This is to certify that the

dissertation entitled

THE CLONING AND CHARACTERIZATION OF TRANSCRIPTION FACTORS HNF-4 AND CREB FROM THE MOSQUITO AEDES AEGYPTI: POSSIBLE ROLES IN THE REGULATION OF VITELLOGENESIS.

presented by

Neal T. Dittmer

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Genetics

Major professor
Dr. Alexander S. Raikhel

Date September 1, 2000

DATE DUE	DATE DUE	DATE DUE

THE CLONING AND CHARACTERIZATION OF TRANSCRIPTION FACTORS HNF-4 AND CREB FROM THE MOSQUITO *AEDES AEGYPTI*: POSSIBLE ROLES IN THE REGULATION OF VITELLOGENESIS.

By

Neal T. Dittmer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Program in Genetics

2000

ABSTRACT

THE CLONING AND CHARACTERIZATION OF TRANSCRIPTION FACTORS HNF-4 AND CREB FROM THE MOSQUITO *AEDES AEGYPTI*: POSSIBLE ROLES IN THE REGULATION OF VITELLOGENESIS.

By

Neal T. Dittmer

Hundreds of millions of people are infected each year by diseases transmitted by mosquitoes. It is while blood feeding that a female mosquito may pick up or transmit a disease-causing organism. The mosquito utilizes the proteins in the blood to synthesize the yolk protein precursors she puts into her eggs, this process is known as vitellogenesis. Vitellogenesis requires the coordinated activity of the midgut (for blood meal digestion), fat body (for yolk protein synthesis), and ovaries (for yolk protein accumulation). Because the insect fat body is the functional equivalent of the liver, it is hypothesized that transcription factors involved in liver-specific gene expression will have a similar role in regulating gene expression in the fat body. Hepatocyte nuclear factor 4 (HNF-4) is a member of the nuclear hormone receptor superfamily, and is involved in the regulation of several genes expressed in the liver. Three isoforms of an *Aedes aegypti* homolog (AaHNF-4a, 4b, and 4c) were cloned and characterized as a beginning to elucidating the role of AaHNF-4 in regulating fat body gene expression. These isoforms contain structural domains characteristic of nuclear hormone receptors, and show high sequence homology to their mammalian counterparts. Cell transfection experiments have shown that both

AaHNF-4a and 4b are transcriptional activators, while AaHNF-4c inhibits this transactivation. Northern blot analysis reveals that these two isoforms show different temporal expression during vitellogenesis, suggesting that each may regulate a different set of genes. DNA-binding sites for HNF-4 have been identified in the upstream regulatory region of two fat body-expressed yolk protein genes, indicating a possible role for AaHNF-4 in controlling the expression of these genes. Finally, electrophoretic mobility shift assays (EMSA) with nuclear extracts revealed highest DNA-binding activity during the pre- and postvitellogenic period, suggesting that AaHNF-4's primary role is in regulating the developmental processes that occur prior to, and following, the synthesis of the yolk protein precursors. The cyclic AMP response element binding (CREB) protein is a member of the basic region/leucine zipper (bZIP) family of transcription factors, and is a mediator of gene regulation induced by extracellular signals. A mosquito isoform (AaCREB) was cloned and has high homology to its mammalian homolog in the DNA-binding and protein dimerization domains. Northern analysis showed AaCREB to be constitutively expressed in the fat body during vitellogenesis. In cell culture experiments, AaCREB was a potent repressor of cAMP-mediated transcriptional activation. EMSA with nuclear extracts detected binding at the late protein synthesis stage and postvitellogenic periods. AaCREB binding sites were also detected in the regulatory regions of yolk protein genes. Taken together, these results suggest that AaCREB may be involved in the termination of yolk protein gene expression.

DEDICATION

I wish to dedicate this dissertation to my parents, Robert and Marcia, who have always encouraged me in all my pursuits, and whose love and support (both emotional and financial!) knows no bounds.

ACKNOWLEDGMENTS

First I wish to thank the friends I made in the Raikhel lab and the Genetics Program for their friendship and advice during all these many years in graduate school. Second, I would like to thank my guidance committee, Drs. Don Jump, Suzanne Thiem, and Rich Schwartz, for their help and suggestions during my studies. I found their critique of my work always to be helpful without squashing the ego of a "budding" scientist. Third I would like to thank Drs. Barb Sears and Jim Hancock for giving me the opportunity to teach genetics. I thoroughly enjoyed being a TA for their "Fundamentals in Genetics" class, and the semester they asked me to be a coinstructor in their absence was a wonderful opportunity and confirmed my desire to pursue a career in academia; I hope I did them proud. Foremost, I wish to thank my major professor, Dr. Alexander Raikhel, for taking a chance on me. Alex always expected the best from his students but understood that mistakes were part of the learning process – believe me I "learned" a lot. Alex also took an interest in my development as a scientist, encouraging (and subsidizing!) my participation in scientific meetings so that I could present my work as well as interact with the top researchers in the field. His mentoring helped me to build confidence in myself to think critically but objectively, and to recognize and follow new avenues of research when the opportunity presents itself.

TABLE OF CONTENTS

LIST OF FIGURES viii

LIST OF ABBREVIATIONS ix

CHAPTER 1 - LITERATURE REVIEW
Medical importance of mosquitoes 2
Vitellogenesis in the mosquito *Aedes aegypti* 4
The previtellogenic period 4
Hormonal regulation of previtellogenic development 7
The vitellogenic period 9
Hormonal regulation during vitellogenesis 13
Termination of vitellogenesis 16
Tissue-specific gene regulation 18
Nuclear hormone receptors 21
Nuclear receptor HNF-4 26
Insect HNF-4 homologs 32
The cyclic AMP signal transduction pathway 35
Transcription factors responsive to the cAMP signaling pathway 39
CRE-binding proteins in insects 46
Significance of current research 49

CHAPTER 2 - MATERIALS AND METHODS
Polymerase chain reaction (PCR) and cDNA library screening 53
Northern blot analysis 55
Preparation of fat body nuclear extracts 56
Electrophoretic Mobility Shift Assays (EMSAs) 57
Transient transfections in cell culture 59
Drosophila S2 cells 59
Human 293T cells 60
CV-1 cells 61
Antibody production 62

CHAPTER 3 - CLONING AND CHARACTERIZATION OF MOSQUITO HNF-4 ISOFORMS
INTRODUCTION 65
RESULTS 66
Cloning of cDNAs for mosquito HNF-4 isoforms 66
Analysis of AaHNF-4 transcripts 73
The distribution of AaHNF-4 expression in tissues of the adult mosquito 74
Stage-specific expression of AaHNF-4 transcripts in the fat body 74
DNA binding by AaHNF-4 proteins 79

CHAPTER 3 - Continued
Identification of HNF-4 binding sites in regulatory regions of fat body-specific YP genes ... 81
AaHNF-4a and –4b are transcriptional activators ... 86
HNF-4 binding activity in mosquito fat body nuclear extracts ... 92
DISCUSSION ... 95

CHAPTER 4 - CLONING AND CHARACTERIZATION OF MOSQUITO CREB
INTRODUCTION ... 105
RESULTS ... 106
Cloning of mosquito CREB cDNA ... 106
Northern blot analysis of AaCREB expression ... 110
Electrophoretic Mobility Shift Assays (EMSA) ... 110
Transient transfection in CV-1 cells ... 113
DISCUSSION ... 116

CHAPTER 5 - SUMMARY AND CONCLUSIONS
Expression of AaHNF-4 transcripts during the first vitellogenic cycle ... 125
The expression of AaHNF-4 proteins in the fat body remains to be resolved ... 126
Affinity purification of α-GSTHNF-4(Com) and α-GSTHNF-4a(NH_2) antibodies ... 127
Possible roles for AaHNF-4 in the mosquito fat body ... 128
Possible regulation of AaHNF-4 expression by 20E ... 130
Possible regulation of YP gene expression by cAMP ... 131
CREB as a potential regulator of YP gene expression ... 133
Generating AaCREB antibodies ... 134

REFERENCES ... 136

LIST OF FIGURES

CHAPTER 1
1.1 Summary of the first vitellogenic cycle..... 5
1.2 Structures of juvenile hormone and ecdysteroids..... 8
1.3 Functional domains of nuclear hormone receptors..... 22
1.4 Schematic drawings of HNF-4 homologs..... 28
1.5 Activation of CREB by different signaling pathways..... 38
1.6 Schematic representation of CREB, CREM, and ATF-1 proteins..... 41

CHAPTER 3
3.1 Restriction maps of AaHNF-4 cDNA clones..... 68
3.2 Alignment of insect and rat HNF-4 sequences..... 69
3.3 Domain comparison of insect and rat HNF-4 homologs..... 72
3.4 Northern hybridization analysis of AaHNF-4 transcripts..... 75
3.5 Expression of AaHNF-4 in the ovary during the first vitellogenic cycle..... 76
3.6 Expression of AaHNF-4 in the fat body during the first vitellogenic cycle..... 78
3.7 Electrophoretic mobility shift assay (EMSA) of HNF-4 isoforms..... 80
3.8 AaHNF-4c inhibits DNA binding of other HNF-4 isoforms..... 83
3.9 AaHNF-4b binds *in vitro* to sequences upstream of the *Aavg* and *Aavcp* genes..... 84
3.10 Transient transfection of AaHNF-4 isoforms in *Drosophila* S2 cells... 88
3.11 Transient transfection of HNF-4 homologs in 293T cells..... 91
3.12 Recognition of AaHNF04 isoforms by the α-GSTHNF-4(Com) antibodies..... 93
3.13 Specificity of the α-GSTHNF-4(Com) antibodies..... 94
3.14 Active binding by AaHNF-4 proteins from fat body nuclear extracts.. 97

CHAPTER 4
4.1 Nucleotide and deduced amino acid sequence of AaCREB..... 107
4.2 Comparison of AaCREB with *Drosophila* and human CREB proteins. 109
4.3 mRNA expression profile of AaCREB in the fat body..... 112
4.4 AaCREB can bind to CREs *in vitro*..... 114
4.5 CRE-binding proteins are present in fat body nuclear extracts..... 115
4.6 Transient transfections of AaCREB in CV-1 cells..... 118

CHAPTER 5
5.1 Comparison of mosquito CREs to the PAR consensus..... 134

ABBREVIATIONS

20E: 20-hydroxyecdysone

Aa: *Aedes aegypti*

Adh: *Alcohol dehydrogenase* gene

APF-1: Apolipoprotein Factor-1 (HNF-4 binding site in *apoCIII*)

apoCIII: *apolipoprotein CIII* gene

apoLp: apolipoprotein

ATF-1: Activating Transcription Factor-1

bZIP: basic region/leucine zipper

C/EBP: CCAAT/Enhancer-Binding Protein

cAMP: cyclic Adenosine 3',5'-Monophosphate

CAT: Chloramphenical Acetyltransferase

CBP: CREB-Binding Protein

cDNA: complementary DNA

CRE: cAMP Response Element

CREB: CRE-Binding protein

CREM: CRE-Modulator protein

DNA: Deoxyribonucleic acid

DR1: PuGGTCA direct repeat (with a 1 nucleotide spacer)

EcR: Ecdysteroid Receptor

EMSA: Electrophoretic Mobility Shift Assay

FA-CoA: Fatty Acyl-Coenzyme A

GST: Glutathione-S Transferase

HNF-1,3,4: Hepatocy Nuclear Factor-1,3,4

HRE: Hormone Response Element

JH: Juvenile Hormone

KID: Kinase Inducible Domain

KIX: KID Interaction Domain

Lp: Lipophorin

LpR: Lipophorin Receptor

Luc: Luciferase

mRNA: messenger RNA

NR: Nuclear hormone Receptor

PAR: Proline/Acid amino acid region

PBM: Post Blood Meal

PCR: Polymerase Chain Reaction

PKA: Protein Kinase A

RNA: Ribonucleic acid

SDS-PAGE: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SomCRE: Somatostatin gene CRE

SREBP: Sterol Response Element Binding Protein

USP: Ultraspiracle

VBP: Vitellogenin Binding Protein

VCB: Vitellogenic Cathepsin-B

VCP: Vitellogenic Carboxypeptidase

VCPCRE: VCP gene cAMP Response Element

VCPHNF4: VCP gene HNF-4 Response Element

Vg: Vitellogenin

VgCRE: Vg gene cAMP Response Element

VgHNF4: Vg gene HNF-4 Response Element

VgR: Vitellogenin Receptor

WHO: World Health Organization

YP: Yolk Protein

CHAPTER 1

LITERATURE REVIEW

LITERATURE REVIEW

Medical importance of mosquitoes.

Millions of people worldwide are infected each year with diseases transmitted by mosquitoes. Malaria, dengue fever, yellow fever, and filariasis are just a few of the diseases that plague the tropical and sub-tropical regions around the globe. Malaria is second only to tuberculosis in the number of deaths caused each year. It is estimated that 2.4 billion people are at risk and that 300 to 500 million will be infected, with over 1 million individuals dying (World Health Organization (WHO), 1998a). Malaria is caused by protozoans of the genus *Plasmodium* and is transmitted by mosquitoes of the genus *Anopheles*. Resistance of the vector to insecticides and the parasite to various anti-malarial drugs has made control of this disease increasingly more difficult.

Dengue fever is now prevalent across all of the tropical and sub-tropical regions of the world. Forty percent of the world's population live in areas where it is endemic, and 50 million cases are expected every year. Dengue fever is caused by four closely related viruses transmitted by the mosquito *Aedes aegypti*. Recovery from infection by one of the viruses confers immunity against that serotype for life, but only partial and transient protection against the other three. There is strong evidence that sequential infection can lead to the more serious dengue haemorrhagic fever. With modern therapy, death rates can be limited to less than 1%. However, without proper treatment fatality can exceed 20% (WHO, 1998b).

Lymphatic filariasis (also known as elephantiasis) is a chronic and debilitating disease that affects more than 120 million people world wide and puts over 1 billion people at risk. It is caused by the parasitic nematodes *Wuchereria bancrofti* and *Brugia malayi*. These worms lodge in the lymph nodes and can cause swelling of an entire leg, arm, genitals, or breast. Even more common than the overt abnormalities caused by these parasites is the internal damage to the kidneys and lymphatic system. They live for 4-6 years and produce millions of immature microfilarie that circulate in the blood (WHO, 1999a).

Yellow fever is viral disease that is transmitted by *Aedes* and *Heamogogus* (South America only) mosquitoes. It is endemic in tropical areas of Africa and Central and South America. Monkeys as well as humans are reservoirs for the virus; thus, it can be maintained even in small isolated populations. In addition, the mosquito can also transmit the virus to its offspring via its eggs (vertical transmission). Deforestation, urbanization, and growing populations are bringing people into contact with vectors and virus in ever increasing frequency. Despite the availability of a safe and highly effective vaccine, there are at least 200,000 cases per year and 30,000 deaths (WHO, 1999b).

Eradication strategies have proven unsuccessful as mosquitoes quickly repopulate regions once control efforts have stopped. Moreover, the increase in resistance to insecticides have made even limited control strategies less effective (Busvine, 1993). Continued efforts to control mosquito populations will likely require novel strategies (Aultman *et al.*, 2000). Both disease transmission and

reproduction are linked to the blood meal. It is while blood feeding that a mosquito make pick up or transmit a disease-causing pathogen. Therefore, a better understanding of the mosquito reproductive cycle may aid in developing new strategies.

Vitellogenesis in the mosquito *Aedes aegypti*.

The female mosquito metabolizes the proteins in the blood and uses the amino acids to synthesize the proteins that are deposited in the egg. The major yolk protein (YP) precursor synthesized is vitellogenin (Vg) and this process is referred to as vitellogenesis. Vitellogenesis requires the coordinated activity of several tissues, namely the midgut for digestion of the blood meal, the fat body for synthesis and secretion of the YP precursors, and the ovaries for accumulation of these proteins (Raikhel, 1992). Vitellogenesis can be divided into two parts, the previtellogenic and the vitellogenic periods (figure 1.1).

The previtellogenic period.

The previtellogenic period is a preparatory stage in which the relevant tissues become competent for vitellogenesis. In the first gonadotropic cycle, this begins when the female ecloses from the puparium. Even upon metamorphosis of the adult, the differentiation of the midgut, fat body, and ovarian tissues is not complete. In the early stages after emergence, midgut epithelial cells are loosely grouped, dissimilar in appearance, and contain few microvilli.

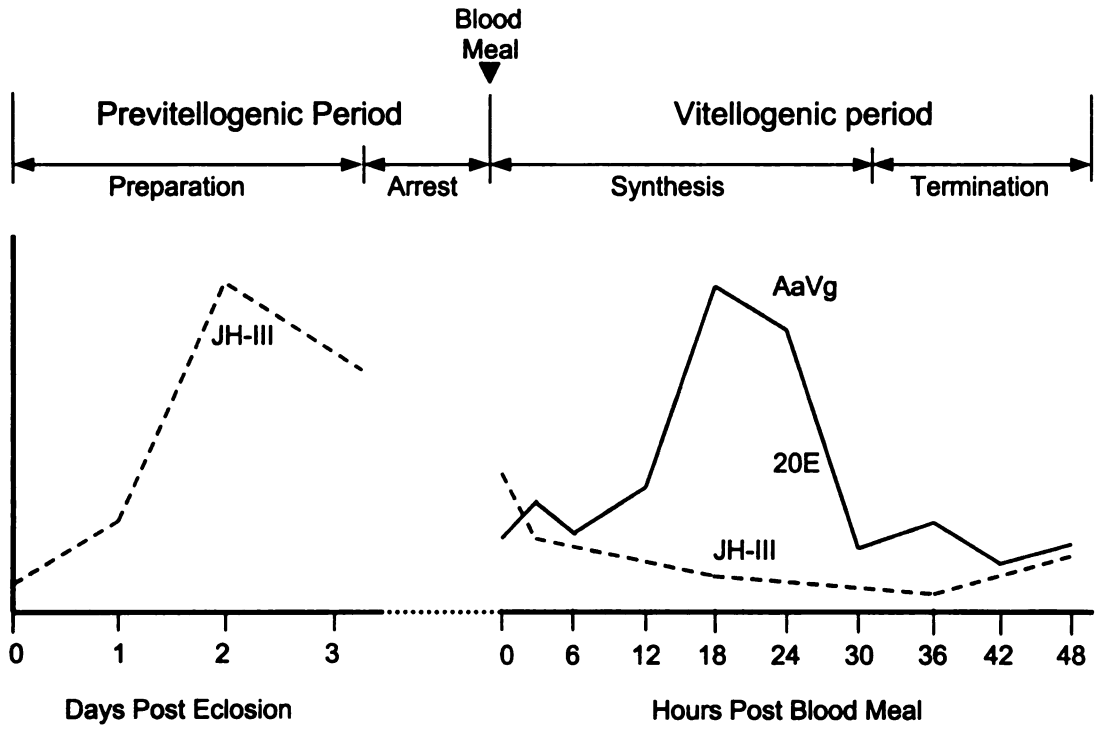


Figure 1.1, Summary of the first vitellogenic cycle. The first cycle proceeds through two developmental periods. The previtellogenic period begins at eclosion of the adult female. Tissue development at this stage is under the control of the terpinoid juvenile hormone III (JH-III, dashed line). The fat body then enters a stage of quiescence (arrest). The vitellogenic period begins when a blood meal is taken, and is divided into a synthesis and termination stage. The synthesis stage is characterized by the massive production and secretion of yolk protein precursors (represented here by AaVg, shaded in gray) by the fat body. Yolk protein gene expression is regulated by the steroid hormone 20-hydroxy-ecdysone (20E, solid line). The termination stage is defined by the ending of yolk protein synthesis, and the selective degradation of biosynthetic organelles by lysosomes. (Redrawn from Raikhel, 1992. JH-III titer is from Shapiro *et al.*, 1986. 20E titer is from Hagedorn *et al.*, 1975.)

By three days post-eclosion, the cells have taken on their regular columnar shape, microvilli are more numerous, and septate desmosomes (necessary for holding the cells together and sealing the gaps between the cells) have formed. In addition, several organelles (mitochondria, rough endoplasmic reticulum (RER), and lysosomes) are in their early stages of formation and continue to develop (Hecker *et al.*, 1971). Thus it seems unlikely that the midgut of newly emerged females would be able to withstand the pressure caused by the ingestion of the blood meal, digest the blood, or absorb the nutrients.

In the fat body, new DNA synthesis results in the ploidy level of trophocytes increasing from diploid to tetraploid or even octoploid (Dittman *et al.*, 1989). An increase in nucleoli activity and the subsequent increase in RNA and ribosomal levels leads to a build up of RER, and the development of Golgi complexes (Raikhel and Lea, 1983); machinery necessary for the massive production and secretion of the YPs. As in the development of the midgut, this preparatory stage occurs over a two to three day period, after which the continued development of fat body ceases until a blood meal is taken (a stage referred to as the state-of-arrest).

The previtellogenic period is also necessary for the ovaries to become competent for uptake of the YPs (Raikhel and Lea, 1985). Endocytosis of extraovarian-produced proteins is a receptor-mediated process (Roth and Porter, 1964, Raikhel and Dhadialla, 1992; Sappington and Raikhel, 1995). Clatherin-coated vesicles and coated pits are hallmarks of receptor-mediated endocytic machinery. However, at eclosion, the oocytes have an undifferentiated cortex

with no coated vesicles or coated pits on the membrane, and only a few undeveloped microvilli. By 48 hours post eclosion, the oocytes have numerous coated pits and coated vesicles, and well developed microvilli (Raikhel and Lea, 1985). The mosquito vitellogenin receptor (AaVgR) has been purified and its cDNA cloned (Sappington *et al.*, 1995, 1996). AaVgR is a 205 kDa protein expressed exclusively in the ovaries. Both mRNA and protein can be detected within the first day post-eclosion. Deduced amino acid sequence reveals AaVgR to be a member of the low density lipoprotein receptor (LDLR) superfamily, and shares significant homology with the *Drosophila* yolk protein and the chicken vitellogenin receptors (Sappington *et al*, 1996; Sappington and Raikhel, 1998).

Hormonal regulation of previtellogenic development.

The previtellogenic development of these tissues is under the control of Juvenile hormone (JH) (reviewed in Hagedorn, 1985; Raikhel, 1992; Dhadialla and Raikhel, 1994). JH is a lipophylic, sesquiterpenoid (figure 1.2) produced by the corpora allata, an endocrine organ. JH-III is the sole JH in adult mosquitoes (Dhadialla and Raikhel, 1994, and references within). Removal of the corpora allata, or abdominal ligation of newly emerged adults, blocks previtellogenic development of these tissues. This effect can be reversed by the reimplantation of the corpora allata, or by topical application of JH-III or an analog (Rossignol *et al.* 1982; Dittman *et al.*, 1989; Raikhel and Lea, 1985, 1990, 1991). However, the mode of action of this hormone is unknown.

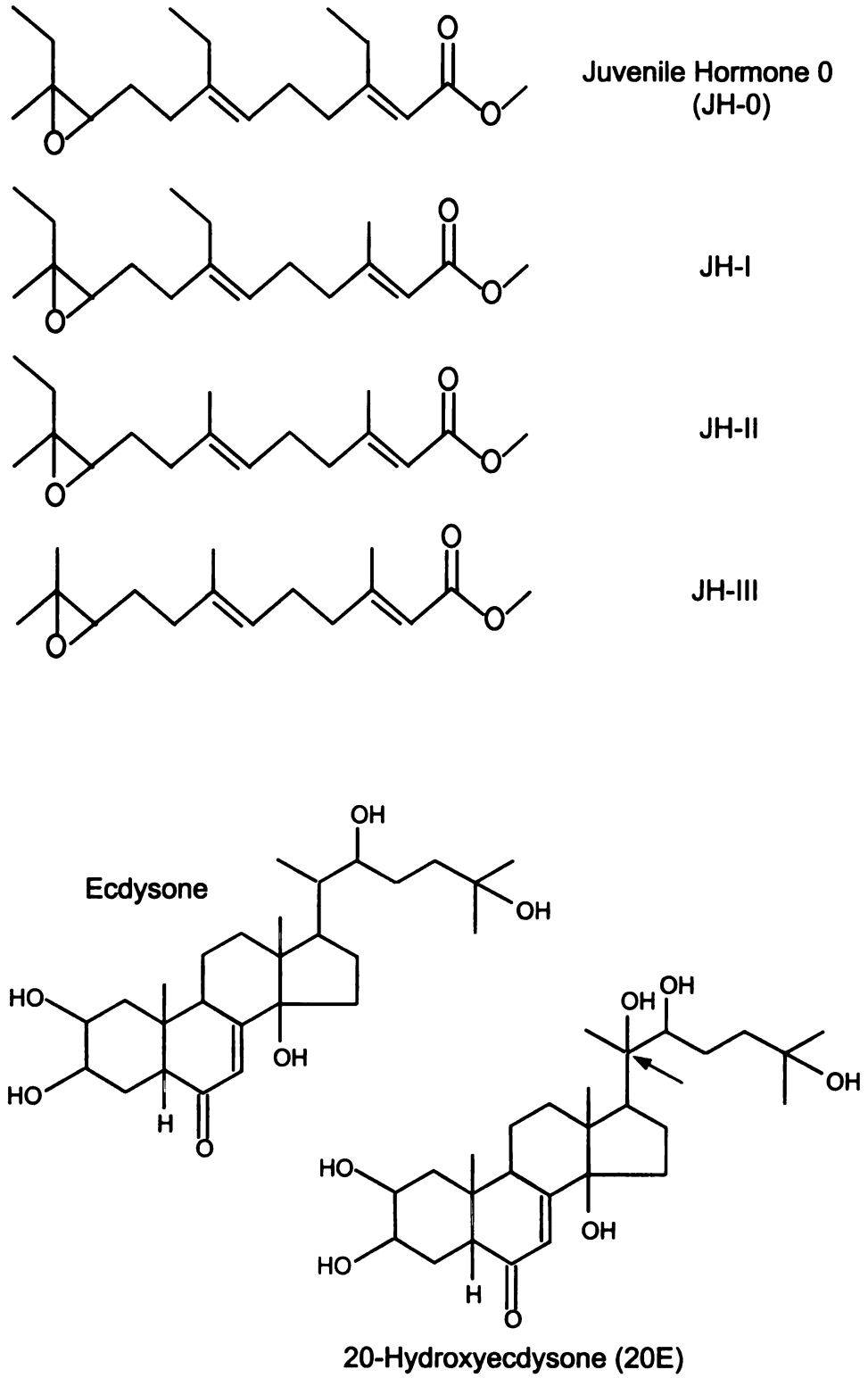


Figure 1.2, Structures of juvenile hormones and ecdysteroids.

Whether JH-III binds to a surface receptor and induces a signal cascade, or binds to a nuclear receptor has yet to be elucidated (Jones, 1995; Riddiford, 1996; Hartfelder, 2000). Recently, Jones and Sharp (1997) demonstrated that the product of the *ultraspiracle* (USP) gene can specifically bind JH-II and JH-III and induce transcription of a reporter gene by a LexA-USP fusion in the yeast *Saccharomyces cerevisiae*, suggesting that USP can function as the JH receptor in insects.

The vitellogenic period.

The vitellogenic period begins with the ingestion of a blood meal. Trypsins are the major enzymes involved in the digestion of the blood meal (Gooding, 1972; Briegel and Lea, 1975). Two distinct types of trypsin are present in the midgut: early trypsins and late trypsins (Graf and Briegel, 1989; Felix *et al.*, 1991). Early trypsin activity begins soon after a blood meal is taken and is essential for inducing the synthesis of late trypsin (Graf and Briegel, 1989; Barillas-Mury *et al.*, 1995; Noriega *et al.*, 1996a). Surprisingly, the early trypsin mRNA is synthesized in the previtellogenic midgut and translation does not occur until after the blood meal is taken (Kalhok *et al.*, 1993; Noriega *et al.*, 1996a, b). Recently, the transcriptional regulation of the early trypsin gene has been shown to be under the control of JH-III (Noriega *et al.*, 1997). Late trypsin is responsible for most of the enzymatic activity found during the peak period of blood meal digestion, 20-30 hours post blood meal (PBM) (Graf and Briegel, 1989). In addition to its regulation by the activity of early trypsin, the level of late trypsin

activity is also determined by the amount and quality of the protein in the blood meal (Briegel and Lea, 1975; Graf and Briegel, 1989; Felix *et al.*, 1991; Noriega *et al.*, 1994). It is postulated that by the use of this two-phase system, the mosquito can assess the quality of the blood meal by early trypsin before committing to the production of late trypsin (Barillas-Mury *et al.*, 1995).

Vitellogenesis in the fat body can be divided into two phases: a synthesis phase, characterized by the massive production and secretion of the YPs, and a termination phase, characterized by the cessation of YP production and the remodeling of the fat body tissue by lysosomes (figure 1.1) (Raikhel, 1986a, b; 1992). Synthesis of AaVg in the fat body begins soon after a blood meal is taken, being detected as early as 1 hour PBM, reaching its peak at 24-28 hours PBM, and then rapidly declining to back ground levels by 36-40 hours PBM (Hagedorn *et al.*, 1973; Raikhel and Lea, 1983). In *Aedes aegypti*, Vg is the product of a multigene family with 4 active genes and one pseudo gene (Gemmill *et al.*, 1986; Hamblin *et al.*, 1987). The AaVg mRNA is translated into a 224 kDa preproprotein (Dhadialla and Raikhel, 1990). Glycosylation and phosphorylation result in a 250 kDa pro-protein. This is then cleaved into two subunits by vitellogenin convertase (AaVC), an endoprotease, and further phosphorylated to produce large and small subunits of 190 and 62 kDa respectively (Raikhel and Bose, 1988; Dhadialla and Raikhel, 1990; Chen and Raikhel, 1996). Sulfation results in a final size of 200 and 66 kDa, respectively. AaVg is secreted into the hemolymph as an oligomeric protein (380 kDa) comprised of these two subunits (non-covalently linked), where it is then taken up by the developing oocytes.

Thus, vitellogenin not only supplies the developing embryo with amino acids, but it also provides carbohydrates, phosphates, and sulfates as well.

In addition to AaVg, several other YP precursors that are synthesized in the fat body have been characterized. Vitellogenic carboxypeptidase (AaVCP), and vitellogenic cathepsin B (AaVCB) are synthesized as inactive proenzymes (Hays and Raikhel, 1990; Cho *et al.*, 1999). Like AaVg, they are expressed in a sex-, stage-, and tissue-specific manner, secreted into the hemolymph and taken up by the maturing oocytes. *In vitro* analysis demonstrated that embryonic protein extracts, containing the activated form of AaVCB, degraded vitellin (the crystalline form of Vg in the egg). The addition of a thiol protease inhibitor, or anti-AaVCB antibodies, prevented this degradation (Cho *et al.*, 1999). Though the substrate of AaVCP is yet to be determined, it is postulated to play a role, either directly or indirectly, in the degradation of other YPs in the developing embryos (Cho *et al.*, 1991a).

In addition to AaVg, AaVCP, and AaVCB, lipophorin (AaLp; a lipid transport protein) is also produced and secreted by the fat body as a YP precursor (Van Heusden *et al.*, 1998; Sun *et al.*, 2000). AaLp is composed of two subunits, apolipoprotein-I (apoLp-I) and apolipoprotein-II (apoLp-II), with molecular weights of 240 and 70 kDa respectively (Capurro *et al.*, 1994; Ford and Van Heusden, 1994; Van Heusden *et al.* 1998). Similar to AaVg and *Locusta migratoria* lipophorin, apoLp-I and apoLp-II are believed to be the products of a single transcript that is post-translationally cleaved to produce the two subunits (Weers *et al.* 1993; Van Heusden *et al.* 1998). AaLp is unique from

the other mosquito YPs described in that is expressed in the previtellogenic fat body. However, the *Aalp* gene shows enhanced expression in response to a blood meal, reaching peak levels at 18 hours PBM before declining to basal levels by 30 hours PBM (Van Heusden *et al.*, 1998; Sun *et al.*, 2000).

As mentioned previously, the ovaries internalize the YP precursors via receptor-mediated endocytosis. While AaVgR protein begins to accumulate during the previtellogenic period, rapid synthesis begins in response to a blood meal. By 24 hours PBM, AaVgR levels are 14-fold greater than in the previtellogenic ovary (Sappington *et al.*, 1995). Endosomes carrying the internalized AaVg coalesce into transitional yolk bodies, and AaVgR is recycled back to the cell surface (Raikhel, 1984; Snigirevskaya *et al.*, 1997b). As the yolk body matures AaVg condenses into its crystalline state, vitellin (AaVn) (Raikhel, 1984).

Unlike AaVg, AaVCP and AaVCB are only localized around the rim of the mature yolk body (Snigirevskaya *et al.*, 1997a; Cho *et al.*, 1999), consistent with the notion that they are involved in the digestion of the yolk proteins. The internalization of the other YPs appears to be similar to that of Vg, though the exact pathway is not currently known (Hays and Raikhel, 1990; Snigirevskaya *et al.* 1997a; Cho *et al.*, 1999). Lately, an ovarian-specific cDNA encoding a putative lipophorin receptor has been cloned and should prove very helpful in elucidating the sequestering of AaLp by the ovaries (S-J Seo, H-M Jun, J Sun, A Raikhel, unpublished results).

Hormonal regulation during vitellogenesis.

The steroid hormone 20-hydroxyecdysone (20E) (figure 1.2) is the primary regulator of vitellogenesis (Hagedorn, 1989; Dhadialla and Raikhel, 1994). Ecdysone is synthesized in the ovary and converted to 20E in the fat body (Hagedorn, 1989; Hagedorn *et al*, 1975). Synthesis of ecdysone by the ovaries is stimulated by factor(s) from the brain (Hagedorn *et al.*, 1979). This factor is presumably egg development neurosecretory hormone (EDNH), which is critical for the development of the eggs (Lea, 1967, 1972). EDNH is produced by the medial neurosecretory cells of the brain and stored in the corpora cardiaca (Lea, 1967, 1972), an organ of the neuroendocrine system. Release of EDNH by the corpora cardiaca is triggered by an "EDNH releasing factor" produced by the ovaries in response to the blood meal (Lea and Van Handel, 1982). The levels of ecdysteroids are low early after a blood meal, with only a small peak at 4 hours PBM. By 10 hours PBM, the hormone titer begins to rise rapidly and reach their peak levels between 18-20 hours PBM, before rapidly dropping to previtellogenic levels by 30 hours PBM (reviewed in Dhadialla and Raikhel, 1994).

The action of 20E is mediated by the binding of this ligand to a nuclear receptor; a heterodimeric complex between two proteins that are the products of the *ecdysteroid receptor* (EcR) and *ultraspiracle* (USP) genes (Oro *et al.*, 1990; Koelle *et al.*, 1991; Yao *et al.*, 1992, 1993). The binding of this complex to specific DNA sequences (hormone response elements, HRE) in the regulatory regions of target genes results in the recruitment of coactivator complexes that help to stabilize the transcriptional machinery for repeated rounds of transcription

(reviewed below in "Nuclear hormone receptors", page 21). In the mosquito, two isoforms of the ecdysteroid receptor (AaEcR, Cho *et al.*, 1995; Wang and Raikhel, unpublished results) and ultraspiracle (AaUSP, Kapitskaya *et al.*, 1996) have been cloned. These isoforms show differential expression in the fat body during both the pre- and vitellogenic periods, suggesting that they each have distinct and separate functions (Wang *et al.*, 2000a). Initially it was believed that ecdysone produced by the ovaries was a precursor and that 20E was the active form. However, Wang *et al.* (2000b) have recently shown that ecdysone can be a potent ligand for the mosquito ecdysteroid receptor complex.

The exact role of 20E in regulating YP gene expression has yet to be discerned. High doses of 20E are needed *in vivo* to stimulate vitellogenesis in non blood-fed females (Fuchs and Kang, 1981; Lea, 1982). Yet, when fat bodies are cultured *in vitro*, physiological doses of 20E are enough to stimulate AaVg synthesis (Fuchs and Kang, 1981; Hagedorn, 1985; Koller *et al.*, 1989; Raikhel *et al.*, 1997). Deitsch *et al.* (1995a) have shown that stimulation of YP gene expression by 20E is dependent on new protein synthesis, suggesting that the mode of action of 20E is indirect and that 20E acts on the YP genes through a cascade. The cloning of the cDNAs for AaEcR and its heterodimeric partner AaUSP, as well as other members of the ecdysteroid regulatory cascade (Pierceall, *et al.*, 1999; Raikhel *et al.*, 1999; Kapitskaya *et al.*, 2000; Zhu *et al.*, 2000) may help elucidate the role of 20E in regulating YP gene expression.

Despite the importance of 20E in regulating vitellogenesis, YP synthesis in the fat body begins before the titer of 20E rises (Hagedorn *et al.*, 1975).

Van Handel and Lea (1984) have shown that the digestion of the blood meal is critical for the initiation of vitellogenesis. Mosquitoes that have had their ovaries removed (and thus unable to synthesize ecdysone), or were decapitated and given blood by enema, were still able to synthesize AaVg. AaVg synthesis could be blocked by the addition of trypsin inhibitors in the blood meal, suggesting that a digestion product of the blood meal or some factor secreted by the midgut signals the fat body to begin vitellogenesis. However, both head and ovarian factors were necessary to maintain the synthesis needed for oocyte maturation (Van Handel and Lea, 1984).

In the ovaries, internalization of the fat body-synthesized YPs shows two distinct phases. In the first 2 hours after a blood meal, accumulation of YPs occurs slowly. This is followed by a more rapid uptake of proteins that peaks between 24 and 30 hours PBM (Koller and Raikhel, 1991). Both of these phases require a head factor(s), as mosquito abdomens that are first isolated and then given blood by enema do not internalize AaVg. For the first phase, release of the head factor is rapid; even mosquitoes decapitated within five minutes of feeding were able to take up AaVg. The second phase, however, required the presence of a head factor(s) for 16-20 hours PBM to achieve levels of endocytosis similar to that of intact mosquitoes (Koller and Raikhel, 1991). Shapiro (1983) demonstrated that ovaries incubated with head extracts have a burst in levels of the second messenger cyclic-AMP (cAMP). The role of the cAMP signal transduction pathway was further substantiated by Sappington *et al.* (1998) who

were able to induce both phases of endocytosis in the ovaries with forskolin, an adenlyate cyclase activator.

Physiological doses of 20E were shown to stimulate the second phase of uptake in the ovary, but not the first, though whether this is a result of a direct effect on the ovaries or indirectly through stimulation of AaVg production remains to be determined (Koller and Raikhel, 1991). However, 20E is necessary for the deposition of the vitellin envelope by the follicular epithelium cells (Raikhel and Lea, 1991). In addition, 20E also regulates the separation of the secondary follicle form the germarium (Beckemeyer and Lea, 1980).

Termination of vitellogenesis.

The termination phase of vitellogenesis begins at about 30 hours PBM (Raikhel, 1986b), coinciding with the rapid decline of YP mRNA and protein levels (Hagedorn *et al.*, 1973; Hays and Raikhel, 1990; Cho *et al.*, 1991a). Several hypotheses have been proposed as to how YP gene expression is terminated. Bohm *et al.* (1978) proposed that the declining 20E titer is responsible for the cessation of YP gene expression. However, as AaVg mRNA levels continue to increase for 6 hours after 20E levels begin to drop, it is likely that other factors are also necessary. It has also been suggested that rising titers of AaVg in the hemolymph induces a negative feed back mechanism that turns off the YP genes (Borovsky 1981). In addition, factors secreted by the ovaries may also play a role, either directly or indirectly, in terminating vitellogenesis. An oostatic hormone (OSH) produced and secreted by the

ovaries during the termination phase (30–48 hours PBM) inhibits uptake of YPs in less developed eggs (Meola and Lea, 1972; Kelly *et al.*, 1986; Lea and Brown, 1990). In addition, other factors secreted by the ovaries inhibit the biosynthesis of digestive enzymes in the midgut (Borovsky, 1988; Borovsky *et al.*, 1993). The genes for both AaVg and AaVCP have been cloned and analysis of their regulation is currently being investigated (Deitsch and Raikhel, 1993; Romans *et al*, 1995).

Also at this time, the activity of several lysosomal enzymes were shown to rapidly increase and remain elevated over the next 12 hours (Raikhel, 1986a). The action of the lysosomes was directed solely at the degradation of RER, Golgi complexes and secretory vesicles necessary for the synthesis and secretion of the YPs (Raikhel, 1986b). To further examine the role of lysosomes in the termination of vitellogenesis, a lysosomal aspartic protease (LAP) was purified and its cDNA cloned (Cho *et al.*, 1991b; Cho and Raikhel, 1992). Surprisingly, the mRNA level for LAP was shown to peak at 24 hours PBM, 12 hours before peak protein and enzymatic activity (Cho and Raikhel, 1992). The rise in protein levels coincided with the drop in the 20E titer. When female mosquitoes were injected with 20E between 18 and 36 hours PBM to keep the hormone titer artificially high, LAP protein levels remained low (Deitsch *et al.*, 1995b). Thus it seems 20E plays an inhibitory role on the translation of LAP mRNA, but the exact mechanism remains unknown.

Tissue-specific gene expression

Because Vg is synthesized only in the fat body, understanding the regulation of gene expression in this tissue is critical to elucidating the mechanisms controlling vitellogenesis. The insect fat body is the functional analog of the vertebrate liver, involved in the synthesis and storage of hemolymph proteins, carbohydrates, lipids, and detoxification enzymes (Friedman, 1985; Chapman 1998). The validity of this biological analogy is further illustrated by the parallels in regulatory mechanisms responsible for fat body- and liver-specific gene expression. For example, the *Drosophila alcohol dehyrogenase* (*Adh*) gene contains an enhancer that regulates expression in the adult fat body (*Adh* adult enhancer: AAE). This enhancer contains binding sites for a repressor protein (adult enhancer factor 1: AEF-1) and an activator protein (Falb and Maniatis, 1992a). While the activator protein has not been identified, its binding site is very similar to the response element of the CCAAT/enhancer-binding protein (C/EBP), a liver-enriched transcription factor. In cell transfection experiments, mammalian C/EBP is able to stimulate transcription from the AAE, and this stimulation is repressed by coexpression with AEF-1 (Falb and Maniatis, 1992b). Surprisingly, similar binding sites are located in the mammalian *Adh* gene, which too can be activated by C/EBP (Potter *et al.*, 1991; Falb and Maniatis, 1992b). Significantly, the Drosophila AEF-1 repressor is also able to inhibit this activation, similar to its action on the Drosophila *Adh* gene. Moreover, electrophoretic mobility shift assays have demonstrated that rat liver nuclear extracts have AEF-1-like binding activity, suggesting that a homolg is present in

mammals (Potter *et al.*, 1994). Thus, a tissue-specific regulatory unit consisting of one positive and one negative element has been conserved between *Drosophila* and mammals.

Liver-specific gene expression in vertebrates relies principally on four families of transcription factors that are liver-enriched but not restricted to this tissue (Lai and Darnell, 1991; Xanthopoulos and Mirkovitch, 1993; Cereghini 1996). These four families are hepatocyte nuclear factors 1, 3, and 4 (HNF-1, 3, 4), and C/EBP. Each of these transcription factors binds to DNA as a dimer. As multiple members and splice variants are present in each family, several heterodimeric combinations are possible, allowing for the diversity necessary to regulate genes of such wide-ranging functions in a tissue-specific manner. Homologs to each of these families, with the exception of HNF-1, have been cloned from insects.

The product of the *slow border cells* (*slbo*) gene of *Drosophila* is a member of the C/EBP family (Montell *et al.*, 1992). The DmC/EBP protein is expressed predominantly during late embryogenesis in the salivary glands, proventriculus, gut, and epidermis, at the time of terminal differentiation of these tissues (Rørth and Montell, 1992). Surprisingly, DmC/EBP is not detected in larval or adult fat body tissue. However, the importance of the C/EBP binding site in the *Adh* gene, as well as a similar element in the fat body enhancer of the *Drosophila yolk protein* genes, for fat body-specific expression indicates that a related protein is likely present (An and Wensink, 1995a, b).

The *Drosophila* gene *forkhead* (*fkh*) is a member of the HNF-3 family (Weigel *et al.*, 1989; Lai *et al.*, 1993). The *forkhead* gene has been studied for its role in development, where it is expressed early in embryogenesis at the terminal ends of the embryo and regulates the development of the gut, salivary glands, and nervous system. Additional genes cloned from Drosophila that are members of the HNF-3 family are *sloppy paired 1* and *2* (*slp1*, *2*: Grossniklaus *et al.*, 1992) and *crocodile* (*croc*: Hacker *et al.*, 1995), both of which are also involved in the development of the embryo. Silk gland factor 1 (SGF-1) is a *fkh* homolog cloned from the silk moth *Bombyx mori* (Mach *et al.*, 1995). SGF-1 has been shown to bind to a sequence in the promoter of a silk protein gene and is expressed in larval silk glands, homologous structures to *Drosophila* salivary glands. The identification of a SGF-1 clone from a *B. mori* embryonic library suggests that it may be required for development of the silk gland, similar to the role of *fkh* in the development of the *Drosophila* salivary glands.

While none of the insect HNF-3 family members cloned to date have been reported to be expressed in the fat body (indeed it remains to be seen whether or not this tissue has been examined for expression of these genes), two partial cDNAs have been cloned by the polymerase chain reaction (PCR) from the vitellogenic fat body of the mosquito *Aedes aegypti* (D. Martin and A. Raikhel, unpublished results). The recent finding that HNF-3 is an important regulator of the *Xenopus vitellogenin B1* gene (Cardinaux *et al.*, 1994; Robyr *et al.*, 2000) further strengthens the hypothesis of the conservation of fat body- and liver-specific gene regulatory mechanisms between insects and vertebrates.

The last family of liver-enriched transcription factors to be considered is HNF-4. As the cloning and characterization of a mosquito HNF-4 homolog is a major portion of this dissertation, it will be discussed in greater detail in the following sections.

Nuclear hormone receptors.

HNF-4 is a member of the nuclear hormone receptor (NR) superfamily. Members of this family are diverse and include receptors for steroid hormones, thyroid hormones, vitamin D, retinoic acids, and a myriad of receptors for which ligands have not yet been identified or may not exist (orphan receptors) (for recent reviews see Tsai and O'Malley, 1994; Kastner *et al.*, 1995; Mangelsdorf *et al.*, 1995; Ribeiro *et al.*, 1995; Laudet *et al.*, 1997; Meier 1997; Whitfield *et al.* 1999). Despite the wide-ranging array of proteins in this family, all NRs have a common modular structure (figure 1.3). At the amino terminus is a domain, commonly referred to as the A/B domain, that is highly variable and has a transactivation function (AF-1) that is ligand-independent. This is followed by a region, the C domain, that contains two C4 zinc fingers that are responsible for DNA-binding and protein dimerization. Adjacent to this is a flexible hinge region (D domain) that often contains a nuclear localization signal. Next is a large hydrophobic region, the E domain, which is involved in ligand binding, protein dimerization, heat-shock protein association (in the case of the vertebrate steroid receptors), nuclear localization, and ligand-dependent transactivation (AF-2). At the very carboxy terminus of the protein is another highly variable region, the F domain. The function of this domain is uncertain, but as many receptor isoforms

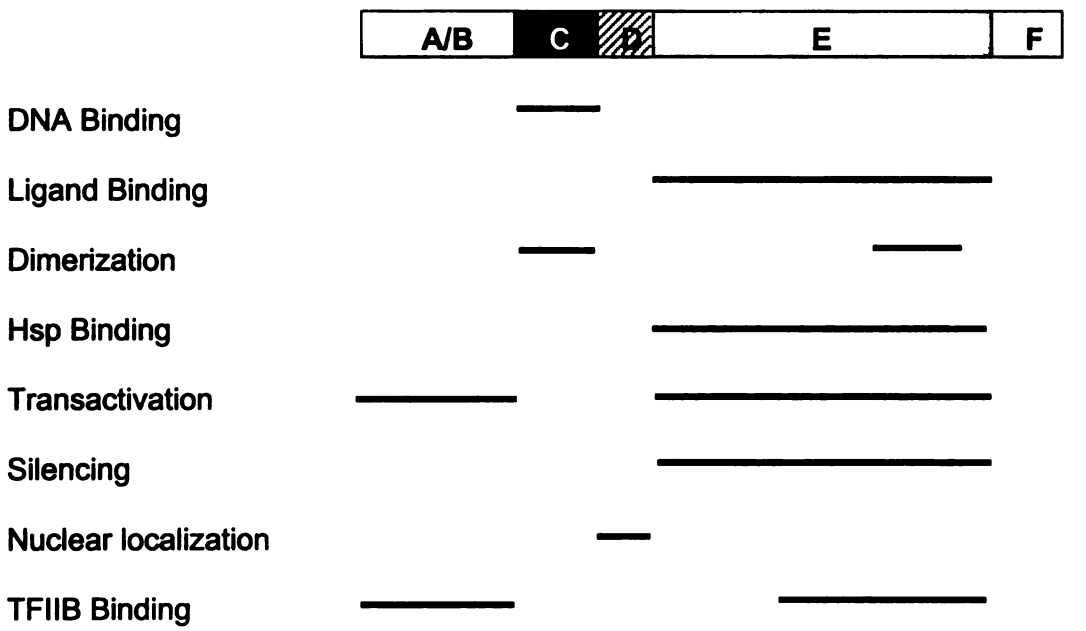


Figure 1.3, Functional domains of nuclear hormone receptors. The structure of nuclear hormone receptors can be divided into five domains: A/B, C, D, E, and F. The function of each domain is indicated by the solid lines (redrawn from Tsai and O'Malley, 1994).

differ in this region it likely plays a role in modulating transactivation (Montano *et al.*, 1995; Tate *et al.*, 1996; Suaud *et al.*, 1999).

Several methods have been proposed for classifying NRs based on specific sequences within the zinc fingers of the DNA-binding domain (Umesono and Evans, 1989; Forman and Samuals, 1990), the position of an intron-exon junction in the DNA-binding domain (Detera-Wadleigh and Fanning, 1994), and phylogenetic analysis based on the alignment of the C, D, and E domains of all published NR sequences (Laudet, 1997). However the NRs can conveniently be divided into 4 general classes based on their DNA-binding and dimerization properties (Mangelsdorf *et al.*, 1995). The first class consists of the vertebrate steroid receptors (i.e. the estrogen (ER), progesterone (PR), glucocorticoid (GR), mineralocorticoid (MR), and androgen (AR) receptors). In the absence of their ligand, these receptors exist as monomers in solution complexed with heat shock proteins (Tsai and O'Malley, 1994; Beato *et al.*, 1995). Upon binding of their ligand, the heat shock proteins dissociate and the proteins dimerize. The receptors bind to DNA as homodimers and recognize inverted repeats of the sequence AGAACA separated by 3 nucleotides (IR3). The ER is an exception to this as it recognizes an IR3 of AGGTCA. Binding of these homodimers to DNA *in vivo* is ligand dependent.

The second class of NRs consists of nonsteroid members that bind as heterodimers with the retinoic X receptor (RXR). This class includes receptors for vitamin D (VDR), retinoic acid (RAR), thyroid hormone (TR), and the peroxisome proliferator-activated receptor (PPAR) (Mangelsdorf and Evans,

1995; Kastner *et al.*, 1995). In contrast to the vertebrate steroid receptors, these proteins are not complexed with heat shock proteins and can bind DNA in the absence of their specific ligand. They recognize elements composed of AGGTCA half-sites that are most often direct repeats (DR), but may also be everted repeats or inverted repeats.

The third and forth classes are composed of the orphan receptors. This class includes RXR, HNF-4, the chicken ovalbumen upstream promoter-transcription factor (COUP-TF), *apolipoprotein AI* regulatory protein (ARP-1), nerve growth factor induced protein B (NGFI-B), and steroidogenic factor 1 (SF-1). These receptors bind as either homodimers to direct repeats of AGGTCA with variable spacer length (class 3), or as monomers (class 4). Numerous homologs for several of these NRs have been cloned from insects as well (reviewed in Segraves, 1994; Henrich and Brown, 1995; Thummel, 1995; Raikhel *et al.*, 1999).

While NRs are often placed in one of these four classes, they often share characteristics with members from other groups. For example, as already mentioned, the ER recognizes half-site sequences of AGGTCA, the same as for the orphan and nonsteroid receptors, but not the AGAACA sequence of the other steroid receptors. *Drosophila* EcR and USP (the insect RXR homolog) interact with chaperones, a novel finding for class 2 receptors (Arbeitman and Hogness, 2000). RXR and COUP-TF can be classified as both class 2 and class 3 receptors as they form both homo- and heterodimers. NGFI-B shares characteristics of both class 2 and class 4 receptors as it can bind either as a

monomer or as a heterodimer with RXR (Mangelsdorf and Evans, 1995). In addition, several NRs can recognize several different response elements, albeit with different affinities (Cooney *et al.*, 1992; Kato *et al.*, 1995; Wang *et al.*, 1998).

Once bound to the DNA, NRs insert their influence on transcription by direct interaction with the basal transcription machinery and recruitment of coactivator proteins (reviewed in Horwitz *et al*, 1996; Glass *et al.*, 1997; Collingwood *et al.*, 1999). The recruitment of coactivators is both ligand and AF-2 domain dependent. Once assembled, coactivators disrupt nucleosome formation by acetylation of histone lysine residues, allowing access of the basal transcription machinery to the promoters of genes. In addition, direct contact with the basal machinery helps to stabilize the preinitiation complex, allowing for more efficient transcription. Some NRs, such as the thyroid hormone, retinoic acid, retinoic X, and vitamin D receptors can bind to DNA in the absence of a ligand. In these situations, the NRs interact with corepressor proteins which recruit histone deacetylase enzymes that stabilize the formation of nucleosomes by deacetylation of histones, and thus exclude the basal transcription machinery from the promoters of target genes. On binding of a ligand, the NR undergoes a conformational change which releases the corepressor and allows for the binding of coactivators. Orphan receptors, which can bind DNA and activate transcription in the absence of a ligand would seem to be constitutively on. However, posttranslational modifications such as phosphorylation could be required for these receptors to undergo the conformational change necessary for interaction with coactivators (Weigel, 1994; Laudet, 1997). In addition,

phosphorylation may also affect DNA-binding as was shown for the orphan receptor HNF-4 (Voillet *et al.*, 1997). Furthermore, competition for binding sites, heterodimeric partners, and coactivators, can all modulate activation by NRs.

Nuclear receptor HNF-4.

HNF-4 was first isolated from rat liver extracts as a protein that bound to similar response elements of two liver-specific genes, *transthyretin* and *apolipoprotein C III* (Costa *et al.*, 1989, 1990). Purification of the protein allowed for partial amino acid sequencing and the design of degenerate primers for use in PCR with a rat cDNA library as template. The amplified PCR product was then used to screen the same library from which a full length clone was isolated (Sladek *et al.*, 1990). Since then, 6 additional splice variants from the *hnf-4α* gene have been identified (figure 1.4A). In addition, a second gene has been isolated from both mammals (*hnf-4γ*; Drewes *et al.* 1996; Taraviras *et al.*, 2000) and *Xenopus* (*hnf-4β*; Holewa *et al.*, 1997) (figure 1.4B). These two genes (γ and β) are considered to be distinct and not homologous based on the fact that they each have unique substitutions in their DNA-binding domain; a domain which is 100% identical in all vertebrate *hnf-4α*s cloned thus far (Holewa *et al.*, 1997).

HNF-4α expression is most abundant in the liver, kidney, and intestine. It is expressed at lower levels in the pancreas, and at very low levels in the skin, stomach, and testis. HNF-4γ transcripts are also detected in the kidney, pancreas, intestine, and testis, but are 10 fold less abundant than HNF-4α,

Figure 1.4, Shematic drawing of HNF-4 homologs. The identity of each domain is listed at the top of the figure. A) Comparison of the splice isoforms of HNF-4α. HNF-4α1, α2, and α3 differ only in their respective F domains: α2 has a 10 amino acid insertion not included in α1, and α3 has a completely different F domain from that of α1 and α2. HNF-4α4 , α5, and α6 all have the same A/B domain, which is the same as for α1, α2, and α3 except for a 30 amino acid insertion. The F domains of α4, α5, and α6 are the same as α1, α2, and α3 respectively. HNF-4α7 is unique for the first 15 amino acids of the A/B domain and then is identical to HNF-4α1 for the remainder of the protein. Reference for the splice variants are: α1, Sladek *et al.* (1990), Chartier *et al.* (1994), Holewa *et al.* (1996);. α2, Hata *et al.* (1992), Chartier *et al.* (1994), Hata *et al.* (1995), Drewes *et al.* (1996); α3, Kritis *et al.* (1996); α4, Drewes *et al.* (1996), Furuta *et al.* (1997); α5 and α6, Furuta *et al.* (1997); α7, Nakhei *et al.* (1998).
B) Homology between HNF-4 homologs encoded by different genes. The amino acid secquenc of each domain for vertebrate HNF-4β, γ, and insect homolgs was compared separately with the corresponding domain of human HNF-4α1. Because of the large size of the A/B domain of HNF-4γ, only the last 50 amino acids of this domain were used for analysis. The percent identity is indicated below each domain. Numbers above each figure represent amino acid positions. References: human (Hs) HNF-4α1, Chartier *et al.* (1994); HsHNF-4γ, Drewes *et al.* (1996); *Xenopus laevis* (Xl) HNF-4β, Holewa *et al.* (1997); *Drosophila melanogaster* (Dm) HNF-4, Zhong *et al.* (1993); *Bombyx mori* (Bm) HNF-4a, Swevers and Iatrou (1998).

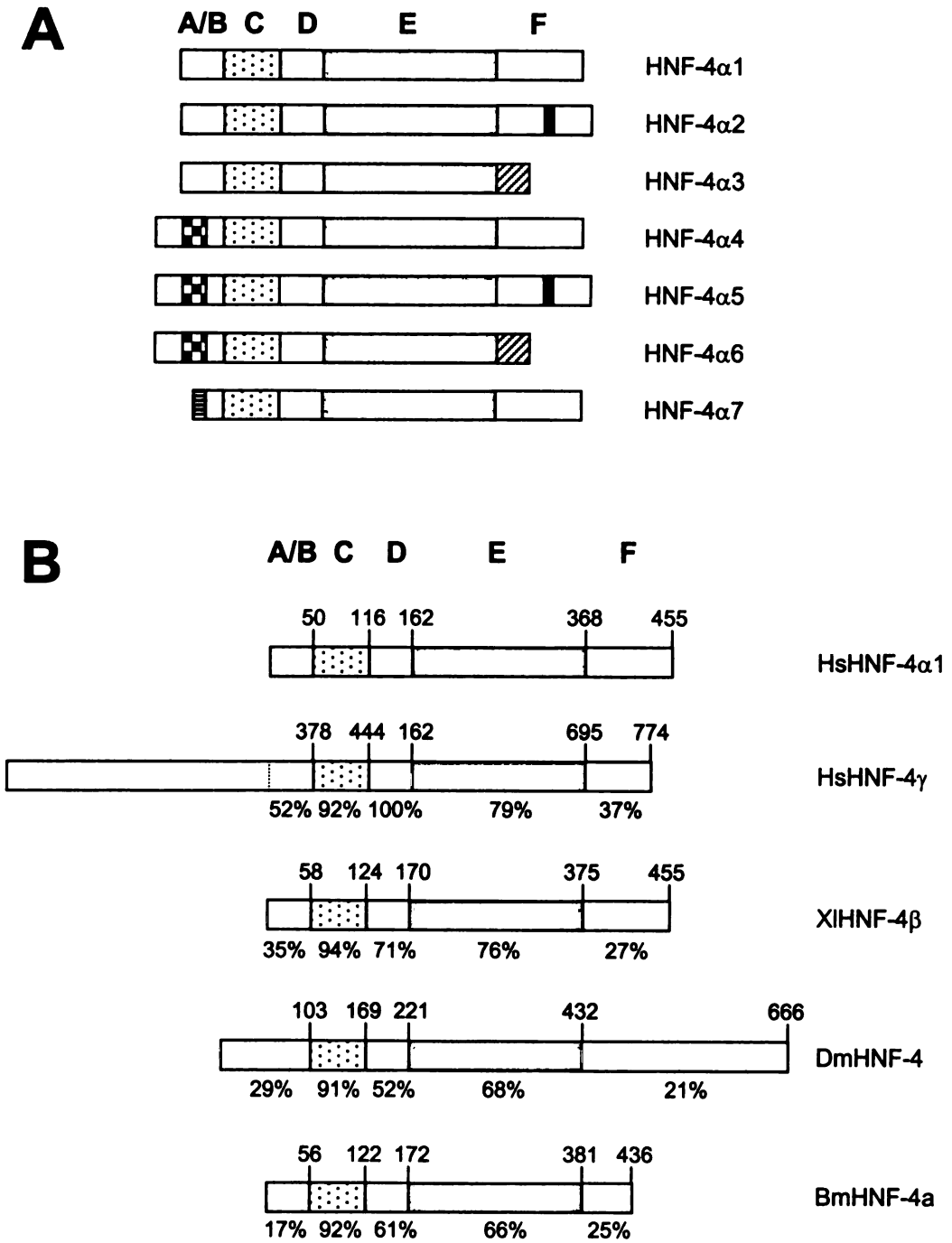


Figure 1.4, Schematic drawing of HNF-4 homologs.

but are not detected in the liver (Sladek, 1994; Drewes, 1996). Similarly, in *Xenopus*, HNF-4α protein is detected in the liver and kidney, while HNF-4β is also found in these tissues, and in the intestine and testis as well. HNF-4β is unique among HNF-4 isoforms in that it was also detected in the stomach, lungs, and ovaries. Both *Xenopus* proteins are maternally expressed and present at constant levels throughout embryogenesis (Holewa *et al.*, 1997). In accordance with the tissues in which it is expressed, HNF-4 regulates genes of great diversity and function. These include genes for serum proteins, blood coagulation factors, metabolism of carbohydrates, lipids, steroids and amino acids, and liver differentiation (reviewed in Sladek, 1994).

HNF-4 has also been shown to be critical for embryonic development. HNF-4 expression is first detected in the extraembryonic visceral endoderm (VE), a tissue that performs many of the same functions as the fetal liver and intestine (Duncan *et al.*, 1994; Taraviras *et al.*, 1994). Mice that are homozygous null for HNF-4 fail to complete gastrulation, suggesting that proteins secreted by the VE are critical at this stage of development (Chen *et al.*, 1994; Duncan *et al.*, 1997). Expression of HNF-4 in embryonic tissues parallels closely that in the adult. Transcripts are first detected around embryonic day 8.5 in the liver and gut primordia, and later in the pancreas primordium. Eventually, expression is detected in the meso- and metanephric tubules that give rise to the kidneys (Duncan *et al.*, 1994; Taraviras *et al.*, 1994).

Analysis of the HNF-4α amino acid sequence shows that it has all of the structural domains typical of the NR superfamily. The AF-1 (ligand-independent)

transactivation domain has been mapped to the first 24 amino acids at the N-terminus, as this region was able to autonomously activate transcription when fused to the GAL4 DNA-binding domain (Hadzopoulou-Cladaras *et al.*, 1997). In contrast, the AF-2 domain required nearly all of the D and E domains, encompassing residues 128 to 366. Deletion of amino acids 128 to 174 of the D domain, or 361 to 366 of the E domain, completely abolished the transactivation potential of AF-2. Since the region spanning amino acids 175 to 360 is sufficient for protein dimerization, this suggests that the flanking residues are involved in interaction with coactivators or other basal transcription factors (Hadzopoulou-Cladaras *et al.*, 1997). Indeed, residues 360 to 366 (LLQEMLL) match a conserved motif ($\phi\phi$XE$\phi\phi$, where ϕ is a hydrophobic residue, X is any residue) shown to be necessary for AF-2 function in many NRs (Danielian *et al.*, 1992; Durand *et al.*, 1994). However, unlike ligand responsive NRs in which the AF-2 function is ligand-dependent, the AF-2 domain of HNF-4α can function in the absence of any known ligand. Surprisingly, a novel finding was the identification of a potent negative regulator region in the F domain (Hadzopoulou-Cladaras *et al.*, 1997; Iyemere *et al.*, 1998). This repressor domain was shown to reduce AF-2 activation by nearly 40-fold, while its affect on the AF-1 domain was only 2-fold. The mechanism of this repression is unknown, however, as the HNFα-4 AF-2 domain can transactivate in an autonomous fashion, this repressor domain may represent a new level of control by masking the AF-2.

HNF-4 isoforms bind to DNA as a homodimer and recognize direct repeats of AGGTCA with a one nucleotide spacer. As no known ligand has been

identified, HNF-4 is classified as an orphan receptor. The ability of HNF-4α, β, and γ to activate transcription in the absence of a known ligand suggests that either one does not exist, or that a metabolite common to cells serves this purpose. In support of this, long chain fatty acyl-CoA thioesters (FA-CoA) have been reported to modulate the transcriptional activity of HNF-4α (Hertz *et al.*, 1998). FA-CoAs with chain lengths of 14 carbon atoms or longer were shown to specifically bind HNF-4α, regardless of the degree of chain saturation. However, chain length and saturation were important in how HNF-4α transactivation was affected. Myristoyl- and palmitoyl-CoAs (14 carbon and 16 carbon saturated chains) both increased the DNA-binding affinity of HNF-4α, but only palmitoyl-CoA was able to increase transactivation by HNF-4α in transient transfections. In contrast, FA-CoAs of 18 carbons or longer (both saturated and unsaturated chains) inhibited DNA-binding and thus suppressed HNF-4α transactivation.

The DNA-binding ability of HNF-4 was also demonstrated to be regulated both positively and negatively by phosphorylation. Phosphorylation of tyrosines and serines was shown to be correlated with increased DNA-binding affinity of HNF-4 (Ktistaki *et al.*, 1995; Jiang *et al.*, 1997). However the specific residues responsible for the increase in DNA-binding were not identified in these studies so the exact nature of this mechanism is unknown. Contrary to this, Viollet *et al.* (1997) were able to demonstrate that phosphorylation of serine 134 (located in the amino half of the D domain, just downstream of the 2nd zinc finger) by protein kinase A abolished DNA binding. Mutation of serine 134 to either a glycine or even a glutamic acid conferred insensitivity to the inhibitory effect of PKA,

suggesting that the phosphate group acted more by steric hindrance and/or conformational modification than by electrostatic repulsion. In addition, competition for binding to response elements by different NRs has also been shown to regulate HNF-4 mediated transcription. For example, HNF-4 is a positive regulator of the *apolipoprotein* (*apo*) *AII*, *B*, and *CIII* genes (Ladias *et al.*, 1992; Mietus-Snyder *et al.*, 1992). Transient transfections in HepG2 cells (a hepatocyte cell line containing endogenous HNF-4) of NRs ARP-1 or COUP-TF, suppresses activation of reporter constructs containing the HNF-4 elements from these genes. This repression can be overcome by coexpressing increasing amounts of HNF-4, indicating that transcriptional control of *apoAII*, *B*, and *CIII* is regulated in part by competition for binding to specific response elements.

Insect HNF-4 homologs.

HNF-4 homologs have been cloned from 3 insects: *Drosophila* (DmHNF-4: Zhong *et al.*, 1993), the silk moth *B. mori* (BmHNF-4: Swevers and Iatrou, 1998), and the mosquito *Ae. aegypti* (Kapitskaya *et al.*, 1998). Both DmHNF-4 and BmHNF-4 show exceedingly high homology with the human HNF-4α1 in both the C and E domains (figure 1.4B). Indeed HNF-4α1 shares higher similarity in the E domain with the insect homologs than it does with any other vertebrate nuclear receptor, with the exception of the HNF-4β and γ receptors. This level of conservation is remarkable considering the evolutionary distance been insects and vertebrates. This point is further illustrated by the ability of HNF4α1, which acts as a homodimer, to form heterodimers with DmHNF-4 (Zhong *et al.*, 1993).

Though only 1 isoform of DmHNF-4 has been cloned, two mRNAs of 3.3 and 4.6 kb are present in embryos, suggesting that splice variant may exist. These transcripts are deposited in the eggs by the nurse cells where they are uniformly distributed. However soon after fertilization the transcripts become restricted to the terminal regions of the blastoderm, and by 3 hours post fertilization no mRNA is detected. The mRNA reappears between 6 and 8 hours post fertilization in the posterior and anterior midgut primordium, followed by expression in the fat body, Malphigian tubules, and salivary glands at later stages. It is striking that DmHNF-4 is expressed in the fat body, midgut, and Malphigian tubules, the insect equivalents of the vertebrate liver, intestine, and kidney. In embryos that are homozygous for a deletion that spans the HNF-4 locus, early midgut, Malphigian tubules and salivary glands begin to develop but then become arrested. This strongly suggests that HNF-4 plays an essential role in early gut formation and organogenesis that has been preserved during evolution from invertebrates to vertebrates.

In contrast to *Drosophila*, two HNF-4 isoforms have been cloned from *B. mori*: BmHNF-4a and BmHNF-4b (Swevers and Iatrou, 1998). These two isoforms differ only at their amino terminus. The first 32 amino acids of BmHNF-4a and the first 6 amino acids of BmHNF-4b are unique to each, with the remainder of the protein identical between the two isoforms. A third isoform containing a different F domain has been reported but not characterized. Similar to observations made in *Xenopus* and *Drosophila*, BmHNF-4 transcripts are maternally produced and deposited in the egg. Both transcripts are strongly

expressed in the follicular cells of the ovary when development of this tissue begins during the transition from the last larval instar to the pupal stage. After induction, BmHNF-4 expression is maintained in these cells throughout development of the ovaries, implying a role for BmHNF-4 in follicular cell-specific gene expression. In agreement with this, nuclear extracts prepared from ovarian follicular cells formed two complexes in an electrophoretic mobility shift assay with a labeled probe containing the HNF-4 binding site of the *apoCIII* gene (Swevers and Iatrou, 1998). However, conclusive proof that these complexes contain BmHNF-4 proteins awaits the generation of specific antibodies.

The effect of 20E on BmHNF-4 expression was examined since the development of the ovaries is controlled by this hormone. Larvae entering the pupal stage were developmentally arrested by placing a ligature between the thorax and abdomen. The isolated abdomens were injected with 20E and the ovaries removed at various times post injection. Northern blot analysis of the RNA prepared from the dissected ovaries showed that BmHNF-4a mRNA levels began to increase at 6 hours post-injection and reached maximal levels after 24 hours post-injection, while BmHNF-4b mRNA levels were unaffected by the 20E treatment (Swevers and Iatrou, 1998). Thus it is likely that two different promoters control the expression of the BmHNF-4 gene. BmHNF-4 expression was also examined in several larval and pupal tissues by Northern analysis. BmHNF-4a was expressed predominantly in the fat body, midgut, Malphigian tubules, and ovaries, while BmHNF-4b was most abundant in the testis, ovaries,

fat bodies and midgut. In addition, BmHNF-4b was also detected at low levels in muscle, Malphigian tubules, brain, and wing epidermis.

The cyclic AMP signal transduction pathway

Vitellogenesis requires coordination between the midgut, fat body and ovarian tissues. As already noted in the sections concerning hormonal regulation and termination of vitellogenesis, several peptide hormones have been implicated or identified that help to regulate this process. Since peptide hormones are hydrophilic in nature, they are unable to pass through the lipid bilayer of the cell surface. Thus, these hormones act by binding to a surface receptor and trigger a cascade of events that can have a pleiotrophic affect on the cell. The cyclic AMP (cAMP) cascade is one of the major signal transduction pathways utilized by cells (for recent reviews see Meyer and Habener, 1993; Montminy, 1997; Daniel *et al.*, 1998; Sassone-Corsi, 1998; De Cesare *et al.*, 1999; Shaywitz and Greenberg, 1999; De Cesare and Sassone-Corsi, 2000). As discussed previously (pages 15-16), activators of this pathways were shown to induce uptake of AaVg by the ovaries. In addition, cAMP was also shown to mediate the secretion of salts and water from the Malphigian tubules (an analog of the vertebrate kidney) of *Ae. aegypti* after a blood meal (Petzel *et al.*, 1987). Thus, links of this pathway to events in vitellogenesis have already been established.

Binding of a hormone to its cognate receptor causes a conformational change that results in the activation of a GTP-binding protein (G-protein). G-

proteins consist of 3 subunits: α, β, and γ. In its inactive state, this complex is associated with the membrane receptor. Upon stimulation by the receptor/ligand complex, the α subunit exchanges a GDP molecule for a GTP. This results in the dissociation of the G protein from the receptor, and the release of the α subunit from β and γ (Kaziro *et al.*, 1991). The α subunit then binds to and activates the enzyme adenylate cyclase which converts ATP to cAMP (Iyengar, 1993). cAMP then acts as the "second messenger" in the cascade. cAMP binds to and activates protein kinase A (PKA). PKA is tetramer composed of two regulatory and two catalytic subunits (Taylor 1989). Each regulatory subunit is capable of binding two cAMP molecules. The binding of cAMP by the regulatory subunits releases and activates the catalytic subunits. The catalytic subunits then phosphorylate a variety of intracellular proteins, including transcription factors (Habener, 1995) which modulate gene transcription (figure 1.5A).

The regulation of this cascade occurs on many levels (Houslay and Milligan, 1997; Daniel *et al.*, 1998). First, phosphorylation of the receptor by kinases impairs their interaction with G-proteins (Freedman and Lefkowitz, 1996; Aragay *et al.*, 1998; Lefkowitz 1998; Pitcher *et al.*, 1998). Second, inhibitory α subunits ($G_{i\alpha}$) can bind to adenylate cyclase and prevent its activation (Hanoune *et al.*, 1997; Houslay and Milligan, 1997). Third, phosphodiesterases hydrolyze cAMP to AMP, decreasing the concentration of cAMP in the cell (Conti *et al.*, 1995). Fourth, small proteins known as PKA inhibitors (PKI), bind to PKA and export it out of the nucleus (Fantozzi *et al.*, 1994).

Figure 1.5, Activation of CREB by different signaling pathways. A) Activation of CREB by the cAMP pathway is as described in the text (page 35). AC, adenylate cyclase; α, β, γ, subunits of receptor coupled G protein; C and R, catalytic and regulatory subunits of PKA; CRE, cAMP response element. B) Activation of CREB by the calcium (Ca^{2+}) and mitogen-activated signaling pathways (reviewed in Shaywitz and Greenberg, 1999). Depolarization of the cell membrane results in an influx of Ca^{2+}. Ca^{2+} can complex with calmodulin and activate the Ca^{2+}/calmodulin-dependent protein kinase-four (CaMK IV) which phosphorlyates CREB on serine 133. Note also that Ca^{2+} can activate PKA through calcium-sensitive adenylate cyclases. The binding of mitogens such as epidermal growth factor (shown; EGF) or nerve growth factor (not shown) to their receptors activates the GTP-binding protein Ras. Ras then activates the serine/threonine kinase Raf. Raf phosphorylates and activates MEK, a dual function mitogen-activated protein kinase kinase (MAPKK). MEK phosphorylates and activates ERK, a MAPK, which in turn phosphorylates and activates the ribosomal S6 kinase-two (RSK-2). RSK-2 then phosphorylates CREB on serine 133.

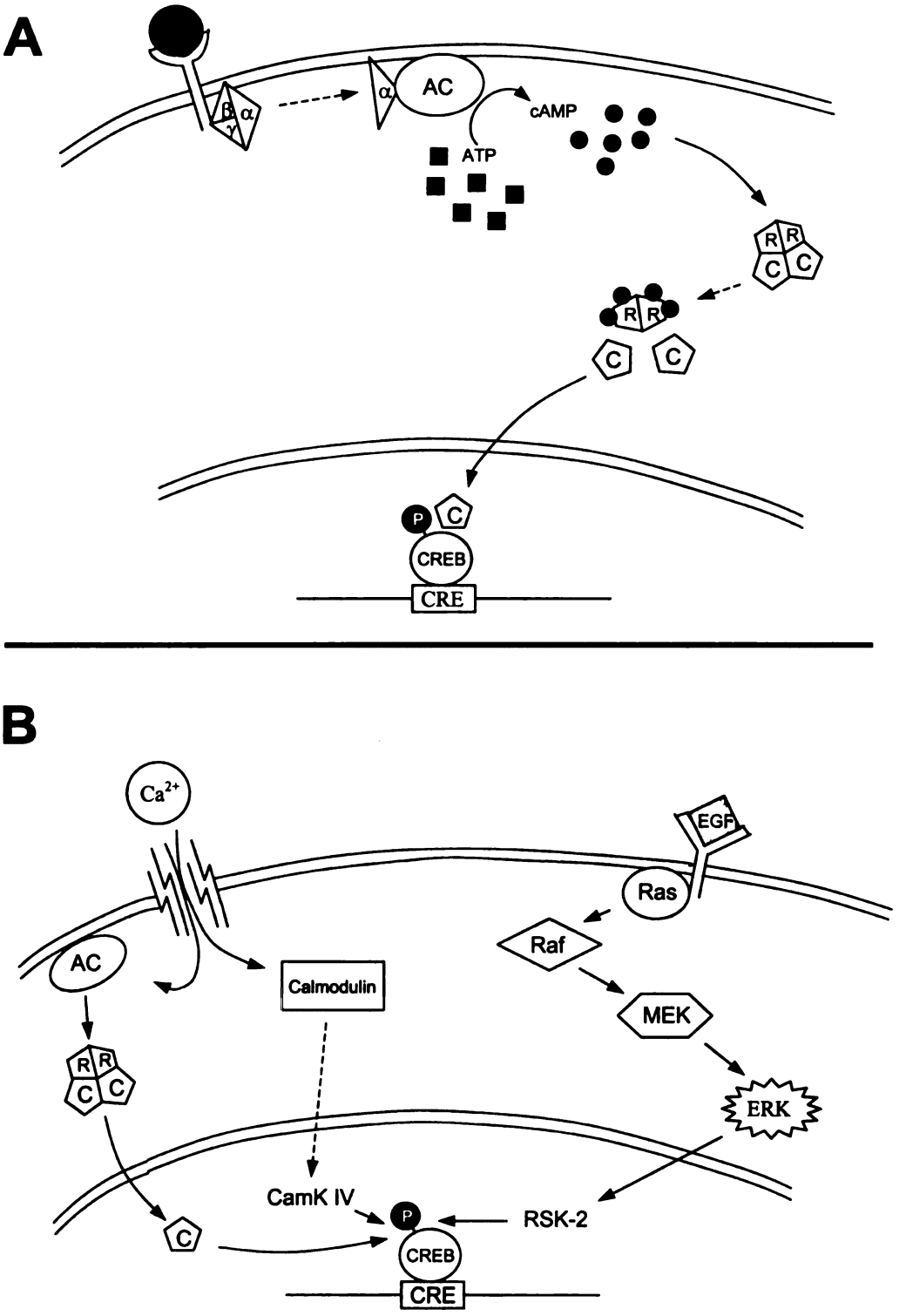


Figure 1.5, Activation of CREB by different signaling pathways.

Last, the inactivation of transcription factors through dephosphorylation by phosphatases (Cohen and Cohen, 1989).

Transcription factors responsive to the cAMP signaling pathway.

The transcription factors best characterized for their regulation by PKA are the cAMP response element binding protein (CREB) (Hoeffler *et al.*, 1988), the cAMP response element modulator (CREM) (Foulkes and Sassone-Corsi, 1992), and activating transcription factor 1 (ATF-1) (Hai *et al.*, 1989). CREB, CREM, and ATF-1 represent a subfamily of a large group of transcription factors known as bZIP proteins (Meyer and Habener, 1993; Hurst, 1994). Members of the bZIP family share a common structural motif at their C-terminus: a region rich in positively charged (basic) amino acids is followed by a sequence containing leucines at every seventh position (a heptad repeat) (Landschulz *et al.*, 1988). The basic region forms two α-helices that fit into the major groove of DNA, and so it is responsible for DNA-binding. The adjacent section forms an amphipathic α-helix with all of the leucines on the same side, and is responsible for protein dimerization. The helices of each member of the dimer wrap around each other in a "coiled coil" structure, with the leucine residues fitting together like the teeth of a zipper (Rasmussen *et al.*, 1991). Thus the name bZIP comes from the **b**asic region/ leucine **zip**per motif.

CREB, CREM, and ATF-1 show a striking degree of homology both in amino acid sequence and gene structure, and are likely the results of gene duplication (Hai *et al.*, 1989; Hoeffler *et al.*, 1990; Meyer and Habener, 1992;

Meyer *et al.*, 1993). Because of the high degree of homology in their bZIP regions, these proteins are able to form heterodimers with each other (reviewed in Meyer and Habener, 1993). Consequently, all three bind to the consensus sequence for the cAMP response element (CRE: TGACGTCA; Montminy *et al.*, 1986). In addition to the bZIP region, these proteins have two other domains necessary to function as transcriptional activators (figure 1.6). The kinase inducible domain (KID) is a region of approximately 60 amino acids that contains consensus phosphorylation sites for PKA, protein kinase C (PKC), Calcium/calmodulin-dependent kinase (CaMK) I, II, and IV, casin kinase II (CKII), glycogen synthase kinase-3 (GSK-3), and ribosomal S6 kinase-2 (RSK-2). Most critical is the serine in the sequence RRPSY, conserved in all 3 proteins (S133 in CREB, S117 in CREM, S63 in ATF-1). A mutation of this serine to an alanine or aspartic acid abolishes all transcriptional ability, demonstrating that phosphorylation of S133 plays an important structural role and is not simply providing a negative charge (Gonzalez and Montminy, 1989). Indeed, phosphorylation of this serine is critical for interaction with the coactivator CREB-binding protein (CBP) (Chrivia *et al.*, 1993; Parker *et al.*, 1996; Radhakrishnan *et al.*, 1997). The interaction of KID with CBP was localized to a 94 amino acid region, termed KIX, near the N-terminus of CBP (Chrivia *et al.*, 1993; Parker *et al.*, 1996). Though CBP was first identified as a nuclear protein that interacted with the phosphorylated, but not the unphosphorylated, form of CREB (Chrivia *et al.*, 1993), it is now known to serve as a coactivator for a wide variety of transcription factors (Chakravarti *et al.*, 1996; Janknecht and Hunter; 1996;

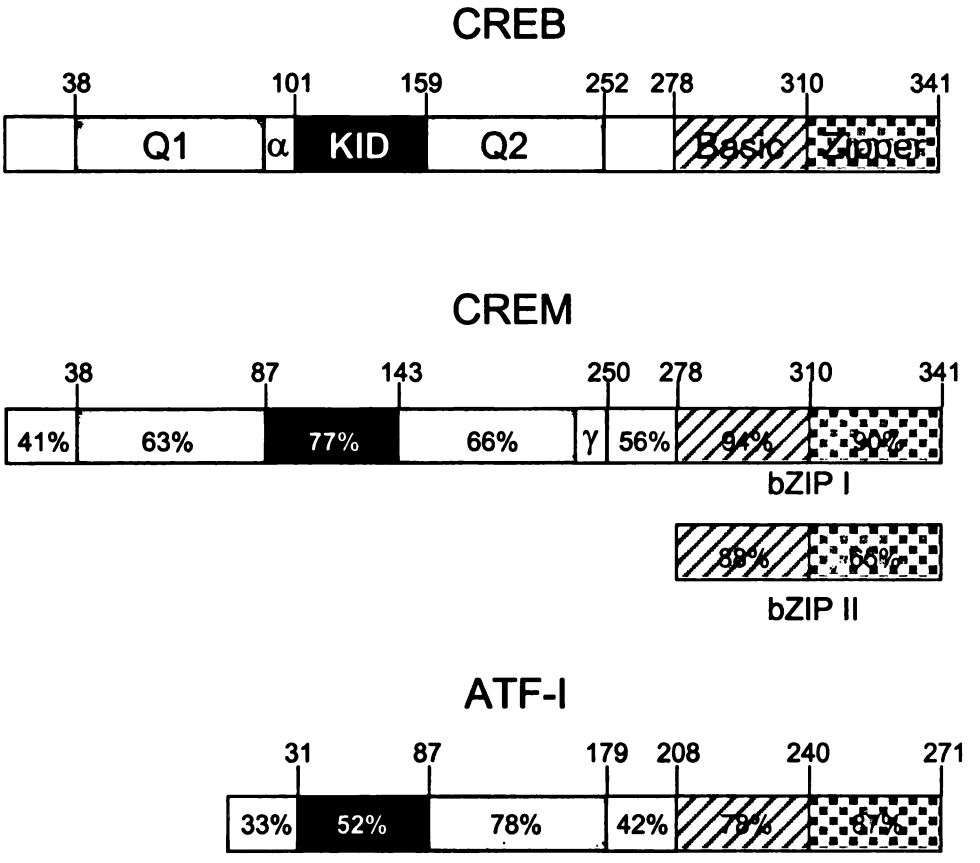


Figure 1.6, Shematic representation of CREB, CREM, and ATF-1 proteins. The identity of each domain is noted in the sketch for CREB: KID, kinase-inducible domain; Q1 and Q2, glutamine rich domains; Basic, region rich in basic amino acids that is responsible for DNA-binding; Zipper, region containing the heptad repeat of leucines responsible for dimerization. The percent identity of human CREM and ATF-I relative to CREB is noted in their respective domains. α is a 14 amino acid exon not present in the CREM or ATF-I genes. The major isoforms of CREB are CREB 341 and CREB 327; in CREB 327 the α exon is spliced out of its transcript. Serine 119 of CREB 327 is equivalent to serine 133 of CREB 341. γ is a 13 amino acid exon not present in the CREB or ATF-I genes. CREM is unique in that it encodes for 2 different basic region/leucine zippers (bZIP I, II) that are regulated by alternative splicing (Foulkes *et al.*, 1991). CREM also contains a second, cAMP-inducible, promoter in an intron upstream of the bZIP region (not shown). The product of this transcript (Inducilbe cAMP early repressor: ICER) is a protein that has only the DNA-binding and protein dimerization domains and so consiquently acts as a repressor of cAMP-regualted genes (Molina *et al.*, 1993).

Gerritsen *et al.*, 1997; Glass *et al.*, 1997; Goldman *et al*, 1997; Collingwood *et al.*, 1999).

The KID region by itself is likely not sufficient to mediate cAMP-stimulated transcriptional activation. CREM isoforms α, β, and γ all contain the KID region but are lacking the flanking glutamine rich (Q) regions and act as repressors (Laoide *et al.*, 1993). In addition, phosphorylation of serine 142 (S142) was known to abrogate cAMP-stimulated activation of CREB even if S133 was phosphorylated (Sun *et al.*, 1994). This effect could be mimicked by mutating S142 to aspartic acid (S142D) without interfering with the binding of CREB to CBP (Sun and Maurer, 1995). However, in the same study, mutation of S142 to glutamic acid (S142E) had no affect on transcriptional activation, thus, the inhibition caused by the phosphorylation of S142 is not due only to the introduction of a negative charge. Circular dichroism and nuclear magnetic resonance have shown that the KID is unstructured in both the phosphorylated and unphosphorylated states (Parker *et al.*, 1996; Richards *et al.*, 1996; Radhakrishnan *et al.*, 1998). Upon binding to KIX, the KID undergoes a structural transition to form two amphipathic α-helices that participate in hydrophobic interactions with CBP (Radhakrishnan *et al.*, 1997). In further experiments, phosphorylation of S142 was shown to disrupt the secondary structures mediated by the interaction of KID with KIX (Parker *et al.*, 1998).

In a recent study, Cardinaux *et al.* (2000) identified a conserved domain between CREB and the sterol-responsive element binding protein (SREBP), which also interacts with the KIX domain of CBP but in a phosphorylation-

independent manner. Substitution of six nonconserved amino acids from SREBP for the RRPSYR sequence of CREB resulted in the recruitment of CBP and transcriptional activation independent of phosphorylation of CREB. This sequence (DIEDML) alone isn't sufficient for the recruitment of CBP. Mutation of the central aspartic and glutamic acid residues had no effect on the binding of SREBP to CBP. Four amino acids flanking this region were shown by mutational analysis to be important for SREBP/CBP interaction (Cardinaux *et al.*, 2000). Significantly, 3 of these were either identical or conservative substitutions in the corresponding KID region of CREB. Additionally, these residues had been identified previously as having direct contact with the KIX domain of CBP (Radhakrishnan *et al.*, 1997). Thus, substitution of the DIEDML sequence served to convert the KID region of CREB into a SREBP-like domain. In fact, this chimeric CREB ($CREB_{DIEDML}$) had a greater affinity for CBP than did the wild-type CREB, and was as strong of an activator. Surprisingly, $CREB_{DIEDML}$ showed an additional 2-fold increase in activation when PKA was present, suggesting that a phosphorylation event following the recruitment of CBP augments CREB-mediated transcription. It should be noted that only 5 of the 11 residues in KID identified to be in contact with KIX are conserved in SREBP. This suggests, in combination with the fact that $CREB_{DIEDML}$ has a stronger affinity for CBP than the wild-type protein, that the binding mechanisms of SREBP and phosphorylated CREB to CBP are not identical (Cardinaux *et al.*, 2000). Thus, it is believed that the role of the KID is to recruit CBP to the promoter, and that this interaction may be modulated by the differential phosphorylation of the KID at

other sites. The recruitment of CBP is dependent on the phosphorylation of S133. Although originally identified as the site of phosphorylation by PKA, S133 is also known to be phosphorylated by CamK IV and RSK-2 (Sun *et al.*, 1994; Xing *et al.*, 1996; Pende *et al.*, 1997). Consequently, CREB is able to mediate signaling of several different pathways (figure 1.5B)

As mentioned previously, the KID alone is not sufficient to mediate transcription. Deletion mutants have shown that the glutamine (Q) rich regions flanking the KID (figure 1.5) are important for CREB activity (Brindle *et al.*, 1993; Quinn, 1993). Of the two, Q2 appears to be the more important. Deletion of the Q2 domain severely affects CREB activation, while deletion of Q1 has only moderate affects. In agreement with these results, 3 splice variants of CREM have different combinations of the Q domains. CREMτ has both Q1 and Q2 and shows the strongest activation of the CREM isoforms. CREMτ1 and CREMτ2 have only the Q1 or Q2 domain respectively and are weaker activators; CREMτ2 is the stronger of the two isoforms, having approximately two-thirds of the activity of CREMτ while CREMτ1 has only one-third the comparable activity (Laoide *et al.*, 1993). Additionally, ATF-1, which has only the corresponding Q2 region, is strong activator (Rehfuss *et al.*, 1991). Furthermore, when fused to the yeast GAL4 DNA binding domain, Q2 can activate transcription in a phosphorylation-independent manner, suggesting that it encodes for a constitutive transactivation domain (Brindle *et al.*, 1993; Quinn, 1993).

Glutamine rich domains have been characterized in other transcription factors and are postulated to interact with components of the basal transcription

machinery (Courey and Tijan, 1988; Triezenberg, 1995). In *in vitro* transcription assays, unphosphorylated CREB when mixed with crude nuclear extracts was able to stimulate transcription of a reporter gene containing 5 copies of the CRE (Ferreri *et al.*, 1994). Further experiments showed that the Q2 region either fused to a GAL4 DNA-binding domain, or left in context with the bZIP domain but lacking the KID and Q1 domains, was able to stimulate transcription. Reconstitution experiments determined that TFIID, but not TBP alone, was necessary for Q2 activity, indicating an interaction between Q2 and a TBP associated factor (TAF) (Ferreri *et al.*, 1994). Affinity chromatography demonstrated that Q2 was able to specifically bind *Drosophila* TAF 110 (homolog to human TAF130), and these results were corroborated by the yeast two-hybrid system (Ferreri *et al.*, 1994).

In vitro transcription assays were also utilized to show the necessity of both the KID and Q2 regions for CREB-induced transcription. CBP was found to be in association with RNA polymerase II in HeLa cell nuclear extracts, suggesting that one role of CBP is the recruitment of basal transcription factors (Kee *et al.*, 1996). Phosphorylated CREB was able to bind the CBP-RNA polymerase complex (CBP-RPC) from a partially purified fraction and mediate transcription in vitro (Nakajima *et al.*, 1997). Pulldown assays with GST fusion constructs determined that the KID region of CREB, and not Q2, was responsible for recruitment of CBP-RPC. However, recruitment of this complex was not sufficient to mediate CREB-induced transcription. Addition of antiserum to the human TAF 130, a component of TFIID, abolished CREB-induced, but not basal,

transcriptional activity (Nakajima *et al.*, 1997). GST pulldown assays demonstrated that the association between CREB and TFIID was mediated by the Q2 domain in a phosphorylation-independent manner. The addition of TAF 130 antiserum severely inhibited this interaction, supporting previous work that Q2 interacts with this TAF of TFIID. Thus, CREB-mediated transcription requires the coordinated activity of two domains: the KID for recruitment of CBP and RNA polymerase II, and the Q2 for the recruitment of TFIID.

CRE-binding proteins in insects.

Thus far, CRE-binding proteins have been identified only in *Drosophila*. Of the two genes identified, dCREB-A (also known as BBF-2) (Abel *et al.*, 1992; Smolik *et al.*, 1992) and dCREB2 (also known as dCREB-B) (Usui *et al.*, 1993; Yin *et al.*, 1995b), dCREB2 shows the highest similarity to mammalian CREB proteins, having conserved 50 of the 58 amino acids in the bZIP region (see figure 4.2, page 109) (Usui *et al.*, 1993; Yin *et al.*, 1995b). In comparison, dCREB-A has only 38% homology with the bZIP domain of mammalian CREB, although this identity increases to 50% if only the DNA-binding (basic) region is considered (Abel *et al.*, 1992; Smolik *et al.*, 1992). In addition, dCREB-A contains the core RRKKKEY sequence common to all CREB/CREM/ATF-1 proteins. Importantly, both proteins have been shown to be able to bind to CREs (Abel *et al.*, 1992; Smolik *et al.*, 1992; Usui *et al.*, 1993; Yin *et al.*, 1995b). Outside of the bZIP domain, these proteins show little homology to mammalian CREB proteins. However dCREB2 does contain a consensus PKA recognition

sequence as well as potential phosphorylation sites for CaMK II and PKC. Moreover, dCREB2 also contains several glutamine rich stretches in its N-terminus. In contrast, dCREB-A does not have a PKA recognition site and is only 5% enriched for glutamine. Nevertheless, dCREB-A does contain potential phosphorylation sites for CaMK II and thus could be regulated through a different signaling pathway.

The two genes also differ in their expression pattern and the number of transcripts produced. Northern blot analysis detects only one transcript of approximately 4.3 kb for dCREB-A (Abel *et al.*, 1992; Smolik *et al.*, 1992). Its expression pattern is limited and perhaps a bit controversial. Abel *et al* (1992) reported dCREB-A (then called BBF-2 for box-b binding factor 2) to be expressed throughout all developmental stages, though highest levels were seen in early to mid embryonic development and in adult males. Northern analysis also revealed the gene to be expressed in the trachea, gut, and fat body of third instar larvae (though only this larval stage was examined for tissue distribution of the transcript). Indeed, the cDNA was cloned by screening a *D. melanogaster* expression library with the box-b element from the fat body enhancer of the *D. mulleri Adh-1* gene. Subsequently, the protein was shown to bind to regulatory elements in both insect fat body and mammalian liver-specific enhancers (Abel *et al.*, 1992). Mutation of the BBF-2 binding site in the *Adh* adult enhacer (AAE; see page 18) resulted in a 4-fold decrease in expression of a β-galactosidase reporter gene in transgenic flies, thus suggesting a role for this protein in regulating fat body gene expression. However, other researchers have not

detected dCREB-A transcripts or protein in the fat body during any of the life stages (Smolik *et al.*, 1992; Andrews *et al.*, 1997). The dCREB-A transcript is expressed primarily in the salivary glands of embryos, and at lower levels in the epidermis, trachea, and foregut. In the adult, transcripts are detected in the brain, midgut epithelial cells, salivary gland, ovarian follicle cells and male reproductive tract (Smolik *et al.*, 1992; Andrews *et al.*, 1997). In agreement with this, Abel *et al.* (1992) detected strong binding with ovarian nuclear extracts to a box-b probe in electrophoretic mobility shift experiments; this binding was abolished by the addition of antisera specific to the protein. Transient transfections in cell culture have shown dCREB-A to be a transcriptional activator, but not a mediator of cAMP-stimulated transcription (Abel *et al.*, 1992; Smolik *et al.*, 1992)

In contrast to dCREB-A, dCREB2 appears to be a homolog of the mammalian CREB and CREM genes (Yin *et al.*, 1995b). The gene is ubiquitously expressed in all life stages. Northern analysis reveals transcripts ranging in size from 650 bp to 3.5 kb, and several spliced variants have been identified by PCR and cDNA library screening (Usui *et al.*, 1993; Yin *et al.*, 1995b). Significantly, isoform dCREB2-a was shown to be a PKA-responsive transcriptional activator (Yin *et al.*, 1995b). Similar to roles of CREB and CREM in mammals, several lines of evidence support the notion that dCREB2 is involved in long-term memory formation in *Drosophila* (Belvin and Yin, 1997). Some flies with impaired learning abilities were found to have mutations in genes for a phosphodiesterase and adenylate cyclase (Chen *et al.*, 1986; Levin *et al.*,

1992). Additionally, transgenic flies that overexpressed either an inhibitor of PKA, or the PKA catalytic subunit, also had impaired learning abilities (Drain *et al.*, 1991). Importantly, in transgenic flies, induced expression of the transcriptional activator dCREB2-a enhanced long-term memory formation, while induced expression of the transcriptional repressor dCREB2-b blocked long-term memory formation (Yin *et al.*, 1994; Yin *et al.*, 1995a).

Recently, dCREB2 was also shown to affect circadian rhythms in *Drosophila*. Through the use of transgenic flies carrying a luciferase reporter construct containing 3 copies of a CRE, enzyme activity was found to cycle over a 24 hour period in both a 12 hour light / 12 hour dark cycle or in constant darkness, indicating that this not simply a response to light (Belvin *et al.*, 1999). A mutation in the dCREB2 gene shortened the circadian locomotor rhythm and dampened the oscillation of *period*, a known clock gene (clock genes are the time keeping mechanisms of organisms, regulating the normal biological rhythms), implying dCREB2 as the regulator. Cycling activity of the reporter construct was also abolished in a *period* mutant, suggesting that dCREB2 and Period affect each other and that the two genes participate in the same regulatory feedback loop (Belvin *et al.*, 1999).

Significance of current research.

The primary focus of Dr. Raikhel's research is to understand how the YP genes are regulated during vitellogenesis. These genes are expressed only in the fat body of female mosquitoes after a blood meal. Their regulation is likely to

be complex as it incorporates sex-, stage-, and tissue-specific factors. Because of the functional analogy between the vertebrate liver and the insect fat body, regulation of liver-specific gene expression was used as a model to investigate fat body-specific gene expression. Of the four classes of liver-enriched transcription factors characterized from vertebrates (HNF-1, 3, 4, and C/EBP), HNF-4 is the only protein for which an insect homolog has been found to be expressed in the fat body. Additionally, since the tissue distribution of HNF-4 in insects is analogous to that in vertebrates, it is hypothesized that the insect homolog performs many of the same functions. HNF-4 has been shown to regulate genes for serum proteins, blood coagulation factors, metabolism of carbohydrates, lipids, steroids and amino acids, and liver differentiation. In concordance with this, the insect fat body is involved in the synthesis and storage of hemolymph proteins, carbohydrates, lipids, and detoxification enzymes. Therefore, a mosquito homolog to HNF-4 was cloned and characterized as a beginning to elucidating the regulation of gene expression in the fat body.

Regulation of the YP genes in the fat body also requires that their expression be coordinated with the function of the midgut and ovaries since the blood meal provides the amino acids for YP synthesis and the oocytes internalize these proteins. Factors from the head and ovaries, presumably peptide hormones, have already been shown to be necessary for maximum YP synthesis by the fat body and uptake by the ovaries. The binding of these hormones to cell surface receptors initiates a cascade of events that profoundly affect the physiology of the cell, including altering gene transcription. As noted earlier, the

cAMP signal transduction cascade is one of the most common pathways used by cells to mediate these changes. Stimulation of this pathway leads to the activation of the CREB family of transcription factors which bind to specific DNA sequences in target genes and modulate their expression. This cascade has already been implicated in the regulation of YP uptake by the ovaries and the excretion of salts and water by the Malphigian tubules after a blood meal. In view of this, a mosquito homolog to CREB was cloned and characterized to examine what role this pathway may play in the regulation of YP gene expression.

CHAPTER 2

MATERIALS AND METHODS

MATERIALS AND METHODS

Polymerase chain reaction (PCR) and cDNA library screening

For AaHNF-4, Degenerate primers were designed for PCR based on conserved residues in the C and E domains of the *Drosophila, B. mori,* and rat HNF-4 proteins (see figure 3.2, page 69). The forward primer, 5'- AGG**TCTAGA** (T/C)GGNTG(T/C)AA(A/G)GGNTT(T/C)TT-3', was designed from the amino acid sequence DGCKGFF and contained an *Xba I* restriction site (shown in bold) in its 5' end to facilitate cloning. The reverse primer, 5'- AGA**GAATTC**GCNAC(T/C) TG(A/G)TC (A/G)TCNAG-3', was designed from the amino acid sequence LLDQVA and contained an *Eco RI* restriction site (shown in bold) in its 5' end. Template cDNA was prepared by reverse transcription of 20 μg of total RNA prepared from the fat bodies of female mosquitoes dissected at 24 hour PBM. PCR conditions were as follows: initial denaturation at 95° C for 4 minutes followed by 10 cycles at 95° C for 30 seconds, 45° C for 30 seconds, 72° C for 45 seconds, and then 30 cycles at 95° C for 30 seconds, 55° C for 30 seconds, 72° C for 45 seconds. A band of approximately 400 bp was amplified and cloned into the *Xba I* and *Eco RI* restriction sites of the plasmid pBluescript SK-. Sequencing of the insert confirmed that it encoded for a portion of a nuclear hormone receptor homologous to HNF-4. This fragment was then used to screen a vitellogenic, fat body, cDNA library (Deitsch and Raikhel, 1993) using standard protocols (Ausuble *et al.*, 1989; Sambrook *et al.*, 1989) Isolated cDNA clones

were sequenced on both strands by the dideoxy method (Sanger *et al.*, 1977) using the T7 sequenase version 2.0 sequencing kit (Amersham Life Sciences).

The same approach was used for isolation of AaCREB cDNA clones. Degenerate primers were designed for PCR based on conserved residues in the bZIP domain of the *Drosophila* and human CREB proteins. The forward primer, 5'- AGG**TCTAGA**A(A/G)AA(T/C)(C/A)GIGA(A/G)GCIGC-3', was designed from the amino acid sequence KNREAA and contained an *Xba I* restriction site (shown in bold) in its 5' end to facilitate cloning. The reverse primer, 5'-AGA**GAATTC**TT(A/G)TT(T/C)TG(A/G)TT(T/C)TCIA(G/A)IACIGC-3', was designed from the amino acid sequence AVLENQNK and contained an *Eco RI* restriction site (shown in bold) in its 5' end. PCR conditions were the same as described above for AaHNF-4, except that the cDNA library was used as template. A band of 112 bp was amplified and cloned into the *Xba I* and *Eco RI* restriction sites of the plasmid pBluescript SK-. Sequencing of the insert confirmed that it encoded for a portion of the bZIP domain of a mosquito CREB protein. This fragment was then used to screen the vitellogenic, fat body, cDNA library using standard protocols (Ausuble *et al.*,1989; Sambrook *et al.*, 1989) Isolated cDNA clones were sequenced by the W. M. Keck facility at Yale University using *Taq* FS DNA polymerase and fluorescent-dideoxy terminators in a cycle sequencing method, and analyzed using an automated Applied Biosytems 377 DNA sequencer.

Northern blot analysis

For analysis of AaHNF-4 expression, total RNA was isolated from female mosquito fat bodies dissected at selected time points pre- and post-blood meal using guanidine thiocyanate (Chomczynski and Sacchi, 1987). Poly(A+) mRNA was purified from the total RNA by Biomag oligo$(dT)_{20}$ magnetic beads, according to the manufacturer's protocols (PerSeptive Diagnostics, Inc.). The RNA was fractionated by electrophoresis in a 1% agarose/formaldehyde gel and transferred to a nitrocellulose membrane by conventional capillary blotting. The DNA fragments used for probes are described in figure 3.1(page 68). Probes were labeled by the random primer method and hybridization was carried out overnight under high stringency conditions (Ausuble *et al.*,1989; Sambrook *et al.*, 1989). Following hybridization, the blot was first washed twice at room temperature in 2 X SSC, 0.1% SDS, then twice at 60° C in the same buffer, followed by a final wash at 60° C in 0.2 X SSC, 0.1% SDS. Autoradiography of the blot was conducted at –80° C with intensifying screens.

To examine AaCREB expression, poly(A+) RNA isolated utilizing the QuickPrep Micro mRNA Purification Kit (Pharmacia Biotech Inc.). The mRNA was fractionated by gel electrophoresis and transferred to a nitrocellulose membrane as described above. A single-strand, antisense DNA probe was generated by PCR using the pBSAaCREB (see below in "Electrophoretic Mobility Shift Assays", page 57) cDNA as template and labeling with [α-^{32}P]dATP. The specific activity of the probe was determined by scintillation counting, and added to the hybridization buffer at a concentration of 1 X 10^6 DPM per milliliter of

solution. Hybridization and washing was performed as described above. This blot had been previously hybridized with probes for mosquito lipophorin, actin, and vitellogenic carboxypeptidase (VCP) (Sun *et al.*, 2000).

Preparation of fat body nuclear extracts

Nuclear extracts were prepared as described by Miura *et al.* (1999) using the low salt (10 mM NaCl) homogenization buffer. Briefly, 200-300 mosquito abdomens with adhering fat body tissue were dissected from female mosquitoes 3-5 days post eclosion or a various time points after a blood meal. The tissues were frozen in liquid nitrogen then ground to a fine powder with a mortar and pestle. The powder was then resuspended in a low salt buffer (10 mM HEPES pH 7.9, 1 mM EDTA, 10 mM NaCl, 0.6% (v/v) Triton X-100, 4 mM DTT, 1 mM 4-(2-aminoethyl)-benzensulfonylflouride (AEBSF), 5 μg/ml each of leupeptin, pepstatin, and antipain, 2 mM benzanidine, and 9 μg/ml aprotinin) and homogenized in a Dounce homogenizer with a loosely fitted pestle. The samples were twice subjected to centrifugation, 800 X g for 30 seconds at 4°C, to remove cell debris and the supernatant transferred to a new, sterile tube each time. The nuclei were recovered by centrifugation at 15,000 X g for 10 minutes at 4°C. The supernatant was discarded and the nuclei were resuspended in extraction buffer (20 mM HEPES pH 7.9, 0.2 mM EDTA, 420 mM NaCl, 1.2 mM $MgCl_2$, 25% glycerol, 1 mM DTT, 1 mM AEBSF, 5 μg/ml each of leupeptin, pepstatin, and antipain, 2 mM benzanidine, and 9 μg/ml aprotinin) at a concentration of 2 mosquito fat bodies equivalents per microliter of buffer. The samples were

incubated on ice for 20 minutes then spun in a microcentrifuge at maximum speed for 5 minutes. The supernatant was dispensed in 10 μl aliquots (equivalent to 20 mosquito fat bodies) and stored at -80° C.

Electrophoretic Mobility Shift Assays (EMSAs)

EMSAs were performed using either *in vitro* synthesized proteins or nuclear extracts. For *in vitro* expression of proteins, the AaHNF-4a and -4b cDNAs were subcloned into the *Eco RI* restriction site of the transcription vector pGEM.3Z (Promega) downstream of the SP6 RNA polymerase promoter. The AaHNF-4c isoform was subcloned into the *Eco RI* restriction site of pBluescript SK- downstream of the T3 RNA polymerase promoter. pBSAaCREB was created by inserting the 1.25 kb *Acc I* fragment, containing the entire CREB coding region, into the *Eco RV* restriction site of pBluescript SK- (Stratagene) with the 5' end of the insert positioned downstream of the T3 promoter. The rat HNF-4α1 and DmGATAb clones were gifts from Drs. F. Sladek and T. Abel (Sladek *et al.*, 1990; Abel *et al.*, 1993). Proteins were synthesized *in vitro* by a coupled transcription and translation rabbit reticulocyte lysate kit (TNT, Promega) according to the manufacturer's directions. One microgram of DNA was used in a total reaction volume of 50 μl. The TNT reactions were conducted at 30° C for 2 hours, then stored at –80° C until needed.

Binding reactions with the *in vitro* synthesized proteins were performed in a total volume of 20 μl containing 3 μl of the appropriate TNT sample, 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM EDTA,

2 μg poly(dIdC), and 4% glycerol. The samples were incubated at room temperature for 15 minutes, then 0.05 picomoles of ^{32}P-labeled DNA probe was added and incubation was continued for an additional 15 minutes. The samples were loaded on a 5% nondenaturing polyacrylamide gel (prerun for 1 hour) in 0.5X TBE buffer and run at 10 V/cm. After electrophoresis, the gel was dried and autoradiographed with an intensifying screen at –80° C.

To initiate heterodimer formation between AaHNF-4c and other HNF-4 isoforms, the TNT samples were first mixed together and heated at 50° C for 1 minute, then cooled at room temperature for 15 minutes. The volume was brought up to 20 μl in shift buffer and the reaction carried out as described above. Control incubations with each HNF-4 isoform without addition of AaHNF-4c showed no or negligible reduction in specific binding due to heating

Binding reactions with nuclear extracts were performed in a similar matter (without heating at 50° C) using 5 μl of nuclear extracts in a total reaction volume of 30 μl. When needed, preimmune serum, α-GSTHNF-4(Com), or anti-phosphoCREB antibody (New England Biolabs) was added at a 1:100 dilution. Initial incubations were carried out on ice for 30 minutes before the addition of the labeled DNA probe. DNA probes were made by annealing together complementary oligonucleotides and labeling the recessed 3' ends with either [α–^{32}P]dATP or [α-^{32}P]dCTP using the Klenow fragment of DNA polymerase I. The oligonucleotides used are listed below (only one strand is shown).

APF-1: 5'-GCGC**TGGGCAAAGGTCA**CCTGC-3'

DR1: 5'-GATCC**AGGTCAAAGGTCA**GGATC-3'

VgHNF4: 5'-ATCGGG**AGGCCAATGGTCG**AGTGAC-3'

VCPHNF4: 5'-ATGCAAA**GGGTCG**TT**AGGTCA**AATCGAT-3'

SomCRE: 5'-GATCCTTGGC**TGACGTCA**GAGAGATC-3'

VgCRE: 5'-GATCTGCCGA**ATACGTCA**TCTTTCGATC-3'

VCPCRE1: 5'-GATCACTGTTA**ATACGTCA**GATTTGATC-3'

VCPCRE2: 5'-GATCCACCGA**TTCCGTCA**TATGTGGATC-3'

Box A: 5'-AGTGGTAT**TGATAA**GACATG-3'

bZIP1: 5'-ATCGTTTATT**ATTGCAAAA**TCGAAAGTC-3'

bZIP2: 5'-ATCGAAATTC**TTGAGCAAG**TACAATTTGCCCATG-3'

Bold face type indicates putative binding sites.

Transient transfection in cell culture

Drosophila S2 cells: The AaHNF-4 cDNAs were cloned into the *Eco RI* sites of pAc5/V5-His C (Invitrogen) to create pAc5AaHNF-4a, -4b, and -4c. The vector pAc5/V5-His/*lacZ* (Invitrogen) is a control vector expressing β-galactosidase and was used to normalize transfections in the S2 cells. The reporter vector pDR1(3X)AdhCAT was derived from the plasmid pD-33CAT, which contains the *Drosophila alcohol dehydrogenase* (*Adh*) promoter from -33 to +53 inserted in front of the *chloramphenical acetyltransferase* (CAT) gene (Krasnow *et al.*, 1989). pD-33CAT was digested with *Xba I* and *Sac I* to release the *Adh*CAT fusion construct. The ends were made blunt with T4 DNA polymerase and the construct was cloned into the *Eco RV* restriction site of pBluescript SK- to create pBS*Adh*CAT. Three copies of the DR1 probe were then inserted into

the *Bam HI* restriction site upstream of the *Adh* promoter in pBS*Adh*CAT to create pDR1(3X)*Adh*CAT.

Drosophila S2 cells were maintained in Schneider's *Drosophila* medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS) plus 50 units of penicillin G and 50 μg of streptomycin sulfate per milliliter of medium. Each transfection contained 100 ng of pDR1(3X)AdhCAT, 50 ng of pAc5/V5-His/*lac*Z, and varying amounts of the pAc5AaHNF-4 constructs. The total amount of DNA was kept constant at 650 ng with empty pAc5/V5-His C vector. For each sample, enough DNA for 3 replicates was complexed with 10 μg of the polycationic lipid CELLFECTIN (Life Technologies) in a total volume of 45 μl. Fifteen microliters was aliquoted into each of 3 wells of a 48-well tissue culture plate and incubated for 30 minutes at room temperature to allow DNA/lipid complexes to form. The DNA/lipid mix was then overlaid with 5 $X10^5$ cells suspended in 150 μl of serum-free medium. At 6 hours after transfection, 150 μl of medium containing twice the normal concentration of FBS and antibiotics was added to each well. The cells were analyzed 48 hours after transfection for enzyme activity using Luciferase and β-galactosidase Assay Reagents (Promega).

Human 293T cells: The AaHNF-4 cDNAs were cloned into the *Eco RI* site of pcDNA3.1/Zeo(+) (Invitrogen) to create pcAaHNF-4a, -4b, and -4c respectively. The reporter vector p*apoCIII*-Luc contains the *apoCIII* promoter from -854 to +22 inserted upstream of the *luciferase* gene. The p*apoCIII*-Luc and the human HNF-4α1expression vectors (Taylor *et al.*, 1996) were gifts from T.

Leff. The vector pCMVβ (Clonetech) is a control vector expressing β-galactosidase and was used to normalize transfections in the 293T and CV-1 cells.

Human embryonic kidney 293T cells were maintained in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% FBS. The day before transfection, the cells were seeded at a density of 1.2 X 10^5 in 12-well tissue culture plates. Each transfection contained 150 ng of p*apoCIII*-Luc, 50 ng of pCMVβ, selected amounts of hHNF-4α1 and/or AaHNF-4c, and empty pcDNA 3.1 vector for a total of 5μg of DNA. For each sample, enough DNA for 3 replicates was complexed with 60 μg of the polycationic lipid LIPOFECTAMINE (Life Technologies) in 300 μl of serum-free DMEM for 30 minutes at room temperature, after which 1.2 ml of serum-free media was added for a total volume of 1.5 ml. While the DNA/lipid complexes were forming, the media was removed from the cells and the cells were washed once with serum-free DMEM. Five hundred microliters of the DNA/lipid complex was then added to each of the three wells per sample. Four hours after transfection, an additional 500 μl of DMEM supplemented with 20% FBS was added to each well. The cells were collected at 48 hours post transfection and analyzed for Luciferase and β-galactosidase activity by the Dual-Light reporter gene assay system (Tropix) using a Wallac-Berthold LB96P-2 luminometer.

CV-1 cells: The AaCREB expression plasmid pcAaCREB was constructed by digesting pBSAaCREB with *Eco RI* and *Xho I*, and cloning the released insert into the *Eco RI* and *Xho I* restriction sites of pcDNA3.1/Zeo(+). The reporter plasmid pCRE-Luc (Stratagene) contains 4 copies of the CRE

consensus sequence upstream of a synthetic promoter fused to the *luciferase* gene. The plasmid pFC-PKA (Stratagene) encodes for the catalytic subunit of protein kinase A.

CV-1 cells were maintained in DMEM supplemented with 10% FBS and 0.1% non-essential amino acids (NEAA). The day before transfections, cells were seeded at a density of 2 X 10^5 cells per well in 6-well tissue culture plates. Each transfection contained 1 μg of the reporter gene pCRE-Luc, 250 ng of pCMVβ, varying amounts of pcAaCREB, and empty pcDNA3.1 vector for a total of 2.5 μg of DNA. CREB activation was stimulated by cotransfection with 50 ng of pFC-PKA. For each sample, enough DNA for 3 replicates was complexed with 30 μg of the polycationic lipid LIPOFECTAMINE (Life Technologies) in 600 μl of serum-free DMEM for 30 minutes at room temperature. While the DNA/lipid complexes were forming, the medium on the cells was replaced with 800 μl of serum-free DMEM supplemented with 0.1% NEAA. Two hundred microliters of the DNA/lipid complex was added to each of the three wells per sample. Five hours after the start of transfection, the media was replaced with fresh DMEM supplemented with 10% FBS and 0.1% NEAA. The cells were analyzed 48 hours after transfection for enzyme activity using Luciferase and β-galactosidase Assay Reagents (Promega).

Antibody production

Polyclonal antibodies against HNF-4 were generated in New Zealand white rabbits using GST-HNF-4 fusion proteins as antigens. A 914 bp *Eco RV*

fragment from the AaHNF-4c cDNA (encoding amino acids 3-306) was cloned into the *Sal I* site (which had been end filled by Klenow) of the GST fusion vector pGEX-4T-1 (Pharmacia) to create pGSTHNF-4(Com). The coding sequence for the first 33 amino acids of HNF-4a was amplified by PCR using the primers 5'-CGT**GGATCC**ATGAAGAGTATTGTGGAAGT-3' and 5'-TCA**CTCGAG**GATGTG TTGGAATAACGT-3' (*Bam HI* and *Xho I* restrictions sites are shown in bold; the ATG start codon is underlined). The amplified DNA was cloned into the *Bam HI* and *Xho I* restriction sites of pGEX-4T-1 to create pGSTHNF-4a(NH_2). The fusion proteins were expressed in the *Eschericia coli* strain BL21 and purified by affinity chromatography using Glutathione Sepharose 4B beads (Pharmacia). After extensive washing of the beads with 1X PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH adjusted to 7.3 with HCl), the fusion proteins were eluted from the matrix with Glutathione Elution Buffer (10 mM reduced glutathione, 50 mM Tris-HCl pH 8.0). The eluate was then dialyzed for 24 hours against 2 liters of 1X PBS with 3 changes of the PBS buffer. The fusion proteins were then sent to Cocalico Biologicals, Incorporated (Reamstown, PA) for production of antibodies. The IgG fraction of the antiserum was purified by precipitation with saturated ammonium sulfate (Ausubel *et al.*, 1989) and the antibodies were resuspended in a buffer of 100 mM sodium phosphate (pH 7.8), 1 mM EDTA, 0.02% sodium azide. Residual ammonium sulfate was removed by dialysis for 24 hours against 2 liters of the same buffer with 3 changes of the wash. The final concentration of the antibodies was adjusted to 1.5 – 2 mg/ml, dispensed in 500 μl aliquots and stored at –20° C.

CHAPTER 3

CLONING AND CHARACTERIZATION OF MOSQUITO HNF-4 ISOFORMS

INTRODUCTION

As described previously in Chapter 1, the insect fat body is the functional equivalent to the vertebrate liver (see page 18). Since conservation in the regulation of fat body and liver gene expression has already been demonstrated for some insect and mammalian genes, the goal of the current research is to identify mosquito homologs to mammalian transcription factors and examine their expression in vitellogenesis. Of the four families of liver-enriched transcription factors previously described, HNF-4 displays the greatest conservation in tissue distribution between mammals and insects. In vertebrates, HNF-4 is most highly expressed in the liver, kidney, and intestine. It is important for both embryonic development and adult liver function (Chapter 1, pages 26-29). Likewise in insects, HNF-4 is most highly expressed in the analogous tissues: the fat body, Malphigian tubules, and the midgut. Additionally, HNF-4 is also necessary for insect development and is expressed in the fat body of both larvae and adults (Chapter 1, pages 32-35). It is with these parallels in mind that it was decided to clone a mosquito homolog to HNF-4 and evaluate its possible role in regulating fat body gene expression in the adult mosquito.

Much of the work described in this chapter has been previously published in which Dr. Marianna Kapitskaya and myself were co-first authors (Kapitskaya *et al.*, 1998). Work described here from that paper is noted in the appropriate figure legends. My major contributions were the *in vitro* expression and EMSA analysis of the mosquito HNF-4 proteins (figures 3.7 - 3.9). Dr. Kapitskaya cloned the

cDNAs and performed the Northern blot analysis (figures 3.1 - 3.6). The expression of mosquito HNF-4 in the human 293T cell line was performed in collaboration with Drs. Todd Leff and David Taylor at Parke Davis Pharmaceuticals (Ann Arbor, MI) (figure 3.11). New results obtain since the publication of the earlier work and described here are the expression of mosquito HNF-4 in the *Drosophila* S2 cell line, the generation of HNF-4 antibodies, and the EMSA experiments with fat body nuclear extracts (figures 3.10, 3.12 - 3.14).

RESULTS

Cloning of cDNAs for mosquito HNF-4 isoforms.

Putative clones encoding the mosquito HNF-4 transcription factor (AaHNF-4) were obtained by a combination of the polymerase chain reaction (PCR) and cDNA library screening. The following strategy was used to design degenerate primers: the sense primer conformed to the P-box sequence, DGCKGFF, of the DNA-binding domain, which is conserved among the members of the nuclear receptor superfamily; the antisense primer was based on the conserved sequence LDDQVA from the dimerization domain of the rat, mouse and *Drosophila* HNF-4 homologs (Sladek *et al.*, 1990; Zhong *et al.*, 1993; Hata *et al.*, 1995). A 440-bp fragment was amplified by PCR and sequenced to confirm that its sequence was similar to that of rodent and *Drosophila* HNF-4s. It was used as a probe to screen a mosquito fat body cDNA library.

Restriction mapping and sequence analysis of positive cDNA clones revealed three distinct mosquito cDNAs present in the library, which were designated as AaHNF-4a, -4b, and -4c (figure 3.1). The AaHNF-4b (1.86 kb) and -4c (1.53 kb) cDNAs share the same sequences at the 5'-terminus (except that the 5'-end of AaHNF-4c is 23 bp longer), whereas the AaHNF-4a (2.43 kb) cDNA has a different 5'-terminal sequence. All three AaHNF-4 cDNAs have the same 3'-end sequences and lack a poly (A) tail, indicating that none of the cDNAs represent a full-length transcript. However, AaHNF-4b has a canonical polyadenylation signal (AATAAA) followed by an additional 14 bp, while the AaHNF-4a and AaHNF-4c cDNA sequences end just before the position of the AATAAA-site in AaHNF-4b. In all three cDNAs, the putative start codons (ATG) are preceded by several in-frame stop-codons, indicating that the open reading frame is full-length in each clone.

The conceptual translation of the AaHNF-4a, -4b, -4c cDNAs show that they encode three different proteins, or isoforms, of 565, 538 and 427 amino acids, respectively (figure 3.2). A sequence comparison of the AaHNF-4a and -4b isoforms show that they encode nearly identical polypeptides, differing only in the amino acids at their N-termini (figures 3.2 & 3.3A), strikingly similar to the *Bombyx mori* isoforms BmHNF-4a and BmHNF-4b (Swevers and Iatrou, 1998). In contrast, while the first six amino acids of the shortest isoform, AaHNF-4c, are identical to those of AaHNF-4b, the rest of the A/B domain and the entire C (DNA-binding) domain are lacking (figures 3.2 & 3.3A).

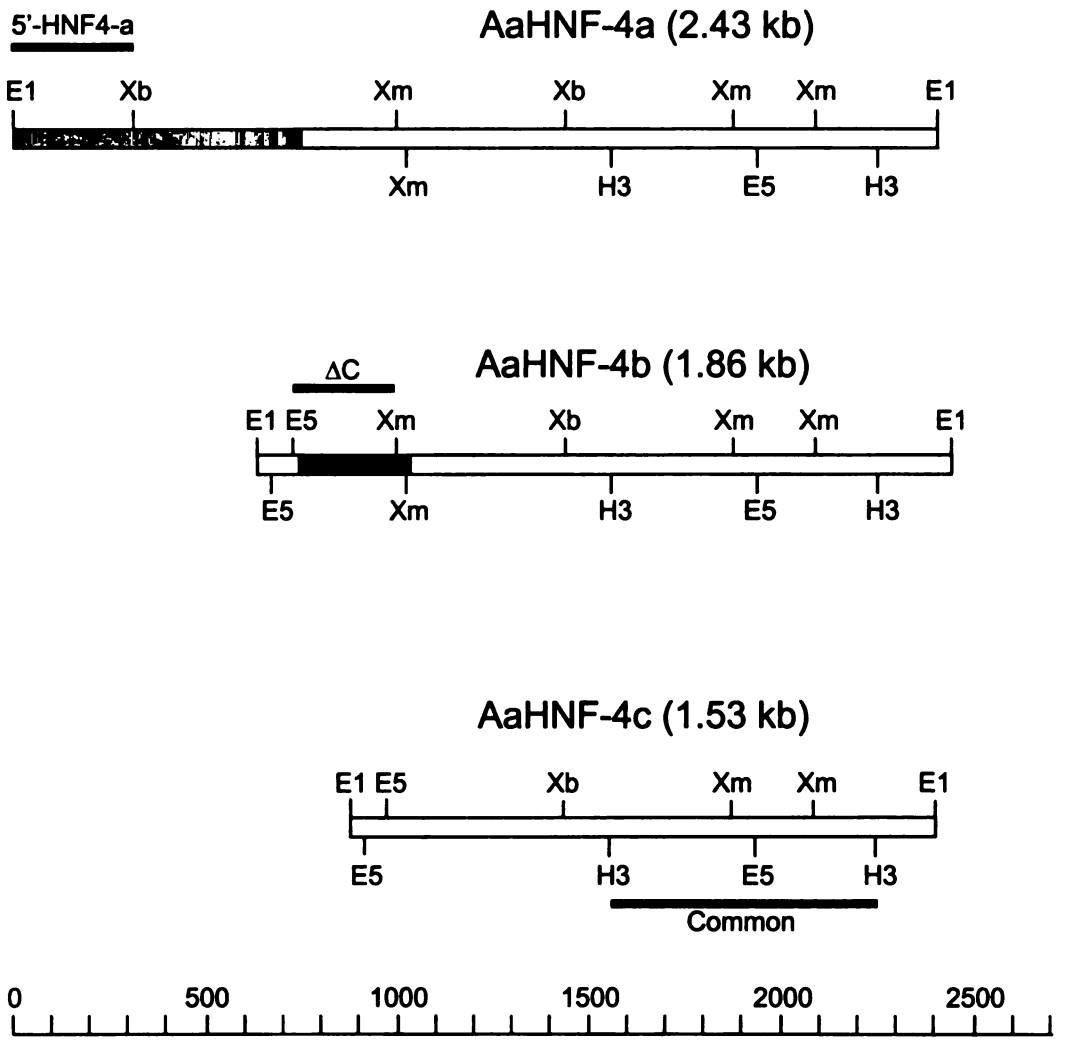


Figure 3.1, Restriction maps of AaHNF-4 cDNA clones. The identity and size of each cDNA is given above each map. Restriction enzyme abbreviations: E1, *Eco RI*; E5, *Eco RV*; H3, *Hind III*; Xb, *Xba I*; Xm, *Xmn I*. The sequence unique to the AaHNF-4a cDNA is shaded in gray. The portion of the AaHNF-4b cDNA that is spliced out of the AaHNF-4c cDNA is shaded in black. Probes used for Northern blot analysis: 5'-HNF-4a, 300 bp *Eco RI/Xba I* fragment unique to the AaHNF-4a cDNA; ΔC, 319 bp *Eco RV/Xmn I* fragment containing a region common to the AaHNF-4a and -4b cDNAs but missing in the AaHNF-4c cDNA; Common, 700 bp *Hind III* fragment common to all three cDNAs. The scale below the restriction maps is in base pairs. Modified from Kapitskaya *et al.* (1998).

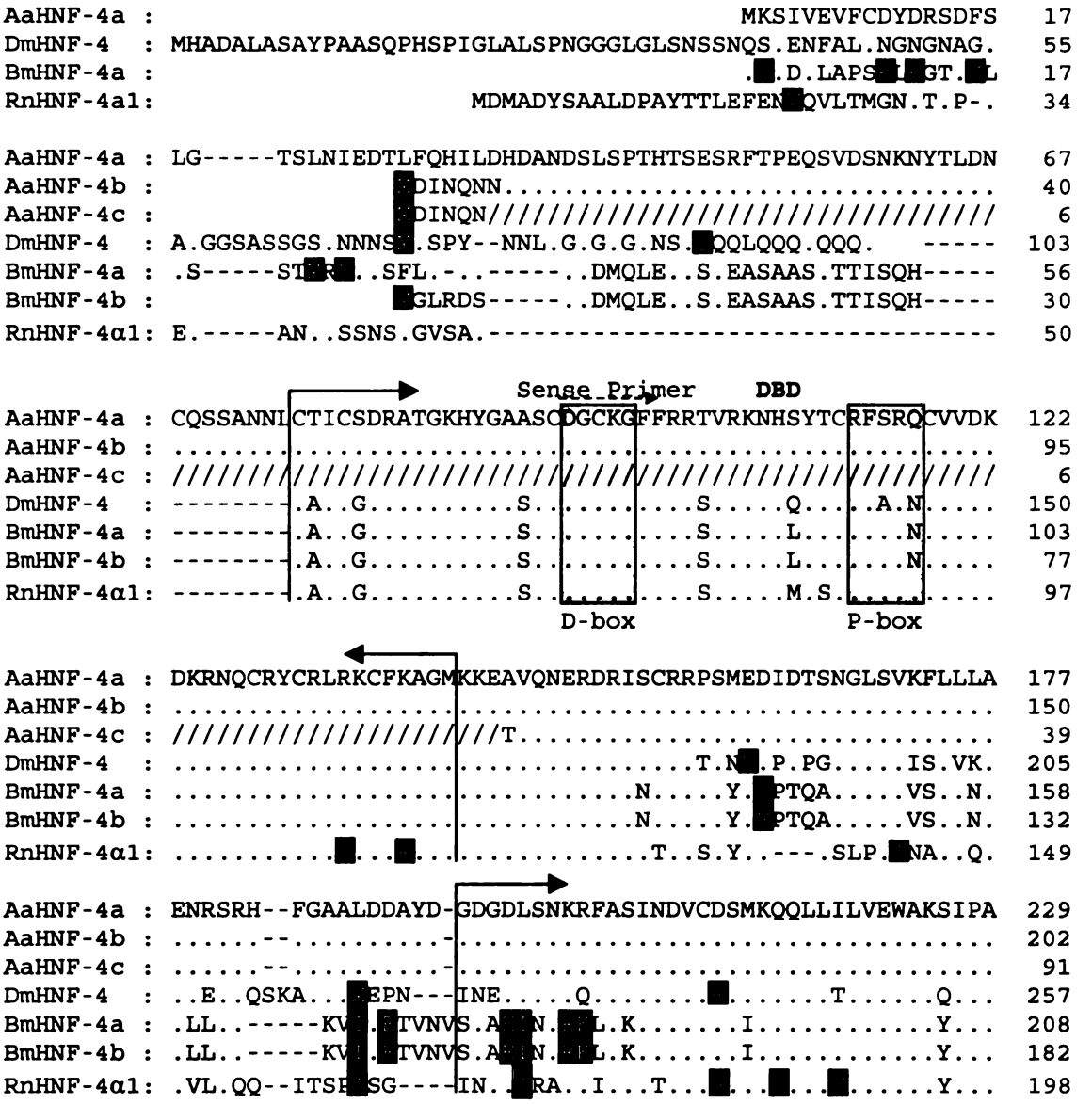


Figure 3.2, Alignment of insect and rat HNF-4 sequences. Amino acid sequences were aligned with AaHNF-4a by the computer program GAP (Wisconsin Package, Genetics Computer Group); all sequences are shown in comparison with AaHNF-4a. Aa, *Aedes aegypti*; Dm, *Drosophila melanogaster* (Zhong *et al.*, 1993); Bm, *Bombyx mori* (Swevers and Iatrou, 1998); Rn, *Rattus norvegicus* (Sladek *et al.*, 1990). Dots indicate identical amino acid with AaHNF-4a. Amino acids boxed in black and in reverse phase type are conservative substitutions. Dashes are gaps introduced to optimize alignment. Forward slants indicate region of AaHNF-4c spliced out. Asterisks denote end of protein. The DNA-binding domain (DBD) and ligand-binding domain (LBD) are delineated by bent arrows. The D-box and P-box of the DBD are boxed in gray. The sequences corresponding to the sense and antisense primers used for PCR are marked with a dashed arrow. Numbers to the right represent amino acid numbers. Modified from Kapitskaya *et al.* (1998).

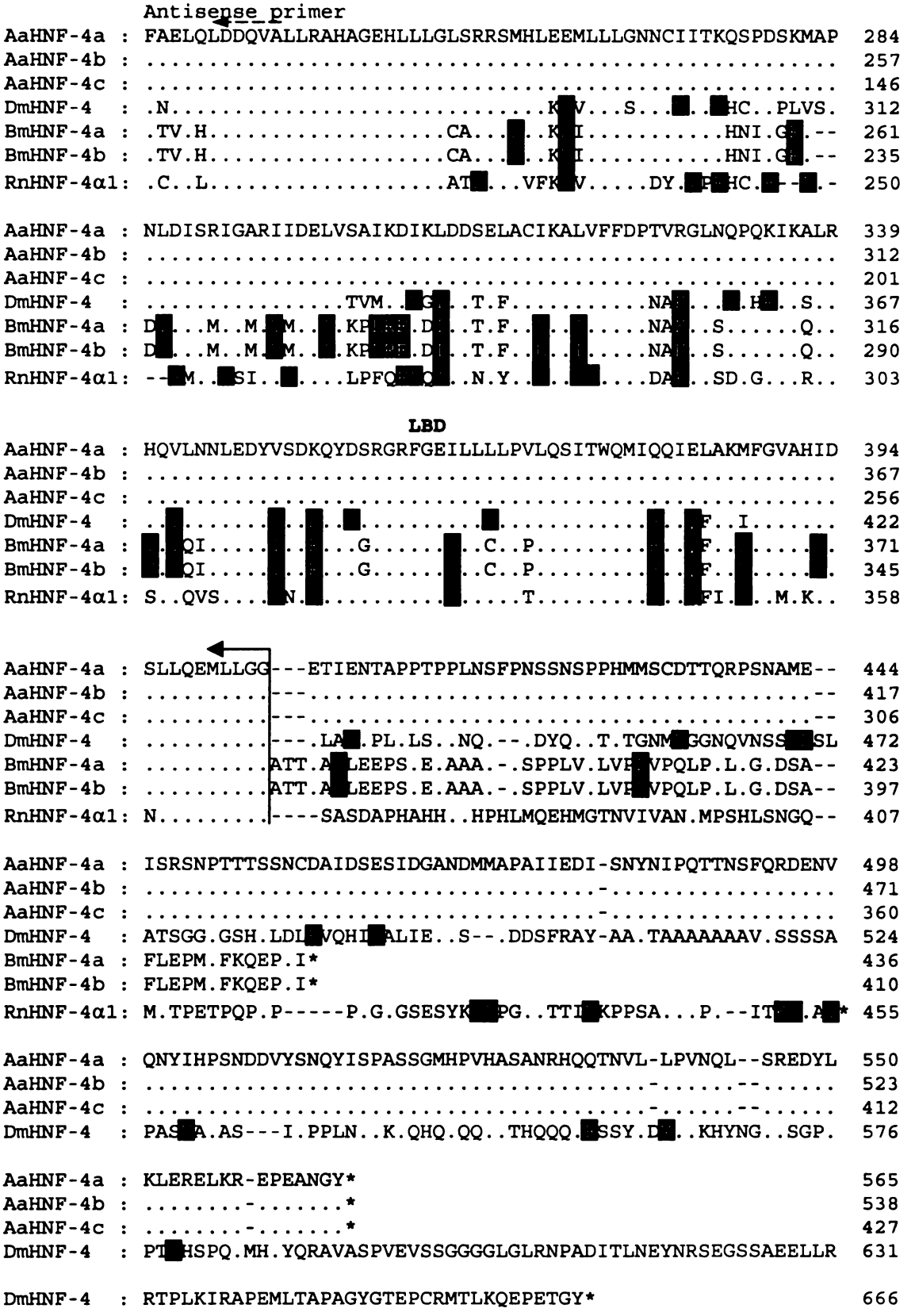


Figure 3.2 continued.

Figure 3.3, Domain comparison of insect and rat HNF-4 homologs.
A) Schematic representation of the mosquito HNF-4 isoforms. The identity of each domain is indicated at the top of the figure. The differences in the N-terminus of the AaHNF-4a and -4b isoforms is indicated by the stippled and striped boxes. The portion of the A/B and C domains that are missing in AaHNF-4c is indicated by a dotted line. B) Percent amino acid identity between HNF-4 homologs. The numbers within each domain represent the percent identity and similarity (in parenthesis) with the corresponding domain in AaHNF-4a. The numbers below each protein indicate the number of amino acids in that domain. The references for DmHNF-4, BmHNF-4a, and RnHNF-4α1 are given in figure 3.2. Modified from Kapitskaya *et al.* (1998).

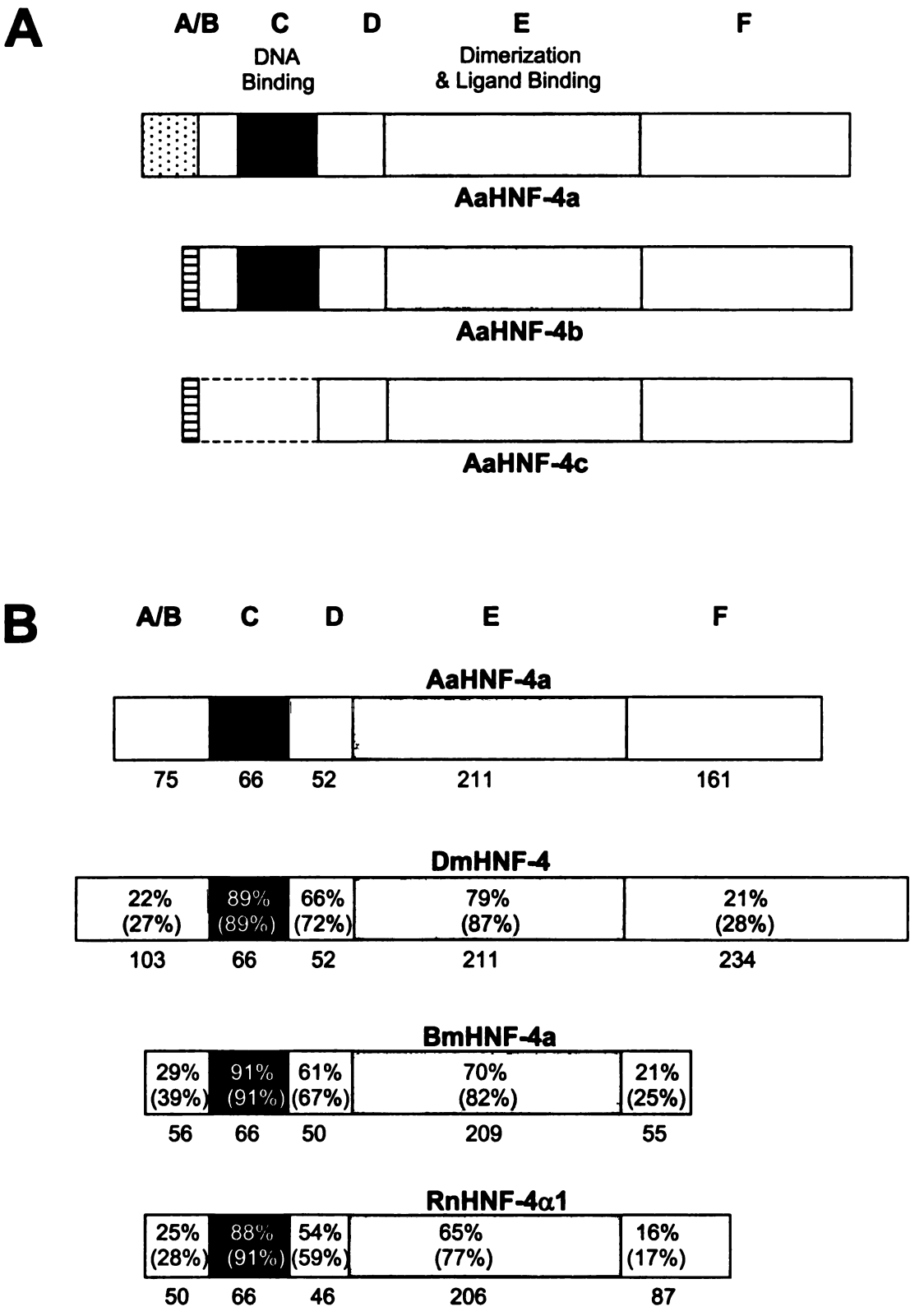


Figure 3.3, Domain comparison of insect and rat HNF-4 homologs.

Thus, the structure of the mosquito AaHNF-4c isoform is unique among insect and vertebrate members of the HNF-4 subfamily that have been reported to date.

The mosquito AaHNF-4a and -4b isoforms exhibit a structural domain organization similar to that of vertebrate and *Drosophila* HNF-4 homologs (figure 3.3B). The C (DNA-binding), D (hinge), and E (dimerization/ligand-binding) domains are highly conserved. For example, AaHNF-4a and -4b share 89%, 65% and 79% identity with the respective domains of *Drosophila* HNF-4, and 91%, 58% and 72% identity with *B. mori* (figure 3.3B). The N-terminal A/B and C-terminal F-domains are poorly conserved in the mosquito AaHNF-4 proteins, a characteristic feature of all members of the HNF-4-subfamily and more generally, the nuclear receptor superfamily (Mangelsdorf *et al.*, 1995).

Analysis of AaHNF-4 transcripts

To determine if the three AaHNF-4 cDNAs cloned represent respective AaHNF-4 mRNA species in the mosquito, Northern hybridizations were performed using particular restriction fragments of the cDNAs as probes. A probe common to all three cDNAs (Common, figure 3.1) detected three transcripts of 2.8, 2.1, and 1.8 kb respectively. A probe specific to the 5'-untranslated region of the AaHNF-4a cDNA (5'-HNF-4a, figure 3.1) hybridized only with the 2.8 kb transcript. A third probe, containing sequences from the A/B and C domains common to the AaHNF-4a and -4b cDNAs but not to AaHNF-4c (ΔC, figure 3.1), recognized the 2.8 and 2.1 kb mRNAs but not the 1.8-kb mRNA. Thus, it was possible to define the 2.8, 2.1, and 1.8 kb mRNA species as

AaHNF-4a, -4b, and -4c transcripts, respectively (figure 3.4). This experiment verifies that one of the transcripts indeed lacks the region encoding for a portion of the A/B and the entire DNA-binding domains, and therefore, AaHNF-4c does not appear to be an artifact of the cDNA library.

The distribution of AaHNF-4 expression in tissues of the adult mosquito.

To determine the distribution of AaHNF-4 expression in tissues of adult mosquitoes, polyadenylated mRNA was isolated from the whole body of males and the fat body, midgut, Malpighian tubules, ovary, and thorax of females, taken at 24 hours post blood meal (PBM) and subjected to Northern blot analysis. A fragment of the cDNA common to all three AaHNF-4 cDNAs was used as a hybridization probe. The mosquito fat body, Malpighian tubules and midgut exhibited significant levels of AaHNF-4 transcription. Two mosquito HNF-4 isoforms (AaHNF-4a and -4b) were detected in these tissues (data not shown). In contrast to the somatic tissues of mosquito females, only the 2.1-kb (AaHNF-4b) transcript was found in the ovaries (figure 3.5). The highest level of ovarian AaHNF-4b mRNA was detected soon after initiation of vitellogenesis by a blood meal (6 hours PBM). It then dropped to an undetectable level by 36 hours PBM, the time of termination of the synthesis phase in the female mosquito (figure 3.5).

Stage-specific expression of AaHNF-4 transcripts in the fat body.

The high level of conservation between the mosquito HNF-4 and its vertebrate counterparts (figures 3.2 and 3.3), as well as the presence of three

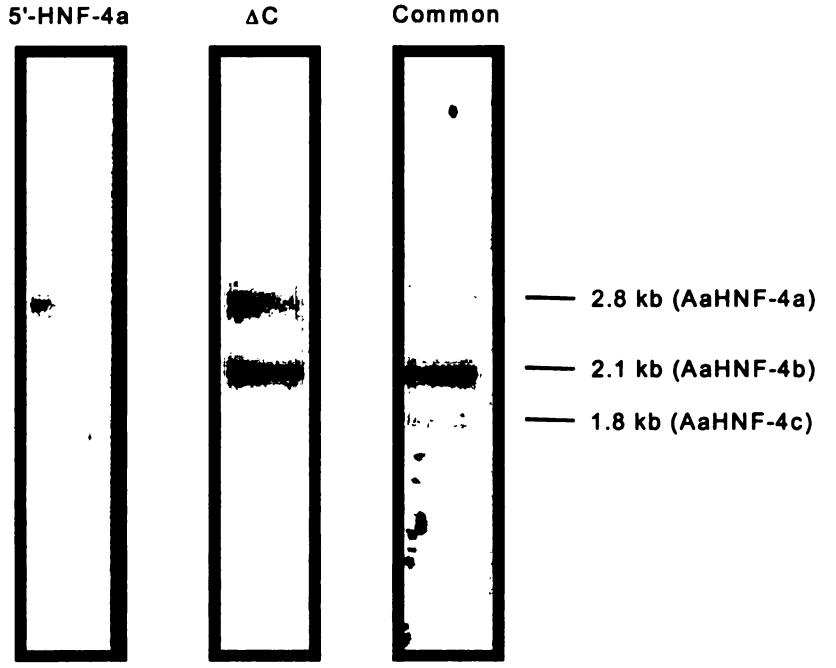


Figure 3.4, Northern hybridization analysis of AaHNF-4 transcripts. Polyadenylated mRNA was isolated from the fat body of female mosquitoes at 26 hours PBM. the mRNA was fractionated by electrophoresis in a 1% agarose/formaldehyde gel. After transfer to a nitrocellulose membrane, the blots were probed with various AaHNF-4 cDNA fragments (noted above each blot). See figure 3.1 for the description of the various probes. From Kapitskaya *et al.* (1998).

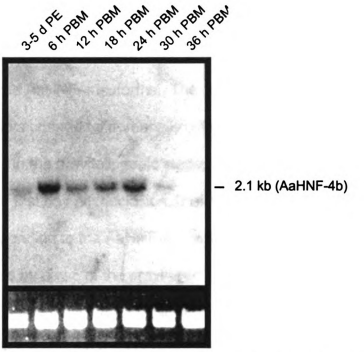


Figure 3.5, Expression of AaHNF-4 in the ovary during the first vitellogenic cycle. Total RNA was extracted from ovaries dissected at the time points indicated: d PE, days post eclosion; h PBM, hours post blood meal. The RNA was fractionated and transferred to a nitrocellulose membrane as described in figure 3.4. A cDNA fragment common to all three isoforms (see figure 3.1 for description) was used as a probe. The bottom panel shows a gel run in duplicate and stained with ethidium bromide to verify equal loading.

AaHNF-4 transcripts in the adult fat body (figures 3.4 and 3.5), suggests that it plays an important role in this metabolic tissue, similarly to that of the vertebrate HNF-4 in the liver. In an attempt to further understand the possible functions of AaHNF-4 in the vitellogenic fat body of adult mosquitoes, its expression was characterized in more detail.

Total RNA from the fat body of female mosquitoes at different stages of vitellogenesis was analyzed by Northern blot hybridization using the cDNA-probe common to all three AaHNF-4 isoforms. The expression pattern of the three AaHNF-4 transcripts changed differentially over the course of the vitellogenic cycle (figure 3.6). In the previtellogenic stage, all three transcripts were clearly present during the first day post eclosion, but in subsequent days only the 2.1-kb transcript, corresponding to the AaHNF-4b isoform, was present. Furthermore, its levels increased by day 5 of the previtellogenic period. During the first 18 hours after a blood meal, only the AaHNF-4b transcript was clearly detectable. Near 24 hours PBM, when the transcription of the yolk protein (YP) genes are nearing their peak (Raikhel, 1992, and references within), the levels of the AaHNF-4a transcript began to rise again, relative to AaHNF-4b. By 36 hours PBM, the levels of HNF-4 mRNA significantly increased, with AaHNF-4a becoming the predominant transcript. This pattern of AaHNF-4 expression was maintained until 48 hours PBM when YP synthesis terminated and the fat body returned to its previtellogenic state (Raikhel, 1992, and references within).

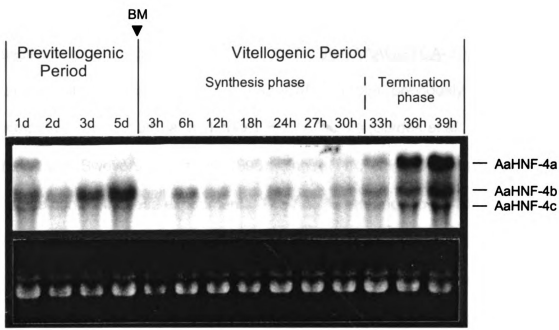


Figure 3.6, Expression of AaHNF-4 in the fat body during the first vitellogenic cycle. Total RNA was extracted from fat bodies dissected at the time points indicated: d, days post eclosion; BM, blood meal; h, hours post blood meal. The RNA was fractionated and transferred to a nitrocellulose membrane as described in figure 3.4. A cDNA fragment common to all three isoforms (see figure 3.1 for description) was used as a probe. The bottom panel shows a gel run in duplicate and stained with ethidium bromide to verify equal loading. From Kapitskaya *et al.* (1998).

DNA binding by AaHNF-4 proteins

Electrophoretic mobility shift assays (EMSAs) were used to determine if the mosquito HNF-4 homologs are functional DNA-binding proteins. The three AaHNF-4 cDNAs were subcloned into transcription vectors, then subjected to coupled transcription/translation (TNT) reactions in a rabbit reticulocyte lysate system. The *in vitro* expressed proteins were examined for their ability to form specific binding complexes with APF-1 (5'-TGGGCAAAGGTCA-3"), a response element from the *apolipoprotein CIII* (*apoCIII*) promoter, previously shown to be a recognition site for both rat and insect HNF-4 factors (Sladek *et al.*, 1990; Zhong *et al.*, 1993; Swevers and Iatrou, 1998). The rat HNF-4α1 cDNA was used as a positive control in these experiments. The AaHNF-4a and -4b proteins formed DNA/protein complexes of similar mobility, which migrated slower than the complex formed by the rat HNF–4α1 (figure 3.7). The binding of these proteins was sequence-specific and could be competed away with an excess of unlabeled APF-1 probe, but not with a nonspecific competitor.

The protein synthesized from the AaHNF-4c cDNA, which is missing the DNA-binding domain, failed to exhibit any binding activity (data not shown). The appearance of its transcript in the fat body during the termination phase of vitellogenesis suggests that it may serve as a repressor by forming inactive heterodimers with other AaHNF-4 isoforms. Utilizing EMSAs, I examined the ability of AaHNF-4c to prevent the other AaHNF-4 isoforms from binding to an HNF-4 response element. To allow heterodimer formation between AaHNF-4c and the other HNF-4 isoforms, the appropriate TNT samples were first mixed

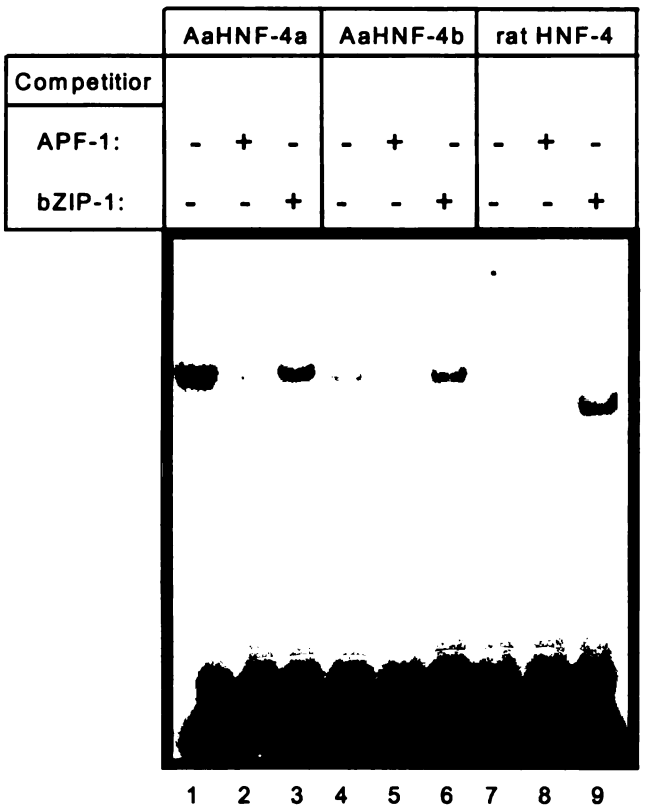


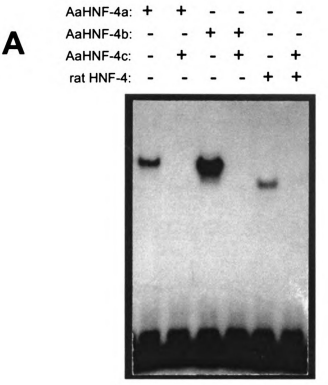
Figure 3.7, Electrophoretic mobility shift assay (EMSA) of HNF-4 isoforms. HNF-4 proteins were synthesized *in vitro* as described in Materials and Methods The HNF-4 binding site in the promoter of the *apolipoprotein C III* gene (APF-1: TGGGCAAAGGTCA, Sladek *et al.*, 1990) was used as a probe. The HNF-4 isoforms are noted above each lane. Competitor DNA was added at a100-fold molar excess to the labeled probe. Unlabeled APF-1 was used as a specific competitor; bZIP-1: ATTTTGCAAT (Dittmer and Raikhel, 1997) binds the CCAAT/enhancer-binding protein and was used as a nonspecific competitor. From Kapitskaya *et al.* (1998).

together and heated, then allowed to cool prior to adding the radiolabeled DNA probe. Incubations of each HNF-4 isoform in the absence of AaHNF-4c showed that heat treatment caused little or no reduction in specific binding (figure 3.8A, lanes 1, 3, and 5). The addition of AaHNF-4c to AaHNF-4a, -4b, and the rat HNF-4α1, prevented the formation of binding complexes with the probe (figure 3.8A, lanes 2, 4, and 6). This inhibition by AaHNF-4c was dose-dependent and specific (figure 3.8B, lanes 2 – 5). In a control experiment, addition of increasing amounts of an unrelated transcription factor (*Drosophila* GATAb) did not interfere with the ability of AaHNF-4b to form the retardation complex (figure 3.8B, lanes 6 – 8). Thus, EMSA analyses suggest that the AaHNF-4c isoform may serve as a transcriptional repressor in the mosquito fat body by forming inactive heterodimers with other AaHNF-4 isoforms.

Identification of HNF-4 binding sites in regulatory regions of fat body-specific YP genes.

In *A. aegypti, vitellogenin* (*Aavg*) (Romans *et al.*, 1995) and *vitellogenic carboxypeptidase* (*Aavcp*) (Deitsch and Raikhel, 1993) are the two major YP genes that are expressed in the fat body of blood-fed females (Raikhel, 1992). Examination of the 5'-regions of both the *Aavg* and *Aavcp* genes revealed the presence of putative HNF-4 binding sites, VgHNF4 and VCPHNF4, upstream of their transcription start sites, respectively (figure. 3.9A). Each sequence was examined by EMSA for the ability to compete with the radiolabeled APF-1 sequence for binding by AaHNF-4b. Both VCPHNF4 and VgHNF4 exhibited

Figure 3.8, AaHNF-4c inhibits DNA binding of other HNF-4 isoforms. EMSA was used to examine the ability of AaHNF-4c to abolish the DNA binding ability of other HNF-4 isoforms via the formation of heterodimers. All proteins were synthesized in separate *in vivo* reactions. The proteins present in each reaction are noted above each lane. The samples were heated at 50° C for 1 minute to allow dimers to break and reform. The APF-1 element of the *apoCIII* gene was used as the probe. A) EMSA showing that the addition of AaHNF-4c abolishes the DNA binding ability of AaHNF-4a, -4b, and rat HNF-4. B) Titration of the effect of AaHNF-4c. A constant amount of AaHNF-4b was mixed with increasing amounts of either AaHNF-4c (lanes 3-5) or DmGATAb (lanes 6-8). Total volume was kept constant with unprogrammed rabbit reticulocyte lysate). DmGATAb, originally named Box A binding factor (ABF; Abel *et al.*, 1993), was used as a nonspecific competitor for dimer formation. To independently demonstrate the DNA binding ability of DmGATAb, the Box A binding site (Abel *et al.*, 1993) was used as a probe (lane 9). From Kapitskaya *et al.* (1998).



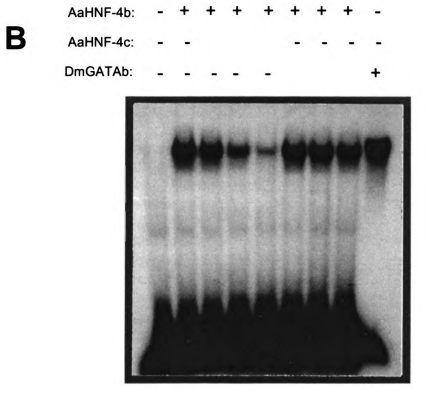


Figure 3.8, AaHNF-4c inhibits DNA binding of other HNF-4 isoforms.

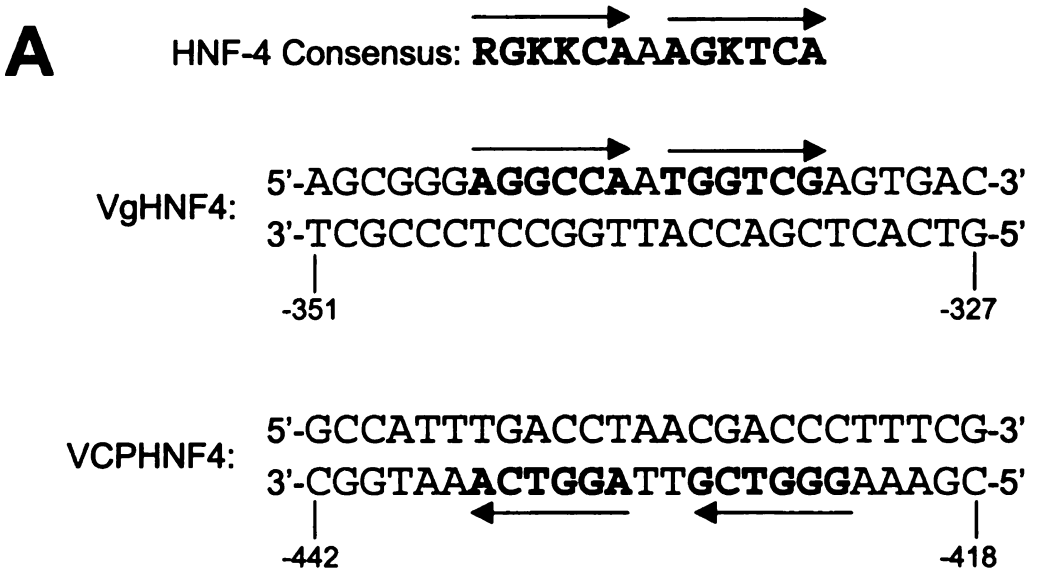
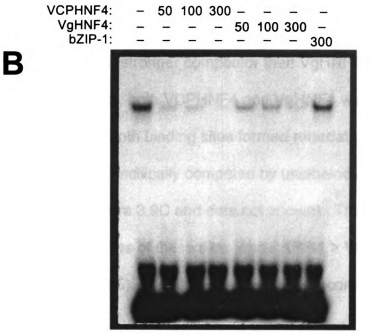


Figure 3.9, AaHNF-4b binds *in vitro* to sequences upstream of the *Aavg* and *Aavcp* genes. A) Putative HNF-4 binding sites in the *Aavg* and *Aavcp* genes. The HNF-4 consensus sequence is from Sladek (1994). Nucleotides in bold type represent sequences similar to the HNF-4 half-sites. The arrows denote the orientation of the half-sites. Numbers below the sequences indicate the distance from the transcription start sites. K, G or T; R, A or G. B) Competition EMSA with VCPHNF4 and VgHNF4. AaHNF-4b was synthesized *in vitro* for binding with the labeled APF-1 probe. Specific competition was performed by adding increasing amounts of unlabeled VCPHNF4 or VgHNF4. Nonspecific competition was with unlabeled bZIP-1 (see figure 3.7 for sequence). Numbers above each lane indicate amount of competitor DNA, expressed as molar excess to the labeled probe. C) EMSA for direct binding to the VCPHNF4 sequence. The VCPHNF4 sequence shown in A) was labeled and tested for binding to AaHNF-4b. Binding complexes were examined by competition with unlabeled APF-1, VCPHNF4, VgHNF4, and bZIP-1 sequences, expressed as molar excess to the labeled probe. From Kapitskaya et al. (1998).



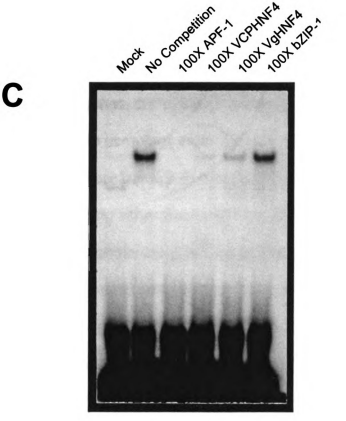


Figure 3.9 continued.

significant specific competition, though binding was weaker than that of APF-1; interestingly, VCPHNF4 was a stronger competitor than VgHNF4 (figure 3.9B). In a second set of experiments, both VCPHNF4 and VgHNF4 were tested for direct binding to AaHNF-4b. Both binding sites formed retardation complexes with AaHNF-4b, which were specifically competed by unlabeled APF-1, VCPHNF4, and VgHNF-4 (figure 3.9C and data not shown). These tests also showed that the binding affinities of these sites were: APF-1 > VCPHNF4 > VgHNF4 (figure 3.9C, lanes 3-5). Addition of a nonspecific competitor had no effect on the retardation complexes (figure 3.9C, lane 6).

AaHNF-4a and -4b are transcriptional activators.

Having established that both AaHNF-4a and -4b are functional DNA-binding proteins, and that putative HNF-4 binding sites are present in the 5' regulatory regions of the *Aavg* and *Aavcp* genes, cell culture transfection assays were utilized to examine if the mosquito homologs could act as transcriptional activators. *Drosophila* S2 cells were transiently transfected with various amounts of an expression plasmid containing either the AaHNF-4a or -4b cDNA. A bacterial chloramphenicol acetyltransferase (CAT) gene fused to the *Drosophila alcohol dehydrogenase* (*Adh*) promoter was used as the reporter; the *Adh*CAT construct contained 3 copies of a perfect DR1 cloned upstream of the *Adh* promoter (pDR1(3X)*Adh*CAT). Both AaHNF-4a and -4b were able to significantly enhance transcription of pDR1(3X)*Adh*CAT above basal levels (figure 3.10A). This activation could be repressed by cotransfection with an expression vector

Figure 3.10, Transient transfection of AaHNF-4 isoforms in *Drosophila* S2 cells. The total amount of DNA transfected was kept constant at 650 ng using empty vector. CAT values shown are with respect to the basal level of the reporter vector (pDR1(3X)*Adh*CAT), set to a value of 1, and normalized to a β-galactosidase coreporter (pAc5/V5-His/*lacZ*) used to monitor transfection efficiency. All data points represent the mean from two independent experiments each with three replicates, ± the standard error. A) AaHNF-4a and -4b are transcriptional activators. *Drosophila* S2 cells were transfected with selected amounts of expression vectors containing the cDNAs for either the AaHNF-4a or -4b isoforms. B) AaHNF-4c represses activation by AaHNF-4a and -4b. Cells were cotransfected with either 100 ng of AaHNF-4a or -4b, and various amounts of AaHNF-4c.

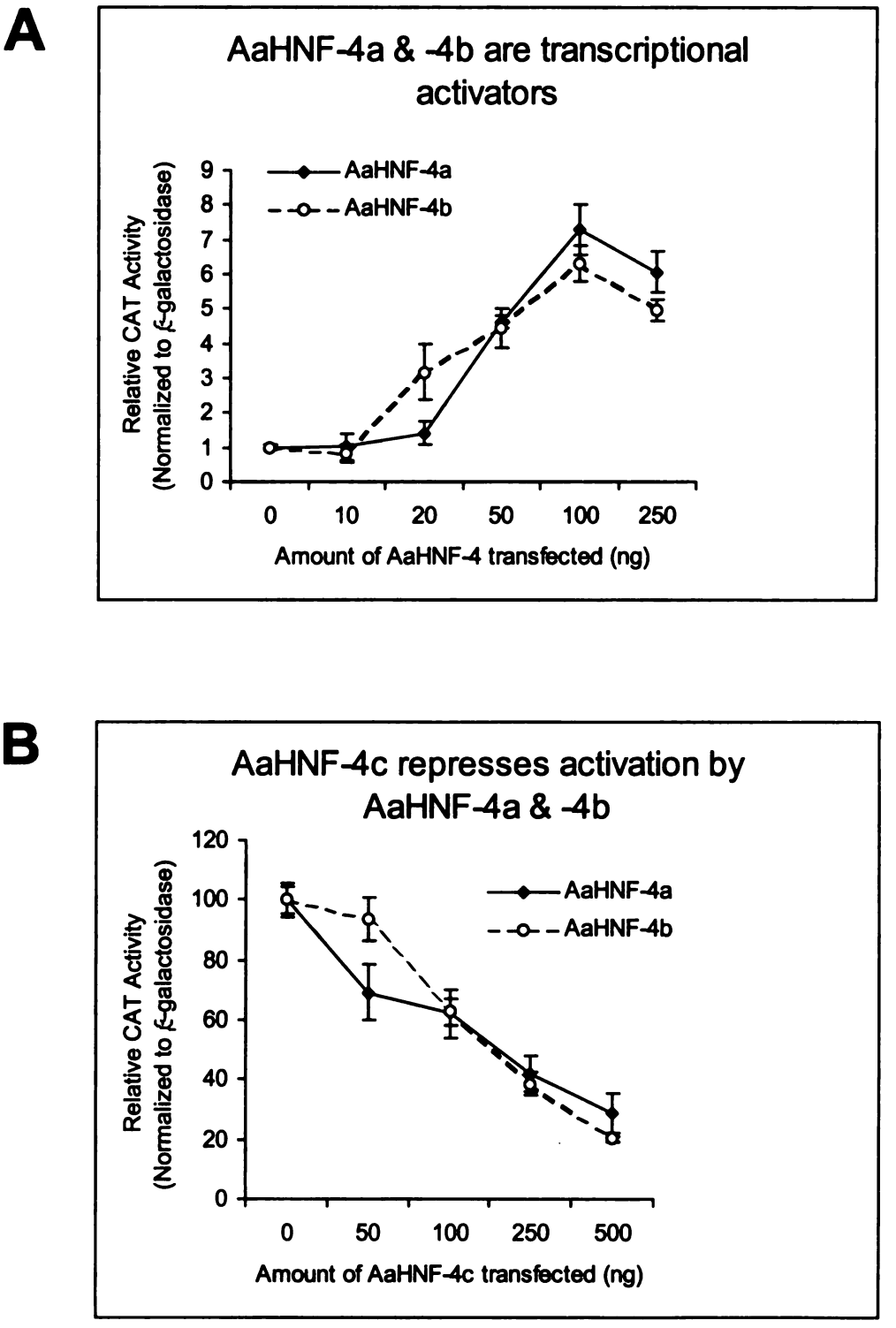
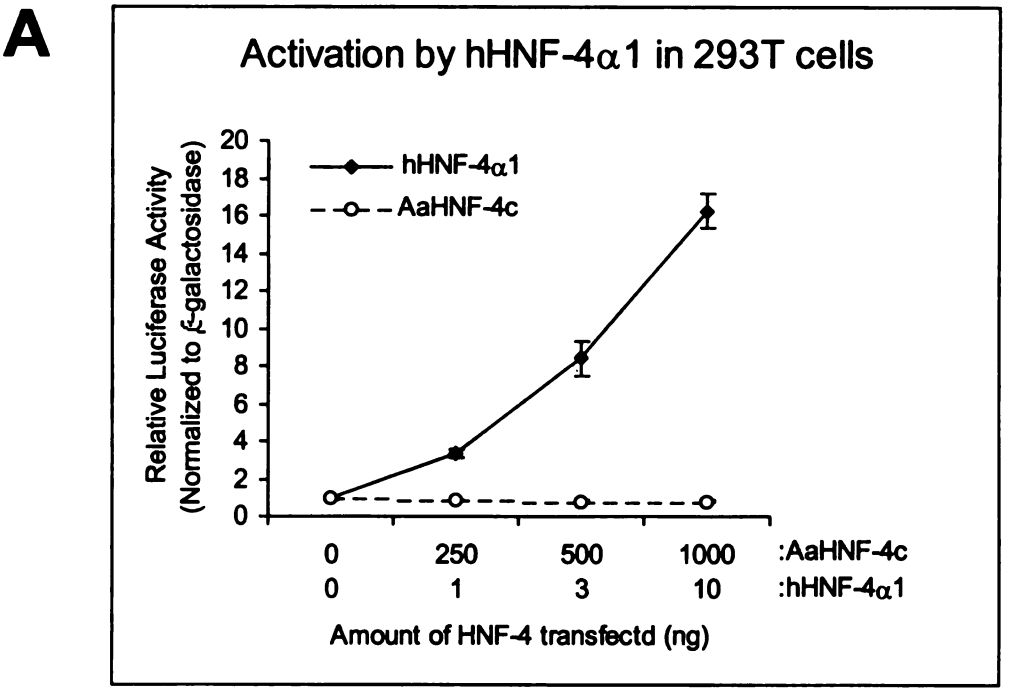


Figure 3.10, Transient transfection of AaHNF-4 isoforms in *Drosophila* S2 cells.

containing the AaHNF-4c cDNA (figure 3.10B). These results support earlier work performed in the human embryonic kidney 293T cell line (Kapitskaya *et al.*, 1998), indicating that this repression was genuine and not due to the sequestering of cofactors caused by the over expression of the HNF-4 isoforms.

The reporter construct used for these experiments consisted of the *luciferase* gene fused to the promoter of the human *apoCIII* gene (p*apoCIII*-Luc), which contains the HNF-4 response element APF-1 (Taylor *et al.*, 1996). In pilot experiments, AaHNF-4a and -4b failed to activate or repress the human *apoCIII* promoter in the 293Tcell line (data not shown); this is most likely due to the incompatibility of the mosquito HNF-4 isoforms with mammalian cofactors. Thus, repression of HNF-4 -mediated transcriptional activation of the p*apoCIII*-Luc reporter by AaHNF-4c in 293T cells could be attributed to the formation of human HNF-4/AaHNF-4c heterodimers incapable of binding to DNA. Therefore, I tested the effect of the AaHNF-4c isoform on the transcriptional activity of the human hHNF-4α1 protein. In the 293T cells, the hHNF-4α1 caused a dose-dependent activation of p*apoCIII*-Luc, while transfected AaHNF-4c elicited no change in activity (figure 3.11A). When the human hHNF-4α1 and mosquito AaHNF-4c were cotransfected, *apoCIII* promoter activity decreased in a dose-dependent manner (figure 3.11B). Thus, AaHNF-4c inhibits transcriptional activation by hHNF-4α1 most likely by forming heterodimers incapable of DNA binding, and thereby causing a dominant-negative effect (Taraviras *et al.*, 1997).

Figure 3.11, Transient transfection of HNF-4 homologs in 293T cells. The total amount of DNA transfected was kept constant at 5 μg using empty vector. Luciferase values shown are with respect to the basal level of the reporter vector (p*apoCIII*-Luc), set to a value of 1, and normalized to a β-galactosidase coreporter (pCMVβ) used to monitor transfection efficiency. Each data point represents the mean of three replicates ± the standard error. A) Human embryonic kidney 293T cells were transfected with selected amounts of expression vectors containing the cDNAs for either the human hHNF-4α1 or mosquito AaHNF-4c isoforms. B) AaHNF-4c inhibits transcriptional activation by hHNF-4α1. Cells were cotransfected with 1 ng of hHNF-4α1 and various amounts of AaHNF-4c. From Kapitskaya *et al.* (1998).



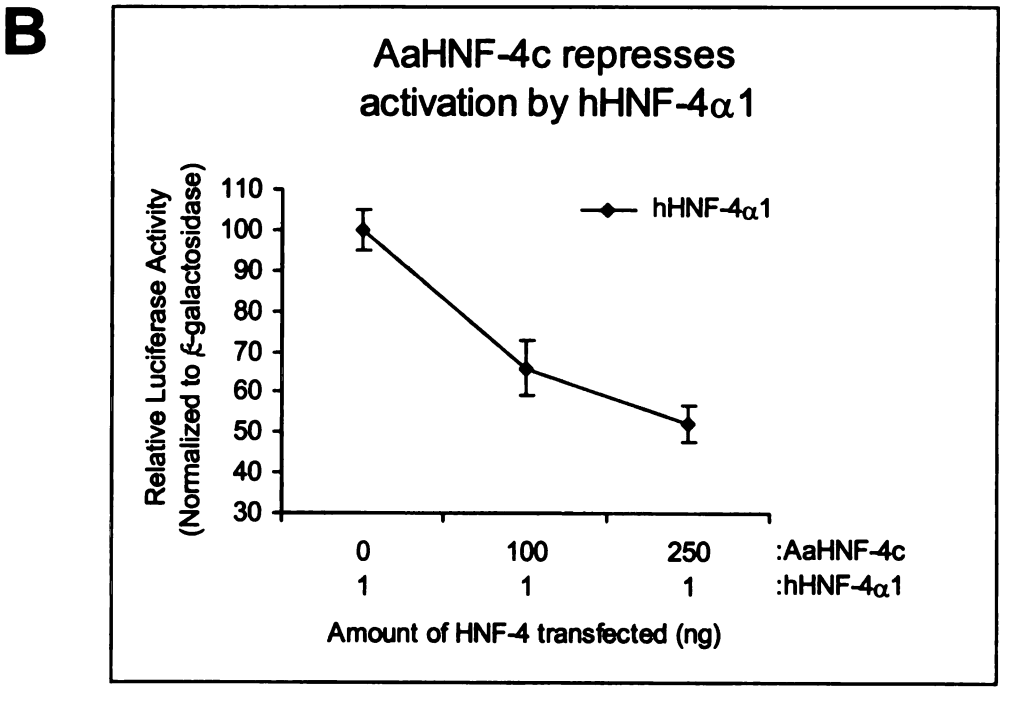


Figure 3.11, Transient transfection of HNF-4 homologs in 293T cells.

HNF-4 binding activity in mosquito fat body nuclear extracts.

Because many nuclear hormone receptors (NR) are able to bind to the same response element, it was first necessary to generate AaHNF-4 specific antibodies to distinguish AaHNF-4/DNA complexes from other NR/DNA complexes. A GSTHNF-4 fusion protein was constructed and used as the antigen for polyclonal antibody production. As seen in figure 3.12, this antibody [α-GSTHNF-4(Com)] recognized all three AaHNF-4 isoforms expressed *in vitro* while the preimmune serum did not. The specificity of the α-GSTHNF-4(Com) antibodies was further demonstrated by EMSA. In addition to AaHNF-4, the mosquito ecdysteroid receptor/ultraspiracle complex (AaEcR/AaUSP; Cho *et al.*, 1995; Kapitskaya *et al.*, 1996; S. Wang, A. Raikhel, unpublished results), AaHR38 (Zhu *et al.*, 2000; homolog to the vertebrate nerve growth factor-induced protein B, NGFI-B), and Seven-up (AaSVP; K. Miura, J. Zhu, N Dittmer, A Raikhel, unpublished results; homolog to the vertebrate chicken ovalbumen upstream promoter transcription factor, COUP-TF) all are capable of binding to a DR1 (figure 3.13, lanes 1, 4, 7, and 10). The addition of α-GSTHNF-4(Com) antibodies abolished the binding of AaHNF-4b to the labeled probe, but did not affect the DNA-binding affinity of AaEcR/AaUSP, AaHR38, or AaSVP (figure 3.13, lanes 3, 6, 9, and 12). The addition of the preimmune serum had no affect on any of the NR/DNA complexes (figure 3.13, lanes 2, 5, 8, and 11).

EMSA experiments using fat body nuclear extracts prepared from various time points throughout the vitellogenic cycle was next employed to examine when AaHNF-4 proteins can actively bind DNA. A strong shift was detected with

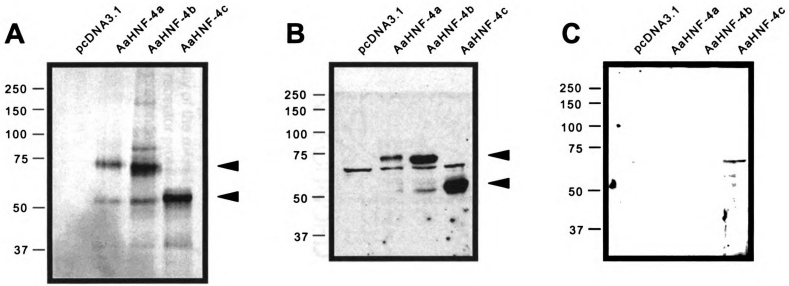


Figure 3.12, Recognition of AaHNF-4 isoforms by the α-GSTHNF-4(Com) antibodies. A) Various AaHNF-4 isoforms, or empty vector, were used for *in vitro* synthesis with ^{35}S-methionine. The translation products were then resolved by electrophoresis in a 10% SDS-polyacrylamide gel and subjected to autoradiography. B) & C) Proteins were synthesized as in A) except without ^{35}S-methionine. Following gel electrophoresis, the proteins were transferred to a nitrocellulose membrane for Western analysis: B) Membrane probed with α-GSTHNF-4(Com) antibodies (1:100 dilution); C) Membrane probed with preimmune serum (1:100 dilution). Molecular weight standards (in kDA) are indicated on the left of each panel. Arrow heads denote positions of specific AaHNF-4 isoforms.

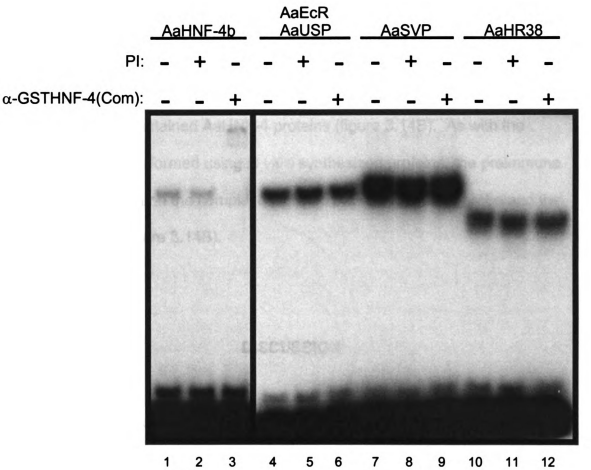


Figure 3.13, Specificity of α-GSTHNF-4(Com) antibodies. EMSA was used to verify the specificity of the α-GSTHNF-4(Com) antibodies. Various mosquito nuclear hormone receptor proteins were synthesized *in vitro* and incubated with a DR1 probe in the presence (+) or absence (-) of preimmune serum (PI) or HNF-4 antibodies.

nuclear extracts from 3-5 days post eclosion and 36 hours PBM, with weak binding detected at 12 hours PBM and moderate binding at 24 hours PBM (figure 3.14A). Since 24 and 36 hours PBM represent the transition from peak YP synthesis to the termination of vitellogenesis, these time points were further examined using the α-GSTHNF-4(Com) antibodies. These antibodies were able to abolish the complexes formed in both samples, indicating that most, if not all, of the complexes contained AaHNF-4 proteins (figure 3.14B). As with the NR/DNA complexes formed using *in vitro* synthesized proteins, the preimmune serum had no affect on the complexes formed between the DR1 probe and the nuclear extracts (figure 3.14B).

DISCUSSION

Comparison of the deduced amino acid sequences of AaHNF-4a and -4b revealed that these isoforms are typical members of the HNF-4 subfamily. They differ only in the N-terminal end of the variable A/B domain, with only the first 33 amino acids of AaHNF-4a and the first 6 amino acids ofAaHNF-4b being the only variation between the two. This is strikingly similar to BmHNF-4a and -4b, which have unique N-terminal sequences of 32 and 6 amino acids respectively (Swevers and Iatrou, 1998). The other domains are identical and show a remarkable level of conservation: the C (DNA-binding), D (hinge), and E (dimerization/ligand-binding) domains share over 90%, 77% and 85% amino acid

Figure 3.14, Active binding by AaHNF-4 proteins from fat body nuclear extracts. A) EMSA was performed with fat body nuclear extracts prepared from the time points indicated (d PE, days post eclosion; h PBM, hours post blood meal) and incubated with a ^{32}P-labeled DR1 as the probe. Ten mosquito fat body equivalents were used per lane. B) Characterization of binding complexes. Nuclear extracts prepared from 24 and 36 hours PBM were further examined by the addition of α-GSTHNF-4(Com) antibodies or preimmune (PI) serum to the binding reaction (1:100 dilution). The presence or absence of antibodies is indicated above each lane.

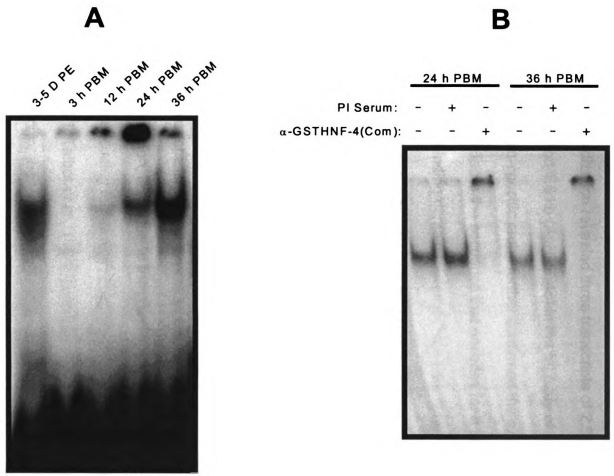


Figure 3.14, Active binding by AaHNF-4 proteins from fat body nuclear extracts.

similarity with the corresponding domains in rat and insect HNF-4s (figure 3.3). The third mosquito isoform, AaHNF-4c, is unique among vertebrate and insect members of the HNF-4 subfamily in that the greater part of the A/B domain and the entire DNA-binding domain are absent; consequently, AaHNF-4c cannot bind DNA. Northern hybridization revealed three transcripts of 2.8 kb (AaHNF-4a), 2.1 kb (AaHNF-4b) and 1.8 kb (AaHNF-4c) in the fat body of adult females. Utilization of a domain-specific probe further confirmed the existence of a transcript that indeed lacks the sequence that encodes for a section of the A/B and the entire DNA-binding domains.

It seems likely that a single gene encodes for the three mosquito isoforms. Comparison of the three mosquito sequences suggests that they are created by usage of unique 5'-exons that are alternatively spliced to a common set of 3'-exons. Amplification of genomic DNA by PCR using primers designed on sequences flanking the C-terminal boundary of the deleted region in AaHNF-4c identified an intron of approximately 3 kb (data not shown). However, the same approach was unsuccessful when applied to the putative splice junction at the N-terminus of the deletion, likely due to the presence of a large intron; a common feature of genes encoding for nuclear receptors (Segraves and Hogness, 1990; Koelle *et al.*, 1991; Talbot *et al.*, 1993; Taraviras *et al.*, 1994; Zhong *et al.*, 1994). Additional support for the mechanism of alternative splicing comes from structural analysis of the mouse HNF-4 gene. It consists of at least 10 exons, with seven of these (exons 2 - 8) encoding the conserved sequences of the DNA-binding, hinge and dimerization domains (Taraviras, *et al.*, 1994). Significantly,

the third intron is located between the glutamic acid at postition 119 (E119) and the alanine at position 120 (A120). These residues are conserved in the mosquito isoforms (E144 and A145 of AaHNF-4a, and E117 and A118 of AaHNF-4b) and map precisely to the C-terminal boundary of the deletion in AaHNF-4c.

In the adult rat and mouse, the expression of HNF-4 is not restricted to the liver, but is fairly high in the kidney and intestine as well (Sladek *et al.*, 1994). Likewise, in the *Drosophila* embryo, HNF-4 mRNA has been detected in the developing fat body (liver analog), Malpighian tubules (kidney analog) and midgut (intestine analog) during organogenesis (Zhong *et al.*, 1993). HNF-4 transcripts were detected in several tissues of *B. mori* with the highest levels in the fat body, gut, ovaries, and testis (Swevers and Iatrou, 1998). Similarly, in the adult mosquito, HNF-4 transcripts were detected in the fat body, midgut, and Malpighian tubules. However, in the ovary, AaHNF-4b is the only transcript detected during the entire vitellogenic cycle. This is in contrast to *B. mori* where both isoforms have been found in the ovary and developing follicles (Swevers and Iatrou, 1998).

The expression pattern of the AaHNF-4 isoforms varies significantly over the course of the vitellogenic cycle in the fat body (figure 3.6). Importantly, these changes in AaHNF-4 expression correspond to essential shifts in the functional activity of this tissue (Raikhel, 1992). As cell culture transfection experiments have established that both AaHNF-4a and -4b are transcriptional activators, the differential expression of their transcripts in the fat body suggest that these

isoforms execute distinct functions required at different stages of the vitellogenic cycle. It is noteworthy that in nuclear receptors, the A/B domain has been implicated in both transactivation and recognition of regulatory sequences of various target genes (Tsai and O'Malley, 1994).

The AaHNF-4a transcript is present in the female fat body during the first day after eclosion, but its level dramatically decreases in the following days of previtellogenic development, a time when the fat body is preparing for vitellogenesis under the control of juvenile hormone. The AaHNF-4a transcript is practically absent during the first 24 hours PBM, during which vitellogenesis is initiated and the expression of YP genes reaches maximal levels under the control of the rising titer of the insect steroid hormone 20E. Transcript levels then gradually rise during the next 12 hours when the titer of 20E drops sharply and the expression of the YP genes are halted. It is intriguing that this is very similar to the profile of another nuclear receptor, AaUSPa, in the mosquito fat body during vitellogenesis (Wang *et al.*, 2000a). AaUSPa also showed its highest level in the first day post eclosion followed by a dramatic drop in mRNA levels that remained low until beginning to rise again at 24 hours PBM. Moreover, this expression of AaUSPa was dependent on the exposure then withdrawal of 20E from the fat body. Thus, AaUSPa was acting like a classical "late gene" in the ecdysone regulatory cascade in which ecdysone stimulates the expression of "early genes" but represses the expression of "late genes" (Ashburner *et al.*, 1974). The "early genes" then stimulate the expression of the "late genes" while repressing the action of ecdysone. This observation is all the more interesting

considering that the BmHNF-4a also behaved like a late gene during 20E-mediated development of the ovary (Swevers and Iatrou, 1998).

The significant elevation in the AaHNF-4a transcript level at 36 hours PBM, corresponds to the time when the synthesis of YPs are terminated and the fat body undergoes remodeling from a protein synthesizing tissue to a storage depot for lipid reserves (Kan and Ho, 1972; Behan and Hagedorn, 1978). Sun *et al.* (2000) determined that while synthesis of lipophorin, a lipid carrier protein, reached peak levels by 18 hours PBM, accumulation of lipophorin by the fat body continued throughout the postvitellogenic period. Recently, a putative lipohorin receptor was cloned from *Ae. aegypti* (S-J Seo, H-M Jun, A Raikhel, unpublished results). Northern blot analysis demonstrated that this gene was expressed in the fat body only between 30 and 48 hours PBM. In agreement with this, the greatest binding activity of nuclear extracts to a DR1 during vitellogenesis was at 36 hours PBM. The abolishing of this binding by the AaHNF-4 specific antibodies indicates that most, if not all, of the proteins in complex with the probe are AaHNF-4. Since HNF-4 is known to regulate genes involved in lipid metabolism in vertebrates (Sladek, 1994), it is likely that this function has been conserved in insects as well. Thus, it is possible that AaHNF-4a may play a role in regulating genes involved in lipogenesis in the mosquito fat body. This idea is further supported by a recent report that fatty acyl-CoA thioesters are potential ligands of HNF-4 which was previously considered an orphan nuclear receptor (Hertz *et al.*, 1998).

In contrast to AaHNF-4a, the AaHNF-4b mRNA accumulates during the previtellogenic period, then drops to a lower level that is maintained throughout the vitellogenic synthesis phase. In agreement with this, strong binding to a DR1 is detected with nuclear extracts prepared from fat bodies 3-5 days after emergence (post eclosion), while only weak binding is detected at 12 hours PBM and only moderate binding at 24 hours PBM, the time of peak YP synthesis. These results suggest that AaHNF-4b is involved in the previtellogenic preparation of the fat body but not in synthesis of the YPs. Nevertheless, the identification of HNF-4 binding sites in the 5'-upstream region of the YP genes *Aavg* and *Aavcp*, suggests the possible role of HNF-4b in regulating the expression of these genes. The putative HNF-4 motif in the *Aavg* gene, VgHNF4, is an imperfect DR1 that closely corresponds to the consensus HNF-4 binding site, while the binding site in the *Aavcp* gene, VCPHNF4, is more unique in that it has two nucleotides that separate the two half-sites, previously reported only for the erythropoietin gene (Blanchard *et al.*, 1992). However, as shown in figure 3.13, several NR are able to bind to a DR1, including AaEcR/AaUSP, the functional receptor for the steroid hormone 20E. As these are the results of *in vitro* experiments, the importance of VGHNF4 and VCPHNF4 *in vivo*, as well as the proteins that bind there, remains to be determined.

Synthesis of YPs reaches maximal levels between 24-27 hours PBM, then decreases dramatically. It is at this time that the level of the AaHNF-4c transcript increases relative to the other two. This pattern of expression raises the possibility that AaHNF-4c plays the role of a transcriptional repressor involved in

switching gene regulation in the mosquito fat body. A number of proteins with abolished DNA-binding function have been identified including E75B (Seagraves and Hogness, 1990), SHP (Seol *et al.*, 1996), Id (Benezra *et al.*, 1990), and ERbA p30 and p27 (Bigler *et al.*, 1992), all belong to families of dimeric transcription factors. It is plausible then that the AaHNF-4c isoform can act as a dominant negative factor by forming heterodimers with other AaHNF-4 isoforms, which are incapable of DNA-binding and subsequent transcriptional activation. In support of this hypothesis, I have shown that AaHNF-4c can abolish the binding of AaHNF-4a, -4b, and even rat HNF-4α1 to target DNA sequences *in vitro*, and to repress HNF-4 mediated transcriptional activation in cell culture. The ability of the mosquito HNF-4c isoform to inhibit the formation of binding complexes between the rat HNF-4α1 and the HNF-4 response element clearly demonstrates dimerization between these HNF-4 homologs, as was also demonstrated between *Drosophila* and rat (Zhong *et al.*, 1993). The precise role of the AaHNF-4c isoform in regulation of the fat body-specific gene expression in the mosquito remains to be elucidated. However, the discovery of an HNF-4 isoform which lacks the DNA binding domain and serves as a transcriptional repressor has expanded our view on potential functions of this important transcription factor.

CHAPTER 4

CLONING AND CHARACTERIZATION OF MOSQUITO CREB

INTRODUCTION

As mentioned previously in Chapter 1, vitellogenesis requires the coordinated activity of several tissues: the midgut where the blood meal is digested, the fat body which synthesizes the yolk protein (YP) precursors, and the ovaries which internalize these proteins through receptor mediated endocytosis. The need to coordinate these activities necessitates the ability of these tissues to communicate with each other. Peptide hormones have been implicated in the regulation of these processes, and some have been identified and characterized (Chapter 1, pages 13-17). The cAMP signal transduction cascade is one of the more common pathways utilized by cells to convey information about the binding of peptide hormones to cell surface receptors, to the nucleus where target genes are activated. Cyclic AMP has already been implicated in the regulation of several vitellogenic related events (Chapter 1, page 35). In combination with the identification of putative cAMP response elements (CREs) in the regulatory regions of the YP genes *Aavg* and *Aavcp* (described below), a mosquito homolog to the cAMP response element binding protein (CREB) was cloned and characterized as a begining to investigate the role of the cAMP signal transduction cascade in the regulation of vitellogenesis.

RESULTS

Cloning of mosquito CREB cDNA

Using degenerate primers designed on the high conservation of amino acids in the DNA binding and protein dimerization domains (bZIP) between *Drosophila* and human CREB proteins, the polymerase chain reaction (PCR) was used to amplify a 112 bp fragment. Sequencing confirmed that the PCR product encoded for a portion of the bZIP region. This fragment was then used to screen and isolate a 4 kb clone from a mosquito cDNA library. This clone had a 1.93 kb 5'-untranslated region and an open reading frame of 885 bp that encoded for a protein of 295 amino acids (figure 4.1). The 3'-untranslated region was 1.2 kb long but lacked a consensus polyadenylation signal sequence. The mosquito homolog, which I have named AaCREB, showed high sequence similarity to both the *Drosophila* and human proteins in the bZIP region, having 97% and 86% identity respectively (figure 4.2A). However, outside of this region the sequence similarity was much reduced, having only 38% identity with the *Drosophila* isoform dCREB2-d (figure 4.2B). Homology between the mosquito and human CREB proteins outside of the bZIP region is restricted mainly to the residues surrounding the PKA-box, however, several residues shown to be critical for interaction with CBP are conserved (Parker *et al.*, 1996; Radhakrishnan *et al.*, 1997) (figure 4.2C); surprisingly, the mosquito sequence shows higher similarity with the human protein than with *Drosophila* in this area.

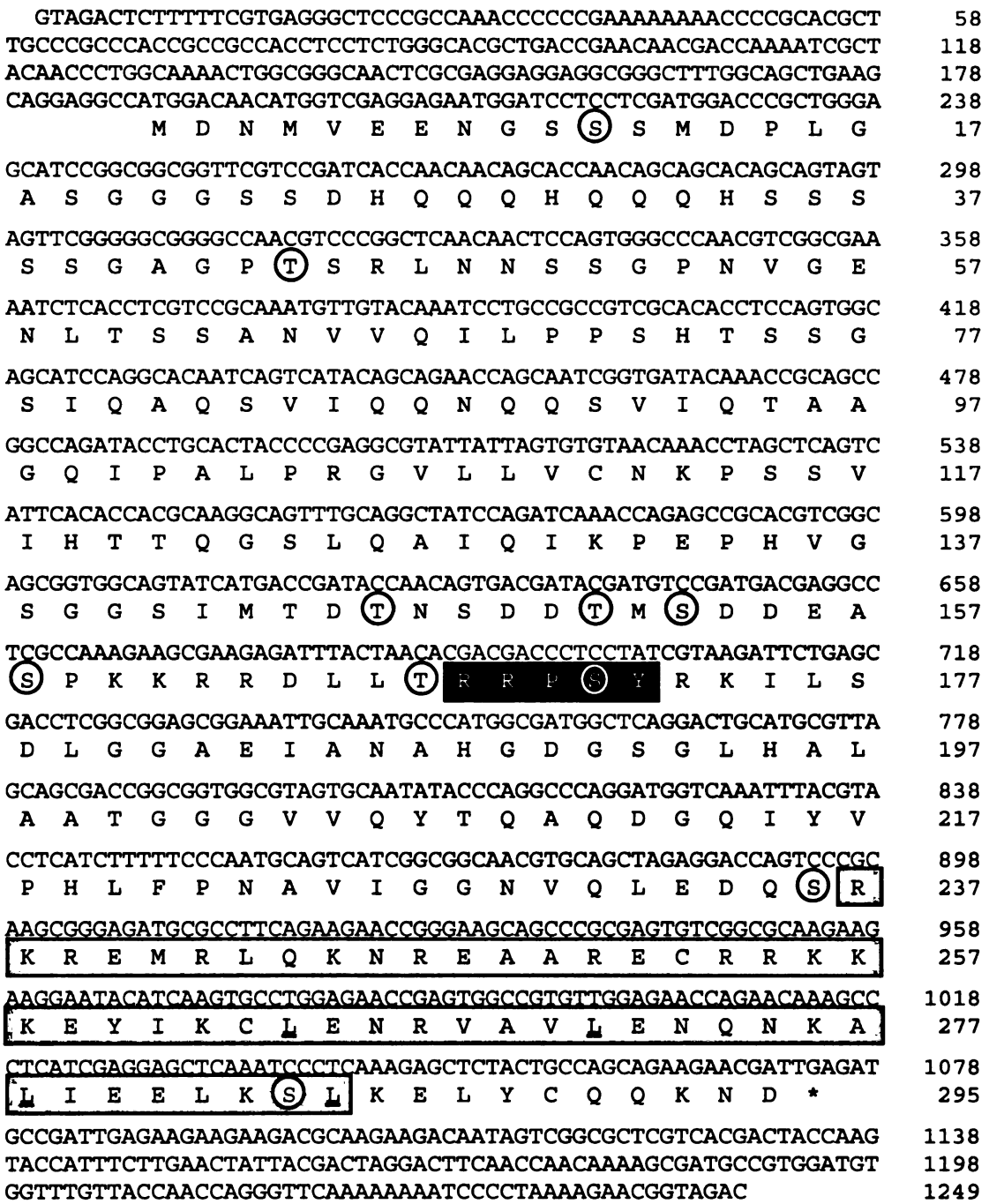


Figure 4.1, Nucleotide and deduced amino acid sequence of AaCREB. The 4 kb AaCREB cDNA was sequenced in its entirety; only the 1.25 kb *Acc I* restriction fragment subcloned for *in vitro* expression is shown here. The PKA phosphorylation site is boxed in black with the single letter abbreviations of the amino acids in reverse phase. The bZIP domain is boxed in gray with the leucines of the leucine zipper region underlined. Potential phosphorylation sites are circled and were identified using the MOTIFS computer program (Wisconsin Package, Genetics Computer Group). Numbers on the right represent nucleotide and amino acid positions.

Figure 4.2, Comparison of AaCREB with *Drosophila* and human CREB proteins. (A) Alignment of the bZIP domains of mosquito (AaCREB), Drosophila (DmCREB, Usui *et al.*, 1993; Yin *et al.*, 1995b), and human (HsCREB, Hoeffler *et al.*, 1988) CREB proteins. Vertical lines represent identical amino acids, colons depict conservative substitutions. (B) Comparison of AaCREB with the *Drosophila* isoform dCREB2-d. Alignment was generated by the GAP computer program (Wisconsin Package, Genetics Computer Group) and adjusted by eye. Dashes indicate gaps introduced to optimize alignment. Vertical lines and colons are the same as in (A). Numbers on the right indicate amino acid positions. (C) Alignment of the sequence surrounding the PKA phosphorylation site of mosquito, *Drosophila*, and human CREB proteins. The PKA recognition site is boxed in. The asterisks above the amino acids in the human CREB sequence are residues shown to make contact with CBP (Parker *et al.*, 1996; Radhakrishnan *et al.*, 1997). Vertical lines and colons are the same as in (A).

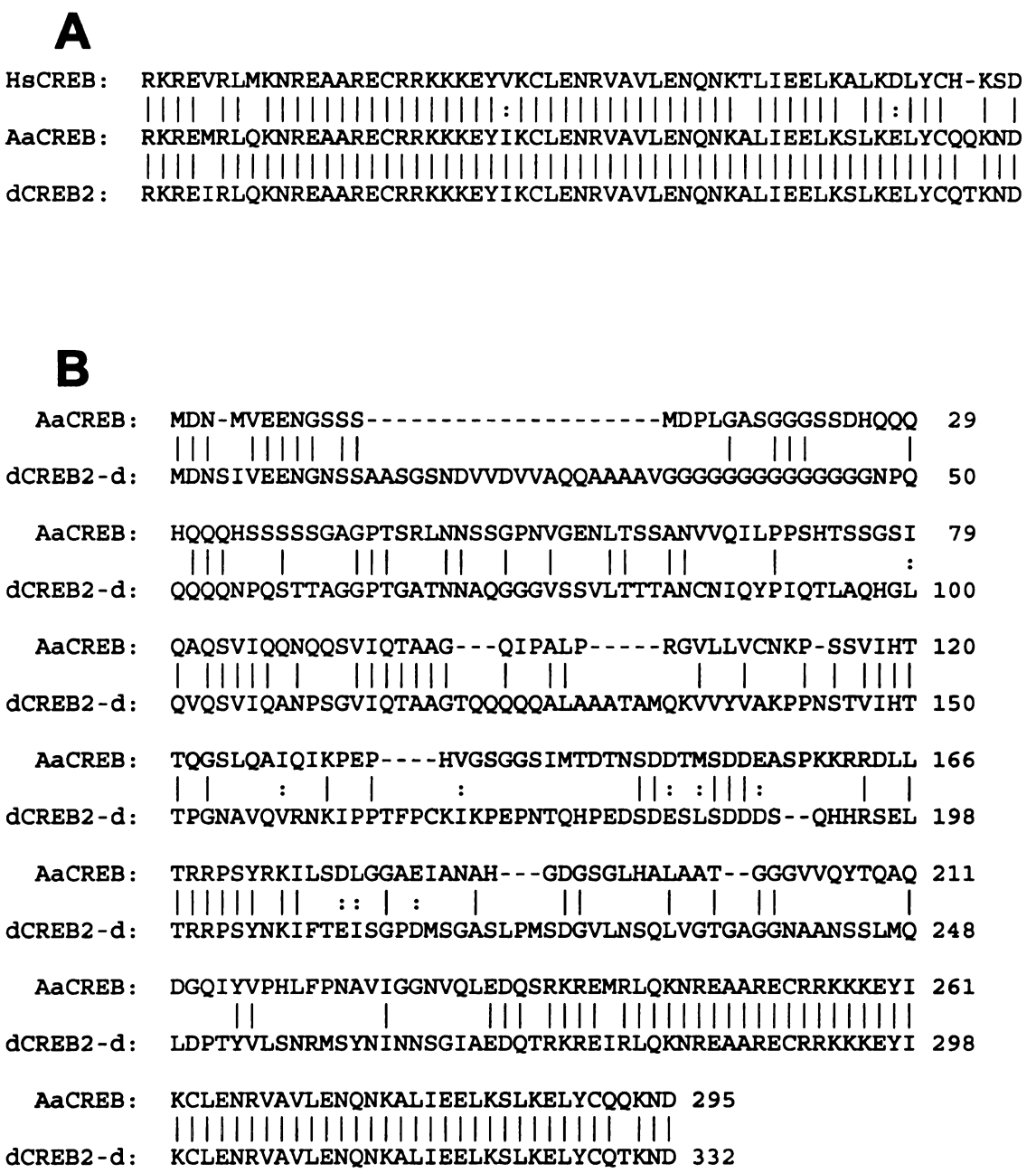


Figure 4.2, Comparison of AaCREB with *Drosophila* and human CREB proteins.

Northern blot analysis of AaCREB expression

Northern blot hybridization was used to analyze the mRNA expression profile of AaCREB. Messenger RNA was extracted from the fat bodies of female mosquitoes at various time points post eclosion and post blood meal. After gel electrophoresis and nucleic acid transfer to a nitrocellulose membrane, the blot was hybridized with a single strand antisense DNA probe. Three transcripts of approximately 5, 4, and 3.5 kb were detected (figure 4.3). The 4 and 3.5 kb transcripts were detected at all time points tested, showing highest levels during the previtellogenic (post eclosion) period and the first 12 hours after a blood meal. The 5 kb transcript was of much lower abundance and was only seen at the time points when the 4 and 3.5 kb messages were at their highest levels. Failure to detect the 5 kb transcript at other time points is likely due to the limitations of the sensitivity of the Northern analysis. For comparison, the blot was also hybridized with probes for mosquito actin (Deitsch *et al.*, 1995a), a constitutively expressed gene, and vitellogenic carboxypeptidase (*Aavcp*), an ecdysteroid responsive yolk protein (YP) gene. It is interesting to note that between 18 and 36 hours post-blood meal, a time when YP synthesis is at its maximum in the fat body, the CREB mRNA is at its lowest levels (compare the expression of CREB with *Aavcp* in figure 4.3C).

Electrophoretic Mobility Shift Assays (EMSA)

In vitro expression of the AaCREB cDNA produces a 41 kDa protein when resolved by sodium dodecyl sulfate - polyacrylamide gel electrophoresis

Figure 4.3, mRNA expression profile of AaCREB in the fat body. Northern blot analysis of mRNA prepared from dissected fat bodies at the indicated times: d, days; h, hours; PE, post eclosion; PBM, post blood meal. Panels A) and C) have 35 mosquito fat body equivalents loaded per lane. Panel B) has 25 mosquito fat body equivalents loaded per lane. This blot had been previously hybridized with probes for mosquito lipophorin, actin, and vitellogenic carboxypeptidase (VCP) (Sun *et al.,* 2000). The bands identified by an asterisk in panel B are a result of a residual signal from the VCP hybridization. An RNA molecular weight marker (in kilobases) is marked to the left of each panel.

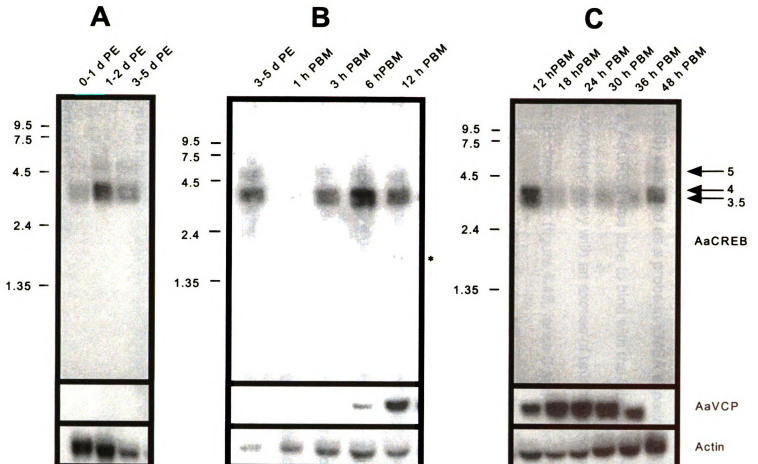


Figure 4.3, AaCREB mRNA expression profile in the fat body.

(SDS-PAGE) (figure 4.4A). Electrophoretic mobility shift assays (EMSAs) were used to examine the DNA-binding ability of this protein. The consensus binding site for CREB proteins is the sequence 5'-TGACGTCA-3', known as the cyclic-AMP response element (CRE). The CRE of the rat somatistatin gene (SomCRE; Montminy *et al.*, 1986) was used as a probe in EMSA experiments. As seen in figure 4.4B, lane 2, AaCREB was able to bind with the labeled SomCRE. This binding could be competed away with an excess of unlabeled SomCRE DNA, but not by a non-specific competitor (figure 4.4B, lanes 3 and 7). In addition, putative CREs identified in the sequences flanking the YP genes *Aavg* and *Aavcp* were also able to compete for binding (figure 4.4B, lanes 4-6). This experiment demonstrated that AaCREB is a DNA-binding protein and may have the potential to regulate *Aavg* and *Aavcp* gene expression through binding to their CREs.

Next, nuclear extracts were used in EMSA experiments to examine at what time points during vitellogenesis AaCREB is actively bound to DNA. As seen in figure 4.5, binding to the labeled probe was only detected at 24 and 36 hours post-blood meal; a time when YP synthesis in the fat body reaches its peak levels then rapidly terminates. This binding could be super-shifted with antibodies that recognize the phosphorylated form of CREB, indicating that the proteins binding in the nuclear extracts are CREB related.

Transient transfection in CV-1 cells

To examine whether AaCREB could act as an activator or repressor of transcription, transient transfections were performed in the African green monkey

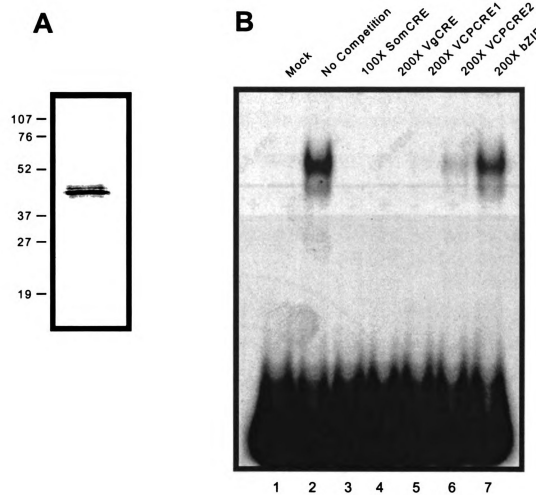


Figure 4.4, AaCREB can bind to CREs *in vitro*. A) Expression of the AaCREB cDNA *in vitro*. The pBSAaCREB clone (see Materials and Methods; Chapter 2, page 57) was expressed *in vitro* and labeled with ^{35}S-methionine; translation products were resolved by electrophoresis in a 10% SDS-polyacrylamide gel. B) EMSA using *in vitro* expressed AaCREB. EMSA was performed using *in vitro* expressed AaCREB and ^{32}P-labeled SomCRE as the probe. Lane 1, SomCRE probe with unprogrammed TNT lysate; lane 2, SomCRE probe with AaCREB TNT lysate; lanes 3-7, same as lane 2 plus the addition of unlabeled competitor DNA, expressed in molar excess with respect to the labeled probe. Competitor DNA sequences are: SomCRE: TGACGTCA; VgCRE: AT**ACGTCA** (located 1.7 kb upstream of the transcription start site); VCPCRE1, AT**ACGTCA** (located 153 bp downstream of the polyadenylation signal); VCPCRE2, **TTCCGTCA** (located 1.3 kb downstream of the polyadenylation signal); bZIP2, CTT**GA**GCAAT (Dittmer and Raikhel, 1997), which binds the CCAAT/enhancer-binding protein (C/EBP) was used as a nonspecific competitor. Bold type indicates matches to the SomCRE.

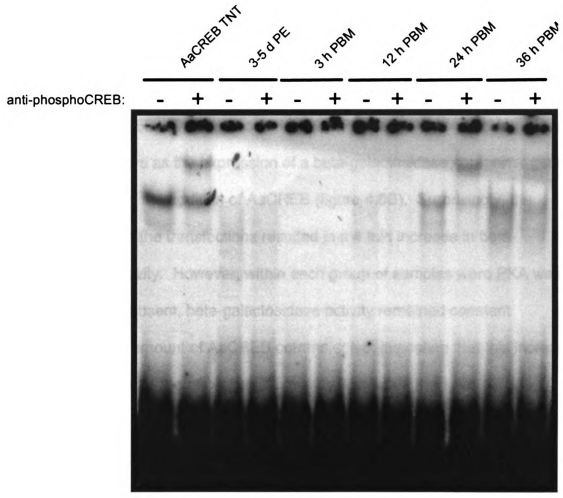


Figure 4.5, CRE-binding proteins are present in fat body nuclear extracts. EMSA was performed with fat body nuclear extracts prepared from the time points indicated (d PE, days post eclosion; h PBM, hours post blood meal) and incubated with a ^{32}P-labeled SomCRE as the probe. Ten mosquito fat body equivalents were used per lane. *In vitro* expressed AaCREB was used as a control for binding. The addition (1:100 dilution), or omission, of anti-phosphoCREB antibodies is indicated above each lane.

kidney cell line (CV-1). Using a luciferase reporter gene containing 4 copies of the consensus CRE, a 5-6 fold increase of luciferase activity was detected upon induction of the cAMP pathway with a plasmid containing the catalytic subunit of protein kinase A (PKA) (figure 4.6A). This activation was repressed by the cotransfection of AaCREB in a dose dependent manner. This repression is likely due to the competition for DNA-binding between the endogenous CREB and the exogenous AaCREB. The repression is unlikely to be due to squelching of basal transcription factors as the expression of a beta-galactosidase coreporter gene was unaffected by the addition of AaCREB (figure 4.6B). Surprisingly, the addition of PKA in the transfections resulted in a 4 fold increase in beta-galactosidase activity. However, within each group of samples were PKA was either present or absent, beta-galactosidase activity remained constant regardless of the amount of AaCREB cotransfected, indicating that this increase in activity was due to the phosphorylation of factors other than AaCREB.

DISCUSSION

In order to begin to elucidate the role of the cyclic-AMP signal transduction pathway on the regulation of gene expression in the fat body during vitellogenesis, I have cloned a mosquito homolog of the CREB/CREM transcription factor. The mosquito clone, AaCREB, shares nearly perfect homology with the *Drosophila* and mammalian CREB proteins in the DNA-binding and protein dimerization (bZIP) domains, but much reduced homology

Figure 4.6, Transient transfections of AaCREB in CV-1 cells. A) AaCREB represses CREB-mediated transcriptional activation. Cells were cotransfected with the reporter plasmid pCRE-Luc, which contains 4 copies of the CRE consensus sequence upstream of a synthetic promoter fused to the *luciferase* gene, and various amounts of pcAaCREB. Endogenous CREB activity was stimulated by the addition of pFC-PKA, which encodes for the catalytic subunit of protein kinase A. B) PKA-enhancement of β-gal expression is CREB independent. Cells were cotransfected with the reporter plasmid pCMVβ, which contains the β-galactosidase gene fused to the human cytomegalovirus immediate early promoter, various amounts of pcAaCREB, and either with or without pFC-PKA.

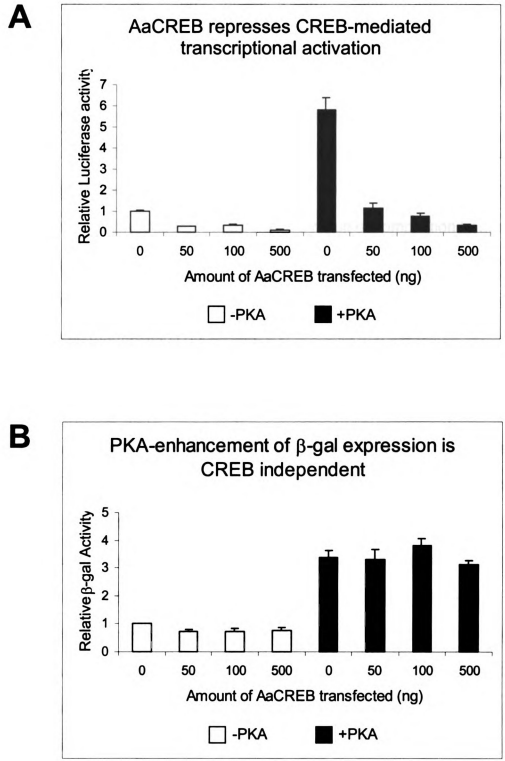


Figure 4.6, Transient transfections of AaCREB in CV-1 cells.

outside of this region. Based on comparison with the several *Drosophila* isoforms thus described, AaCREB is most related to dCREB2-d; initial characterization of dCREB2-d reports this isoform not to be an activator of gene transcription (Yin *et al.*, 1995b). Transient transfection of the AaCREB cDNA into CV-1 cells likewise failed to activate gene transcription of a reporter construct and instead acted as a potent inhibitor of PKA-mediated transcriptional activation. Presumably, this repression was the result of competition between the exogenous AaCREB and the endogenous CREB for binding to CREs. However, because of the lack of homology outside of the bZIP domain between AaCREB and its mammalian counterpart, it is possible that the failure of AaCREB to activate gene transcription may be due to its inability to interact with other cofactors in mammalian cells.

Further support for the role of AaCREB as a repressor is suggested by the time at which active DNA-binding complexes are observed. When several time points pre- and post-blood meal were examined by EMSA using fat body nuclear extracts, protein/DNA complexes were detect only at 24 and 36 hours after a blood meal; a time when vitellogenesis is in the transition from synthesis to termination of YP gene expression. The use of nuclear extracts collected at more frequent time points during vitellogenesis should provide a more detailed analysis of the DNA-binding profile of AaCREB.

Northern blot analysis shows the *Aacreb* gene to be constitutively expressed throughout vitellogenesis. Interestingly, when the expression of the YP gene *Aavcp* is at its highest, AaCREB mRNA levels are at their lowest (figure

4.3). As detection of DNA-binding by EMSA falls within this time frame, this suggests a feedback mechanism whereby AaCREB may down regulate its own expression. Three asymmetrical CREs have been identified in the promoter of the human *creb* gene and it is postulated that this allows for transcriptional regulation, both positive and negative, by CREB and/or CREM isoforms (Foulkes *et al.*, 1991; Meyer *et al.*, 1993). The identification of putative CREs from the flanking regions of the *Aavg* and *Aavcp* genes, and the binding to these sequences by AaCREB *in vitro*, also suggests that these genes could be regulated by AaCREB as well.

The exact role of AaCREB in vitellogenesis is, as yet, unknown. However, AaCREB could repress *Aavg* and *Aavcp* gene transcription by preventing other transcription factors that recognize CREs, or overlapping response elements, from binding at these sites. Interestingly, the putative mosquito CREs identified also show strong sequence similarity to the binding sites of another group of bZIP transcription factors, the PAR subfamily (Drolet *et al.*, 1991; Johnson 1993; Hurst 1994). When the flanking nucleotides are included in the analysis, VCPCRE1 matches 7 of 10 nucleotides with RTTACGTAAY, the PAR binding site consensus sequence (Haas *et al.*, 1995; Falvey *et al.*, 1996). VCPCRE2 and VgCRE each match 8 of 10 nucleotides with this consensus. Significantly, one member of the PAR family is the chicken vitellogenin binding protein (VBP), which binds to a control element in the promoter of the chicken *vitellogenin II* gene (Iyer *et al.*, 1991; Burch and Davis, 1994). When the amino acid sequence of the basic (DNA-binding) region of the chicken VBP was used to search the

Drosophila genome database (Adams *et al.*, 2000), three matches were found. One match was the previously identified *giant* gene, which binds to two regulatory sequences in the upstream region of the *Krüppel* gene (Capovilla *et al.*, 1992) which show similarity to both CREB and PAR response elements. However, the other two matches represent previously uncharacterized *Drosophila* genes: CG17888 (GenBank accession number AAF50513) shares 82% identity and 91% similarity with the basic region of VBP, while CG7786 (GenBank accession number AAF58037) shares 64% identity and 88% similarity with VBP; both of which are greater then the homology between Giant and VBP. Thus it is possible that mosquito members of the CREB and PAR subfamilies of bZIP proteins may compete for binding to these elements; an AaVBP homolog which acts as a transcriptional activator, and AaCREB which represses transcription.

If AaCREB functions as a repressor, then post-translational modifications that affect DNA-binding, either directly or through interference of protein dimerization, could allow for additional regulation of its action. Phosphorylation of transcription factors can affect DNA-binding either positively or negatively (Hunter and Karin, 1992). For example, c-Jun (Lin *et al.*, 1992), the vitamin D receptor (VDR; Hsieh *et al.*, 1993), NF-IL6 (Trautwein *et al.*, 1994), Sp-1 (Armstrong *et al.*, 1997), and HNF-4 (Viollet *et al.*, 1997) all have reduced affinity for their respective response elements when phosphorylated at specific residues, while the serum response factor (SRF; Janknecht *et al.*, 1992) and the adenovirus E4 gene promoter binding protein (E4BP4; Chen *et al.*, 1995) showed

enhanced DNA-binding upon phosphorylation. How DNA-binding by CREB is regulated is poorly understood and controversial. Some studies have demonstrated that phosphorylation by PKA increases DNA-binding to symmetrical and/or asymmetrical CREs (Nichols *et al.*, 1992; Bullock and Habener, 1998), while others have shown that DNA-binding is unaffected (Richards *et al.*, 1996; Wu *et al.*, 1998). However, it should be noted that Nichols *et al.* (1992) reported that the PKA-enhanced DNA-binding of CREB could be attenuated by phosphorylation of serine 115 in CREB 327 (this corresponds to serine 129 in CREB 341) by glucogen synthase kinase-3. In addition, *in vivo* footprinting and *in vivo* crosslinking experiments have shown that CREB binding to CREs is enhanced after stimulation of cells with forskolin, an adenylate cyclase activator, though the mechanism by which this is achieved is unknown and may be indirect (Boshart *et al.*, 1991; Wölfl *et al.*, 1999). One possible mechanism proposes that the phosphorylation of CREB by PKA allows for the recruitment of CBP which remodels the chromatin at the CRE, via its intrinsic histone acetylase (HAT) activity, thereby granting CREB access to the DNA (Wölfl *et al.*, 1999). Nevertheless, CREB binding *in vivo* appears to be a signal-mediated event. Since multiple kinase recognition motifs are present in the CREB protein sequence, a complex pattern of phosphorylation that may affect DNA-binding or protein dimerization is possible. Of significance, the *Drosophila* dCREB2 protein when purified from head extracts, embryos, or Schneider cells, fails to bind to a CRE *in vitro* unless first treated with phosphatases (J. Yin, personal communication). A similar result was observed for c-Jun in human

epithelial and fibroblastic cells; activation of protein kinase C results in the dephosphorylation of c-Jun at specific residues and is concurrent with an increase in AP-1 binding activity (Boyle *et al.*, 1991). Interestingly, in AaCREB, serine 236 (immediately amino terminal of the basic region) is a possible target for protein kinase C, and serine 284 (in the leucine zipper region) can potentially be phosphorylated by protein kinase C and casein kinase II (figure 4.1). Both of these residues are alanines in mammalian CREB/CREM proteins and therefore are not subject to phosphorylation (though threonine 324, located in the leucine zipper, is a possible target for casein kinase II). Recently, Szilak *et al.* (1997) have reported that phosphorylation within the leucine zipper of a bZIP protein destabilizes the alpha helix and results in a reduction of protein dimerization. Thus, phosphorylation of these serines may present an additional mode of regulation of insect CREB proteins either through direct interference of binding to CREs, or by disrupting the formation of dimers. Mutational analysis of these and other putative phosphorylation sites should help address the role of phosphorylation in regulating AaCREB DNA-binding.

CHAPTER 5

SUMMARY AND CONCLUSIONS

SUMMARY AND CONCLUSIONS

HNF-4 is a liver-enriched transcription factor involved in a wide array of liver gene expression. These include genes for serum proteins, blood coagulation factors, metabolism of carbohydrates, lipids, steroids and amino acids, and liver differentiation (reviewed in Sladek, 1994). Its high amino acid sequence identity and tissue distribution in insects suggests that it plays a similar role in invertebrates. Three isoforms of a mosquito homolog have been cloned and characterized (Kapitskaya *et al.* (1998) and this dissertation). The AaHNF-4a and 4b isoforms differ only at their amino terminus, while AaHNF-4c is missing a greater portion of the A/B domain and all of the DNA binding domain. The three isoforms show differential expression of their mRNA throughout the first vitellogenic cycle.

Expression of AaHNF-4 transcripts during the first vitellogenic cycle.

During the first day post-eclosion, AaHNF-4a and 4b transcripts are approximately equally abundant, but AaHNF-4b clearly becomes the predominant transcript for the rest of the previtellogenic period (see figure 3.6, page 78). This pattern remains the same for the first 18 hours PBM. From 24 to 33 hours PBM, the two are again in approximately equal proportion, and then AaHNF-4a becomes the more abundant of the two. Similar to AaHNF-4a, the AaHNF-4c isoform appears to be absent for most of the previtellogenic period and during the synthesis stage of the YPs. It is only during the peak of YP

synthesis just prior to termination of these genes that AaHNF-4c transcripts are detected, and remain during the termination stage of vitellogenesis.

The expression of AaHNF-4 proteins in the fat body remains to be resolved.

While the mRNA profile of AaHNF-4 suggests that each isoform may be involved in different aspects of the first vitellogenic cycle in the fat body, the expression of the proteins may not necessarily mirror that of the mRNA and thus it is critical to determine the protein profiles of the isoforms. The specificity of the α-GSTHNF-4(Com) antibodies was demonstrated by both Western analysis of *in vitro* synthesized proteins, and EMSA experiments with nuclear extracts (figures 3.12 and 3.13, pages 93 and 95). However, in Western analysis with nuclear extracts, these antibodies were unable to detect the AaHNF-4 proteins. This can likely be attributed to the relatively low amount of the antibodies with respect to the total amount of protein in the antisera. As detailed in the Chapter 2 (Materials and Methods), the antibodies were purified by ammonium sulfate precipitation. Since this method serves to enrich the IgG component of the serum, the resulting fraction is a heterogeneous mixture of antibodies. Thus, at dilutions low enough to be able to identify AaHNF-4 proteins, nonspecific binding by the other antibodies in the mixture resulted in high background when detected by chemiluminescence. This problem was very prevalent with the antibodies specific to AaHNF-4a [α-GSTHNF-4a(NH_2)], even when used against *in vitro* synthesized proteins (not shown). Since the GST-HNF-4a fusion construct contained only the first 33 amino acids of AaHNF-4a, it can be expected that

fewer epitopes were presented to the rabbit's immune system as compared to the GST-HNF-4(Com) which encompassed 300 amino acids, and thus fewer antibodies were produced. Therefore, affinity chromatography should be used to further purify the AaHNF-4 specific antibodies. A general scheme for purification is discussed below.

Affinity purification of α-GSTHNF-4(Com) and α-GSTHNF-4a(NH_2) antibodies.

In the first step, a bacterial culture expressing GST will be lysed by sonication and the cellular proteins immobilized on a CNBr-activated Sepharose 4B matrix. The antisera will be incubated with the matrix to remove antibodies against GST and bacterial proteins that may be present. Having removed the antiGST antibodies, the GSTHNF-4 fusion proteins will then be coupled to the matrix and used to bind the α-GSTHNF-4 antibodies. After extensive washing to remove unbound proteins, the α-GSTHNF-4 antibodies can be eluted from the matrix. Once purified, these antibodies should allow for Western analysis to be performed with nuclear extracts. Because the AaHNF-4a and -4b proteins are of such similar size (see figure 3.12), it is unlikely that it will be possible to differentiate between them by SDS-PAGE and Western analysis with the α-GSTHNF-4(Com) antibodies. Thus, analysis using the two antibodies in parallel will be invaluable in deducing the profile of AaHNF-4b. In addition, Western analysis should be able to determine when and at what levels (with respect to the other isoforms) AaHNF-4c is expressed.

Possible roles for AaHNF-4 in the mosquito fat body.

The combination of EMSA with nuclear extracts and HNF-4 specific antibodies indicates that HNF-4 binding activity is highest during the previtellogenic and termination stages in the fat body (figure 3.14, page 97). This suggests that the main role of AaHNF-4 is to regulate the developmental events occurring in the fat body prior to and after YP gene expression. With respect to the types of genes known to be regulated by HNF-4 in mammals, perhaps those most relevant here are genes involved in the metabolism of lipids. Most of the apolipoproteins synthesized in the vertebrate liver are regulated at least in part by HNF-4. A direct role has been demonstrated for HNF-4 in the regulation of *apoAII*, *apoB*, *apoCII*, and *apoCIII* (Sladek *et al.*, 1990; Ladias *et al.*, 1992; Mietus-Snyder *et al.*, 1992; Metzger *et al.*, 1993; Kardassis *et al.*, 1998; Vorgia *et al.*, 1998). In addition, mouse embryos generated from HNF-4$\alpha^{-/-}$ embryonic stem cells failed to express *apoAI* and *apoAIV* (Duncan *et al.*, 1997; Li *et al.*, 2000). The striking conservation of the tissue distribution of HNF-4 between vertebrates and invertebrates suggests that HNF-4 likely regulates many of the same classes of genes in insects as well. Thus a probable candidate in mosquitoes is the *lipophorin* (*Aalp*) gene. *Aalp* could also prove to be an excellent example for studying tissue and hormonal regulation. In addition to its normal levels of expression, *Aalp* is up regulated by 20E during vitellogenesis (Van Heusden *et al.*, 1998; Sun *et al.*, 2000). As both AaHNF-4 and AaEcR/AaUSP can bind to a DR1 (figure 3.13), it would be interesting to examine if the same binding site is used for hormonal and tissue specific

regulation, or if separate elements are required. Since AaHNF-4b appears to be constitutively expressed in the fat body, I would hypothesize that it is the isoform more likely to regulate lipophorin gene expression. The recent cloning of a partial cDNA for AaLp (Van Heusden *et al.*, 1998) should allow for the isolation of the gene and the subsequent characterization of its promoter.

Another candidate gene for HNF-4 regulation is the *lipophorin receptor* (*Aalpr*). As mentioned in Chapter 3, a putative AaLpR was cloned from a mosquito fat body cDNA library. Northern hybridization has shown this transcript to be expressed only between 30 and 48 hours PBM, a time when the fat body is increasing storage of lipid reserves. EMSA with nuclear extracts prepared from fat bodies at 36 hours PBM detected strong binding with the DR1 probe (figure 3.14A). The abolishing of the shift complex with α-GSTHNF-4(Com) antibodies indicates that this complex contains mostly, if not entirely, AaHNF-4 proteins (figure 3.14B). This correlation between *Aalpr* gene expression and AaHNF-4 binding activity in the fat body strongly suggests that AaHNF-4 may be involved in its regulation. Given that *Aalpr* expression was detected only during the termination period, AaHNF-4a would seem to be the more likely of the two to be involved in its regulation. The (potential) identification of two genes showing differential expression but both being regulated by HNF-4 would present an excellent opportunity to study isoform specific gene regulation.

Possible regulation of AaHNF-4 expression by 20E.

The complete amino acid identity of the AaHNF-4 isoforms, with the exception of their amino terminus, suggests that they are the products of differential splicing and/or alternative promoter usage of a single gene. A logical question is to ask what is responsible for the differential expression of these isoforms during vitellogenesis. The mRNA expression profile of AaHNF-4a bears a striking resemblance to that of AaUSPa. Since 20E was shown to regulate the expression of AaUSPa (Wang *et al.*, 2000a), and BmHNF-4a (Swevers and Iatrou, 1998), similar experiments should be carried out for AaHNF-4. In these experiments, mosquito fat bodies are cultured *in vitro* either continuously in the presence of 20E, or for only a short time followed by culturing in media without 20E. At various time points during incubation, mosquito fat bodies are collected and RNA is isolated from the tissue. The RNA is reverse transcribed, and the synthesized cDNA is used as template for PCR (RT-PCR) with isoform specific primers. Following amplification, the products are separated by gel electrophoresis and transferred to a nitrocellulose membrane for Southern hybridization. Phosphorimaging can then be used to quantitate the activity of each sample. In parallel, PCR can be performed with primers specific for AaUSPa, AaVCP, and actin as controls. AaUSPa should only show stimulation in samples where 20E was exposed and then withdrawn from the fat body. AaVCP should show activation only in the presence of 20E. Actin levels should remain constant regardless of the presence or absence of 20E.

Possible regulation of YP gene expression by cAMP.

Though it is believed that 20E is the main regulator of YP gene expression, it is likely that peptide hormones also play an important role. This is further supported by the affect stimulators of the cAMP pathway have on vitellogenin accumulation by the ovaries (Sappington *et al.*, 1998) and the secretion of salts and water from the Malphigian tubules after a blood meal (Petzel *et al.*, 1987). Because the cAMP pathway is a very common signal transduction cascade utilized by cells and a link has already been established with other vitellogenic events, I was interested in the possible role this pathway may play in regulating YP gene expression. Towards this goal, a cAMP response element binding protein (CREB) transcription factor was cloned. As described in Chapter 4, the mosquito homolog contains the characteristic domains of the CREB family of transcription factors; mainly a putative KID region and a highly conserved bZIP domain responsible for DNA binding and protein dimerization. The gene is constitutively expressed and produces two major transcripts of 3.5 and 4 kb, and a minor transcript of 5 kb (figure 4.3, page 113).

Functional analysis in cell culture indicates that AaCREB can be a potent repressor of transcription. Activation of a CRE-responsive luciferase reporter gene by endogenous CREB in CV-1 cells could be blocked by coexpression with AaCREB (see figure 4.6, page 119). Presumably, this repression is by competition for binding to response elements in the DNA. However, as negative results can be misleading, AaCREB expression should be tried in another cell line. The *Drosophila* dCREB2-a transfections were performed in F9 cells,

demonstrating that an insect CREB protein can work in mammalian cells. Since CREB is a ubiquitous protein, it seems unlikely that cell specific accessory factors would allow for an insect homolog to function as an activator in one cell line but not in another, and at the same time show no such discrimination towards the mammalian protein. Nevertheless, transfections should be tried in F9 cells since this cell line was shown to support dCREB2-a mediated transcription (Yin *et al.*, 1995b).

Additionally, current work has demonstrated that the β-galactosidase reporter construct used to normalize transfection efficiency is affected by PKA (figure 4.6B). Since this activation was consistent, regardless of the presence or absence of AaCREB, it can be concluded that the affect of PKA was due to its phosphorylation of factors other than CREB. Nevertheless, it is less than ideal to have a "control" vector affected by the treatment. Therefore, a vector utilizing a different promoter to drive β-galactosidase gene expression should be used. The CMVβ plasmid used in the experiments reported here contains the cytomegalovirus (hence CMV) immediate early gene promoter. Alternatively, luciferase reporter activity can be normalized against total protein content of the cells. However, since this does not directly reflect the efficiency of transfection, it will require several more repetitions to be performed to reduce the variation caused by differences in transfection efficiency of independent experiments.

CREB as a potential regulator of YP gene expression.

Transient transfection in CV-1 cells demonstrates that AaCREB can function as a transcriptional repressor. This notion is further supported by the detection of shift complexes only at 24 and 36 hours PBM, when YP gene expression is terminated, by EMSA with nuclear extracts (figure 4.5, page 116). Examination of the regulatory regions flanking the *Aavg* and *Aavcp* genes identified putative CREs (figure 4.4, page 115). EMSA has established that these are indeed functional CREB-binding sites. I proposed that AaCREB may function as a repressor by competing with transcriptional activators for binding to the *Aavg* and *Aavcp* CREs. I believe that shifted complexes were only detected at 24 and 36 hours PBM with nuclear extracts because the consensus CRE was used as a probe. In such a case, the *Aavg* and *Aavcp* CREs may serve as response elements for other transcription factors during the synthesis stage of vitellogenesis. As discussed in Chapter 4 (page 121), the *Aavg* and *Aavcp* CREs also share strong sequence homology with the consensus sequence for the PAR family of bZIP proteins. As shown in figure 5.1, the SomCRE matches only 6 of 10 nucleotides with the PAR consensus sequence. Thus it could be expected that PAR family proteins would not recognize the CRE consensus sequence. I propose the EMSA experiments utilizing nuclear extracts and the *Aavg* and *Aavcp* CREs as probes should be performed. If these CREs were to bind proteins from nuclear extracts during the YP synthesis stage of vitellogenesis, and these complexes were shown not to contain CREB by the use of anti-CREB antibodies in the binding reaction, this would strengthen the

hypothesis that AaCREB acts as a repressor to terminate vitellogenesis. Regardless, EMSA with nuclear extracts covering more time points than was used in previous experiments, for example every 6 hours as opposed to every 12 hours PBM, should be performed to get a more accurate profile of AaCREB binding activity.

PAR Consensus: RTTAYGTAAY
VgCRE: **AA**TA**CGT**C**AT**
VCPCRE2: **ATTCCGTCAT**
VCPCRE1: **AATACGTCAG**
SomCRE: CT**GACGTCAG**

Figure 5.1, Comparison of mosquito CREs to the PAR consensus. The mosquito Vg and VCP CREs , and the SomCRE used as a probe in EMSA experiments, with the PAR family consensus sequence (Haas *et al.*, 1995; Falvey *et al.*, 1996). Matches to the PAR sequence are shown in bold.

Generating AaCREB antibodies.

EMSA with nuclear extracts detected binding with the SomCRE probe only between 24 and 36 hours PBM. However, it should be remembered that this is more accurately an indication on the functional ability of DNA-binding and not a true reflection of the protein expression profile. In addition, EMSA was not able to conclusively establish the identity of the proteins in the complex, though the use of anti-PhosphoCREB antibodies indicates that the proteins are CREB related. the antibodies used in this study were purchased commercially. Western analysis showed that they recognized *in vitro* expressed AaCREB with high specificity (figure 4.5 and data not shown). However, several additional higher molecular weight bands were detected when nuclear extracts were used.

There was no correlation between these bands and the binding detected with nuclear extracts. Since EMSA utilized two methods for detecting CREB proteins, recognition of the DNA probe by the binding proteins and recognition of these proteins by the antibody, I believe that the shifts accurately reflect binding by CREB and not other unrelated proteins. A comparison of the protein profile during vitellogenesis with DNA-binding ability by EMSA may provide further insights on the regulation of CREB activity. For example, identification of CREB proteins by Western analysis during the previtellogenic and YP synthesis stages of vitellogenesis when DNA-binding activity is absent further strengthens the hypothesis that DNA-binding is regulated by post translational modification (i.e. phosphorylation) of the protein. Furthermore, Western analysis with nuclear extracts using anti-AaCREB antibodies may reveal if other CREB isoforms are present (the higher molecular weight proteins detected with the anti-PhosphoCREB antibodies are much larger than any other CREB isoforms described). Additionally, these antibodies, being specific for AaCREB, would be more conclusive in the identification of proteins in the EMSA complexes. Therefore I propose that a GST-AaCREB fusion protein should be constructed for the purpose of antibody production.

REFERENCES

REFERENCES

Abel T, Bhatt R, and Maniatis T (1992) A *Drosophila* CREB/ATF transcriptional activator binds to both fat body- and liver-specific regulatory elements. *Genes Dev.* **6,** 466-480.

Abel T, Michelson AM, and Maniatis T (1993) A *Drosophila* GATA family member that binds to *Adh* regulatory sequences is expressed in the developing fat body. *Development* **119,** 623-633.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF *et al.* (2000) The genome sequence of *Drosophila melanogaster. Science* **287,** 2185-2195.

An W, and Wensink PC (1995a) Three protein binding sites form an enhancer that regulates sex- and fat body-specific transcription of *Drosophila* yolk protein genes. *EMBO J.* **14,** 1221-1230.

An W, and Wensink PC (1995b) Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer. *Genes Dev.* **9,** 256-266.

Andrew DJ, Baig A, Bhanot P, Smolik SM, and Henderson KD (1997) The *Drosophila dCREB-A* gene is required for dorsal/ventral patterning of the larval cuticle. *Development* **124,** 181-193.

Aragay AM, Ruiz-Gomez A, Penela P, Sarnago S, Elorza A, Jimenez-Sainz MC, Mayor F Jr (1998) G protein-coupled receptor kinase 2 (GRK2): mechanisms of regulation and physiological functions. *FEBS Lett.* **430,** 37-40.

Arbeitman MN, and Hogness DS (2000) Molecular chaperones activate the *Drosophila* ecdysone receptor, an RXR heterodimer. *Cell* **101,** 67-77.

Armstrong SA, Barry DA, Leggett RW, and Mueller CR (1997) Casein kinase II-mediated phosphorylation of the C terminus of Sp1 decreases its DNA binding activity. *J. Biol. Chem.* **272,** 13489-13495.

Ashburner M, Chihara C, Meltzer P, and Richards G (1974) Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* **38,** 655-662.

Aultman KS, Walker ED, Gifford F, Severson DW, Beard CB, and Scott W (2000) Managing risks of arthropod vector research. *Science* **288,** 2321-2322.

Ausubel FM. Brent R, Kingston RE, Moore DD, Seidman JD, Smith JA, and Struhl K (1989) *Current Protocols in Molecular Biology, Vol. 1.* Greene Publishing associates and Wiley-Interscience, New York.

Baniahmad A, Steiner C, Kohne AC, and Renkawitz R (1990) Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. *Cell* **61,** 505-514.

Barillas-Mury CV, Noriega FG, and Wells MA (1995) Early trypsin activity is part of the signal transduction system that activates transcription of the late trypsin gene in the midgut of the mosquito *Aedes aegypti*. *Insect Biochem. Molec. Biol.* **25,** 241-246.

Beato M, Herrlich P, and Schültz G (1995) Steroid hormone receptors: many actors in search of a plot. *Cell* **83,** 851-857.

Beckemeyer EF, and Lea AO (1980) Induction of follicle separation in the mosquito by physiological amounts of ecdysterone. *Science* **209,** 819-821.

Behan M, and Hagedorn HH (1978) Ultrastructural changes in the fat body of adult female *Aedes aegypti* in relationship to vitellogenin synthesis. *Cell Tissue Res.* **186,** 499-506.

Belvin MP, and Yin JCP (1997) *Drosophila* learning and memory: recent progress and new approaches. *Bioessays* **19,** 1083-1089.

Belvin MP, Zhou H, and Yin JC (1999) The *Drosophila dCREB2* gene affects the circadian clock. *Neuron* **22,** 777-787.

Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H (1990) The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61,** 49-59.

Bigler J, Hokanson W, Eisenman RN (1992) Thyroid hormone receptor transcriptional activity is potentially autoregulated by truncated forms of the receptor. *Mol Cell Biol.* **12,** 2406-2417.

Blanchard KL, Acquaviva AM, Galson DL, and Bunn HF (1992) Hypoxic induction of the human *erythropoietin* gene: cooperation between the promoter and enhancer, each of which contains steroid receptor response elements. *Mol Cell Biol.* **12,** 5373-5385.

Bohm MK, Behan M, and Hagedorn HH (1978) Termination of vitellogenin synthesis by mosquito fat body: a programmed response to ecdysterone. *Physiol. Ent.* **3,** 17-25.

Borovsky D (1981) Feedback regulation of vitellogenin synthesis in *Aedes aegypti* and *Aedes atropalpus*. *Insect Physiol.* **2,** 333-349.

Borovsky D (1988) Oostatic hormone inhibits biosynthesis of midgut proteolytic enzymes and egg development in mosquitoes. *Arch. Insect Biochem. Physiol.* **7,** 187-210.

Borovsky D, Carlson DA, Griffin PR, Shabanowitz J, and Hunt DF (1993) Sequencing and characterization of trypsin modulating oostatic factor. *In* Borovsky D and Spielman A (eds:) *Proceedings of the Third Symposium on Host Regulated Developmental Mechanisms in Vector Arthropods,* Vero Beach, Florida, February 1993. Institute of Food and Agricultural Sciences, University of Florida, Vero Beach. pp. 36-47.

Boshart M, Weih F, Nichols M, and Schütz G (1991) The tissue-specific extinguisher locus TSE1 encodes a regulatory subunit of cAMP-dependent protein kinase. *Cell* **66,** 849-859.

Boyle WJ, Smeal T, Defize LH, Angel P, Woodgett JR, Karin M, and Hunter T (1991) Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. *Cell* **64,** 573-584.

Briegel H, and Lea AO (1975) Relationship between protein and proteolytic activity in the midgut of mosquitoes. *J. Insect Physiol.* **21,** 1597-1604.

Brindle P, Linke S, and Montminy M (1993) Protein-kinase-A-dependent activator in transcription factor CREB reveals new role for CREM repressors. *Nature* **364,** 821-824.

Bullock BP, and Habener JF (1998) Phosphorylation of the cAMP response element binding protein CREB by cAMP-dependent protein kinase A and glycogen synthase kinase-3 alters DNA-binding affinity, conformation, and increases net charge. *Biochemistry* **37,** 3795-3809.

Burch JBE, and Davis DL (1994) Alternative promoter usage and splicing options result in the differential expression of mRNAs encoding for isoforms of chicken VBP, a member of the PAR subfamily of bZIP transcription factors. *Nucleic Acids Res.* **22,** 4733-4714.

Busvine JR (1993) *Disease transmission by insects: its discovery and 90 years of effort to prevent it.* Springer-Verlag Press. New York.

Capovilla ME, Eldon ED, and Porrotta V (1992) The *giant* gene of *Drosophila* encodes a b-ZIP DNA-binding protein that regulates the expression of other segmentation gap genes. *Development* **114,** 99-112.

Capurro MdeL, De Bianchi AG, and Marinotti O (1994) *Aedes aegypti* lipophorin. *Comp. Biochem. Physiol.* **108B,** 35-39.

Cardinaux J-R, Chapel S, and Wahli W (1994) Complex organization of CTF/NF-I, C/EBP, and HNF3 binding sites within the promoter of the liver-specific vitellogenin gene. *J. Biol. Chem.* **269,** 32947-32956.

Cardinaux J-R, Notis JC, Shang Q, Vo N, Craig JC, Fass DM, Brennan RG, and Goodman RH (2000) Recruitment of CREB binding protein is sufficient for CREB-mediated gene activation. *Mol. Cell. Biol.* **20,** 1546-1552.

Carlberg C, Bendik I, Wyss A, Meier E, Sturzenbecker LJ, Grippo JF, and Hunziker W (1993) Two nuclear signaling pathways for vitamin D. *Nature* **361,** 657-660.

Cereghini S (1996) Liver-enriched transcription factors and hepatocyte differentiation. *FASEB J.* **10,** 267-282.

Chakravarti D, LaMorte VJ, Nelson MC, Nakajima T, Schulman IG, Juguilon H, Montminy M, and Evans R (1996) Role of CBP/p300 in nuclear receptor signaling. *Nature* **383,** 99-103.

Chapman RF (1998) *The insects: structure and function* - 4th ed. Cambridge University Press, Cambridge, UK.

Chartier FL, Bossu J-P, Laudet V, Fruchart J-C, and Laine B (1994) Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147,** 269-272.

Chen CN, Denome S, and Davis RL (1986) Molecular analysis of cDNA clones and the corresponding genomic coding sequences of the *Drosophila dunce* gene, the structural gene for cAMP phosphodiesterase. *Proc. Natl. Acad. Sci. USA* **83,** 9313-9317.

Chen J-S, and Raikhel AS (1996) Subunit cleavage of mosquito pro-vitellogenin by a subtilisin-like convertase. *Proc. Natl. Acad. Sci. USA.* **93,** 6186-6190.

Chen WJ, Lewis KS, Chandra G, Cogswell JP, Stinnet SW, Kadwell SH, and Gray JG (1995) Characterization of human E4BP4, a phosphrylated bZIP factor. *Biochim. Biophys. Acta.* **1264,** 388-396.

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, and Darnell JE Jr (1994) Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes Dev.* **8,** 2466-2477.

Cho W-L, and Raikhel AS (1992) Cloning of cDNA for mosquito lysosomal aspartic protease. *J. Biol. Chem.* **267,** 21823-21829.

Cho W-L, Deitsch KW, and Raikhel AS (1991a) Extraovarian protein accumulated in mosquito oocytes is a serine carboxypeptidase activated in embryos. *Proc. Natl. Acad. Sci. USA.* **88,** 1082-10824.

Cho W-L, Dhadialla TS, and Raikhel AS (1991b) Purification and characterization of a lysosomal aspartic protease with cathepsin D activity from the mosquito. *Insect Biochem.* **21,** 165-176.

Cho W-L, Kapitskaya MZ, and Raikhel AS (1995) Mosquito ecdysteroid receptor: analysis of the cDNA and expression during vitellogenesis. *Insect Biochem. Molec Biol.* **25,** 19-27.

Cho W-L, Tsao S-M, Hays AR, Walter R, Chen J-S, Snigirevskaya ES, and Raikhel AS (1999) Mosquito cathepsin B-like protease involved in embryonic degradation of vitellin is produced as a latent extraovarian precursor. *J. Biol. Chem.* **274,** 13311-13321.

Chomczynski P, and Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162,** 156-159.

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, and Goodman RH (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365,** 855-859.

Cohen P, and Cohen PT (1989) Protein phosphatases come of age. *J. Biol. Chem.* **264,** 21435-21438.

Collingwood TN, Urnov FD, and Wolffe AP (1999) Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. *J. Mol. Endocrin.* **23,** 255-275.

Conti M, Nemoz G, Sette C, and Vicini E (1995) Recent progress in understanding the hormonal regulation of phosphodiesterases. *Endocr. Rev.* **16,** 370-389.

Cooney AJ, Tsai SY, O'Malley BW, Tsai M-J (1992) Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. *Mol Cell. Biol.* **12,** 4153-4163.

Costa RH, Grayson DR, and Darnell JE Jr (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α1-antitrypsin genes. *Mol. Cell. Biol.* **9,** 1415-1425.

Costa RH, Van Dyke TA, Yan C, Kuo F, and Darnell JE Jr (1990) Similarities in transthyretin gene expression and differences in transcription factors: liver and yolk sac compared to choroid plexus. *Proc. Natl. Acad. Sci. USA* **87,** 6589-6593.

Courey AJ, and Tjian R (1988) Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55,** 887-898.

Daniel PB, Walker WH, and Habener JF (1998) Cyclic AMP signaling and gene regulation. *Annu. Rev. Nutr.* **18,** 353-383.

Danielian PS, White R, Lees JA, and Parker MG (1992) Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J.* **11,** 1025-1033.

De Cesare D, and Sassone-Corsi P (2000) Transcriptional regulation by cyclic AMP-responsive factors. *Prog. Nucleic Acid Res. Mol. Biol.* **64,** 343-369.

De Cesare D, Fimia GM, and Sassone-Corsi P (1999) Signaling routes to CREM and CREB: plasticity in transcriptional activation. *Trends Biochem. Sci.* **24,** 281-285.

Deitsch KW, and Raikhel AS (1993) Cloning and analysis of the locus for mosquito vitellogenic carboxypeptidase. *Insect Molec. Biol.* **2,** 205-213.

Deitsch KW, Chen J-S, and Raikhel AS (1995a) Indirect control of yolk protein genes by 20-hydroxyecdysone in the fat body of the mosquito, *Aedes aegypti. Insect Biochem. Molec. Biol.* **25,** 449-454.

Deitsch KW, Dittmer NT, Kapitskaya MZ, Chen J-S, Cho W-L, and Raikhel AS (1995b) Regulation of gene expression by 20-hydroxyecdysone in the fat body of *Aedes aegypti* (Diptera: Culicidae). *Eur. J. Entomol.* **92,** 237-244.

Detera-Wadleigh SD, and Fanning TG (1994) Phylogeny of the steroid receptor superfamily. *Mol. Phylogenet. Evol. 3,* 192-205.

Dhadialla TS, and Raikhel AS (1990) Biosynthesis of mosquito vitellogenin. *J. Biol. Chem.* **265,** 9924-9933.

Dhadialla TS, and Raikhel AS (1994) Endocrinology of mosquito vitellogenesis. *In* Davey KG, Peter RE and Tobe SS (eds): *Perspectives in Comparative Endocrinology.* National Research Council of Canada. pp 275-281.

Dittman F, Kigan PH, and Hagedorn HH (1989) Ploidy levels and DNA synthesis in the fat body cells of the adult mosquito, *Aedes aegypti*: The role of juvenile hormone. *Arch. Insect Biochem. Physiol.* **12,** 133-143.

Dittmer NT, and Raikhel, AS (1997) Analysis of the mosquito lysosomal aspartic protease gene: an insect housekeeping gene with fat body-enhanced expression. *Insect Biochem. Molec. Biol.* **27,** 323-335.

Drain P, Flokers E, and Quinn WG (1991) cAMP-dependent protein kinase and the disruption of learning in transgenic flies. *Neuron* **6,** 71-82.

Drewes T, Senkel S, Holewa B, and Ryffel GU (1996) Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. *Mol. Cell. Biol.* **16,** 925-931.

Drolet DW, Scully KM, Simmons DM, Wegner M, Chu K, Swanson LW, and Rosenfeld MG (1991) TEF, a transcription factor expressed specifically in the anterior pituitary during embryogenesis, defines a new class of leucine zipper proteins. *Genes Dev.* **5,** 1739-1753.

Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, and Darnell JE Jr (1994) Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proc. Natl. Acad. Sci. USA* **91,** 7598-7609.

Duncan SA, Nagy A, and Chan W (1997) Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of *Hnf-4*$^{-/-}$ embryos. *Development* **124,** 279-287.

Durand B, Saunders M, Gaudon C, Roy B, Losson R, Chambon P (1994) Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of response element on AF-2 activity. *EMBO J.* **13,** 5370-5382.

Falb D, and Maniatis T (1992a) *Drosophila* transcriptional repressor protein that binds specifically to negative control elements in fat body enhancers. *Mol. Cell. Biol.* **12,** 4093-4103.

Falb D, and Maniatis T (1992b) A conserved regulatory unit implicated in tissue-specific gene expression in *Drosophila* and man. *Genes Dev.* **6,** 454-465.

Falvey E, Marcacci L, and Schibler U (1996) DNA-binding specificity of PAR and C/EBP leucine zipper proteins: a single amino acid substitution in the C/EBP DNA-binding domain confers PAR-like specificity to C/EBP. *Biol. Chem.* **377,** 797-809.

Fantozzi DA, Harootunian AT, Wen W, Taylor SS, Feramisco JR, Tsien RY, Meinkoth JL. (1994) Thermostable inhibitor of cAMP-dependent protein kinase enhances the rate of export of the kinase catalytic subunit from the nucleus. *J Biol Chem.* **269,** 2676-2686.

Felix CR, Betschart B, Billingsley PF, and Freyvogel TA (1991) Post-feeding induction of trypsin in the midguts of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochem.* **21,** 197-203.

Ferreri K, Gill G, and Montminy M (1994) The cAMP-regulated transcription factor CREB interacts with a component of the TFIID complex. *Proc. Natl. Acad. Sci. USA* **91,** 1210-1213.

Ford PS, and Van Heusden MC (1994) Triglyceride-rich lipophorin in *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **31,** 435-431.

Forman BM, and Samuels HH.(1990) Interactions among a subfamily of nuclear hormone receptors: the regulatory zipper model. *Mol. Endocrinol.* **4,** 1293-1301.

Foulkes NS, and Sassone-Corsi P (1992) More is better: activators and repressors from the same gene. *Cell* **68,** 411-414.

Foulkes NS, Borrelli E, and Sassone-Corsi P. (1991) CREM gene: use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. *Cell* **64,** 739:749.

Freedman NJ, Lefkowitz RJ (1996) Desensitization of G protein-coupled receptors. *Recent Prog. Horm. Res.* **51,** 319-351.

Friedman S (1985) The fat body as a metabolic source. *In* Blum MS (ed:) *Fundamentals of insect physiology*. John Wiley & Sons, Inc. New York. pp 471-571.

Fuchs MS, and Kang S-H (1981) Ecdysone and mosquito vitellogenesis: a critical appraisal. *Insect Biochem.* **11,** 627-633.

Furuta H, Iwasaki N, Oda N, Hinokio Y, Horikawa Y, Yamagata K, Yano N, Sugahiro J, Ogata M, Ohgawara H, Omori Y, Iwamoto Y, and Bell GI (1997) Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* **46,** 1652-1657.

Gemmill RM, Hamblin M, Glaser RL., Racioppi JV, Marx JL., White BN., Calvo JM, Wolfner MR, and Hagedorn HH (1986) Isolation of mosquito vitellogenin genes and induction of expression by 20-hydroxyecdysone. *Insect Biochem.* **16,** 761-774.

Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, and Collins T (1997) CREB-binding protein/p300 are transcriptional coactivators of p65. *Proc. Natl. Acad. Sci. USA* **94,** 2927-2932.

Glass CK, Rose DW, and Rosenfeld MG (1997) Nuclear receptor coactivators. *Curr. Opin. Cell Biol.* **9,** 222-232.

Goldman PS, Tran VK, and Goodman RH (1997) The multifunctional role of the co-activator CBP in transcriptional regulation. *Recent Prog. Horm. Res.* **52,** 103-119.

Gonzalez GA, and Montiminy MR (1989) Cyclic AMP stimulates *somatostatin* gene transcription by phosphorylation of CREB at serine 133. *Cell* **59,** 675-680.

Gooding RH (1972) Digestive processes of haematophagous insects. I. A literature review. *Quaest Entomol.* **8,** 5-60.

Graf R, and Briegel (1989) The synthetic pathway of trypsin in the mosquito, *Aedes aegypti* L. (Diptera: Culicidae) and *in vitro* stimulation in isolated midguts. *Insect Biochem.* **19,** 129-137.

Grossniklaus U, Pearson RK and Gehring WJ (1992) The *Drosophila sloppy paired* locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes Dev.* **6,** 1030-1051.

Habener JF (1995) Cyclic AMP second messenger signaling pathway. *In* DeGrood LJ (ed:) *Endocrinology.* Saunders - 3rd edition. Philadelphia, PA, pp. 77-92.

Hacker U, Kaufmann E, Hartmann C, Jurgens G, Knochel W, and Jackle H (1995) The *Drosophila* fork head domain protein crocodile is required for the establishment of head structures. *EMBO J.* **14,** 5306-5317.

Hadzopoulou-Cladaras M, Kistanova E, Evagelopoulou C, Zeng S, Cladaras C, and Ladias JA (1997) Functional domains of the nuclear receptor hepatocyte nuclear factor 4. *J. Biol. Chem.* **272,** 539-550.

Hagedorn HH (1985) The role of ecdysteroids in reproduction. *In* Kerkut GA, and Gilbert LI (eds): *Comprehensive insect Physiology, Biochemistry, and Pharmacology. Vol. 8.* Pergamon Press. Oxford. pp 205-261.

Hagedorn HH (1989) Physiological roles of hemolymph ecdysteroids in the adult insect. *In* Koolman J (ed): *Ecdysone from Chemistry to Mode of Action.* Thieme Medical Publishers, Inc. New York. pp 279-289.

Hagedorn HH, Fallon AM, and Laufer H (1973) Vitellogenin synthesis by the fat body of *Aedes aegypti. Develop. Biol.* **31,** 285-294.

Hagedorn HH, O'Connor JD, Fuchs MS, Sage B, Schlaeger DA, and Bohm MK (1975) The ovary as a source of α-ecdysone in an adult mosquito. *Proc. Natl. Acad. Sci. USA.* **72,** 3255-3259.

Hagedorn HH, Shapiro JP, and Hanaoka K (1979) Ovarian ecdysone secretion is controlled by a brain hormone in an adult mosquito. *Nature* **282,** 92-94.

Hai TW, Liu F, Coukos WJ, and Green MR (1989) Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes Dev.* **3,** 2083-2090. Erratum (1990) *Genes Dev.* **4,** 682.

Hamblin MT, Marx JL, Wolfner MF, and Hagedorn HH (1987) The vitellogenin gene family of *Aedes aegypti. Mem. Inst. Oswaldo Cruz* **82,** 109-114.

Hanoune J, Pouille Y, Tzavara E, Shen T, Lipskaya L, Miyamoto N, Suzuki Y, and Defer N (1997) Adenylyl cyclases: structure, regulation and function in an enzyme superfamily. *Mol. Cell. Endocrinol.* **128,** 179-194.

Hartfelder K (2000) Insect juvenile hormone: from "status quo" to high society. *Braz. J. Med. Biol. Res.* **33,** 157-177.

Haas NB, Cantwell CA, Johnson PF, and Burch JBE (1995) DNA-binding specificity of the PAR basic leucine zipper protein VBP partially overlaps those of the C/EBP and CREB/ATF families and is influenced by domains that flank the core basic region. *Mol. Cell. Biol.* **15,** 1923-1932.

Hata S, Tsukamoto T, and Osumi T (1992) A novel isoform of rat hepatocyte nuclear factor 4 (HNF-4). *Biochim Biophys Acta.* **1131,** 211-213.

Hata S, Inoue T, Kosuga K, Nakashima T, Tsukamoto T, and Osumi T (1995) Identification of two splice isoforms of mRNA for mouse hepatocyte nuclear factor 4 (HNF-4). *Biochim Biophys Acta.* **1260,** 55-61.

Hays AR, and Raikhel AS (1990) A novel protein produced by the vitellogenic fat body and accumulated in mosquito oocytes. *Roux's Arch. Dev. Biol.* **199,** 114-121.

Hecker H, Freyvogel TA, Briegel H, and Steiger R (1971) Ultrastructural differentiation of the midgut epithelium in female *Aedes aegypti* (L.) (Insecta, Diptera) imagines. *Acta Tropica* **28,** 80-104.

Henrich VC, and Brown NE (1995) Insect nuclear receptors: a developmental and comparative perspective. *Insect. Biochem. Mol. Biol.* **25,** 881-897.

Hertz R, Magenheim J, Berman I, and Bar-Tana J (1998) Fatty acyl-CoA thioesters are ligands of hepatic nuclear factor-4α. *Nature* **392,** 512-516.

Hoeffler JP, Meyer TE, Yun Y, Jameson JL, and Habener JF (1988) Cyclic AMP responsive DNA-binding protein: structure based on a cloned placental cDNA. *Science* **242,** 1430-1433.

Hoeffler JP, Meyer TE, Waeber G, and Habener JF (1990) Multiple adenosine 3',5'-cyclic monophosphate response element DNA-binding proteins generated by gene diversification and alternative exon splicing. *Mol Endocrinol.* **4,** 920-930.

Holewa B, Stradmann EP, Zapp D, Lorenz P, and Ryffel GU (1996) Transcriptional hierarchy in *Xenopus* embryogenesis: HNF4 a maternal factor involved in the developmental activation of the gene encoding the tissue specific transcription factor HNF1α (LFB1). *Mech. Dev.* **54,** 45-57.

Holewa B, Zapp D, Drewes T, Senkel S, and Ryffel GU (1997) HNF4β, a new gene of the HNF4 family with distinct activation and expression profiles in oogenesis and embryogenesis of *Xenopus laevis*. *Mol. Cell. Biol.* **17,** 687-694.

Horwitz KB, Jackson TA, Bain DL, Richer JK, Takimoto GS, and Tung L (1996) Nuclear receptor coactivators and corepressors. *Mol. Endo.* **10,** 1167-1177.

Houslay MD, and Milligan G (1997) Tailoring cAMP-signaling responses through isoform multiplicity. *Trends Biochem. Sci,* **22,** 217-224.

Hsieh JC, Jurutka PW Nakajima S, Galligan MA, Haussler CA, Shimizu Y, Shimizu N, Whitfield GK, and Haussler MR (1993) Phosphorylation of the human vitamin D receptor by protein kinase C. Biochemical and functional evaluation of the serine 51 recognition site. *J. Biol. Chem.* **268,** 15118-15126.

Hunter T and Karin M (1992) The regulation of transcription by phosphorylation. *Cell* **70,** 375-387.

Hurst HC (1994) Transcription factors 1, bZIP proteins. *Protein profiles* **1,** 123-168.

Iyanger R (1993) Molecular and functional diversity of mammalian Gs-stimulated adenylyl cyclases. *FASEB J.* **7,** 768-775.

Iyemere VP, Davies NH, and Brownlee GG (1998) The activation function 2 domain of hepatic nuclear factor 4 is regulated by a short C-terminal proline-rich repressor domain. *Nucleic Acids Res.* **26,** 2098-2104.

Iyer SV, Davis DL, Seal SN, and Burch JBE (1991) Chicken vitellogenin gene-binding protein, a leucine zipper transcription factor that binds to an important control element in the chicken vitellogenin II promoter, is related to rat DBP. *Mol. Cell. Biol.* **11,** 4863-4875.

Janknecht R, and Hunter T (1996) A growing coactivator network. *Nature* **383,** 22-23.

Janknecht R, Hipskind RA, Houthaeve T, Nordheim A, and Studdenberg HG (1992) Identification of multiple SRF N-terminal phosphorylation sites affecting DNA binding properties. *EMBO J.* **11,** 1045-1054.

Jiang G, Nepomuceno L, Yang Q, and Sladek FM (1997) Serine/threonine phosphorylation of orphan receptor hepatocyte nuclear factor 4. *Arch. Biochem. Biophys.* **340,** 1-9.

Johnson PF (1993) Identification of C/EBP basic region residues involved in DNA sequence recognition and half-site spacing preference. *Mol. Cell. Biol.* **13,** 6919-6930.

Jones G (1995) Molecular mechanisms of action of juvenile hormone. *Annu. Rev. Entomol.* **40,** 147-169.

Jones G and Sharp PA (1997) Ultraspiracle: An invertebrate nuclear receptor for juvenile hormones. *Proc. Natl. Acad. Sci. USA* **94,** 13499-13503.

Kalhok SE, Tabak LM, Prosser DE, Brook W, Downe AER, and White BN (1993) Isolation, sequencing and characterization of two cDNA clones coding for trypsin-like enzymes from the midgut of *Aedes aegypti. Insect Molec. Biol.* **2,** 71-79.

Kan S-P, and Ho B-C (1972) Development of *Breinlia sergenti* (Dipetalonematidae) in the fat-body of mosquitoes. II. Ultrastructural changes in the fat body. *J. Med. Entomol.* **9,** 255-261.

Kapitskaya MZ, Wang S, Cress DE, Dhadialla TS, and Raikhel AS (1996) The mosquito *ultraspiracle* homologue, a partner of ecdysteroid receptor heterodimer: cloning and characterization of isoforms expressed during vitellogenesis. *Molec. Cell. Endocrin.* **121,** 119-132.

Kapitskaya MZ, Dittmer NT, Deitsch KW, Cho W-L, Taylor DG, Leff T, and Raikhel AS (1998) Three isoforms of a hepatocyte nuclear factor-4 transcription factor with tissue- and stage-specific expression in the adult mosquito. *J. Biol. Chem.* **273,** 29801-29810.

Kapitskaya MZ, Li C, Miura K, Segraves W, and Raikhel AS (2000) Expression of the early-late gene encoding the nuclear receptor HR3 suggest its involvement in regulating the vitellogenic response to ecdysone in the adult mosquito. *Mol. Cell. Endocrinol.* **160,** 25-37.

Kardassis D, Sacharidou E, and Zannis VI (1998) Transactivation of the human *apolipoprotein CII* promoter by orphan and ligand-dependent nuclear receptors. *J. Biol. Chem.* **273,** 17810-17816.

Kastner P, Mark M, Chambon P (1995) Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* **83,** 859-69.

Kato S, Sasaki H, Suzawa M, Masushige S, Tora L, Chambon P, and Gronemeyer H (1995) Widely spaced, directly repeated PuGGTCA elements act as promiscuous enhancers for different classes of nuclear receptors. *Mol. Cell. Biol.* **15,** 5858-5867.

Kaziro Y, Itoh H, Kozasa T, Nakafuku M, and Satoh T (1991) Structure and function of signal-transducing GTP-binding proteins. *Annu. Rev. Biochem,* **60,** 349-400.

Kee BI, Arias J, and Montminy MR (1996) Adaptor-mediated recruitment of RNA polymerase II to a signal-dependent activator. *J. Biol. Chem.* **271,** 2373-2375.

Kelly TJ, Masler PE, Schwartz MB. and Haught SB (1986) Inhibitory effects of oostatic hormone on ovarian maturation and ecdysteroid production in Diptera. *Insect Biochem.* **16,** 273-279.

Koelle MR, Talbot WS, Segraves WA, Bender MT, Cherbas P, Hogness DS (1991) The *Drosophila EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67,** 59-77.

Koller CN and Raikhel AS (1991) Initiation of vitellogenin uptake and protein synthesis in the mosquito (*Aedes aegypti*) ovary in response to a blood meal. *J. Insect Physiol.* **37,** 703-711.

Koller CN, Dhadialla TS, and Raikhel AS (1989) Selective endocytosis of vitellogenin by oocytes of the mosquito, *Aedes aegypti*: an *in vitro* study. *Insect Biochem.* **19,** 693-702.

Krasnow MA, Saffman EE, Kornfeld K, and Hogness DS (1989) Transcriptional activation and repression by Ultrabithorax proteins in cultured *Drosophila* cells. *Cell* **57,** 1031-1043.

Kritis AA, Argyrokastritis A, Moschonas NK, Power S, Katrakili N, Zannis VI, Cereghini S, and Talianidis I (1996) Isolation and characterization of a third isoform of human hepatocyte nuclear factor 4. *Gene*. **173,** 275-280.

Ktistaki E, Ktistaki NT, Papadogeorgaki E, and Talianidis I (1995) Recruitment of hepatocyte nuclear factor 4 into specific intranuclear compartments depends on tyrosine phosphorylation that affects its DNA-binding and transactivation potential. *Proc. Nat. Acad. Sci. USA* **92,** 9876-9880.

Ladias JA, Hadzopoulou-Cladaras M, Dardassis D, Cardot P, Cheng J, Zannis V, and Cladaras C (1992) Transcriptional regulation of human apolipoprotein genes *apoB*, *apoCIII*, and *apoAII* by members of the steroid hormone receptor superfamily HNF-4, ARP-1, EAR-2, and EAR-3. *J Biol. Chem*. **267,** 15849-15860.

Lai E, and Darnell JE Jr (1991) Transcriptional control in hepatocytes: a window on development. *Trends Biochem. Sci.* **16,** 427-430.

Lai E, Clark KL, Burley SK, and Darnell JE Jr (1993) Hepatocyte nuclear factor3/fork head or "winged helix" proteins: a family of transcription factors of diverse biologic function. *Proc. Natl. Acad. Sci. USA* **90,** 10421-10423.

Landschulz WH, Johnson PF, and McKnight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240,** 1756-1764.

Laoide BM, Foulkes NS, Schlotter F, and Sassone-Corsi P (1993) The functional versatility of CREM is determined by its modular structure. *EMBO J.* **12,** 1179-1191.

Laudet V (1997) Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J. Mol. Endocrin.* **19,** 207-226.

Lea AO (1967) The medial neurosecretory cells and egg maturation in mosquitoes. *J. Insect Physiol.* **13,** 419-429.

Lea AO (1972) Regulation of egg maturation in the mosquito by the neurosecretory system: the role of the corpus cardiacum. *Gen. Comp. Endocrinol.* **Supp. 3,** 602-608.

Lea AO (1982) Artifactual stimulation of vitellogenesis in *Aedes aegypti* by 20-hydroxyecdysone. *J. Insect Physiol.* **28,** 173-176.

Lea AO, and Brown MR (1990) Neuropeptides of mosquitoes. *In* Hagedorn HH, Hildebrand JG, Kidwell MG, and Law JH (eds:) *Molecular Insect Science*. Plenum Press, New York, pp. 147-154.

Lea AO, and Van Handel E (1982) A neurosecretory hormone-releasing factor from ovaries of mosquitoes fed blood. *J. Insect Physiol.* **28,** 503-508.

Lefkowitz RJ (1998) G protein-coupled receptors. *J. Biol. Chem.* **273,** 18677-18680.

Levin LR, Han PL, Hwang PM, Feinstein PGk, Davis RL, and Reed RR (1992) The *Drosophila* learning and memory gene *rutabaga* encodes a Ca2+/calmodulin-responsive adenylyl cyclase. *Cell* **68,** 479-489.

Li J, Ning G, and Duncan SA (2000) Mammalian hepatocyte differentiation requires the transcription factor HNF-4. *Genes Dev.* **14,** 464-474.

Lin A, Frost J, Deng T, Smeal T, al-Alawi N, Kikkawa U, Hunter T, Brenner D, and Karin M (1992) Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. *Cell* **70,** 777-789.

Mach V, Takiya S, Ohno K, Handa H, Imai T, and Suzuki Y (1995) Silk gland factor-1 involved in the regulation of *Bombyx sericin-1* gene contains fork head motif. *J. Biol. Chem.* **270,** 9340-9346.

Mangelsdorf DJ, and Evans RM (1995) The RXR heterodimers and orphan receptors. *Cell* **83,** 841-850.

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, and Evans RM (1995) The nuclear receptor superfamily: the second decade. *Cell* **83,** 835-839.

Meier CA (1997) Regulation of gene expression by nuclear hormone receptors. *J. Recept. Signal Transduct. Res.* **17,** 319-335.

Meola R, and Lea AO (1972) Humoral inhibition of egg development of mosquitoes. *J. Med. Entomol.* **9,** 99-103.

Metzger S, Halaas JL, Breslow JL, and Sladek FM (1993) Orphan receptor HNF-4 and bZip protein C/EBPα bind to overlapping regions of the *apolipoprotein B* gene promoter and synergistically activate transcription. *J. Biol. Chem.* **268,** 16831-16838.

Meyer TE, and Habener JF (1992) Cyclic AMP response element binding protein CREB and modulator protein CREM are products of distinct genes. *Nucleic Acids Res.* **20,** 6106.

Meyer TE, and Habener JF (1993) Cyclic adenosine 3',5'-monophosphate response element binding protein (CREB) and related transcription-activating deoxyribonucleic acid-binding proteins. *Endo. Reviews.* **14,** 269-290.

Meyer TE, Waeber G, Beckmann w, Lin J, and Habener JF (1993) The promoter of the gene encoding cAMP-response element binding protein CREB contains cAMP response elements: evidence for positive autoregulation of gene transcription. *Endocrinology* **132,** 770-780.

Mietus-Snyder M, Sladek FM, Ginsburg GS, Kuo CF, Ladias JA, Darnell JE Jr, and Karathanasis SK (1992) Antagonism between apolipoprotein AI regulatory protein 1, Ear3/COUP-TF, and hepatocyte nuclear factor 4 modulates *apolipoprotein CIII* gene expression in liver and intestinal cells. *Mol. Cell. Biol.* **12,** 1708-1718.

Miura K, Wang S-F, and Raikhel AS (1999) Two distinct subpopulations of ecdysone receptor complex in the female mosquito during vitellogenesis. *Mol. Cell. Endo.* **156,** 111-120.

Molina CA, Foulkes NS, Lalli E, and Sassone-Corsi P (1993) Inducibility and negative autoregulation of CREM: an alternative promoter directs the expression of ICER, an early response repressor. *Cell* **75,** 875-886.

Montano MM, Muller V, Trobaugh A, and Katzenellenbogen BS (1995) The carboxy-terminal F domain of the human estrogen receptor: role in the transcriptional activity of the receptor and the effectiveness of antiestrogens as estrogen antagonists. *Mol. Endocrinol.* **9,** 814-825.

Montell DJ, Rørth P, and Spradling AC (1992) *slow border cells*, a locus required for a developmentally regulated cell migration during oogenesis, encodes *Drosophila* C/EBP. *Cell* **71,** 51-62.

Montminy M (1997) Transcriptional regulation by cyclic AMP. *Annu. Rev. Biochem.* **66,** 807-822.

Montminy MR, Sevarino KA, Wagner JA, Mandel G, and Goodman RH. (1986) Identification of a cyclic-AMP-responsive element within the rat *somatostatin* gene. *Proc Natl Acad Sci U S A.* **83,** 6682-6686.

Nakajima T, Uchida C, Anderson SF, Parvin JD, and Montminy M (1997) Analysis of a cAMP-responsive activator reveals a two-component mechanism for transcriptional induction via signal-dependent factors. *Genes Dev.* **11,** 738-747.

Nakhei H, Lingott A, Lemm I, and Ryffel GU (1998) An alternative spliced variant of the tissue specific transcription factor HNF4α predominates in undifferentiated murine cell types. *Nucleic Acids Res.* **26,** 497-504.

Nichols M, Weih F, Schmid W, DeVack C, Kowenz-Leutz E, Luckow B, Boshart M, and Schütz G (1992) Phosphorylation of CREB affects its binding to high and low affinity sites: implications for cAMP induced gene transcription. *EMBO J.* **11,** 3337-3346.

Noriega FG, Barillas-Mury CV, and Wells MA (1994) Dietary control of late trypsin gene transcription in *Aedes aegypti. Insect Biochem. Molec. Biol.* **24,** 627-631.

Noriega FG, Wang X-Y, Pennington JE, Barillas-Mury CV, and Wells MA (1996a) Early trypsin, a female-specific midgut protease in *Aedes aegypti*: isolation, amino-terminal sequence determination, and cloning and sequencing of the gene. *Insect Biochem Molec. Biol.* **26,** 119-126.

Noriega FG, Pennington JE, Barillas-Mury CV, Wang X-Y, and Wells MA (1996b) Early trypsin, an *Aedes aegypti* female specific protease, is post-transcriptionally regulated by the blood meal. *Insect Molec Biol.* **5,** 25-29.

Noriega FG, Shah DK, and Wells MA (1997) Juvenile hormone controls early trypsin gene transcription in the midgut of *Aedes aegypti. Insect Molec. Biol.* **6,** 63-66.

Oro AE, McKeown M, and Evans RM (1990) The *Drosophila* retinoid X receptor homolog ultraspiracle functions in both female reproduction and eye morphogenesis. *Nature* **347,** 298-301.

Parker D, Ferreri K, Nakajima T, LaMorte VJ, Evans R, Koerber SC, Hoeger C, and Montminy MR (1996) Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism. *Mol. Cell. Biol.* **16,** 694-703.

Parker D, Jhala US, Radhakrishnan I, Yaffe MB, Reyes C, Shulman AI, Cantley LC, Wright PE, and Montminy M (1998) Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Mol. Cell* **2,** 353-359.

Pende M, Fisher TL, Simpson PB, Russell JT, Blenis J, and Gallo V (1997) Neurotransmitter- and growth factor-induced cAMP response element binding protein phosphorylation in glial cell progenitors: role of calcium ions, protein kinase C, and mitogen-activated protein kinase/ribosomal S6 kinase pathway. *J Neurosci.* **17,** 1291-1301.

Petzel DH, Berg MM, and Beyenbach KW (1987) Hormone-controlled cAMP-mediated fluid secretion in yellow-fever mosquito. *Am. J. Physiol.* **253 (5 Pt 3),** R701-R711.

Pierceall, WE, Li C, Biran A, Miura K, Raikhel AS, and Segraves WA (1999) E75 expression in mosquito ovary and fat body suggests reiterative use of ecdysone-regulated hierarchies in development and reproduction. *Mol. Cell. Endocrinol.* **150,** 73-89.

Pitcher JA, Freedman NJ, and Lefkowitz RJ (1998) G protein-coupled receptor kinases. *Annu. Rev. Biochem.* **67,** 653-692.

Potter JJ, Mezey E, Christy RJ, Crabb DW, Stein PM, and Yang VW (1991) CCAAT/enhancer binding protein binds and activates the promoter of the rat class I *alcohol dehydrogenase* gene. *Arch Biochem Biophys.* **285,** 246-251.

Potter JJ, Mezey E, and Yang VW (1994) The adult enhancer factor-1, a *Drosophila melanogaster* transcriptional repressor, modulates the promoter activity of the rat class-I *alcohol dehydrogenase*-encoding gene. *Gene* **149,** 325-330.

Quinn PG (1993) Distinct activation domains within cAMP response element-binding protein (CREB) mediate basal and cAMP-stimulated transcription. *J. Biol. Chem.* **268,** 16999-17009.

Radhakrishnan I, Pérez-Alvarado GC, Perker D, Dyson HJ, Montminy MR, and Wright PE (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* **91,** 741-752.

Radhakrishnan I, Pérez-Alvarado GC, Dyson HJ, and Wright PE (1998) Conformational preferences in the Ser^{133}-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett.* **430,** 317-322.

Raikhel AS (1984) The accumulative pathway of vitellogenin in the mosquito oocyte; a high resolution immuno- and cytochemical study. *J. Ultrastr. Res.* **87,** 285-302.

Raikhel AS (1986a) Role of lysosomes in regulating of vitellogenin secretion in the mosquito fat body. *J. Insect Physiol.* **32,** 597-604.

Raikhel AS (1986b) Lysosomes in the cessation of vitellogenin secretion by the mosquito fat body: selective degradation of Golgi complexes and secretory granules. *Tissue Cell.* **18,** 125-142.

Raikhel AS (1992) Vitellogenesis in mosquitoes. *Adv. Dis. Vect. Res.* **9,** 1-39.

Raikhel AS, and Bose SG (1988) Properties of mosquito yolk protein: A study using monoclonal antibodies. *Insect Biochem.* **18,** 565-575.

Raikhel AS, and Dhadialla TS (1992) Accumulation of yolk proteins in insect oocytes. *Annu. Rev. Entomol.* **37,** 217-251.

Raikhel AS, and Lea AO (1983) Previtellogenic development and vitellogenin synthesis in the fat body of a mosquito: an ultrastructural and immumocytochemical study. *Tissue Cell.* **15,** 281-300.

Raikhel AS, and Lea AO (1985) Hormone-mediated formation of the endocytic complex in mosquito oocytes. *Gen. Comp. Endocrinol.* **53,** 424-435.

Raikhel AS, and Lea AO (1990) Juvenile hormone controls previtellogenic proliferation of ribosomal RNA in the mosquito fat body. *Gen. Comp. Endocrinol.* **77,** 423-434.

Raikhel AS, and Lea AO (1991) Control of follicular epithelium development and vitelline envelope formation in the mosquito; role of juvenile hormone and 20-hydroxyecdysone. *Tissue Cell.* **23,** 577-591.

Raikhel AS, Deitsch KW, and Sappington TW (1997) Culture and analysis of the insect fat body. *In* Crampton JM, Beard CB, and Louis C (eds): *Molecular Biology of Insect Disease Vectors: A methods manual.* Chapman and Hall. New York. pp 507-522.

Raikhel AS, Miura K, and Segraves WA (1999) Nuclear receptors in mosquito vitellogenesis. *Amer. Zool.* **39,** 722-735.

Rasmussen R, Benvegnu D, O'Shea EK, Kim PS, and Alber T (1991) X-ray scattering indicates that the leucine zipper is a coiled coil. *Proc. Natl. Acad. Sci. USA* **88,** 561-564.

Rehfuss RP, Walton KM, Loriaux MM, Goodman RH (1991) The cAMP-regulated enhancer-binding protein ATF-1 activates transcription in response to cAMP-dependent protein kinase A. *J Biol Chem.* **266,** 18431-18434.

Ribeiro RC, Kushner PJ, Baxter JD (1995) The nuclear hormone receptor gene superfamily. *Annu. Rev. Med.* **46,** 443-453.

Richards JP, Bächinger HP, Goodman RH, and Brennan RG (1996) Analysis of the structural properties of cAMP-responsive element-binding protein (CREB) and phosphorylated CREB. *J. Biol. Chem.* **271,** 13716-13723.

Riddiford LM (1996) Juvenile hormone: the status of its "status quo" action. *Arch. Insect. Biochem. Physiol.* **32,** 271-286.

Robyr D, Gegonne A, Wolffe AP, and Wahli W (2000) Determinants of vitellogenin B1 promoter architecture: HNF3 and estrogen responsive transcription within chromatin. *J Biol. Chem.* (In press)

Romans P, Tu Z, Ke Z, and Hagedorn HH (1995) Analysis of a vitellogenin gene of the mosquito, *Aedes aegypti* and comparisons to vitellogenins from other organisms. *Insect Biochem Molec. Biol.* **25,** 939-958.

Rørth P, and Montell DJ (1992) *Drosophila* C/EBP: a tissue-specific DNA-binding protein required for embryonic development. *Genes Dev.* **6,** 2299-2311.

Rossignol PA, Spielman A, and Jacobs MS (1982) Rough endoplasmic reticulum in midgut cells of mosquitoes (Diptera: Culicidae): aggregation stimulated by juvenile hormone. *J. Med. Entomol.* **19,** 719-721.

Roth TF and Porter KR (1964) Yolk protein uptake in the oocyte of the mosquito *Aedes aegypti. J. Cell. Biol.* **20,** 313-332.

Sambrook J, Fritsch EF, and Maniatis T (1989) *Molecular Cloning: a laboratory manual, 2nd edition.* Cold Spring Harbor Laboratory Press.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* **74,** 5463-5467.

Sappington TW and Raikhel AS (1995) Receptor-mediated endocytosis of yolk proteins by insect oocytes. *In* Ohnishi E, Sonobe H, and Takahashi SY (eds:) *Recent Advances in Insect Biochemistry and Molecular Biology.* Magoya, Japan. Nagoya University Press. pp 235-257.

Sappington TW and Raikhel AS (1998) Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochem. Mol. Biol.* **28,** 277-300.

Sappington TW, Hays AR, and Raikhel AS (1995) Mosquito vitellogenin receptor: purification, developmental and biochemical characterization. *Insect Biochem. Molec. Biol.* **25,** 807-817.

Sappington TW, Kokoza VA, Cho W-L, and Raikhel AS (1996) Molecular characterization of the mosquito vitellogenin receptor reveals unexpected high homology to the *Drosophila* yolk protein receptor. *Proc. Natl. Acad. Sci. USA.* **93,** 8934-8939.

Sappington TW, Helbling P, Koller CN, Raikhel AS (1998) Activation *in vitro* of vitellogenin uptake by the oocytes of the mosquito, *Aedes aegypti. Physiol. Entom.* **23,** 158-164.

Sassone-Corsi P (1998) Coupling gene expression to cAMP signaling: role of CREB and CREM. *Int. J. Biochem. Cell Biol.* **30,** 27-38.

Segraves WA (1994) Steroid receptors and orphan receptors in *Drosophila* development. *Semin. Cell. Biol.* **5,** 105-113.

Segraves WA, Hogness DS (1990) The E75 ecdysone-inducible gene responsible for the 75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily. *Genes Dev.* **4,** 204-219.

Seol W, Choi HS, and Moore DD (1996) An orphan nuclear hormone receptor that lacks a DNA binding domain and heterodimerizes with other receptors. *Science* **272,** 1336-1339.

Shapiro AB, Wheelock GD, Hagedorn HH, Baker FC, Tsai LW, and Schooley DA (1986) Juvenile hormone and juvenile hormone esterase in adult females of the mosquito *Aedes aegypti*. *J. Insect Physiol.* **32,** 867-877.

Shapiro JP (1983) Ovarian cyclic AMP and response to a brain hormone from the mosquito *Aedes aegypti*. *Insect Biochem.* **13,** 273-279.

Shaywitz AJ, and Greenberg ME (1999) CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annu. Rev. Biochem.* **68,** 821-861.

Sladek FM (1994) Hepatocyte nuclear factor 4 (HNF4). *In* Tronche F, and Yaniv M (eds:) *Liver Gene Expression*. RG Landes Company. Austin, TX, pp. 207-230.

Sladek FM, Zhong W, Lai E, and Darnell JE Jr (1990) Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* **4,** 2353-2365.

Smolik SM, Rose RE, and Goodman RH (1992) A cyclic AMP-responsive element-binding transcriptional activator in *Drosophila melanogaster*, dCREB-A, is a member of the leucine zipper family. *Mol. Cell. Biol.* **12,** 4123-4131.

Snigirevskaya ES, Hays AR, and Raikhel AS (1997a) Secretory and internalization pathways of mosquito yolk protein precursors. *Cell Tiss. Res.* **290,** 129-142.

Snigirevskaya ES, Sappington TW, and Raikhel AS (1997b) Internalization and recycling of vitellogenin receptor in the mosquito oocyte. *Cell Tiss. Res.* **290,** 175-183.

Suaud L, Formstecher P, and Laine B (1999) The activity of the activation function 2 of the human hepatocyte nuclear factor 4 (HNF-4) is differently modulated by F domains from various origins. Biochem. J. 340, 161-169.

Sun J, Hiraoka T, Dittmer NT, Cho K-H, and Raikhel AS (2000) Lipophorin as a yolk protein precursor in the mosquito *Aedes aegypti. Insect Biochem. Mol. Biol.* (In Press).

Sun P, Enslen H, Myung PS, and Maurer RA (1994) Differential activation of CREB by Ca2+/calmodulin-dependent protein kinases type II and type IV involves phosphorylation of a site that negatively regulates activity. *Genes Dev.* **8,** 2527-2539.

Sun P, and Maurer RA (1995) An inactivating point mutation demonstrates that interaction of cAMP response element binding protein (CREB) with the CREB binding protein is not sufficient for transcriptional activation. *J. Biol. Chem.* **270,** 7041-7044.

Swevers L, and Iatrou K (1998) The orphan receptor BmHNF-4 of the silkmoth *Bombyx mori*: ovarian and zygotic expression of two mRNA isoforms encoding polypeptides with different activating domains. *Mech. Dev.* **72,** 3-13.

Szilak L, Moitra J, Krylov D, and Vinson C (1997) Phoshorylation destabilizes alpha-helices. *Nat. Strut. Biol.* **4,** 112-114.

Talbot WS, Swyryd EA, Hogness DS (1993) *Drosophila* tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms. *Cell.* **73,** 1323-1337.

Taraviras S, Monaghan AP, Schütz G, and Kelsey G (1994) Characterization of the mouse HNF4 gene and its expression during mouse embryogenesis. *Mech. Dev.* **48,** 67-79.

Taraviras S, Mantamadiotis T, Song-Si T, Mincheva A, Lichter P, Drewes T, Ryffel GU, Monaghan AP, and Schütz G (2000) Primary structure, chromosomal mapping, expression and transcriptional activity of murine hepatocyte nuclear factor 4γ. *Biochim. Biophys. Acta* **1490,** 21-32.

Tate BF, Allenby G, Perez JR, Levin AA, and Grippo JF (1996) A systematic analysis of the AF-2 domain of human retinoic acid receptor alpha reveals amino acids critical for transcriptional activation and conformational integrity. *FASEB J.* **10,** 1524-1531.

Taylor DG, Haubenwallner S, and Leff T (1996) Characterization of a dominant negative mutant form of the HNF-4 orphan receptor. *Nucleic Acids Res.* **24,** 2930-2935.

Taylor SS (1989) cAMP-dependent protein kinase. Model for an enzyme family. *J Biol Chem.* **264,** 8443-8446.

Thummel CS (1995) From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor superfamily members. *Cell* **83,** 871-877.

Trautwein C, Van der Geer P, Karin M, Hunter T, and Chojkier M (1994) Protein kinase A and C site-specific phosphorylations of LAP (NF-IL6) modulate its binding affinity to DNA recognition elements. *J. Clin. Invest.* **93,** 2554-2561.

Triezenberg SJ (1995) Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5,** 190-196.

Tsai MJ and O'Malley BW (1994) Molecular mechanism of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* **63,** 451-486.

Umesono K, and Evans RM (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57,** 1139-1146.

Umesono K, Giguere V, Glass CK, Rosenfeld MG, and Evans RM (1988) Retinoic acid and thyroid hormone induce gene expression through a common responsive element. *Nature* **336,** 262-265.

Usui T, Smolik SM, and Goodman RH (1993) Isolation of *Drosophila* CREB-B: a novel CRE-binding protein. *DNA Cell Biol.* **12,** 589-595.

Van Handel E, and Lea AO (1984) Vitellogenin synthesis in blood-*fed Aedes aegypti* in the absence of the head, thorax and ovaries. *J. Insect Physiol.* **30,** 871-875.

Van Heusden MC, Thompson F, and Dennis J (1998) Biosynthesis of *Aedes aegypti* lipophorin and gene expression of its apolipoproteins. *Insect Biochem. Mol. Biol.* **28,** 733-738.

Viollet B, Kahn A, and Raymondjean M (1997) Protein kinase A-dependent phosphorylation modulates DNA-binding activity of hepatocyte nuclear factor 4. *Mol. Cell. Biol.* **17,** 4208-4219.

Vorgia P, Zannis VI, and Kardassis D (1998) A short proximal promoter and the distal hepatic control region-1 (HCR-1) contribute to the liver specificity of the human *apolipoprotein CII* gene. *J. Biol. Chem.* **273,** 4188-4196.

Wang S-F, Miura K, Miksicek RJ, Segraves WA, and Raikhel AS (1998) DNA binding and transactivation characteristics of the mosquito ecdysone receptor-ultraspiracle complex. *J. Biol. Chem.* **273,** 27531-27540.

Wang S-F, Li C, Zhu J, Miura K, Miksicek RJ, and Raikhel AS (2000a) Differential expression and regulation by 20-hydroxyecdysone of mosquito ultraspiracle isoforms. *Dev. Biol.* **218,** 99-113.

Wang S-F, Ayer S, Segraves WA, Williams DR, and Raikhel AS (2000b) Molecular determinants of differential ligand sensitivities of insect ecdysteroid receptors. *Mol. Cell. Biol.* **20,** 3870-3879.

Weers, PMM, Van Marrewijk WJA, Beenakkers AMTh, and Van Der Horst DJ (1993) Biosynthesis of locust lipophorin. *J. Biol. Chem.* **268,** 4300-4303.

Weigel D, Jürgens G, Küttner F, Seifert E, and Jäckle H (1989) The homeotic gene *forkhead* encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell* **57,** 645-658.

Weigel NL (1994) Receptor phosphorylation. *In* Tsai M-J, and O'Malley BW (eds:) *Molecular biology intelligence unit: mechanism of steroid hormone regulation of gene transcription.* RG Landes Company. Austin, TX, pp. 93-110.

Whitfield GK, Jurutka PW, Haussler CA, Haussler MR (1999) Steroid hormone receptors: evolution, ligands, and molecular basis of biologic function. *J. Cell Biochem.* **Suppl 32-33,** 110-122.

Wölfl S, Martinez C, and Majzoub JA (1999) Inducible binding of cyclic adenosine 3',5'-monophosphate (cAMP)-responsive element binding protein (CREB) to a cAMP-responsive promoter in vivo. *Mol. Endocrinol.* **13,** 659-669.

World Health Organization (1998a) Malaria. Fact sheet No. 94. [Online] Available http://www.who.int/inf-fs/en/fact094.html.

World Health Organization (1998b) Dengue and Dengue haemorrhagic fever. Fact sheet No. 117. [Online] Available http://www.who.int/inf-fs/en/fact117.html.

World Health Organization (1999a) Yellow fever. Fact Sheet No. 100. [Online] Available http://www.who.int/inf-fs/en/fact100.html.

World Health Organization (1999b) Lymphatic filariasis. Fact sheet No. 102. [Online] Available http://www.who.int/inf-fs/en/fact102.html.

Wu X, Spiro C, Owen WG, and McMurray CT (1998) cAMP response element-binding protein monomers cooperatively assemble to form dimers on DNA. *J. Biol. Chem.* **33,** 20820-20827.

Xanthopoulos KG, and Mirkovitch J (1993) Gene regulation in rodent hepatocytes during development, differentiation and disease. *Eur. J. Biochem.* **216,** 353-360.

Xing J, Ginty DD, and Greenberg ME (1996) Coupling of the RAS-MAPK pathway to gene activation by RSK2, a growth factor-regulated CREB kinase. *Science* **273,** 959-963.

Yamamoto KK, Gonzalez GA, Biggs WH, III, and Montminy MR (1988) Phosphorylation-induced binding and transcriptional efficacy of nuclear factor CREB. *Nature* **334,** 494-498.

Yao T-P, Segraves WA, Oro AE, McKeown M, Evans RM (1992) *Drosophila* ultraspiracle modulates ecdysone receptor function via heterodimer formation. *Cell* **71,** 63-72.

Yao T-P, Forman BM, Jiang Z, Cherbas L, Chen J-D, McKeown M, Cherbas P, and Evans RM (1993) Functional ecdysone receptor is the product of *EcR* and *ultraspiracle* genes. *Nature* **366,** 476-479.

Yin JCP, Wallach JS, Del Vecchio M, Wilder EL, Zhou H, Quinn WG, and Tully T (1994) Induction of a dominant negative CREB transgene specifically blocks long-term memory in *Drosophila. Cell* **79,** 49-58.

Yin JCP, Del Vecchio M, Zhou H, and Tully T (1995a) CREB as a memory modulator: induced expression of a dCREB2 activator isoform enhances long-term memory in *Drosophila. Cell* **81,** 107-115.

Yin JCP, Wallach JS, Wilder EL, Klingensmith J, Dang D, Perrimon N, Zhou H, Tully T, and Quinn WG (1995b) A *Drosophila* CREB/CREM homolog encodes multiple isoforms, including a cyclic AMP-dependent protein kinase-responsive transcriptional activator and antagonist. *Mol. Cell. Biol.* **15,** 5123-5130.

Zhong W, Sladek FM, and Darnell JE Jr (1993) The expression pattern of a *Drosophila* homolog to the mouse transcription factor HNF-4 suggests a determinative role in gut formation. *EMBO J.* **12,** 537-544.

Zhong W, Mirkovitch J, Darnell JE Jr (1994) Tissue-specific regulation of mouse hepatocyte nuclear factor 4 expression. *Mol Cell Biol.* **14,** 7276-7284.

Zhu J, Miura K, Chen L, and Raikhel AS (2000) AHR38, a homolog of NGFI-B, inhibits formation of the functional ecdysteroid receptor in the mosquito *Aedes aegypti. EMBO J.* **19,** 253-262.