	16.8%	0.004%	0.006%	162.8%

The hemolymph from the same insects was collected along with the fat body at day 14. Comparison of the protein pattern by SDS PAGE (Figure 3-5) shows the reduction or complete disappearance of the VG bands identified in Chapter 3 (~ 200, 120, 100 and 60 kDa) in all of the RNAi treated insects. The band at 120 kDa, normally the most prominent in control insects, while still detectable, was much fainter in the treated animals.

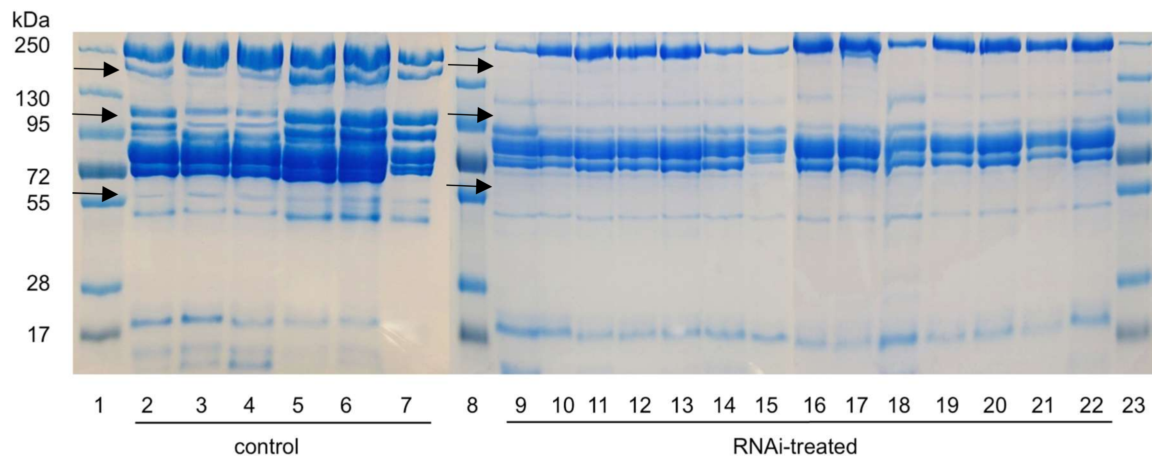


Figure 3-5 SDS PAGE of hemolymph proteins from14 day old adult females
Lane 1, 8, 23: Molecular weight markers; Lane 2-7: control animals injected with water; Lane 8-22: animals injected with Met RNAi.

3.3.3. The effect of RXR dsRNAi on vitellogenin expression

Similar experiments were carried out to determine if the RXR nuclear receptor is required as well for vitellogenin production. We used RXR dsRNAi to knock down the

expression of this potential nuclear receptor. Adult females were injected by the same protocol as used for Met for three times with 10 µg of RXR dsRNAi, and quantitative RT-PCR with primers targeting the transcripts of both RXR isoforms was used to confirm the effectiveness of the knockdown. After three injections with RXR dsRNAi, the RXR's mRNA transcript in the fat body was knocked down to less than 30% of its control level, with no mortality rate and no external abnormalities, The reduced levels of RXR also seem to influence the expression of VG mRNAs, which were reduced to 4% of the control values (Table 3-5).

Table 3-5 Transcript changes after treatment with RXR dsRNAi
Relative expression of different genes in control and RXR dsRNAi-treated insects on day 14 after the adult eclosion using qRT-PCR. Expression levels were determined by the ΔΔCt method with normalization against EF1α and are expressed as a percentage relative to the control insects.

Experiment	RXR	VG-A	VG-B	Met
Ct RXR	30.49±1.11	23.05±3.19	23.58±3.22	26.02±1.11
ΔCt RXR	10.90±1.06	3.46±3.00	3.98±2.99	6.42±1.25
Ct H_2O	28.84±1.18	18.55±3.22	18.95±3.32	26.21±0.93
ΔCt H_2O	9.08±1.03	-1.21±3.03	-0.81±3.12	6.45±0.91
ΔΔCt	1.82	4.67	4.79	-0.03
$2^{\Delta\Delta Ct}$	3.52	25.40	27.65	0.98
% of control	28%	4%	4%	102%

Although the mRNA for both VGs was present in the fat body of 14-day old females, the corresponding protein bands were quite weak on the SDS PAGE gels of the hemolymph (Figure 3-6). We observed a similar pattern as seen for Met knockdown, where the three bands around 200, 100 and 60 kDa almost completely disappeared. The band at 120 kDa, normally the strongest of the VG bands, was also much weaker than in the controls, indicating that much less vitellogenin was present in the hemolymph of the treated insects at day 14.

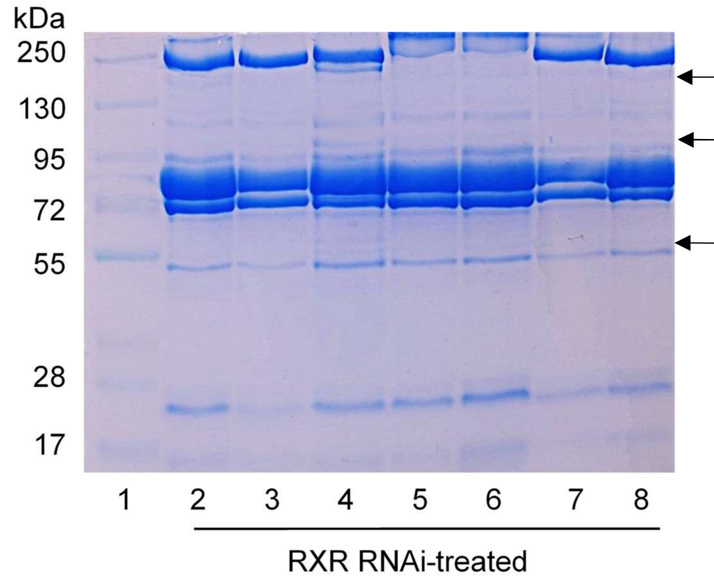


Figure 3-6 SDS PAGE of hemolymph proteins from RXR RNAi treated 14-day old adult females

Lane 1: Molecular weight marker; Lane 2-8: hemolymph extracted from treated insects injected with RXR dsRNAi

3.4. Discussion

The transcriptional regulation of the VG genes in insects has been shown to be under the control of endocrine factors (Wyatt and Davey, 1996; Engelmann, 2002). In orthopteran insects, it is well established that JH acts as a primary regulating hormone during the adult fertility stage, controlling reproduction mechanisms including the transcription of several vitellogenesis-related genes. In *Locusta migratoria,* JH was found to induce and control the expression of VG genes in a tissue, stage, and sex-dependent manner (Chen *et al*., 1979; Chinzei *et al.,* 1982; Dhadialla and Wyatt, 1983). In the lubber grasshopper; *Romalea microptera*, immunoassay analysis of JH III in the hemolymph along with the quantification of VG and VG mRNA during the vitellogenesis cycle have shown that they follow the same pattern as seen in the migratory locust, suggesting the involvement of JH III in the synthesis pathway of VG (Borst *et al.,* 2000). In other hemimetabolous species, such as the German cockroach *Blattella germanica*, treatment of the fat body of immature insects with juvenile hormone stimulated the production of VG protein (Kunkel, 1981).

It has long been postulated that juvenile hormone acts as a ligand to a nuclear receptor, in a mechanism similar to steroid hormones. The search for JH-receptors in insects, however, has been difficult and controversial, due to the polymorphic actions of JH as well as its unique chemical structure that makes it different from other invertebrate and vertebrate hormones. Therefore, it is possible that the actions of juvenile hormone are mediated by various factors, depending on the stages, pathways, and specific genes on which the hormone acts. The RNA interference technique, which is very effective in locust species, has made it possible to knockdown potential receptors and to look for the effects on development, reproduction, and gene expression.

In order to evaluate the role of the nuclear receptor Met in the transcription of VGs in the desert locusts, we administered RNAi targeting the Met mRNA transcript. RNAi was administered with a high dosage (10 μg in 10 μl injection volume) on the first day of the adult eclosion, to precede the first gonadotropic cycle. The first injection was followed by two additional doses with the same concentration of RNAi. The fat body and the hemolymph were collected from treated and control insects to examine the effect of the Met dsRNAi on its corresponding transcript, and its influence on VG expression at the RNA and protein level. Met RNAi was effective, knocking down its transcript to 16% of its normal level. As a consequence of this, the expression of both VG genes was prevented, and their transcripts and corresponding potential protein bands in the hemolymph were almost undetectable, suggesting a direct involvement of Met in controlling VG gene expression.

Our data are in agreement with previous studies in both hemi- and holometabolous insects that showed the involvement of Met in this pathway, both directly and indirectly. In the migratory locust, it was demonstrated that the Met nuclear receptor forms a complex with a steroid receptor co-activator to mediate the JH-molecular action in activating some of minichromosome maintenance (Mcm) genes which prepare the fat body to produce large amounts of VGs through genome duplication, and the knockdown of these genes block VG expression similar to the Met knockdown or JH-deprivation (Guo *et al.,* 2014). In the absence of the JH, it was found that the Met nuclear receptor forms a homodimer; upon binding with its ligand through the C-terminal PAS domain (which appears to be a potential domain for binding JH), the Met-homodimer dissociates allowing Met to bind to another nuclear receptor, Taiman, and forming a ligand-dependent complex that

transduces the JH actions (Charles *et al*., 2011 and references therein: Godlewski *et al*., 2006; Li *et al*., 2011).

It is well established that the action of juvenile hormone leads to dramatic changes in fat body at the cellular level, including the induction of extensive DNA replication. During the maturation of the adult female, fat body cells become polyploid, which is believed to contribute to the rapid synthesis of the large amounts of vitellogenin required (Nair *et al.,* 1981; Oishi *et al.,* 1985). Recent studies by Guo *et al.* (2014) showed that *Mcm4* and *Mcm7*, the two genes that induce DNA replication and polyploidization, are also under the control of JH.

An indirect effect of JH to upregulate VG production has been shown in the red flour beetle, *Tribolium casteneum,* where JH binds to Met forming the ligand-receptor complex that activates the insulin-like peptides 2 (ILP2) while the feeding activates ILP3, both ILP2 and ILP3 through insulin signaling pathway phosphorylates the repressive transcription factors; FOXO which in turns stimulates the production of VG2 (Sheng *et al*., 2011). On the other hand, Bitra and Palli (2009) found that Met binds USP, the lepidopteran ortholog of RXR, however independent of the presence of juvenile hormone.

To evaluate the potential role of RXR in the regulation of vitellogenin expression, we carried out a similar silencing study to that of Met-knockdown with RXR dsRNAi. Real-time qPCR data analysis revealed a statistically significant reduction of both VGs mRNA transcripts in 14-day old adult females to low, but clearly measurable levels comparable to the levels seen in untreated animals two or three days earlier (day 11 or 12 after adult ecdysis). Although RXR and Met are reduced by similar margins, the effect of Met is much more pronounced, a finding that could suggest that Met is indeed acting directly on the promoters of the vitellogenin genes. RXR, in contrast, may have an indirect effect, perhaps influencing maturation or related developmental processes.

While it is clear from our study that RXR is necessary for the full expression of both VGs, it is impossible to conclude how RXR exerts its action, and whether it involves interaction with juvenile hormone, as previously proposed by Jones and Sharp (1997) for the lepidopteran ortholog USP. While these authors cautioned that the binding affinity of USP is far too low to serve as a nuclear receptor for JH III (Jones *et al*., 2001), later studies have shown that USP binds the JH III precursor; methyl farnesoate with a higher binding

affinity (Jones *et al*., 2006). Thus, it may be possible that USP or RXR by interacting with the juvenile hormone signaling pathway, mediates the expression of vitellogenin.

In the red flour beetle, *Tribolium castaneum*, it was found that the ecdysteroid receptors EcR and USP are required for egg production and affecting VG transcription (Parthasarathy *et al*., 2010). It was proposed that the knockdown of EcR or USP would inhibit the ovarian growth which mediated by 20E and indirectly inhibit the VGs transcription in the fat body. Another recent study (Elgendy *et al.,* 2014) on the American cockroach; *Periplaneta americana,* quantitatively analyzed the expression of the two different RXR isoforms spatially and temporally. It revealed a possible dual role of RXR during vitellogenic cycles with JH and ecdysteroids. These authors found that during early vitellogenesis, there was a high expression of RXR in the female fat body with the high JH-titer in the hemolymph, while a second peak in RXR expression was noticed during late vitellogenesis that coincided with a high titer of ecdysteroids. These results would suggest that RXR induces the expression of vitellogenin in the early vitellogenesis while it reduces the production at the end of vitellogenesis process.

Our results, combined with the evidence obtained from the previous research, suggests a crucial role for each nuclear receptor, Met, and RXR, in the vitellogenin synthesis pathway of the desert locust.

Connecting statement 3: Lipids are needed for reproduction and dispersal

The proliferation of locusts depends both on reproduction and dispersal, and efficient lipid transport and mobilization is essential for both processes. It is well established in many insect species that a reduction of the yolk protein vitellogenin leads to reduced reproductive success. In the previous chapters, I demonstrated that the expression of vitellogenin in adult females could be virtually eliminated by knocking down the transcription factors Met and RXR, thus preventing the delivery of nutrients to the developing eggs. In the following chapters, I will focus on the effect of eliminating proteins involved in intra- and extracellular lipid transport during dispersal. Lipids mobilized during flight are transported through the hemolymph to the flight muscles, where the intracellular fatty acid binding protein (FABP) is thought to facilitate the delivery of fatty acids to the mitochondria for beta-oxidation. While FABP is a prominent protein in all muscles that metabolize lipids, it has not been demonstrated in any species that the protein is required for efficient fatty acid utilization to fuel these muscles. The next chapter describes the knockdown of FABP by RNAi techniques in adult locusts and its impact on endurance flight.

Chapter 4. FABP silencing impedes extended flight in the desert locust, *Schistocerca gregaria*

Authors

Sanjeeva Rajapakse, David Qu, Ahmed Sayed Ahmed, Jutta Rickers-Haunerland, and Norbert H. Haunerland

SR was an Undergraduate Research Assistant in our laboratory who carried out most of the molecular work presented in this paper. ASA trained, and directly supervised SR carried out the protein analytical work and helped DQ in insect rearing and treatment. JRH assisted in the training of experimental techniques and troubleshooting, and NHH conceived and directed the study, and carried out the flight experiments

Abstract

During migratory flight, desert locusts rely on fatty acids as the predominant source of energy for their flight muscles. Lipids mobilized in the fat body, are transported to the flight muscles and enter the muscle cells as free fatty acids. It has been postulated that muscle fatty acid-binding protein (FABP) needed for the efficient translocation of fatty acids through the aqueous cytosol towards mitochondrial β-oxidation. To assess whether FABP is required for this process, dsRNAi was injected into freshly emerged adult males to knockdown the expression of FABP in the flight muscles. Three weeks after injection, FABP and its mRNA were undetectable in the flight muscles, indicating efficient silencing of the FABP gene. The treated animals exhibited no morphological or behavioral differences from control animals at rest. In tethered flight experiments, both control and treated insects were able to fly continually in the initial, carbohydrate fueled phase of flight, and in both lipids were mobilized and released in the hemolymph. Flight periods exceeding thirty minutes, however, when carbohydrate resources are depleted, and fatty acids become the predominant source of energy, were not possible for FABP-depleted animals, while control insects were able to fly for more than 3 h. These results provide unambiguous proof for an essential role of FABP in skeletal muscle energy metabolism *in vivo*.

4.1. Introduction

For many centuries, locusts have inflicted severe damage to human populations in African and Asian countries. Every few years, when weather conditions are favorable, locusts that normally develop dispersed in their solitary stage accumulate in large numbers and undergo a phase transformation to their gregarious form (Pener and Simpson, 2009). As adults, gregarious locusts form gigantic swarms that can migrate in a coordinated manner for several hundred kilometers, touching down for feeding and eradicating much of the vegetation along their path. Migratory flight of locusts is among the most energy demanding activities, and insects have developed an efficient mechanism to fuel this metabolic activity (Wegener, 1996). In the initial phase of flight, the readily available disaccharide trehalose serves as the main energy source for muscle contraction, but within 30-60 minutes lipids become the exclusive metabolic fuel (Mayer and Candy, 1969).

Lipids are stored as triglycerides in the fat body. Their mobilization is initiated by the release of adipokinetic hormone (AKH), which activates a signal transduction pathway that triggers the action of a lipase in the fat body. One fatty acid chain is cleaved from the triacylglycerol molecule, and the resulting diacylglycerol, which is the major transport form of lipids in insects, is released into the hemolymph (Van der Horst and Rodenburg, 2010).

Locusts use an effective transport system, often referred to as the "lipophorin shuttle," to assure sustained delivery of diacylglycerol (DAG) to the flight muscle. (Van der Horst and Rodenburg, 2010). In resting insects, the predominant hemolymph lipoprotein is the high density form of lipophorin (HDLp), a protein composed of two the apoproteins apoLp-I (~250 kDa) and apoLp-II (~80 kDa), as well as phospholipids, diacylglycerol, and smaller amounts of other lipids, which together amount to around 20% of the mass of the lipophorin molecule. Upon their release from the fat body, numerous DAG molecules associate with HDLp, and the lipid-enriched particle is stabilized by the binding of several molecules of a third apoprotein, apoLp-III (~18 kDa). The resulting low-density lipophorin (LDLp) has a density of ~1.02 g/ml and contains more than 40% lipid, mostly in the form of DAG. A lipoprotein lipase located at the flight muscle membrane hydrolyzes DAG; free fatty acids enter the flight muscle, while glycerol and apoLp-III are released into the hemolymph. Lipophorin returns to the high-density form HDLp, which remains in the hemolymph and can continue to transport DAG from the fat body to the flight muscle (Van der Horst and Rodenburg, 2010).

While the transport of lipids through the hemolymph has been studied extensively, less is known about how fatty acids enter the flight muscle cells and translocate through the aqueous cytosol to the mitochondria where beta-oxidation takes place. It is widely believed that fatty acid binding proteins (FABPs) play a role in intracellular transport of fatty acids, especially in muscle cells (Haunerland and Spener, 2004). FABPs belong to an ancient family of genes now called the intracellular lipid binding proteins (iLBPs) that has originated more than a billion years ago. The first gene duplication appears to have occurred approx. 900 mya, long before the vertebrate-invertebrate divergence, and hence all animals seem to have at least two distinct FABPs, reflecting the two major branches of the phylogenetic tree. Subsequent gene and genome duplications gave rise to the variety of FABP found today (Schaap *et al.*, 2002). In mammals, more than 14 different members of the gene family have been identified, with distinct differences in tissue-specific expression patterns. In contrast, fewer paralogues have been characterized in insects, which appear to express only one or two isoforms on each of the two branches (Haunerland and Thakrar, 2009). In locusts, only one FABP has been characterized to date, but recent EST or genome sequencing projects suggest a potentially larger number of paralogues. FABP was first discovered in the flight muscle of adults of the desert locust, *Schistocerca gregaria* (Haunerland and Chisholm, 1990), and later in *Locusta migratoria* (Van der Horst, 1990; Maatman *et al.,* 1994) as well. FABP is the most prominent cytosolic protein in mature adult locusts, amounting to almost 20% of all soluble proteins (Haunerland *et al*., 1992). Interestingly, the protein and its mRNA is completely absent in mesothorax muscles of nymphs. Its expression starts right after adult ecdysis and continues strongly for the following 7-8 days. FABP mRNA levels decline afterward, while the protein reaches its maximum 2 weeks after ecdysis (Haunerland *et al*., 1992). FABP is a very stable protein, and its level in the flight muscle remains constant for several weeks afterward, requiring only a low amount of mRNA to replace degrading FABP. The three-dimensional structure and the binding affinity of locust FABP are very similar to vertebrate FABP3 (Haunerland, 1994), which is expressed strongly in cardiac muscles and other lipid-dependent muscles.

In mammals, skeletal muscle contains relatively small amounts of FABP (1-2% of all cytosolic proteins), while in cardiac muscle, the most lipid-dependent tissue in mammals, FABP amounts to up to 5% (Kaikaus *et al*., 1990; Al-Hadi *et al*., 2009). In contrast, the flight muscles of the Western Sandpiper, a migratory shorebird, or the desert

locust encounter that metabolic rates two- or three fold higher than mammalian heart also possess twice or three times as much FABP, respectively. There appears to be a clear correlation between the cellular FABP content and the rate of beta-oxidation encountered by a muscle, suggesting that FABP facilitates the uptake and transport of fatty acids for energy production in those muscles (Veerkamp and Vanmoerkerk, 1993). However, in spite of many attempts, a conclusive proof that FABP is necessary for sustained muscle activity has not yet been obtained. Ablation of the muscle FABP gene in mice did not result in a distinct phenotype, and knockout mice remained active and viable, with the possible exception of somewhat reduced exercise tolerance and a slight reduction in fatty acid uptake rates (Binas and Erol, 2007). This, however, cannot be interpreted as an argument against an essential role of FABP, as compensatory overexpression of other members of the ILBP family appears to provide other proteins with similar characteristics (Haunerland and Spener, 2004).

In contrast to gene ablation, RNAi-mediated knockdown of proteins is unlikely to result in compensatory expression of other proteins. It has recently been demonstrated that locusts are excellent candidates for RNAi knockdown experiments (Wynant *et al.*, 2012). Taking advantage of the adult-specific nature of locust FABP and its well-defined expression pattern, the current study was carried out to investigate the consequences of a reduction or elimination of FABP on the flight performance of the desert locust.

4.2. Material and Methods

4.2.1. Insects

Locusts (*Schistocerca gregaria*) were reared in crowded conditions at 30 °C as described in Chapter 2. To determine FABP expression in adult locusts; freshly emerged adults were removed and reared separately until used at the specified age. Individuals between 0 and 24 hours after adult molting are referred to as day 1 adults, with each subsequent day representing an additional 24-hour period.

4.2.2. Flight experiments

Sustained flight was induced by mounting locusts at 30 °C in front of two large fans. The prothorax of the insect was fixed by means of a low-temperature glue stick to the tip of a wooden stick, which was attached to a stand in such a way that the mounted insects were suspended in the air and faced the fan directly. The wind speed was measured with an anemometer and adjusted to 4 m/s. Locusts generally engaged in flight as soon as the fans were turned on but sometimes needed to be gently nudged to initiate sustained flight. Insects that were not able to continuously fly for 15 min were discarded. Flight performance was measured at 15 min intervals and scored by a numerical scale, where 5 indicated continuous, uninterrupted flight, and 0 no flight activity even after gentle stimulation (1-3 touches with a pencil tip). The values of 4, 3, 2, and 1 were assigned if stimulation led to continuing flight for at least 5 min, 1-5 min, 10-60 s, and <10 s, respectively.

4.2.3. Protein extraction and analysis

Locust mesothorax muscles were dissected under phosphate buffered saline. The tissue was cleaned of adhering fat body and washed in saline, yielding approx. 30 mg of muscle tissue per locust. The tissue was rapidly frozen under liquid nitrogen and lyophilized for later usage. For protein analysis, 50 µg of the lyophilate were homogenized with 125 µl PBS and centrifuged at 12,000xg to remove insoluble materials, and the supernatant analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), following the protocol by Laemmli (1970). A 4 µl aliquot of the supernatant was boiled for 2 min with 6 µl Laemmli sample buffer (Bio-Rad), loaded on a stacking gel (5% T, 20% C), and separated on a 15% T, 5% C stacking gel. The Spectra Multicolor Broad Range Protein Ladder (Thermo Scientific) was used to identify the size of the proteins. Gels were run for 1 h at 180 V (200 mA) in 1x Tank buffer. Gels were stained overnight with EZblue Gel Staining Reagent (Sigma-Aldrich).

4.2.4. Lipid analysis

Following the completed flight experiment, hemolymph was collected from both control, and RNAi treated insects. Lipids were extracted according to Bligh and Dyer (1959). Lipids were separated by two-step thin layer chromatography on a normal phase

silica gel plate. The plate was developed with hexane-ether-acidic acid 60:40:1, followed by a second solvent chloroform-methanol-water 65:40:5 that was allowed to move 4 cm up the plate. Lipids were visualized by exposure to iodine vapor, and the intensities of the spots for phosphatidylcholine (PC) and 1,3 diacylglycerol determined densitometrically with the Image J software (Schneider *et al.,* 2012).

4.2.5. Total RNA preparation

Total RNA was extracted from the tissue homogenate with Qiagen RNeasy Mini kit according to the manufacturer's instructions. Freshly dissected tissue was ground up to a fine powder with a mortar and pestle under liquid nitrogen. Muscle powder (30 mg) were transferred into 600 µl RLT buffer vial, homogenized with 7 strokes through 21 gauge needle in a 5 ml syringe, and centrifuged for 3min at 12,000xg. The supernatant was transferred to a new tube and mixed with the same volume of ethanol, and separated on an RNeasy spin column as described by the manufacturer. Total RNA concentration and purity were determined photometrically on a Nanodrop 2,000C UV-Vis spectrophotometer (Thermo Scientific).

4.2.6. cDNA synthesis

Total RNA (1 µg) was transcribed into cDNA with a Bio-Rad iScriptTM cDNA synthesis (Cat # 170-8890) kit in an MJ Mini thermal cycler (cDNA synthesis program: 5 min at 25 °C, 30 min at 42 °C, 5 min at 85 °C, cooled down to 4 °C). A parallel control reaction for detecting genomic DNA contamination was carried out without the addition of reverse transcriptase.

4.2.7. Synthesis of dsRNA

The dsRNA constructs were prepared using T7 RiboMAX^{TM} Express RNAi System (Promega) which is designed for the synthesis of dsRNA molecules of > 200bp. Using cDNA as a template, the 400 bp target region of the FABP transcript (Gen Bank Accession AH010557) was first amplified in a simple PCR reaction with forward primer sgFABP2F (5'-TGGTGAAGGAATTCGCAGGCAT-3') and reverse primer sgFABP401R (5'-TGGGCCTTGTATATTCTCGTTGCC-3') using RedTaq DNA polymerase (Sigma-Aldrich) (95 °C for 3 min, followed by 40 cycles of 95 °C for 30 s, 57 °C for 45 s and 72 °C for 45

s), then 7 min at 72 °C. The PCR product was purified by agarose gel electrophoresis and sequenced. For the synthesis of single-stranded sense and antisense RNA, a T7 promoter sequence (5‘-GGATCCTAATACGACTCACTATAGG-3’) added to 5’-end of the forward or reverse primer, respectively, and two separate PCR reactions with the respective unmodified reverse or forward primer were carried out to generate two separate single promoter templates (95 °C for 3 min, followed by 10 cycles of 95 °C for 30 s, 57 °C for 45 s and 72 °C for 45s, then 35 cycles of 95 °C for 30 s, 65 °C for 45 s and 72 °C for 45s), and eventually kept for 7 min at 72 °C. These templates were directly used in a single high-yield in vitro transcription reaction. Remaining ssRNA and DNA was removed by nuclease digestion, and the 400 bp dsRNA was further purified according to the manufacturer’s instructions. Following quantification in a NanoDrop spectrophotometer, dsRNA was diluted to 1 µg/µl in DEPC treated water, analyzed on a 1% agarose gel, and aliquoted for storage at -20 °C.

4.2.8. Injection of dsRNA

Double-stranded siRNA (4 µg in 4 µl H_20) was injected with a 5 µl Hamilton syringe into the hemolymph of adult male locusts within 12 hours of adult eclosion. Control male locusts were injected with an equivalent volume of DEPC treated water.

4.2.9. Quantitative real-time PCR (qPCR)

FABP transcript in the cDNA was quantified by real-time PCR with primers FABP127F (5’-GAGCTGGAGATCCTGGACGGT-3’) and FABP273R (5’-TCCGTCCTGAGTGATGGT GGA-3’) that yielded a 147 bp product. Elongation factor 1 α (EF1α) which has been shown before to be a reliable reference gene in *Schistocerca gregaria* brain (Van Hiel *et al.*, 2009) was used as reference gene in this study (forward primer EF1α271F (5’-GATGCTCCAGGCCACAGAGA-3’), reverse primer SgEF1α336R (5’-TGCACAGTCGGCCTGTGAT-3’). The efficiency of qPCR and correlation coefficient (R2) was measured for each primer pair. All reactions were performed in duplicate in 48-well plates on MiniOpticon Real-Time PCR System (Bio-Rad). Each reaction contained 2 µl cDNA, 0.5 µl forward and reverse primer (10 µM), 5 µl 2X SsoFast EvaGreen Supermix (Bio-Rad) and 2 µl water. For all q-RT-PCR reactions, the following thermal cycling profile was used: 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. A

melt curve analysis was performed to check for primer dimers. For all transcripts, only a single melting peak was found during the dissociation protocol.

4.3. Results

4.3.1. FABP real-time PCR

Earlier work carried out with Northern Blots (Haunerland *et al*., 1992 & 1993) had established the expression pattern for locust muscle FABP, and provided a semi-quantitative estimation of the mRNA levels throughout adult maturation. These data were verified by real-time PCR with a primer pair specific for FABP. As shown in Table 4-1, FABP mRNA is undetectable in mesothorax muscle of last instar nymphs but shows up shortly after adult ecdysis. When normalized to the expression of the housekeeping gene EF1α and expressed relative to the maximal level of mRNA seen at day 7, the results from the real-time PCR closely resemble the data obtained before from the Northern Blots after normalization against β-actin (Figure 4-1). FABP mRNA is near the detection limit until 6 h after ecdysis, and increases more than 100-fold between day 1 and day 7 to its maximal value, before gradually declining over the next two weeks to the very low level needed to maintain the FABP concentration in the muscle. When 4 µg double-stranded RNA were injected within the first 24 h after ecdysis, FABP mRNA did not increase and remained undetectable for at least 21 days. There were very low rates of mortality associated with the injection: less than 10% of all injected locusts died prior to being analyzed. As similar mortality was seen for both RNAi injected or H_2O injected control animals, it is likely due to injury by the injection process, rather than the injected compound. Subsequent analysis of the protein expression pattern by SDS PAGE revealed the complete absence of the strong band at 15 kDa that represents FABP (Figure 4- 2), confirming that the RNAi indeed results in a nearly complete knockdown of the FABP gene expression.

Table 4-1 Expression of FABP mRNA in adult locusts.

ΔCt is the average of 3-4 independent replicates of Ct of the housekeeping gene subtracted from the Ct of FABP; ΔΔCt= average ΔCt of each day - average ΔCt of day zero; copy number was calculated based on experimentally determined amplification efficiencies of 91.8% (FABP) and of 92.5% (EF1 α).

Day	ΔCT	ΔΔCT	Relative Expression	% of Maximum
0	6.67	0	0	0
1	-1.84	8.51	365	4
3	-3.77	10.4	1389	14
7	-6.60	13.27	9878	100
8	-6.55	13.22	9541	97
14	-4.45	11.12	2226	23
18	-3.41	10.08	1052	11
21	-1.77	8.44	344	4

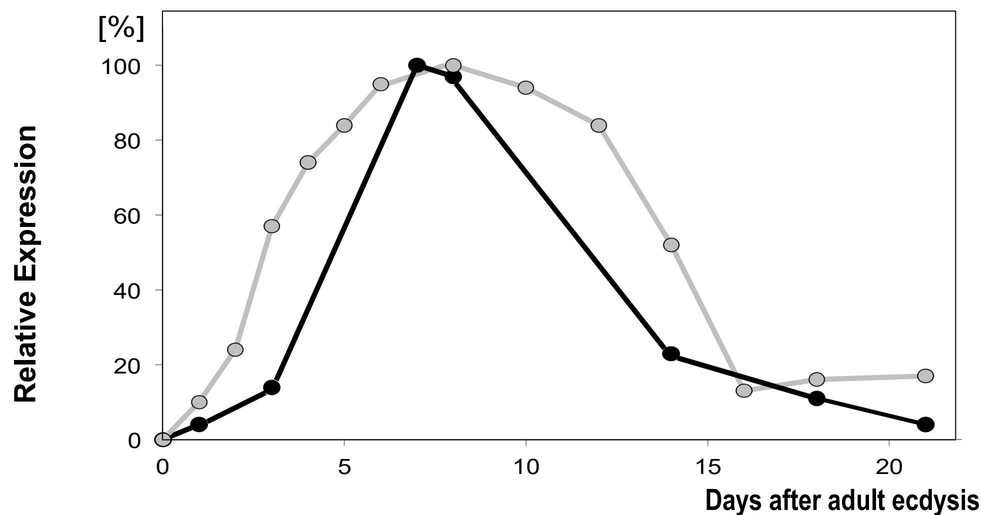


Figure 4-1 FABP expression in adult locusts.
FABP mRNA was quantified by qRT-PCR (black circles), and the expression levels relative to the expression at day 7 after adult eclosion was compared to previously published results from Northern blots (grey circles, Haunerland *et al*., 1992)

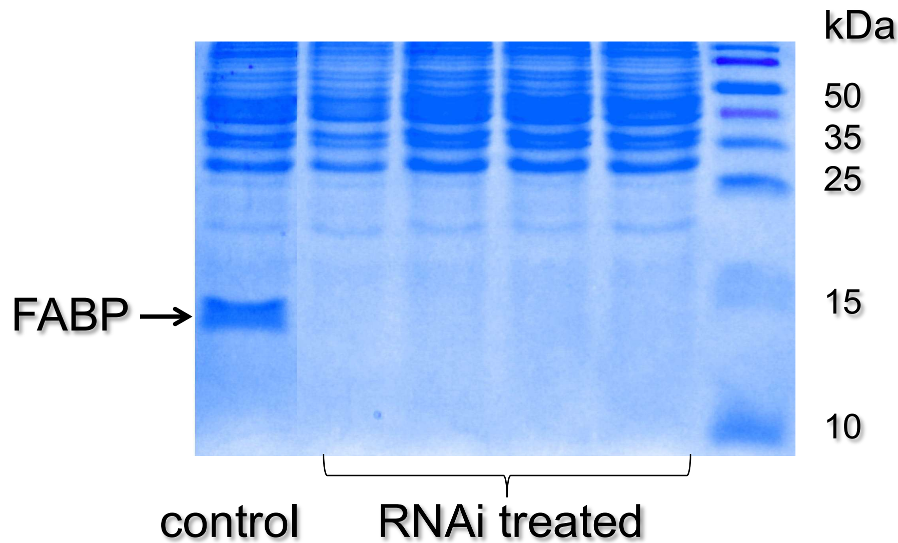


Figure 4-2 SDS PAGE of soluble muscle proteins.
Newly emerged adult locusts were injected with 4 µg FABP RNAi or water (control), and muscle proteins were analyzed 21 days later by SDS-PAGE as described in Materials and Methods.

4.3.2. Flight experiments

The dsRNAi treated insects and control locusts that had been injected with H_2O were mounted in a suspended position on the flight stand, and flight was induced by a headwind of 4 m/s. Insects engaged in uninterrupted flight within 5 minutes were observed at 15 min intervals and scored for their flight performance. As shown in Figure 4-3, there was no noticeable difference between control and treated insect for the first 30 min. Within the following 60 min, however, RNAi treated animals showed a distinct decline, as noticeable in more frequent rest periods and shorter flight durations. Control animals usually remained in their migratory flight position for more than three hours, as shown in Figure 4-4; any insect that stopped flying could be easily stimulated to continue its flight. Attempts to stimulate resting FABP knockdown locusts to continue their flight were rarely successful, and flight did not continue for more than a few minutes. Sometimes these insects bent their forelegs, abdomen, or wings, as shown in Figure 4-5, until their tarsi had contact with some part of the locust body.

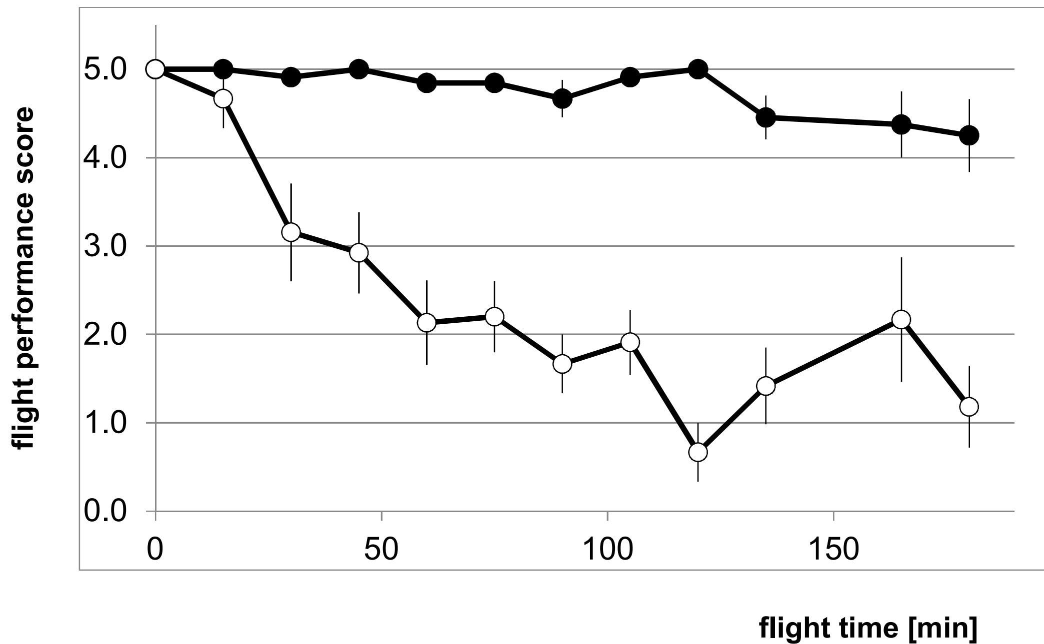


Figure 4-3 Flight performance of adult locusts.
Newly emerged adult locusts were injected with 4 µg FABP dsRNAi or water (control), and subjected to tethered flight 21 days later, as described in Materials and Methods. Flight performance was measured at 15 min intervals and scored on a numerical scale, as follows: 5: continuous, uninterrupted flight; gentle stimulation was required to restart flight of a duration of 4:>5 min; 3: 1-5 min; 2: 10-60 s; 1:<10 s; 0: no flight. Open circles: RNAi treated locusts; closed circles: control locusts.

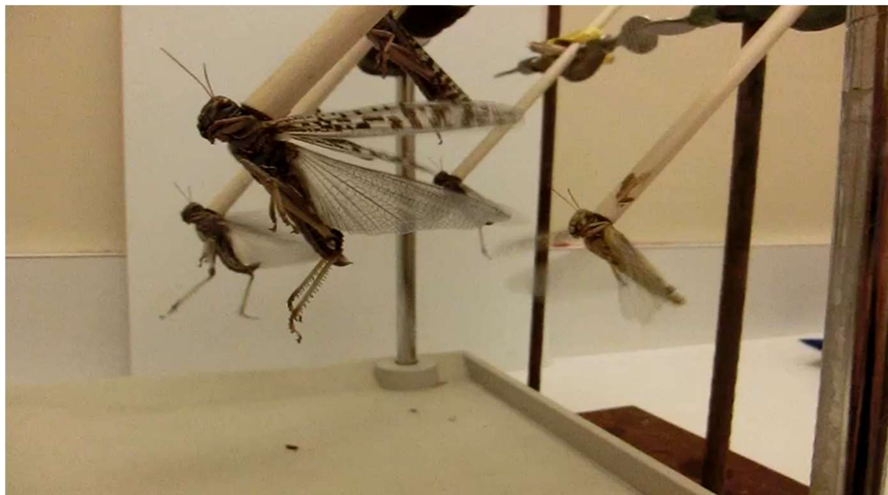


Figure 4-4 Tethered flight of locusts.
Control and dsRNAi treated adult males at 21 days old mounted in a suspended position for induced flying.

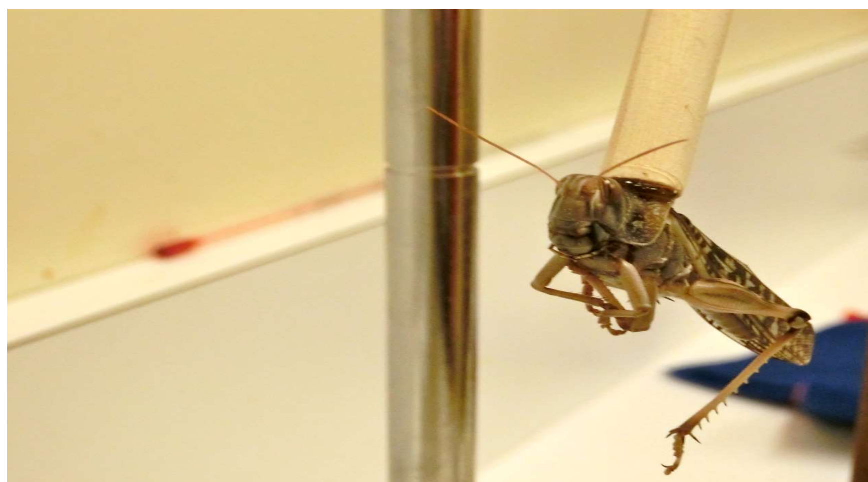


Figure 4-5 Behavioral changes of exhausted FABP knockdown locusts.
The insect folds its appendages around the body with the tarsi touching the ventral-abdominal side.

4.3.3. Lipids content analysis

Flight muscles and hemolymph were collected after the experiments. SDS analysis of flight muscle proteins confirmed the complete absence of FABP in the knockdown animals while showing the strong band at 15 kDa for all controls. Hemolymph samples of control and RNAi treated locusts collected three hours after the initiation of flight were analyzed for their lipid content and compared to animals that had not been allowed to fly (Figure 4-6). By densitometric analysis of thin layer chromatograms, DAG in the hemolymph was normalized to phosphatidylcholine, which does not change when DAG is mobilized during flight. In both, the control and RNAi treated animals, flight resulted in a 4- to 5-fold increase in the amount of DAG over the un-flown insects, confirming that the mobilization of lipids is not affected by RNAi treatment (Table 4-2).

Table 4-2 Hemolymph lipid content of resting and flying locusts.

	Control		RNAi	
	Rest	Flight	Rest	Flight
DAG/PC ratio	1.6±0.2	7.1±1.4	1.6±0.3	6.9±2.4
DAG (-fold increase)	4.4		4.3	

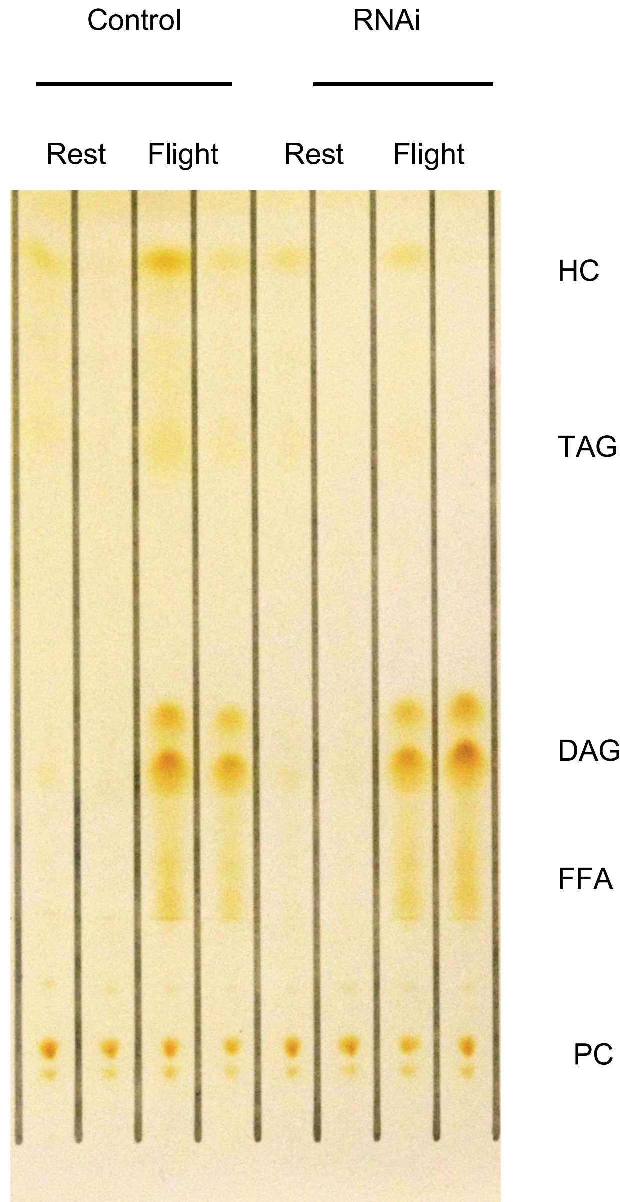


Figure 4-6 TLC separation of hemolymph lipids.
Total lipids were extracted from the hemolymph of locusts at rest and immediately after the flight experiments. Lipids were separated by thin layer chromatography, as described in Material and Methods. The lipids content analysis in the hemolymph shows no difference in the lipid content between the control and dsRNA treated insects. PC: phosphatidylcholine, FFA: Free fatty acid, DAG: diacylglycerol, TAG: Triacylglycerol, HC: Hydrocarbons.

4.4. Discussion

RNA interference (dsRNAi) is used increasingly as an effective tool for functional studies of individual proteins, as this technique allows gene specific degradation of mRNA at any time point in the life of the organism. While its mechanism is well understood, the effectiveness of this technique varies widely, due to variances in delivery, uptake, and stability of the small interfering RNA molecules. In insects, this technique has been used with mixed results, as the effectiveness of the dsRNAi depends on various factors, such as the concentration of dsRNAi, the target species, the corresponding transcript, and the delivery procedure (reviewed by Huvenne and Smagghe, 2010; Terenius *et al*., 2011; Gu and knipple, 2013; Scott *et al.,* 2013). However, studies by the laboratory of Vanden Broeck and others (Wynant *et al*., 2012; Santos *et al*., 2014) have demonstrated that RNA interference works extraordinarily well in locust species. A single injection of small amounts of double-stranded RNAi generally causes a strong systemic response that persists for 10 days or more (Wynant *et al*., 2012; Santos *et al*., 2014), in a variety of tissues including muscle, fat body, midgut, and the central nervous system. Nevertheless, it was not clear at the onset of our study whether it was possible to obtain a similar reduction of FABP, which is the most prominent soluble protein in the flight muscle of mature adult locusts. As muscle FABP expression is specific for adult flight muscle, we injected dsRNAi within 24 h of adult ecdysis, when FABP mRNA was minimal. A single injection of 4 µg dsRNAi very effectively prevented the accumulation of FABP mRNA and of FABP itself. There were no apparent anatomical or behavioral differences during maturation and adult development between the dsRNAi treated and control animals. However, FABP and its mRNA were almost undetectable in treated animals, even three weeks after adult ecdysis when muscle FABP levels normally reach their maximum and insects are fully capable of migratory flight. This confirms that locusts are excellent organisms for dsRNAi studies targeted at prominently expressed proteins.

Because FABP was essentially absent in the flight muscle, treated animals were ideal for testing the physiological significance of FABP. While it has generally been assumed that muscle FABP fulfills an important role in the uptake and transportation of fatty acids, unambiguous proof for this function has never been obtained for any animal species. Our study clearly demonstrates that the absence of FABP in the flight muscle of locusts prevents sustained flight activity, while it can be easily induced in control animals

in a tethered flight setup. Locust flight is initiated when their legs are lifted off the ground, and wind-sensitive hairs on the head detect air movement (Weis-Fogh and Jensen, 1956). Flight is maintained as long as there is sufficient wind resistance (>2 m/s) and the wings continue to oscillate in the wind. All insects used in this study initiated flight once mounted on the flight stand, and, after occasional stops in the first few minutes, engaged in uninterrupted flight for at least 30 minutes. There was no apparent difference between control and treated animals, which could be expected as the initial phase of locust flight is fueled by trehalose and easily mobilized carbohydrate stores. In the following hour, however, after carbohydrate resources had been depleted, and lipids gradually became the sole fuel, the flight performance of insects lacking FABP rapidly declined. Once these insects stopped moving their wings they could initially be stimulated to continue flying for short durations, but eventually, it was apparent that they struggled to keep moving their wings. While there were individual differences in length of time until uninterrupted flight ceased, none of the insects lacking FABP was capable of flying after 2 hours, while all control animals continued to fly for 3 hours and beyond.

Treated insects eventually kept hanging motionless in the flight stand, and occasionally we observed unusual movements, such as bending a foreleg so that it would touch the mouth, or folding one of the inner wings so that it would rest under the body of the insect. It is likely that these motions were aimed at providing tactile contact at the tarsi, in order to suppress the flight reflex which otherwise would force the insects to continue moving their wings, even though no substrate was available to fuel the muscle activity.

The current study clearly demonstrates that FABP plays an essential role in lipid uptake and transport in the flight muscle of locusts, and that its absence prevents long-distance flight activity. FABP may act as a sink for fatty acids in the cytosol, thus maintaining a concentration gradient across the plasma membrane, as well as a transport vehicle to allow for the rapid transport of the hydrophobic fatty acid molecules through the aqueous cytosol. Our findings add to earlier circumstantial evidence from vertebrate and invertebrate studies that suggest a role of FABP in sustained muscle activity, including the correlation between FABP content and lipid-fueled metabolic rates and the observations that FABP expression can be induced by increased lipid supply (Qu *et al*., 2007). Endurance training in rodents, humans, birds, and insects leads to increased FABP expression in skeletal muscle (Haunerland, 1994; Clavel *et al*., 2002; as do pathological conditions with elevated plasma levels (diabetes, obesity) (Maeda *et al*., 2003; Choi *et al*.,

2009; Atshaves *et al*., 2010). We have demonstrated before that endurance flight in fully mature locusts, in spite of its already very high levels of FABP, up-regulates FABP expression in the flight muscle, and that this effect can be mimicked in resting locusts by either inducing lipid release and LDLp production by injecting AKH, or simply injecting externally provided LDLp (Chen and Haunerland, 1994). In the present study, our experiments indicate that AKH signaling and lipid release is unaffected by RNAi and the resulting lack of FABP, and that there were no apparent differences in availability of lipids in the hemolymph: following the initial, carbohydrate fueled phase of flight, lipids were mobilized and released in the hemolymph in both, treated and control animals, with a similar rise in the DAG content, indicating that the flight-induced mobilization of lipids in the fat body and the lipophorin shuttle in the hemolymph are fully functional.

The present study provides the first compelling proof for an essential role of FABP in skeletal muscle energy metabolism in vivo, possible due to advantageous locust model system with the stage-specific, very high levels of FABP, their total dependency on lipid as an energy source for sustained flight, and the extraordinarily strong response to dsRNA treatment. The structural and functional similarities between vertebrate and invertebrate muscle and their FABP structure and regulation suggests that FABP fulfills similar functions in vertebrates as well.

Connecting Statement 4: apolipophorin-III and lipid transport in the hemolymph

Efficient transport of lipids from their storage site, the fat body, to the mitochondria of the flight muscles is necessary for migratory flight of locusts. In chapter 4, it was demonstrated that the intracellular transport of fatty acid through the flight muscle cytosol requires the presence of its transport protein FABP and that FABP is essential for extended flight activity. While the initial, carbohydrate-fuelled phase of flight was not altered, elimination of this protein by RNAi techniques prevented flight activities exceeding 30 minutes, when carbohydrate stores are depleted, and lipids become the dominant energy source. In this chapter, I focus on the extracellular part of the delivery of lipids through the hemolymph, from the fat body to the flight muscle. This transport is mediated by the major hemolymph lipoprotein, lipophorin (Lp). Lipophorin is composed of three apoproteins, apoLp-I, apoLp-II, and apoLp-III. ApoLp-I and II, encoded by a single gene that is part of the LLTP gene family, form the core high-density form of lipophorin (HDLp) that is predominant at rest. During prolonged flight, DAG is released from the fat body and taken up by HDLp, to form the lipid-enriched low-density lipophorin (LDLp) particle which is stabilized by binding to several molecules of the small, water-soluble apoLp-III. This process is reversed at the flight muscle membrane, where DAG is hydrolyzed; free apoLp-III and HDLp remain in the hemolymph, ready to repeat the cycle and pick up more DAG at the fat body. As apoLp-III seems to play an essential role in this process, I investigated the consequences of knocking down this apoprotein. In the next chapter, I describe the primary structure and expression profile of apoLp-III from the desert locust, its knockdown by RNAi techniques, and the effect of this treatment on induced flight experiments.

Chapter 5. Apolipophorin III: sequence, expression pattern, and role in sustained flight

Abstract

The complete cDNA sequence of the hemolymph apolipoprotein, apolipophorin III (apoLp-III) from the desert locust, *Schistocerca gregaria,* has been sequenced using 3` and 5` RACE techniques. The sequencing results yielded a nucleotide sequence of 735 bp with a deduced amino acid sequence of 179 residues. The temporal expression pattern in the fat body was analyzed using real-time qPCR, starting in the last nymphal instar and proceeding to 2 month-old adults. We found that apoLp-III is constitutively expressed in the fat body over the entire period examined. To investigate the role of this gene during sustained flight, dsRNAi of apoLp-III was used to knockdown the corresponding mRNA in the fat body. Unlike control animals, most insects lacking apoLp-III appeared sluggish and were not able to engage in flight activities longer than 5 min, and none of the treated insects could engage in extended flight periods that are known to be fuelled by lipids. Therefore, it appears that apoLp-III is critical not only for effective lipid shuttling during flight but may play additional physiological roles in insects.

5.1. Introduction

In insects, the metabolic rate during flight activities is among the highest levels found in living organisms (Van der Horst, 2003). In locusts, lipids are the preferred substrate to meet this demand during prolonged flight activities; due to the physical characteristics of lipids that hinder their solubility and mobilization in water-based environments, a delivery system is needed for the translocation of lipid molecules from their storage sites towards the energy-consuming muscle tissues. The active translocation of lipids through the hemolymph is especially crucial for long-distance migratory flight, as carried out by the desert locust, *Schistocerca gregaria*, which forms enormous swarms of millions of insects that move for hundred kilometers or more without food intake.

Shortly after the start of sustained flight activity, peptide neurohormones called adipokinetic hormones (AKHs) are released from the *corpora cardiaca* to trigger the conversion of the triacylglycerol (TAG) to diacylglycerol (DAG) in the fat body through a cascade of intracellular reactions (reviewed by Van der Horst *et al.,* 2001). In the hemolymph, DAG is released and incorporated into the major hemolymph lipoprotein, lipophorin. This abundant lipoprotein is composed of two apolipoproteins, apolipophorin I and apolipophorin II, with corresponding molecular weights ~ 250 and ~85 kDa, respectively (Ryan *et al.,* 1984). In the resting state, high-density lipoprotein (HDLp, density ~1.10 g/ml) is the predominant protein, in which these two apoproteins during lipid transport are associated with ~50% lipids, mostly DAG and phospholipids. During flight, DAG released from the fat body associates with the HDLp particle and several molecules of a third, small apolipoprotein, apolipophorin III (apoLp-III), which shields the hydrophobic lipids at the surface of the HDLp particle from the aqueous environment. Due to its increased lipid content and association with apoLp-III, the resulting lipophorin particle that has a lower density (~1.04 g/ml) is called low-density lipophorin (LDLp).

At the flight muscle, the large LDLp disassembles, releasing the DAG, apoLp-III, and returns to the original HDLp. While HDLp and apoLp-III are ready for another cycle of DAG uptake (Weers and Ryan, 2003), the DAG is hydrolyzed at the flight muscles cell membranes into glycerol and free fatty acid. Subsequently, the free fatty acids are taken up by the flight muscle cells and transported intracellularly to the mitochondria for β-oxidation by another binding protein, fatty acid binding protein (FABP) (Haunerland, 1997).

ApoLp-III has a molecular weight of ~20 kDa and belongs to a large family of exchangeable transport proteins (Narayanaswami and Ryan, 2000). Members of this family are characterized by the presence of multiple amphipathic α-helix segments (Segrest *et al*., 1992). In the absence of lipids, they form a globular structure that is stabilized by helix-helix interactions. Upon exposure to fats, the globular helix opens exposing the hydrophobic side to load the lipid molecules inside. Beyond its role in the lipid transport, apoLp-III is thought to be a multifunctional protein with additional, different roles as well, as it was found to be involved in the innate immunity in the larvae of *Galleria mellonella* (Wiesner *et al*., 1997; Dettloff *et al*., 2001a, b; Whitten *et al*., 2004). Also, it has been proposed that apoLp-III plays a role in membrane disintegration of nervous and muscle tissues during developmentally-regulated programmed cell death (Sun *et al*., 1995).

In Chapter 4, we have shown that the efficient transport of lipids within locust flight muscle cells is essential for sustained flight. Knockdown of the flight muscle FABP did not have a noticeable effect on the ability of the insects to engage in flight, as the quickly mobilized carbohydrate trehalose fuels the initial phase of flight. However, these insects could not sustain flight durations exceeding 30 minutes, after which beta-oxidation of fatty acids has become the sole source of energy. As apoLp-III appears to be an essential element of lipid mobilization and delivery of lipid to the flight muscle, reduction or elimination of this protein should affect flight ability in a similar manner.

The present study was carried out to evaluate the effect of reduced, or absent, levels of apoLp-III on the flight ability of the desert locust. We determined the sequence and the temporal expression profile of apoLp-III, knocked down the expression by RNAi techniques, and evaluated the physiological consequences.

5.2. Materials and Methods

5.2.1. Insect treatment and sample collection

As previously described, the desert locust, *Schistocerca gregaria* was kept under a crowded condition at 30 ºC. The male insects were isolated on the first day of adult eclosions and marked on the ventral side for the knockdown experiment. The dsRNA treatment followed the same protocol as described in Chapter 4, with the first dose (10 µg dsRNA) applied on the first day of the adult stage supported by another boosting dose at day 14 after the adult eclosion. The sample collection was carried out at three weeks adult old. For the expression profile study, the fat body from 3-5 untreated female adults were collected at different time points between 0 and 60 days after adult eclosion. Total RNA and complementary cDNA were prepared as previously described in Chapter 2.

5.2.2. RACE libraries

The SMARTer® RACE cDNA Amplification kit (Clontech, cat.: 634860) was used to generate 3` and 5` RACE libraries from the total RNA extracted from the desert locust fat body, following the manufacturer's protocol as described in the second chapter. An optimized PCR kit SeqAmpTM DNA polymerase from Clontech (Cat#: 638504) was used to amplify the 3` and 5` ends of the apoLp-III transcript. The forward and reverse internal gene-specific primers (Table 5-1) were used along with Universal primer mix (UPM) at 65 °C annealing temperature to amplify and sequence the 3`and 5` ends which contain the untranslated regions (UTR), transcription start site (TSS) and the termination signals.

5.2.3. ApoLp-III knockdown

ApoLp-III dsRNA was produced using the primer sets shown in Table 5-1 to knock down the corresponding apoLp-III mRNA using the T7 RiboMAX^{TM} Express RNAi System, Promega (Catalogue # P1700) according to the manufacturer's protocol.

Table 5-1 Primers used to amplify apoLp-III cDNA and synthesize the dsRNAi.

T7 promoter sequence is underlined.

Primers	Sequence 5`-----------3`
ApoLp-III Forward	AGCTGAATCACACCATCGTC
ApoLp-III Reverse	GTCTTGGTTGCGATGTCCT
T7-Sg-ApoLp-III Reverse	GGATCCTAATACGACTCACTATAGGGTCTTGGTTGCGATGTCCT
T7-Sg-ApoLp-III Forward	GGATCCTAATACGACTCACTATAGGGCTGAATCACACCATCGTC

5.2.4. Flight experiments

The sustained flight was provoked by mounting the treated and control insects in front of two large fans as described in detail in Chapter 4, with the same system that was used to estimate the insects' performance during long induced flight.

5.2.5. Gene expression

The real-time qPCR was used to measure the temporal gene expression pattern of apoLp-III and the effect of dsRNAi knockdown. A specific real-time primer set was designed based on the obtained sequence (table 5-2). The efficiency of the primer set was assessed using a series of dilution of the amplified amplicon. GAPDH and EF1α were used as an internal housekeeping gene to normalize the expression of apoLp-III. The ΔCt method was used to assess the expression pattern of apoLp-III in the fat body at different time intervals as described in chapter 2. The $\Delta\Delta$Ct method (Livak and Schmittgen, 2001) as detailed in the third chapter was used to test the effect of the dsRNA treatment.

Table 5-2 Real-time primers used to assess apoLp-III expression level

Name	Primer 5`-----------3`	Efficiency
ApoLp-III-Forward	CTGAATCACACCATCGTCAAC	94.3%
ApoLp-III-Reverse	GATCTTGGTCTTGAAGGCATTAG	
GAPDH- Forward	CCCAATGTATGTTGTTGGTGTAA	92%
GAPDH- Reverse	GTGCCAGGCAATTTGTAGTG	
EF1α- Forward	GATGCTCCAGGCCACAGAGA	92.5%
EF1α- Reverse	TGCACAGTCGGCCTGTGAT	

5.3. Results

5.3.1. cDNA nucleotide sequence of apoLp-III

The apoLp-III cDNA sequence from the migratory locust, *Locusta migratoria* (Genbank accession J03888.1; Kanost *et al.*, 1988) was used to design primers in an attempt to obtain a partial sequence of the homologous gene from the fat body of the desert locust. Gene-specific forward and reverse primers (Table 5-1) were used in a PCR reaction against ~200 ng total RNA extracted from the adult male fat body. The PCR reaction yielded a strong band of the anticipated size ~ 300 bp (Figure 5-1). The PCR product was excised from the gel and sent for sequencing. The sequencing result was confirmed to be apoLp-III, with 88% identity after alignment with apoLp-III from the migratory locust.

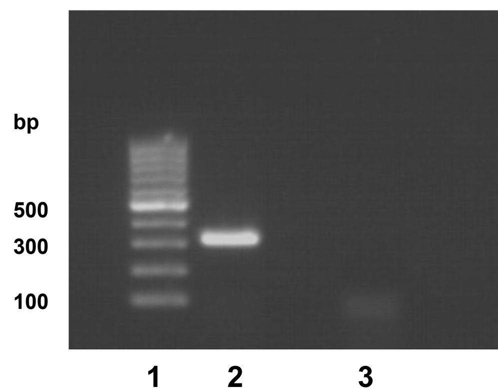


Figure 5-1 Amplification of apoLp-III using primers from the migratory locust. cDNA from adult male locust was used to amplify apoLp-III. Lane 1: 100 bp DNA ladder, Lane 2: ~300 bp of the amplified amplicon, Lane 3: no template control

To get the complete cDNA sequence, the 3` and 5` RACE techniques were used to amplify the 3` and 5`ends of the apoLp-III. The universal primers that anneal with the ends of the RACE libraries were used with the internal gene-specific primers to amplify and sequence the remeaning nucleotides in addition to the 3` and 5` UTRs (Appendix C; sequence C1). The 735 bp complete cDNA nucleotide sequence (Figure 5-2) contains an ORF of 537 bp that starts with the ATG start codon at the position +64. The polyadenylation signal (AATAAA) was identified in the 3` UTR following the termination signal TGA with 52 bp and 46 bp before the poly A tail.

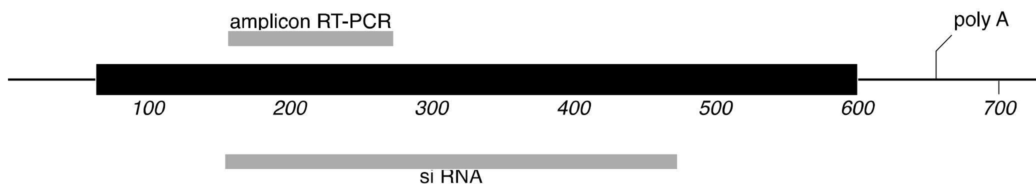


Figure 5-2 An illustration of the complete cDNA sequence of *S. gregaria* apoLp-III
The graph illustrates the coding sequence and the targeted areas for real-time PCR and dsRNAi.

5.3.2. Deduced amino acid sequence

The predicted ORF codes for 179 amino acids (Figure 5-3) with a molecular weight of 19.237 kDa and an isoelectric point 5.30. The Predisi server online website (http://www.predisi.de/home.html) was used to predict potential cleavage sites of the signal peptide which is removed from the polypeptide sequences of the excretory proteins; the highest possibility is after 16 residues, with a nother less likely alternative after residue 18.

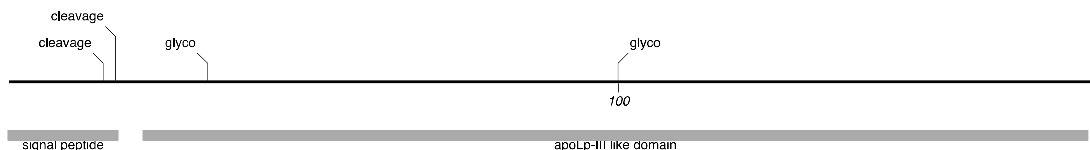


Figure 5-3 Schematic graph of the deduced amino acids sequence of apoLp-III.
The signal peptide cleavage sites, glycosylation sites, and the apoLp-III like domain were depicted on the schematic graph of apoLp-III.

The deduced amino acids sequence was found to include the typical conserved apolipophorin-III like-domain between the residues (23-176) which are observed in similar exchangeable apolipoproteins family members. The sequence was also found to comprise two potential glycosylated sites at the amino acids residues 33 and 100.

The sequence is highly similar to apoLp-III from *Locusta migratoria*, with 88% sequence identity and an additional 8% of conservative substitutions. Van der Horst *et al*. (1991) reported the existence of two isoforms in *L. migratoria*, apoLp-IIIa, and apoLp-IIIb,

distinguished by the cleavage site of the N-terminal signal peptide. While the authors did not report the entire sequence for apoLp-IIIb, a later submission to Genbank (accession P10762) shows identical sequence with apoLp-IIIa except for residues 95, 111-114, and 133-141 (Figure 5-4).

```
S.g.apoLp-III     1 MKTLLAVLMLAVACQARPDA-GQVNIAETVQQLNHTIVNAAHELRETLGLPTQDEALNLL  59
L.m.apoLp-IIIa    1 MNTLLAVLMLAVAAQARPDAAGHVNIAEAVQQLNHTIVNAAHELHETLGLPTPDEALNLL  60
L.m.apoLp-IIIb    1 MNTLLAVLMLAVAAQARPDAAGHVNIAEAVQQLNHTIVNAAHELHETLGLPTPDEALNLL  60
                    * ***********.****** *.*****.***************.******* *******

S.g.apoLp-III    60 TEQANAFKTKIAEVTSTLKQEAEKHQGTVSEQLNAFARNLNNSIHDAATSLNLEEQLNSL 119
L.m.apoLp-IIIa   61 TEQANAFKTKIAEVTTSLKQEAEKHQGSVAEQLNAFARNLNNSIHDAATSLNLQDQLNSL 120
L.m.apoLp-IIIb   61 TEQANAFKTKIAEVTTSLKQEAEKHQGSVAEQLNRFARNLNNSIHDAATSAQPADQLNSL 120
                    ***************..**********.*.**** *************** .  .*****

S.g.apoLp-III   120 QSALNNVGHQWQDIASKTQAAAQDAWGPVQSALQEAAEKTKQAAANLQNSIQSAVQKPAN 179
L.m.apoLp-IIIa  121 QSALTNVGHQWQDIATKTQASAQEAWAPVQSALQEAAEKTKEAAANLQNSIQSAVQKPAN 180
L.m.apoLp-IIIb  121 QSALTNVGHQWQTSQPR-PSVAQEAWAPVQSALQEAAEKTKEAAANLQNSIQSAVQKPAN 179
                    ****.*******    .  . **.** **************.******************
```

Figure 5-4 Alignment of locust apoLp-III sequences

While no information is provided about the method used to obtain this sequence, the fact that the sequences of apoLp-IIIa and apoLp-IIIb are otherwise identical makes it very unlikely that these forms originate from different genes. However, it could be possible that the differences are the result of alternative splicing of the primary transcript, with the replacement of one exon by a different one contained within the same genomic sequence. To investigate this possibility, we assembled the *Locusta* apoLp-III gene following Blast searches of both isoforms against the *Locusta* genome database. Only one gene was found, yielding a transcript that is 100% identical to the reported cDNA and the corresponding protein sequence for apoLp-IIIa. The gene spans over 16 kb contained in contigs AVCP010784081-AVCP010784078, as shown in Figure 5-5. Thus, the only difference between the two isoforms seems to be the posttranslational cleavage of the signal peptide, occurring at either of the two predicted cleavage sites at a position which of the two possible cleavage sites for the signal peptide.

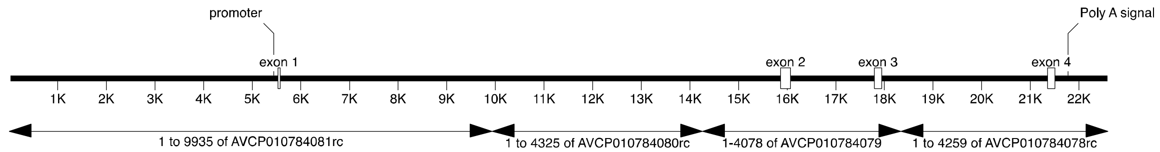


Figure 5-5 Graphical illustration of the *L. migratoria* apoLp-III gene.
The gene assembly displays the distribution of the different apoLp-III exons in the genomic DNA.

The sequence coding for residues 95 and 111-114 is contained in exon 3, while exon 4 contains residue 133-141. Since neither the second nor the third intron contains any sequence similarity to the coding sequence for apoLp-III, one can exclude the possibility of alternative splicing. Instead, a close inspection of the coding sequence suggests that the apoLp-IIIb sequence published in Genbank (accession P10762) contains a number of simple sequencing errors that through frameshifts and substitutions result in the observed differences, as shown in Figure 5-6.

```
apoLp-IIIa

130                  150     360        370        380            430        440        450        460 bp
GlnLeuAsnAlaPheAlaArgAs      AlaThrSer LeuAsnLeuGlnAspGlnLeu      TrpGlnAspIleAlaThrLysThrGlnAlaSerAlaGln
CAGCTGAACGCGTTCGCGCGCAA      GCCACGTCG CTCAACCTGCAGGACCAGCTC      TGGCAGGACATCGCAACCAAGACCCAGGCGTCCGCCCAG
CAGCTGAACCGGTTCGCGCGCAA      GCCACGTCGGCTCAACCTGCA-GACCAGCTC      TGGCAG-ACATCGCAACCAAGACCCA-GCGTC-GCCCAG
GlnLeuAsnArgPheAlaArgAs      AlaThrSer AlaGlnProAlaAspGlnLeu      TrpGln ThrSerGlnProArgProS erVal AlaGln
AA       95                          110             115          131           135            140

apoLp-IIIb
```

Figure 5-6 Alignment of the nucleotide sequences and their deduced amino acid sequences between apoLp-III a & b from *L. migratoria*
Shown are the 5 potential sequencing errors that could result in the different sequence for aopLp-IIIb

5.3.3. Temporal expression analysis

ApoLp-III expression in the fat body was measured and analyzed using real-time qPCR. Specific real-time qPCR primers yielding a small target amplicon were designed from the generated sequence. The efficiency of these primers was calculated from real-time qPCR runs against a dilution series of the target amplicon, using the Bio-Rad program software, as previously described (Chapter 2). The GAPDH was used as a reference gene to normalize the expression of apoLp-III using ΔCt method.

As illustrated in Figure 5-7, apoLp-III was detected in both nymphs and throughout adult life. The expression was the lowest in the immature stage and the first week of the adult stage compared to the older adults. Although there was a relative variability between different insects of the same age, we found an increase over the first three weeks of adult life.

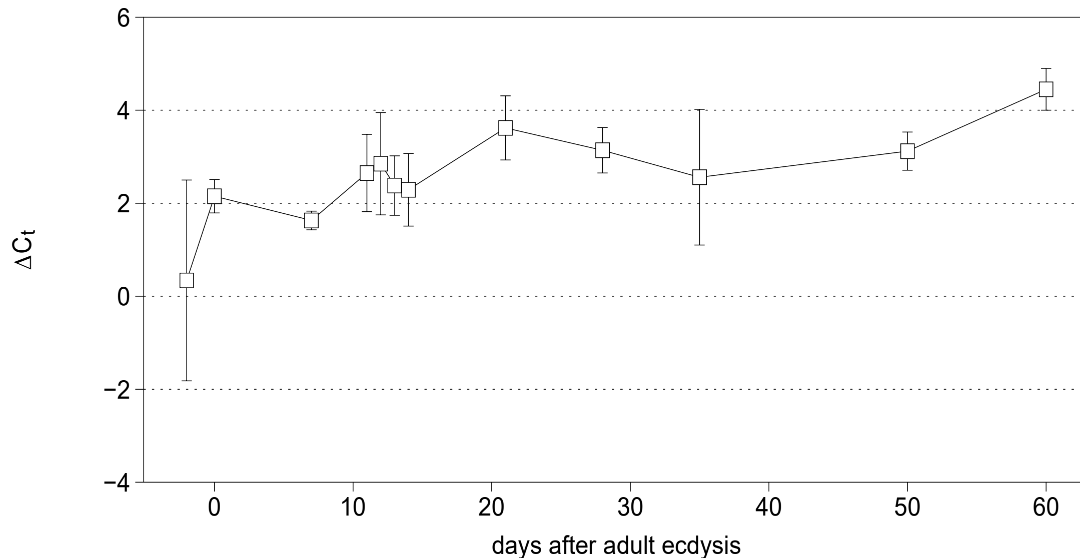


Figure 5-7 The expression of ApoLp-III measured using real-time qPCR in the total RNA of the fat body.
The average ΔCt value ± S.D. of the mean of 3-5 control untreated insects used to represent the expression level at each time point.

5.3.4. Knockdown of apoLp-III expression in the fat body

To knockdown the expression of apoLp-III in the fat body, 10 µg of dsRNAi of the apoLp-III was prepared and injected into the hemocoel of adult males, starting on the first day after adult eclosion. Repeat injection after 7 days, as carried out successfully for the nuclear transcription factors RXR and Met (chapter 3) resulted in 90% mortality. Therefore, repeat injection with the same amount of dsRNAi was carried out at day 14 instead of day 7; this injection was tolerated well by the insects and did not result in any mortality.

The fat body was collected one week after the last injection at 21 days adult old from treated and control insects. Real-time qPCR was used to confirm the apoLp-III

knockdown in the fat body. Both GAPDH and EF1a were tested as reference genes; both remained stable through development and did not seem to be affected by the knockdown and flight experiments carried out (Table 5-3). The Livac method (Livak and Schmittgen, 2001) was used to calculate the expression differences between the treated and control samples. One-way ANOVA was used to test the statistical significance of the differences between the values obtained with the different housekeeping genes. The corresponding calculated P-values are 0.142 and 0.523 for GAPDH and EF1α, respectively, suggesting that the treatments are not significantly affecting the expression of these internal controls.

Table 5-3 Ct values of the housekeeping genes GAPDH and EF1α in control and treated insects.

The numbers represent the average Ct values ± S.D of each treatment group: control males without flying (n=5), dsRNAi treated insects without flying (n=4), control after flying for two hours (n=6), and treated males engaged in a flying experiment (n=6), "n" represents the number of insects in each experimental group.

Experiment	GADPH	EF1α
Control (H_2O)	21.20±0.92	21.16±1.60
apoLp-III dsRNAi	20.24±0.67	19.69±0.39
Flight (H_2O)	21.03±0.72	20.71±1.00
Flight treated	20.43±0.53	20.29±0.48
Average	20.73±0.46	20.46±0.63

We have found that the corresponding expression of apoLp-III was suppressed to 2% of its normal expression (Table 5-4), as assessed with control animals treated in the same manner with H_2O. All insects survived the treatments, without any apparent phenotypical differences.

Table 5-4 The relative expression of apoLp-III in the treated insects with apoLp-III dsRNA.

The apoLp-III expression is normalized to the internal housekeeping genes using the ΔΔCt method. % of the control shows the expression percentage of treated insects (n=4) relative to that of the control (n=6), where n is the number of samples used in each experiment. One-way ANOVA and Tukey HSD were carried out using ΔCt. Significance was considered if P< 0.05*, P<0.01**; P<0.001***

Experiment	GAPDH Normalized	EF1α Normalized
Ct H_2O	19.93±1.59	19.93±1.59
ΔCt H_2O	-1.27±0.80	-1.68±0.22
Ct RNAi	24.74±0.39	24.74±0.39
ΔCt RNAi	4.50±0.66**	5.05±0.56**
ΔΔCt	5.77	6.73
% of control	1.83%	0.94%

5.3.5. Flight experiments

Flight experiments analogous to those reported for FABP (Chapter 4) were carried out to investigate the role of apoLp-III in transferring lipids during flight. Insects were tethered in an elevated position in front of strong fans as described in Chapter 4, and their flight ability was measured at set intervals over a period of 2 h.

All control insects were able to fly regularly for the two hours without much struggle. They commenced flight immediately after being mounted to the flight stand and flew without interruption for at least 30 min, after which they occasionally needed to be gently stimulated by nudging them with a pencil tip to re-engage in flight, as observed in our earlier experiments. Insects with reduced amounts of apoLp-III, however, were not able to participate in the flight experiment for more than five minutes maximally, with one exception of one treated insect which was able to fly for around 25 minutes but with struggling.

After the flight, we confirmed the knockdown of apoLp-III in all treated insects by real-time qPCR (Table 5-5). We collected the fat body from the examined treated and control insects and extracted the total RNA from them. We found that the expression of apoLp-III for each of the control insects was clearly elevated after flight (2-3-fold), but apoLp-III remained essentially absent in the RNAi treated insects, even after the unsuccessful attempts to engage these insects in flight.

Table 5-5 The relative expression of apoLp-III after flight experiment.

The apoLp-III expression is normalized to GAPDH using the $\Delta\Delta$Ct method. % of the control shows the expression percentage of apoLp-III after flying relative to that of the control at rest without flying. Four to six insects were used in each experimental group. One-way ANOVA and Tukey HSD were carried out using ΔCt. Significance was considered if $P< 0.05^{*}$, P<0.01**; P<0.001***. No significance after flying.

Experiment	H_2O	dsRNA
Ct rest	19.93±1.59	24.74±0.39
ΔCt rest	-1.27±0.80	4.50±0.66
Ct flight	18.33±1.45	24.70±0.51
ΔCt flight	2.70±0.86	4.24±0.66
ΔΔCt	-1.43	-0.24
% of control	269%	118%

5.4. Discussion

ApoLp-III is a small, abundant protein in insect hemolymph with a molecular weight of ~20 kDa (Ogoyi *et al.,* 1995). It plays a significant role in transferring lipids between different tissues. Its role was first discovered in the hemolymph of the migratory locust, *Locusta migratoria,* and subsequently identified, isolated and further studied in other numerous insect species (Shapiro and Law, 1983; Van der Horst *et al.,* 1984; Haunerland *et al*., 1986; Ryan *et al*., 1990; Smith *et al*., 1994; Weise *et al*., 1998; Chung and Ourth, 2002).

Only one manuscript described apoLp-III from the desert locust *Schistocerca gregaria*, however without any sequence information (Ogoyi *et al*., 1995). In our study, the *S. gregaria* apoLp-III nucleotide sequence was amplified and sequenced from fat body total RNA using the 5` and 3` RACE method. The sequencing results revealed a complete cDNA sequence of 735 bp. It contains a starting codon ATG at the position +64 and the stop codon TGA at the position +601 with an ORF 537 bp, coding for a protein of 179 residues with a typical amino-terminal signal sequence that targets the protein for export into the hemolymph. Two potential cleavage sites for the signal peptide were predicted, which would result in a mature protein of 163 or 161 amino acids, with a molecular weight of ~17,579 kDa or 17,326 kDa, respectively. As expected, the protein is highly similar to its ortholog from *Locusta migratoria,* which was first sequenced from its cDNA (Kanost *et*

al., 1988); the primary and tertiary structure of the mature protein was later confirmed by x-ray crystallography and NMR spectroscopy (Breiter *et al*., 1991; Wang *et al*., 1997; Fan *et al.,* 2003). Van der Horst *et al*. (1991) reported that two isoforms of apolipophorin exist in *Locusta migratoria*, distinguished by different cleavage of the signal peptide, without an effect on the lipid binding ability (Van der Horst *et al.,* 1991; Weers *et al*., 1993). The reported amino acid sequence of these isoforms, however, is also different in various positions in the carboxy-terminal area. We have searched for apoLp-III sequences in the *Locusta migratoria* genome and transcriptome and found only one sequence, identical with the originally reported apoLp-IIIa, and conclude that the reported sequence for apoLp-IIIb contains some errors, most likely due to simple sequencing errors in GC rich areas of the cDNA. *S. gregaria* apoLp-III has two potential sites for N-glycosylation, in identical positions where the *L. migratoria* protein is glycosylated; the mass of the native protein from the migratory locust was found to contain ~14% carbohydrate with unique carbohydrate moieties linked to the residues 18 and 85 (Hård *et al*., 1993). The function of the associated sugar moieties with apoLp-III is still unclear. Despite the similarity in function of apoLp-III in different insect species, lepidopteran insects are not glycosylated, while glycosylation was found in three additional hemimetabolous orthopteran species (Weers and Ryan, 2003).

The protein structure of apoLp-III was studied in several insect species both by X-ray crystallography and by NMR spectroscopy, which is virtually identical to the 3-D protein structure for the *S. gregaria* protein predicted by SWISS-MODEL website (https://swissmodel.expasy.org/). The protein is composed of 5 elongated α-helices (Fig. 5-8), which is similar to that of the migratory locust (see; Fan *et al*., 2003), a finding not surprising given the high degree of sequence identity between the two polypeptide sequences in the close evolutionary related insect species. The previous comparison by Weers and Ryan (2003) between the 3-D structure of this protein in the migratory locust and the tobacco hornworm, *Manduca sexta*, showed a high degree of structural resemblance despite the relatively low degree of similarity between the amino acid sequences. These authors found that both proteins are composed of 5-helices, similar to our model of the apoLp-III structure of the desert locust. In addition to the main 5-helices, in *M. sexta* there was another short helix in between the third and the fourth helices, and it was linked to a lipid initiation binding (Wang *et al*., 1997; Narayanaswami *et al*., 1999).

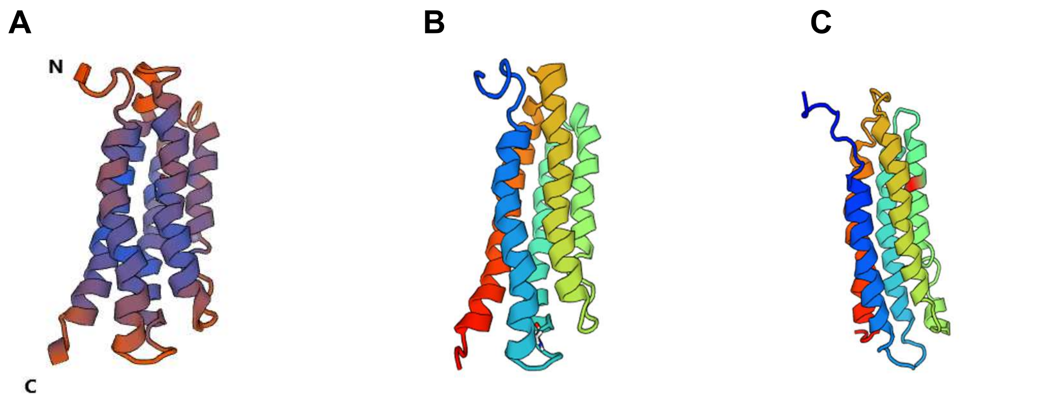


Figure5-8 3-D structures of insect apoLp-III
A: Predicted 3-D structure of apoLp-III of *S. gregaria* compared to B: the NMR structure of apoLp-III in *L. migratoria* (Fan *et al*., 2003) and C: NMR of *Manduca sexta* apoLp-III (Wang *et al*., 2002)

ApoLp-III is constitutively expressed in adult and immature locusts, however at different levels. We measured the amount of its mRNA in the fat body by quantitative real-time PCR, from the last nymphal stage throughout the adult lifespan from the first day of the adult eclosion until 60 days later. In the nymphal instar, the expression of apoLp-III was lower than in the adult; the mRNA started to rise after adult eclosion until it reached the maximum level around day 21, and it remained high throughout the adult stadium.

In general, our results are in good agreement with the findings for mRNA expression and the protein level in the hemolymph of the migratory locust, *Locusta migratoria* (De Winther *et al*., 1996). In the migratory locust, the concentration of the apoLp-III in the hemolymph was at the highest level at day 19 after the adult ecdysis with 17 mg/ml, while the concentration gradually decreased with aging, to a low of 6 mg/ml at day 42. The mRNA expression pattern in the fat body was similar to that of the corresponding protein in the hemolymph, and the maximum level was reported just before the rise of the protein concentration.

The rise of apoLp-III during adult maturation reflects its primary function in the transport of lipids during prolonged flight. The insects depend on the lipophorin shuttle system, in which HDLp takes up mobilized lipids and forms together with 9 molecules of apoLp-III the low-density form of this lipoprotein (LDLp). At the flight muscle where lipids are needed as energy to fuel flight, diacylglycerols are hydrolyzed to glycerol and free fatty

acids, which in turn enter the flight muscle cell. ApoLp-III dissociates from the lipid-depleted particle, which again becomes HDLp, and the cycle of uptake at the fat body and release at the flight muscle can start again (Beenakkers *et al*., 1985; Wells *et al*., 1987; Shapiro *et al*., 1988; Van der Horst, 1990; Ryan, 1990; Blacklock and Ryan, 1994).

We tested the role of apoLp-III in transporting the lipids from the fat body to the muscle tissues during sustained muscles activities experimentally, by knocking down the expression of apoLp-III in adult males. While these knockdown experiments were similarly effective as the experiments described earlier in this thesis, resulting in >98% reduction of apoLp-III mRNA, the treatment had a noticeable impact on the health and fitness of the treated animal, unlike treatment with dsRNAi for FABP, RXR, and Met. Following initial treatment just after adult ecdysis, the insects did not tolerate a second injection 7 days later, requiring us to postpone the second treatment by 7 days and administer it at two weeks old adults. The insects appeared sluggish, moving slowly and failing to quickly react to stimuli.

At three weeks old, we attempted to engage these insects into extended flight activities, as previously done for FABP-depleted males. While the control animals were able to fly easily for more than 2 hours, with the occasional stimulation to re-engage insects in continuous flight, treated insects were not able to fly more than 5 min. This was unexpected as in the initial phase of flight, hemolymph carbohydrates are the main energy source for the flight muscle, and lipids become the predominant substrate only after 30 min of flight. The fact that the treated insects were unable to engage even in carbohydrate-fueled flight suggests that knockdown of apoLp-III had other effects than only preventing the lipophorin shuttle during flight. Taken together with the lethargic behavior of the animals, these findings could mean that apoLp-III is also required for the uptake and storage of nutrients in flightless stages, or point to other, different roles for this protein. Indeed, others have suggested that apoLp-III plays different roles in the innate immunity in various insect species, in a comparable manner as their mammalian homologs analogs such as apolipoprotein E (apo-E) (see: Cole *et al*., 1987; Witten *et al*., 2004). These were identified as pattern recognition molecules (Peiser *et al*., 2002; Yu *et al*., 2002; Witten *et al*., 2004), as well as participants in the Lipopolysaccharide (LPS) detoxification mechanism (Kato *et al.*, 1994; Feingold *et al*., 1995). Additionally, they were found to enhance the phagocytic activity of the blood cells (Wiesner *et al*., 1997; Carvalho *et al*., 2000).

While it is impossible from our study to predict in which pathways apoLp-III involved, it indicates more roles than only to enable effective lipid transport, and it is clear that the apoLp-III protein is essential for the overall fitness of locusts. Thus, it could provide another potential target for insect control strategies.

Chapter 6. General Conclusions

In this thesis, I studied intra- and extracellular proteins involved in lipid transport in the desert locust, *Schistocerca gregaria*. I used quantitative real-time PCR techniques to measure the expression of these proteins and their potential regulators, and RNA inference to knockdown these genes. Finally, functional studies were carried out with respect to lipid mobilization and utilization during flight.

I was able to obtain the full-length sequences of the two vitellogenins, VG-A and VG-B, expressed in the adult female of *S. gregaria*; these two proteins were identified in *Locusta migratoria* more than 30 years ago (Wyatt *et al*., 1986; Locke *et al*., 1987), but to date, only partial sequence information for these proteins has been obtained. By searching the draft of the genome and transcriptome data from *L. migratoria,* the complete cDNA sequence for VG-A and VG-B from this species was deduced, and the entire genomic sequence for both proteins was established. Sequence similarity between the two locust species was very high (>80%), as these species are evolutionary very close. Therefore, I believe that our regulatory and functional conclusions likely are valid for both locust species.

I demonstrated that both proteins are co-expressed in similar amounts in the fat body of mature females, commencing approximately 11 days after adult eclosion, and continuing at high levels throughout the entire adult life. It has long been knowing that the expression of VGs is under the control of juvenile hormone (Chen *et al*., 1979; Chinzei *et al.,* 1982; Dhadialla and Wyatt, 1983), and potential receptors for JH have been postulated (Riddiford, 2008; Parthasarathy *et al.,* 2010; Charles *et al*., 2011; Hult *et al.,* 2015). We were able to partially knockdown two of the candidate receptors, Met and RXR. Reduction of Met expression by 85% during the first two weeks post adult eclosion resulted in the complete disappearance of vitellogenins, reducing the levels of VG-A and VG-B to less than 0.04% of their control values. In contrast, RXR knockdown resulted in a smaller reduction of vitellogenins, to approx. 4% of their control levels, suggesting that this transcription factor may act in a less specific manner than Met.

The cytosolic fatty acid binding protein FABP shows a similar expression profile in adult muscle as VG in the fat body, albeit in both males and females. We were able to directly knockdown this strongly expressed protein by RNA inference, reducing its levels to < 2% of what is normally found in adults 3 weeks after adult eclosion. In a series of flight experiments, we found that in the absence of FABP the insects are incapable of engaging

in flight longer than 30 min, i.e., the time when carbohydrate resources have been depleted, and the locusts switch to lipids as the sole fuel for energy production in the muscle. Short-term flight performance of FABP knockdown locusts was identical to control insects, suggesting that the lack of FABP does not interfere with carbohydrate metabolism. Moreover, the mobilization of lipids in the fat body and their transport through the hemolymph was indistinguishable in FABP knockdown and control animals. Thus, our study provides conclusive evidence for an essential role of FABP in energy substrate transport in the flight muscle for migratory flight activity.

Prevention of lipid mobilization and delivery to the flight muscle should have a similar effect on migratory flight as the blockage of intracellular transport. In order to limit lipid delivery through the hemolymph, I knocked down apolipophorin III (apoLp-III), the apolipoprotein that is necessary for the formation of the DA-enriched low-density lipoprotein (LDLp) that transports lipids from the fat body to the flight muscle. Removal of apoLp-III did prevent extended flight, but also affected short-term flight activities and resulted in animals that appeared much less active than control insects. Therefore, it is likely that apoLp-III plays other essential roles in the life of the insects, as it has been suggested before by various researchers (Sun *et al*., 1995; Wiesner *et al*., 1997; Dettloff *et al.,* 2001a, b; Whitten *et al*., 2004).

Taken together, my work has highlighted the essential role of lipid transport proteins for locust reproduction and dispersal and has identified possible targets for strategies that may mitigate the damage done to human habitats by these highly successful insect species.

References

Adams, T. S., & Filipi, P. A. (1988). Interaction between juvenile hormone, 20-hydroxyecdysone, the corpus cardiacum-allatum complex, and the ovaries in regulating vitellogenin levels in the housefly, *Musca domestica*. *Journal of Insect Physiology*, *34*, 11–19.

Agui, N., Shimada, T., Izumi, S., & Tomino, S. (1991). Hormonal control of vitellogenin mRNA levels in the male and female housefly, *Musca domestica*. *Journal of Insect Physiology*, *37*, 383–390.

Al-hadi, H. A., William, B., & Fox, K. A. (2009). Serum Level of Heart-Type Fatty Acid-Binding Protein in Patients with Chronic Renal Failure. *Sultan Qaboos University Medical Journal, 9,* 311–314.

Anderson, T. A., Levitt, D. G., & Banaszak, L. J. (1998). The structural basis of lipid interactions in lipovitellin, a soluble lipoprotein. *Structure*, *6*, 895–909.

Arrese, E. L., & Soulages, J.L. (2010). Insect Fat Body: Energy, Metabolism, and Regulation. *Annual Review of Entomology*, *55*, 207–225.

Ashok, M., Turner, C., & Wilson, T. G. (1998). Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators. *Proceedings of the National Academy of Sciences*, *95*, 2761–2766.

Athenstaedt, K., & Daum, G. (2006). The life cycle of neutral lipids: Synthesis, storage and degradation. *Cellular and Molecular Life Sciences*, *63*, 1355–1369.

Atshaves, B. P., Martin, G. G., Hostetler, H. A., McIntosh, A. L., Kier, A. B., & Schroeder, F0. (2010). Liver fatty acid-binding protein and obesity. *The Journal of Nutritional Biochemistry, 21,* 1015–1032.

Azeez, O. I., Meintjes, R., & Chamunorwa, J. P. (2014). Fat body, fat pad and adipose tissues in invertebrates and vertebrates: the nexus. *Lipids in Health and Disease*, *13*, 71. https://doi.org/10.1186/1476-511X-13-71

Azevedo, D. O., de Paula, S. O., Zanuncio, J. C., Martinez, L. C., & Serrão, J. E. (2016). Juvenile hormone downregulates vitellogenin production in *Ectatomma tuberculatum* (Hymenoptera: Formicidae) sterile workers. *The Journal of Experimental Biology*, *219*, 103–108. https://doi.org/10.1242/jeb.127712

Badisco, L., Huybrechts, J., Simonet, G., Verlinden, H., Marchal, E., Huybrechts, R., Schoofs, L., de Loof, A., & Vanden Broeck, J. (2011). Transcriptome analysis of the desert locust central nervous system: Production and annotation of a Schistocerca gregaria EST database. *PLoS ONE, 6,* https://doi.org/10.1371/journal.pone.0017274

Baillie, A. G. S., Coburn, C. T., & Abumrad, N. A. (1996). Reversible binding of long-chain patty acids to purified FAT, the adipose CD36 homolog. Journal of Membrane Biology, *153,* 75–81.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., & Jones, A. T. (1994). Lipid-Binding Proteins: A Family of Fatty Acid and Retinoid Transport Proteins, *45*, 89–151.

Barr, P. J. (1991). Mammalian subtilisins: The long-sought dibasic processing endoproteases. Cell, *66,* 1–3.

Beenakkers, M. T., Van der Horst, D. J., & Van Marrewijk, W. J. A. (1985). Insect lipids and lipoproteins, and their role in physiological processes. *Progress in Lipid Research*, *24*, 19–67.

Belles, X. (1998). Endocrine effectors in insect vitellogenesis. In: *Recent advances in Arthropod Endocrinology* (Coast, G. M. & Webster, S.G. eds). Cambridge University Press, Cambridge, pp. 71–90.

Belles, X., & Piulachs, M. D. (2015). Ecdysone signalling and ovarian development in insects: From stem cells to ovarian follicle formation. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms, 1849,* 181–186.

Binas, B., & Erol, E. (2007). FABPs as determinants of myocellular and hepatic fuel metabolism. *Molecular and Cellular Biochemistry, 299,* 75–84.

Bitra, K., & Palli, S. R. (2009). Interaction of proteins involved in ecdysone and juvenile hormone signal transduction. *Archives of Insect Biochemistry and Physiology*, *70*, 90–105.

Blacklock, B. J., & Ryan, R. O. (1994). Hemolymph lipid transport. *Insect Biochemistry and Molecular Biology*, *24*, 855–873.

Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology*, *37*, 911–917.

Borst, D. W., Eskew, M. R., Wagner, S. J., Shores, K., Hunter, J., Luker, L., Hatle, J. D., & Hecht, L. B. (2000). Quantification of juvenile hormone III, vitellogenin, and vitellogenin-mRNA during the oviposition cycle of the lubber grasshopper. *Insect Biochemistry and Molecular Biology*, *30*, 813–819.

Breiter, D. R., Kanost, M. R., Benning, M. M., Wesenberg, G., Law, J. H., Wells, M. A., Rayment, I., & Holden, H. M. (1991). Molecular structure of an apolipoprotein determined at 2.5-.ANG. resolution. *Biochemistry*, *30*, 603–608.

Brelivet, Y., Kammerer, S., Rochel, N., Poch, O., & Moras, D. (2004). Signature of the oligomeric behaviour of nuclear receptors at the sequence and structural level. *EMBO Reports*, *5*, 423–429.

Burmester, T., & Scheller, K. (1999). Ligands and receptors: common theme in insect storage protein transport. *Naturwissenschaften.*, *86*, 468-474.

Camacho, J. P. M., Ruiz-Ruano, F. J., Martín-Blázquez, R., López-León, M. D., Cabrero, J., Lorite, P., Cabral-de-Mello, D.C., & Bakkali, M. (2015). A step to the gigantic genome of the desert locust: chromosome sizes and repeated DNAs. *Chromosoma*, *124*, 263–275.

Canavoso, L. E., Jouni, Z. E., Karnas, K. J., Pennington, J. E., & Wells, M. A. (2001). Fat metabolism in insects. *Annual Review of Nutrition*, *21*, 23–46.

Carvalho, M., Tobias, V., Vendrame, C., Shimabukuro, A., Gidlund, M., & Quintão, E. (2000). Lipoproteins modify the macrophage uptake of triacylglycerol emulsion and of zymosan particles by similar mechanisms. *Lipids*, *35*, 55–9.

Ceccato, P., Cressman, K., Giannini, A., & Trzaska, S. (2007). The desert locust upsurge in West Africa (2003-2005): Information on the desert locust early warning system and the prospects for seasonal climate forecasting. *International Journal of Pest Management, 53*, 7–13.

Chapman, R.E. (1998). Reproductive system: female. In: *The Insect*, *Structure and Function* (Chapman, R. F. & Reginald, F. eds) Cambridge University Press, Cambridge, 4th edn, pp 295–324.

Charles, J. P., Iwema, T., Epa, V. C., Takaki, K., Rynes, J., & Jindra, M. (2011). Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant. *Proceedings of the National Academy of Sciences*, *108*, 21128–21133.

Chen, J. S., & Raikhel, A. S. (1996). Subunit cleavage of mosquito pro-vitellogenin by a subtilisin-like convertase. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 6186–6190.

Chen, T. T., Couble, P., Abu-Hakima, R., & Wyatt, G. R. (1979). Juvenile hormone-controlled vitellogenin synthesis in *Locusta migratoria* fat body. Hormonal induction in vivo. *Developmental Biology*, *69*, 59–72.

Chen, T. T., Strahlendorf, W., & Wyatt, R. (1978). and Vitellogenin from Locusts (*Locusta migratoria*. *The Journal of Biological Chemistry*, *253*, 5325–5331.

Chen, X., & Haunerland, N. H. (1994). Fatty acid binding protein expression in locust flight muscle. Induction by flight, adipokinetic hormone, and low density lipophorin. *Insect Biochemistry and Molecular Biology*, *24*, 573–579.

Chino, H., Downer, R. G., & Takahashi, K. (1986). Effect of adipokinetic hormone on the structure and properties of lipophorin in locusts. *Journal of Lipid Research*, *27*, 21–9.

Chino, H., Hirayama, Y., Kiyomoto, Y., Downer, R. G. H., & Takahashi, K. (1987). Spontaneous aggregation of locust lipophorin during hemolymph collection. *Insect Biochemistry*, *17*, 89–97.

Chino, H., Murakami S., & Harashima K. (1969). Diglyceride-carrying lipoproteins in insect haemolymph: isolation, purification and properties*. Biochim Biophys Acta*, *176*,1-26

Chinzei, Y., & Wyatt, G. R. (1985). Vitellogenin titre in haemolymph of *Locusta migratoria* in normal adults, after ovariectomy, and in response to methoprene. *Journal of Insect Physiology*, *31*, 441–445.

Chinzei, Y., White, B. N., & Wyatt, G. R. (1982). Vitellogenin mRNA in locust fat body: identification, isolation, and quantitative changes induced by juvenile hormone. *Canadian Journal of Biochemistry*, *60*, 243–251.

Choi, I. K., & Hyun, S. (2012). Conserved microRNA miR-8 in fat body regulates innate immune homeostasis in Drosophila. *Developmental and Comparative Immunology*, *37*, 50–54.

Choi, K. M., Kim, T. N., Yoo, H. J., Lee, K. W., Cho, G. J., Hwang, T. G., Baik, S. H., Choi, D. S., & Kim, S. M. (2009). Effect of exercise training on A-FABP, lipocalin-2 and RBP4 levels in obese women. *Clinical Endocrinology, 70,* 569–574*.*

Chung, K. T., & Ourth, D. D. (2002). Purification and characterization of apolipophorin III from immune hemolymph of *Heliothis virescens* pupae. *Comparative Biochemistry and Physiology – B, 132*, 505–514.

Clavel, S., Farout, L., Briand, M., Briand, Y., & Jouanel, P. (2002). Effect of endurance training and/or fish oil supplemented diet on cytoplasmic fatty acid binding protein in rat skeletal muscles and heart. *European Journal of Applied Physiology*, *87*, 193–201.

Cole, K. D., Fernando-Warnakulasuriya, G. P., Boguski, M. S., Freeman, M., Gordon, J. I., Clark, W. A., Law, J. H., Wells, M. A. (1987). Primary structure and comparative sequence analysis of an insect apolipoprotein. Apolipophorin-III from *Manduca sexta. Journal of Biological Chemistry*, *262*, 11794–11800.

Comas, D., Piulachs, M.-D., & Bellés, X. (2000). Vitellogenin of *Blattella germanica* (L.) (Dictyoptera, Blattellidae): Nucleotide Sequence of the cDNA and Analysis of the Protein Primary Structure. *Archives of Insect Biochemistry and Physiology*, *45*, 1–11.

Cox, R. A., & García-Palmieri, M. R. (1990). Cholesterol, Triglycerides, and Associated Lipoproteins. In: *Clinical Methods: The History, Physical, and Laboratory Examinations* (Walker, H. K., Hall, W. D. & Hurst J. W. eds). Boston: Butterworths, 1990, 3rd edition, Chapter 31. https://www.ncbi.nlm.nih.gov/books/NBK351/

Cui, L. (2008). Control of locust FABP expression by RXR. M.Sc. thesis, Simon Fraser University, Burnaby, Canada. Unpublished.

Davey, K. G. (1996). Hormonal control of the follicular epithelium during vitellogenin uptake. *Invertebrate Reproduction & Development*, *30*, 249–254.

De Loof, A., Baggerman, G., Breuer, M., Claeys, U., Cerstiaens, A., Clynen, E., Janssen, T., Schoofs, L., & Broeck, J. Vanden. (2001). Gonadotropins in insects: An overview. *Archives of Insect Biochemistry and Physiology*, *47*, 129–138.

De Winther, M. P. J., Weers, P. M. M., Bogerd, J., & Van Der Horst, D. J. (1996). Apolipophorin III levels in *Locusta migratoria*. Developmental regulation of gene expression and hemolymph protein concentration. *Journal of Insect Physiology*, *42*, 1047–1052.

Dettloff, M., Kaiser, B., & Wiesner, A. (2001a). Localization of injected apolipophorin III in vivo - New insights into the immune activation process directed by this protein. *Journal of Insect Physiology*, *47*, 789–797.

Dettloff, M., Wittwer, D., Weise, C., & Wiesner, A. (2001b). Lipophorin of lower density is formed during immune responses in the lepidopteran insect *Galleria mellonella. Cell and Tissue Research*, *306*, 449–458.

Dhadialla, T. S., & Raikhel, A. S. (1990). Biosynthesis of mosquito vitellogenin. *Journal of Biological Chemistry*, *265*, 9924–9933.

Dhadialla, T. S., & Wyatt, G. R. (1983). Juvenile hormone-dependent vitellogenin synthesis in *Locusta migratoria* fat body: Inducibility related to sex and stage. *Developmental Biology*, *96*, 436–444.

Downer, R. G. H., & Matthews, J. R. (1976). Patterns of lipid distribution and utilisation in insects. *Integrative and Comparative* Biology, *16*, 733–745.

Edwards, G. C., Braun, R. P., & Wyatt, G. R. (1993). Induction of vitellogenin synthesis in *Locusta migratoria* by the juvenile hormone analog, pyriproxyfen. *Journal of Insect Physiology*, *39*, 609–614.

Elgendy, A. M., Elmogy, M., & Takeda, M. (2014). Molecular cloning, characterization, and expression pattern of the ultraspiracle gene homolog (RXR/USP) from the hemimetabolous insect *Periplaneta americana* (Dictyoptera, Blattidae) during vitellogenesis. *Molecular Biotechnology*, *56*, 126–135.

Engelmann, F. (2002). Ecdysteroids, Juvenile Hormone and Vitellogenesis in the Cockroach *Leucophaea maderae*. *Journal of Insect Science*, *2*, 20-27.

Fan, D., Zheng, Y., Yang, D., & Wang, J. (2003). NMR solution structure and dynamics of an exchangeable apolipoprotein, *Locusta migratoria* apolipophorin III. *Journal of Biological Chemistry*, *278*, 21212–21220.

Feingold, K. R., Funk, J. L., Moser, A. H., Shigenaga, J. K., Rapp, J. H., & Grunfeld, C. (1995). Role for circulating lipoproteins in protection from endotoxin toxicity. *Infection and Immunity*, *63*, 2041–2046.

Fernandis, A. Z., & Wenk, M. R. (2007). Membrane lipids as signaling molecules. *Current Opinion in Lipidology*, *18*, 121–128.

German, J. B., Smilowitz, J. T., & Zivkovic, A. M. (2006). Lipoproteins: When size really matters. *Current Opinion in Colloid and Interface Science*, *11*, 171–183.

Geva, N., Guershon, M., Orlova, M., & Ayali, A. (2010). Memoirs of a locust: Density-dependent behavioral change as a model for learning and memory. *Neurobiology of Learning and Memory*, *93*, 175–182.

Giguere, V. (1999). Orphan Nuclear Receptors: From Gene to Function. *Endocrine Reviews*, *20*, 689–725.

Gilbert, L.I., & Goodman, W. (1981). Chemistry, Metabolism, and Transport of Hormones Controlling Insect Metamorphosis. In: *Metamorphosis: A problem in developmental biology* (Gilbert, L.I., & Frieden, E. eds), 2nd edn. Plenum, New York, p 139-175.

Giorgi, F., Cecchettini, A., Falleni, A., Masetti, M., & Gremigni, V. (1998). Vitellogenin is glycosylated in the fat body of the stick insect *carausius morosus* and not further modified upon transfer to the ovarian follicle. *Micron*, *29*, 451–460.

Glatz, J. F., Van Breda, E., & Van der Vusse, G. J. (1998). Intracellular transport of fatty acids in muscle. Role of cytoplasmic fatty acid-binding protein. *Advances in Experimental Medicine and Biology*, *441*, 207–218.

Glinka, A. V., & Wyatt, G. R. (1996). Juvenile hormone activation of gene transcription in locust fat body. *Insect Biochemistry and Molecular Biology*, *26*, 13–18.

Godlewski, J., Wang, S., & Wilson, T. G. (2006). Interaction of bHLH-PAS proteins involved in juvenile hormone reception in Drosophila. *Biochemical and Biophysical Research Communications*, *342*, 1305–1311.

Griffin, B. A. (1999). Macronutrient Group Symposium on “Dietary determinants of lipoprotein-mediated cardiovascular risk” Lipoprotein atherogenicity: an overview of current mechanisms. *Proceedings of the Nutrition Society*, *58*, 163–169.

Gu, L., & Knipple, D. C. (2013). Recent advances in RNA interference research in insects: Implications for future insect pest management strategies. *Crop Protection, 45,* 36–40.

Guershon, M., & Ayali, A. (2012). Innate phase behavior in the desert locust, *Schistocerca gregaria*. *Insect Science*, *19*, 649–656.

Guidugli, K. R., Nascimento, A. M., Amdam, G. V., Barchuk, A. R., Omholt, S., Simões, Z. L. P., & Hartfelder, K. (2005). Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *FEBS Letters*, *579*, 4961–4965.

Guo, W., Wu, Z., Song, J., Jiang, F., Wang, Z., Deng, S., Walker, V. K., & Zhou, S. (2014). Juvenile Hormone-Receptor Complex Acts on Mcm4 and Mcm7 to Promote Polyploidy and Vitellogenesis in the Migratory Locust. *PLoS Genetics*, *10*(10). https://doi.org/10.1371/journal.pgen.1004702.

Hafer, J., Fischer, A., & Ferenz, H. J. (1992). Identification of the yolk receptor protein in oocytes of *Nereis virens* (Annelida, Polychaeta) and comparison with the locust vitellogenin receptor. *Journal of Comparative Physiology B*, *162*, 148–152.

Hård, K., Van Doorn, J. M., Thomas-Oates, J. E., Kamerling, J. P., & Van der Horst, D. J. (1993). Structure of the Asn-Linked Oligosaccharides of Apolipophorin III from the Insect *Locusta migratori*a. Carbohydrate-Linked 2-Aminoethylphosphonate as a Constituent of a Glycoprotein. *Biochemistry*, *32*, 766–775.

Haunerland, N. H. & Bowers, W. S. (1989). Comparative studies on arthropod lipoproteins. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *92*, 137–141.

Haunerland, N. H. (1994). Fatty acid binding protein in locust and mammalian muscle. Comparison of structure, function and regulation. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, *109*, 199-208.

Haunerland, N. H. (1996). Insect storage proteins: gene families and receptors. *Insect Biochemistry and Molecular Biology*, *26*, 755-765.

Haunerland, N. H. (1997). Transport and utilization in insect flight muscles. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, *117*, 475–482.

Haunerland, N. H., & Bowers, W. S. (1986a). A larval specific lipoprotein: Purification and characterization of a blue chromoprotein from *Heliothis zea*. *Biochemical and Biophysical Research Communications*, *134*, 580-586.

Haunerland, N. H., & Bowers, W. S. (1986b). Arylphorin from the corn earworm, *Heliothis zea*. *Insect Biochemistry*, *16*, 617–625.

Haunerland, N. H., & Chisholm, J. M. (1990). Fatty acid binding protein in flight muscle of the locust, *Schistocerca gregaria*. *Biochimica et Biophysica Acta*, *1047*, 233–238.

Haunerland, N. H., & Spener, F. (2004). Fatty acid-binding proteins - insights from genetic manipulations. *Progress in Lipid Research, 43, 328–349.*

Haunerland, N. H., & Thakrar, D. R. (2009). Insect fatty acid-binding proteins. In: *Fatty Acid-Binding Proteins* (Esteves, A. ed), Trivandrum: Transworld Research Network, pp. 31–44.

Haunerland, N. H., Andolfatto, P., Chisholm, J. M., Wang, Z., & Chen, X. (1992). Fatty-acid-binding protein in locust flight muscle: Developmental changes of expression, concentration and intracellular distribution. *European Journal of Biochemistry*, *210*, 1045–1051.

Haunerland, N. H., Chen, X., Andolfatto, P., Chisholm, J. M., & Wang, Z. (1993). Developmental changes of FABP concentration, expression, and intracellular distribution in locust flight muscle. *Molecular and Cellular Biochemistry*, *123*, 153–158.

Haunerland, N. H., Jagschies, G., Schulenberg, H., & Spener, F. (1984). Occurence of two fatty-acid-binding proteins in bovine liver cytosol and their binding of fatty acids, cholesterol, and other lipophilic ligands. *Hoppe-Seyler's Zeitschrift Fur Physiologische Chemie*, *365*, 365–376.

Haunerland, N. H., Ryan, R. O., Law, J. H., & Bowers, W. S. (1986). Lipophorin from the grasshopper, *Gastrimargus africanus*. Isolation and properties of apolipophorin III. *Insect Biochemistry*, *16*, 797–802.

Hayward, D. C., Bastiani, M. J., Trueman, J. W., Truman, J. W., & Riddiford, L. M. & Ball, E. E. (1999). The sequence of Locusta RXR, homologous to Drosophila Ultraspiracle, and its evolutionary implications. *Development Genes and Evolution*, *209*, 564–571.

Hayward, D. C., Dhadialla, T. S., Zhou, S., Kuiper, M. J., Ball, E. E., Wyatt, G. R., & Walker, V. K. (2003). Ligand specificity and developmental expression of RXR and ecdysone receptor in the migratory locust. *Journal of Insect Physiology*, *49*, 1135–1144.

Hughes, A. L. (2010). Evolutionary Conservation of Amino Acid Composition in Paralogous Insect Vitellogenins. *Gene, 467*, 35–40.

Hult, E. F., Huang, J., Marchal, E., Lam, J., & Tobe, S. S. (2015). RXR/USP and EcR are critical for the regulation of reproduction and the control of JH biosynthesis in *Diploptera punctata*. *Journal of Insect Physiology, 80, 48–60.*

Huvenne, H., & Smagghe, G. (2010). Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: A review. *Journal of Insect Physiology, 56,* 227–235.

Jia, Z., Pei, Z., Maiguel, D., Toomer, C. J., & Watkins, P. A. (2007). The fatty acid transport protein (FATP) family: Very long chain acyl-CoA synthetases or solute carriers? *Journal of Molecular Neuroscience*, *33*, 25–31.

Jindra, M., Uhlirova, M., Charles, J. P., Smykal, V., & Hill, R. J. (2015). Genetic Evidence for Function of the bHLH-PAS Protein Gce/Met as a Juvenile Hormone Receptor. *PLoS Genetics*, *11*(7), e1005394.

Jones, G., & Sharp, P. A. (1997). Ultraspiracle: An invertebrate nuclear receptor for juvenile hormones. *Proceedings of the National Academy of Sciences*, *94*, 13499–13503.

Jones, G., Jones, D., Teal, P., Sapa, A., & Wozniak, M. (2006). The retinoid-X receptor ortholog, ultraspiracle, binds with nanomolar affinity to an endogenous morphogenetic ligand. *FEBS Journal*, *273*, 4983–4996.

Jones, G., Wozniak, M., Chu, Y., Dhar, S., & Jones, D. (2001). Juvenile hormone III-dependent conformational changes of the nuclear receptor ultraspiracle. *Insect Biochemistry and Molecular Biology*, *32*, 33–49.

Kaikause, R. M., Bass, N. M., & Ockner, R. K. (1990). Functions of fatty acid binding proteins. *Experientia, 46*, 617–630.

Kanost, M. R., Boguski, M. S., Freeman, M., Gordon, J. I., Wyatt, G. R., & Wells, M. A. (1988). Primary Structure of Apolipophorin-III from the Migratory Locust, *Locusta-Migratoria* - Potential Amphipathic Structures and Molecular Evolution of an Insect Apolipoprotein. *Journal of Biological Chemistry*, *263*, 10568–10573.

Kanost, M. R., Kawooya, J. K., Law, J. H., Ryan, R. O., Heusden, M. C. van, & Ziegler, R. (1990). Insect Haemolymph Proteins*. Advances in Insect Physiology*, *22*, 299-396.

Kato, Y., Motoi, Y., Taniai, K., Kadono-Okuda, K., Yamamoto, M., Higashino, Y., Shimabukuro, M., Chowdhury, S., Xu, J., Sugiyama, M., Hiramatsu, M., & Yamakawa, M. (1994). Lipopolysaccharide-lipophorin complex formation in insect hemolymph: a common pathway of lipopolysaccharide detoxification both in insects and in mammals. *Insect Biochemistry and Molecular Biology*, *24*, 547–555.

Kawooya, J. K., Keim, P. S., Ryan, R. O., Shapiro, J. P., Samaraweera, P., & Law, J. H. (1984). Insect apolipophorin III. Purification and properties. *Journal of Biological Chemistry*, *259*, 10733–10737.

Kawooya, J. K., Meredith, S. C., Wells, M. A., Kézdy, F. J., & Law, J. H. (1986). Physical and surface properties of insect apolipophorin III. *The Journal of Biological Chemistry*, *261*, 13588–13591.

Kawooya, J. K., Osir, E. O., & Law, J. H. (1988). Uptake of the major hemolymph lipoprotein and its transformation in the insect egg. *Journal of Biological Chemistry*, *263*, 8740–8747.

Kewley, R. J., Whitelaw, M. L., & Chapman-Smith, A. (2004). The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *International Journal of Biochemistry and Cell Biology*, *36*, 189–204.

Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P., & Hogness, D. S. (1991). The *Drosophila EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell*, *67*, 59–77.

Kuczer, M., Rosiński, G., & Konopińska, D. (2007). Insect gonadotropic peptide hormones: some recent developments. *Journal of Peptide Science: An Official Publication of the European Peptide Society*, *13*, 16–26.

Kunkel, J.G. (1981). A Minimal Model Of Metamorphosis: Fat Body Competence to Respond to Juvenile Hormone. In: *Current Topics in Insect Endocrinology and Nutrition, 107. Plenum Publishing Corporation* (Bhaskaran, G., Friedman, S., & Rodriguez, J.G. eds), New York, pp. 107-129. In: Cruz, J., Martín, D., Pascual, N., Maestro, J. L., Piulachs, M. D., & Bellés, X. (2003). Quantity does matter. Juvenile hormone and the onset of vitellogenesis in the German cockroach. *Insect Biochemistry and Molecular Biology*, *33*, 1219–1225.

Laemmli, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, *227*, 680-685.

Levenbook, L. (1985). Storage proteins. In: Kerkut, G.A., Gilbert, L.I. (eds.), Comprehensive Insect Physiology, Biochemistry, and Pharmacology, vol. 10. Pergamon Press, Oxford, New York, pp. 307-346.

Li, M., Mead, E. A., & Zhu, J. (2011). Heterodimer of two bHLH-PAS proteins mediates juvenile hormone-induced gene expression. *Proceedings of the National Academy of Sciences*, *108*, 638–643.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*, 402–408.

Locke, J., White, B., & Wyatt, G. R. (1987). Cloning and 5’ end nucleotide sequences of two juvenile hormone-inducible vitellogenin genes of the African migratory locust. *DNA, 6*, 331–342.

Maatman, R. G., Degano, M., van Moerkerk, H. T., van Marrewijk, W. J., van der Horst, D. J., Sacchettini, J. C., & Veerkamp, J. H. (1994). Primary structure and binding characteristics of locust and human muscle fatty acid binding proteins. *European Journal of Biochemistry, 221,* 801–810.

Maeda, K., Uysal, K. T., Makowski, L., Görgün, C. Z., Atsumi, G., Parker, R. A., Brüning, J., Hertzel, A. V., Bernlohr, D. A., & Hotamisligil, G. S. (2003). Role of the fatty acid binding protein mal1 in obesity and insulin resistance. *Diabetes*, *52*, 300–307.

Mahamat, H., Hassanali, A., & Ferenz, H. J. (1997). Vitellogenin titres in normal and accelerated maturation of Gragarious-phase *Schistocerca gregaria*. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, *116*, 447–451.

Mayer, R. J. and Candy, D. J. (1969). Changes in Energy Reserves during Flight of Desert Locust, *Schistocerca Gregaria*. *Comparative Biochemistry and Physiology 31,* 409-418.

Miura, K., Oda, M., Makita, S., & Chinzei, Y. (2005). Characterization of the Drosophila Methoprene-tolerant gene product: Juvenile hormone binding and ligand-dependent gene regulation. *FEBS Journal*, *272*, 1169–1178.

Nair, K. K., Chen, T. T., & Wyatt, G. R. (1981). Juvenile hormone-stimulated polyploidy in adult locust fat body. *Developmental Biology*, *81*, 356–360.

Narayanaswami, V., & Ryan, R. O. (2000). Molecular basis of exchangeable apolipoprotein function. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1483*, 15–36.

Narayanaswami, V., Wang, J., Schieve, D., Kay, C. M., & Ryan, R. O. (1999). A molecular trigger of lipid binding-induced opening of a helix bundle exchangeable apolipoprotein. *Proceedings of the National Academy of Sciences*, *96*, 4366–4371.

Nelson, C. M., Ihle, K. E., Fondrk, M. K., Page, R. E., & Amdam, G. V. (2007). The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biology*, *5*, 0673–0677.

Nose, Y., Min Lee, J., Ueno, T., Hatakeyama, M., & Oishi, K. (1997). Cloning of cDNA for vitellogenin of the parasitoid wasp, *Pimpla nipponica* (Hymenoptera: Apocrita: Ichneumonidae): Vitellogenin primary structure and evolutionary considerations. *Insect Biochemistry and Molecular Biology*, *27*, 1047–1056.

Ogoyi, D. O., Osir, E. O., & Olembo, N. K. (1995). Lipophorin and apolipophorin-III in solitary and gregarious phases of *Schistocerca gregaria*. *Comparative Biochemistry and Physiology, 112B*, 441–449.

Oishi, M., Locke, J., & Wyatt, G. R. (1985). The ribosomal RNA genes of *Locusta migratoria*: copy number and evidence for underreplication in a polyploid tissue. *Canadian Journal of Cell Biology*, *63*, 1064–1070.

Owen, G., & Zelent, A. (2000). Origins and evolutionary diversification of the nuclear receptor superfamily. *Cellular and Molecular Life Sciences*, *57*, 809–827.

Pan, M.L., Bell, W.J., & Telfer, W.H. (1969). Vitellogenic Blood Protein Synthesis by Insect Fat Body. *Science*, *165*, 393–394.

Parthasarathy, R., Sheng, Z., Sun, Z., & Pall, S. R. (2010). Ecdysteroid regulation of ovarian growth and oocyte maturation in the red flour beetle, *Tribolium castaneum*, *40*, 429–439.

Peiser, L., Mukhopadhyay, S., & Gordon, S. (2002). Scavenger receptors in innate immunity. *Current Opinion in Immunology*, *14*, 123–128.

Pener, M. P., & Simpson, S. J. (2009). Locust Phase Polyphenism: An Update. *Advances in Insect* Physiology*, 36,* 1-272*.*

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods, 8*, 785–786. https://doi.org/10.1038/nmeth.1701

Piulachs, M. D., Guidugli, K. R., Barchuk, A. R., Cruz, J., Simões, Z. L. P., & Bellés, X. (2003). The vitellogenin of the honey bee, Apis mellifera: Structural analysis of the cDNA and expression studies. *Insect Biochemistry and Molecular Biology*, *33*, 459–465.

Qu, H., Cui, L., Rickers-Haunerland, J., & Haunerland, N. H. (2007). Fatty acid-dependent expression of the muscle FABP gene - Comparative analysis of gene control in functionally related, but evolutionary distant animal systems. *Molecular and Cellular Biochemistry*, *299*, 45–53.

Richard, D. S., Gilbert, M., Crum, B., Hollinshead, D. M., Schelble, S., & Scheswohl, D. (2001). Yolk protein endocytosis by oocytes in *Drosophila melanogaster*: Immunofluorescent localization of clathrin, adaptin and the yolk protein receptor. *Journal of Insect Physiology*, *47*, 715–723.

Riddiford, L. M. (2008). Juvenile hormone action: A 2007 perspective. *Journal of Insect Physiology*, *54*, 895–901.

Riddiford, L. M., Cherbas, P., & Truman, J. W. (2001). Ecdysone Receptors and Their Biological Actions. *Vitamins and Hormones*, *60*, 1–73. https://doi.org/10.1016/S0083-6729(00)60016-X

Robinson, R. (2008). For mammals, loss of yolk and gain of milk went hand in hand. *PLoS Biology*, *6*, 421. https://doi.org/10.1371/journal.pbio.0060077

Roehrkasten, A., Ferenz, H.-J., Buschmann-Gebhardt, B., & Hafer, J. (1989). Isolation of the vitellogenin-binding protein from locust ovaries. *Archives of Insect Biochemistry and Physiology*, *10*, 141–149.

Romans, P., Tu, Z., Ke, Z., & Hagedorn, H. H. (1995). Analysis of a vitellogenin gene of the mosquito, *Aedes aegypti* and comparisons to vitellogenins from other organisms. *Insect Biochemistry and Molecular Biology*, *25*, 939–958.

Roth, Z., Parnes, S., Wiel, S., Sagi, A., Zmora, N., Chung, J. S., & Khalaila, I. (2010). N-glycan moieties of the crustacean egg yolk protein and their glycosylation sites. *Glycoconjugate Journal*, *27*, 159–169.

Ryan, R. O. (1990). Dynamics of insect lipophorin metabolism. *Journal of Lipid Research*, *31*, 1725–1739.

Ryan, R. O., Law, J. H., & Schmidt, J. O. (1984). Chemical and immunological properties of lipophorins from seven insect orders. *Archives of Insect Biochemistry and Physiology*, *1*, 375–383.

Ryan, R. O., Ziegler, D. J., Van der Horst, D. J., & Law, J. H. (1990). Characterization of apolipophorin III from *Barytettix psolus* and *Melanoplus differentialis. Insect Biochemistry*, *20*, 127–133.

Santos, D., Vanden Broeck, J., & Wynant, N. (2014). Systemic RNA interference in locusts: Reverse genetics and possibilities for locust pest control. *Current Opinion in Insect Science, 6,* 9–14.

Sappington, T. W., & S. Raikhel, A. (1998). Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochemistry and Molecular Biology*, *28*, 277–300.

Schaap, F. G., van der Vusse, G. J., & Glatz, J. F. C. (2002). Evolution of the family of intracellular lipid binding proteins in vertebrates*. Molecular and Cellular Biochemistry, 239,* 69–77.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*, 671–675.

Schneider, W. J. (1996). Vitellogenin Receptors: Oocyte-Specific Members of the Low-Density Lipoprotein Receptor Supergene Family. *International Review of Cytology*, *166*, 103–137.

Schoonhoven, L. M., Van Loon, J. J. A., & Dicke, M. (2005). Insect-Plant Biology. 2nd ed., Oxford: University Press, UK., pp. 237.

Scott, J. G., Michel, K., Bartholomay, L. C., Siegfried, B. D., Hunter, W. B., Smagghe, G., Zhu, K. Y., & Douglas, A. E. (2013). Towards the elements of successful insect RNAi. *Journal of Insect Physiology, 59,* 1212–1221.

Segrest, J. P., Jones, M. K., De Loof, H., Brouillette, C. G., Venkatachalapathi, Y. V, & Anantharamaiah, G. M. (1992). The amphipathic helix in the exchangeable apolipoproteins: a review of secondary structure and function. *Journal of Lipid Research*, *33*, 141–166.

Shapiro, J. P., & Law, J. H. (1983). Locust adipokinetic hormone stimulates lipid mobilization in *Manduca sexta. Biochemical and Biophysical Research Communications*, *115*, 924–931.

Shapiro, J. P., Law, J. H., & Wells, M. A. (1988). Lipid transport in insects. *Annual Review of Entomology*, *33*, 297–318.

Sheng, Z., Xu, J., Bai, H., Zhu, F., & Palli, S. R. (2011). Juvenile hormone regulates vitellogenin gene expression through insulin-like peptide signaling pathway in the red flour beetle, *Tribolium castaneum*. *Journal of Biological Chemistry*, *286*, 41924–41936.

Smith, A. F., Owen, L. M., Strobel, L. M., Chen, H., Kanost, M. R., Hanneman, E., & Wells, M. A. (1994). Exchangeable apolipoproteins of insects share a common structural motif. *Journal of Lipid Research*, *35*, 1976–1984.

Song, J., Guo, W., Jiang, F., Kang, L., & Zhou, S. (2013). Argonaute 1 is indispensable for juvenile hormone mediated oogenesis in the migratory locust, *Locusta migratoria*. *Insect Biochemistry and Molecular Biology*, *43*, 879–887.

Steiner, B., Pfister-Wilhelm, R., Grossniklaus-Bürgin, C., Rembold, H., Treiblmayr, K., & Lanzrein, B. (1999). Titres of juvenile hormone I, II and III in *Spodoptera littoralis* (Noctuidae) from the egg to the pupal moult and their modification by the egg-larval parasitoid *Chelonus inanitus* (Braconidae). *Journal of Insect Physiology*, *45*, 401–413.

Sum, H., & Haunerland, N. H. (2007). VHDL, a larval storage protein from the corn earworm, *Helicoverpa zea*, is a member of the vitellogenin gene family. *Insect Biochemistry and Molecular Biology*, *37*, 1086–1093.

Sun, D. H., Ziegler, R., Milligan, C. E., Fahrbach, S., & Schwartz, L. M. (1995). Apolipophorin-III Is Dramatically Up-Regulated During the Programmed Death of Insect Skeletal-Muscle and Neurons. *Journal of Neurobiology*, *26*, 119–129.

Symmons, P.M., & Cressman, K. (2001). Desert locust Guidelines. FAO, Rome. Available: http://www.fao.org/ag/locusts/common/ ecg/347_ en_DLG1e.pdf, p 43.

Telfer W.H., Huebner E., & Smith D.S. (1982). The Cell Biology of Vitellogenic Follicles in *Hyalophora* and *Rhodnius*. In: *Insect Ultrastructure* (King R.C., & Akai H. eds). Springer, Boston, MA.

Telfer, W.H. (1954). Immunological studies of insect metamorphosis. II. The role of a sex-limited blood protein in egg formation by the Cecropia silkworm. *The Gournal of General Physiology*, *37*, 539–558.

Terashima, J., Takaki, K., Sakurai, S., & Bownes, M. (2005). Nutritional status affects 20-hydroxyecdysone concentration and progression of oogenesis in *Drosophila melanogaster. Journal of Endocrinology*, *187*, 69–79.

Terenius, O., Papanicolaou, A., Garbutt, J. S., Eleftherianos, I., Huvenne, H., Kanginakudru, S., Albrechtsen, M., An, C., Aymeric, J. L., Barthel, A., Bebas, P., Bitra, K., Bravo, A., Chevalier, F., Collinge, D. P., Crava, C. M., de Maagd, R. A., Duvic, B., Erlandson, M., Faye, I., Felföldi, G., Fujiwara, H., Futahashi, R., Gandhe, A. S., Gatehouse, H. S., Gatehouse, L. N., Giebultowicz, J. M., Gómez, I., Grimmelikhuijzen, C. J. P., Groot, A. T., Hauser, F., Heckel, D. G., Hegedus, D. D., Hrycaj, S., Huang, L., Hull, J. J., Iatrou, K., Iga, M., Kanost, M. R., Kotwica, J., Li, C., Li, J., Liu, J., Lundmark, M., Matsumoto, S., Meyering-Vos, M., Millichap, P. J., Monteiro, A., Mrinal, N., Niimi, T., Nowara, D., Ohnishi, A., Oostra, V., Ozaki, K., Papakonstantinou, M., Popadic, A., Rajam, M. V., Saenko, S., Simpson, R. M., Soberón, M., Strand, M. R., Tomita, S., Toprak, U., Wang, P., Wee, C. W., Whyard, S., Zhang, W., Nagaraju, J., Ffrench-Constant, R. H., Herrero, S., Gordon, K., Swevers, L., & Smagghe, G. (2011). RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology, 57,* 231–245.

Thomas, H. E., Stunnenberg, H. G., & Stewart, A. F. (1993). Heterodimerization of the Drosophila ecdysone receptor with retinoid X receptor and ultraspiracle. *Nature*, *362*, 471–475.

Thompson, J., Winter, N., Terwey, D., Bratt, J., & Banaszak, L. (1997). The crystal structure of the liver fatty acid-binding protein. A complex with two bound oleates. *Journal of Biological Chemistry*, *272*, 7140–7150.

Tufail, M., & Takeda, M. (2008). Molecular characteristics of insect vitellogenins. *Journal of Insect Physiology*, *54*, 1447-1458.

Tufail, M., Hatakeyama, M., & Takeda, M. (2001). Molecular evidence for two vitellogenin genes and processing of vitellogenins in the American cockroach, *Periplaneta americana*. *Archives of Insect Biochemistry and Physiology*, *48*, 72–80.

Tufail, M., Lee, J. M., Hatakeyama, M., Oishi, K., & Takeda, M. (2000). Cloning of vitellogenin cDNA of the American cockroach, *Periplaneta americana* (Dictyoptera), and its structural and expression analyses. *Archives of Insect Biochemistry and Physiology*, *45*, 37–46.

Tufail, M., Nagaba, Y., Elgendy, A. M., & Takeda, M. (2014). Regulation of vitellogenin genes in insects. *Entomological Science*, *17*, 269–282.

Uryu, O., Ameku, T., & Niwa, R. (2015). Recent progress in understanding the role of ecdysteroids in adult insects: Germline development and circadian clock in the fruit fly *Drosophila melanogaster*. *Zoological Letters*, *1*, 32. https://doi.org/10.1186/s40851-015-0031-2

Valle, D. (1993). Vitellogenesis in insects and other groups: a review. *Memórias Do Instituto Oswaldo Cruz*, 88, 1-26.

Van der Horst, D. J. & Rodenburg, K. W. (2010). Locust flight activity as a model for hormonal regulation of lipid mobilization and transport. *Journal of Insect Physiology 56*, 844-853.

Van der Horst, D. J. (1990). Lipid transport function of lipoproteins in flying insects. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, *1047*, 195–211.

Van der Horst, D. J. (2003). Insect adipokinetic hormones: Release and integration of flight energy metabolism. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, *136*, 217–226.

Van der Horst, D. J., Van Doorn, J. M., & Beenakkers, A. M. T. (1984). Hormone-induced rearrangement of locust haemolymph lipoproteins. The involvement of glycoprotein C2. *Insect Biochemistry and Molecular Biology*, *14*, 495–504.

Van der Horst, D. J., Van Hoof, D., Van Marrewijk, W.J.A., & Rodenburg, K.W. (2002). Alternative lipid mobilization: The insect shuttle system. *Molecular and Cellular Biochemistry*, *239*, 113–119.

Van der Horst, D. J., Van Marrewijk, W. J. A., & Diederen, J. H. B. (2001). Adipokinetic hormones of insect: Release, signal transduction, and responses. *International Review of Cytology*, *211*, 179–240.

Van der Horst, D. J., Vandoorn, J. M., Voshol, H., Kanost, M. R., Ziegler, R., & Beenakkers, A. M. T. (1991). Different Isoforms of an Apoprotein (Apolipophorin-Iii) Associate with Lipoproteins in *Locusta-Migratoria*. *European Journal of Biochemistry*, *196*, 509–517.

Van der Vusse, G. J., Glatz, J. F., Van Nieuwenhoven, F. A., Reneman, R. S., & Bassingthwaighte, J. B. (1998). Transport of long-chain fatty acids across the muscular endothelium. *Advances in Experimental Medicine and Biology*, *441*, 181–91.

Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. Vanden, & Simonet, G. (2009). Identification and validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. *BMC Molecular Biology*, *10*(1), 56. https://doi.org/10.1186/1471-2199-10-56

Veerkamp, J. H., & Vanmoerkerk, H. T. B. (1993). Fatty-Acid-Binding Protein and Its Relation to Fatty-Acid Oxidation. *Molecular and Cellular Biochemistry, 123,* 101–106.

Wang, J., Gagné, S. M., Sykes, B. D., & Ryan, R. O. (1997). Insight into lipid surface recognition and reversible conformational adaptations of an exchangeable apolipoprotein by multidimensional heteronuclear NMR techniques. *Journal of Biological Chemistry*, *272*, 17912–17920.

Wang, J., Sykes, B. D., & Ryan, R. O. (2002). Structural basis for the conformational adaptability of apolipophorin III, a helix-bundle exchangeable apolipoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 1188–1193.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C., Wang, Y., He, J., Luo, Y., Wang, Z., Guo, X., Guo, W., Wang, X., Zhang, Y., Yang, M., Hao, S., Chen, B., Ma, Z., Yu, D., Xiong, Z., Zhu, Y., Fan, D., Han, L., Wang, B., Chen, Y., Wang, J., Yang, L., Zhao, W., Feng, Y., Chen, G., Lian, J., Li, Q., Huang, Z., Yao, X., Lv, N., Zhang, G., Li, Y., Wang, J., Zhu, B., & Kang, L. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, *5*, 2957. https://doi.org/10.1038/ncomms3957

Weber, J. M. (2011). Metabolic fuels: regulating fluxes to select mix. *Journal of Experimental Biology*, *214*, 286–294.

Weers, P. M. M., & Ryan, R. O. (2003). Apolipophorin III: A lipid-triggered molecular switch. *Insect Biochemistry and Molecular Biology*, *33*, 1249–1260.

Weers, P. M. M., Van Marrewijk, W. J. A., Beenakkers, A. M. T., & Van der Horst, D. J. (1993). Biosynthesis of locust lipophorin. Apolipophorins I and II originate from a common precursor. *Journal of Biological Chemistry*, *268*, 4300–4303.

Wegener, G. (1996). Flying insects: Model systems in exercise physiology. *Experientia 52,* 404-412*.*

Weise, C., Franke, P., Kopácek, P., & Wiesner, A. (1998). Primary structure of apolipophorin-III from the greater wax moth, *Galleria mellonella*. *Journal of Protein Chemistry*, *17*, 633–641.

Weis-Fogh, T. & Jensen, M. (1956). Biology and Physics of Locust Flight. I. Basic Principles in Insect Flight. Philosophical Transactions of the Royal Society of London. Series B, *Biological Sciences, 239,* 415–458.

Wells, M. A., Ryan, R. O., Kawooya, J. K., & Law, J. H. (1987). The role of apolipophorin III in in vivo lipoprotein interconversions in adult *Manduca sexta*. *Journal of Biological Chemistry*, *262*, 4172–4176.

Whitten, M. M. A., Tew, I. F., Lee, B. L., & Ratcliffe, N. A. (2004). A Novel Role for an Insect Apolipoprotein (Apolipophorin III) in -1,3-Glucan Pattern Recognition and Cellular Encapsulation Reactions. *The Journal of Immunology*, *172*, 2177–2185.

Wiesner, A., Losen, S., kopacek, P., Weise, C., & Gotz, P. (1997). Isolated Apolipophorin III from *Galleria mellonella* Stimulates the Immune Reactions of This Insect. *Journal of Insect Physiology*, *43*, 383–391.

Wigglesworth, V. B. (1934). The Physiology of Ecdysis in *Rhodnius Prolixus* (Hemiptera). II. Factors controlling Moulting and “Metamorphosis.” *Quarterly Journal of Microscopical Science*, *77*, 191-222.

Wigglesworth, V. B. (1936). The function of the Corpus Allatum in the Growth and Reproduction of *Rhodnius Prolixus* (Hemiptera). *Quarterly Journal of Microscopical Science*, *79*, 91-121.

Wyatt, G. R. (1988). Vitellogenin synthesis and the analysis of juvenile hormone action in locust fat body. *Canadian Journal of Zoology*, *66*, 2600–2610.

Wyatt, G. R., & Davey, K. G. (1996). Cellular and Molecular Actions of Juvenile Hormone. II. Roles of Juvenile Hormone in Adult Insects. *Advances in Insect Physiology*, 26, 1-155.

Wyatt, G. R., Kanost, M. R., Locke, J., & Walker, V. K. (1986). Juvenile hormone-regulated locust vitellogenin genes: Lack of expression after transfer into Drosophila. *Archives of Insect Biochemistry and Physiology*, *1*, 35–46.

Wynant, N., Verlinden, H., Breugelmans, B., Simonet, G., & Vanden Broeck, J. (2012). Tissue-dependence and sensitivity of the systemic RNA interference response in the desert locust, *Schistocerca gregaria*. *Insect Biochemistry and Molecular Biology*, *42*, 911–917.

Yu, X.-Q., Zhu, Y.-F., Ma, C., Fabrick, J. A., & Kanost, M. R. (2002). Pattern recognition proteins in *Manduca sexta* plasma. *Insect Biochemistry and Molecular Biology*, *32*, 1287–1293.

Zhang, J., & Wyatt, G. R. (1996). Cloning and upstream sequence of a juvenile hormone-regulated gene from the migratory locust. *Gene*, *175*, 193–197.

Appendix A. Supplemental data for Chapter 2

Sequence A1: Nucleotide and deduced protein sequence of *S. gregaria* VG-A (GenBank accession # MK206969)

Underlined are transcription start sit and the signal peptide sequence. The poly-adenylation signal is doubled underlined.

```
acatggggaccagctgctccagtccacacgatcatgtgggccgtgatcctcttggggctc
 -  -  -  -  -  -  -  -  -  -  -  M  W  A  V  I  L  L  G  L
ctcgtcggggcaactgcgtccgaacagccaagcctatggaaaccgggatatgaatacgtg
 L  V  G  A  T  A  S  E  Q  P  S  L  W  K  P  G  Y  E  Y  V
tacaacgttactgccacaagtttgaccgcactaaaccagctagcaaaccaatggagtggt
 Y  N  V  T  A  T  S  L  T  A  L  N  Q  L  A  N  Q  W  S  G
ctaacctacaatgtaaccctggttgccaggccgagatcgccgagagagcttcagctgcgg
 L  T  Y  N  V  T  L  V  A  R  P  R  S  P  R  E  L  Q  L  R
atctacaacgcaatctacagcagaattcatgcaaatctgtctggaggatttgacaatcct
 I  Y  N  A  I  Y  S  R  I  H  A  N  L  S  G  G  F  D  N  P
atcccagaagaatatttgcattggaagcctctgaggctatcgaaggctcctttcgacgtg
 I  P  E  E  Y  L  H  W  K  P  L  R  L  S  K  A  P  F  D  V
ctctacaaggatggagagatcgtcgatatgctcgtcgactgttcgctcactgatgcagaa
 L  Y  K  D  G  E  I  V  D  M  L  V  D  C  S  L  T  D  A  E
gttaatcagctgaaagctcttgccagtatgttccaagtagacatagatccagagaaaggg
 V  N  Q  L  K  A  L  A  S  M  F  Q  V  D  I  D  P  E  K  G
tcgatatttaagaagtacgagcaaactgtaactggagactgtatgactcaatatgccaga
 S  I  F  K  K  Y  E  Q  T  V  T  G  D  C  M  T  Q  Y  A  R
actcagattccgaggtacttcctccagagtgaacaaaagattgatctgcccgaaatcatc
 T  Q  I  P  R  Y  F  L  Q  S  E  Q  K  I  D  L  P  E  I  I
gacggtaaagccaaactaatagacgtcacgaaagagctaaacaactcaaggtgccgtgac
 D  G  K  A  K  L  I  D  V  T  K  E  L  N  N  S  R  C  R  D
gaccgcctttactacttcggtctggacgagaaacccttcaacagtagctcagcaccgtct
 D  R  L  Y  Y  F  G  L  D  E  K  P  F  N  S  S  S  A  P  S
ctggctcattcgtccaacagcagagctattgtggctaacatgtcacaaaacaactacctc
 L  A  H  S  S  N  S  R  A  I  V  A  N  M  S  Q  N  N  Y  L
ctgggttacgcggagtcaacagaaaagattgttgtcagtcctttcttctacaaccaacag
 L  G  Y  A  E  S  T  E  K  I  V  V  S  P  F  F  Y  N  Q  Q
aagggaatgattgcgagcacagtgcgtgcggttctgactgaacacttccgcctccgcaag
 K  G  M  I  A  S  T  V  R  A  V  L  T  E  H  F  R  L  R  K
gtagcccaagctcctacccccgaatgccgcaccggaggtctcctgtacagccacgttgac
 V  A  Q  A  P  T  P  E  C  R  T  G  G  L  L  Y  S  H  V  D
tcttctgagggcagagctcatcttactgagtctggcagcgtagaagatcgtgaacaacag
 S  S  E  G  R  A  H  L  T  E  S  G  S  V  E  D  R  E  Q  Q
tcagacgaagcactagaaagttcagaaagtgccgcagctagttacgaggcgccgaggaga
 S  D  E  A  L  E  S  S  E  S  A  A  A  S  Y  E  A  P  R  R
taccggcgagcagttagtccataccaacaacatctgccagagcctgaaatgaaaaagcca
 Y  R  R  A  V  S  P  Y  Q  Q  H  L  P  E  P  E  M  K  K  P
ccacaccacccgtatgaggctcttttcgtggctgtaggaggtaaatctgtactctatgac
 P  H  H  P  Y  E  A  L  F  V  A  V  G  G  K  S  V  L  Y  D
cgacaagtaaatgcctctcgcgtagctatttcattggccagagatattggcgaagcactc
 R  Q  V  N  A  S  R  V  A  I  S  L  A  R  D  I  G  E  A  L
tacgacccactcttgattgcaccaaaagaagccttatcaaaattcccgctattgataagc
 Y  D  P  L  L  I  A  P  K  E  A  L  S  K  F  P  L  L  I  S
cttctccgagtattcgacgagatcagtcttaaacgcattactgaaaaactatacagtgaa
 L  L  R  V  F  D  E  I  S  L  K  R  I  T  E  K  L  Y  S  E
```

```
cagtcaccagaaaacccaagctcggagtatgcatggagagcgtactatatggctctcgca
 Q  S  P  E  N  P  S  S  E  Y  A  W  R  A  Y  Y  M  A  L  A
caagctggtactggccctgctctcctgaatctcatccagatcatcaatgaagacatgatc
 Q  A  G  T  G  P  A  L  L  N  L  I  Q  I  I  N  E  D  M  I
agtggaaaagccgccgccgaagcttttgcatacatccctgaagcagctcgtacacctatt
 S  G  K  A  A  A  E  A  F  A  Y  I  P  E  A  A  R  T  P  I
ccgggatacgtcaagcaggtcttcaaatttgccacccgagatgcatctgacaagtacccg
 P  G  Y  V  K  Q  V  F  K  F  A  T  R  D  A  S  D  K  Y  P
ctgaacatcactgcaatactggccgccactgaactcatccatttggcccaagtgagcaat
 L  N  I  T  A  I  L  A  A  T  E  L  I  H  L  A  Q  V  S  N
gtatccaattcacgctatccgcaagataactatgaagacttgctgggtgaatataacccc
 V  S  N  S  R  Y  P  Q  D  N  Y  E  D  L  L  G  E  Y  N  P
gaatacgtgaaacagttggaagacaagctgaggcgcgctgtagaagacgcttacagcccg
 E  Y  V  K  Q  L  E  D  K  L  R  R  A  V  E  D  A  Y  S  P
aatattcaactgtacatcagggccctcggcaacattggacatccagacatcgtgaaagtc
 N  I  Q  L  Y  I  R  A  L  G  N  I  G  H  P  D  I  V  K  V
tttaagccttacttggaaggacgcagacccgtttctaccttccagcgatttatgatggtc
 F  K  P  Y  L  E  G  R  R  P  V  S  T  F  Q  R  F  M  M  V
atgtcactggacaaacttgccgaagcttatccagagatagcgcagaacgtattcctgcag
 M  S  L  D  K  L  A  E  A  Y  P  E  I  A  Q  N  V  F  L  Q
ctcttccagaacactggagaaacgcacgagatcagatgtgcttcggctctgttgttgatg
 L  F  Q  N  T  G  E  T  H  E  I  R  C  A  S  A  L  L  L  M
gcgacgaacccaccagctagcgtcctgcagaaactggctcatttctcaaacctcgaccat
 A  T  N  P  P  A  S  V  L  Q  K  L  A  H  F  S  N  L  D  H
aacacacaggtgaaagcagcagtcaagagcataatacagtcagccgtcagcctacttcgg
 N  T  Q  V  K  A  A  V  K  S  I  I  Q  S  A  V  S  L  L  R
cccgaggattctgagcttgctgtggctgccaaatctgcccttaatttgttgacaacggaa
 P  E  D  S  E  L  A  V  A  A  K  S  A  L  N  L  L  T  T  E
gaatacggactggacaagtcctttggagatatcagcagcgccgttgttgaggacttcgac
 E  Y  G  L  D  K  S  F  G  D  I  S  S  A  V  V  E  D  F  D
ttggcatacaagcaattatcggcatttattggcagcgtggacagttacgtacacagtgcg
 L  A  Y  K  Q  L  S  A  F  I  G  S  V  D  S  Y  V  H  S  A
gcttcagtgaaacttcagaagttccacggcccgctcatcagtcaccgttcagaggcaacc
 A  S  V  K  L  Q  K  F  H  G  P  L  I  S  H  R  S  E  A  T
tcgatggtcagcagcgttcggcagctggtaagcgttttactaaatcagacgagcgaatcc
 S  M  V  S  S  V  R  Q  L  V  S  V  L  L  N  Q  T  S  E  S
gaaccttcatcggccagcgaccctgcaagaaagctaaatatacagagaaacgtcagcgaa
 E  P  S  S  A  S  D  P  A  R  K  L  N  I  Q  R  N  V  S  E
cctctcgagggcaacatccttctcgatctttttgcatcgaaatcatttgtagcgttcgac
 P  L  E  G  N  I  L  L  D  L  F  A  S  K  S  F  V  A  F  D
aacaccaccattgacttgcttaaagaatttatccaggacgagctccaagacctacgcgtt
 N  T  T  I  D  L  L  K  E  F  I  Q  D  E  L  Q  D  L  R  V
ggcaacagctacaacaaaacaattctggcgaactcagcatcggccagcataggtcttcct
 G  N  S  Y  N  K  T  I  L  A  N  S  A  S  A  S  I  G  L  P
tcaatactgggtttgccatgcaatataaaattcgacgcccctgcattagtttcgattcag
 S  I  L  G  L  P  C  N  I  K  F  D  A  P  A  L  V  S  I  Q
gtaaatgcgacggctgagattcacccagacattgagaagctcttgtcggatccccaacag
 V  N  A  T  A  E  I  H  P  D  I  E  K  L  L  S  D  P  Q  Q
cgtctccagtccctaagtgtgaagtctgaaatccaggccacgtacgcttcacaaatgagc
 R  L  Q  S  L  S  V  K  S  E  I  Q  A  T  Y  A  S  Q  M  S
tccaagatcaatatcatcttcccactcgagaagaaagtgtacgagtcatcgctcgtgcag
 S  K  I  N  I  I  F  P  L  E  K  K  V  Y  E  S  S  L  V  Q
aatctacaggtttacgttccggtgaagttgcatatttcccacaatctgccgcagaatctt
 N  L  Q  V  Y  V  P  V  K  L  H  I  S  H  N  L  P  Q  N  L
acgacgattaaggtccagcctttagacaagaaccaaaaatacccacttgtccgtgcgagc
 T  T  I  K  V  Q  P  L  D  K  N  Q  K  Y  P  L  V  R  A  S
cgcgcgcctaccaccgctgtgtggccggcgatcccactcatcaacgccggtgaaaaccgc
```

```
 R  A  P  T  T  A  V  W  P  A  I  P  L  I  N  A  G  E  N  R
cgcatcgtgcacgtgaggccgctggtgcaggctcagcgcgtgtttggtaaggactcaact
 R  I  V  H  V  R  P  L  V  Q  A  Q  R  V  F  G  K  D  S  T
ggcatggccttcatcgtagactacaagtctgaagccgaattcattgacaggaaggctgca
 G  M  A  F  I  V  D  Y  K  S  E  A  E  F  I  D  R  K  A  A
gccgatcacgcaaagaaattcgacgttttgtcgagtgtatacttcccatgggcttcggat
 A  D  H  A  K  K  F  D  V  L  S  S  V  Y  F  P  W  A  S  D
gaaatttattactacaaccacacactctactacgcaccagaagaatcgtcggccaaggac
 E  I  Y  Y  Y  N  H  T  L  Y  Y  A  P  E  E  S  S  A  K  D
gttacgatacaaattagagctctcaagaaggagtacacgaaatactttccgggcccggag
 V  T  I  Q  I  R  A  L  K  K  E  Y  T  K  Y  F  P  G  P  E
agcaggagtgacgagcagtctttggagtcagatgatgaagaggagccacgaagctacagg
 S  R  S  D  E  Q  S  L  E  S  D  D  E  E  E  P  R  S  Y  R
tccatgcagagacgtcatccaaaattccatcacaagagggcagctgattcttctgacgag
 S  M  Q  R  R  H  P  K  F  H  H  K  R  A  A  D  S  S  D  E
gactcttatgaacagcttgtgaatcaaatgaagaacaggaagcttgccaacgccgatttg
 D  S  Y  E  Q  L  V  N  Q  M  K  N  R  K  L  A  N  A  D  L
tttgtaaaaccgctggaacagagtctggaatcctttgaaaagggcattttagaacgtgga
 F  V  K  P  L  E  Q  S  L  E  S  F  E  K  G  I  L  E  R  G
acaacttacatccttgaaatgagagtgatgttccaacagcctcgtggtggagactaccgt
 T  T  Y  I  L  E  M  R  V  M  F  Q  Q  P  R  G  G  D  Y  R
ttgtcactgatgcactctcgtaacactctcagatctgttacattatcagacgcggaatgg
 L  S  L  M  H  S  R  N  T  L  R  S  V  T  L  S  D  A  E  W
tggtccaaaccagcaagaggacgcatttaccaggcgtctgctgaagcgattggtttattc
 W  S  K  P  A  R  G  R  I  Y  Q  A  S  A  E  A  I  G  L  F
ccactgttgcccgaaatgaacttcgaaaaggcgtataattccaaccctaactcaactata
 P  L  L  P  E  M  N  F  E  K  A  Y  N  S  N  P  N  S  T  I
tcattttcgtatctctcagaatatgaaggtgccaagaagtcgcacatctttgcacagctc
 S  F  S  Y  L  S  E  Y  E  G  A  K  K  S  H  I  F  A  Q  L
aaagcagggcgcagcgaaaggttgaaggtacagatcggcaagagtgacatcgccaagaaa
 K  A  G  R  S  E  R  L  K  V  Q  I  G  K  S  D  I  A  K  K
tgtcgtgaacaaatgcgagaaggaaacaacattttgtacgcatgccgaaacgccacagag
 C  R  E  Q  M  R  E  G  N  N  I  L  Y  A  C  R  N  A  T  E
gagtcaggattgctcgactacttttccgttaagggaagctttaacgacttgtctctggaa
 E  S  G  L  L  D  Y  F  S  V  K  G  S  F  N  D  L  S  L  E
ttcattaacagcacttaccaagtatactcatatgccaggcacctattgtacccgttcatt
 F  I  N  S  T  Y  Q  V  Y  S  Y  A  R  H  L  L  Y  P  F  I
tatgaaaaaccggatctgagagctgcccaccaggtgacgagcggcaaatttgacgggcat
 Y  E  K  P  D  L  R  A  A  H  Q  V  T  S  G  K  F  D  G  H
ctgtggttctccccgaaatgggattctttcaacgcctccctgaacgtaccagccctgtcg
 L  W  F  S  P  K  W  D  S  F  N  A  S  L  N  V  P  A  L  S
acgaacttcactaacgtcagaggatggccaaaaggcaaagtccaactgtggctgccgcat
 T  N  F  T  N  V  R  G  W  P  K  G  K  V  Q  L  W  L  P  H
ccatcacgacacgttgctgagaagctctccagaggataccaacatcgttacgtcgctcct
 P  S  R  H  V  A  E  K  L  S  R  G  Y  Q  H  R  Y  V  A  P
gcatgttccttggacaagcgcaacatgacaacattcgacaacctgactataccgctcgag
 A  C  S  L  D  K  R  N  M  T  T  F  D  N  L  T  I  P  L  E
ttgcccgaatgttggactctgattttgtcgcttgtaccgaaaccagtctatgacgaagaa
 L  P  E  C  W  T  L  I  L  S  L  V  P  K  P  V  Y  D  E  E
ttaagctacgtaaacgtcagtgttctcgcaatgcaggaacagtcaggcgcaaggaaattc
 L  S  Y  V  N  V  S  V  L  A  M  Q  E  Q  S  G  A  R  K  F
aagatacaagtcggaaaacactatattgagcaagtttcaacgaagtacattgtagtagac
 K  I  Q  V  G  K  H  Y  I  E  Q  V  S  T  K  Y  I  V  V  D
cgcatgaagcataacctacgttctgatgaaacaacgcacgtgaatgtggatggagcagtg
 R  M  K  H  N  L  R  S  D  E  T  T  H  V  N  V  D  G  A  V
ttcgcagtgggtgcgctcctacccagtggcagcagctggctgtccgtgccgctgtacggc
 F  A  V  G  A  L  L  P  S  G  S  S  W  L  S  V  P  L  Y  G
```

```
ctaagtctgcactacgatggagaacgcgccctcttgctggtcaacgagacctacaaggac
 L  S  L  H  Y  D  G  E  R  A  L  L  L  V  N  E  T  Y  K  D
tctgccagaggcctttgcggtacctacgatggcgaaacggtgaccgacttgaccacccc
 S  A  R  G  L  C  G  T  Y  D  G  E  T  V  T  D  L  T  T  P
gagaactgtatcttgaaggacccccgccatttcgtcgcacgctacgtcattccggagacc
 E  N  C  I  L  K  D  P  R  H  F  V  A  R  Y  V  I  P  E  T
tgccaaatcgacagtgtcaagcgcttgcagtcagaggcgaataacgttccctgctacccg
 C  Q  I  D  S  V  K  R  L  Q  S  E  A  N  N  V  P  C  Y  P
cacgaagtctttccggctgacaacatcatcgccaggcaatatcccagcccttacagagct
 H  E  V  F  P  A  D  N  I  I  A  R  Q  Y  P  S  P  Y  R  A
gctgcacgtagcgtggacagtgacgaatcatcctctagcgacgaagaatctgattctcgt
 A  A  R  S  V  D  S  D  E  S  S  S  S  D  E  E  S  D  S  R
gaatcgcgcaactacaagtctccgggcaacctccgtccggtgaaggtcgtcaaggcactg
 E  S  R  N  Y  K  S  P  G  N  L  R  P  V  K  V  V  K  A  L
gagcacgcgggcaagctggcattcagcgttgagaaggttaaacagtgcttgccgcctcta
 E  H  A  G  K  L  A  F  S  V  E  K  V  K  Q  C  L  P  P  L
gtggctgacaaggtggagtggaggaaagtcccataccacttcatcaacaaaacctctgct
 V  A  D  K  V  E  W  R  K  V  P  Y  H  F  I  N  K  T  S  A
ggcctccactggttccaacaagtgaagaagtcgccacagaacttcgatttcagccagaag
 G  L  H  W  F  Q  Q  V  K  K  S  P  Q  N  F  D  F  S  Q  K
tcgaagaacgccgaaatacgcaccgaagtccacttgtcatgcagacaaccatgaaatact
 S  K  N  A  E  I  R  T  E  V  H  L  S  C  R  Q  P  -  -  -
gacaaaatgtaaagaacttctaagtgtgcaaagatgttcacatggttagaaagatgaaat
aaatcttttctctcaatctgcaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Sequence A2: Nucleotide and deduced protein sequence of *S. gregaria* VG-B (GenBank accession # MK206970)

Underlined are transcription start sit and the signal peptide sequence. The poly-adenylation signal is doubled underlined.

```
acatggggaccagctgctccccagtcctcacagccatgtgggcgctgatcctcttggggctc
   -  -  -  -  -  -  -  -  -  -  -  M  W  A  L  I  L  L  G  L
ctcgttggggcaactgtgtccgctgccgaaccagatgcgtggtccaagggctacctctac
 L  V  G  A  T  V  S  A  A  E  P  D  A  W  S  K  G  Y  L  Y
cagtatgaattgaagtccaaggcacttacggctctacatgaagtcgccaaccagttttca
 Q  Y  E  L  K  S  K  A  L  T  A  L  H  E  V  A  N  Q  F  S
ggtgtgcagatagaaggaaacctacatgtccagctcataaattcgtcaactgtggccttc
 G  V  Q  I  E  G  N  L  H  V  Q  L  I  N  S  S  T  V  A  F
cagctggatgatttccgcgtggctgaggttcaccaaaatctgtctggtggctgggacgct
 Q  L  D  D  F  R  V  A  E  V  H  Q  N  L  S  G  G  W  D  A
gacattcctgagagccagttatctttccagagttttaaactcccctccaatgtgttctac
 D  I  P  E  S  Q  L  S  F  Q  S  F  K  L  P  S  N  V  F  Y
ctcaactacacaaacggagttatttctaacgtaactgttagtaagaaagtgccattatgg
 L  N  Y  T  N  G  V  I  S  N  V  T  V  S  K  K  V  P  L  W
ttcgtcaactgggtcaagggacaggccagtgtattccagttggacgtccggggctccaac
 F  V  N  W  V  K  G  Q  A  S  V  F  Q  L  D  V  R  G  S  N
agagtcagtaggcccgtcccaggcggcgacagtgatatctactggacgatggaggacgtc
 R  V  S  R  P  V  P  G  G  D  S  D  I  Y  W  T  M  E  D  V
attaccggcaaatgtcaaacggcatacgtgtacgctccctatcccaaatatgctgctctg
 I  T  G  K  C  Q  T  A  Y  V  Y  A  P  Y  P  K  Y  A  A  L
tccgacaatgacatcgaacccgagaactgggaggacgacagcaaactgtacaccgtatct
```

```
 S  D  N  D  I  E  P  E  N  W  E  D  D  S  K  L  Y  T  V  S
aaagcagtgaactacacgaaatgtgagaaacgagctggttacacctggggcttgaacact
 K  A  V  N  Y  T  K  C  E  K  R  A  G  Y  T  W  G  L  N  T
gtgacccgctgggagcctattgaaaactccgtaggacaaactatatcgcgctcgtcgaac
 V  T  R  W  E  P  I  E  N  S  V  G  Q  T  I  S  R  S  S  N
acagaggccgttatcagtggttccattcggaagtacacaatccagaagtcggtaacgacg
 T  E  A  V  I  S  G  S  I  R  K  Y  T  I  Q  K  S  V  T  T
aacagagtcattctgaacccactggtatacaaccaacagaagggcatgctcggaagtcag
 N  R  V  I  L  N  P  L  V  Y  N  Q  Q  K  G  M  L  G  S  Q
gtgaaaatcagtttgcagaggaagatgaagtctgacaaacccattccaaaaatctcggac
 V  K  I  S  L  Q  R  K  M  K  S  D  K  P  I  P  K  I  S  D
cccgtgacctacgacgaccttgtgtacgagtacgaggggatgtttccactgaaacacgct
 P  V  T  Y  D  D  L  V  Y  E  Y  E  G  M  F  P  L  K  H  A
gacgaagctcagggcagaggctacagcgacgaacccagcgaagagtctgaggaacgggct
 D  E  A  Q  G  R  G  Y  S  D  E  P  S  E  E  S  E  E  R  A
tacaggcctaggagacagttcaagcgtgcggtttgggacaatgccgacgaatccgacagc
 Y  R  P  R  R  Q  F  K  R  A  V  W  D  N  A  D  E  S  D  S
agcagcagcagcagcagcagcgaagaggatcttgaaattcccgagcccgatcttttgcaa
 S  S  S  S  S  S  S  E  E  D  L  E  I  P  E  P  D  L  L  Q
ccgcctttgctaccgttcagcaacatggagcagcctcaagacaagccagacatcagcgct
 P  P  L  L  P  F  S  N  M  E  Q  P  Q  D  K  P  D  I  S  A
atcattgaattggttgagaagattgctcgtcagattcaagagcctagcagcatcccatca
 I  I  E  L  V  E  K  I  A  R  Q  I  Q  E  P  S  S  I  P  S
agcgacacactagcctcattccagatcctcaacagagagctgatgaaatacggcgtaacg
 S  D  T  L  A  S  F  Q  I  L  N  R  E  L  M  K  Y  G  V  T
tctctgcgccaaatagctaacaagctttatgcaccgtacccggcagattcgcacagctac
 S  L  R  Q  I  A  N  K  L  Y  A  P  Y  P  A  D  S  H  S  Y
ccgcaatctgtaaaatacacccagtgggttgtcttcatcaatgccctttcacagtcgggc
 P  Q  S  V  K  Y  T  Q  W  V  V  F  I  N  A  L  S  Q  S  G
actggaccagcaatcagaactattaacgagcagatcaaactgcagaagcttcgaggacgc
 T  G  P  A  I  R  T  I  N  E  Q  I  K  L  Q  K  L  R  G  R
cttgccgccatcgccatcggaaatattccgcgcagtgctctgtaccccaacaaagcattc
 L  A  A  I  A  I  G  N  I  P  R  S  A  L  Y  P  N  K  A  F
ctcaaggaattcttcatgatagtgaagaacacagagaagtatagtcctgagaacgcaact
 L  K  E  F  F  M  I  V  K  N  T  E  K  Y  S  P  E  N  A  T
gctcttctgggatttggtaacctcctgcgcctcgctatcgttgacagagacagttctcac
 A  L  L  G  F  G  N  L  L  R  L  A  I  V  D  R  D  S  S  H
aatctctacccagttcacttatacggcaagttagacagtccgagcgagtctgagctccaa
 N  L  Y  P  V  H  L  Y  G  K  L  D  S  P  S  E  S  E  L  Q
gcatacgtgaagtacctgggtaacctcctcaaccgtgctgttcaaaatgctgacagtgtc
 A  Y  V  K  Y  L  G  N  L  L  N  R  A  V  Q  N  A  D  S  V
ggtattcaagtatacaccagagccctgggaaatatcggacacccatacatcctgaaacct
 G  I  Q  V  Y  T  R  A  L  G  N  I  G  H  P  Y  I  L  K  P
ctacttccgtatatcctgggagagaagcaagtctccaacttccaacgacttttgatggta
 L  L  P  Y  I  L  G  E  K  Q  V  S  N  F  Q  R  L  L  M  V
ctggctctcgacagggttgctgaactctatccccaaatagtgagaccagttttggtccag
 L  A  L  D  R  V  A  E  L  Y  P  Q  I  V  R  P  V  L  V  Q
atttaccagagcactggtgaaacacaccagatccgcagcactgcagttttgcttatcatg
 I  Y  Q  S  T  G  E  T  H  Q  I  R  S  T  A  V  L  L  I  M
ggttccaaaccttcggcaagtgtcctgcagaggctggctcagttttcgaagcaagatccc
 G  S  K  P  S  A  S  V  L  Q  R  L  A  Q  F  S  K  Q  D  P
aaccctcacgtcgtctctgttgtcaagacagctattcagagtgctgccagactgacaacc
 N  P  H  V  V  S  V  V  K  T  A  I  Q  S  A  A  R  L  T  T
cctgaaaatgaggatctagccagaagtgctatggcagccgtcaatatgctgagccaaaat
 P  E  N  E  D  L  A  R  S  A  M  A  A  V  N  M  L  S  Q  N
aaaaccgcggtacagtactcggccaagcatttacaagattacttaattcgtgagatggca
 K  T  A  V  Q  Y  S  A  K  H  L  Q  D  Y  L  I  R  E  M  A
```

```
ttatcatacaatctgaagtacagccagctcggcagtaaggacacacttatacctaatgca
 L  S  Y  N  L  K  Y  S  Q  L  G  S  K  D  T  L  I  P  N  A
ctgtacgctggtctgaaactgaacattggaggattccagcaccttgcaggagaagcacaa
 L  Y  A  G  L  K  L  N  I  G  G  F  Q  H  L  A  G  E  A  Q
gccgcagttagcagcattaagaccctacaacatctgcttagatacatcattgtagaggga
 A  A  V  S  S  I  K  T  L  Q  H  L  L  R  Y  I  I  V  E  G
gatgaaaatgagaatcaaccaacgaaggatccagttatcaacatcgtatctgacttgagg
 D  E  N  E  N  Q  P  T  K  D  P  V  I  N  I  V  S  D  L  R
gaacaacttgagggtcaaatagcaggaaacgttatggcccatccgtctataattgccttt
 E  Q  L  E  G  Q  I  A  G  N  V  M  A  H  P  S  I  I  A  F
gacaacaacacaattaactcatgggtggaagctatcaggaacaatgtcaacaagctacag
 D  N  N  T  I  N  S  W  V  E  A  I  R  N  N  V  N  K  L  Q
aatgggctgccttttaactgcactagtattgtaagcgcagttgatgtcaggatctttttc
 N  G  L  P  F  N  C  T  S  I  V  S  A  V  D  V  R  I  F  F
cccaacgctctgggattcccaagttccatcgtatacgagactccagttctctactcagtt
 P  N  A  L  G  F  P  S  S  I  V  Y  E  T  P  V  L  Y  S  V
ggtggagaactgcgcctgaagacttcaccgaaactcaacagtgctcctaagggccacctg
 G  G  E  L  R  L  K  T  S  P  K  L  N  S  A  P  K  G  H  L
gcctggccgaccgtatggaatatcactgcagatgtgagggctgtatactcaagaagaagc
 A  W  P  T  V  W  N  I  T  A  D  V  R  A  V  Y  S  R  R  S
tcgggatccattttattcactgtcctccctctcgacaagagttattctgccggatatgtt
 S  G  S  I  L  F  T  V  L  P  L  D  K  S  Y  S  A  G  Y  V
aagaaccagcaatatcaagtaccagtaagattcctcgtcaacgttgatctcgaagacaat
 K  N  Q  Q  Y  Q  V  P  V  R  F  L  V  N  V  D  L  E  D  N
agcacatatgtgaaactacagcccatcaacaagaaccacaggtaccagctgtcacacatg
 S  T  Y  V  K  L  Q  P  I  N  K  N  H  R  Y  Q  L  S  H  M
agcagcgttccttacacaacgatctaccgcatgtcccccgtggaaagtgctgctcttacg
 S  S  V  P  Y  T  T  I  Y  R  M  S  P  V  E  S  A  A  L  T
gcgccagaaactgaaatagttcacactcgcaagccgcgctcgtggaaggctcagtacgga
 A  P  E  T  E  I  V  H  T  R  K  P  R  S  W  K  A  Q  Y  G
agatcaactggtatggtgtattccgtagagtacaactcggaaaaggattacgatgacatg
 R  S  T  G  M  V  Y  S  V  E  Y  N  S  E  K  D  Y  D  D  M
tacgagagataccaaaaatatgcaaaccgcgatgccctttctgctatcttgttcttcaat
 Y  E  R  Y  Q  K  Y  A  N  R  D  A  L  S  A  I  L  F  F  N
gcggagcagcaggtctactacagcaacatcagcctttattatgagccagaaaagtccatc
 A  E  Q  Q  V  Y  Y  S  N  I  S  L  Y  Y  E  P  E  K  S  I
tccaaagccgtggaactcagtgtccactattatgacgaactctcttcgtcgtcatcacgc
 S  K  A  V  E  L  S  V  H  Y  Y  D  E  L  S  S  S  S  S  R
cgtcgacctgagcaatctgaagagtcacaggagtctgaagaccagcccgagtccagggta
 R  R  P  E  Q  S  E  E  S  Q  E  S  E  D  Q  P  E  S  R  V
gcgtcggaggagatgtctgacgaaccagagagcgctgagtcaagccaagagaagtactac
 A  S  E  E  M  S  D  E  P  E  S  A  E  S  S  Q  E  K  Y  Y
agatccaaacgaggagcaatgttctggggaagagccaagcacaggagagaagcctccggt
 R  S  K  R  G  A  M  F  W  G  R  A  K  H  R  R  E  A  S  G
caaaacagagccgcactgagtaacgtgaatcccaacgaatcacataaggaactggtacag
 Q  N  R  A  A  L  S  N  V  N  P  N  E  S  H  K  E  L  V  Q
agattgaaggaacaaggacttgcgaagagtgtcaccacgaaggcgctgttggttggagtg
 R  L  K  E  Q  G  L  A  K  S  V  T  T  K  A  L  L  V  G  V
aattttgtagaaggcaaaggcgcttcttcgctgctggcaatcgtctggagctccagtcca
 N  F  V  E  G  K  G  A  S  S  L  L  A  I  V  W  S  S  S  P
gtgtccagcaaatcgcaactgctcgctttcgcgtcagttaaaggagcacaggctagcaca
 V  S  S  K  S  Q  L  L  A  F  A  S  V  K  G  A  Q  A  S  T
ccttaccaggtttgtgtggaggcacaatccgtgttccccaacgttccattcatgaacttg
 P  Y  Q  V  C  V  E  A  Q  S  V  F  P  N  V  P  F  M  N  L
gagaaagctctgaagaccaatccaaactcgtctgtatctgtcagactagattctggagaa
 E  K  A  L  K  T  N  P  N  S  S  V  S  V  R  L  D  S  G  E
ggctcgtgcttgtcaggaatgtcaattcaggccaatcttgatattggaagaagccaaagc
```

```
 G  S  C  L  S  G  M  S  I  Q  A  N  L  D  I  G  R  S  Q  S
atgtaccaatacctgcagaagtctagactcgtggcgcaatgcagatctcaaatgaaggag
 M  Y  Q  Y  L  Q  K  S  R  L  V  A  Q  C  R  S  Q  M  K  E
gataactacgtcctgcctgcatgtcgtaacgccacaattaaggctggtgtgctcaacaat
 D  N  Y  V  L  P  A  C  R  N  A  T  I  K  A  G  V  L  N  N
tacgacctagaggtcgagtatcagaatattccagaagcactgaagcattcgatttacaga
 Y  D  L  E  V  E  Y  Q  N  I  P  E  A  L  K  H  S  I  Y  R
gcttactccaacatggattacgctcatttcctttatggatcacaaaacgtcgtaaatggt
 A  Y  S  N  M  D  Y  A  H  F  L  Y  G  S  Q  N  V  V  N  G
accggtgctcctggaaaactgtatgccaaccttatagtcgcacccaacatgcgctcgctc
 T  G  A  P  G  K  L  Y  A  N  L  I  V  A  P  N  M  R  S  L
aacttctccttctcatctgcatacttggacgccgcctacgacaacgttcgcctcgtgatg
 N  F  S  F  S  S  A  Y  L  D  A  A  Y  D  N  V  R  L  V  M
cccgtagcaaaagccatcgtcagccatccactcttgcccagagcagaacgcgtcgctaac
 P  V  A  K  A  I  V  S  H  P  L  L  P  R  A  E  R  V  A  N
tactacaccaactggcagtataacgcaacctgcagcgtggattcttcatacattaagact
 Y  Y  T  N  W  Q  Y  N  A  T  C  S  V  D  S  S  Y  I  K  T
tacgacaatctaacatacccgttcgagtcgaagagcccctgctggagaatcctgctcgcc
 Y  D  N  L  T  Y  P  F  E  S  K  S  P  C  W  R  I  L  L  A
acagctcgtcagcgcgacgtgaaaccatccccacttgtcccgtcacctaacgtcacagtg
 T  A  R  Q  R  D  V  K  P  S  P  L  V  P  S  P  N  V  T  V
atggttcgggaaatctcagacaagagggaagtcagagtgctggtcgatgacaacgttgtg
 M  V  R  E  I  S  D  K  R  E  V  R  V  L  V  D  D  N  V  V
tcactagcataccagtcgggcaagtacatgctgagagccaacagcaaggccgctgccgtc
 S  L  A  Y  Q  S  G  K  Y  M  L  R  A  N  S  K  A  A  A  V
tctcagtcagaggtgacgcaactgaacgacaaggaaggagagccgctggtgctggcctac
 S  Q  S  E  V  T  Q  L  N  D  K  E  G  E  P  L  V  L  A  Y
gccctcccagactctgtcaggatggaaatacaggatctcatcatctactttgaccagcag
 A  L  P  D  S  V  R  M  E  I  Q  D  L  I  I  Y  F  D  Q  Q
cgtgtactgttacaaccgagcgacatttacaggcaagcagtgcgcggtctgtgcggtacc
 R  V  L  L  Q  P  S  D  I  Y  R  Q  A  V  R  G  L  C  G  T
ttcgacggcaaacgacagacggacttcaaacttccggccaactgcattgtgcgaaacgtc
 F  D  G  K  R  Q  T  D  F  K  L  P  A  N  C  I  V  R  N  V
acggcgttcatcgaagcctacacgtatggtgacaaatgcgcagctcctcgcagagcacag
 T  A  F  I  E  A  Y  T  Y  G  D  K  C  A  A  P  R  R  A  Q
ccgaagcaatgctatgaggagaagatcattcctggaaatcaaactactcccgagccacta
 P  K  Q  C  Y  E  E  K  I  I  P  G  N  Q  T  T  P  E  P  L
actgacgcaacatgcgtctcactccagaacagggttgtgctgcggggtcagaaagtctgc
 T  D  A  T  C  V  S  L  Q  N  R  V  V  L  R  G  Q  K  V  C
ataagccgaaagcgcctgccgatgtgcaaatgggaatgcaaggcagtagacaccgtcgac
 I  S  R  K  R  L  P  M  C  K  W  E  C  K  A  V  D  T  V  D
aaggtggttccattccgctgctactccagaactgagactgttgaaactgtcgccgctcaa
 K  V  V  P  F  R  C  Y  S  R  T  E  T  V  E  T  V  A  A  Q
gctcaacggggacgccacgttgacctctacagcataccaattgaaggtgacgataccaat
 A  Q  R  G  R  H  V  D  L  Y  S  I  P  I  E  G  D  D  T  N
atcgaggtttcggtagcaactgcctgcacacgtctggaggagtcagaagaaagcgacagt
 I  E  V  S  V  A  T  A  C  T  R  L  E  E  S  E  E  S  D  S
gacgaaaattaacagcttcatgttaaacgctaacgaaaaagcactattcgaccaaaacaa
 D  E  N  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
tatgtactaaaactgtgcaaataaaatctctttagacgaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaa
```

Table A1: List of primers used to obtain the sequence of VG-A

Primer Name	5`- Sequence -3`	Source	Annealing temperature °C
Sg-VG-F Sg-VG-R	AAGCAGCTCGTACACCTATTC GCTGACGGCTGACTGTATTAT	Est database LC01027B2G10.f1	52
Sg-VG-F Sg-VG-R	ATGAAGACTTGCTGGGTGAATA GGAATACGTTCTGCGCTATCT	Est database LC01027B2G10.f1	52
Lm-VG-F Lm-VG-R	CTCTGACAAGTACCCACTCAAC CTGTCCACGCTTCCGATAAA	*L. migratoria* sequence	51
Lm-VG-F Lm-VG-R	GCTACAAGTCTGACCGCATTA TAGCCCAGGAGGTAGATGTT	*L. migratoria* sequence	51
Lm-VG-F Lm-VG-R	AGACCTGCTGGGTGACTATAA CTGTCCACGCTTCCGATAAA	*L. migratoria* sequence	51
Lm-VG-F Lm-VG-R	CGTCCAACAGCCGAGTTATT CGAATACCCGAAGAACGCTTAT	*L. migratoria* sequence	51
Lm-VG-F Lm-VG-R	GTGGCTGTAGGAGGTAAATCTG GTCTTCCAACTGTTTCACGTATTC	*S. gregaria* sequence	52
Sg-VGA-F Sg-VGA-R	CTGTACTCTATGACCGACAAGTAAAT ACCCAGCAAGTCTTCATAGTTATC	*S. gregaria* sequence	52
Sg-VGA-F Sg-VGA-R	ATACCCACTTGTCCGTGCGAGCCG TGTAGAGCACGTCGAAAGGAGCC	*S. gregaria* sequence	65
Sg-VGA-F LUP	CGGTACCTACGATGGCGAAACGG	*S. gregaria* sequence	60
Sg-VGA-F LUP	CGAAACGGTGACCGACTTGACC	*S. gregaria* sequence	60
Sg-VGA-F LUP	GCACAGCTCAAAGCAGGGCG	*S. gregaria* sequence	62
Sg-VGA-R LUP	GCTCGCAATCATTCCCTTCTG	*S. gregaria* sequence	60
Sg-VGA-F Sg-VGA-R	CAATACTGGGTTTGCCATGC CTGATTCTTCTGACGAGGACTC	*S. gregaria* sequence	52.5
Sg-VGA-F Sg-VGA-R	TGGTCCAAACCAGCAAGAG CCTCTAGCGACGAAGAATCTG	*S. gregaria* sequence	52.5

LUP is the Long universal primer from SMARTer RACE kit.

Table A2: List of primers used to obtain the sequence of VG-B

Primer Name	5`- Sequence -3`	Source	Annealing temperature °C
Lm-VGB-F Lm-VGB-R	AGG AAA CCT GTA TGC CCA GC ACT TTG ACC TGG CTT CCG AG	*L. migratoria* sequence	53
Lm-VGB-F Lm-VGB-R	CGACCAGATCAAACAGCAGA ATAAGCAAAACTGCCGTGCT	*L. migratoria* sequence	53
Lm-VGB-F Lm-VGB-R	GCCGTTATCAGTGGTTCCATTC GCCGTTATCAGTGGTTCCATTC	*L. migratoria* sequence	54
Lm-VGB-F Lm-VGB-R	TGAAACACACCAGATCCGCA TGGTTCTTGACATACCCGGC	*L. migratoria* sequence	55
Sg-VGB-R LUP	TGGCTCTCAGGAATGTCAGC	*S. gregaria* sequence	62
Sg-VGB-F Sg-VGB-R	GGGAAATATCGGACACCCATAC GTCTGGCAGCACTCTGAATAG	*S. gregaria* sequence	54
Sg-VGB-F LUP	TTGGTGGAGAACTGCGCCTGAAGA	*S. gregaria* sequence	65
Sg-VGB-F LUP	TCCAAACGAGGAGCAATGTTCTGG	*S. gregaria* sequence	60
Sg-VGB-F LUP	ACTACACCAACTGGCAGTATAAC	*S. gregaria* sequence	60
Sg-VGB-F LUP	ATGGGAATGCAAGGCAGTAG	*S. gregaria* sequence	60
Sg-VGB-F LUP	CACGTTGACCTCTACAGCATAC	*S. gregaria* sequence	60
Sg-VGB-F Sg-VGB-R	GCTGGTGTGCTCAACAATTACG CTCCTTCCTTGTCGTTCAGTTGC	*S. gregaria* sequence	55

LUP is the Long universal primer from SMARTer RACE kit.

Sequence A3: Assembled CDS of *L. migratoria* VG-A

>L. migratoria VG-A, complete coding sequence

ATGTGGGCCGTGATCCTCTTGGGGCTCCTCGTCGGGGCAACTGCGTCCGAACAGCCAAGCCTATG
GAAACCGGGATATGCGTACGTGTACAATGTTACTGCTACAAGTCTGACCGCATTACACCAGCTAG
CTGACCAATGGAGTGGTCTAATCTACAACGCCACCCTGGTTGCCATCCCGAGATCGCCGTCAGAG
GTTCAGTTGCGGATCAACAGCGCAATCTATAGCAAAATTCATGCAAACCTCTCTGGAGGATTTGA
CAATCCTATCCCAGAGGAATACTTGAATTGGAAGCCTCTGCGGCTGTCGAAGGCCCCTTTCGACG
TGATCTACAAGGACGGAGAGATCGTCGATATGCTCGTCGACTGTTCGCTCTCTGATGCTGAAGTC
AATCAGCTGAAAGGTCTTGCAAGTATGTTCCAAGTTGACATCGATGAAGATAAGGGTCCGATATT
TAAGAAGTACGAGCAAACTGTAACGGGAGACTGTATGACTCAGTACGCCAGGACTGTAATCCCGA
GGTACATTCTCCAGAGTGAGCAAAAGTTGGAACTACCCGGACTCATCGACGATAAGGCCAATATA
CTCATGGACATCACGAAGGAGCGGAATATCTCAAAGTGCCACGACGACCGTATTTCCTACTTCGG
TCTGGATCAGAAACCATTCAACAGCAGCTCGGCACCTTCTCTGGCTCTTTCGTCCAACAGCCGAG
TTATTGTGGCCAACAGTTCGAACATCTACCTCCTGGGCTACGCAGAGACAACAGAAAAGACTGTC
ATTAGTCCTTTCTTGTACAACCAGCTGAAGGGAATGGTTGCGAGCACCATGCGTGTAGTGTTGAC
TAAACATCACAGGCTCGTCCAGCCAGAAAAAGGTGCCCAACCCCAATGCAGCACCGGAGGTCTCT
TGTACCACCCCTTGAGCTCACCTGAGGGCACCAACACCAAAGCCAGCAACTACGGCAGTGTAGAA
AATCGCGAACAAGCCTCAGATGAAGCATCAGACAGTTCAGAGAGTAGCGCAGCTAGTTACGAGGC
GCCGTCGAGAAGGTACCGGCGAGCAGTTAGGCCATACCAACAATTCTTGCCAGAGCCTCAACTCC
AAAGCGCACCTCACCAGCCGTATGACGCTCTTTCCGTAGCTGTAAAAGGCCAATCAATCCGCTAT
GACCAACAAGTAAATGCCTCTAAAGTAGCTATTTCACTGGCAAGAGATATTGGTCAAGCCCTCTA
CGAGCCAAGCCTGATTGCACCGAAAGAAGCTCTGACAAAATTCCCGCTGCTGATAAGCGTTCTTC
GGGTATTCGGACACGTCAGTCTCCAACGCCTTGCTGACCAACTGTACGAGGAGCAGTCTCCAGAA
AATCCGAGCTCGCAGTATGCATGGACTGCCTACGTCATGGCTGTAGCCCAGGCCGGTACTGGCCC
CGCTCTCCTGAACCTCTTCTACATGATCAATGAGGGCATGCTCAGTAACAGGGTCGCTGCAAGCG
CTTATGCGTACATCCCTCAAGCAGCTCGCACACCAACTCCGCAATACGTCGAGAAAGTCTTCAAC
TTTGCCCTCCGTGACGCCTCTGACAAGTACCCACTCAACGTCACCGCAGTTCTAGCAGCCACTGA
ACTCATCCATTTGGCTCAAGTAAGCAACGTATCCAGTTCACGCTACCCCGTAGAAACCTTTGCAG
ACCTGCTGGGTGACTATAACCCCGACTACGTTCAACGATTGGCAGATAAGTTGAAGCGCGCTGTA
GATGAAGCTTACAGTCCGAATATTCAACTGTACATCAGAGCCCTGGGTAACATTGGACATCCAGA
CATCGTGAAAGTCTTCAGGCCTTACTTAGAAGGAAGCAAACCTGTATCTACCTTCCAGCGCTTTA
TGATGGTCATGTCCCTGGACAAACTTGCTGAAGCTTACCCACAAGAAGCACAGAACGTATTCATG
CTGCTCTTCCAGAACACTGGAGAGACTCACGAGATCAGATGTGCTTCAGCTCTTCTGTTGATGGC
GACCAACCCACCCCCGAGCGTCCTTCAGAAGCTGGCCCACCTCACAAACAGTGAGCAAAACACTC
AGGTGAAAGCCGCAGTAAAGAGCATCATACAGTCAGCCGTGACCCTTCTTCGCCCCGAAGACTCT
GAGCTGGCTCAAGCTGCTCAATCTGCCATTCACATGTTGACCACGGAAGAATACGGAATGGATAG
ATCCTTTGGAGATGTCAGCAGTTCCGTTGTTGAAGACTTTGGATTGGCATACAAGCAGCTCGCGG
CATTTATCGGAAGCGTGGACAGCTACGTACACAGTTCGGCTTCAGTGAAACTCCAGAGGTTCCAT
GGCCCGCTCATGAGTCACAGTTCAGAGGTAACTTCGATGGTCAGCAGCATTCGCCAGCTGGCACA
CGTTTTGCTGAACAATACGATCGAATACCAACCTTACTCCTCTAAGGACCCTGCAGGCAGGCTGA
ACATACAGAGAAACATCAGTGAACCACTCGAAGGCAACATCCTTCTTGACGTTTTAGCATCGAAA
TCATTTGTAGCCTTCGACAACACCACCATTGAACTGTTTAAAGAAGTCCTCCTGGAAGAGCTTCT
GGACCTGAAATTGTGTCAATGCTACAACAAAACTATTCTGGCAAACTCAGCGTCGGCCAGCATCG
GCTTCCCTACAATCTTGGGATTGCCAGCTAATGTAAAATTCGATGCACCTGTACTAATTTCGATA
CAGGTAAATGCGTCGGTTGAAGTTGAGCCAGACATCGAGGAGGTCTTTGTCAACGCCAAGAAGCA
CCTTCACTACGTCAATGTAACGTCTGAAATCAAAGCTGCATACGCCTCACAAATGAACTCCAAGA
TCAATATCATCTTCCCACTCGAGTCGAAGGTGTTCGAGTCATCGCTCGTGCAGAATCTCGAGCTG
TACCTTCCACTGCAACTGAACATTTCCCACGACCTGCCGCAGAACATGACAACGTTCGAGTTGGA
ACCACTTGACGCGGAACAGAAATACCCGATTTTCCGTACAAGTCGCCAGCCCACCACTGCTGTGT
GGTCGCCGATGCCTCTCACCAACGCCGAACGAAACCGACGCATTGTGCACGTCAGACCATTGAAG

CAGGCTCAATACGTATTTGGTAAACAAACTACTGGCATGGTCTTCATCGGAGAGTACAAATCGGA
AGGCGAATTCAATGACAAGAAAACTCTGGCTGATCACGCAAACAAGTTCGACGCTTTATCTGCTG
TATTATTCCCATGGGCTTCAGTGGAAACCTATTATTACAACCATACACTCTACTATTCACCACAA
CAATCGTCGGCCAAGAACGTCGTCATTCAGACCAGAACCGTCGTTTTGAAGAATGTGAAATTTAT
TGCGGCGCCAGAGAGCAGGAGTGACGAACAGTCTCTCGAGTCTGACGATGAAGAGCAGCCAAGGA
GCTACAGACGCACACAGAGACCTCATCCGAAATTACGACCAAGGAGGGATGTATCTCCCACTGAA
CTTGAATATTTTGAGAAGCTCGTGCTCCAAATGAAGAACAAGGAACTCGACAACTCGGATCTCTT
TGAACAACCTCTGGAAGAGAATCTGGATGACTTTGAACAGGACATCTCAGAACGTGGAACTACCT
ACATCTTAGAAACGAAAGTGACGTTCCAGCAGCCTCGTGGTGGAGAGTACCGTTTGTCACTGATG
CATTCCCACAACAGTATTGGGACTGTTACACTGACCAACGCAAAATATTGGGCCAAATCAGTACT
TGGACGCGTATTTCAGGGGTCCGTTGACACGTTGGGTTTCTTCCCACTGTTGCCCGAAATGAACT
TCGAACAGGCGATCAACTCCAATCCCAACTCGAGCATAGGAATTTCGATCATCACAGAACATGAA
GTTGGAAAGGAGTCAAGGATCTATGTTGGGATCAAAGCAGGACGCAGCGAAAGGCTGAAGAAACA
GATCGACAAGGGCGACGTCGCTAAGCAATGCCGCCAACAAATGAGGCAAGGAAACAACATTTTGT
ACGCATGCCGAAACGCCACGGAGCAGTCAGGATTGCTCGACAACTTTTCCATTAAGGGCAACTTC
AAGAACTTGTCTCCAGAGTTCATTAACGCCACTTACAAAGCATACTCAGCTGTCAGGTACCTGCT
GTACCCATTCATTTCTCAGAGGGCAAGTCACAGCCCTGTCCCGCCGGTGAGGACTGGCGAATTTG
ATGGCTATCTGTGGCTCTCGCCGCAACTGGATTCCTTCAATGCCTCCCTGAACGTCCCATCCCTG
TCAGTGAACTTCACTAACGTGAGAGGATGGCCAGCAGGCAAAGTTCCTCTGGTGTTGCCACACCC
AGCACGACACGTCGCTGAGAAAATCTCCAGTGGATACGAGCATCGTTATGTCCCTCCTACTTGCT
CTTTGGACAGGCGCAACATGACAACATTCGACAACCTGACTATACCGCTCGACCTGCCCAAATGC
TGGACCCTCGTCCTGTCGCTTGTACCGAAACCATTGTACGACGAAGAGTTGAGCTACACGAACGT
CAGCGTTCTCGCAATGCAGGCGCAGGCAGGCTCAAGGAAATACCAGATACAAGTGGGAAGTCACA
TCGTTGAGCAAGTTTCACCGAAGTACATTGTAGTGAACCACACGAAGTACAACCTGCAGCAGCAC
GAAACGACGCGCGTGATGGTGGACGGCGTAGAGTTGCTCGAGGGTGCAATTCTACCGAGCGGCAG
CAGCTGGCTCTCGTTGCCACGATACGGCCTCACACTGCACTTCGACGGAGAACGAGCCCTATTGC
AGACCAACGAGACCTACAAGGGCGCCGCCAGAGGTCTCTGCGGAACCTACGACGGCGAAACCGTG
ACCGACCTGACCGTCCCCAGAAACTGCATCTTGAAGGATCCCAACAGCTTCGTCGCCCAGTACGT
CGTCCCGGAGACCTGCCAGGACGAAAGCGCCAGGAGCTTGCTGTCAGAGGCCAAGAATGCTCCCT
GCTACCCGTTTGAGATCTTCCCGGCTTCCAACATCAACGCCAGACCGGGCCCCGGCCAGTACGGC
CCGTACGGACCTTCGGCACGCAGCGCAGACAGCGACGAGCTGTCCGCCAGCGACGAGGATTCCGC
TTCCGACGAAGCACCCAGGTCCCGCAGCTCTCGACCAAGCCTCCGTCCGGTGAAGATCGTCAAGG
CTGTGGACCGAGGCGAGAAGGTGGCATTCAGCATTGAAAAGGTGCCGGTGTGCCAGAAACCTCTC
GTGCCTGACCAAGTCGAATGGAGGAAGATCCCATACCACGAAATCAGGAAAACCGAGGCTGGTCG
CCACTGGTTAGAGAAGGTCCAGAACGCACCGCATAACTTCGACTTCAGCAAGAAGTCCAAGAGCA
CCGAGATCCGCACCAAGTTCCACCTGTCATGCAAGCAACCA

Sequence A4: Assembled CDS of *L. migratoria* VG-B

>L. migratoris VG-B, complete coding sequence

ATGTGGGCGCTGATCCTCTCGGGGCTCCTCGCTAACACTAAGGACGGCACAAATATCCATGCCCG
AGGGGGGATTCGAACCCCCGACAGTAGTGGCCGTGCAGACCGTGACTGGAGCGCCTCAGTCCCCT
CGGCCACCGCGGCCGGCAGTGGTGACACACTGGACTCGCATTCGGGAGGAGGGCGTTCTCATATT
AGTAACGCGTTCATTCAAAATTCTTCAGCCGTAGCGTTACAGGTGACTACCCTACGCGGACAGAC
TCAAAGAATCCGACGGCGAGGTTACAACAGGCGCGTTGTAGAAGAACACACTCTTCAAATTTTGG
TTTGCAAGTCTTGGGGCAACTGTGTCCGCTGCCGATCAAGATGCCACGCGTTCAGTTACAGACAG
GACTGTCAGAAACTGCTGAATAATCATGTTATGTGTGTAATTCCAGCTTGGTCCAAGGGCTACCT
GTACCAGTATGAAGTTAAGTCCAAGGCACTTACGGCTCTTCATGAAGTCGCCAACCAGTTCGCTG
GAGTGCAAATAGAAGGAAACCTGTATGCCCAGCTCATTAATTTGACAACTTTGGCTGTCCAGCTG
GAAGACTTCCGCGTAGCTGAAGTTCACCAGAATTTGTCTGGCGGCTGGGACGCTGAGCTTCCTGA
GGGCCAGTTCCAGAAGTTCCCACTCCCCTCCAATGTGTTCTACCTCAACTACACAAATGGAGTGA
TTACTAACATCACTGTTAGTAAGGAAGTGCCGTTATGGTTCATCAACTGGGTCAAGGGACAGGCT
AGTGTATTCCAGTTGGACATCAGGGGTACCAACAGAATCAACAGACAGGTCCCAGGCGGCGACAG
TGATATTTACTGGACGATGGAGGACGTCGTTACCGGCCAATGTCAAACGCAATACGTTTACGCTC
CCTACCCCAAATATGCCGCTCTATCCGACGATGACATCAGCCACGAGAACTGGAAGAATGACAAC
AACTTGTACACCGTGTCCAAATCTGTCAACCTAACGAAATGTGAAAAGCGACCCGGCTACACCTG
GGGCTTGAACAGCGTAACACGCTTGGAGCCTGTTGAAAACTCAATGGGAGAAACTTTGTCGCGCT
CGTTGAACGCAGAGGCCGTTATCAGTGGTTCCATTAGCAAATACACCATCCAGAAGTCGGTGACG
ACGAACAGAGTCATTCTGAACCCACTGCTCTACAACCAACAGAAGGGCATGCTCGGAAGCCAGGT
CAAAGTCAGCTTGCAGAAGAAGATTAAGAGCAACAGACCCATTCCTCAAATCAAGGACCCACTGC
ACATCGATGACCTGGTTTACCAGTACTATGAGGAAAGTATAGTGAATCCCATTAACAAGGCTCAG
CGCAGAAGCGACACCATCGAAGACAGCGAAGACTCTTCTGAAGAGGCTTACAGGCCTAGAAGGCA
GTTCAAGCGCGCAGCTTGGGATAATTTCGACGAATCTGACAGCAGCAGCAGCAGTACCAGCGAGG
AATATCCTGCAGTGCCCCAGCCTGACCTCGTGCAACCACCTTTCCTGCCATTCAGCGATGTGGAG
CTGCCTGAAAACAACCCAAACAAAATTCGCATTGTCAGCGAACTTGTTGAGAAGATTGCTCGTCA
GCTGCAAGAGCCTAACGGCATCCCATCAAGCGACACACTGGCCTTATTCCAAATCCTCTGCAAAG
AACTGATGGGATACAACGTAGGGACTCTGCGCCAACTGGCTAGCAAGTTCTATGCACAATACCCG
GAAAATCCGCACAGCTACCCGCAAGCTGTAAAATACACCCAGTGGGTCGTCTTCATTAACGCCCT
TTCACAGTCGGGTACAGGCCCAGCAATCAAAACTATTAGCGACCAGATCAAACAGCAGAAACTTC
AAGGACCCCTTGCCGCCAGCGCCATCGGAAATATTCCACGCAGTGCTCTATACCCCAACAAAGAA
TACCTCCAGGAATTCTTCATGATTGTGAAGAACACAGAGAAGAACAGTCCTGAGAACTCGACCGC
TCTCGCTCTTCTGGCGTATGGTAACCTCCTGCGTACCGCTGTTGTTGACAGAGATAGTGCCCACA
ACCTCTTCCCAGTTCACATATACGGCAAGCTGGACCCTCCGAGCCAGTCTGAACCACTCCAATCA
TACGTGAAGTACCTGACTAACCTCCTCAATCGCGCCGTTAAAAACGCCGACAGTGTTGGTATTCA
AGTATACACCAGAGCCCTGGGAAATATTGGACACCCATCCATCCTGAAAGCGCTACTGCCGTACG
TCTTTGCGGAGAAGCAGGTCTCTCACTTCCAGCGACTTCTGATGGTACTGGCACTCGATCGGGTT
ACTGAGCTCTACCCCAACGTATTGAGACCACTTCTGATCCAGATTTACCAGAGCACTGGTGAAAC
ACACCAGATCCGCAGCACGGCAGTTTTGCTTATTATGGGTTCCAACCCTTCAGGAAGTGTCCTGC
AGAGGCTTGCTCAGTTTTCTAAGCAAGATCCCAGCCCGCAAGTCGCCTCTGTTGTCAAGACAGCG
ATTCAAAGTGCTGCCCAACTGTCAAACCCGGAAAATCAGGAGCTAGCCCAAAGTGCTATGGCTGC
AGTCAATATGCTGAACCAAAATAAAACCGCGGTCCAGTACTCTCTCAAACATTTGCAAGATTACG
TCGTACGTGAGATGGCATTGTCGTACAATCTGAAGTACAGCCACCTCGGCAGCAGGGATTCTATT
ATACCTAACTCACTGTTTTCTGCTCTGAACGTGAATATTGGAGGATTGCAGCAGCTTACTGGAGA
AGCCCAAGCCGCACTTAGCAGCATTAAGACCCTACAACATCTTATTAACCACATCGTAGTTGAGG
GCGAAGAAGACGAACAGCAGCCACCGAAAGATCCCCTTGTCAGCATGGTAACTGACTTAAGGCAA
CAACTCGAGGGTCAAGTCGTAGGAAACCTTTTGTCCTATCCATCTATCGTTGCCTTTGACAACAA
CACAATTCACTCTTGGGTGGAAGATCTCAGAAAGAATGTCAGCAAGCTGGAGAGAGGACTGCCTT
TCAACTGCACCAATATTGTAAGTGCAGTTGATATCAAGCTGTCTTTCCCCAATGCTCTTGGATTC

CAAAGCTCCATCGCATACGAGACTCCAGTGCTCTACTCAGTTGGTGGTGAAGTGAACCTGAAGAC
TTCGCCAAAACTGAACAGTGCTCCTCAGGGCCATCTGGCAGCGCCAACCACATGGAATATCACTG
CAGATATCCAGGCTGTCTATTCAAGGCGAAGCTCAGGAACCGTATCTTTCACTGTACCCCCTCTT
GTAAAGACTTACTCCGCCGGGTATGTCAAGAACCAGCAAATTCAAGTACCACTAAGAGTTGTCCT
CGACATCGATGTCGAATACAATAACGCATATGTGAAACTGCAGCCCATCGATAAGAACCAGAGGT
ATCAGTGGAGGGCTGAATACGGAAAACGCTCTACTGGAATGGTGTATTCTGTGGAGTACAGTTCA
GAAAAGCATTACGATGACCTGTTCGACAAGTACGAACGTCTTGGTAAGCACGATGAACTGTCTGC
TATCTTGTTCCCCACCGAGGAGCAGCAGATCTATGCTAGCAACATCAGCATTTACTACGAGCCCG
AGAAGTCGACCTCCAAGGCTGTGGAACTTTGGACCCAATACTACGATGAACTCGTAACGCCCCCA
CCGCAACGGCAAGGTGCGCAGGACGACGAATACCAGCCCGAGTCCAGGGCAGCGTCTGACGAAGA
CGAGGCTTCTGATCAGGAGAGCGCTGAGGCAAGCGAGGAGAAATACTACCGATCCAAGCGTGGAT
CACTCTACGCGGGAAGAGGCAAGAGCAGGAGACAGAGCGACAGTCAGAACAGAGCAGTCCTGAGC
AGCGTCGGCCCCAATGAAGCGCACAAGGCACTGATTCAGAAATTACTGGCACTAGGATCCGGCAA
CGGTTACAATGTGAACGCGCTGTCGTTCGAAGCGAAGTTTGTAGAAGGCAAGGGTGCTTCATCGC
TGTTTTTGGCTGCCTGGAGAAAGAGTCCAGTGGCCAGCAAAGCCCAACTGCTCATCTTCGCGAAT
GCTAAGGCAGCACAGGCTAGCTCGCGCTATGAGGTCTGCGTGGACGCAGAATCCGATTTCCCCAA
CGTTCCCTACATGAACTTGGAGAAAGCCTTGAAGAACAATCCAAACTCGTCTGTGTCTATCACAC
TGGATGCTGGAGAAGGCTCCTGCTCGTCGGGAATGAACATTCAGGCAAATCTCGATATTGGAAGA
AGCCAAAGCATGTACCGGTACCTGCAGAACTCAAGGGTCGTGGCACAATGCAAATCTCAGATGAA
GAATGATAACTACATTCTGCCTGCATGTCGCAACGCCACAATTAAGGCGAGCGTGCTGAACAACT
ACGACCTAGAGGTCGAGTACCAGAACGTTCCTGAAGCTGTTAAGCAAGCTCTTTCCAAAGCTTAC
TCGGCCGTGGACTACGCTCTGTACATTTATGGATCACAAAACGTTGTGAATGGTACCGGTGCTCC
TGGAAAACTGTCTGCCAACCTTATCGTCGCACCAAACCTGCGCTCTCTGAATGCCTCCATCTCGA
CTGCATATTTCGACGCTATCTACGAGAACGTTCGCCTTGCGGTACCCGTAGCTAAAGCCATCGTC
ACCCACCCACTCTTGCCCAGAGCTGAACGCGTGGCTAACTACTACACGAACTGGAAGTACAACGC
AACCTGCAGTGTGGACTCTTCCTACATCAAGACTTACGACAATCTGACGTACCCGTTCGAGTCGG
AGAGCCCCTGCTGGAGAATTCTGCTGGTCACAGCTCGCCCGCGCGATTGGAAACCGTCCCCACTT
ATCCAGGCACCAAACGTCACTGTGATGGTGCGGGAAGCCTCTGGCAAGAGGGAAATCCGAGTGCA
GGTCGGTGAGGACGTCGTATCAGTATCGTACCAGTCGGGCAAGTACATGCTCAAGGCCAACAAAA
AATCCGTCCCCATCACGCAGTCAGGAGTGGCCGTCCTGTACGGCAAGCAAGGACAGCAGCTGGTG
CTGGCGTACGGCCTCGCAGACTCTGTCAGGGTCGAAATTGCAGACCTGATCATGCACTTTGACCA
GAAACGAGTCCTGTTGCAACCGGGCGACGACTACAGGCGAGCAGTGCGCGGTCTCTGCGGTACCT
TCGACGGCCTGCGTCAGACGGACTTCAAGCTCCCGGCCAACTGCATCGTAAGAAACGTCACCTCA
TTCATCGAAGCCTACACGTACGGCGACAAATGCGCAGCTCCCCGCAAGTCGCAGCCCAAGCAGTG
CTACCCGGAGCGGATCATCCCCGGAAACTCAACGGAACCGTGGCCCGTTGACCCAGTTGACCCCA
AATGCATCTCTCTGCGGTACAGGGTCCTGCTGCGCGGTCAAAGAGTCTGCATCAGCCACAAGCCT
CTGCCGATGTGCAAGCAGAACTGCAAGGCAGTACACACCGTCGAGAAAGTGGTTCCATTCCGCTG
CTTCACCCAGAGCGACACAATTGCAACCGTCGCCCTTACAGCCAAACGGGGAGGCCAGGTGAACG
TGAAGGATATACCAGTTGAAGGAGAAGACACCAGACTGGAAGTCACATTGGCAACTGGCTGCGCA
CGTGCGGACGACGACAGCGACGAGAGCAGCGAGAGTGACGAAAATTAA

Appendix B. Supplemental data for Chapter 3

Sequence B1: Partial cDNA sequence of Met nuclear receptor of *S. gregaria*.

AAGACCTCGTGCTGTTTCTGAAGGAGATGACACTACTGGATTCAATTTAC
AGAATTTTGTAAGTCTCTTCAAGGTGCCTGAAGAGGAAGAAAATAGTGTT
GAAGTTCAGCTGTCACCTCTTGCCAATAGTGCTGGCATTACCACAGACTT
AGAAGAAAAAGAACAGCTCAGAAATCCACTTTCTCCTTCATCACCTCCAT
CGAGGCACCCATACACCGCTGCCCAGAGGATTAGGGCAGCATCAGAAAGT
GAACTACAGACTACCAATAGTGGATCGACTTTTATTCAGTCAGATCCAGC
GAGCAGAGGGGGCAACTGGCAGTTTCCGAGGCCCGGAGACATGCCTCATT
CGGGAACTACGGATAGAAGGCCCCCGGAAAGACCTACGGTTATTACCGGA
CGGTGCTGTGCTGGATCTGAGAGGATGAAGTGCAGTAGATTGTCAGTTAT
GTCAGTCACAGGAGATGAGAAGAGGATTACCAGTGTGCGTGGTGAAGTTA
CTGTGTTAAAAAGGACTATGAGTGCAGATGTAGAAAATCCGAGCACTACA
AAGAAACTGAATGTTGGACCCCGGTACTGTAGNGCAGCGGCACAGAGNTG
AGCATCCTCCTACGGAAGCTCTGAGTGTGGAAGTAAATACACTTCCTCCC
GATGAACTGCAGCATATAATATTGTCGAGGGACAGTGGCATGTACAATTC
GCAGCCACAGTTCTCGTATCACAGCACAGCCTGTTTCGGACATCACAGCC
AGCACTTCAGACTGTGACACAAAATGTGCCTGCATCTTCTTTGAGCCCCC
CTCAATTAGCCGAAATGCAGAGCACGTCTGTTACACCACCATCGTTTATT
ATGGAAGAACCTAGTGGCCTGCCTCTTGTGAACTTAAGTGATGCCAACAC
ATTATCAGTTGGTTGCCAGGATCCTCTTATCAGCCCTGATCATGGAGACA
GTGACTCATTTTCATCTCTAGGATTACCTGATCCAGAGAATAATCCAGAG
GTCTTAGACCTGTTAGAAGCATC

Appendix C. Supplemental data for Chapter 5

Sequence C1: cDNA sequence of apoLp-III of *S. gregaria*

The start codon and termination signal are underlined

ACATGGGGAACGCACTCAGCGCTCCCCAACCACACGACCGCAGAAGTCGCCACCGCAAGCGC
CATGAAGACTCTGCTGGCTGTCCTGATGTTGGCCGTCGCCTGCCAGGCGCGACCAGATGCTG
GCCAGGTGAACATCGCAGAGACGGTGCAGCAACTGAACCACACGATCGTCAACGCCGCCCAT
GAGCTCCGCGAGACACTGGGCCTGCCAACGCAGGACGAGGCCCTCAACCTGCTCACCGAGCA
GGCTAATGCCTTCAAGACCAAGATCGCCGAGGTCACCTCCACGCTCAAGCAGGAGGCCGAGA
AGCACCAGGGCACCGTGTCGGAACAGCTGAACGCCTTTGCGCGAAACCTGAACAACAGCATC
CACGACGCCGCCACGTCCCTCAACCTGGAAGAACAGCTCAACTCTCTCCAGTCGGCACTCAAC
AACGTTGGACACCAGTGGCAGGACATCGCGAGCAAGACGCAGGCGGCGGCGCAGGACGCG
TGGGGTCCGGTGCAGTCGGCGCTGCAGGAGGCCGCCGAGAAGACGAAGCAGGCAGCGGCC
AACCTGCAGAACTCCATCCAGTCTGCCGTCCAGAAGCCTGCCAACTGAGCCACGCTGCCTTCT
GCCCCCTCTACCACCTTCCTCTCTGCCTGTCTTAGCCAATAAATTACATGTACAGTTTATCGTAA
AATACAACTTTTTCCTCTCTATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA