This is to certify that the

dissertation entitled

The Structure and Expression of the gene for Mosquito Vitellogenic Carboxypeptidase

presented by

Kirk W. Deitsch

has been accepted towards fulfillment
of the requirements for

PhD degree in Genetics

Alexander S. Raikhel
Major professor

Date 8/3/94

O-12771

THE STRUCTURE AND EXPRESSION OF THE GENE FOR MOSQUITO VITELLOGENIC CARBOXYPEPTIDASE

By

Kirk W. Deitsch

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Graduate Program in Genetics

1994

ABSTRACT

THE STRUCTURE AND EXPRESSION OF THE GENE FOR MOSQUITO VITELLOGENIC CARBOXYPEPTIDASE

By

Kirk W. Deitsch

At the heart of the reproductive cycle of the blood-feeding mosquito, *Aedes aegypti,* is vitellogenesis, the production of yolk proteins by the fat body and their subsequent storage in the developing eggs. The genes involved in yolk protein synthesis are regulated by a complicated system of control that ultimately leads to their sex-, stage- and tissue-specific expression. Several hormones have been implicated in this control, but the exact nature of their action remains unclear. To understand this system further, a small 53kDa yolk protein was chosen as the object of a detailed investigation into the regulation of its production. A putative cDNA was cloned for this protein and verified through direct N-terminal sequencing of the purified protein. This cDNA hybridized to a 1.5 kb mRNA present only in the fat bodies of vitellogenic females. The deduced amino acid sequence and experiments using the serine protease inhibitor, [^{3}H] diiisopropyl fluorophosphate, indicated that this protein was a serine carboxypeptidase. This finding that an enzyme is produced by an extraovarian tissue and accumulated by oocytes for use in embryonic development represents a biological phenomenon not previously known for oviparous animals. The protein was therefore name vitellogenic carboxypeptidase (VCP).

Two alleles for the gene encoding VCP were cloned from an *Aedes aegypti* genomic library, sequenced and analyzed. Several putative regulatory

elements were found, including a 16 bp repeated imperfect palindrome, a TATA box, a region resembling the fat body enhancer found in *Drosophila*, and an arthropod transcriptional initiator. Primer extension analysis showed a single start site within this initiator. The palindromic element found repeated upstream of the gene resembled several defined steroid hormone response elements. The gene for VCP can be induced by treatment with the steroid hormone 20-hydroxyecdysone (20E), however, this putative response element did not meet the consensus thought to be necessary for binding of the 20E/receptor complex, leading to the hypothesis that the control of 20E on yolk protein synthesis in mosquito may be indirect. Experiments done with the protein synthesis inhibitor cycloheximide confirmed this hypothesis. A 20E response cascade therefore probably exists controlling the expression of the genes involved in mosquito vitellogenesis. Analysis of the structure and expression of the gene for VCP represents a first step in the illucidation of this important hormone response hierarchy.

AKNOWLEDGMENTS

Special thanks to my professor, Dr. Alexander Raikhel, for his advice and guidance throughout the course of this study and to my committee, Dr. Thomas Freidman, Dr. Zachary Burton and Dr. Jerry Dodgson, for their help in meeting the requirements for my degree. In addition, thanks to Wen-Long Cho and Tolarchan Dhadialla for their assistance in the early stages of this project. All of my colleagues in the lab, Marianna Kapitskaya, Vladimir Kokoza, Tom Sappington, Alan Hays, Katya Snigerevskaya, Neal Dittmer, and Jeng-Shong Chen, as well as Dr. Suzanne Theim and Chi-Zu Chen have had a hand in this work through discussions and lab talks, and have been extremely helpful.

I would also like to thank Jeng-Shong Chen and Neal Dittmer for their friendship and camaraderie. They helped make the last four and one half years an enjoyable experience both inside and outside the lab. Their friendship will always be valued.

Particular acknowledgment is owed to my parents, William and Connie Deitsch. All that I achieve in life is ultimately a credit to their patience, guidance, encouragement, and steadfast loyalty. They have always been there to provide me with the advice and assistance needed to help me overcome whatever obstacles I encountered. I only hope to be as good a parent to their grandchildren as they were to me.

Lastly, I must thank Kelly Hallman; my wife, confidant, counselor, playmate, best friend and ultimate source of sanity and inspiration. She provided me with the motivation and desire needed at times when things did not go well, and helped me enjoy the successes that came along the way. Marrying her has been my greatest achievement.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 4

CHAPTER 5

ABBREVIATIONS

AInr .....arthropod initiator

ATF .....activating transcription factor

BBF-2 .....*Drosophila* Box B-binding factor-2

bp .....base pair

cAMP .....cyclic adenine monophosphate

cDNA .....complementary DNA

Chx .....cycloheximide

CREB .....cAMP responsive transcriptional enhancer binding protein

DFP .....diisopropyl fluorophosphate

DNA .....deoxyribonucleic acid

DSX .....*Drosophila* double sex protein

20E 20-hydroxyecdysone

EcR .....ecdysone receptor

EcRE .....ecdysone response element

EDTA .....ethylenedinitrilo tetraacetic acid

GCG .....Compusoftware from University of Wisconsin Genetics Computer Group

JHIII .....juvenile hormone III

kDa .....kiloDalton

kb .....kilobase pairs

M_r .....molecular weight

mRNA .....messenger RNA

PBM .....post blood meal

PCR .....polymerase chain reaction

RNA .....ribonucleic acid

SDS .....sodium dodecyl sulfate

SDS/PAGE .....SDS polyacrylamide gel electrophoresis

TCA .....trichloroacetic acid

UGAL .....University of Georgia at Athens laboratory strain

VCP .....vitellogenic carboxypeptidase

VCP^1 .....first allele of the VCP gene

VCP^2 .....second allele of the VCP gene

Vg .....vitellogenin

WCP .....wheat carboxypeptidase

YP1 .....*Drosophila* yolk protein 1

YP2 .....*Drosophila* yolk protein 2

CHAPTER 1

INTRODUCTION

1. *Historical Perspective:*

Mosquitoes are the most prominent insect pests of human populations throughout the world. By acting as a biting nuisance and as vectors of human disease, these animals take a great toll in both human health and financial resources. In California, almost 55 million dollars per year are spent on mosquito control and research for a region with a population of 22 million people. An additional 45 million dollars are spent in Florida each year to combat the problem. Worldwide, hundreds of millions of dollars have been spent in an attempt to control mosquito populations to reduce malaria and other diseases (Table 1) (McClelland, 1992).

Attempts to control the spread of mosquito-borne diseases have often focused on control of insect populations. In the mid 1950's, a global malaria eradication campaign was begun, largely funded and organized by the World Health Organization. This campaign relied primarily on the use of DDT as an insecticide and chloroquine to treat infected individuals. Initial results of the campaign were encouraging; in India, the number cases of malaria fell from 75 million in 1953 to 50,000 in 1958. Similar results were obtained in other regions of the world. However, as pesticide resistance grew in the mosquito population and chloroquine resistance of the parasite increased, incidence of disease rapidly rose to levels seen before the campaign began. Indeed, in 1977, 50 million cases of malaria were detected in India (McClelland, 1992).

In recent decades, much effort and emphasis has been placed on the development of a vaccine against the malaria parasite. This search has been largely unsuccessful thus far, with no substantially effective vaccine likely available in the near future. Such a development would be highly beneficial to people traveling into endemic areas or residents of cities in areas with high

Table 1: A sample of the costs of recent epidemics of mosquito-borne diseases in several regions of the world. (From McClelland, 1992).

Disease	No of Cases	Medical cost	Control cost	Indirect cost	Overall cost	Per patient
Dengue						
Puerto Rico 1977	390,503	$ 2.3 million	$ 1.2 million	$ 7.2 million	$ 10.8 million	$ 27.64
Thailand 1980	43,000	$ 4.3 million	$ 0.5 million	$ 1.07 million	$ 7.0 million	$ 162.00
Cuba 1981	10,000	?	?	?	$ >100 million	$ >10,000
St. Louis Encephalitis						
Dallas, Texas 1966		$ 0.2 million	$ 0.35 million	$ 0.25 million	$ 0.8 million	$ 4,631

malaria rates. However, in much of the malarious zones of the third world, problems of poverty, shortages, communications, transportation and political unrest make the feasibility of high levels of vaccination doubtful. Thus the benefits of vaccination to a substantial portion of the at risk population may be unrealistic (Anderson and May, 1991).

Numerous other illnesses are carried by mosquitoes in various parts of the world. Mosquitoes from the genus *Aedes* carry the viruses responsible for Dengue, yellow fever, and numerous varieties of encephalitis as well as the parasite responsible for filariasis. *Anopheles* mosquitoes are most known as malaria vectors, but also carry filariasis parasites. Control of major outbreaks and epidemics of these diseases also depends on control of the vector population. Because of the rapidly rising incidence of resistant populations of mosquitoes to current pesticides and growing dislike of large scale application of chemical insecticides, new and innovative control strategies are needed. The development of such methods are dependent on basic research into the biology of these insects. The rapidly developing field of molecular biology provides new tools for use in this investigation.

2. *Basic Physiology of Female Mosquito Reproduction:*

In anautogenous (blood feeding) mosquitoes, egg production depends on a blood meal. Over the lifetime of an individual female, multiple blood meals are taken, one or more for each clutch of eggs. Pathogens are acquired by the mosquito during initial blood meals and passed on to future hosts through subsequent feeding. In this way, the propagation of disease causing organisms and the spread of mosquito-borne diseases are both intimately linked to the behavior and physiology of the female. For this reason, a great deal of research has been conducted to build an understanding of how this process works in

hopes of developing ways to interrupt either the reproduction of the insect or the propagation of the pathogen.

The cyclical nature of egg production in mosquitoes differs significantly from that of *Drosophila,* in insect more familiar to geneticists. Where as in fruit flies adult females produce eggs continuously, mosquitoes instead produce large numbers of completely synchronous eggs at distinct time intervals. As a result, the insect undergoes rather drastic, tightly controlled physiological changes that facilitate large scale synthesis of the many components of the eggs, followed by remodeling of the synthetic tissue for subsequent egg producing cycles. This process, called vitellogenesis, is regulated by various hormones that act specifically at different points in the cycle.

The major synthetic tissue during vitellogenesis is the fat body, the insect analog to the vertebrate liver. Tissue of this type is found primarily in the abdomen, arranged as sheets or lobes which are attached to the epidermis and surround the midgut. Additional fat body cells are also found in the head and thorax. The activity of this tissue during egg production can be divided into three distinct stages: 1) a preparatory phase in which the female develops the competence to produce eggs, 2) an intensely synthetic phase during which large amounts of yolk protein molecules are produced and secreted into the hemolymph for subsequent storage in the rapidly maturing oocytes and 3) a termination phase where the biosynthetic organelles involved in yolk protein production are broken down by lysosomes (for review, see Raikhel, 1992).

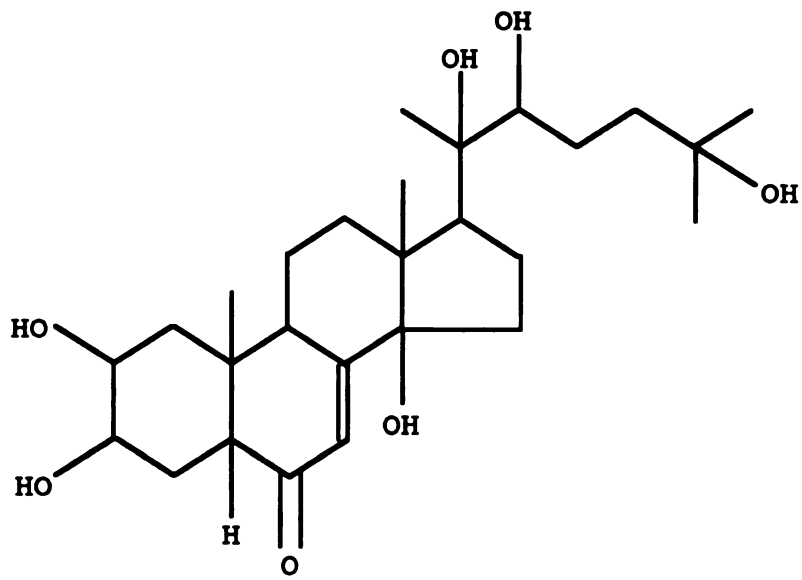
The preparatory phase of development is absolutely necessary for vitellogenesis to occur. This stage lasts for about three days beginning from the time the adult ecloses from the pupa. During this time, the insect is unwilling to take a blood meal and cannot produce eggs. Two major cellular events take place in the fat body during this period: 1) an increase in ploidy (Dittman *et al.*,

1989), and 2) a proliferation of ribosomes (Raikhel and Lea, 1990). These processes appear to be controlled by juvenile hormone III (JHIII)(Fig. 1). The titer of this hormone rises over the first 2 days after eclosion, then slowly declines over the next five days in females not given a blood meal (Shapiro *et al.*, 1986). The rate of RNA synthesis and the amount of ribosomal RNA in fat body cells as well as synthesis of DNA directly follow the increase in JHIII. After three days of preparatory development, the insect goes into a state of arrest. During this time her reproductive cycle remains static, without major changes until a blood meal is obtained.

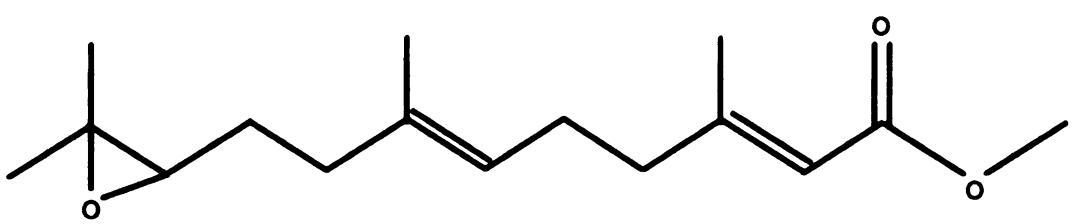
With the acquisition of a blood meal, the synthetic phase of vitellogenesis begins. During this time the fat body produces massive quantities of yolk proteins which are secreted into the hemolymph and taken up by the developing oocytes. This period lasts from the time of the blood meal until approximately 36 hrs later, when yolk protein synthesis and uptake have ceased (Koller *et al.*, 1989). During this time, the levels of JHIII fall to very low levels, while the titer of the steroid hormone 20-hydroxyecdysone (20E)(Fig. 1) increases dramatically (Shapiro *et al.*, 1986; Hagedorn, 1983). The amount of 20E peaks at 18-24 hrs after a blood meal, then begins to fall rapidly in the later stages of the synthetic phase. This rise and fall in 20E titer is correlated with the changes in the rate of yolk protein synthesis, leading to the initial hypothesis that 20E controls production of the yolk proteins, probably at the level of transcription (Fallon *et al.,* 1973). This topic has been the subject of a considerable amount of research, and remains unclear.

Once the titer of 20E has fallen to pre-blood meal levels, the termination phase of vitellogenesis begins. This stage lasts from approximately 30 hrs to 48 hrs post blood meal, after which the mosquito is ready to lay eggs and begin the entire cycle anew. Termination involves two distinguishable steps: 1) turning off

Fig. 1: The structure of the two major hormones controlling mosquito vitellogenesis. (Oberlander, 1985).



20-hydroxyecdysone



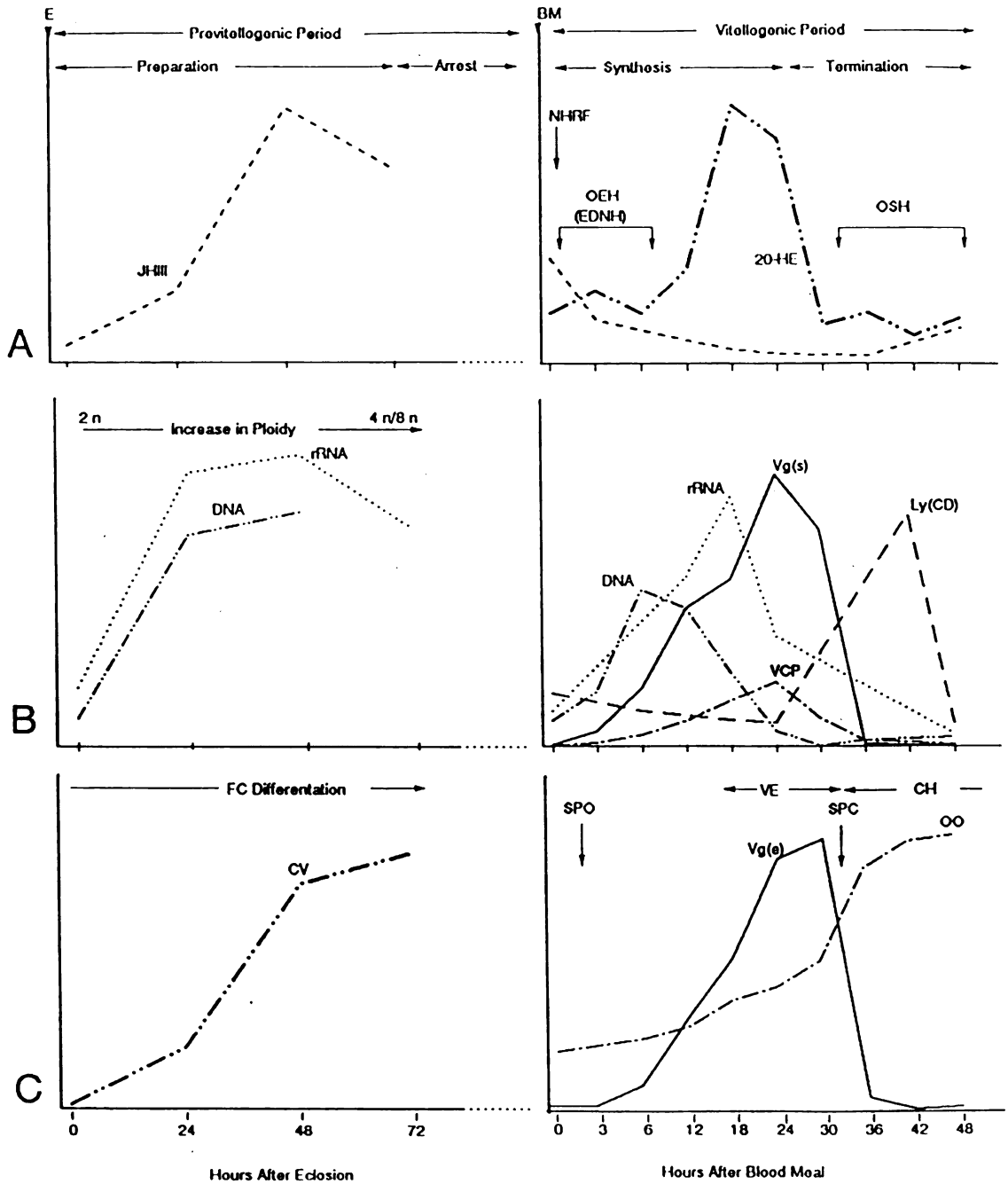
Juvenile Hormone III

the expression of the yolk protein genes and 2) destruction of the biosynthetic machinery (rough endoplasmic reticulum, Golgi complexes and secretory vesicles) responsible for yolk protein secretion (Raikhel 1986a; Raikhel, 1986b). This is accompanied by a large proliferation of lysosomes within the cells. The fat body cells are remodeled and readied to re-enter the preparatory phase to initiate a new round of the reproductive cycle. This entire cycle is summarized in Fig. 2.

3. *Control of Yolk Protein Synthesis:*

Perhaps the most dramatic aspect of the vitellogenic cycle is the transition of the fat body from its state-of-arrest, to 36 hrs of extremely intense yolk protein synthesis, then its subsequent return to a non-synthetic state. The exact nature of how this tightly regulated process is controlled has been the subject of a great deal of speculation and research. The initial hypothesis that the titer of 20E directly controls the transcription of the yolk protein genes and, as a result, the rate of yolk protein synthesis, was simple and intuitive. To test this hypothesis, a number of *in vivo* and *in vitro* experiments were done. When fat bodies of mature females were cultured *in vitro*, physiological doses of 20E were found to stimulate yolk protein synthesis (Fuchs and Kang, 1981; Hagedorn, 1985). However, *in vivo* injections of 20E required large doses to generate similar levels of yolk protein production (Fuchs and Kang, 1981; Lea, 1982). In addition, yolk protein induction at both the protein and mRNA levels is detectable in females only 1 hr after a blood meal, significantly before increases in 20E titers can be measured (Raikhel and Lea, 1983; Racioppi *et al.*, 1986). Thus, 20E is clearly a significant factor in the initiation of vitellogenesis, however, exactly how this hormone exerts control over its target genes remains unclear.

Fig. 2: Summary of events during the first cycle of egg maturation in the anautogeneous mosquito, *Aedes aegypti.* (A) Hormonal Events. EDNH, egg development neurosecretory hormone; JHIII, juvenile hormone III; NHRF, neurosecretory hormone-releasing factory; OEH, ecdysteroidogenic hormone; OSH, oostatic hormone; 20E, 20-hydroxyecdysone. (B) Events in the fat body. DNA, rate of DNA synthesis; VCP, rate of synthesis of vitellogenic carboxypeptidase; Vg(s), rate vitellogenin synthesis. (C) Events in the ovarian primary follicle. CV, number of coated vesicles per unit of the oocyte cortex; CH, chorion deposition; FC, follicle cell; OO, oocyte yolk length; SPC, closing of interfollular spaces; SPO, openning of interfollicular spaces; VE, vitelline envelope formation; Vg(e), rate of Vg endocytosis. (From Raikhel, 1992).


A
E
Previtellogenic Period
Preparation
Arrest
JHIII
BM
Vitellogenic Period
Synthesis
Termination
NHRF
OEH
(EDNH)
OSH
20-HE
B
2 n
Increase in Ploidy
4 n/8 n
rRNA
DNA
Vg(s)
rRNA
Ly(CD)
DNA
VCP
C
FC Differentation
CV
VE
CH
SPO
SPC
OO
Vg(e)
0
24
48
72
0
3
6
12
18
24
30
36
42
48
Hours After Eclosion
Hours After Blood Meal

4. *The Ashburner Model of 20-hydroxyecdysone Control of Gene Expression:*

Ashburner and his colleagues proposed a model to describe how 20E regulates gene expression in *Drosophila* salivary glands (Ashburner *et al.*, 1974). Working with polytene chromosomes that were exposed to 20E under a variety of conditions, they observed puffing at the loci of genes specifically regulated by this hormone. In this model, 20E enters the cell and is bound by its receptor. The activated hormone/receptor complex then differentially regulates two classes of genes. The first class, referred to as the early genes, is a small set of loci encoding transcription factors that repress their own transcription and induce transcription of the late genes. The late genes are a large class of loci that represent the bottom of the hierarchy. The model provides for regulation at the levels of the hormone, the receptor, the early genes and the late genes. This control is exhibited in sex-, tissue- and stage-specific ways.

In recent years, a number of molecular studies have been published that support this model of 20E action in *Drosophila* salivary glands. Three of the early puffing loci have been cloned and analyzed. These include the *E75* gene (Segraves and Hogness, 1990), the *E74* gene (Thummel *et al.*, 1990) and a member of the *Broad-Complex* (DiBello *et al.*, 1991). As predicted, transcription of each of these genes is induced by 20E in the initial response, then subsequently repressed as the late genes are activated. These genes encode DNA binding proteins that show significant homology to transcription factors. In addition, E74 and E75 proteins were shown to bind to both the early and late classes of ecdysone responsive puff sites on the polytene chromosomes. These observations are consistent with the idea that these early gene products repress their own expression while inducing late gene transcription.

In the mosquito, such detailed studies of the action of 20E have not been done. During vitellogenesis, control of the yolk proteins by 20E could follow two potential pathways. The hormone, bound by its receptor, may bind directly to the control regions of these genes, inducing high levels of expression. Alternatively, it may act through a hormone response cascade, with the yolk protein genes at the bottom of the hierarchy. In this way, the yolk protein genes would represent "late" loci in the Ashburner model. Either way, the control of expression of these genes is likely to be quite complex, with numerous factors determining the sex-, stage-, and tissue-specific control exhibited.

5. *Yolk Protein Expression in* Drosophila:

Considerable research has been done on the regulation of the yolk protein genes in *Drosophila*. Experiments using hormone injections and organ culture have shown that yolk protein synthesis occurs in response to either 20E or JHIII (Jowett and Postlethwait, 1980; Bownes, 1982; Bownes *et al.*, 1983). In fruit flies, however, yolk proteins are produced by both the fat body and the follicle cells of the ovary. In addition, males can be induced to produce yolk proteins by injections of 20E (Kozma and Bownes, 1986). The sexual differentiation pathway has also been implicated in the control of yolk protein synthesis. Genetic analysis using mutants of the sex determination regulatory gene *transformer-2* have shown that proper expression of this gene is required for expression of the yolk protein genes (Belote *et al.*, 1985). The nutritional status of the female also influences yolk protein production (Bownes and Blair, 1986).

Three yolk proteins have been found in *Drosophila*, named YP1 (46 kDa), YP2 (45 kDa) and YP3 (44 kDa) (Warren and Mahowald, 1979). The proteins are encoded by different, single copy genes, all located on the X-chromosome

(Barnett *et al.*, 1980; Barnett and Wensink, 1981). YP1 and YP2 are divergently transcribed, sharing a 1227 bp upstream regulatory region, while YP3 is located at least 10^6 bp away. All three genes have been cloned and sequenced, and considerable research has been done on their regulatory regions. Using P-element mediated germline transformation, deletion analysis and site-directed mutagenesis, several specific regulatory elements have been defined. These include a 127 bp sequence that induces fat body specific expression of a reporter gene (Shepherd *et al.*, 1985; Garabedien *et al.*, 1986), a 301 bp region directing ovarian specific expression (Logan *et al.*, 1989) and several specific binding sites for a member of the sex determination pathway, doublesex (Coshigano and Wensink, 1993). In addition, binding sites for two tissue-specific transcription factors, BBF-2 (a member of the CREB/ATF family of transcriptional regulatory proteins) and AEF-1 (adult enhancer factor 1) have been identified between YP1 and YP2 (Falb and Maniatis, 1992; Abel *et al.*, 1992) This research provides a model for comparison when similar regulatory regions in mosquito become available.

Yolk Protein Expression in* Aedes aegypti*:

The study of vitellogenesis in mosquitoes is most advanced in the species *Aedes aegypti*. This species is easily reared in a laboratory setting and provides the best model for studying the physiology and molecular biology of mosquitoes in general. Most research has centered on synthesis and secretion of the most abundant yolk protein, termed vitellogenin (Vg). This is a 380 kDa protein produced exclusively in the fat body and selectively stored in the developing oocytes. The protein consists of 2 subunits, one of 200 kDa and the second of 66 kDa (Raikhel and Bose, 1988). Synthesis, processing and secretion of this protein has been extensively studied. Recently, a second, much smaller yolk

protein was observed that was produced in a manner identical to Vg (Hays and Raikhel, 1990). Unlike Vg, this protein is not phosphorylated or sulfated, and shows small levels of glycosylation. Because of its smaller size and simpler post-translational modifications, this 53 kDa protein has been considered a more convenient marker for molecular studies of protein processing in the fat body and internalization by the oocyte.

Research into yolk protein production at the gene level has been limited. A multigene family consisting of 5 genes encoding Vg has been identified (Hamblin *et al.*, 1987). Four of these are thought to be actively transcribed and all hybridize to a 6.5 kb mRNA present only in vitellogenic fat bodies. Preliminary sequence data from the regulatory region of one of these gene has identified a number of elements resembling a TATA box, a putative capsite and sequence motifs similar to hormone response elements (Hagedorn, unpublished). Complete analysis of the structure of these genes is somewhat cumbersome due to the large size of the coding region and the number of copies in the genome. This combined with the lack of an expression system for vitellogenic genes have slowed research in this area. As a result, nothing has been published regarding the genes for *A. aegypti* Vg since their initial identification. Analysis of the locus for the recently identified 53 kDa yolk protein may represent an alternative research pathway, providing a smaller, perhaps single copy gene for analysis.

Beyond the basic research interests for investigating the regulatory regions responsible for yolk protein expression, there is also a desire to obtain such DNA sequences for use in possible genetic engineering of mosquitoes against disease causing parasites (Aldhous, 1993). Numerous laboratories are currently attempting to develop methods of germline transformation of disease-vector species, as well as finding ways to move the genes of interest throughout

a natural population (Kidwell and Ribeiro, 1992; Ribeiro and Kidwell, 1994). Candidate parasite-inhibiting genes for use in such a strategy are also being developed (Huber *et al.*, 1993). When potential anti-parasite genes have been found, powerful promoters and regulatory DNA sequences will be needed to express these genes in such a way as to make them effective against the disease causing organism. Elements that control yolk protein gene transcription may provide the regulatory sequences desired. Yolk protein genes are expressed at extremely high levels in the major synthetic organ of the adult, at the specific time period when the insect is likely to come into contact with the parasite. Thus, the regulatory region of the gene for Vg or the 53 kDa protein could be of great value to the development of a biological control strategy involving mosquito genetic engineering.

REFERENCES

Abel, T., Rupal B. and Maniatis, T., 1992, A *Drosophila* CREB/ATF transcriptional activator binds to both fat body- and liver-specific regulatory elements, *Genes Develop.* **6:** 466-480.

Aldhous, P., 1993, Malaria: focus on mosquito genes, *Science*, **261:** 546-548.

Anderson, R.M. and May, R.M., 1991, *Infectious Diseases of Humans*, Oxford University Press, Oxford, United Kingdom.

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:** 645-654.

Barnett, T., Pachl C., Gergen, J.P. and Wensink, P.C., 1980, The isolation and characterization of *Drosophila* yolk protein genes, *Cell* **21:** 729-738.

Barnett, T. and Wensink, P.C., 1981, Trancription and translation of yolk protein mRNA in fat bodies of *Drosophila*, In: Brown D.D. and Fox, C.F. (eds) *Developmental Biology Using Purified Genes*: UCLA-ICN Symposium, University of California Press, Berkely, **23:** 99-106.

Belote, J.M., Handler, A.M., Wolfner, M.F., Livak, K.J. and Baker, B.S., 1985, Sex-specific regulation of yolk protein gene expression in *Drosophila*. *Cell* **40:** 339-348.

Bownes, M., 1982, The role of 20-hydroxyecdysone in yolk-polypeptide synthesis by male and female fat bodies of *Drosophila melanogaster*, *J. Insect Physiol.* **28:** 317-328.

Bownes, M. and Blair, M., 1986, The effects of a sugar diet and hormones on the expression of the *Drosophila* yolk-protein genes, *J. Insect Physiol.* **32:** 493-501.

Bownes, M., Blair, M., Kozma, R. and Dempster, M., 1983, *J. Embryol. Exp. Morph.* **78:** 249-268.

Coschigano, K.T. and Wensink, P.C., 1993, Sex-specific transcriptional regulation by the male and female doublesex proteins of *Drosophila*, *Genes Develop.* **7:** 42-54.

DiBello, P.R., Withers, D.A., Bayer, C.A., Fistrom, J.W. and Guild, G.M., 1990, Spatial and temporal patterns of *E74* transcription during *Drosophila* development. *Cell* **61:** 101-111.

Dittman, F., Kogan, P.H., and Hagedorn, H.H., 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.

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

Fallon, A.M., Hagedorn, H.H., Wyatt, G.R., and Laufer, H., 1974, Activation of vitellogenin synthesis in the mosquito *Aedes aegypti* by ecdysone, *J. Insect Physiol.* **20:** 1815-1823.

Flanagan, T.R., and Hagedorn, H.H., 1977, Vitellogenin synthesis in the mosquito: the role of juvenile hormone in the development of responsiveness to ecdysone, *physiol. Entomol.* **2:** 173-178.

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

Garabedien, M.J., Shepherd, B.M. and Wensink, P.C., 1986, A tissue-specific transcription enhancer from the Drosophila yolk protein 1 gene, *Cell* **45:** 859-967.

Hagedorn, H.H., 1983, The role of ecdysteroids in the adult insect, in Downer, R.G.H., and Laufer, H., (eds): Endocrinology of Insects, Alan R. Liss, Liss, Inc., New York, pp. 271-304.

Hagedorn, H.H., 1985, The role of ecdysteroids in reproduction, in Kerkut, G.A. and Gilbert, L.I., (eds): Comprehensive Insect Physiology, Biochemistry, and Pharmacology, vol. 8, Pergamon Press, Oxford, pp.205-261.

Hagedorn, H.H., Fallon, A.M., and Laufer, H., 1973, Vitellogenin synthesis by the fat body of *Aedes aegypti*, *Develop. Biol.* **31:** 285-294.

Hamblin, M.T., Marx, J.L., Wolfner, M.F., and Hagedorn, H.H., 1987, The vitellogenin gene family of *Aedes aegypti*, *Mem. Inst. Oswaldo Cruz* **82:** 109-114.

Hays, A.R., and Raikhel, A.S., 1990, A novel protein produced by the vitellogenic fat body and accumulated in mosquito oocytes, *Roux's Arch. Develop. Biol.* **199:** 114-121.

Huber, M., Cabib, E., and Miller, L.H., 1991, Malaria parasite chitinase and penetration of the mosquito peritrophic membrane, *Proc. Natl. Acad. Sci. USA* **88:** 2807.

Kidwell, M.G. and Ribeiro, J.M.C., 1992, Can transposable elements be used to drive disease refractoriness into vector populations?, *Parasitol. Today* **8:** 325.

Kozma, R. and Bownes, M., 1986, Yolk protein induction in males of several *Drosophila* species, *Insect Biochem.* **16:** 263-271.

Lea, A.O., 1982, Artifactual stimulation of vitellogenesis in *Aedes aegypti* by 20-hydroxyecdysone, *J. Insect Physiol.* **28:** 173-176.

Logan, S.K., Garabedien, M.J. and Wensink, P.C., 1989, DNA regions that regulate the ovarian transcriptional specificity of *Drosophila* yolk protein genes, *Genes Develop.* **3:** 1453-1461.

Ma, M., Zhang, J-Z., Gong, H., and Gwarz, R., 1988, Permissive action of juvenile hormone on vitellogenin production by the mosquito, *Aedes aegypti, J. Insect Physiol.* **34:** 593-596.

McClelland, G.A.H., 1992, *Medical Entomology: an Ecological Perspective*, University of California Press, Davis, California, USA.

Oberlander, H., 1985, Hormonal action during insect development, In: Blum M. (ed), *Fundementals of Insect Physiology*, John Wiley and Sons, inc., New York, 513-514.

Postlethwait, J.H. and Jowett, T., 1980, Genetic analysis of the horomonally regulated yolk polypeptide genes in *D. Melanogaster*, *Cell* **20:** 671-678.

Racioppi, J.V., Gemmill, R.M., Kogan, P.H., Calvo, J.M., and Hagedorn, H.H., 1986, Expression and regulation of vitellogenin messenger RNA in the mosquito *Aedes aegypti, Insect Biochem.* **16:** 255-262.

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

Raikhel, A.S., 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, A.S., 1992, Vitellogenesis in mosquitoes, in: *Advances in Disease Vector Research*, Springer-Verlag, New York, Inc., vol. 9, pp.1-39.

Raikhel, A.S., and Bose, S.G., 1988, Properties of mosquito yolk protein: a study using monoclonal antibodies, *Insect Biochem.*, **18:** 565-575.

Raikhel, A.S., and Lea, A.O., 1983, Previtellogenic development and vitellogenin synthesis in the fat body of a mosquito: an ultrastructural and immunocytochemical study, *Tissue Cell* **15:** 281-300.

Raikhel, A.S., and Lea, A.O., 1990, Juvenile hormone controls previtellogenic proliferation of ribosomal RNA in the mosquito fat body, *Gen. Comp. Endocrinol.* **77:** 423-434.

Ribeiro, J.M.C. and Kidwell, M.G., 1994, Transposable elements as population drive mechanisms: specification of critical parameter values, *J. Med. Ent.* **31:** 10.

Seagraves, W.A., and Hogness, D., 1990, the *E75* ecdysone-inducible gene responsible for the E75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily, *Genes Develop.* **4:** 204-219.

Shapiro, A.B., Wheelock, G.D., Hagedorn, H.H., Baker, F.C., Tsai, L.W., and Schooley, D.A., 1986, Juvenile hormone and juvenile hormone esterase in adult females of the mosquito *Aedes aegypti*, *J. Insect Physiol.* **32:** 867-877.

Shepherd, B., Garabedien, M.J., Hung, M.-C. and Wensink, P.E., 1985, Developmental control of *Drosophila* yolk protein 1 gene by cis-acting DNA elements, *Cold Spring Harbor Symp. Quant. Biol.* **50:** 521-526.

Thummel, C.S., Burtis, K.C., and Hogness, D.S., 1990, Spatial and temporal patterns of *E74* transcription during *Drosophila* development, *Cell* **61:** 101-111.

Warren, T.G. and Mahowald, A.P., 1979, Isolation and partial chemical characterization of the three major yolk polypeptides from *Drosophila melangaster*, *Proc. Natl. Acad. Sci. USA* **76:** 2848-2852.

CHAPTER 2*

AN EXTRAOVARIAN PROTEIN ACCUMULATED IN MOSQUITO OOCYTES IS A CARBOXYPEPTIDASE ACTIVATED IN EMBRYOS

*Originally published in Proceedings of the National Academy of Sciences

Reference: Cho, W.-L., Deitsch, K.W., and Raikhel, A.S. (1991) Proc. Natl. Acad. Sci., **88**, 10821-10824.

ABSTRACT

We report a phenomenon previously unknown for oviparous animals; in *Aedes aegypti* mosquitoes a serine carboxypeptidase is synthesized extraovarially and then internalized by oocytes. The cDNA encoding mosquito vitellogenic carboxypeptidase (VCP) was cloned and sequenced. The VCP cDNA hybridizes to a 1.5 kilobase mRNA present only in the fat body of vitellogenic females. The deduced amino acid sequence of VCP shares significant homology with members of the serine carboxypeptidase family. Binding assays using a serine protease inhibitor, [^{3}H]diisopropyl fluorophosphate, showed that VCP is activated in eggs at the onset of embryonic development. Activation of VCP is associated with the reduction in its size from 53 kDa (inactive proenzyme) to 48 kDa (active enzyme). The active, 48 kDa, form of VCP is maximally present at the middle of embryonic development and disappears by the end.

INTRODUCTION

In oviparious animals, development of an embryo depends upon utilization of yolk proteins accumulated by oocytes during vitellogenesis. In most oviparous animals, e.g., nematodes, arthropods, and vertebrates, the task of producing massive amounts of yolk protein precursors is performed by extraovarian tissues (1-5). The developing oocytes of these animals are highly specialized for the specific accumulation of extraovarian yolk protein precursors. The process of receptor-mediated endocytosis by which yolk protein precursors are accumulated in oocyte yolk bodies is well characterized for insects and vertebrates (6,7). In contrast, the mechanisms of degradation of yolk proteins have received considerably less attention. Several enzymes that participate in hydrolysis of yolk proteins in insect embryos have been recently described (8-11). However, the origin of these enzymes and the time of their delivery into yolk bodies are unknown.

In this paper, we report a biological phenomenon previously unknown for insects or other oviparous animals. In vitellogenic female mosquitoes, an extraovarian tissue, the fat body, synthesizes not only the yolk protein precursor vitellogenin but also a proenzyme of a serine carbosypeptidase. We have named it vitellogenic carboxypeptidase (VCP). This enzyme is internalized by developing oocytes, accumlated in yolk bodies, and activated during embryonic development.

MATERIALS AND METHODS

Animals:

Aedes aegypti mosquitoes were reared as described by Hays and Raikhel (12). The embryonic stages of *A. aegypti* were determined according to Raminani and Cupp (13).

Cloning and Sequencing of cDNA:

A cDNA clone encoding VCP was isolated by immunoscreening of a λ gt11 library generated from fat bodies of vitellogenic female mosquitoes. This cDNA clone, pVCP-1, was used to isolate another cDNA clone, pVCP-2, encoding the 5'-end 1020 base pairs (bp) of the VCP cDNA from the same library. The pVCP-2 cDNA clone was sequenced in both direction with the dideoxy chain termination method (ref. 14, pp. 13.65-13.69).

The missing part of the 3'-end sequence was obtained by the polymerase chain reaction, using 20 nucleotides at the 3'end of pVCP-1 and $(dT)_{17}$ as primers and including the *Xba* I site at the 5' ends of both primers (15). Amplification was achieved by using cDNA reverse transcribed from 20 μg of total RNA prepared from whole bodies of mosquitoes 24 hrs after a blood meal as templates. The specific amplified band, identified by Northern blot hybridization to 1.5 kilobase (kb) mRNA, was subcloned in pUC119 for sequencing. Like the pVCP-2 clone, the 3'-end VCP clone was sequenced in both directions.

Protein Purification for Partial Sequencing:

Mosquito ovaries, 48 hr after a blood meal, were homogenized in 20mM Tris·HCl, pH 7.5/150 mM NaCl/0.02 % sodium, azide. VCP was isolated from

the protein extract by anion-exchange, affinity, and hydrophobic chromatographic steps using DEAE-Sepharose CL-4B, Concanavalin A, and phenyl-Sepharose columns, respectively. A fraction with nearly pure protein was resolved by SDS/PAGE and subjected to peptide sequencing using the method and facilities described by Cho *et al.* (16).

Purification and Analysis of RNA:

Isolation of total and polyadenylated RNA [poly(A)$^+$ RNA] from mosquito tissues was performed as previously described (17).

In vitro translation of mosquito poly(A)$^+$ RNA was performed as described by Bose and Raikhel (17), using rabbit reticulocyte lysate (Promega). The VCP precursor was identified in *in vitro* translation products of poly(A)$^+$ RNA by immunoprecipitation with anti-VCP polyclonal antibodies (12).

The 840-bp insert from pVCP-1, a putative VCP cDNA clone, was used for hybrid selection of VCP mRNA from fat body poly(A)$^+$ RNA according to Sambrook *et al.* (ref. 14, pp. 8.50-8.51). The hybrid-selected mRNA was translated *in vitro* in rabbit reticulocyte lysate and the protein product was resolved by SDS/PAGE. The pUC119 plasmid without an insert was used for hybrid selection of mRNA as control.

For Northern analysis, total RNA was separated by 1.2 % formaldehyde/agarose gel electrophoresis, transferred to nitrocellulose paper, and hybridized first to a ^{32}P-labeled 840-bp VCP cDNA insert from pVCP-1 under high-stringency conditions. RNA hybridized to the VCP cDNA probe was visualized by autoradiography of the blot. The probe was then stripped from the blot (ref. 14, p.9.58) and the blot was rehybridized with a ^{32}P-labeled 1.8-kb fragment of vitellogenin gene A1 (gift of H.H. Hagedorn, University of Arizona) under high stringency conditions.

Immunoblot Analysis:

Preparation of proteins, SDS/PAGE, and immunoblotting were performed as described by Hays and Raikhel (12).

Binding to Protein Inhibitor:

Protein extracts, prepared as for immunoblotting but without any protease inhibitors, were incubated with [^{3}H]diisopropyl fluorophosphate ([^{3}H]DFP; DuPont) at 0.1 μCi/μl (1 μCi = 1 kBq) for 1 hr at room temperature. These mixtures were then separated by SDS/PAGE under reducing conditions and processed for fluorography.

Analysis of Deduced Amino Acid Sequence:

Analysis of the deduced amino acid sequence was performed by using the FASTA program (University of Wisconsin Genetics Computer Group software) according to the algorithm of Lipman and Pearson (18). The deduced amino acid sequence of VCP was compared with amino acid sequences from the National Biomedical Research Foundation Protein Data Base (release 21.0) and with deduced amino acid sequences translated from GenBank (release 66.0) and European Molecular Biology Laboratory (release 19.0) data bases. The initial score was calculated from the best subsequence alignment of two sequences. The optimized score, calculated by considering insertions or deletions, gives a better indication about functional relation between proteins. The statistical significance of the scores was evaluated by Z value [(similarity score- mean of random scores); $Z > 10$ indicates statistical significance]. The similarities between VCP and other sequences were calculated as the percentage of amino acid number having identical residues or functional substitutions relative to the total number of amino acids in VCP (441).

RESULTS AND DISCUSSION

Recently, a female-specific protein (M_r = 53,000), initially called 53KP, has been found in *A. aegypti*. Like vitellogenin, this protein is synthesized by the fat body of vitellogenic females under the control of 20-hydroxyecdysone. The kinetics of 53KP production by the vitellogenic fat body is also similar to that of vitellogenin: it is produced as early as 4 hr and reaches its peak near 24 hr after the initiation of vitellogenesis. Synthesis then drops to low levels by 36 hr and declines to background levels by 48 hr. This protein is secreted to the hemolymph and is selectively accumulated in yolk bodies of developing oocytes (12). Data presented here show that this protein (VCP) is a serine carboxypeptidase.

The cDNA encoding VCP was cloned by a combination of immunoscreening of a λgt11 cDNA library and the polymerase chain reaction. The identity of the cDNA was confirmed by direct sequencing of the N terminus of the purified VCP and by *in vitro* translation of hybrid selected mRNA (data not shown).

Northern blot analysis demonstrated that transcription of the VCP mRNA is limited to female fat bodies and is intiated after a blood meal. Similar to vitellogenin mRNA, the VCP mRNA in the fat body is maximal at the peak of the protein production, 24 hr after a blood meal (figs. 1 and 2).

Fig. 1: Northern analysis of sex- and stage-specific expression of the VCP mRNA transcript. Total RNA was extracted from whole bodies of male mosquitoes (lane 1) and fat bodies of female mosquitoes before a blood meal (lane 2) or 24 and 48 hr after a blood meal (lanes 3 and 4 respectively). An RNA ladder (BRL) was resolved in lane M. (*A*) Agarose gel stained with ethidium bromide. (*B*) Hybridization with 0.84-kb fragment of VCP cDNA. (*C*) Hybridization with 1.8 kb fragment of mosquito vitellogenin A1 gene.

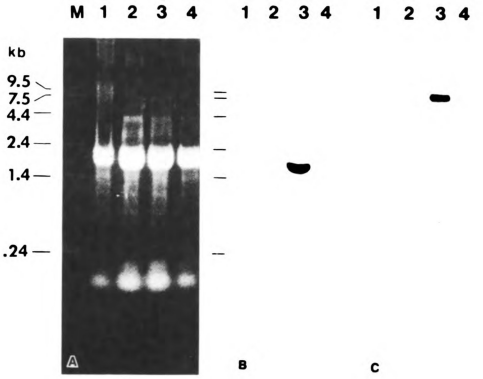
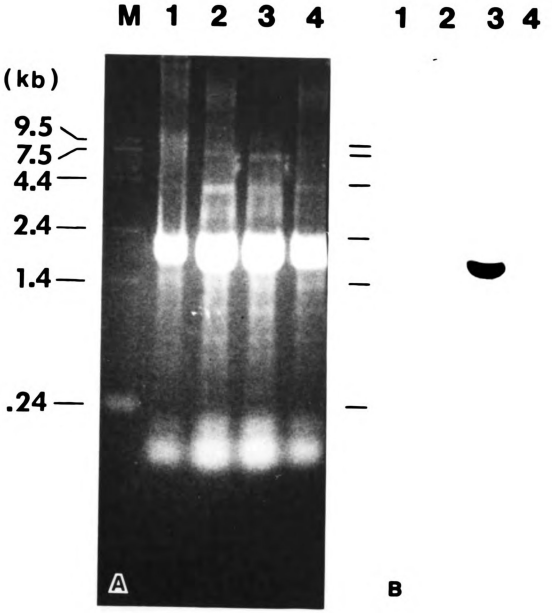

M 1 2 3 4
1 2 3 4
1 2 3 4
kb
9.5
7.5
4.4
2.4
1.4
.24
A
B
C

Fig. 2: Northern blot analysis of tissue-specific expression of the VCP mRNA transcript. Total RNA was extracted from ovaries (lane 1), fat bodies (lane 2), and midgut (lane 3) of female mosquitoes 24 hr after a blood meal. An RNA ladder (BRL) was resolved in lane M. (*A*) Agarose gel stained with ethidium bromide. (*B*) Hybridization with 0.84-kb fragment of VCP cDNA.


M 1 2 3 4
1 2 3 4
(kb)
9.5
7.5
4.4
2.4
1.4
.24
A
B

The sequence of the full length cDNA encoding VCP, confirmed by sequencing of cDNA clones in both directions, is presented in Fig. 3. The size of the VCP mRNA of 1.5 kb, estimated by Northern blot analysis, is in agreement with the 1511-bp mRNA estimated from the cDNA sequence. The VCP cDNA has a single open reading frame that encodes a protein of 441 amino acids with a deduced molecular mass of 50,153 Da. Cell-free translation of VCP mRNA revealed that the VCP precursor is a 50-kDa polypeptide (data not shown).

Hydropathy analysis (19,20) of the deduced amino acid sequence has shown that VCP exhibits properties typical of a secretory protein. The amino acid sequence of VCP has only one potential glycosylation site, at position Asn-135. Glycosylation of VCP by mannose polysaccharide, which accounts for about 2 kDa of the molecular mass, was demonstrated previously (12).

In other insects, small extraovarian proteins are also involved in formation of yolk protein reserves. Microvitellogenin (31 kDa) of *Manduca sexta* and *Hyalophora cecropia* and the 30-kDa protein of *Bombyx mori* are produced by the fat body and deposited in oocytes (21-23). The nucleotide sequence of the *Manduca* microvitellogenin cDNA has 70% similarity to the cDNA sequence coding for the *Bombyx* 30-kDa protein, indicating a close evolutionary relationship between these proteins (24-26).

Comparison of amino acid sequences has not revealed any similarity between mosquito VCP and either *Manduca* microvitellogenin or *Bombyx* 30-kDa protein. Unexpectedly, the VCP sequence exhibited significant similarity with members of a family of serine carboxypeptidases (Table 1). The similarity between the amino acid sequences of mosquito VCP and these carboxypeptidases is highest at the N-terminal portion, which includes two conserved domains (27-32). The highest identity (28.1%) was found in a stretch of 392 amino acids of VCP and wheat carboxypeptidase Y homolog (WCP)

Fig. 3: Nucleotide and deduced amino acid sequences of mosquito VCP. The amino acid sequence matching the N terminus determined from purified VCP is underlined by a solid line. The signal peptide is boxed. The circled amino acid (N) is a potential glycosylation site. A putative polyadenylation signal, AATAAA, is underlined by a broken line.

```
   1  TTCCCACTCGTTGTAAAGTGTAATCGAACAAGCATGGTGAAATTCCATTTACTAGTGCTG   60
                                          M  V  K  F  H  L  L  V  L     9

  61  ATTGCGTTCACCTGCTATACATGCAGCGACGCAACCTTATGGAATCCGTACAAGAAGCTA  120
  10   I  A  F  T  C  Y  T  C  S  D  A  T  L  W  N  P  Y  K  K  L    29

 121  ATGCGAGGATCGGCGTCTCCTCGTCGTCCAGGTGAAAGTGGTGAACCTTTGTTCCTGACT  180
  30   M  R  G  S  A  S  P  R  R  P  G  E  S  G  E  P  L  F  L  T    49

 181  CCACTGTTGCAGGATGGCAAAATTGAAGAGGCTCGCAACAAAGCCCGCGTCAACCATCCC  240
  50   P  L  L  Q  D  G  K  I  E  E  A  R  N  K  A  R  V  N  H  P    69

 241  ATGTTGAGCTCAGTGGAAAGCTACTCCGGTTTTATGACCGTTGATGCCAAGCACAACTCC  300
  70   M  L  S  S  V  E  S  Y  S  G  F  M  T  V  D  A  K  H  N  S    89

 301  AATTTGTTCTTCTGGTATGTTCCAGCGAAGAACAACCGCGAACAAGCGCCCATTCTTGTT  360
  90   N  L  F  F  W  Y  V  P  A  K  N  N  R  E  Q  A  P  I  L  V   109

 361  TGGCTGCAAGGAGGTCCAGGTGCGTCATCGCTGTTTGGAATGTTCGAAGAGAATGGACCG  420
 110   W  L  Q  G  G  P  G  A  S  S  L  F  G  M  F  E  E  N  G  P   129

 421  TTCCATATTCACAGGAACAACTCAGTGAAGCAACGTGAATATTCCTGGCATCAGAACCAT  480
 130   F  H  I  H  R  N  N  S  V  K  Q  R  E  Y  S  W  H  Q  N  H   149

 481  CACATGATCTACATCGATAATCCAGTTGGAACGGGATTCAGTTTCACCGATAGCGATGAA  540
 150   H  M  I  Y  I  D  N  P  V  G  T  G  F  S  F  T  D  S  D  E   169

 541  GGTTACTCCACCAACGAAGAACACGTGGGTGAAAATCTGATGAAATTCATTCAACAATTC  600
 170   G  Y  S  T  N  E  E  H  V  G  E  N  L  M  K  F  I  Q  Q  F   189

 601  TTCGTGCTGTTCCCCAATCTGTTGAAGCATCCATTCTACATCTCCGGTGAATCTTATGGT  660
 190   F  V  L  F  P  N  L  L  K  H  P  F  Y  I  S  G  E  S  Y  G   209

 661  GGAAAGTTTGTTCCTGCCTTCGGTTATGCCATCCACAACTCCCAAAGCCAACCTAAAATC  720
 210   G  K  F  V  P  A  F  G  Y  A  I  H  N  S  Q  S  Q  P  K  I   229

 721  AATCTGCAAGGATTGGCCATTGGTGATGGCTACACTGATCCGCTGAACCAACTTAACTAC  780
 230   N  L  Q  G  L  A  I  G  D  G  Y  T  D  P  L  N  Q  L  N  Y   249

 781  GGAGAATACCTGTATGAGCTGGGCCTGATCGATTTGAACGGAAGAAAGAAGTTCGACGAG  840
 250   G  E  Y  L  Y  E  L  G  L  I  D  L  N  G  R  K  K  F  D  E   269

 841  GATACGGCTGCTGCCATCGCCTGTGCCGAACGTAAGGACATGAAGTGCGCCAACCGCCTT  900
 270   D  T  A  A  A  I  A  C  A  E  R  K  D  M  K  C  A  N  R  L   289

 901  ATCCAAGGTCTGTTCGATGGACTCGATGGACAGGAATCGTACTTCAAGAAGGTCACCGGA  960
 290   I  Q  G  L  F  D  G  L  D  G  Q  E  S  Y  F  K  K  V  T  G   309

 961  TTCTCGTCCTACTACAACTTCATCAAGGGAGACGAGGAGAGCAAACAGGATAGTGTCCTG 1020
 310   F  S  S  Y  Y  N  F  I  K  G  D  E  E  S  K  Q  D  S  V  L   329

1021  ATGGAGTTCCTCAGCAACCCGGAGGTACGTAAGGGCATCCACGTTGGTGAACTGCCGTTC 1080
 330   M  E  F  L  S  N  P  E  V  R  K  G  I  H  V  G  E  L  P  F   349

1081  CACGACTCTGACGGTCACAACAAGGTCGCGGAAATGCTGTCCGAAGACACTCTGGACACC 1140
 350   H  D  S  D  G  H  N  K  V  A  E  M  L  S  E  D  T  L  D  T   369

1141  GTGGCTCCATGGGTCAGCAAGCTGCTCTCGCACTACCGCGTGCTGTTCTACAACGGTCAG 1200
 370   V  A  P  W  V  S  K  L  L  S  H  Y  R  V  L  F  Y  N  G  Q   389

1201  TTGGACATCATCTGCGCCTACCCGATGACGGTCGACTTCCTGATGAAGATGCCTTTCGAT 1260
 390   L  D  I  I  C  A  Y  P  M  T  V  D  F  L  M  K  M  P  F  D   409

1261  GGCGATAGCGAGTACAAGCGGGCCAATCGTGAGATCTACCGCGTGGATCGGAAATCGCCG 1320
 410   G  D  S  E  Y  K  R  A  N  R  E  I  Y  R  V  D  R  K  S  P   429

1321  GGTACAAGAAGCGGGCTGGTCGTCTGCAAGAGGTGCTGATCAGAAACGCCGGACACATGG 1380
 430   G  T  R  S  G  L  V  V  C  K  R  C

1381  TTCCGCGGGACCAGCCGAAATGGGCTTTCGACATGATCACCTCGTTCACTCACAAAAACT 1440

1441  ACTTGTGAGGGAAATGGTAATGATTTGGATGAATAAAGCTTTAAGCTGTAATTAAAAAAA 1500

1501  AAAAAAAAAAA
```

Table 1. Homology of mosquito VCP to several serine carboxypeptidases.

TABLE 1. Homology of mosquito VCP to serine carboxypeptidases.

		Score		Z-value	
		I	O	I	O
Wheat:	Carboxypeptidase Y homolog	292	523	45.4	81.1
Yeast:	Carboxypeptidase Y	209	387	30.4	43.2
Human:	Protective protein	194	422	38.4	59.2
Mouse:	Protective protein	190	434	25.9	54.8
Yeast:	KEX1 carboxypeptidase	182	339	37.3	40.0
Barley:	Serine carboxypeptidase I	139	394	18.5	42.9
Barley:	Serine carboxypeptidase II	116	380	16.5	49.1

I - initial score, O - optimized score, Z-value - statistical significance.

(Fig. 4). However, considering both identical and conservative replacements, similarity between VCP and WCP is 62.3%. In mosquito VCP the conserved domains are at positions Trp-110 to Ser-119 and Gly-205 to Gly-210 (Fig. 4). The Ser-207 of the VCP second domain corresponds to Ser-257 of yeast carboxypeptidase Y, in a sequence that was shown to be the catalytic center of serine carboxypeptidases (33). In contrast to mosquito VCP and WCP (Fig. 4), all other members of the serine carboxypeptidase family have three conserved domains (27-32). Functional implications of these differences are not clear. Unlike the situation for serine endopeptidases, the importance of three domains for carboxypeptidase activity has not been established (34). Similar to all other serine carboxypeptidases in the family (27-32), mosquito VCP does not have any significant homology with any known sequences of trypsin-like serine endopeptidases or esterases.

Internalization of VCP in mosquito oocytes occurs without any changes in its size (12). The molecular mass of VCP, however, decreases by 0.5-1.0 kDa at the onset of embryonic development (Fig. 5). As embryogenesis progresses, VCP undergoes further reduction in its size and another, immunologically related 48 kDa band appears (Fig. 5). By the end of embryonic development, VCP degrades into smaller peptides, which disappear once the first-instar larva hatches (Fig. 5).

Radiolabeled serine protease inhibitor, [^{3}H]DFP, which binds to the active center of serine proteases (35), binds weakly to VCP in oocytes (Fig. 5, lanes 1). The intensity of binding to VCP, however, increases at the onset and reaches maximum in the middle of embryogenesis, when the inhibitor binds to a VCP band of 48 kDa (Fig. 5). The binding of DFP to VCP can be inhibited by prior treatment of protein extracts with other serine protease inhibitors, phenylomethylsulfonyl fluoride or leupeptin (data not shown). On the basis of

Fig. 4: Alignment of amino acid sequences of mosquito VCP and WCP. Amino acid sequences deduced from corresponding cDNAs were aligned by FASTA computer program. Vertical lines indicate identical residues and colons denote functional substitutions. Only portions of both amino acid sequences, flanking the overlapping regions, are presented. Two conserved domains found in sequences of serine carboxypeptidases are boxed. The serine protease catalytic center is marked by an asterisk.

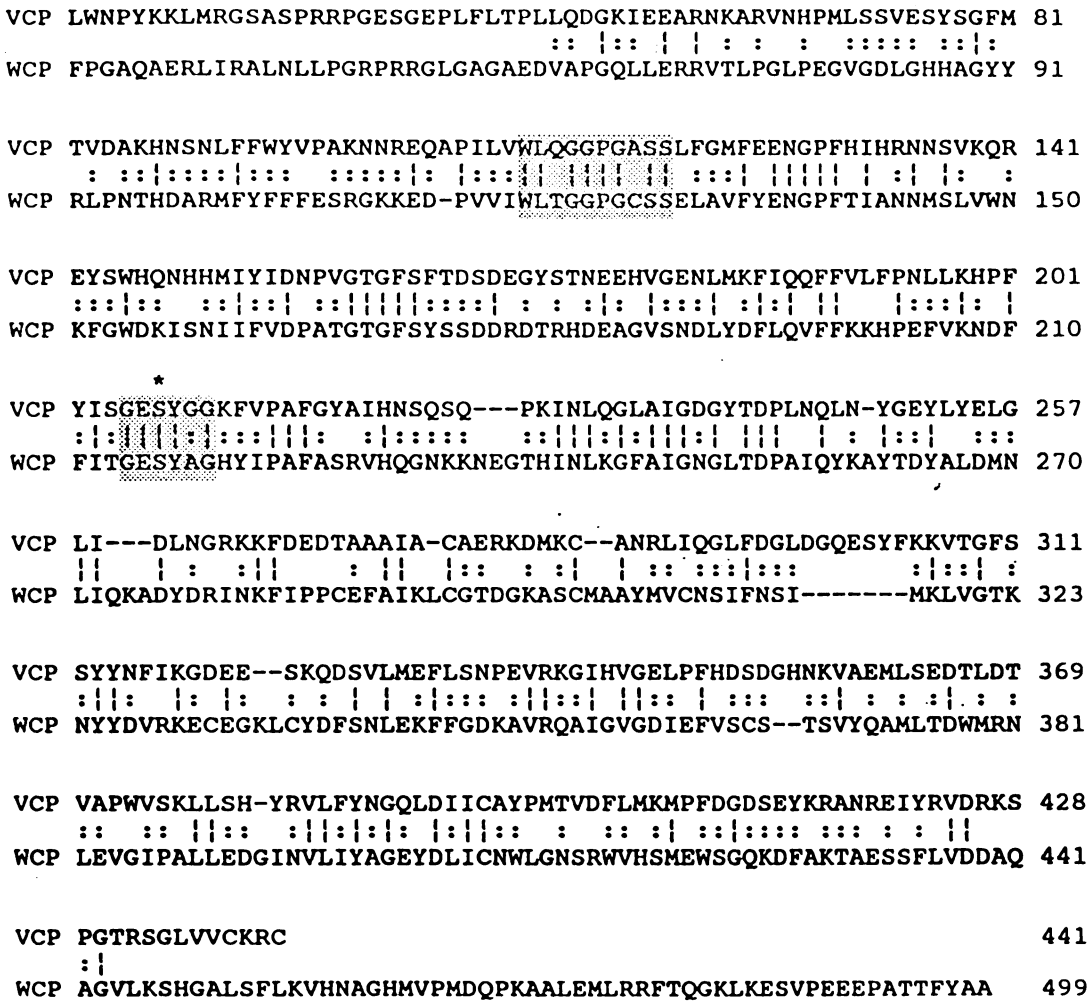
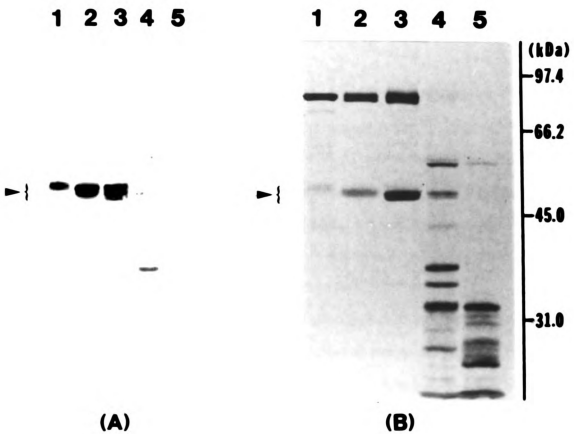

VCP LWNPYKKLMRGSASPRRPGESGEPLFLTPLLQDGKIEEARNKARVNHPMLSSVESYSGFM 81
WCP FPGAQAERLIRALNLLPGRPRRGLGAGAEDVAPGQLLERRVTLPGLPEGVGDLGHHAGYY 91
VCP TVDAKHNSNLFFWYVPAKNNREQAPILVWLQGGPGASSLFGMFEENGPFHIHRNNSVKQR 141
WCP RLPNTHDARMFYFFFESRGKKED-PVVIWLTGGPGCSSELAVFYENGPFTIANNMSLVWN 150
VCP EYSWHQNHHMIYIDNPVGTGFSFTDSDEGYSTNEEHVGENLMKFIQQFFVLFPNLLKHPF 201
WCP KFGWDKISNIIFVDPATGTGFSYSSDDRDTRHDEAGVSNDLYDFLQVFFKKHPEFVKNDF 210
VCP YISGESYGGKFVPAFGYAIHNSQSQ---PKINLQGLAIGDGYTDPLNQLN-YGEYLYELG 257
WCP FITGESYAGHYIPAFASRVHQGNKKNEGTHINLKGFAIGNGLTDPAIQYKAYTDYALDMN 270
VCP LI---DLNGRKKFDEDTAAAIA-CAERKDMKC--ANRLIQGLFDGLDGQESYFKKVTGFS 311
WCP LIQKADYDRINKFIPPCEFAIKLCGTDGKASCMAAYMVCNSIFNSI-------MKLVGTK 323
VCP SYYNFIKGDEE--SKQDSVLMEFLSNPEVRKGIHVGELPFHDSDGHNKVAEMLSEDTLDT 369
WCP NYYDVRKECEGKLCYDFSNLEKFFGDKAVRQAIGVGDIEFVSCS--TSVYQAMLTDWMRN 381
VCP VAPWVSKLLSH-YRVLFYNGQLDIICAYPMTVDFLMKMPFDGDSEYKRANREIYRVDRKS 428
WCP LEVGIPALLEDGINVLIYAGEYDLICNWLGNSRWVHSMEWSGQKDFAKTAESSFLVDDAQ 441
VCP PGTRSGLVVCKRC 441
WCP AGVLKSHGALSFLKVHNAGHMVPMDQPKAALEMLRRFTQGKLKESVPEEEPATTFYAA 499

Fig. 5: Immunoblot (*A*) and [^{3}H]DFP binding (*B*) analyses of VCP during mosquito embryonic development. Both panels contain protein extracts from the following: lane 1, mosquito ovaries removed 24 hrs after a blood meal during peak yolk accumulation; lane 2, 0- to 3-hr postoviposition eggs at the onset of embryogenesis; lane 3, 44- to 47-hr postoviposition eggs during midembryogenesis; lane 4, 94- to 97-hr postoviposition eggs at the end of embryogenesis; lane 5, first-instar larva. The mosquito VCP bands are shown by arrows. The high molecular mass polypeptide bound to DFP (*B*) is not immunologically related to VCP. In both panels, proteins were resolved by SDS/PAGE on 9% gels under reducing conditions. The molecular mass standards in order of decreasing mass are phosphorylase b, bovine serum albumin, ovalbumin, and carbonic anhydrase (Bio-Rad).


1 2 3 4 5
1 2 3 4 5
(kDa)
97.4
66.2
45.0
31.0
(A)
(B)

the above results, we conclude that VCP is synthesized by the fat body and internalized by oocytes as a proenzyme, which is then activated in eggs at the onset of embryogenesis. The activation of VCP is associated with an increase in its electrophorectic mobility. It is not clear, however, whether this increase in VCP mobility is due to its proteolytic cleavage or deglycosylation. The mouse carboxypeptidase (protective protein) is activated as result of cleavage and DFP binds only to the activated carboxypeptidase subunit that contains the serine catalytic center (31, 32).

Some of the serine carboxypeptidases with which VCP shares significant homology, such as human and mouse protective proteins and yeast KEX1 carboxypeptidase, are implicated in proteolytic activation of a number of enzymes or other biologically active molecules (29-32). The mosquito VCP could play a similar role by acitvating hydrolytic enzymes that are involved in degradation of yolk proteins in developing embryos. Alternatively, it could function as an exopeptidase in sequential degradation of vitellogenin.

The mosquito VCP is different from proteases known to hydrolyze yolk proteins in insect embryos: an acidic cathepsin B-like protease of *Drosophila* and a thiol protease or cathepsin L-like protease of *Bombyx* (8-11). Similar to mosquito VCP, all these proteases are deposited as proenzymes in yolk bodies of eggs and activated in embryos. *Drosophila* cathepsin B-like protease is activated as a result of proteolysis by a serine protease (25 kDa) which is also present in egg yolk bodies (9). It is not known, however, whether any of these enzymes are of extraovarian origin or whether they are synthesized by developing oocytes themselves. The mosquito VCP, therefore, is, to our knowledge, the first example among oviparous animals of a proteolytic enzyme shown to be produced by an extraovarian tissue and accumulated by oocytes for use in embryonic development.

ACKNOWLEDGEMENTS

We thank Dr. T.S. Dhadialla and Mr. A.R. Hays for their advice in protein purification, Dr. H.H. Hagedorn for his generous gift of vitellogenin gene probe, and Drs. T.S. Dhadialla, T.B. Friedman, J.R. Miller, and N.V. Raikhel for their critical reading of the manuscript. This work was supported by grants from the National Institutes of Health and from the Biotechnology Center of Michigan State University to A.S.R.

REFERENCES

1. Kunkel, J.G. and Nordin, J.H. (1985) in *Comprehensive Insect Physiology, Biochmistry and Pharmocology,* eds. Kerkut, G.A. and Gilbert, L.I. (Plenum, New York), Vol. 1, pp.83-111.

2. Wyatt, G.R. (1988) *Can. J. Zool.* **66,**2600-2610.

3. Dhadialla, T.S. and Raikhel, A.S. (1990) *J. Biol. Chem.* **265,** 9924-9933.

4. Kanost, M.R., Kawooya, J.K., Law, J.H., Ryan, R.O., VanHeusden, M.C. and Ziegler, R. (1990) *Adv. Insect Physiol.* **22,** 299-396.

5.Whali, W. (1988) *Trends Genet.* **4,** 227-232.

6. Ferenz, H.-J. (1990) in *Molecular Insect Science*, eds Hagedorn, H.H., Hildebrand, J.G., Kidwell, M.G. and Law, J.H. (Plenum, New York), pp131-138.

7. Raikhel, A.S. and Dhadialla, T.S. (1992) *Annu. Rev. Entomol.* **37,** 217-251.

8. Medina, M., Leon, P., and Vallejo, C.G. (1988) *Arch Biochem. Biophys.* **263,** 355-363.

9. Medina, M. and Vallejo, C.G. (1989) *Insect Biochem.* **19,** 687-691.

10. Indrasith, L.S., Sasaki, T. and Yamashita, O. (1988) *J Biol. Chem.* **263,** 1045-1051.

11. Kageyama, T. and Takahashi, S. (1990) *Eur. J. Biochem.* **193,** 203-210.

12. Hays, A.R. and Raikhel, A.S. (1990) *Roux's Arch. Dev. Biol.* **199,** 114-121.

13. Raminani, L.N. and Cupp E.W. (1978) *Int. J. Insect Morphol. Embryol.* **7,** 273-296.

14. Sambrook, J., Fritsch, E.F., and Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.

15. Lee, C.C., Wu, X., Gibbs, R.A., Cook, R.G., Muzny, D.M. and Caskey, C.T. (1988) *Science* **239,** 1288-1291.

16. Cho, W.-L., Dhadialla, T.S. and Raikhel, A.S. (1991) *Insect Biochem.* **21,** 165-176.

17. Bose, S.G. and Raikhel, A.S. (1988) *Biochem. Biophys. Res. Commun.* **155,** 436-442.

18. Lipman, D.J. and Pearson, W.R. (1985) *Science* **227,** 1435-1441.

19. Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* **157,** 105-132.

20. von Heijne, G. (1983) *Eur. J. Biochem.* **133,** 17-21.

21. Kawooya, J.K., Osir, E.O. and Law, J.H. (1986) *J. Biol. Chem.* **261,** 10844-10849.

22. Zhu, J., Indrasith, L.S. and Yamashita, O. (1986) *Biochim. Biohpys. Acta* **882,** 427-436.

23. Kulakosky, P.C. and Telfer, W.H. (1987) *Insect Biochem.* **17,** 845-858.

24. Sakai, N., Mori, S., Izumi, S., Haino-Fukishima, K., Ogura, T., Maikawa, H. and Tomino, S. (1988) *Biochem. Biophys. Acta* **949,** 224-232.

25. Wang, X.-Y., Cole, K.D. and Law, J.H. (1988) *J. Biol. Chem.* **263,** 8851-8855.

26. Wang, X.-Y., Cole, K.D. and Law, J.H. (1989) *Gene* **80,** 259-268.

27. Baulcombe, D.C., Backer, R.F. and Jarvis, M.G. (1987) *J. Biol. Chem.* **262,** 13726-13735.

28. Sorenson, S.B., Svendsen, I.B. and Breddam, K. (1987) *Carlsberg Res. Commun.* **52,** 285-295.

29. Dmochowska, A., Dignard, D., Henning, D., Thomas, D.Y. and Bussey, H. (1987) *Cell* **50,** 573-584.

30. Valls, L.A., Hunter, C.P., Rothman, J.H. and Stevens, T.H. (1987) *Cell* **48,** 887-897.

31. Galijart, N.J., Gillemans, N., Harris, A., van der Horst, G.T.J., Verheijen, F.W., Glajaard, H. and d'Asso, A. (1988) *Cell* **54,**755-764.

32. Galijart, N.J., Gillemans, N., Meijer, D. and d'Asso, A. (1990) *J. Biol. Chem.* **265,** 4678-4684.

33. Hayashi, R., Moore, S. and Stein, W.H. (1973) *J. Biol. Chem.* **248,** 8366-8396.

34. Wilson, K.P., Liao, D.-I., Bullock, T., Remington, S.J. and Breddam, K. (1990) *J. Mol. Biol.* **211,** 301-303.

35. Scheiner, C.J. and Quigley, J.P. (1982) *Anal. Biochem.* **122,** 58-69.

CHAPTER 3*

CLONING AND ANALYSIS OF THE LOCUS FOR MOSQUITO VITELLOGENIC CARBOXYPEPTIDASE

*Originally published in Insect Molecular Biology

Reference: Deitsch, K.W. and Raikhel, A.S. (1993) Insect Mol. Biol. **2**(4), 205-213.

ABSTRACT

Vitellogenic carboxypeptidase is a 53-kDa yolk protein produced by the fat body of the female mosquito, *Aedes aegypti*, in response to a blood meal. Its expression is sex-, stage-, and tissue-specific and is identical to that of the major yolk protein, vitellogenin. The gene is intronless and two alleles have been cloned and sequenced, including more than 1.5 kb on both sides of the coding region. A capsite consensus recently identified as an arthropod initiator was identified at the start site of transcription. Upstream of this capsite is a 16 bp imperfect palindrome repeated 4 times showing strong homology to defined hormone-response elements. In addition, a region that closely resembles the fat body enhancer and double sex binding site from the *Drosophila* yolk protein genes and several potential fat body-specific regulatory protein binding sites were found.

INTRODUCTION

The reproduction of mosquitoes is a subject of intense interest due to the vectoral capacity of these insects. A key process in egg maturation is vitellogenesis, the tightly regulated mass production of yolk protein precursors by the fat body, their secretion into the hemolymph and subsequent uptake by the developing eggs. This process is unique as it is held in a state-of-arrest in adult females of anautogenous mosquitoes until a blood meal can be obtained, at which time the vitellogenic genes are expressed exclusively in the fat body at extremely high levels, after which expression ceases (for review, see Raikhel, 1992). This provides an excellent model system for the study of tissue-, sex- and stage-specific gene expression. Analysis of vitellogenic genes should elucidate the molecular mechanisms behind mosquito egg production and lead to possible novel methods of mosquito control.

The onset of vitellogenesis in the mosquito *Aedes aegypti* is marked by falling titers of juvenile hormone (JH) and rapidly rising titers of 20-hydroxyecdysone (20E) (Hagedorn, 1983, 1985; Raikhel, 1992). The production of yolk proteins in the fat body is correlated with these hormonal changes, increasing with rising titers of 20E and ceasing when this steroid hormone's levels decrease to pre-blood meal levels. *In vitro* experiments have shown that physiological doses of 20E are capable of stimulating synthesis of the major yolk protein, vitellogenin (Vg) in previtellogenic fat bodies (Fuchs and Kang, 1981; Hagedorn, 1985). However, *in vivo*, large doses of 20E are needed to generate similar levels of Vg synthesis (Fuchs and Kang, 1981; Lea, 1982). In addition, Vg production is detectable only 1 hr after a blood meal, before 20E titers begin

to rise, at both the protein (Raikhel and Lea, 1983) and the mRNA levels (Racioppi *et al.*, 1986). These results indicate that the regulation of expression of the vitellogenic genes is complex, and that 20E plays a significant, although unclear role in this regulation. Little research into the control of expression of the genes involved in vitellogenesis has been published. Gemmill *et al.* (1986) reported the isolation of the mosquito Vg genes and their induction of expression by 20E. However, the analysis of the genes for this protein is difficult due to their large size (6.5 kb mRNA) and the fact that they are transcribed from a four or five member gene family.

Recently, a unique vitellogenic protein was discovered in the mosquito *A. aegypti* that is synthesized in a manner indistinguishable from that of Vg (Hays and Raikhel, 1990). The cDNA encoding this protein was cloned and sequenced (Cho *et al.*, 1991). The deduced amino acid sequence shares significant homology with members of the serine carboxypeptidase family. The protein is synthesized and secreted as a 53 kDa proenzyme by the vitellogenic fat body. It is internalized by developing oocytes and then stored in the egg until embryogenesis, at which time it undergoes a reduction in size to the active form of 48 kDa. The highest levels of the active enzyme are found at midembryogenesis, after which the protein disappears. These results show that an enzyme involved in embryogenesis is produced by an extraovarian tissue, stored in eggs and activated later in development; a phenomenon not previously known for oviparious animals. Reflecting these unique characteristics, the protein was named vitellogenic carboxypeptidase (VCP).

Here we describe the analysis of the locus for VCP, a gene that shows identical temporal, tissue and levels of expression to vitellogenin, but that is single copy, intronless and much smaller. Two alleles for this gene have been cloned and sequenced, including ~1.5 kb on both sides of the coding region.

Analysis of the regulatory regions provides a starting point for molecular characterization of the hormone responsiveness of the genes involved in vitellogenesis.

RESULTS

Cloning and Sequencing of VCP Genomic Fragments

Previously, the cDNA encoding VCP had been cloned and sequenced (Cho *et al.*, 1991). Using an 840 bp fragment of this sequence as a probe, several genomic clones were obtained by screening a λ Fix II *Aedes aegypti* genomic library. Restriction mapping indicated that these DNA fragments could be classified into two groups, arbitrarily named VCP^1 and VCP^2 (Fig. 1). Genomic blots probed with the VCP cDNA showed bands corresponding to both VCP^1 and ^{2}. Because preparations of genomic DNA are typically extracted from large numbers of mosquitoes which are not genetically homogeneous, it was not clear whether VCP^1 and VCP^2 represented two copies of the VCP gene in the *A. aegypti* haploid genome, or simply two separate alleles existing within the population. To differentiate between these two possibilities, genomic blots of DNA extracted from individual mosquitoes were probed with the VCP cDNA. The resulting autoradiograph showed clearly that VCP^1 and VCP^2 are two single copy alleles and that homozygotes for both alleles as well as heterozygotes exist (Fig. 2). The results were identical regardless of whether DNA was extracted from male or female individuals. The intensity of hybridization was similar when comparing lanes of male DNA to DNA extracted from females or female fat bodies, indicating that specific amplification was unlikely. To confirm the copy number determination, 0.5, 1.0 and 2.0 copy equivalents of the VCP cDNA were blotted next to 10 μg of genomic DNA and the intensity of the hybridization with the VCP probe compared. This method also indicated a single copy gene (data not shown). Recently, using restriction fragment length polymorphisms, Severson *et al.* (1993) have mapped VCP to a

Fig. 1: Restriction maps of the two alleles of VCP. The open box represents the coding region. Hatched box represents the inserted DNA unique to VCP2. The striped white bars show the downstream non-homologous region of the two alleles. The length of each restriction fragment is indicated in base pairs above the map. The approximate site of several putative transcriptional control elements is also indicated.

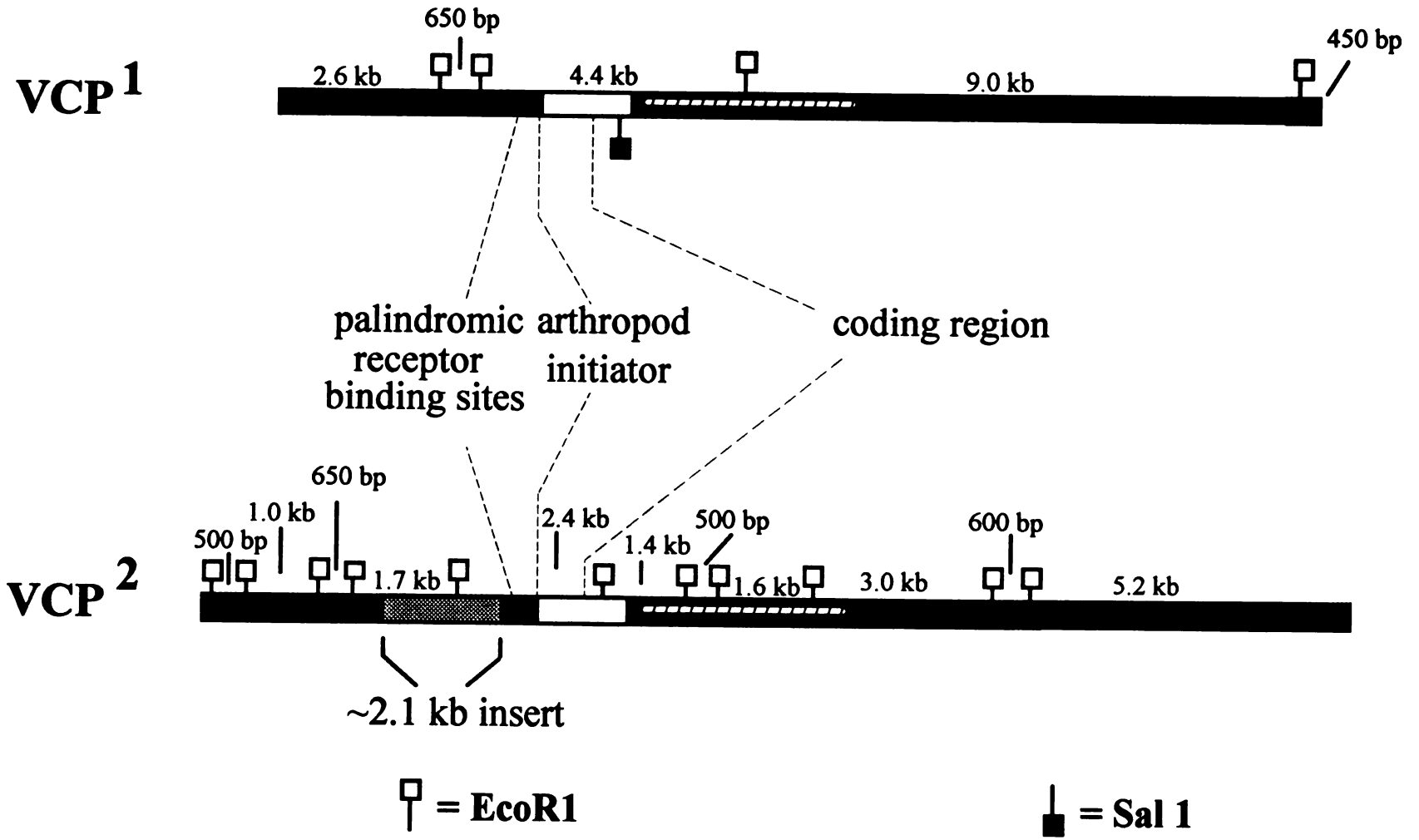
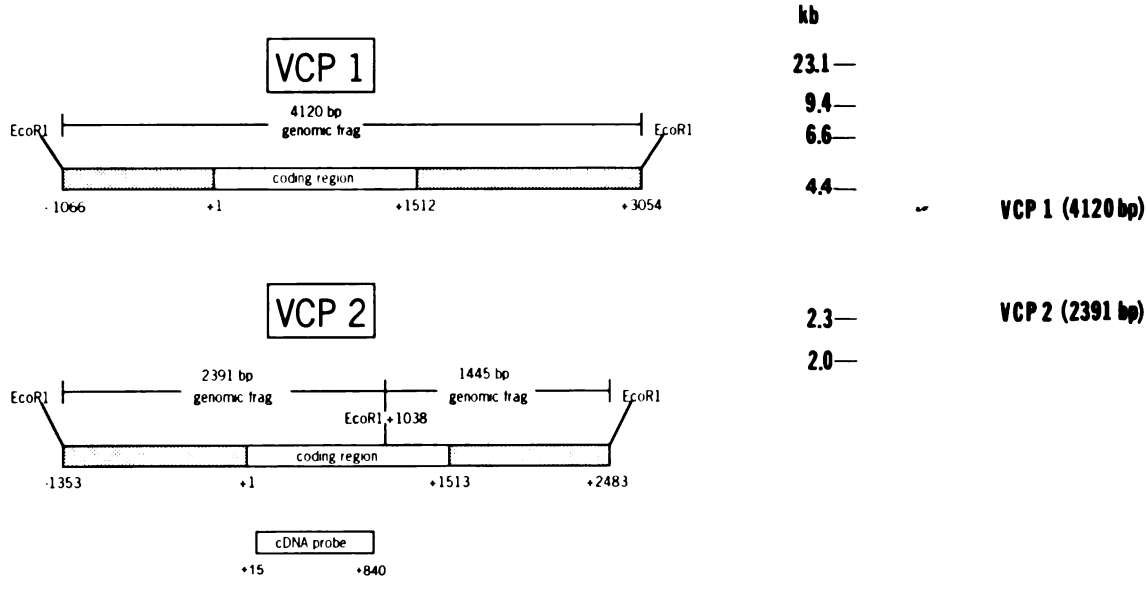

VCP 1
650 bp
450 bp
2.6 kb
4.4 kb
9.0 kb
palindromic receptor binding sites
arthropod initiator
coding region
VCP 2
650 bp
1.0 kb
500 bp
2.4 kb
500 bp
600 bp
1.4 kb
1.7 kb
1.6 kb
3.0 kb
5.2 kb
~2.1 kb insert
= EcoR1
= Sal 1

Fig. 2: Genomic blots of DNA from individual mosquitoes. (A) Diagram showing the hybridizating fragments distinct to the two alleles. (B) Blots showing individuals homozygous for VCP^2 (lane 1), VCP^1 (lane 2) and the corresponding heterozygotes showing both alleles (lane 3).

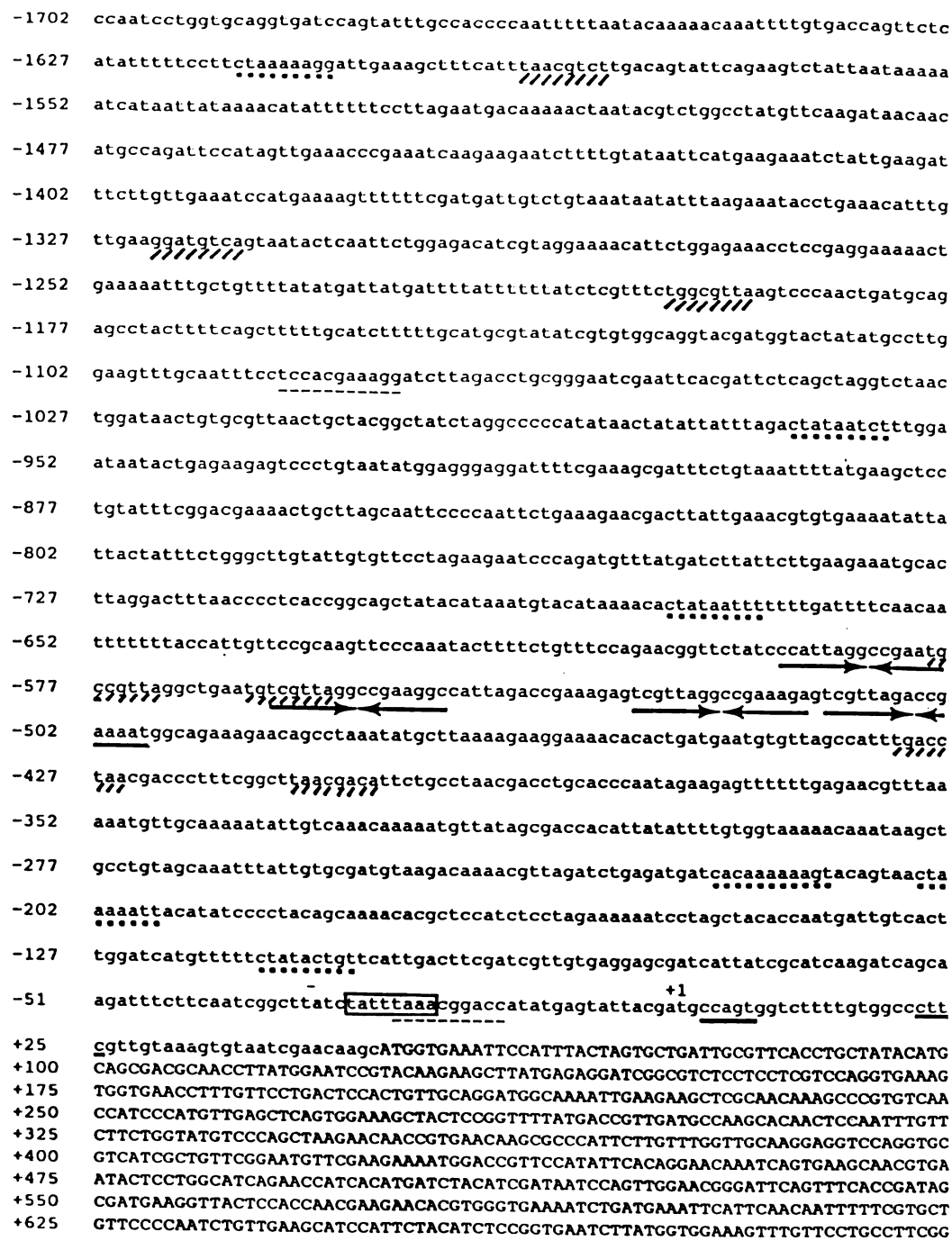

VCP 1
4120 bp
genomic frag
EcoR1
EcoR1
coding region
-1066
+1
+1512
+3054
VCP 2
2391 bp
genomic frag
1445 bp
genomic frag
EcoR1
EcoR1
EcoR1 +1038
coding region
-1353
+1
+1513
+2483
cDNA probe
+15
+840
kb
23.1
9.4
6.6
4.4
2.3
2.0
VCP 1 (4120 bp)
VCP 2 (2391 bp)

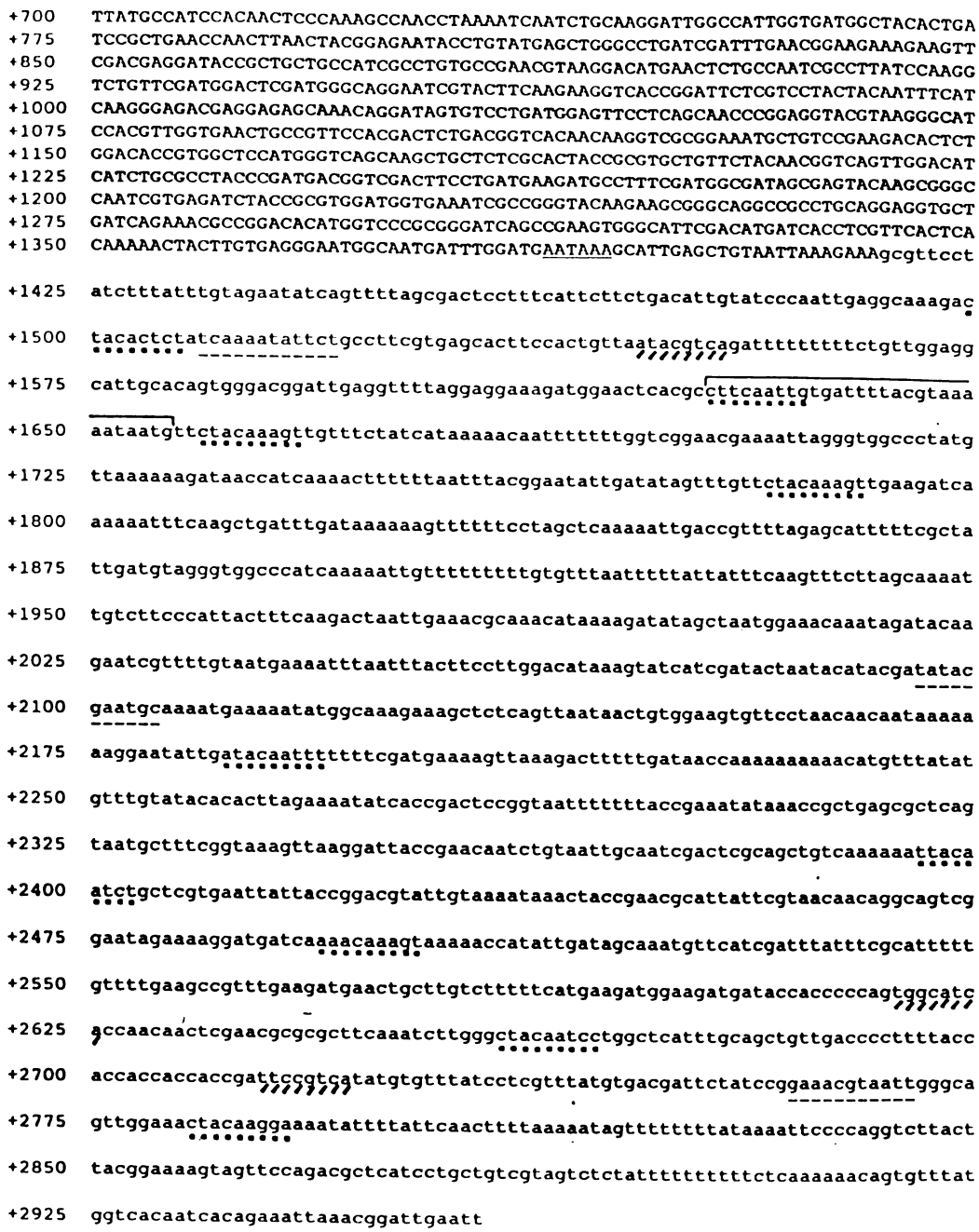
single site in the mosquito genome between the marker LF181 and the γ-aminobutyric receptor subunit A receptor gene.

The differences in the restriction maps of the two alleles for VCP are much greater than one would anticipate for two alleles found within a single inbred population. When a second gene was cloned from this population of mosquitoes, that encoding lysosomal aspartic protease, a number of distinctly different alleles were also found (Dittmer *et al.*, unpublished). This lab strain of *Aedes aegypti* was founded approximately 20 years ago, and the origin of the founding mosquitoes is not known. It is possible that individuals from distinctly different populations were brought together in the initial founding of the colony. This could have led to the rather stark differences in the alleles shown here for VCP.

1,716 bp upstream and 1,542 bp downstream of the coding region in VCP^1 and 1,353 bp upstream and 970 bp downstream of the coding region in VCP^2 were sequenced. The coding regions in both VCP^1 and VCP^2 included no introns and encoded 100% identical polypeptides. The nucleotide sequence of the two alleles is also highly conserved (~97% identity) with the exception of two unique, non-homologous areas (Fig. 1). Sequence analysis has shown that 614 bp upstream of the start site of VCP^2 a ~2.1 kb fragment exists which is not found in VCP^1. This represents either an insertion in VCP^2 or a deletion in VCP^1. Also, cross-hybridization analysis of the two alleles indicates that downstream of the open reading frames there is a ~3 kb region with considerably less homology. The complete nucleotide sequence of VCP^1 is shown in Fig. 3, including a putative TATA box and ribosomal binding site as defined by Hagenbuchle *et al.* (1978).

Fig. 3: Complete genomic sequence of VCP[1]: Inverted arrows represent the putative hormone response elements (-561 to -499); the single black bar shows the arthropod initiator (+2 to +6); the TATA sequence is boxed (-22 to -29); +1 indicates the start site of transcription; capital letters show the translated portion of the gene; double underline is the putative ribosomal binding site (+22 to +25); single underline shows the polyadenylation signal (+1389 to +1394); ⌐——¬ identifies the region of similarity with the *Drosophila* yolk protein fat body enhancer (+1627 to +1657); • • • • shows the consensus double sex binding sequence; ------ identifies sequences with similarity to the BBF-2 binding site; /// shows sequences with similarity to CREB/ATF binding sites.


-1702 ccaatcctggtgcaggtgatccagtatttgccaccccaatttttaatacaaaaacaaattttgtgaccagttctc
-1627 atatttttccttctaaaaaggattgaaagctttcatttaacgtcttgacagtattcagaagtctattaataaaaa
-1552 atcataattataaaaacatatttttttccttagaatgacaaaaactaatacgtctggcctatgttcaagataacaac
-1477 atgccagattccatagttgaaacccgaaatcaagaagaatcttttgtataattcatgaagaaatctattgaagat
-1402 ttcttgttgaaatccatgaaaagtttttttcgatgattgtctgtaaataatatttaagaaatacctgaaacatttg
-1327 ttgaaggatgtcagtaatactcaattctggagacatcgtaggaaaacattctggagaaacctccgaggaaaaact
-1252 gaaaaatttgctgttttatatgattatgattttatttttttatctcgtttctggcgttaagtcccaactgatgcag
-1177 agcctacttttcagctttttgcatctttttgcatgcgtatatcgtgtggcaggtacgatggtactatatgccttg
-1102 gaagtttgcaatttcctccacgaaaggatcttagacctgcgggaatcgaattcacgattctcagctaggtctaac
-1027 tggataactgtgcgttaactgctacggctatctaggcccccatataactatattatttagactataatctttgga
-952 ataatactgagaagagtccctgtaatatggagggaggattttcgaaagcgatttctgtaaatttttatgaagctcc
-877 tgtatttcggacgaaaactgcttagcaattccccaattctgaaagaacgacttattgaaacgtgtgaaaatatta
-802 ttactatttctgggcttgtattgtgttcctagaagaatcccagatgtttatgatcttattcttgaagaaatgcac
-727 ttaggactttaacccctcaccggcagctatacataaatgtacataaaacactataatttttttgattttcaacaa
-652 ttttttttaccattgttccgcaagttcccaaatacttttctgtttccagaacggttctatcccattaggccgaatg
-577 ccgttaggctgaatgtcgttaggccgaaggccattagaccgaaagagtcgttaggccgaaagagtcgttagaccg
-502 aaaatggcagaaagaacagcctaaatatgcttaaaagaaggaaaacacactgatgaatgtgttagccatttgacc
-427 taacgacccttttcggcttaacgacattctgcctaacgacctgcacccaatagaagagtttttttgagaacgtttaa
-352 aaatgttgcaaaaatattgtcaaacaaaaatgttatagcgaccacattatattttgtggtaaaaacaaataagct
-277 gcctgtagcaaatttattgtgcgatgtaagacaaaacgttagatctgagatgatcacaaaaaagtacagtaacta
-202 aaaattacatatcccctacagcaaaacacgctccatctcctagaaaaaatcctagctacaccaatgattgtcact
-127 tggatcatgtttttctatactgttcattgacttcgatcgttgtgaggagcgatcattatcgcatcaagatcagca
-51 agatttcttcaatcggcttatctatttaaacggaccatatgagtattacgatgccagtggtcttttgtggcccctt
+25 cgttgtaaagtgtaatcgaacaagcATGGTGAAATTCCATTTACTAGTGCTGATTGCGTTCACCTGCTATACATG
+100 CAGCGACGCAACCTTATGGAATCCGTACAAGAAGCTTATGAGAGGATCGGCGTCTCCTCCTCGTCCAGGTGAAAG
+175 TGGTGAACCTTTGTTCCTGACTCCACTGTTGCAGGATGGCAAAATTGAAGAAGCTCGCAACAAAGCCCGTGTCAA
+250 CCATCCCATGTTGAGCTCAGTGGAAAGCTACTCCGGTTTTATGACCGTTGATGCCAAGCACAACTCCAATTTGTT
+325 CTTCTGGTATGTCCCAGCTAAGAACAACCGTGAACAAGCGCCCATTCTTGTTTGGTTGCAAGGAGGTCCAGGTGC
+400 GTCATCGCTGTTCGGAATGTTCGAAGAAAATGGACCGTTCCATATTCACAGGAACAAATCAGTGAAGCAACGTGA
+475 ATACTCCTGGCATCAGAACCATCACATGATCTACATCGATAATCCAGTTGGAACGGGATTCAGTTTCACCGATAG
+550 CGATGAAGGTTACTCCACCAACGAAGAACACGTGGGTGAAAATCTGATGAAATTCATTCAACAATTTTTCGTGCT
+625 GTTCCCCAATCTGTTGAAGCATCCATTCTACATCTCCGGTGAATCTTATGGTGGAAAGTTTGTTCCTGCCTTCGG


+700 TTATGCCATCCACAACTCCCAAAGCCAACCTAAAATCAATCTGCAAGGATTGGCCATTGGTGATGGCTACACTGA
+775 TCCGCTGAACCAACTTAACTACGGAGAATACCTGTATGAGCTGGGCCTGATCGATTTGAACGGAAGAAAGAAGTT
+850 CGACGAGGATACCGCTGCTGCCATCGCCTGTGCCGAACGTAAGGACATGAACTCTGCCAATCGCCTTATCCAAGG
+925 TCTGTTCGATGGACTCGATGGGCAGGAATCGTACTTCAAGAAGGTCACCGGATTCTCGTCCTACTACAATTTCAT
+1000 CAAGGGAGACGAGGAGAGCAAACAGGATAGTGTCCTGATGGAGTTCCTCAGCAACCCGGAGGTACGTAAGGGCAT
+1075 CCACGTTGGTGAACTGCCGTTCCACGACTCTGACGGTCACAACAAGGTCGCGGAAATGCTGTCCGAAGACACTCT
+1150 GGACACCGTGGCTCCATGGGTCAGCAAGCTGCTCTCGCACTACCGCGTGCTGTTCTACAACGGTCAGTTGGACAT
+1225 CATCTGCGCCTACCCGATGACGGTCGACTTCCTGATGAAGATGCCTTTCGATGGCGATAGCGAGTACAAGCGGGC
+1200 CAATCGTGAGATCTACCGCGTGGATGGTGAAATCGCCGGGTACAAGAAGCGGGCAGGCCGCCTGCAGGAGGTGCT
+1275 GATCAGAAACGCCGGACACATGGTCCCGCGGGATCAGCCGAAGTGGGCATTCGACATGATCACCTCGTTCACTCA
+1350 CAAAAACTACTTGTGAGGGAATGGCAATGATTTGGATGAATAAAGCATTGAGCTGTAATTAAAGAAAgcgttcct
+1425 atctttatttgtagaatatcagttttagcgactcctttcattcttctgacattgtatcccaattgaggcaaagac
+1500 tacactctatcaaaatattctgccttcgtgagcacttccactgttaatacgtcagatttttttttctgttggagg
+1575 cattgcacagtgggacggattgaggttttaggaggaaagatggaactcacgccttcaattgtgattttacgtaaa
+1650 aataatgttctacaaagttgtttctatcataaaaacaatttttttggtcggaacgaaaattagggtggccctatg
+1725 ttaaaaaagataaccatcaaaacttttttaatttacggaatattgatatagtttgttctacaaagttgaagatca
+1800 aaaaatttcaagctgatttgataaaaaagtttttttcctagctcaaaaattgaccgttttagagcatttttcgcta
+1875 ttgatgtagggtggcccatcaaaaattgtttttttttgtgtttaatttttattatttcaagtttcttagcaaaat
+1950 tgtcttcccattactttcaagactaattgaaacgcaaacataaaagatatagctaatggaaacaaatagatacaa
+2025 gaatcgttttgtaatgaaaatttaatttacttccttggacataaagtatcatcgatactaatacatacgatatac
+2100 gaatgcaaaatgaaaaatatggcaaagaaagctctcagttaataactgtggaagtgttcctaacaacaataaaaa
+2175 aaggaatattgatacaatttttttcgatgaaaagttaaagacttttgataaccaaaaaaaaaacatgtttatat
+2250 gtttgtatacacacttagaaaatatcaccgactccggtaatttttttaccgaaatataaaccgctgagcgctcag
+2325 taatgctttcggtaaagttaaggattaccgaacaatctgtaattgcaatcgactcgcagctgtcaaaaaattaca
+2400 atctgctcgtgaattattaccggacgtattgtaaaataaactaccgaacgcattattcgtaacaacaggcagtcg
+2475 gaatagaaaaggatgatcaaaacaaagtaaaaaccatattgatagcaaatgttcatcgatttatttcgcatttttt
+2550 gttttgaagccgtttgaagatgaactgcttgtctttttcatgaagatggaagatgataccacccccagtggcatc
+2625 gccaacaactcgaacgcgcgcttcaaatcttgggctacaatcctggctcatttgcagctgttgacccctttttacc
+2700 accaccaccaccgattccgtcatatgtgtttatcctcgtttatgtgacgattctatccggaaacgtaattgggca
+2775 gttggaaaactacaaggaaaatattttattcaacttttaaaaatagttttttttataaaattccccaggtcttact
+2850 tacggaaaagtagttccagacgctcatcctgctgtcgtagtctctatttttttttttctcaaaaaacagtgtttat
+2925 ggtcacaatcacagaaattaaacggattgaatt

Identification of the "Arthropod Initiator" and Capping Site

Examination of VCP^1 and VCP^2 also showed the sequence TCAGT, recently identified as an arthropod initiator, 24 bp downstream of the putative TATA box, indicating that this sequence may also be located at or near the transcriptional start site in this gene (Fig. 3). The capping sites of the *Drosophila* yolk protein genes 1 and 2 show a sequence very similar to this AInr (Hung *et al.*, 1983) (A; Fig. 4). To confirm that the true start site is localized in this region, a primer extension experiment was done, showing that +1 is in fact immediately upstream of this consensus capsite sequence (B; Fig. 4).

Putative Hormone Response Elements

Inspection of the untranscribed regions surrounding the VCP gene revealed the existence of a 16 bp imperfect palindrome repeated 4 times (Fig. 3). All four repeated elements are located in close proximity to one another upstream of the VCP start site and show strong similarity to previously described hormone response elements (Fig. 5). The arms of the palindromes are 8 bp long and extend directly to the center of the element, leaving no spacer.

Homology to Fat Body Enhancers and Protein Binding Sites from Drosophila

Several regulatory elements and protein binding sites have been identified in *Drosophila* that confer fat body specific expression to the genes under their control. These include the 127 bp fat body enhancer and doublesex (DSX) binding site originally identified between the YP1 and YP2 genes (Burtis *et al.*, 1991;Garabedien *et al.*, 1986), several recently identified fat body enhancers also found between YP1 and YP2 (Abrahamsen *et al.*, 1993) and the

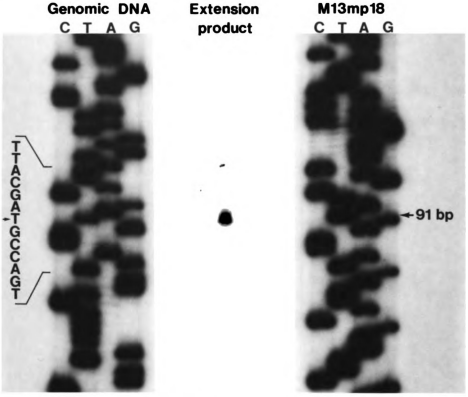
Fig 4. Identification of the transcription start site for VCP as well as a putative Arthropod initiator. (A) Comparison of capsites from two alleles of VCP with those from the *Drosophila* yolk protein genes 1 and 2 and the arthropod initiator consensus sequence. (B) Primer extension analysis showing the start site of transcription. The extension product is run next to two sequencing ladders for markers; one from VCP genomic DNA and the second from an M13 standard. (C) Diagramatic representation of the extension reaction.

A.

Consensus Capsites; Putative Arthropod Initiators

Consensus:		T C A G T
VCP1	:	C C A G T G G
VCP2	:	C C A G T T G
*Dros.*YP2	:	C C A G T T C
*Dros.*YP1	:	C C A A T C G

B.



C.

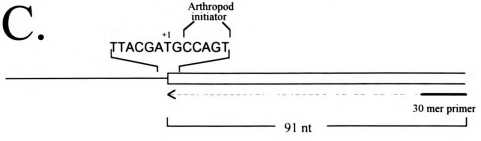
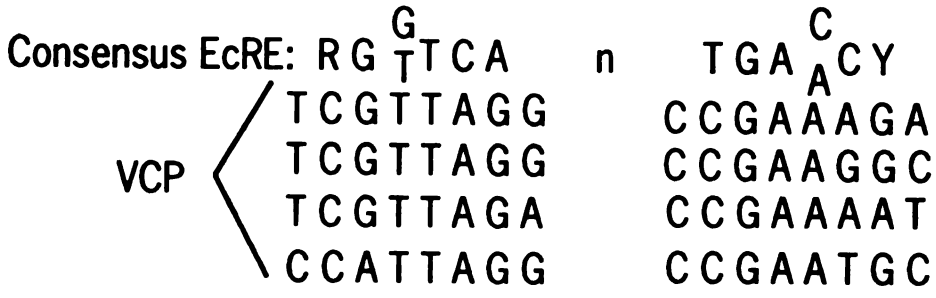


Fig. 5: Response element comparisons. Comparison of the putative hormone receptor binding sites from VCP with the consensus EcRE from *Drosophila* (Cherbas and Cherbas, 1993).


Consensus EcRE: R G G/T T C A n T G A C/A C Y
VCP
T C G T T A G G C C G A A A G A
T C G T T A G G C C G A A G G C
T C G T T A G A C C G A A A A T
C C A T T A G G C C G A A T G C

binding sites for the box B-binding factor-2, CREB (cAMP responsive transcriptional enhancer binding protein) and ATF (activating transcription factor) defined for fat body specific expression of the alcohol dehydrogenase gene (Abel *et al.*, 1992; Hai *et al.*, 1989; Hoeffler *et al.*, 1988). Using GCG/Fasta, these enhancer regions and consensus sequences were compared to the untranscribed portions of the VCP gene.

In both alleles of VCP, a sequence downstream of the coding region was identified as having similarity to the 127 bp fat body enhancer identified by Garabedien *et al.* (1986) (Fig. 6). Several short stretches of homology were also identified when the fat body specific enhancers described by Abrahamsen *et al.* (1993) were used for comparison. These areas tended to be AT rich and less similar than the previous fat body enhancer, indicating they maybe of less significance. Six potential binding sites for BBF-2 and 12 sites matching the CREB/ATF binding sequences were identified scattered throughout the VCP untranscribed DNA (Fig. 3).

Fig. 6: Fat body enhancer comparison. Region of similarity identified through computer comparison of the fat body enhancer from *Drosophila* yolk proteins 1 and 2 (Garabedien *et al.*, 1986) and the untranscribed region of both alleles of VCP. The consensus binding sequence for double sex is identified (DSX) (Burtis *et al.*, 1991).

Possible Fat Body Enhancers

DISCUSSION

Using the cDNA for VCP as a probe to screen an *Aedes aegypti* genomic library, genomic clones representing two alleles for the VCP locus have been identified, sequenced and analyzed. The sequences show several elements indicative of the hormone responsiveness and sex-, stage- and tissue-specific expression of this gene. These include 4 putative hormone receptor binding sites, an arthropod initiator, a region of similarity to a *Drosophila* fat body enhancer, and several potential regulatory protein response elements. The exact start site of transcription was mapped to within two base pairs of this initiator.

Cherbas and Cherbas (1993) theorized that the capsite consensus found in the promoters of many arthropod genes, termed the "arthropod initiator" (AInr), is a functional homologue to the TdT family of mammalian initiators. Experiments done with the control regions of the 20E inducible gene Eip 28/29 and the ecdysone response element (EcRE) of hsp27 in *Drosophila* indicated that hormone response elements and the AInr work in conjunction with one another, and are dependent on each other to induce expression (Cherbas and Cherbas, 1993). The expression of VCP is likely also dependent on the binding of a hormone/receptor complex to several putative binding sites upstream of the capsite. The function of these binding sites may depend on the presence of the AInr in much the same way as the 20E binding sites are dependent on this sequence in Eip28/29 hormone induced expression.

The promoter and enhancer regions of steroid hormone responsive genes defined in other organisms show many conserved motifs. In general, these regulatory regions contain hormone response elements, which tend to be imperfect palindromes 12-16 bp in length, often present in multiple copies,

called hormone response units (for review, see Martinez and Wahli, 1991). The structural features critical to the activity of EcREs in *Drosophila* have now been rigorously examined (Riddihough and Pelham, 1987; Cherbas *et al.*, 1991). Recently, Antoniewski *et al.*, (1993) have shown through mutational analysis using the hsp 27 and Fbp genes that the integrity of each half site of the palindrome as well as the 1 bp spacing is critical for binding of the ecdysone receptor (EcR). They offer a revised version of a previously proposed EcRE consensus sequence and state that variations in this sequence may result in binding sites having different affinities for the receptor.

The repeated imperfect palindrome found upstream of the gene for VCP resembles the described EcRE from *Drosophila* and also strongly resembles a retinoic acid response element (de Thé *et al.*, 1990) (Fig. 5). These repeated palindromes may represent a hormone response unit responsible for the tightly regulated, high levels of expression of this gene and may shed light on the expression of several vitellogenic gene products. The potential binding sites are similar to but not exactly like the consensus EcRE described by Antoniewski *et al.* (1993). In particular, the palindromes found in the VCP upstream region have no spacer, while the responsiveness of the EcRE from *Drosophila* was shown to be dependent upon a 1 bp spacer. These differences, combined with the recent research by Cho *et al.* (1993) showing that EcR in mosquito is expressed only at very low levels in the vitellogenic fat body indicates that this gene may be responding to a cellular signal other than 20E. The Ashburner model explains how 20E in complex with its receptor directly causes the expression of certain early response genes, many of which encode transcriptional activators that act on genes downstream in the pathway (Ashburner *et al.*, 1974). It is possible that the VCP gene of mosquito is regulated by a similar cascade, and that the

transcriptional activator binding upstream of VCP may be a protein further downstream in the 20E response hierarchy.

Due to a lack of germline transformation or vitellogenic cell lines which would enable us to investigate transcriptional regulation by specific promoter and enhancer regions directly, computer aided comparison of sequenced regulatory regions was used to identify possible control elements within the VCP gene. This comparison identified several possible control elements scattered throughout the untranscribed region, including several in close proximity to one another. Computer comparisons should be viewed with caution. Abrahamsen *et al.* (1993) showed how searches for protein binding sites can identify consensus sequences in nearly every sequence compared, including those not regulated by the binding proteins. However, these searches can provide useful information for further analysis including gel retardation and footprinting assays.

The vitellogenic genes of the female mosquito are expressed in a sex-, stage- and tissue-specific manner. The gene for VCP is a small, concise gene representative of this important developmental cycle. It is hoped that the identification of possible hormone response elements as well as potential regulatory protein binding sites will prove useful in determining the molecular mechanisms controlling this expression.

EXPERIMENTAL PROCEDURES

Animals

Mosquitoes (UGAL strain of *Aedes aegypti*) were reared at 27° C and 85% relative humidity. Larvae were fed on a standard diet (Lea, 1964). Adults were fed 10% sucrose by wick continuously. Three to five days after eclosion adult females were allowed to feed on white rats to initiate vitellogenesis.

Genomic Library Construction and Screening

Genomic DNA was extracted from approximately 0.5 g of mosquito whole bodies using a standard DNA extraction procedure (Ausubel, *et al*, 1989). The tissue was frozen in liquid N and ground to a fine powder using a mortar and pestle. It was then suspended in 5 ml of digestion buffer (100 mM NaCl, 10 mM Tris-Cl pH 8.0, 25 mM EDTA, 0.5 % SDS, 0.1 mg/ml fresh proteinase K) and incubated at 50° C overnight. The solution was then extracted gently with 1 vol. phenol/chloroform/isoamyl alcohol (25:24:1) and centrifuged at 10k rpm (Beckman J-20 rotor) for 10 min. The aqueous phase was saved and the extraction repeated until no interphase was present. The DNA was then precipitated with the addition of 0.5 vol. 7.5 M NH_4-acetate and 2 vol. ethanol. This was spun down at 5k rpm for 2 min., rinsed with 70% ethanol and allowed to air dry. The resulting DNA pellet was gently resuspended in 1 ml of TE (10mM Tris-Cl, 1mM EDTA, pH 8.0). This DNA was sent to the Stratagene Corp. and a λFIX II genomic library constructed. This library was screened according to the manufacturers instructions using an 840 bp [α-^{32}P] random primer labeled VCP cDNA fragment (Cho, *et al.*, 1991). Resulting positive clones were verified through Southern analysis, restriction mapped and selected DNA fragments subcloned into pUC 118 and pUC 119 for sequencing.

DNA Sequencing

Subcloned genomic fragments were sequenced in both directions using the Sequenase (USB Corp.) sequencing kit and the manufacturer's protocols. Unidirectional deletions were generated using the method of Dale (1985) with M13 derived single stranded DNA. The resulting sequences were analyzed using Dotplot, Fasta, and Compare programs (University of Wisconsin Genetics Computer Group software) and through visual inspection.

Genomic Blots

Genomic DNA was isolated as previously described, digested to completion with restriction enzymes, and fractionated using 0.8% agarose gel electrophoresis. The DNA was then transferred to nitrocellulose membrane using the method of Southern (1975) and probed under high stringency conditions (5X SSC, 50 mM phosphate buffer, 50 % formamide, 5X Denhardt's, 0.01 % SDS, 200 μg/μl salmon sperm DNA). The resulting blots were washed in 2X SSC, 0.1% SDS at room temp twice, in the same buffer at 60° C once, and in 0.2X SSC, 0.1% SDS at 60° C once. The blots were allowed to air dry and were used to expose film at -70° C using an intensifying screen. Genomic DNA from individual mosquitoes was isolated by homogenizing single mosquitoes in 400 μl digestion buffer, followed by the previously mentioned protocol scaled accordingly.

Primer Extension

The primer extension experiment was done following the methods described by Sambrook *et al.* (1989) and Ausubel *et al.* (1989). 150 ng of a 30 bp synthetic primer derived from the VCP cDNA sequence was end labeled with

[γ-^{32}P] ATP by T4 polynucleotide kinase. This was hybridized to 10 µg of total RNA extracted from vitellogenic fat bodies isolated 20 - 24 hr after a blood meal. AMV reverse transcriptase was then used to reverse transcribe labeled DNA from the RNA template. This was extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with 3 vol. ethanol in a dry ice/ethanol bath. The DNA was spun down in a microcentrifuge at 4^{o} C, rinsed with 70% ethanol and allowed to air dry. The dry pellet was taken up in 4 µl H_2O and 6 µl formamide loading buffer (80% formamide, 10mM EDTA pH 8.0, 1 mg/ml xylene cyanol FF, 1mg/ml bromophenol blue). 5 µl of this was run on a 6% polyacrylamide denaturing sequencing gel. A sequencing reaction done with single stranded DNA derived from a VCP genomic clone using the same synthetic primer was used as a marker. An additional marker was derived by sequencing M13mp18 single stranded DNA. The resulting autoradiograph shows the start site of VCP mRNA transcription.

Computer Analysis

The VCP untranscribed regions were compared with the untranscribed regions of the *Drosophila* yolk protein genes using the Fasta program (University of Wisconsin Genetics Computer Group software). To identify possible DSX binding sites, 7 bp homology to either the suggested consensus sequence (CTACAAAGT) or to the sequence CTACAATGT which was shown to exhibit the strongest affinity for the DSX protein was searched for. Similar searches were done for the consensus binding sites of CREB/ATF (TGACGTCA) and BBF-2 (TA/CNACGTANT/GC).

ACKNOWLEDGEMENTS

We would like to thank Dr. Wen-long Cho for his advice throughout the course of this study. We also thank Dr.Wen-long Cho, Jeng-Shong Chen and Neal Dittmer for their critical reading of the manuscript. This research was supported by grant AI-24716 from the National Institutes of Health to ASR.

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:** 466-488.

Abrahamsen, N., Martinez, A., Kjaer, T., Sondergaard, L. and Bownes, M. (1993) Cis-regulatory sequences leading to female-specific expression of yolk protein genes 1 and 2 in the fat body of *Drosophila melangaster*. *Mol. Gen. Genet.* **237:** 41-48.

Antoniewski C., Laval M. and Lepesant J. (1993) Structural features critical to the activity of an ecdysone receptor. *Insect Biochem. Mol. Biol.* **23:** 105-114.

Ashburner M., Chihara C., Meltzer P. and Richards, G. (1974) Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp.* **38:** 645-654.

Ausubel F. M., Brent R., Kingston R. E., Moore D. D., Seidman J. G., Smith, J. A. and Struhl, K. (1989) *Current Protocols in Molecular Biology*, 2.2.1-2.2.2. Greene Publishing Associates and Wiley-Interscience, New York.

Burtis, K. C., Coschigano, K. T., Baker, B. S. and Wensink, P. C. (1991) The doublesex proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. *EMBO* **10:**2577-2582.

Cherbas L. and Cherbas P. (1993) The arthropod initiator: the capsite consensus plays an important role in transcription. *Insect Biochem. Mol. Biol.* **23:** 81-90.

Cherbas, L., Lee, K. and Cherbas, P. (1991) Identification of ecdysone response elements by analysis of the Eip28/29 gene. *Genes Dev.* **5:** 120-131.

Cho W., Deitsch K. W. and Raikhel A. S. (1991) An extraovarian protein accumulated in mosquito oocytes is a carboxypeptidase activated in embryos. *Proc. Natl. Acad. Sci. USA.* **88:** 10821-10824.

Cho W., Kapitskaya M. Z. and Raikhel A. S. (1993) Mosquito ecdysteroid receptor: analysis of the cDNA and transcription during vitellogenesis. *Proc. Natl. Acad. Sci. USA.* In Press.

Dale R. M. K., McClure B. A. and Houchins J. P. (1985) A rapid single-stranded cloning strategy for producing sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18s rDNA. *Plasmid* **13:** 31-40.

de Thé H., Vivanco-Ruiz M.M., Tiollais P., Stunnenberg H. and Dejean A. (1990) Identification of a retinoic acid responsive element in the retinoic acid receptor β gene. *Nature* **343:** 177-180.

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

Garabedien M. J., Shepherd B. M. and Wensink P.C. (1986) A tissue-specific transcription enhancer from the *Drosophila* yolk protein 1 gene. *Cell* **45:** 859-867.

Gemmill R. M., Hamblin M., Glaser R. L., Racioppi J. V., Marx J. L., White B. N., Calvo J. M., Wolfner M. F. and Hagadorn H. H. (1986) Isolation of mosquito vitellogenin genes and induction of expression by 20-hydroxyecdysone. *Insect Biochem.* **16:** 761-774.

Hagedorn, H. H. (1983) The role of ecdysteroids in the adult insect, in Downer, R. G. H. and Laufer, H. (eds): *Endochrinology of Insects,* Alan R. Liss, Inc., New York, pp. 271-304.

Hagedorn, H. H. (1985) The role of ecdysteroids in reproduction, in Kerkut, G. A., Gilbert, K. I. (eds): *Comprehensive Insect Physiology, Biochemistry, and Pharmacology,* Volume 8, Pergamon Press, Oxford, pp. 205-261.

Hagenbuchle O., Santer M., Steitz J. A. and Mans R. (1978) conservation of the primary structure at the 3' end of 18s rRNA from eukaryotic cells. *Cell* **13:** 551-563.

Hai, T., Liu, F., Coukos, W. and Green, M. (1989) Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes and Dev.* **3:** 2083-2090.

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

Hoeffler, J. P., Meyer, T. E., Yun, Y., Jameson, J. L. and Habener, J. F. (1988) Cyclic AMP-responsive DNA-binding proteins: structure based on cloned placental cDNA. *Science* **242:** 1430-1432.

Hung M. and Wensink P.C. (1983) Sequence and structure conservation in yolk proteins and their genes. *J. Mol. Biol.* **164:** 481-492.

Lea, A. (1964) Studies on the dietary and endocrine regulation of autogenous *Aedes taeniorhynchus*. *J. Med. Ent.* **1:** 40-44.

Martinez E. and Wahli W. (1991) Characterization of hormone response elements. *Nuclear hormone receptors: molecular mechanisms, cellular functions, clinical abnormalities*, 125-153. Academic Press, New York.

Racioppi J.V., Gemmill R. M., Kogan P. H., Calvo J. M., and Hagedorn H. H. (1986) Expression and regulation of vitellogenic messenger RNA in the mosquito *Aedes aegypti*. *Insect Biochem.* **16:** 255-262.

Raikhel A. (1992) Vitellogenesis in mosquitos. *Advances in disease vector research*. **9**, 1-39. Springer-Verlag, New York.

Raikhel A. S. and Lea A. O. (1983) Previtellogenic development and vitellogenin synthesis in the fat body of a mosquito: an ultrastructural and immunocytochemical study. *Tissue Cell* **15:** 281-300.

Riddihough, G. and Pelham, H. R. B. (1987) An ecdysone response element in the *Drosophila* hsp27 promoter. *EMBO* **6:** 3729-3734.

Sambrook J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Severson, D. W., Mori, A., Zhang, Y. and Christensen, B. M. (1993) Linkage map for *Aedes aegypti* using restriction fragment length polymorphisms. *J. Heredity* **84:** 241-247.

Southern E. M. (1975) Detection of specific sequences among DNA fragments seperated by gel electrophoresis. *J. Mol. Biol.* **98:** 503-517.

CHAPTER 4*

INDIRECT CONTROL OF YOLK PROTEIN GENES BY 20-HYDROXYECDYSONE IN THE FAT BODY OF THE MOSQUITO, *AEDES AEGYPTI*

*Submitted for publication to Insect Biochemistry and Molecular Biology

Reference: Deitsch, K.W., Chen, J.-S., and Raikhel, A.S. (1994) Insect Biochem. Mol. Biol., Submitted.

ABSTRACT

In response to a blood meal, the fat body of the female mosquito, *Aedes aegypti,* begins massive production of several yolk proteins which are subsequently stored in the developing oocytes. Although 20-hydroxyecdysone (20E) is important for initiation and maintenance of expression of the yolk protein genes, the exact nature of 20E action has not been clearly defined. A primary question is whether this hormone directly stimulates yolk protein mRNA transcription or if it acts indirectly through a hormone response cascade. We have demonstrated that 10^{-4} M cycloheximide (Chx) reversibly inhibits >98% of protein synthesis in *in vitro* fat body primary culture. 10^{-5} M 20E stimulates high levels of yolk protein gene transcription, but initiation of this transcription is eliminated by Chx. Thus, our results indicate that protein synthesis is required in response to 20E before transcription of the yolk protein genes is initiated. We therefore conclude that the action of 20E is indirect.

INTRODUCTION

In the anautogenous mosquito, *Aedes aegypti*, vitellogenesis is initiated by a blood meal. The fat body transcribes large quantities of yolk protein mRNA and massive amounts of these proteins are produced and transported to the developing oocytes. The titer of 20-hydroxyecdysone (20E) is tightly correlated with this cycle, rising with yolk protein production and falling when vitellogenesis ceases (Hagedorn, 1983, 1985; Raikhel, 1992). Initial experiments on the role of 20E in yolk protein regulation have been inconclusive. *In vitro*, physiological doses of 20E can stimulate synthesis of the major yolk protein, vitellogenin (Vg) in previtellogenic fat bodies (Fuchs and Kang, 1981; Hagedorn, 1985). However, large doses of 20E are needed to generate similar levels of Vg production *in vivo* (Fuchs and Kang, 1981; Lea, 1982). In addition, induction of Vg synthesis is detectable at both the protein (Raikhel and Lea, 1983) and the mRNA levels (Racioppi *et al*, 1986) only 1 hr after a blood meal, before 20E titers begin to rise. These results show how complex the regulation of vitellogenic gene expression is, and that 20E plays a significant, although unclear role in this regulation.

A second yolk protein has been described called vitellogenic carboxypeptidase (VCP) (Hays and Raikhel, 1990; Cho *et al.*, 1991). *In vivo* and *in vitro* experiments indicate that this protein is synthesized in the same tissue-, stage- and sex-specific manner as Vg. Recently, cDNAs encoding both Vg and VCP have been cloned and sequenced (Cho *et al.*, 1991; Chen *et al.*, 1994), providing molecular probes for investigating the role of 20E in the control of these two genes.

Ashburner *et al.* (1974) proposed a model for the mode of action of 20E in *Drosophila*. They hypothesized that 20E, complexed with its receptor, directly

regulates two classes of genes: a small class of early genes that are transcribed when bound by this complex and a large class of late genes that are repressed. The early genes encode transcription factors that serve two functions. They repress early gene transcription and they induce late gene transcription. Recent research at the molecular level has revealed many genes whose regulation is consistent with this model (Segraves and Hogness, 1990; Burtis *et al.*, 1990; Koelle *et al.*, 1991). Thus, 20E is thought to operate through a cascade or hierarchy of genes initially induced by the 20E/receptor complex and ultimately leading to expression of numerous downstream target genes. It is important to determine whether 20E and its receptor act directly to initiate yolk protein gene transcription in mosquito or if they act through a hierarchy of upstream transcriptional activators in a response cascade. The results presented here support the latter hypothesis, indicating that the Ashburner model also likely applies to the genes involved in mosquito vitellogenesis.

MATERIALS AND METHODS

Insects

Mosquitoes were reared at 27° C and 85% relative humidity in a 16/8 (light/dark) cycle. Larvae were fed on a standard diet (Lea, 1964) and adults were fed 10% sucrose continuously by wick. Three to 5 days after eclosion, adult females were allowed to feed on white rats to initiate vitellogenesis. All dissections were performed in *Aedes* physiological saline (Hagedorn *et al.*, 1977) at room temperature. For culture, isolated abdominal cuticle with adhering fat bodies were used immediately after dissection. Extractions for protein or total RNA isolation were either done immediately after culture or after quick freezing in liquid Nitrogen and storage at -80° C.

Cloning of Actin cDNA

A partial genomic clone encoding the *Drosophila* actin 5C gene was kindly provided by Dr. Eric Fryberg (Fryberg *et al.*, 1980). This DNA fragment was labeled with [^{32}P] using the random primer method and used to screen a λ gt11 *Aedes aegypti* cDNA library made from vitellogenic fat body mRNA (Cho *et al.*, 1991) as outlined by Sambrook *et al.* (1989). Several positive clones were obtained and subcloned into pUC119 for sequencing. Northern blots of mosquito total RNA probed with the putative mosquito actin cDNA or the *Drosophila* actin gene showed hybridization to the same band. In addition, the deduced amino acid sequence showed > 85% identity with the *Drosophila* actin gene, confirming the identity of the mosquito actin cDNA.

In vitro *Fat Body Culture*

Abdomens from blood-fed or three day old sugar-fed females were isolated and placed in a tissue culture system as previously described (Koller *et al.*, 1989; Dhadialla and Raikhel, 1990). Protein was labeled by including [^{35}S] methionine in the culture media. 20E (Calbiochem) was dissolved in ethanol and stored at -80^{o} C until immediately before use. Chx (Calbiochem) was dissolved in H_2O immediately before use. Radiochemicals were purchased from Amersham and New England Nuclear. All tissue culturing was performed in a hood using sterile instruments to reduce contamination. Cultures were incubated at 27^{o} C for the duration of the experiments. Three repetitions of 9 fat bodies were cultured for each treatment.

Protein Analysis

Cellular proteins were isolated by homogenizing cultured fat bodies in B1 buffer (50mM Tris pH 8.4, 50mM NaCl, 0.2% Triton-X-100, 0.02% sodium azide). Vg was immunoprecipitated using monoclonal antibodies specific to both the large and small Vg subunits. Vg protein levels were estimated by measuring the level of [^{35}S] in the precipitate using a scintillation counter (LKB Wallace) (Raikhel and Bose, 1988; Dhadialla and Raikhel, 1990). Total protein levels were calculated by trichloroacetic acid precipitation and subsequent scintillation counting of incorporated [^{35}S]. All experiments were done in triplicate to ensure reproducibility.

mRNA Analysis

Cultured fat bodies were homogenized and total RNA isolated as previously described (Raikhel and Bose, 1988; Cho *et al.*, 1991). RNA concentration was measured using a spectrophotometer. Equal amounts of

RNA were blotted onto 0.45μm nitrocellulose membrane (MSI) using a Bio-Rad slot blot apparatus and the protocol described by Sambrook *et al.* (1989). All RNA samples were divided into three equal portions and blotted separately. These replicate blots were probed with random primer labeled cDNA clones for Vg, VCP or actin as described (Cho *et al.*, 1991) and used to expose film at -70^{o} C using an intensifying screen. After autoradiography, the blots were cut into pieces and each individual sample placed in a scintillation counter for quantification. All experiments were done in triplicate to ensure reproducibility.

To ensure that the above procedure accurately quantifies different levels of probe hybridization, the following pilot experiment was done. Serial dilutions of the cDNAs for VCP, Vg and actin were slot blotted, baked, then hybridized to radiolabeled cDNA probes. The blots were autoradiographed, cut into pieces and each sample placed in a scintillation counter for quantification. The results showed that this method accurately determined the level of cDNA blotted onto the membrane to within + or - 7 %.

RESULTS

Initiation of Transcription and Synthesis of Vitellogenin and VCP by Cultured Fat Bodies in Response to 20E

The effect of 20E on the initiation of yolk protein production in the fat body of female mosquitoes was first investigated. The level of Vg production was measured in cultured previtellogenic fat bodies in response to increasing amounts of 20E added to the culture media. Fat bodies from sugar-fed females were incubated for 6 hrs in culture media containing [^{35}S] methionine and different concentrations of 20E with a change of media every 2 hrs. Fat bodies from sugar-fed females cultured in media containing no 20E were used to establish a background level of expression while fat bodies from blood-fed females 6 hrs after feeding were used to show a 100% response. Vg and total protein production reached maximum levels at 10^{-5} M 20E (Fig. 1). Vg levels reached ~60% of that obtained from blood-fed controls while total protein levels reached ~87% of control. Concentrations greater than 10^{-5} M did not give any further increases in Vg or total protein production (not shown).

The effect of 20E on the initiation of yolk protein gene transcription in cultured previtellogenic fat bodies was then examined. This was done by measuring the amount of Vg and VCP mRNA present in fat bodies from sugar-fed females incubated for 6 hrs in culture media (changed every 2 hrs) containing different concentrations of 20E. Vg and VCP mRNA produced by fat bodies from blood-fed females 6 hrs post blood meal was measured to determine *in vivo* levels of mRNA. The amount of actin mRNA produced also was measured to show a basal level of transcription of a non-vitellogenic gene. Both Vg and VCP mRNA are transcribed in response to 20E, reaching maximum transcription rates at 10^{-5} M (Fig. 2), concentrations similar to those inducing

Fig. 1. Effects of 20E concentrations on Vg and total protein production in cultured fat bodies. Previtellogenic fat bodies were incubated for 6 hrs in culture media containing 0 to 10^{-5}M 20E and [^{35}S] methionine as described in Materials and Methods. Fat bodies taken from blood-fed females were used as a positive control. After culturing, proteins were extracted from the tissue and Vg immunoprecipitated using monoclonal antibodies specific to this protein. Total protein was precipitated with TCA. The activity was determined using a scintillation counter. The vertical line at the top of each bar indicates the standard error of three repetitions.

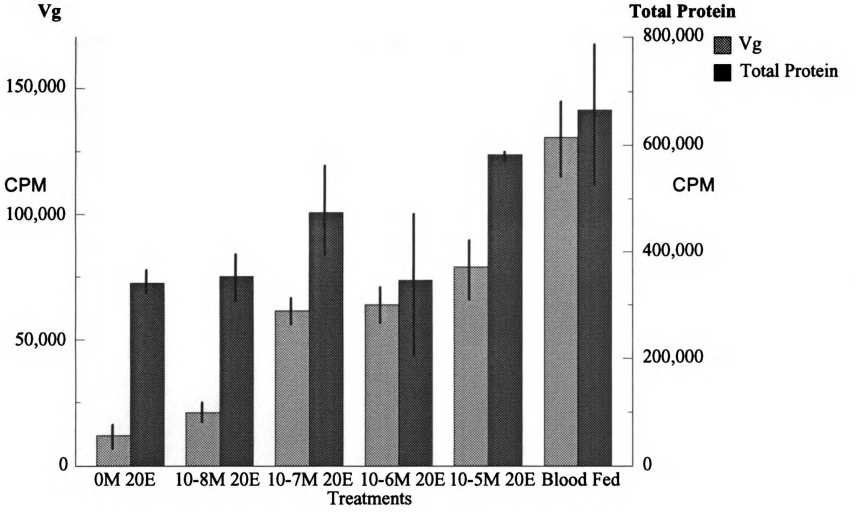
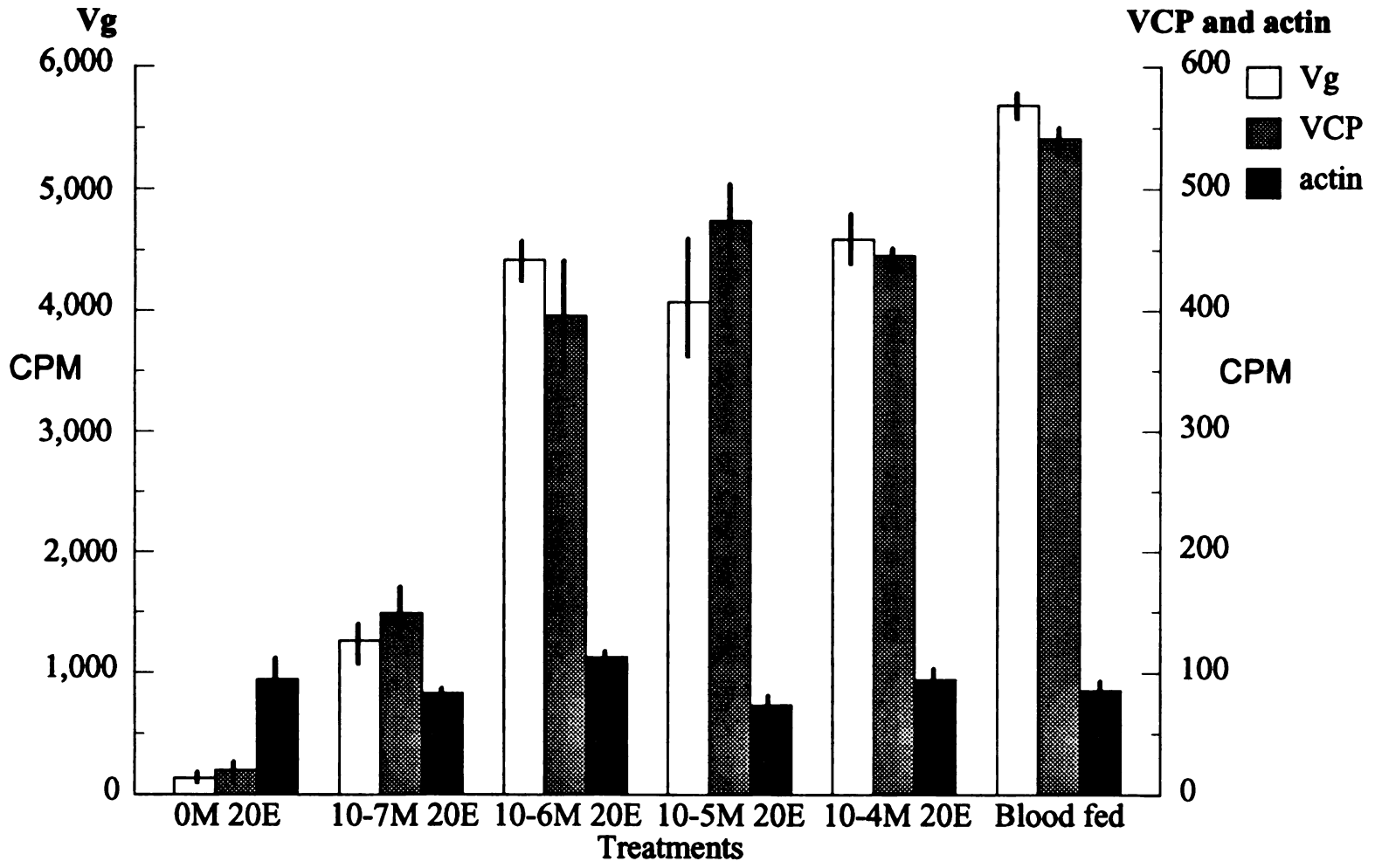

Vg
Total Protein
Vg
Total Protein
150,000
100,000
50,000
0
CPM
800,000
600,000
400,000
200,000
0
CPM
0M 20E
10-8M 20E
10-7M 20E
10-6M 20E
10-5M 20E
Blood Fed
Treatments

Fig. 2. Effect of 20E concentration on Vg, VCP and actin mRNA production in cultured fat bodies. Cultured previtellogenic fat bodies were incubated for 6 hrs with 0 to 10^{-4} M 20E. Total RNA was extracted and levels of Vg, VCP and actin mRNA measured by slot blot and hybridization to the appropriate radiolabeled cDNA probes. Fat bodies from females 6 hrs after a blood meal were used to show an *in vivo* response. Hybridization was quantified using a scintillation counter. Vertical lines indicate the standard error of the mean of three repetitions.


Vg
VCP and actin
6,000
5,000
4,000
3,000
2,000
1,000
0
CPM
600
500
400
300
200
100
0
CPM
Vg
VCP
actin
0M 20E
10-7M 20E
10-6M 20E
10-5M 20E
10-4M 20E
Blood fed
Treatments

maximum production of Vg protein. At this concentration, Vg and VCP mRNA levels reached ~81% and 82% of that obtained from blood-fed controls, respectively. Background levels of Vg mRNA were always found to be higher than VCP levels in previtellogenic fat bodies and controls untreated with 20E, indicating that a low level of Vg transcription is present in non-induced fat bodies. As expected, actin levels remained relatively constant.

Effect of Cycloheximide on Cultured Fat Bodies

The minimum concentration of Chx capable of inhibiting the majority of protein synthesis was determined using a dose response curve of protein production in response to increasing concentrations of Chx. Fat bodies from females 18-20 hrs after a blood meal (near the peak of yolk protein production) were treated with different doses of Chx for 1 hr, then incubated in media containing Chx and [^{35}S] methionine for 1 hr to label newly synthesized proteins. At 10^{-4} M Chx, >98% of protein synthesis is inhibited (Fig. 3), providing a treatment concentration for subsequent experiments.

The effect of Chx on yolk protein mRNA transcription was now tested by culturing sugar-fed female fat bodies in media containing 10^{-4}M Chx for 1 hr, then in media containing 10^{-5}M 20E and 10^{-4}M Chx for 6 hrs. The media was changed every 2 hrs. 10^{-4} M Chx virtually eliminated induction Vg and VCP gene transcription in response to 20E (Fig. 4). At the same time, Chx appears to have little effect on the transcription rate of the actin gene. Background Vg transcription was again higher than that of VCP. To ensure that this inhibition was not due to a decay of cellular proteins needed for the 20E response over the 6 hr incubation, RNA was extracted from fat bodies cultured as above after only 2 hrs of 20E stimulation. This experiment also showed elimination of 20E induction of VCP and Vg transcription (not shown). Thus, these experiments

Fig. 3. Inhibition of Vg and total protein synthesis by cycloheximide. Vitellogenic fat bodies 18 hrs post blood meal were cultured for 1 hr in 0 to 10^{-3}M Chx followed by 1 hr with Chx and [^{35}S] methionine. Levels of newly synthesized Vg and total protein were measured as described in the text. Values are expressed as percentages of that produced by fat bodies cultured without Chx (0 M Chx, control).

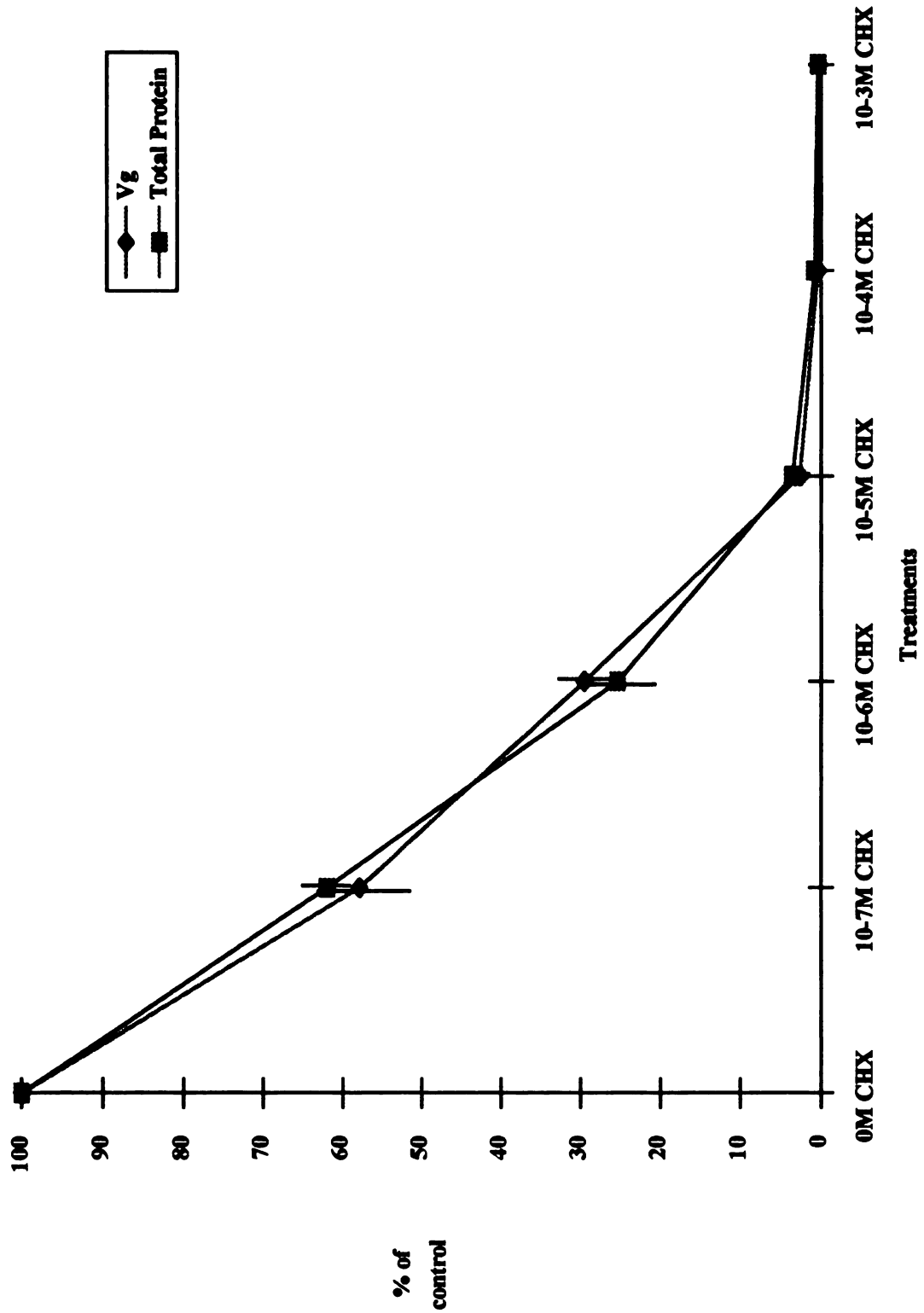
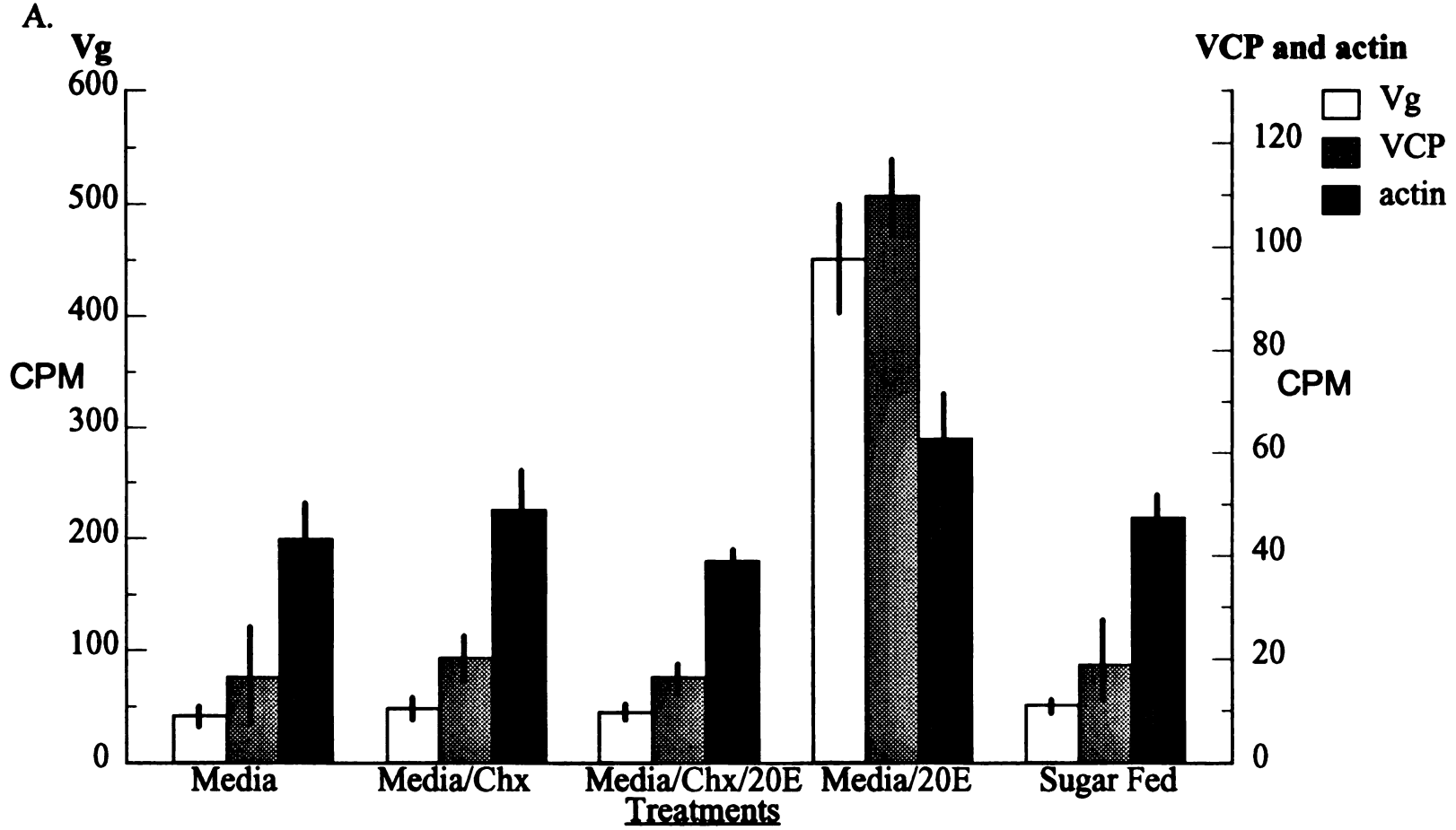
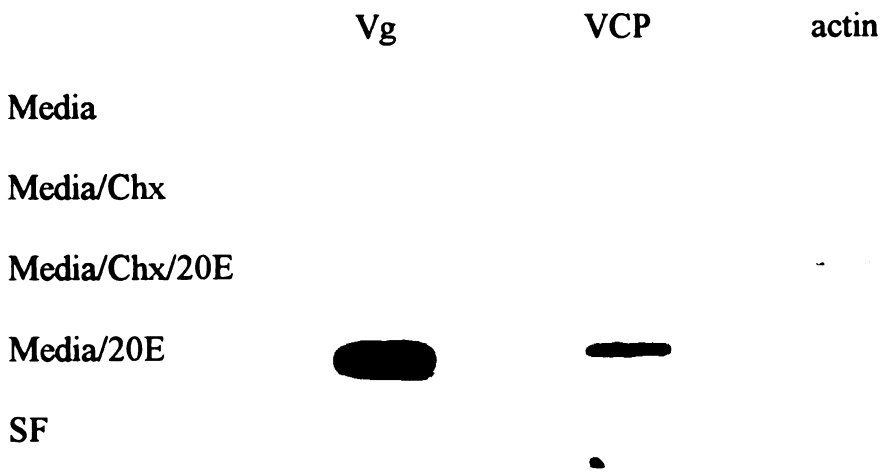

Vg
Total Protein
100
90
80
70
60
50
40
30
20
10
0
% of control
0M CHX
10-7M CHX
10-6M CHX
10-5M CHX
10-4M CHX
10-3M CHX
Treatments

Fig. 4. Effect of Chx on Vg, VCP and actin mRNA production. Previtellogenic fat bodies were preincubated with or without Chx for 1hr, then cultured for 6 hrs in media containing different combination of Chx and 20E. Fat bodies taken from previtellogenic sugar-fed females were used as a control. After culturing, total RNA was extracted and levels of Vg, VCP and actin mRNA measured by slot blot and hybridization to the appropriate radiolabeled cDNA probes. A. Quantitation of hybridization using a scintillation counter. Vertical lines show the standard error of the mean of three repetitions. B. Autoradiograph of slot blots probed with cDNAs for Vg, VCP and actin.


A.
Vg
600
500
400
CPM
300
200
100
0
VCP and actin
Vg
VCP
actin
120
100
80
CPM
60
40
20
0
Media
Media/Chx
Media/Chx/20E
Media/20E
Sugar Fed
Treatments

B.



indicate that inhibition of protein synthesis abolishes the ability of 20E to induce yolk protein transcription.

Recovery of Yolk Protein Production and Transcription After Cycloheximide Treatment

To ensure that the effect of Chx on protein synthesis is not simply the result of cell death, the following recovery experiments were done. First, fat bodies from females 18 hrs after a blood meal were incubated in media containing Chx for 1 hr, rinsed twice with culture media, then incubated for 1 hr in standard culture media, followed by the labeling of newly synthesized proteins for 1 hr in culture media containing [^{35}S] methionine. Both total protein synthesis and Vg synthesis were recovered by washing the Chx from the tissue (Fig. 5). At concentrations of both 10^{-4} and 10^{-5}M Chx, >90% of protein synthesis is inhibited, yet significant synthetic capability can be recovered within 2 hr of removing the Chx from the culture media, indicating that these concentrations of Chx induce a reversible inhibition of protein synthesis. The cultures treated with 10^{-5} M Chx recovered more rapidly or more completely (~60% of control) than those treated with 10^{-4} M (~40% of control). This is likely because Chx is more readily removed from cultures where the initial concentration is lower.

Initiation of transcription of Vg and VCP mRNA in previtellogenic fat bodies can also be recovered by removal of Chx from the culture media. Previtellogenic fat bodies were incubated in culture media with 10^{-4}M Chx for 1 hr, rinsed twice in standard culture media then incubated in culture media with 10^{-5}M 20E for 6 hrs, changing the media every 2 hrs. The cultures rinsed of Chx showed significant levels of Vg and VCP mRNA, indicating the tissue can

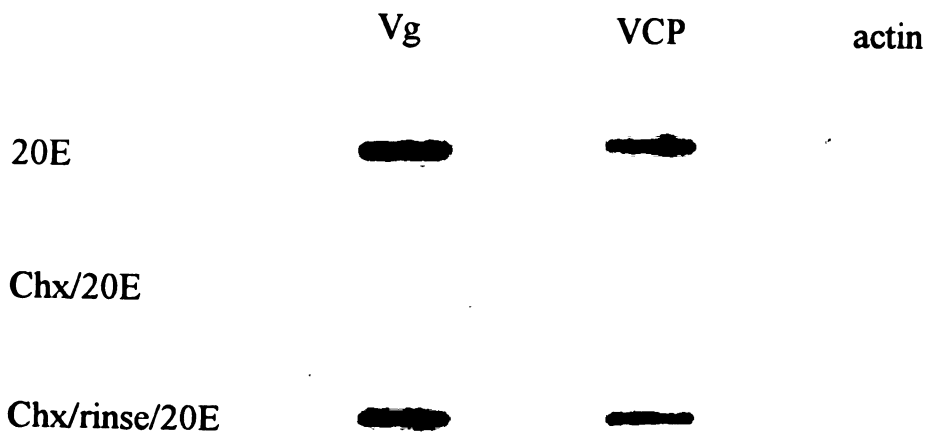
Fig. 5. Recovery of vitellogenesis after removal of Chx. Vg and total protein produced by cultured vitellogenic fat bodies 18 hrs post blood meal under different conditions: 1, culture media plus 10^{-4}M Chx for 3 hrs; 2, culture media plus 10^{-4} M Chx for 1 hr, 0M Chx for 2 hrs; 3, culture media plus 10^{-5}M Chx for 3 hrs; 4, culture media plus 10^{-5}M Chx 1 hr, 0M Chx for 2 hrs; 5, culture media with no Chx for 3 hrs. In all treatments, protein production was measured for the third hour of incubation. Values are expressed as a percentage of levels produced by tissue cultured without Chx (treatment 5, control). Vertical lines indicate the standard error of the mean of three repetitions.

Vg
Total Protein
% of Control
120
100
80
60
40
20
0
1
2
3
4
5
Treatments

initiate transcription of these genes in response to 20E after Chx has been removed (not shown). Actin mRNA levels again appear relatively constant.

Last, recovery of the ability of cultured previtellogenic fat bodies to respond to 20E after 6 hrs of Chx treatment was considered. Fat bodies from previtellogenic females were incubated in media containing Chx for 6 hrs, rinsed three times with culture media containing 10^{-5}M 20E over a 1 hr period, then incubated for 1 hr in culture media containing [^{35}S] methionine/10^{-5}M 20E, followed by a 1 hr incubation in standard culture media/10^{-5}M 20E to obtain labeled secreted proteins. The culture media was TCA precipitated to measure protein production and the tissue homogenized for RNA analysis. Both protein synthesis and yolk protein mRNA transcription were recovered by washing the Chx from the tissue. Within 3 hrs of removal of the Chx from the media, protein production had returned to 70% of the untreated control. Induction of both Vg and VCP mRNA was detected in cultures rinsed of Chx (Fig. 6), indicating the cultured tissue is still capable of responding to 20E even after the 6 hr Chx treatment. Actin levels again appear relatively constant.

Fig. 6. Recovery of Vg and VCP mRNA production after removal of Chx. Autoradiograph of slot blots probed with cDNAs for Vg, VCP and actin. Previtellogenic fat bodies were incubated in culture media containing 10^{-4}M Chx for 1 hr, rinsed of Chx, then incubated in culture media free of Chx and containing 10^{-5}M 20E for 6 hrs. Total RNA was extracted and levels of Vg, VCP and actin mRNA measured by slot blot and hybridization to the appropriate cDNA probes. Top lane: RNA from cultured fat bodies treated with 10^{-5}M 20E only; center lane: RNA from cultured fat bodies treated with 10^{-4}M Chx and then in 10^{-5}M 20E in the presence of Chx; last lane: RNA from cultured fat bodies treated with 10^{-4} M Chx, rinsed in culture media without Chx and then treated with 10^{-5}M 20E only.


Vg
VCP
actin
20E
Chx/20E
Chx/rinse/20E

DISCUSSION

The effect of hormones on gene expression is of fundamental interest to the field of developmental biology. A great deal of research has focused on the mode of action of numerous hormones as they illicit transcription of their target genes or as they initiate a hormone response cascade. Within such a cascade, the hormone acts on particular genes both directly and indirectly. Direct effects on gene expression can be defined as those which are rapid and can occur even though synthesis of new proteins has been inhibited (Yamamoto and Alberts, 1976). An example of this is the glucocorticoid induction of MMTV RNA (Ringold *et al.*, 1977). Conversely, an indirect response is one in which the initial effect of the hormone is thought to be mediated or amplified by production of other regulatory proteins. Such control would consequently be sensitive to inhibition of protein synthesis. Examples of indirect responses include the induction of the late gene products in *Drosophila* salivary glands by 20E (Ashburner, 1973) and the estrogen induction of ovalbumin mRNA in the chick oviduct (McKnight and Palmiter, 1979).

In *Xenopus leavis,* the liver produces high levels of Vg in response to estrogen. Experiments using Chx showed that the effect of estrogen on Vg mRNA production is direct because inhibition of liver protein synthesis does not prevent activation of Vg gene transcription (Hayward *et al.*, 1982). The action of 20E on the yolk protein genes in flies appears to be quite different. In *Drosophila*, Bownes *et al.* (1987) used Chx to determine the nature of the effect of 20E on yolk protein transcription. Their results indicated that protein synthesis is required for 20E to initiate and maintain yolk protein transcription. However, these studies were done *in vivo* and secondary effects of Chx on the insect cannot be ruled out. By using a well defined *in vitro* system, the relevant

tissue can be isolated and a more controlled experiment done. Our studies indicate that in the mosquito 20E is acting on the genes for Vg and VCP indirectly. 20E induces transcription of these genes in a dose dependent manner (Fig. 2), yet protein synthesis inhibition completely eliminates this induction (Fig. 4). It is unlikely that the effect of Chx on Vg and VCP transcription is due to decay of the 20E receptor over the course of the experiment, rather than the interuption of the 20E response pathway because Chx treatment leads to an immediate elimination of Vg and VCP transcription in response to 20E rather than a slow decline, and this response is rapidly recovered upon removal of Chx. In addtion, similar experiments with cultured *Drosophila* salivary glands showed that proteins required for direct responses to 20E are present through 8 hrs of Chx treatment, indicating that the 20E receptor persists throughout such an experiment (Segraves and Hogness, 1990). Thus, a 20E response cascade is likely being induced by rising titers of the hormone with the genes for Vg and VCP downstream in the hierarchy.

This conclusion is consistent with the recent finding that upstream of the VCP gene exists a putative response element that shows limited similarity to the consensus ecdysone receptor binding site (Deitsch and Raikhel, 1993). This palindromic sequence is found tandemly repeated four times and has no spacing between half sites. This is in contrast to the one base pair spacing shown to be necessary for binding of the ecdysone receptor in *Drosophila* (Antoniewski *et al.*, 1993). The cDNA for the ecdysone receptor from mosquito has recently been cloned and shows very high similarity to that of *Drosophila*, particularly in the P-box domain (100% amino acid identity), indicating that its binding specificity will also likely depend on a one base pair spacer (Kapitskaya *et al.*, 1993). Therefore, it is likely that the putative response element upstream of VCP is

bound by a transcription factor other than 20E. Elucidation of this activating protein awaits further research.

The conclusion that the control of the yolk protein genes by 20E in mosquito is indirect provides the necessary groundwork for future research in this area. By eventually uncovering the entire hierarchy of target genes and activator proteins, a more complete understanding of the effect of 20E on the development of mosquitoes and insects in general can be acquired.

ACKNOWLEDGEMENTS

We thank Dr. Tom Sappington and Dr Suzanne Thiem for their critical reading of the manuscript. This research was supported by grant AI-24716 from the National Institutes of Health to A.S.R.

REFERENCES

Antoniewski C., Laval M. and Lepesant J. (1993) Structural features critical to the activity of an ecdysone receptor. *Insect Biochem. Mol. Biol.* **23,** 105-114.

Ashburner M. (1973) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster. Dev. Biol.* **35,** 47-61.

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.

Bownes M., Scott A. and Blair M. (1987) The use of an inhibitor of protein synthesis to investigate the roles of ecdysteroids and sex-determination genes on the expression of the genes encoding the *Drosophila* yolk proteins. *Development* **101,**931-941.

Burtis K., Thummel C., Jones C., Karmin F. and Hogness D. (1990) The *Drosophila* E74F early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-related proteins. *Cell* **61,** 85-99.

Chen J-S., Cho W-L. and Raikhel A. (1994) Analysis of the cDNA for mosquito vitellogenin: similarity with vertebrate phosvitins and arthropod serum proteins. *J. Mol. Biol.* **237,** 641-647.

Cho W., Deitsch K. and Raikhel A. (1991) An extraovarian protein accumulated in mosquito oocytes is a carboxypeptidase activated in embryos. *Proc. Natl. Acad. Sci. USA* **88,** 10821-10824.

Deitsch K. and Raikhel A. (1993) Cloning and analysis of the locus for mosquito vitellogenic carboxypeptidase. *Insect Mol. Biol.* **2**(4) In press.

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

Fryberg E., Kindle K., Davidson N. and Sodja A. (1980) The actin genes of *Drosophila*: a dispersed multigene family. *Cell* **19,** 366-378.

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

Hagedorn H. (1983) The role of ecdysteroids in the adult insect. In *Endocrinology of Insects* (edited by Downer R. and Laufer H.), pp. 271-304. Alan R. Liss, Inc., New York.

Hagedorn H. (1985) The role of ecdysteroids in reproduction. In *Comprehensive Insect Physiology, Biochemistry and Pharmocology* (edited by Kerkut G. and Gilbert K.), vol. 8, pp. 205-261. Pergamon Press, Oxford.

Hagedorn H., Turner S., Hagedorn E., Pontecorvo D., Greenbaum P., Pfeiffer D., Wheelock G. and Flanagan T. (1977) Post-emergence growth of the ovarian follicles of *Aedes aegypti*. *J. Insect Phys.* **23,** 203-206.

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

Hayward M., Brock M. and Shapiro D. (1982) Activation of vitellogenin gene transcription is a direct response to estrogen in *Xenopus laevis* liver. *Nucleic Acids Res.* **10,** 8273-8284.

Kapitskaya M., Cho W-L. and Raikhel A. (1993) Analysis of mosquito ecdysteroid receptor cDNA and expression during vitellogenesis. *Second Int. Symp. on Mol. Insect Sci.* pg. 87.

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

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

Lea A. (1964) Studies on the dietary and endocrine regulation of autogenous *Aedes taeniorhynchus*. *J. Med. Ent.* **1,** 40-44.

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

McKnight G. and Palmiter R. (1979) Transcriptional regulation of the ovalbumin and conalbumin genes by steroid hormones in chicken oviduct. *J. Biol. Chem.* **254,** 9050-9058.

Racioppi J., Gemmill R., Kogan P., Calvo J. and Hagedorn H. (1986) Expression and regulation of vitellogenic messenger RNA in the mosquito *Aedes aegypti*. *Insect Biochem.* **16,** 255-262.

Raikhel A. (1992) Vitellogenesis in mosquitoes. In *Advances In Disease Vector Research*. vol. 9, pp. 1-39. Springer-Verlag, New York.

Raikhel A. and Bose S. (1988) Properties of the mosquito yolk protein: a study using monoclonal antibodies. *Insect Biochem.* **18,** 565-576.

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

Ringold G., Yamamoto K., Tomkin G., Bishop J. and Varmus H. (1977) Glucocorticoid-stimulated accumulation of mouse mammary tumor virus RNA: increased rate of synthesis of viral RNA. *Proc. Natl. Acad. Sci. USA* **74,** 2879-2883.

Sambrook J., Fritsch E. and Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

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

Yamamoto K. and Alberts B. (1976) Steroid receptors: elements for modulation of eukaryotic transcription. *Ann. Rev. Biochem.* **45,** 721-746.

CHAPTER 5

SUMMARY AND FUTURE RESEARCH PROSPECTS

SUMMARY

1. *VCP cDNA cloning and analysis:*

Using antibodies to a purified 53 kDa yolk protein from the female mosquito, *Aedes aegypti*, a partial cDNA was cloned from a λgt11 cDNA library made from mRNA isolated from vitellogenic fat bodies. This initial 840 bp clone was used to rescreen the library and a 1020 bp clone containing the 5' end of the cDNA was isolated. The remaining 3' end of the cDNA was obtained using PCR. The complete nucleotide sequence of the cDNA was determined using the dideoxy sequencing method. Northern analysis was done using the isolated cDNA fragment as a probe to demonstrate the stage-, sex- and tissue- specific expression expected for a gene encoding a mosquito yolk protein. Final confirmation of the identity of the clone was obtained by comparison of the deduced amino acid sequence from the cDNA to that provided by direct N-terminal sequencing of the purified protein. Comparison of the complete deduced amino acid sequence to those found in the data base indicated strong homology to members of a carboxypeptidase family. This strongly indicated that this yolk protein was an enzyme made by the maternal fat body and stored in the developing oocyte for use during embryonic development. Experiments using the serine protease inhibitor DFP confirmed that this protein is in fact a proenzyme produced by the vitellogenic fat body, stored in the egg, then subsequently activated in the developing embryo, showing its highest level of activity during mid-embryogenesis. This protein was therefore named Vitellogenic Carboxypeptidase (VCP).

2. *VCP gene cloning and analysis:*

In order to investigate the regulation of expression of the gene for VCP, the cDNA was used as a probe to screen a λ phage library made from genomic DNA extracted from the colony of mosquitoes maintained in our laboratory. Numerous clones were obtained and restriction mapped. The maps indicated that the cloned came from two distinct groups, representing either two separate alleles of the VCP gene within the population or two copies of the gene within the genome. Genomic blots showed fragments representing both VCP clones in equal proportions. However, blots made from DNA taken from individual mosquitoes showed that the two different clones clearly represent alleles, as homozygotes for each allele exist as well as the corresponding heterozygotes. Both alleles of the gene were sequenced in both directions using the dideoxy method and the nucleotide sequence analyzed both visually and with the aid of a computer. Several potential control elements were found surrounding the coding region. A putative TATA box and arthropod initiator were found near a proposed transcription start site. Primer extension analysis confirmed that transcription initiation does indeed begin in this region. Upstream of this site exists a 16 bp imperfect palindrome repeated 4 times. This sequence shows similarity to previously defined steroid hormone response elements. In addition, several other sequences were identified that are similar to regulatory elements found in other organisms, including a downstream element with similarity to a *Drosophila* fat body enhancer. The function of these putative regulatory elements awaits further analysis.

3. *Control of yolk protein expression by 20E:*

The vitellogenic cycle of female mosquitoes has been shown to be initiated with a rise in the titer of the steroid hormone 20-hydroxyecdysone (20E).

How this hormone affects the rate of transcription of the yolk protein genes has been a matter of debate. Using the protein synthesis inhibitor cycloheximide (Chx), and an *in vitro* culture system of mosquito fat bodies, we attempted to determine whether the effect of 20E on Vg and VCP transcription was direct or indirect. First, the optimum titer of 20E needed in culture to illicit maximum levels of yolk protein production was determined. Then the concentration of Chx needed to reversibly inhibit > 98% of protein production by the cultured fat body was established. With these experimental conditions known, transcription of the genes for Vg and VCP in the presence of 20E with or without Chx in the media was measured. The results indicated that protein synthesis is required in response to 20E before transcription of the yolk protein genes is initiated. This is indicative of a hormone response cascade existing with one or more steps between the 20E/receptor complex and the genes for Vg and VCP. Such a cascade may resemble that proposed by Ashburner when considering the activation of 20E responsive genes in *Drosophila* salivary glands.

FUTURE RESEARCH PROSPECTS

1. *DNA structure surrounding the gene for VCP:*

It was discovered in the late 1960's that eukaryotic genomes contain highly repeated, moderately repeated and unique genomic sequences (reviewed in Britton and Kohne, 1968). The structure of the repetitive DNA can be arranged in three different classes; 1) simple repetitive DNA, usually shorter than 600 bp, interspersed with single- or few-copy segments, 2) DNA organized as long tandem arrays of identical repeating units, and 3) stretches of repetitive DNA in which short sequences are interspersed with one another in complex permutations (Kumar and Rai, 1991). These repeated sequences seem to be maintained through evolution and many such DNA fragments have been found to hybridize to genomic DNA isolated from a number of diverse species. This has been suggested to infer a sequence-dependent function for such families of repetitive DNA (Besansky, 1992). It has also been suggested that the sequences are parasitic elements capable of independently replicating in the genome and survive the pressure of natural selection without fulfilling any useful function (Orgel and Crick, 1980). Some repetitive elements interspersed in the genome are mobile and their ability to transpose can lead to gross evolutionary changes through genomic rearrangements (Georgiev *et al.*, 1982).

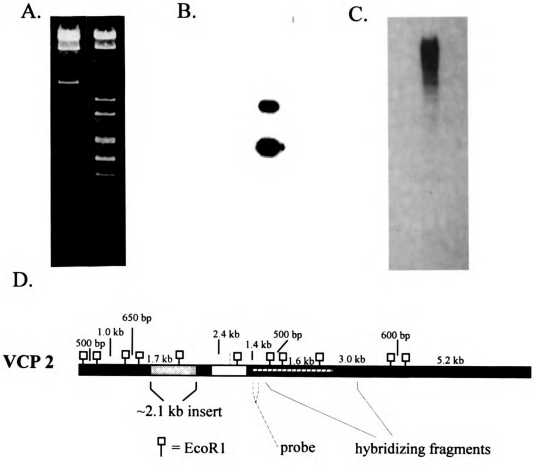
Mosquitoes show a wide range of repetitive DNA content within their genomes, ranging from 20% in *Anopheles quadrimaculatus* to 84% in *Aedes triseriatus* (Black and Rai, 1988). It was found that the proportion of the genome occupied by repeated DNA increases as genome size increases. It was also found that the amplification of short, interspersed repeats accounts for, at least in part, the increased genome size. This is in contrast to *Drosophila melanogaster* in which repetitive DNA is often clustered (Manning *et al.*, 1975).

The species of mosquito used in our laboratory, *Aedes aegypti*, has a relatively large genome, and as a result, would be expected to show a fairly high level of interspersed repetitive DNA.

When the 2 alleles of VCP were cloned and analyzed, a region of very low homology was found downstream of the coding region in both VCP^1 and VCP^2. Partial sequence analysis showed virtually no homology between these two elements, and cross hybridization of the two clones to one another showed that these unique regions begin 110 bp downstream of the polyadenylation signal and extend ~3 kb, after which homology resumes. At first, these non-homologous elements were thought to be artifacts of cloning (i.e. recombinant clones that occurred when the library was made) that did not reflect the actual structure of the mosquito genome. However, two lines of evidence argue against this: 1) numerous, overlapping clones of various lengths were obtained for both VCP^1 and VCP^2. All clones contain these unique elements. It is unlikely that such a recombination event would have occurred in each clone independently, and 2) Southern blots of genomic DNA digested with enzymes that cut within the coding region and within the non-homologous region produced bands of the size predicted by the restriction maps of the cloned DNA. This indicates that the clones have the same restriction map as the isolated genomic DNA. Thus it appears that the clones VCP^1 and VCP^2 do in fact reflect the structure of the locus for VCP.

Next, the structure of these unique elements themselves were considered. When a small fragment of the 5' end of the element found in VCP^2 was labeled and hybridized with Southern blots of both VCP^1 and VCP^2, it hybridized only weakly with a fragment from VCP^1, but hybridized with 2 fragments in VCP^2 (Fig. 1). This indicates that this DNA fragment is repeated within the unique region downstream of the VCP^1 coding region. Similar results were obtained

Fig. 1: Southern blots showing the repetitive nature of the DNA in the region downstream of the VCP locus. A. EcoRI digests of VCP^1 (left) and VCP^2 λ clones run on a 1% agarose gel and stained with EtBr. B. Southern blot probed with the 5' end of the unhomologous region from VCP^2. C. Genomic DNA cut with EcoRI, run on a 0.8% gel and probed with the 5' end of the nonhomologous region from VCP^2. D. Restriction map of VCP^2 showing the probe and the hybridizing fragments.


A.
B.
C.
D.
VCP 2
650 bp
1.0 kb
500 bp
1.7 kb
2.4 kb
1.4 kb
500 bp
1.6 kb
3.0 kb
600 bp
5.2 kb
~2.1 kb insert
= EcoR1
probe
hybridizing fragments

when the same experiment was done using a small fragment of the unique region downstream of VCP1. These experiments suggest that these three kb unique regions are either simple sequences of tandemly repeated DNA, or perhaps that they have repeated DNA at both ends of the element, much like the LTR's of many transposable elements. When fragments from the regions were labeled and used to probe genomic blots, high molecular weight smears resulted, indicating that the sequence is highly repeated within and dispersed throughout the genome (Fig. 1) (Kumar and Rai, 1991).

One of these elements is currently being sequenced to determine its exact structure. Such repetitive elements are often used in the determination of species and in studies of genome evolution (Besansky, 1992). These unique DNA elements may prove to be very interesting fragments of DNA.

2. *Development of an expression system:*

One of the most powerful methods of analyzing the regulatory regions of a particular gene is through the use of an expression system. In this method, a piece of DNA containing the regulatory region of the gene of interest is introduced into the system and the level of transcription measured. Once transcription is established, either the reaction conditions or the DNA itself can be altered to determine the important components involved in transcriptional regulation of the gene being investigated. Many systems have been developed for this purpose, including *in vitro* transcription (Morton and Sprague, 1984), stable and transient cell culture transformation (Fallon, 1986) and germline transformation (Spradling and Rubin, 1982). In insects, by far the most advanced system developed thus far is in *Drosophila*, where the transposable element P is used to introduce foreign DNA directly into the germline. This results in a stabley transformed line of transgenic flies, with the introduced DNA

being inherited in a Medellian fashion. This system has been used to investigate the regulation of numerous genes in fruit flies.

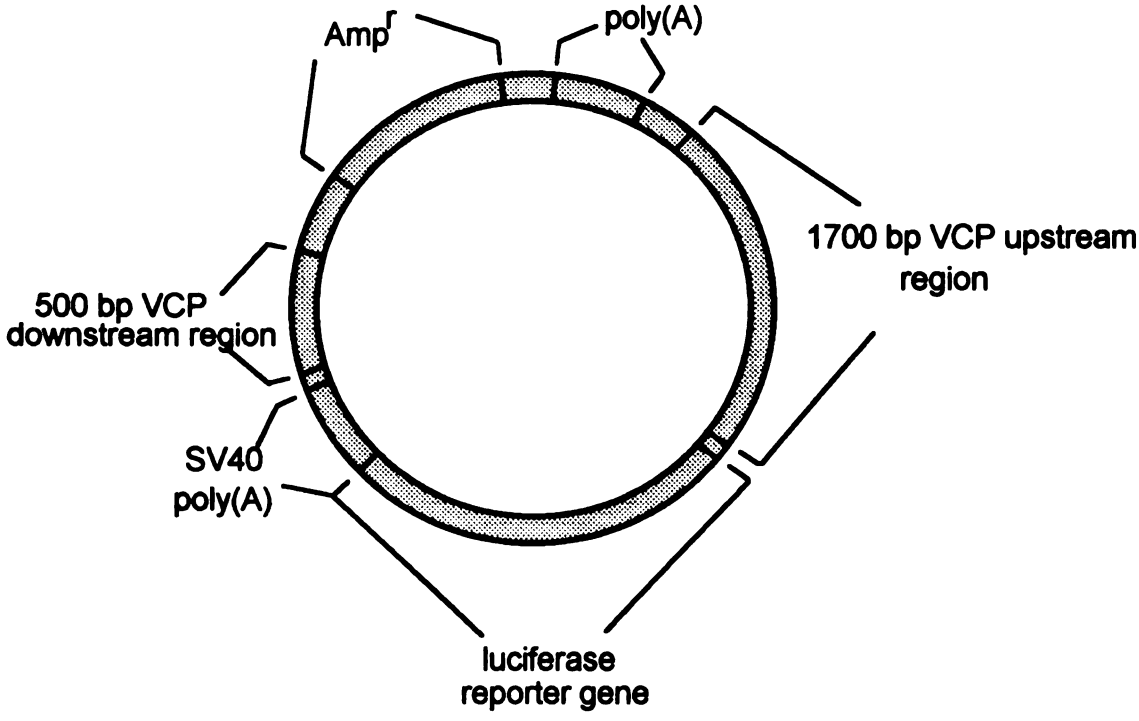
Outside of *Drosophila*, no other system of routine germline transformation has been developed in insects. Alternatively, some expression of exogenous DNA has been observed in cell culture (Fallon, 1986). This system is somewhat limited, however, because the cell lines thus far available are embryonic in origin and as a result, undifferentiated. Thus, any sex-, tissue-, or sex-specific genes are unlikely to be correctly expressed if introduced into such cells. In fact, a construct containing the regulatory region of the mosquito Vg gene introduced into a mosquito cell line showed very low expression, even in response to treatment with 20E (Fallon and Hagedorn, unpublished). Until a cell line expressing female vitellogenic genes has been established, it is unlikely that use of a cell culture system will be an attractive method for analysis of the VCP regulatory region. As an alternative, organ culture systems have been developed as a way to separate differentiated tissue from the organism for more detailed biochemical and physiological experiments. It is possible that transformation of such cultured cells will provide a method for transient expression of exogenous DNA.

Liposomes have been used as a way to introduce foreign DNA into cells (Fraley *et al.*, 1981). In this method, a synthetic cationic lipid is used to make small, unilamellar liposomes that spontaneously interact with DNA to form lipid-DNA complexes. These complexes subsequently fuse with the plasma membrane of the cells, introducing the encapsulated DNA. While providing no basis for DNA integration or transcriptional regulation, this provides an efficient and convenient method for getting a DNA construct into target cells. As a result, it is now widely used to introduce DNA into cultured cells for transient expression assays where stable transformation is either not yet possible or not convenient.

Recently, liposome transfection has been used to introduce foreign DNA into mouse cells *in vivo* (Brigham, *et al.*, 1989). In this particular experiment, the liposome-DNA complex was injected into the blood stream of the animal, and subsequently showed expression in the lungs. This seems to indicate that this method may be applicable to intact animals or tissue culture systems.

In mosquitoes, it was recently reported that liposome-transfection was used to introduce a construct into cultured salivary glands (Morris and James, 1993). A *Drosophila* heat-shock promoter, previously shown to be active in mosquito cells, was used to drive a luciferase reporter gene. The cultured glands were first treated with elastase to remove the basement membrane, then transfected using the liposome-mediated technique. Luciferase activity was detected, indicating the foreign DNA was taken-up by the cells and expressed. This technique may be applicable to VCP analysis. A fat body tissue culture system has been well established and the conditions needed to induce expression of VCP have been defined (Chapter 4). The VCP regulatory region should be an excellent candidate for developing such an expression system do to its tight regulation and its high level of expression when induced. For this purpose, the construct shown in Fig. 2 has been made and will be used in experiments aimed at developing a transcriptional assay. The construct contains 1700 bp of DNA upstream of the start site, including the arthropod initiator, TATA box, putative ribosomal binding site and the 4 repeated imperfect palindromes resembling hormone response elements. Downstream of the reporter are 800 bp of the VCP downstream region, including a small stretch of sequence resembling the *Drosophila* fat body enhancer (Garabedien, *et. al.*, 1986). If transcription can be established using this system, deletions and/or site directed mutagenesis will be used to define those sequences surrounding the VCP gene that are responsible for control of its expression.

Fig. 2: Diagram of the construct to be used in lipofection experiments. This construct was made by inserting the 1700 bp 5' upstream region from VCP1 into the (name of vector) polycloning site. Downstream of the luciferase reporter gene was cloned the 3' region of VCP1 containing a region of similarity to the *Drosophila* fat body enhancer. Also including in the vector are and SV40 poly(A) signal, an ampicillan resistence gene for propagation in *E. coli*, and a poly(A) signal downstream of the amp^r gene to prevent readthrough transcription from the amp^r promoter.


Amp^r
poly(A)
1700 bp VCP upstream region
500 bp VCP downstream region
SV40 poly(A)
luciferase reporter gene

3. *Analysis of DNA/protein interactions:*

Once a region of DNA has been shown to be active in inducing transcription of a particular gene, often it is interesting to know exactly what base pairs are bound by a particular protein, which are important for DNA binding, how tightly the protein binds to the DNA, and what the identity of the regulatory protein is. Often these questions are addressed using footprinting (Carthew *et al.*, 1986) or gel-retardation (Fried and Crothers, 1981) assays. In gel-retardation assays, a labeled fragment of DNA that is thought to be bound by a regulatory protein is incubated in either a crude nuclear extract or in the presence of the purified DNA binding protein. After the protein is allowed to bind to the DNA, the complex is loaded onto an acrylamide gel and electrophoresed to separate bound and unbound DNA. DNA bound by the regulatory protein has its migration through the gel "retarded". Competition assays using cold competitor DNA can be used to show specificity of the interaction and to determine binding affinity of the protein. Usually, a synthetic nucleic acid such as polydI·polydC is used to eliminate non-specific binding. In this way, gel-retardation assays can be a useful technique in identifying important regulatory regions controlling transcription of a particular gene.

Footprinting involves much the same idea as above, except that rather than electrophoresing a DNA/protein complex, this interaction is detected by digesting the DNA with either a nuclease, such as DNAseI (Carthew *et al.*, 1985), or with a chemical such as hydroxyl radicals (Tullius, 1990). The DNA bound by the protein is protected from this digestion, and the protected region is detected on a subsequent sequencing gel. The DNA is end-labeled before the reaction, thus when cleaved, it will produce a ladder of bands on a sequencing gel. Spaces, or "footprints", in this ladder are left where the DNA has been protected by bound protein. In this way, this technique can detect the specific

base pairs bound by the regulatory protein. Altering the binding conditions or mutating the sequence of the DNA can provide information about what base pairs or conditions are necessary for binding.

Once a specific DNA sequence has been identified as being bound by a specific regulatory protein, it is possible to use this sequence to either purify the protein (Rutberg and Ronai, 1992) or to directly clone the cDNA encoding this protein from an expression library (Vinson, *et al.*, 1988; Shea, *et al.*, 1990). Affinity columns can be made using the binding sequence from the regulatory region of a gene as the ligand. Nuclear extracts from tissue actively transcribing the gene are then passed over the column under conditions previously shown to promote DNA/protein interaction. The column is then washed and the highly enriched protein eluted. The characteristics of the purified protein can then be studied or it can be used in the first steps of cloning. Alternatively, a labeled fragment of DNA identified as a binding site can be used as a probe to screen nitrocellulose lifts of proteins produced by clones in an expression library. The DNA should be bound by the regulatory protein of interest and, in this way, the cDNA for this protein is obtained. This is limited to proteins that bind as monomers or homopolymers, and also tends to identify numerous non-specific DNA binding proteins. It has, however, been used to clone several regulatory proteins.

Some of these techniques will undoubtedly be used in the future analysis of the regulation of the VCP gene. Perhaps the most immediately available procedure would be to use several fat body specific regulatory proteins recently cloned from *Drosophila*, including a member of the CREB/ATF family (Abel, *et al.*, 1992) and AEF-1 (Falb and Maniatis, 1992), and to attempt gel-retardation assays using them. The conditions for binding of these proteins, produced in *E. coli*, have been determined. Also, the mosquito homologue of a liver-specific

regulatory protein (HNF-4) has also been recently cloned (Kapitskaya and Raikhel, unpublished). Binding of this protein to the VCP regulatory region could also be tried. Alternatively, crude nuclear extracts could be used to identify important sequences involved in VCP regulation. None of the above techniques have been used in mosquitoes, and their development will undoubtedly require effort. However, with the examples from *Drosophila* available, such methods of analysis will certainly become common.

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 and Develop.*, **6:** 466-480.

Besansky, N.J., Finnerty, V. and Collins, F.H., 1992, Molecular genetics of mosquitoes, In *Advances in Genetics*, **30:** 123-184.

Black, W.C., IV, and Rai, K.S., 1988, Genome evolution in mosquitoes: Intraspecific and interspecific variation in repetitive DNA amounts and organization, *Genet. Res.*, **51:** 185-196.

Brigham, K.L., Meyrick, B., Christman, B., Magnuson, M., King, G. and Berry, L.C., 1989, Rapid Communication: *In vivo* transfection of murine lungs with a functioning prokaryotic gene using a liposome vehicle, *American Journal of the Medical Sciences,* **298:** 178-281.

Britten, R.J. and Kohne, D.E., 1968, Repeated sequences in DNA, *Science,* **161:** 529-540.

Carthew, R.W., Chodosh, L.A. and Sharp, P.A., 1985, An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter, *Cell*, **43:** 439-448.

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

Fallon, A.M., 1986, Factors affecting polybrene-mediated transfection of cultured *Aedes albopictus* (mosquito) cells, *Experimental Cell Research*, **166:** 535-542.

Fraley, R., Straubinger, R.M., Rule, G., Springer, E.L. and Paphadjopoulos, D., 1981, Liposome-mediated delivery of deoxyribonucleic acid to cells: enhanced efficiency of delivery related to lipid composition and incubation conditions, *Biochemistry,* **20:** 6978-6987.

Fried, M.G. and Crothers, D.M., 1981, Equilibrium and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res., **9:** 6505-6525.

Garabedien, M.J., Shepherd, B.M. and Wensink, P.C., 1986, A tissue-specific transcription enhancer from the *Drosophila* yolk protein 1 gene, *Cell*, **45:** 859-867.

Georgiev, G.P., Kramerov, D.A., Ryskov, A.P., Skryabin, K.G. and Lukanidin, E.M., 1982, Dispersed repetitive sequences in eukaryotic genomes and their possible biological significance. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 1109-1121.

Kumar, A. and Rai, K.S., 1991, Organization of a cloned repetitive DNA fragment in mosquito genomes (Diptera: Culicidae), *Genome,* **34:** 998-1006.

Manning, J.E., Schmid, C.W. and Davidson, N., 1975, Interspersion of repetitive and nonrepetitive DNA sequences in the *Drosophila melanogaster* genome, *Cell,* **4:** 141-155.

Morris, A.C. and James, A.A., Transfection of salivary glands from the mosquito, *Aedes aegypti*, *In* Second International Symposium on Molecular Insect *Science,* pg. 113.

Morton, D.G. and Sprague, K.U., 1984, *In vitro* transcription of a silkworm 5S RNA gene requires an upstream signal, *Proc. Natl. Acad. Sci. USA*, **81:** 5519-5522.

Orgel, L.E. and Crick, F.H.C., 1980, Selfish DNA: the ultimate parasite, *Nature,* **284:** 604-607.

Rutberg, S.E. and Ronai, Z., 1992, A simple and efficient method for the purification of specific DNA binding proteins, *Nuc. Acids Res.*, **20:** 1815.

Shae, M.J., King, D.L., Conboy, M.J., Mariani, B.D. and Kafatos, F.C., 1990, Proteins that bind to *Drosophila* chorian cis-regulatory elements: a new C_2H_2 zinc finger protein and a C_2C_2 steroid receptor-like component, *Genes and develop.*, **4:** 1128-1140.

Spradling, A.C. and Rubin, G.M., 1982, Transposition of cloned P elements into *Drosophila* germline chromosomes, *Science,* **218:** 341-347.

Tullius, T.D., 1989, Physical studies of protein-DNA complexes by footprinting, *Annu. Rev. Biophys. Biophys. Chem.*, **18:** 213-237.

Vinson, C.R., LeMarco, K.L., Johnson, P.F., Landschultz, W.H. and McKnight, S.L., 1988, *In situ* detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage, *Genes and Development,* **2:** 801-806.