INTERACTION OF THE LOCUST FABP PROMOTER WITH RXR

by

Li Cui
B.Sc., Nan Kai University, 1988

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Department of Biological Sciences

SIMON FRASER UNIVERSITY

Spring 2008

APPROVAL

Name: Li Cui

Degree: Master of Science

Title of Thesis:

Interaction of the locust FABP promoter with RXR

Examining Committee:

Chair: Dr. J. Mattsson, Assistant Professor

Dr. N. Haunerland, Professor, Senior Supervisor
Department of Biological Sciences, S.F.U.

Dr. C. Beh, Assistant Professor
Department of Molecular Biology and Biochemistry, S.F.U.

Dr. C. Lowenberger, Associate Professor
Department of Biological Sciences, S.F.U.

Dr. I. Novales Flamarique, Associate Professor
Department of Biological Sciences, S.F.U.
Public Examiner

---

25 March 2008
Date Approved

ABSTRACT

The expression of the fatty acid binding protein (FABP) in locust flight muscle is up-regulated by fatty acids *via* a fatty acid response element in the promoter of its gene that binds to unknown nuclear receptor proteins. In this study, the involvement of the retinoid-X receptor was studied. A cDNA library of *Schistocerca gregaria* flight muscle was constructed, and two RXR isoforms were cloned. Yeast one-hybrid interaction screens were used to assess the interaction between the FABP fatty acid response element and RXR. Vectors expressing RXR/GAL4-activation domain fusion proteins were constructed using yeast homologous recombination. RXR/GAL4 transcriptional activation was then tested using vectors containing the wild-type and altered fatty acid response element sequences. These experiments demonstrated that both RXR isoforms interact with the element, suggesting RXR involvement. Further research is needed to reveal the molecular mechanism of FABP gene regulation.

Keywords:

Fatty acid activated receptor; retinoic X receptor; fatty acid response element; locust fatty acid binding protein; heart-type fatty acid binding protein; yeast one-hybrid assay

To my husband Lei and our parents,

and to

my daughters Cindy and Ashley

ACKNOWLEDGEMENTS

I would like to acknowledge many people who made this thesis possible.

My sincere thanks go to my senior supervisor, Dr. Norbert Haunerland, for his guidance, encouragement and patience for the past years. His understanding and supervision helped me to overcome difficulties in my study and research. I am very grateful to the members of my supervisory committee, Dr. Christopher Beh and Dr. Carl Lowenberger, for their consistent guidance and valuable advice. My appreciation is also extended to Dr. Iñigo Novales Flamarique for being my examiner.

Special thanks to Dr. Jutta Rickers-Haunerland for her inspiration and support in my research. Many thanks go to Gabe Alfaro and Giselle Duamel for their assistance in yeast one-hybrid analysis, to Ying Zeng for his help in gene cloning, and to Dipen Thakrar for his help in my research. I appreciate all the present and past members of the Haunerland Lab for an enjoyable, dynamic working environment, especially Dr. Jutta Rickers-Haunerland, David Qu, Jin Wu, Dipen Thakrar, Herbert Sum and Sunny Chahal.

I would finally like to acknowledge the great support from my friends and family. Special thanks to my husband Lei for his understanding and full support, and to our parents for their unconditional love, to my daughters for making my life so complete and wonderful.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3-AT	3-amino-1,2,4-trizole
AD	Activation domain
AKH	Adipokinetic hormone
apoLp-III	Apolipophorin III
CoA	Coenzyme A
CRABP	Cellular retinoic acid binding protein
CRBP	Cellular retinol binding protein
DAG	Diacylglycerol
DBD	DNA binding domain
DR	Direct repeat
EcR	Ecdysone receptor
EMSA	Electrophoretic mobility-shift assay
FA	Fatty acid
FAAR	Fatty acid activated receptor
FABP	Fatty acid binding protein
FARE	Fatty acid response element
FXR	Farnesoid X receptor
HDLp	High density lipophorin
iLBP	Intracellular lipid binding protein
IR	Inverted repeat
LBD	Ligand binding domain
LDLp	Low density lipophorin
LmRXR	*Locusta migratoria* RXR
LTP	Lipid transfer particle
MBS	Mutated bait sequence
MCS	Multiple cloning site
NLS	Nuclear localization signal
NR	Nuclear receptor

PPAR	Peroxisome proliferator acitivated receptor
PPRE	Peroxisome proliferator response element
RA	Retinoic acid
RAR	Retinoic acid receptor
RARE	Retinoic acid response element
RXR	Retinoic acid X receptor
SgRXR	*Schistocerca gregaria* RXR
TAG	Triacylglycerol
T_C	Constitutive TATA box
TMTC	Too many to count
T_R	Regulated TATA box
UAS	Upstream activating sequence
USP	Ultraspiracle
VLDL	Very low density lipoprotein

CHAPTER 1:
GENERAL INTRODUCTION

1.1 Fatty acid transport and utilization

Various different fuels are used by animals to fuel muscle contraction. Generally, carbohydrates are used for immediate energy needs because they can be quickly mobilized and metabolized. In contrast, fatty acid mobilization is slower, but fatty acids (FAs) are preferred as energy source when muscle activity needs to be sustained for long periods, or when carbohydrates are limited. Since most animal muscles store only small amounts of lipids in their cytoplasmic compartments, they must rely heavily on exogenous fatty acids. These are mostly stored in adipocytes as triacylglycerols (TAGs), and will be transported to the target tissues through the circulatory system. Within the muscle cells, fatty acids serve as substrates to provide large amount of energy through their oxidation in mitochondria or peroxisomes. While this general concept is true for all animals, there are some noteworthy differences between mammals and insects with regard to fatty acid transport and utilization.

1.1.1 Mammalian heart and skeletal muscle

In mammals, the cytoplasm of adipose cells is the major site for the synthesis and storage of triacylglycerols. The mobilization and utilization of fatty acids in mammalian heart and skeletal muscles have been well studied (van der Vusse et al., 1992 and 2002).

When fatty acids are needed, triacylglycerols are hydrolyzed into fatty acids and glycerol within the adipose cell by a hormone-regulated lipase. Fatty acids are released into the bloodstream where they are bound by serum albumin and transported to the muscle tissue. Muscles may also take up fatty acids that are contained in lipoproteins, such as very low-density lipoproteins (VLDLs) or chylomicrons (van der Vusse et al., 1992). These particles consist of a triacylglycerol core surrounded by a shell of polar lipids and apoproteins, but they are different in lipid content, density, and origin. Chylomicrons are synthesized from digested lipids and released into the circulation from the intestine. VLDLs are synthesized in the liver and exported directly into the blood. Prior to their transport into cardiomyocytes and skeletal muscle cells, fatty acids are liberated from the triacylglycerol molecules by the action of a lipoprotein lipase, which is attached to the luminal side of the endothelial cell membrane (van der Vusse et al., 2000).

After passing through the endothelial cells and interstitial compartment, fatty acids must cross the sarcolemma and the sarcoplasm before they can be linked to coenzyme A (CoA) at the outer mitochondrial membrane or the sarcoplasmic reticulum (van der Vusse et al., 2002). In a carnitine-mediated process, the activated fatty acids are carried across the inner mitochondrial membrane and degraded via the β-oxidation pathway and the citric acid cycle (van der Vusse et al., 2000). For continuous contraction, as encountered during endurance exercise or in the constantly beating heart, mammalian muscles generate most of the needed energy through fatty acid oxidation. A variety of

other substrates can also provide fuels for their muscle activities including glucose, pyruvate, lactate, ketone bodies and several amino acids (Schaap et al., 1998). However, fatty acid oxidation contributes up to 70% of the overall energy produced in the healthy heart, and may exceed 90% under fasting or diabetic conditions when carbohydrate availability is limited (van der Vusse et al., 1992).

1.1.2 Insect flight muscle

Just like in mammals, insect muscles rely on fatty acids to fuel extended muscle activity, such as that encountered during migratory flight. Migrating locusts, for example, can fly continuously for 10 hours or more over long distances, sometimes exceeding 500 km (van der Horst 1990). The energy required in the initial period of insect flight is mostly provided by quickly mobilized carbohydrates, but after a short period, gradually switches to fatty acids. Sustained flight activity depends almost completely on a highly efficient fatty acid transport system and a high rate of fatty acid oxidation (Haunerland 1997; Weers and Ryan 2003). Based on studies in lepidopteran and orthopteran species (Beenakkers et al., 1985; Haunerland 1997; Weers and Ryan 2003; van der Horst 1990), a unique lipoprotein shuttle system was proposed that assures efficient fatty acid transport during the sustained flight (Figure 1-1).

Fatty acids are stored as triacylglycerols in the insect fat body. As a major metabolic centre, fat body constitutes up to 50% of the fresh weight of insects, and triacylglycerols take up about 90% of the total fat body lipid (Gilbert and Chino 1974), providing sufficient fatty acids to supply the energy for extended flight activity. In response to flight, adipokinetic hormone (AKH) is released from

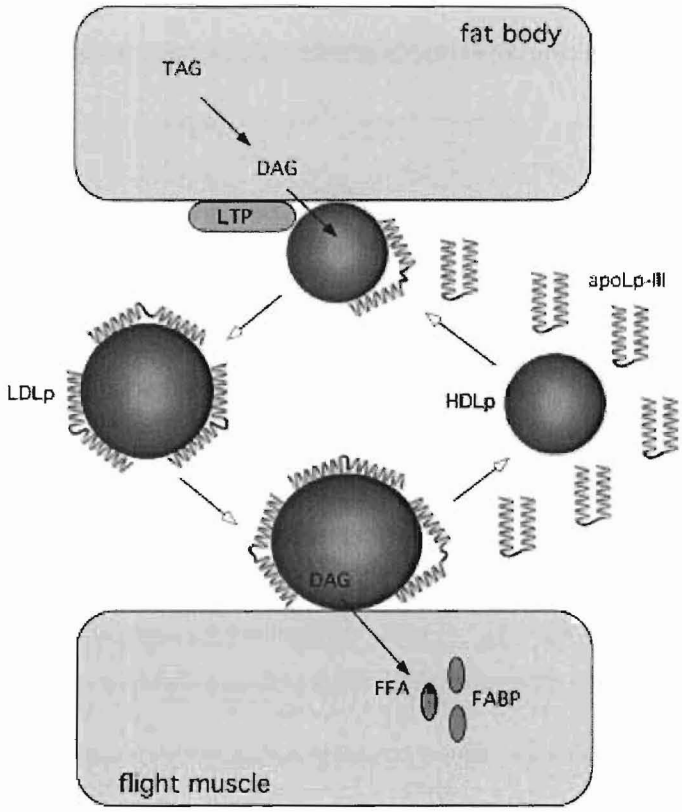


Figure 1-1 Schematic representation of fatty acid transport in insects

TAG — triacylglycerol; DAG — diacylglycerol; LTP — lipid transfer particle; HDLp — high-density lipophorin; apoLp-III — apolipophorin III; LDLp — low-density lipophorin; FFA — free fatty acid; FABP — fatty acid binding protein

the corpus cardiacum (van der Horst et al., 2001); it activates a triacylglycerol lipase in the fat body, which catalyzes the hydrolysis of triacylglycerols and the release of diacyglycerols (DAGs) into the hemolymph. DAG is associated with the core high-density lipophorin (HDLp) particle with the aid of a lipid transfer particle (LTP) (Ryan et al. 1986). As the DAG content in HDLp increases, the binding of free apolipophorin III (apoLp-III) is induced and a DAG-enriched low-density lipoprotein (LDLp) molecule is formed. Low-density lipophorin particles, which are structurally unrelated to vertebrate lipoproteins, deliver DAG to the flight muscle.

Outside the flight muscle cell, DAG is released from LDLp, and hydrolyzed into free fatty acids and glycerol by a yet unknown lipophorin lipase at the flight muscle cell membrane. The DAG-depleted HDLp and free apoLp-III can be reused for future DAG delivery. The free fatty acid molecules are taken up into the muscle cells, where intracellular fatty acid binding proteins (FABPs) facilitate their transport to the outer mitochondrial membrane (Weers and Ryan 2003). Fatty acids are activated by esterification with coenzyme A (CoA), converted to their carnitine esters and shuttled across the inner mitochondrial membrane. In the mitochondrial matrix, FA-carnitine esters are re-transformed to their CoA-esters and oxidized *via* the β-oxidation pathway, providing the energy needed to fuel sustained flight activity (Haunerland 1997).

1.2 Cytoplasmic fatty acid binding proteins

1.2.1 Overview

Once fatty acids enter the muscle cells, the intracellular FABPs are responsible for the efficient transport of fatty acids to the mitochondria. The binding of fatty acids to FABPs results in a large fatty acid pool, thus enhancing the substrate availability for energy production *via* β-oxidation. FABPs can also increase fatty acid uptake into cells by increasing their concentration gradient through minimizing unbound fatty acids in cells (Weisiger 1996). In most cells, the FABP content is proportional to their rate of fatty acid metabolism, and increased fatty acid exposure results the increase of FABP expression. Cytoplasmic FABPs can act as a cytoprotective mechanism against an overload of fatty acids, as free fatty acids have detergent effects (Das et al., 1991) and may form lipid radicals (Yamomoto and Niki, 1990), which damage the membranes. Therefore, by increasing the solubility of fatty acids, cytoplasmic FABPs enlarge the cellular fatty acid pool without damage to cellular structures. In addition, FABPs may also play important roles in fatty acid signalling and gene regulation, thus affecting cellular growth and differentiation (Zimmerman and Veerkamp, 2002).

Since cytoplasmic FABP was first discovered in liver (Ockner et al., 1972), several such proteins have been found in many tissues of different organisms including mammals, fish, birds and insects. They are especially prominent in tissues with active fatty acid metabolism, such as cardiac and skeletal muscle, liver, intestine and adipose tissues (Haunerland and Spener, 2004a). All FABPs

are members of a conserved multigene family of 14-16 kDa intracellular lipid binding proteins (iLBPs) which include FABPs and binding proteins for retinoids and other lipophilic ligands (Glatz and van der Vusse, 1996).

Members of the iLBP family and their ligands are listed in Table 1-1(Haunerland and Spener, 2004a and 2004b). The proposed ligands for members of subfamily I are either retinoic acid or retinol, and thus the binding proteins have been called cellular retinoic acid binding proteins (CRABPs) or cellular retinol-binding proteins (CRBPs), respectively. Besides long-chain fatty acids, some types of FABPs can also bind to acyl-CoA, heme, bile acids or docosahexacenoic acid. FABPs were originally named after the tissue from which they were discovered or where they are expressed abundantly. However, the expression of most types of FABPs is not limited to a certain tissue, and most tissues express various types of FABPs because tissues usually contain different cell types. Heart-type FABP (H-FABP) is very abundant in mammalian heart, but it has also been found, with low content, in skeletal muscles, brain, kidney, lung, mammary, placenta, testis, ovary and stomach. In addition to H-FABP, cardiac and skeletal muscles also contain E-FABP (Haunerland and Spener, 2004a and 2004b).

1.2.2 Heart-type fatty acid binding protein (H-FABP)

The expression of FABP in muscle tissue has been demonstrated in a wide range of animals, from vertebrates (mainly mammals, birds, fish) to invertebrates (mainly insects). Interestingly, H-FABPs from different species are more similar to each other than to other types of FABPs within the same species.

Table 1-1 FABPs as members of the intracellular lipid binding protein family

iLBP-type	Mammalian expression	Non-mammalian expression	Ligands
Subfamily I			
CRBP I, II, III, and IV	Ubiquitous in mammalian cells		Retinol
CRABP I and II	Ubiquitous in mammalian cells		Retinoic acids
Subfamily II			
L-FABP (liver)	Liver, intestine, kidney, lung , and pancreas		Long-chain fatty acids, Acyl-CoA , and heme
I-BABP (intestinal)	Ileum		Bile acids
Lb-FABP (liver-basic)		Fish and bird liver	Long-chain fatty acids
Subfamily III			
I-FABP(intestinal)	Intestine, liver		Long-chain fatty acids
Subfamily IV			
H-FABP (heart)	Heart, skeletal muscle, brain, kidney, Lung, mammary, placenta, testis, ovary and stomach	Fish muscle, bird muscle, insect muscle, fish ovary	Long-chain fatty acids
A-FABP (adipocyte)	Adipose tissue, and macrophages liver	Fish muscle (?)	Long-chain fatty acids
E-FABP (epidermal)	Skin, adipose tissue, lung, brain, heart, skeletal muscle, testis, retina, and kidney		Long-chain fatty acids
B-FABP (brain)	Brain, glia cells, and retina	Bird brain, retina	Long-chain fatty acids and docosahexacenoic acid
M-FABP (myelin)	Brain and Schwann cells		Long-chain fatty acids
T-FABP (testis)	Testis		Long-chain fatty acids

Because of their important functions in fatty acid transport and oxidation during sustained muscle activities, H-FABPs have been well studied in mammalian hearts and insect flight muscles (Haunerland and Spener, 2004a and 2004b).

H-FABP found in mammalian heart contains 132 amino acid residues with a molecular weight of 14.5 kDa (Schaap et al., 1998). It contains a high-affinity fatty acid binding site, and the binding affinity and specificity appear to depend on fatty acid chain length and degree of saturation. H-FABPs are highly conserved even between evolutionary distant species, maybe because the structure and function of muscles are conserved in different species. The FABPs found in the flight muscle of desert (*Schistocerca gregaria*) and migratory (*Locusta migratoria*) locusts are highly similar to their mammalian counterpart and shares many characteristics in amino acid sequence (42% identity and a total of 82% similarity), tertiary structure and binding behaviour (Haunerland 1994). The high similarity between locust FABP and mammalian H-FABP makes locust flight muscle a good model to study FABP gene regulation in a muscle that uses fatty acids at very high rates.

1.3 Locust flight muscle as experimental model

Desert locusts (*Schistocerca gregaria*) have been known since biblical times as major problems in many African and Middle Eastern countries. Their success and destructive potential are to a large part due to their migratory capacity. Their flight muscles are extremely active, and demand huge amount of energy supply during flight. Fatty acids are the main energy source for their flight,

and their unique and efficient lipid transport system constantly transports fatty acids to the cytoplasm, where they are bound by FABP, translocated into mitochondria and metabolized *via* β-oxidation. The metabolic activity and the rate of β-oxidation in locust flight muscle during flight are extremely high, perhaps 20 times as high as in resting animals (Kammer and Heinrich, 1978). The maximal rates of fatty acid utilization in locust flight muscle (0.9 μmol palmitate/min/g muscle weight) are considerably higher than those reported for birds (pigeon pectoral muscle, 0.6 μmol palmitate/min/g) or mammals (rat heart, 0.3 μmol palmitate/min/g) (Crabtree and Newsholme, 1975).

Moreover, the FABP content in the cytosol of different muscles has a strong correlation with their fatty acid oxidation capacity (Haunerland 1994). Very low levels of H-FABP are present in smooth muscles that depend largely on carbohydrates as energy source. Cardiac tissue that depends largely on lipid for energy supply has the highest FABP content in mammals (up to 5% of all cytosolic proteins). Muscles with even higher metabolic rates contain even higher FABP content, such as the flight muscles of Western sandpiper (9% of all cytosolic proteins) and desert locust (18% of all cytosolic proteins).

It has been demonstrated that endurance training or increased fatty acid utilization leads to elevated levels of locust FABP expression (Chen and Haunerland, 1994). As an adult specific protein, locust FABP is absent in immature locusts and newly emerged adults. However, FABP expression levels rise rapidly to adapt to the extreme high metabolic rate when the locust acquires flight ability within the first two weeks of adulthood. The fact that the levels of

FABP expression increase in line with the increasing levels of fatty acid utilization may indicate that flight muscle cells can sense fatty acid levels and respond by enhancing the expression of their binding proteins. It has been proven that this up-regulation occurs at the level of transcription initiation (Zhang and Haunerland, 1998). This adaptation makes locust FABP an attractive model to study the mechanism of fatty-acid-mediated gene control. The locust FABP gene has been cloned, and distinct regulatory regions have been identified in its promoter region (Wu et al., 2001; Wu and Haunerland, 2001).

1.4 Regulation of locust FABP gene expression

1.4.1 The locust FABP gene

The gene structure of the FABP family is highly conserved. All members share an identical gene organization of four conserved exons and three introns of various sizes, and the exon/intron boundaries are in similar positions (Veerkamp and Maatman, 1995), with the only exception that several, but not all insect FABPs lack the second intron.

The gene coding for locust muscle FABP together with a 1.2 kb of upstream sequence have been cloned (Figure 1-2) (Wu et al., 2001). The coding sequence with 607 bp is interrupted by two introns of 12.7 and 2.9 kb. Compared to its mammalian homologues, the two introns are inserted in analogous positions as the first and third introns of the mammalian H-FABP. The total size of the locust FABP gene is 16 kb, but only 4% of the total size encodes FABP, and the sizes of the introns are large. The promoter region includes a canonical

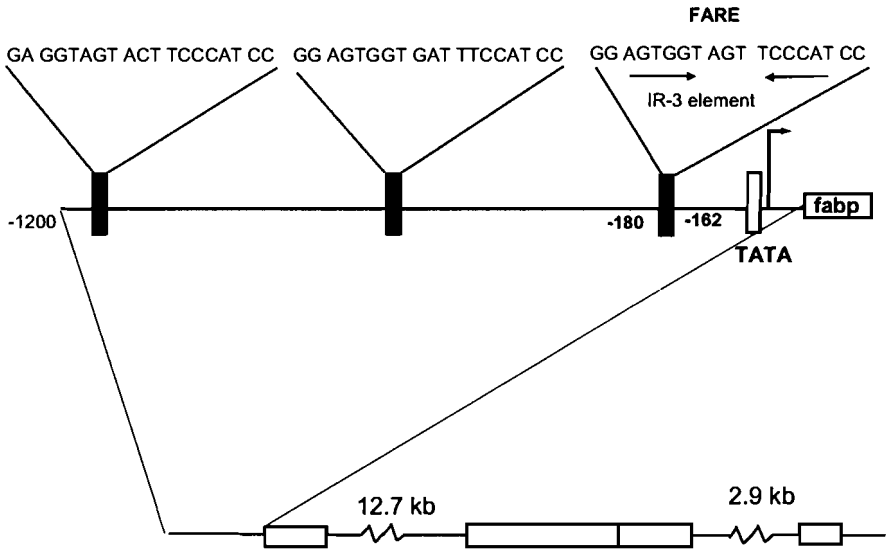


Figure 1-2 Schematic representation of the locust FABP gene

The gene structure of Locust FABP gene is composed of three exons and two introns. A fatty acid response element (FARE) is identified in the promoter region located at -180/-162.

TATA box located 24 bp upstream of the transcriptional start site. Flanking the TATA box are GC-rich areas, which often confer strong transcriptional activation. Three copies of a 19-bp imperfect inverted repeat sequence were found in the promoter region (Figure 1-2); the one located at (-180/-162) was identified as a fatty acid response element (FARE) necessary for the induction of FABP gene expression by fatty acids (Wu and Haunerland, 2001). This locust FARE is an imperfect palindromic sequence, reminiscent of a steroid hormone response element, containing two hexanucleotide half sites (AGTGGT, ATGGGA) separated by three nucleotides. Thus, the FARE can be classified as an IR-3 element.

By gel-shift analysis, it has been shown that nuclear proteins from locust flight muscles bind specifically to the FARE element (Wu et al., 2002). In the presence of nuclear proteins, oligonucleotides containing the locust FARE exhibit retarded mobility during gel electrophoresis. The addition of fatty acids results in a further band shift, indicating a change in protein binding induced by fatty acids. As symmetrical elements often bind two transcription factors, it is possible that two transcription factors, at least one of which is activated by fatty acids, regulate FABP gene expression by binding to the locust FARE. Those transcription factors are most likely members of the nuclear receptor superfamily, because nuclear receptors are transcription factors that regulate gene expression by binding to specific regulatory DNA sequences upon ligand activation.

1.4.2 The nuclear receptor superfamily

Nuclear receptors (NRs), as ligand-activated transcription factors, regulate the expression of target genes by binding to specific cis-acting sequences. Through ligand binding, nuclear receptors respond directly to chemical or hormonal triggers, attach to specific regulatory DNA sequence of target genes and alter the gene transcription (Owen and Zelent, 2000). Their ligands are often small, hydrophobic regulatory signalling molecules, such as steroid hormones, vitamin D, ecdysone, retinoic acids, thyroid hormones, fatty acids, farnesol metabolites, etc. (Laudet 1997). However, the nuclear receptor superfamily also includes a subset of "orphan" receptors that are not known to bind any ligands, as well as those receptors that bind unrecognized ligands (Robinson-Rechavi et al., 2003).

Both structure and function of nuclear receptors have been found to be significantly conserved throughout evolution (Thornton 2003). Phylogenetic analysis suggests they originated from a common ancestral transcription factor with pre-existent DNA binding functionality. Through critical mutations and related evolutionary changes, this original transcription factor acquired the ability to bind various ligands, hold cofactors and create functional dimers, thus giving rise to the variety of factors that exist today in the nuclear receptor superfamily (Owen and Zelent, 2000).

Nuclear receptors share a common structural organization, and are composed of five modular domains (A/B, C, D, E, F) (Laudet and Gronemeyer, 2002). The N-terminal A/B domain is highly variable in length, and shows the

weakest evolutionary conservation. The A/B domain contains at least one autonomous transactivation domain which activates transcription in a constitutive manner when linked to a heterologous DNA binding domain. The C domain (DNA binding domain, DBD) is the most conserved region, and composed of a distinct dual-zinc finger motif (Cys-X2- Cys- X13- Cys- X2- Cys)-(X15-17)-(Cys-X5- Cys- X9- Cys- X2- Cys) and a carboxy-terminal extension(CTE)(Figure 1-3). The four cysteine residues in each motif coordinate with one Zn^{2+} ion. The two zinc figure structures play complementary roles in the DNA binding process. The position of D domain (hinge) follows after A/B and C domain, and is poorly conserved. The D domain allows nuclear receptors to change conformation and subsequently change the activity in response to ligand binding. The hinge region also harbours at least some elements of a functional nuclear localization signal (NLS) that may function to direct nuclear receptors to nucleus. The E domain (ligand binding domain, LBD) is the largest and moderately conserved region. It is the most important region due to its ability to mediate ligand binding, dimerization, nuclear localization and transactivation functions. It has been demonstrated that the ligand binding of LBD can induce a conformational change and allow the binding of coactivators (Renaud et al., 1995). The C-terminal F region shows little evolutionary conservation, and might play a role in coactivator binding.

The ligand-activated nuclear receptors regulate gene expression by binding regulatory DNA element as monomers, homodimers and heterodimers (Giguere, 1999). Most heterodimeric complexes regulate gene expression by

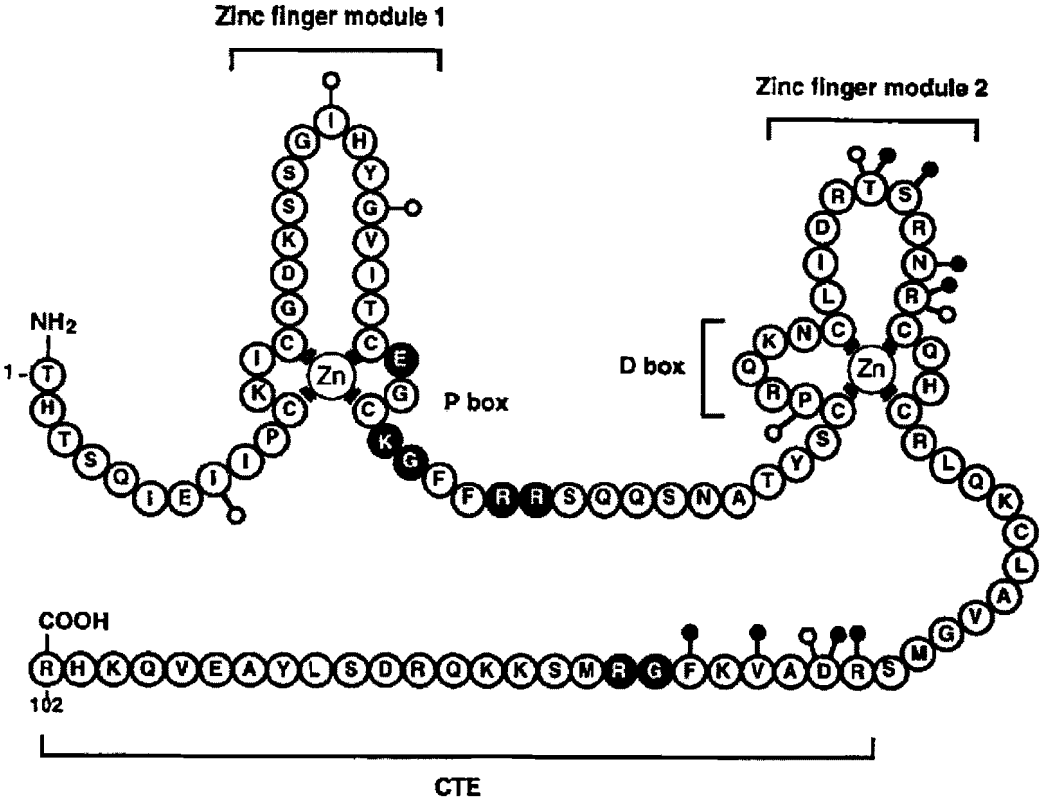


Figure 1-3 Schematic structure of a nuclear receptor DNA binding domain

The nuclear receptor DNA binding domain (DBD) is composed of two zinc finger and a carboxy-terminal extension (CTE).

partnering with one of the retinoid X receptors (RXRs) (Glass 1994; Mangelsdorf and Evans, 1995). The DNA binding sites for nuclear receptors in the gene promoter region are called response elements. Response elements frequently contain two consensus core half-site sequences (A/GGGTCA), with the two half-sites arranged as inverted, everted or direct repeats separated by a variable number of spacer nucleotides between each half site (Laudet and Gronemeyer, 2002). For a dimeric response element, a single conserved half-site is usually sufficient to form homo- or heterodimer complexes (Giguere 1999). Inverted repeats are often found to bind homodimers, while direct repeats and everted repeats seem to bind to both homo- and heterodimers (Laudet and Gronemeyer, 2002). For some nuclear receptors that are capable of forming both homo- or heterodimers, such as the thyroid hormone and vitamin D receptors, the formation of RXR heterodimer is more favourable upon ligand binding (Cheskis and Freedman, 1994; Collingwood et al., 1997; Kakizawa et al., 1997). However, there are many exceptions to these trends; for example, the RXR/ecdysone receptor heterodimer is capable of interacting with natural ecdysone response elements, which are asymmetric elements composed of either imperfect palindromes or direct repeats (Perera et al., 2005).

The nuclear receptors function by binding to DNA response elements in a ligand-dependent manner. The specific binding of nuclear receptors to response elements and the dimerization of nuclear receptors are dependent on both the nature of response elements and the structure of the nuclear receptors. The dimerization of nuclear receptors is primarily controlled by their ligand binding

domains because ligand binding can affect the dimerization properties of nuclear receptors (Laudet and Gronemeyer, 2002). While receptors for hormones have been studied in much more detail, several nuclear receptors have been found to regulate gene expression upon the ligand binding of fatty acids or their derivatives.

1.4.3 Nuclear receptors regulate FABP gene expression

Genes of various FABPs are up-regulated at the transcriptional level by fatty acids, and their gene expression is enhanced by increased fatty acid uptake and metabolism (Haunerland and Spener, 2004a). It was proposed that intracellular fatty acids act as ligands that activate the nuclear receptors, thus modulating the transcription of FABP genes (Duplus et al., 2000). These nuclear receptors may bind as a monomer, homodimer or as heterodimer with RXR to their respective response elements, and thus regulate the expression of their target genes in response to fatty acid binding.

1.4.3.1 Peroxisome proliferator activated receptors (PPARs)

The best known fatty acid activated nuclear receptors are the peroxisome proliferators activated receptors (PPARs). PPARs are involved in the regulation of several muscle-specific genes. Three major forms of PPARs have been identified in mammals: PPARα, β (also called δ) and γ. What the natural ligands of PPARs are is not entirely clear; they bind various polyunsaturated long-chain fatty acids, branched fatty acids, and several eicosanoids (Laudet and Gronemeyer, 2002). It is generally assumed that PPARs bind as heterodimers

with RXR to a direct repeat element in the promoter region of the gene they regulate. This peroxisome proliferator response element (PPRE, consensus sequence AACT AGGNCA A AGGTCA) (Desvergne and Wahli, 1999) consists of a direct repeat of two AGGTCA half-sites separated by a single nucleotide. Such DR1 elements are also target sequences for other nuclear hormone receptor complexes, such as hepatocyte nuclear factor 4(HNF4), the thyroid receptors, or the estrogen receptors (Jump and Clarke, 1999). Generally, PPARs appear to interact with the 5' extended half-site of DR1, whereas RXR binds to the 3' half-site (Desvergne and Wahli, 1999; DiRenzo et al., 1997).

The regulatory function of PPARs in mammalia A-FABP and L-FABP has been well studied. PPARγ2 (an adipocyte-specific N-terminal isoform of PPARγ) binds with RXRα to a direct-repeat element 5.2 kb upstream of the murine A-FABP gene thus regulating its gene expression (Tontonoz et al.,1994). In the liver, the heterodimer of PPARα and RXR binds to the PPRE around 110 bp upstream of the transcriptional start site of the rodent L-FABP gene and stimulates gene expression (Issemann et al., 1992). The L-FABP gene is regulated in the intestine in the same manner, except that PPARβ instead of PPARα is involved (Poirer et al., 2001). Although it has been shown that FABP expression in cardiac and skeletal muscles is also up-regulated by fatty acids, it appears that a PPAR is not involved (Duplus and Forest, 2002). Perhaps other nuclear receptors acting alone or with RXR affect FABP expression in these tissues.

1.4.3.2 Retinoic X receptors

As nuclear receptors of retinoic acid (RA), RXRs function in gene regulation either as heterodimer partners of various other nuclear receptors, or as homodimers (Bouton et al., 2005). There are three major forms of RXRs found in vertebrates: RXRα, β and γ, which all bind 9-cis RA. RXR homologues were also found in variety of metazoans, including insects. The RXR homologue found in *Drosophila melanogaster* is called ultraspiracle (USP). Generally, the RXR homologues in all dipteran and lepidopteran species have been named USP while the name RXR is retained for the homologous proteins in other insect orders as well as other metazoan classes, due to their greater structural similarity to the vertebrate RXRs (Laudet and Gronemeyer, 2002). In addition to the natural ligand 9-cis RA, methoprene (an insect juvenile hormone analog) and docosahexaenoic acid (a long-chain polyunsaturated fatty acid highly enriched in mammalian brain) have been reported to activate mammalian RXRs (Harmon et al., 1995). However, no ligands have yet been identified for *Drosophil*a USP or other insect RXRs, perhaps because ligands for insect RXRs may be non-specific, or because they have lost their ligand-binding ability.

As indicated above, RXRs bind target response element either as homodimers or as heterdimers of other nuclear receptors. Like most nuclear receptors, RXRs recognize the consensus sequence AGGTCA. While a DR1 element is the common target site of RXR homodimers, RXR heterodimers can recognize distinct types of response elements, such as direct repeat element separated by 2-5 nucleotides, or palindromic sequences (Rastinejad 2001). In

insects, the USP/Ecdysone receptor (EcR) heterodimer can bind a palindromic sequence as well as a direct repeat based on AGGTCA, with repeats separated by a unique central base A or T (Laudet and Gronemeyer, 2002).

Since RXRs are common partners of numerous nuclear receptors, they are involved in the regulation of numerous different genes. Within the iLBP family, analogous mechanisms to the above mentioned PPRE-mediated regulation of A- and L-FABP have been reported for the ileal bile acid binding protein (I-BABP) and the cellular retinoic acid binding proteins (CRABPs) (Hwang et al., 2002; Noy, 2000). Farnesoid X receptor (FXR), the bile acid receptor, regulates the gene expression of I-BABP as a heterodimer of RXRα, by binding to the conserved inverted repeat (IR-1) located within 200 bp of the transcriptional start site in the promoter of the human, rabbit, and mouse I-BABP genes (Hwang et al., 2002). Similarly, the genes for cellular retinoic acid binding protein are upregulated by their ligands. As binding proteins for retinoic acids, CRABPs transport retinoic acids to the nucleus, and their genes are under the control of retinoic acid response elements (RAREs) to which the heterodimer of RXR and the retinoic acid receptor (RAR) is bound (Noy, 2000).

Hence, a highly analogous mechanism of ligand-activated gene regulation exists for various members of the iLBP family. Fatty acids activate A-and L-FABP expression via PPAR/RXR dimers, while bile acids stimulate the I-BABP gene through FXR/RXR, and retinoic acid regulates the expression of its binding protein (CRABP) via RAR/RXR. Thus, one can predict that the observed up-regulation of FABP in muscle is also due to fatty acid binding to a complex of

RXR and another nuclear receptor. As outlined above, this up-regulation is present not only in vertebrates, but in an even more pronounced way in the locust flight muscle as well.

1.4.4 Hypothetical mechanism for locust FABP gene regulation

Based on our current knowledge, a hypothetical model (Figure 1-4) has been proposed to explain the up-regulation of locust FABP expression by fatty acids (Qu et al., 2007). Cytoplasmic FABP acts as a fatty acid sensor, binds and transports fatty acids to the nucleus. When fatty acid levels are high, a hypothetical nuclear transcription factor (fatty acid activated receptor, FAAR) is activated by fatty acids, binds to the conserved control element (fatty acid response element, FARE) upstream of the locust FABP gene as a hetero- or homodimer, and enhances the transcription of the FABP gene. Because RXR is a common partner of heterodimers, it is possible that FAAR regulates FABP gene expression by binding to FARE as a RXR heterodimer.

To date it remains unclear which transcription factors mediate the up-regulation of the H-FABP gene in mammals or the analogous FABP gene in locust flight muscle. A PPAR/ RXR heterodimer, however, appears to be not involved the regulation of the H-FABP gene (Schachtrup et al., 2004), and hence one might predict a different nuclear receptor. This is certainly true for insects, which do not possess homologs of PPAR, as these receptors are believed to have evolved relatively recently (Owen and Zelent, 2000). Consequently, functional PPREs have not been found in the muscle FABP genes of vertebrates or insects. The identified locust fatty acid response element is an IR-3 element,

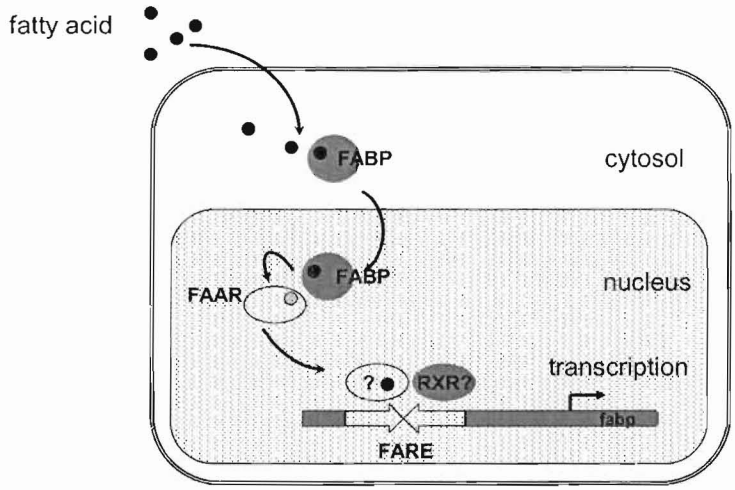


Figure 1-4 Hypothetical mechanism of locust FABP gene expression

FABPs bind and transport fatty acids to the nucleus, where fatty acids activate the potential fatty acid activated receptor (FAAR). The RXR heterodimer of FAAR up-regulate the FABP gene expression by binding to the fatty acid response element (FARE).

and different from PPREs (Haunerland and Spener, 2004a). There are far fewer members of the nuclear receptor superfamily known in insects than in mammals. Most of the identified receptors are involved in insect development. In locusts, only two nuclear receptors have been found to date, the ecdysone receptor and RXR. These act as heterodimers in the regulation of ecdysone-controlled genes (Hayward et al. 1999 and 2003).

1.5 Research objectives of thesis

The aim of this thesis was to identify and clone the potential FAARs in the locust flight muscle, and to detect their interaction with the identified locust FARE. While locust flight muscle is an attractive model for this research from a physiological point of view, the lack of genomic resources makes this a challenging project. Only two nuclear receptors have been found in locusts, and EcR is unlikely bind to fatty acids. RXR, on the other hand, is one likely binding partner, if indeed the protein is expressed in flight muscle of adult locusts where the fatty acids mediated up-regulation has been observed (Chen and Haunerland, 1994). So far, RXR has not been identified in *Schistocerca gregaria*, but only in the closely related species *Locusta migratoria* (Hayward et al. 1999 and 2003). Based on that sequence, RXR from *S. gregaria* was cloned, and its expression in muscle established. Since the mechanism of the up-regulation involves protein-DNA interaction, the interaction of RXR with the response element can be investigated with a yeast one-hybrid analysis system. Moreover, other candidate fatty acid activated receptors may be identified from a one-hybrid screen of a cDNA library of locust flight muscle genes.

CHAPTER 2:
CONSTRUCTION OF A CDNA LIBRARY FOR YEAST ONE-HYBRID SCREENING

2.1 Introduction

Yeast one-hybrid library screening is a method to identify and characterize proteins that bind to a target DNA sequence. To find potential DNA-binding proteins, a cDNA library must be constructed, expressed, and screened in yeast cells in the presence of bait vector containing the target DNA sequence. With the BD Matchmaker Library construction and screening kit (Clontech, Mountain View, CA), a double-stranded (ds) cDNA library is synthesized from total RNA or mRNA, then co-transformed into yeast cells with a bait vector and a *Sma* I-linearized GAL4 AD vector (a GAL4-activation-domain-expressing vector). In the yeast cotransformants, cDNAs will recombine *in vivo* into the linearized AD-expressing vector through homologous recombination, and will be expressed as fusion proteins with the GAL4 activation domain. The bait vector contains tandem copies of the target sequence upstream of the *HIS3* minimal promoter ($P_{minHIS3}$) and the *HIS3* nutritional reporter gene. The interaction between a DNA-binding protein and the target DNA activates the transcription of *HIS3* reporter gene, so the histidine auxotrophic host becomes histidine prototrophic and grows on the minimal media lacking histidine. Colonies growing under these conditions are potential positive clones indicative of a possible DNA-protein interaction. By isolating cDNA-containing plasmids from positive clones, one can isolate the

cDNA sequences for candidate proteins, and confirm the DNA–protein interaction by further experiments.

Building on our earlier discovery of a fatty acid response element (FARE) upstream of the locust FABP gene (Wu et al., 2002; Wu and Haunerland, 2001), we decided to employ the one-hybrid screening technique to identify the transcription factor(s) that interact with this element. The library screening procedure involves the steps of cDNA library synthesis, FARE-bait vector construction and screen optimization, yeast co-transformation as well as library screening.

2.1.1 cDNA library synthesis

The synthesis of a ds cDNA library with the Clontech kit employs the unique BD SMART technology (Figure 2-1). The CDSIII primer used for the first-strand cDNA synthesis from mRNA is a modified oligo (dT) primer. It has two degenerate nucleotides at the 3' end, which positions the primer at the start of the poly A tail and prevents mispriming at an internal site within the tail. The MMLV reverse transcriptase exhibits terminal transferase activity upon reaching the 5' terminus on the RNA template and adds a few additional nucleotides, primarily deoxycytidine to the 3' end of the first-strand cDNA. The addition of the BD SMART III primer (with an oligo (G) sequence at the 3' end) creates an extended template by base-pairing with deoxycytidine. The reverse transcriptase continues the replication process until the end of the oligonucleotides, producing the single-stranded (ss) cDNA with the complete 5' end of the mRNA, plus the sequence complementary to the BD SMART III oligo. In the subsequent

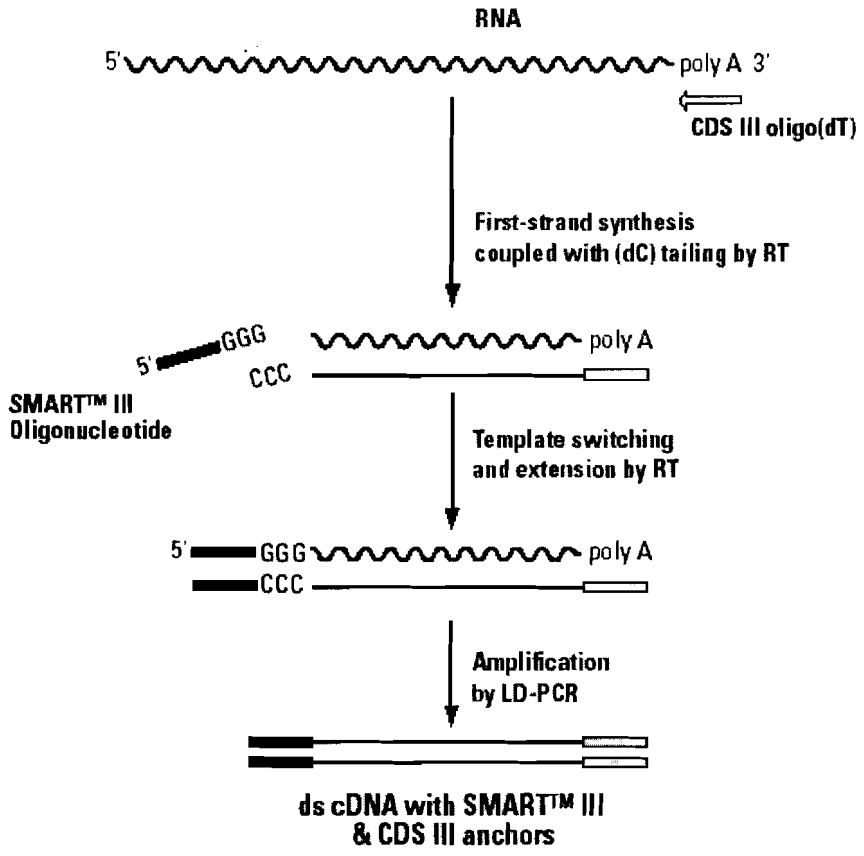


Figure 2-1 Schematic representation of double-stranded cDNA synthesis

amplification of ss cDNA by long-distance PCR (LD-PCR), all the ds cDNA produced will have a BD SMART III anchor at one side and a CDSIII anchor at the other side. The presence of both anchors is mandatory for the recombination-mediated cloning of cDNA into the AD-expressing vector and the construction of a cDNA expression library.

The expression of the cDNAs in this library depends on their *in vivo* cloning into the AD-expressing vector in yeast cells. The screen of the cDNA expression library in yeast requires the presence of the FARE-containing reporter vector. These vectors have some extra features compared to bacterial vectors to permit their propagation in yeast.

2.1.2 Characteristics of yeast vectors

Plasmid vectors used in yeast are commonly constructed from bacterial plasmids with the addition of some special features. They are shuttle vectors, and contain different selective markers for selection and different specific sequences for proper propagation in both yeast and bacteria (Sikorski and Hieter, 1989; Rose and Broach, 1991). The yeast vectors are unique in that they contain both yeast origin of replication and bacterial origin of replication, so they can propagate properly both in yeast and in bacteria. They also contain markers that allow the selection of plasmid-bearing transformants both in yeast and in bacteria. Their selection in bacteria usually depends on the selectable markers with antibiotic resistance, while their selection in yeast most commonly depends on nutritional markers such as *URA3, HIS3, LEU2* and *TRP1.* These nutritional markers complement specific auxotrophic mutations in yeast, such as *URA3* for

mutations with uracil auxotrophy, *HIS3* for histidine auxotrophy, *LEU2* for leucine auxotrophy and *TRP1* for tryptophan auxotrophy (Burke et al., 2000).

The pHIS2 vector used for constructing the FARE-bait is a YCp vector. It contains an *ARS4* element (autonomous replication sequence) derived from the centromere of chromosome IV (Gunge 1983) and a *CEN6* element (centromeric sequence) derived from yeast chromosome VI (Panzeri and Philippsen 1982). The presence of *CEN6/ARS4* cassette makes the vector mitotically and meiotically stable, ensuring the proper replication and segregation during cell division in yeast. The combination of *CEN6* and *ARS4* also function in maintaining low copy level of the plasmid in yeast host cells (Sikorski and Hieter 1989; Rose and Broach 1991). The *TRP1* gene in pHIS2 vector functions as a selective nutritional marker for the selection of the yeast transformants. Its gene product, phosphoribosyl-anthranilate (PRA) isomerase, is an enzyme required for the biosynthesis of the amino acid tryptophan (Braus et al., 1988). The Col E1 origin and the kanamycin resistance gene function in the proper propagation and the selection of the plasmid in *E. coli*.

Importantly, the *HIS3* gene in the pHIS2 vector serves as a reporter gene in the yeast one-hybrid assay. The control of *HIS3* gene regulation makes pHIS2 an ideal reporter vector for the identification and characterization of DNA-binding proteins. The insertion of the FARE into the pHIS2 vector results in a FARE-bait vector, which will serve as a bait to screen the cDNA library for FARE-binding proteins.

2.1.3 *HIS3* reporter gene in FARE-bait vector

A target bait vector is constructed by inserting tandem copies of cis-acting target element into the one-hybrid reporter vector pHIS2 (Figure 2-2). The target copy number should be between one and six (Ghosh et al., 1993). The multiple cloning site (MCS) of pHIS2 is just upstream of the *HIS3* minimal promoter ($P_{minHIS3}$) and the *HIS3* nutritional reporter gene. This reporter gene encodes His3 protein, i.e., imidazoleglycerol-phosphate dehydratase (IGP dehydratase), which catalyzes an essential step in the biosynthesis of histidine (Fink 1964). Without this enzyme, the yeast strain used is histidine auxotrophic; it needs histidine as a nutrient. With the expression of the His3 protein, yeast strain becomes histidine prototrophic. The level of its expression can be measured semi-quantitatively by growing the strain on media containing different concentrations of 3-amino-1, 2, 4-trizole (3-AT), a competitive inhibitor of His3 protein (Horecka and Sprague, 2000).

The growth selection due to the *HIS3* reporter is a crucial step to detect the DNA-protein interactions in the yeast one-hybrid assay. By restriction digestion and ligation, tandem copies of FARE sequence can be inserted into the MCS. The link of the target DNA sequence to the *HIS3* reporter makes it possible to detect the trans-acting factors that regulate the gene expression. The interaction of the GAL4 AD fused-proteins with the target FARE sequence enhances the basal transcription of *HIS3* reporter gene, so the yeast host strain will grow well on the minimal medium lacking histidine.

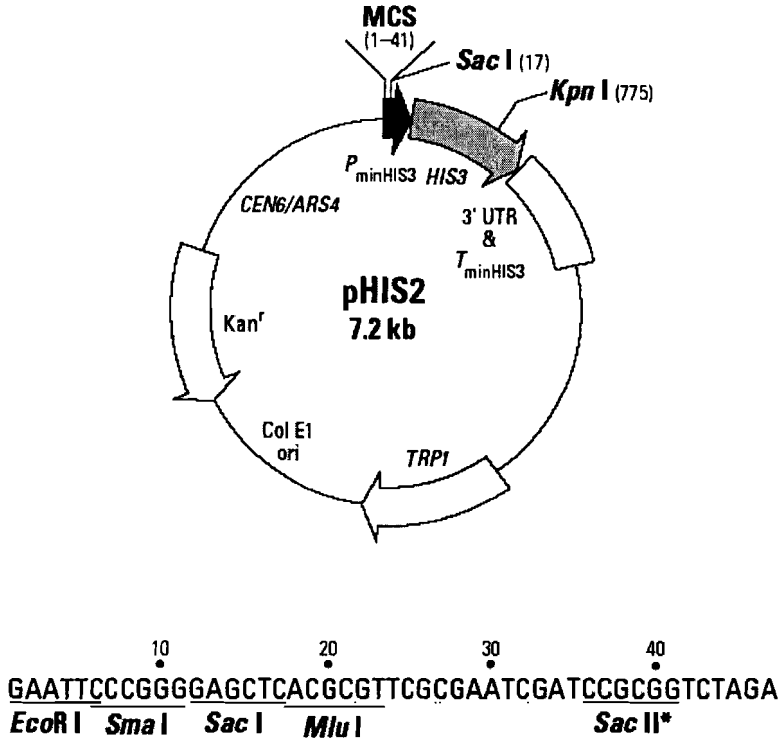


Figure 2-2 Map of pHIS2 reporter vector

The minimal promoter of the *HIS3* locus ($P_{minHIS3}$) and the *HIS3* nutritional reporter gene are located just downstream of the MCS. The restriction sites in MCS are shown below the map of the vector

The basal expression of the *HIS3* gene in yeast is under the control by its yeast-specific promoter. Generally, yeast promoters are composed of initiator element TATA box and associated upstream cis-regulatory elements (Stargell and Struhl, 1996). The TATA sequence defines the transcriptional start site and determines the basal transcription level of a gene, so it is also referred to minimal promoter. Transcription factors or other trans-acting regulatory proteins regulate the transcription level by binding to the cis-regulatory elements. Upstream activating sequence (UAS) is one common cis-regulatory element which enhances the transcription after being activated by specific transcriptional activators.

The native yeast *HIS3* promoter contains two TATA boxes. They are T_R (regulated TATA box) and T_C (constitutive TATA box). It has been reported that the transcriptional activator proteins GCN4 or GAL4 activate the transcription of the *HIS3* gene by binding to the UAS, thus affecting T_R, but not T_C. In contrast, T_C is responsible for the low-level constitutive *HIS3* expression, resulting in a low background or "leaky" expression of *HIS3* (Mahadevan and Struhl, 1990). The pHIS2 reporter vector has both T_R and T_C in the *HIS3* minimal promoter, but it does not contain UAS elements. The inserted target response elements upstream of $P_{minHIS3}$ function as potential UAS and are used as bait for "fishing out" the candidate DNA-binding proteins. Because all proteins encoded by the cDNA library are fused with a GAL4 AD, the binding to their target DNA sequence will bring the GAL4 AD in proximity of the promoter, which results in the activation of *HIS3* gene expression through the regulated TATA box (T_R).

Due to the characteristics of the *HIS3* promoter, low levels of His3 protein are constitutively expressed in yeast even without activation. In order to detect the FARE-binding proteins, which results in the activation of *HIS3* gene expression, the FARE bait vector must be transformed into Y187 yeast strain and tested for the background *HIS3* expression in the absence of activation.

2.1.4 Optimization of *HIS3*-based FARE one-hybrid screens

The *Saccharomyces cerevisiae* Y187 strain functions as the host in the yeast one-hybrid assay. The genotype for Y187 is *MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met –, gal80Δ, URA3::GAL1UAS-GAL1TATA-lacZ, MEL1* (BD Matchmaker Library Construction & Screening Kits User Manual). The Y187 strain cannot grow on 5 types of Synthetic Complete (SC) selection media (SC/-His, SC/-Ade, SC/-Trp, SC/-Leu and SC/-Met) due to the alteration or deletion of the corresponding genes required for their growth on above specific media. However, Y187 grows on SC /-Ura or YPDA media.

Several methods are commonly used for introducing plasmid DNA into yeast. The spheroplast transformation involves the preparation of spheroplast and the incubation with DNA, polyethylene glycol (PEG) and CaCl2 (Hinnen et al., 1978). The lithium acetate (LiAc) transformation eliminates the step of spheroplast preparation by treating intact yeast cells with LiAc followed by incubating the treated cells with DNA and PEG (Ito et al., 1983). The electroporation transformation has been developed to introduce macromolecules into yeast at high efficiencies (Becker and Guarente 1991). LiAc-mediated transformation is the most convenient method due to its ease and high

reproducibility. Five alkali cations (Li+ (most effective), Cs+, Rb+, K+, and Na+) have been proven to induce the competence of the yeast cells and facilitate their uptake of plasmid DNA, and transformations with the LiAc method obtain the highest efficiency. Negatively charged PEG promotes the transformation by changing membrane charges. The efficiency can be further increased by heat pulse or higher concentrations of plasmid DNA (Ito et al., 1983).

After the FARE-bait vector is transformed into yeast cells, its minimal *HIS3* promoter will express low levels His3 protein in the absence of activation. However, this constitutively expressed enzyme is sufficient for the yeast transformants to grow on the minimal medium lacking histidine. When the fusion protein (composed of candidate DNA-binding protein fused to a GAL4 activation domain) binds to the inserted target response element, T_R is affected and the *HIS3* expression level is enhanced beyond the basal level. Therefore, the target bait constructs should be tested for background *HIS3* expression before the one-hybrid screening of a cDNA library. Background growth due to the leaky *HIS3* expression can be reduced by adding 3-AT (3-amino-1, 2, 4-trizole) to the selection medium. The concentration of 3-AT needed to fully eliminate leaky *HIS3* expression must be experimentally determined for each DNA target element. Only one-hybrid positive clones should grow on the selection medium containing the optimal concentration of 3-AT, because all *HIS3* expression is due to the binding of the fusion protein to the target element. With the FARE-bait inserted into the promoter of the *HIS3* reporter gene, it should thus be possible to screen a cDNA library for FARE-binding proteins.

2.1.5 Principle of yeast one-hybrid library screening

Many eukaryotic transcription factors have two structurally and functionally distinct domains. The DNA binding domain (DBD) binds to the target response element while the transcription activating domain (AD) stimulates the downstream gene expression. Based on above fact, DNA-protein interactions can be detected by performing the one-hybrid assay in the Y187 strain that is auxotrophic for histidine (Figure 2-3). The target bait vector, the linearized AD-expression vector pGADT7-rec-2 and the ds cDNA library can be co-transformed into this strain for the purpose of one-hybrid cDNA library screening. The target bait vector is constructed by inserting tandem copies of target response element into pHIS2 vector upstream of the $P_{minHIS3}$ and *HIS3* reporter gene. It functions as a bait to screen a GAL4 AD/cDNA fusion library for proteins that interact with the target sequence.

To isolate potential binding proteins specific to the target response element, gene fusions are constructed by in vivo cloning the cDNA library into the AD-expressing vector pGADT7-rec-2 in yeast though homologous recombination. With the SMART technology, all the ds cDNAs in the library have a BD SMART III anchor and a CDS III anchor. The Sma I-linearized pGADT7-rec-2 has a BD SMART III anchor and a CDS III anchor at each end, too. Therefore, cDNAs can be cloned into pGADT7-rec-2 by homologues recombination in yeast, resulting GAL4/cDNA prey vectors. The expressed proteins in yeast are composed of library proteins fused with the GAL4 activation domain. When the candidate proteins bind to the target response element, the fused GAL4 activation domain

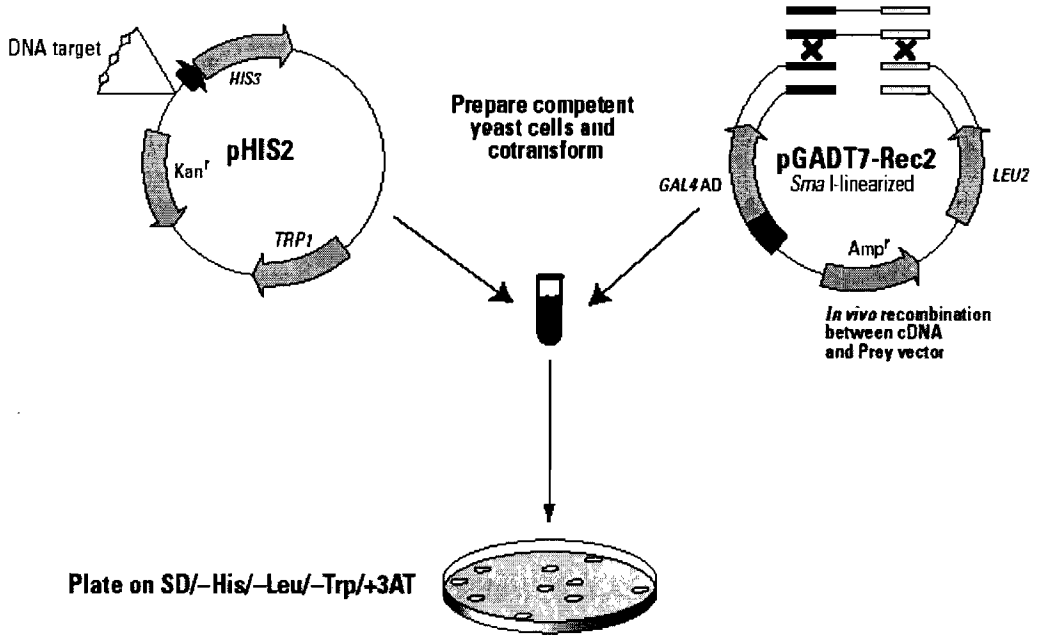


Figure 2-3 Flow chart of library screening with the yeast one-hybrid assay

will simultaneously activate the *HIS3* gene expression beyond the basal level by affecting T_R of the *HIS3* minimal promoter. The enhanced *HIS3* expression due to the positive one-hybrid interactions can be detected by plating the transformants on the selection plates (SC/-Trp/-His/-Leu + optimal 3-AT). After the prey vector in the positive clones is isolated, genes encoding the binding proteins can be analyzed by PCR colony screening and the gene sequence can be obtained by DNA sequencing.

One-hybrid positive and negative controls can be performed before the step of cDNA library screening. The prey vector pGAD-Rec2-53 encodes murine p53 protein fused to the GLA4 AD. P53 is a DNA-binding protein which binds to its target response element contained in the positive control reporter vector p53HIS2. So the prey vector pGAD-Rec2-53 serves together with p53HIS2 in yeast as a one-hybrid positive control, and together with "empty" reporter vector pHIS2 (i.e., without inserted bait sequences) as a negative control.

2.2 Materials and methods

2.2.1 Total RNA extraction

The flight muscle of desert locusts (7- 8 days after adult ecdysis) was homogenized in 1 ml of TRIzol reagent (Invitrogen, Burlington, ON). The addition of 200 μl chloroform, followed by centrifugation, separated the homogenate into three phases. The aqueous phase was mixed well with 600 µl ethanol, loaded onto an RNeasy mini column and processed according to the RNeasy Mini kit protocol (Qiagen, Mississauga, ON). Total RNA was eluted twice with 30 μl of

RNase-free water each, and precipitated overnight at -20 °C after the addition of 6 µl sodium acetate (3 M) plus 75 µl ethanol. After centrifugation, the pellet was washed with 75% and 100% ethanol, then air-dried and re-suspended in 30 µl of RNase-free water. The concentration and purity of RNA were determined by checking the absorbance at 260 nm and 280 nm. RNA integrity was assessed by eletrophoresis in a denaturing formaldehyde agarose gel.

2.2.2 cDNA Library synthesis

A ds cDNA library was synthesized according to the protocol of BD Matchmaker Library Construction& Screening Kit (Clontech, Mountain View, CA). Total RNA (1 to 2 μg) was used to synthesize the first-strand cDNA from the CDIII primer (ATTCTAGAGGCCGAGGCGGCCGACATG$(T)_{30}$VN) and BD SMART III Oligo (AAGCAGTGGTATCAACGCAGGTGGCCATTATGGCCGGG). A 2 μl aliquot from the first-strand synthesis was used to amplify ds cDNA by long distance PCR with 5' PCR primer TTCCACCCAAGCAGTGGTATCAACGC AGAGTGG and 3' PCR primer GTATCGATGCCCACCCTCTAGAGG CCGAGG CGGCCGACA. The cycling parameters were: denaturation 5 min at 95 °C, followed by 20 cycles of 10 s at 95 °C, 6 min plus 5 s per each successive cycle at 68°C, and a final elongation time of 5 min at 68 °C. PCR products were analyzed on a 1.2% agarose gel, purified with BD CHROMA SPIN TE-400 Column and used for *in vivo* cloning into the linearized AD-expressing vector in yeast transformants.

2.2.3 Synthesis of FARE target element

Two complementary oligonucleotides were designed and ordered from Sigma (Oakville, ON): Oligo-1 (AATTCGGAGTGGTAGTTCCCATCCGGAGTGG TAGTTCCCATCCGGAGTGGTAGTT CCCATCCA) and Oligo-2 (CGCGTGGAT GGGAACTACCACTCCGGATGGGAACTACCACTCCGGATGGGAACTACCACT CCG).The two oligonucleotides (0.5 μg each) are mixed with 50 μl of 50 mM NaCl and incubated for 15 minutes at 100 °C, then cooled down slowly to room temperature. The size of the band (supposed to be 63bp) on 1.5% agarose gel confirmed the success of the annealing procedure (Figure 2-4).

2.2.4 Construction of FARE bait vector

pHIS2 vector (0.5 μg) was digested with EcoR I and Mlu I (Invitrogen, Burlington, ON) in a total volume of 20 μl, then the digest was purified with the QIAquick PCR purification kit (Qiagen, Mississauga, ON) and eluted in 18 μl water. Electrophoresis of the digest product showed the sharp band at the size of 7 kb, indicating the complete digestion (Figure 2-5). The double-stranded FARE target element was ligated into the digested pHIS2 vector by T4 ligase (Invitrogen Burlington, ON). MAX Efficiency DH5α competent cells (Invitrogen Burlington, ON) were transformed with the resulting FARE bait vector, using the 42 °C heat-shock method. The transformants were plated onto the LB/kan agar medium and incubated overnight at 37 °C. Isolated colonies were picked and grown overnight at 37 °C in 5 ml of LB/kan broth. Plasmid DNA was prepared from the cell culture with GenElute Plasmid Miniprep Kit (Sigma Oakville, ON). The plasmid was first analyzed by double digestion with *Eco*R I and *Mlu* I, and

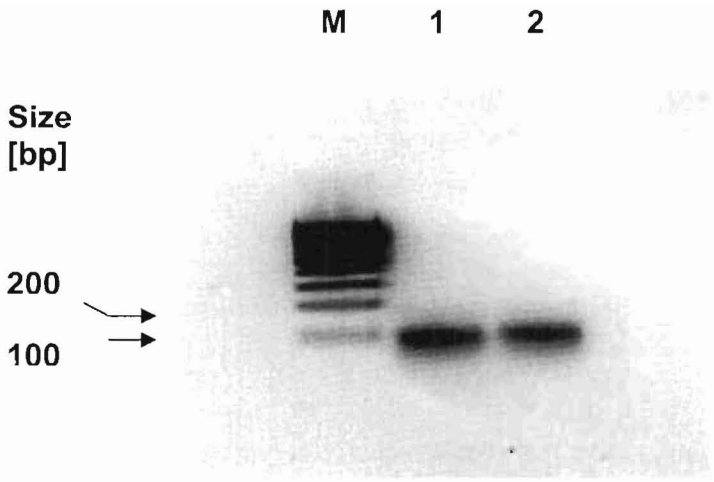


Figure 2-4 Analysis of the FARE tandem repeat

The synthesized sense and antisense oligonucleotides of the FARE tandem repeat were annealed, and analyzed on a 1.5% agarose gel. Lane 1&2, samples from two independent preparations; M1, 100 bp DNA ladder.

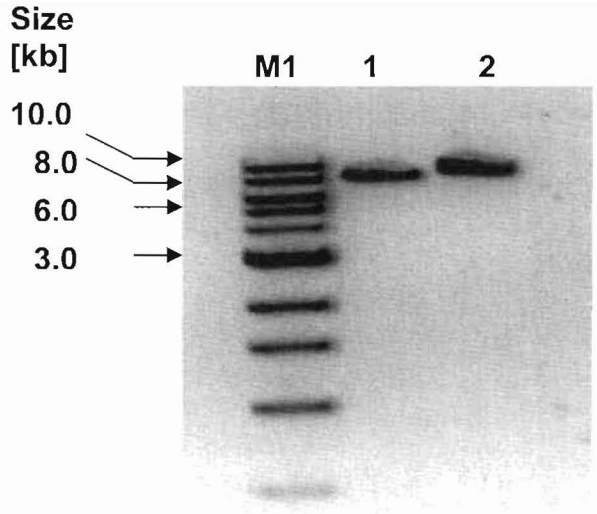


Figure 2-5 Double digestion of pHIS2

pHIS2 vector was digested with *Eco*R I and *Mlu* I, and analyzed on a 1% agarose gel. M1, 1 kb ladder; lane 1&2, digests from two independent preparations.

then amplified by PCR with primers GGCGAAAGGGGGATGTGCGT and TAGGGCTTTCTGCTCTGTCA. The PCR product was sequenced to ascertain the correct construct of FARE bait.

2.2.5 Optimization of *HIS3*-based FARE one-hybrid screens

The FARE bait vector was transformed into Y187 yeast cells according to the protocol in the manual of Matchmaker library construction and screening kits. Asmall amount of frozen yeast stock was streaked on a YPAD agar plate, and healthy colonies were picked for the transformation. Through steps of inoculation and incubation, the yeast cells reached mid-log-phase and were harvested by centrifugation. After washing with sterile water, the yeast cell pellets were made competent through treatment with TE/LiAc solution. FARE-bait plasmid and denatured Herring testes Carrier DNA were added to the competent yeast cells, mixed with PEG/LiAC solution and incubated for 30 min at 30 °C. After DMSO was added to promote the transformation, the mixture was incubated for 15 min at 42 °C. The yeast cells were pelleted and re-suspened in YPD plus medium (Clontech, Mountain View, CA). After 90 min incubation at 30 °C while shaking, the yeast cell pellets were re-suspended in 1 ml 0.9% NaCl; 100 μl of 1:10 and 1:100 diluted suspension were spread on a series SC/-Trp/-His plates containing different concentration of 3-AT.

The one-hybrid positive and negative control vectors were transformed into competent yeast cells according to the small-scale protocol in the manual. The yeast cell pellets from transformation procedure were re-suspended in 1 ml 0.9% NaCl, then 100 μl of undiluted and 1:10 diluted mixtures were spread on

SC/-Trp, Sc/-Leu, SC/-Trp/-Leu and a series SC/-Trp /-His /-Leu plates containing different concentration of 3-AT.

2.3 Results

2.3.1 cDNA library synthesis

The RNA isolation procedure resulted in total RNA at a concentration of 2.56 μg/µl, and an A260/A280 absorbance ratio of 1.8, indicating pure RNA. The total yield was about 75 μg/80 mg tissue. On a denaturing agarose gel, two predominant bands appeared at approximately 2.2 kb and 1.9 kb, representing the large and small ribosomal RNA subunits. There were also some bands present lower than 1.9 kb, probably comprised of mRNA and tRNA (Figure 2-6). The locust flight muscle cDNA library synthesized from this RNA preparation appeared as a smear from 0.1 kb to 3 kb on 1.2% agrose gel (Figure 2-7), with more intense bands between 0.5 kb to 2 kb, indicating a good representation of transcripts $\leq$2 kb.

2.3.2 Construction of FARE bait vector

The FARE-bait vector was constructed by directional cloning of the FARE target element into the pHIS2 reporter vector, and amplified in the DH5α cells. Plasmids from 10 transformant colonies were prepared and digested with *EcoR* I and *Mlu* I (Figure 2-8). Five plasmid constructs (Lane 4, 5, 7, 8, 9) that showed more than one band were discarded. The other 5 plasmid constructs (Lane 1, 2, 3, 6, 10) that showed single sharp bands of around 7 kb were further analyzed by PCR with primers annealing upstream and downstream of the multiple cloning

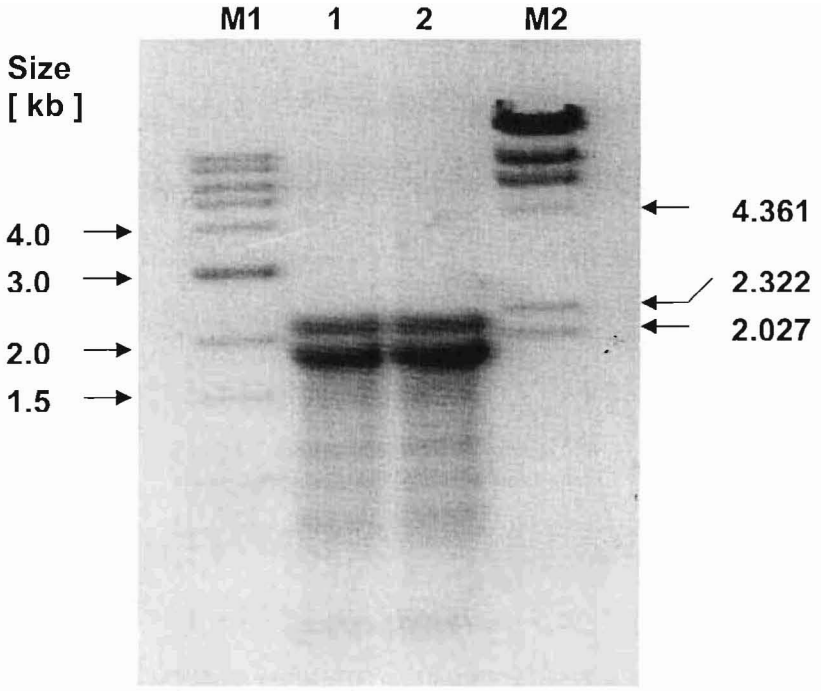


Figure 2-6 Total RNA from locust flight muscle

Total RNA was isolated from locust flight muscle and analyzed on a 1.2% denaturing agarose gel. Lane 1&2, RNA extracts from two independent preparations; M1, 1 kb ladder; M2, λDNA/*Hind* III fragments.

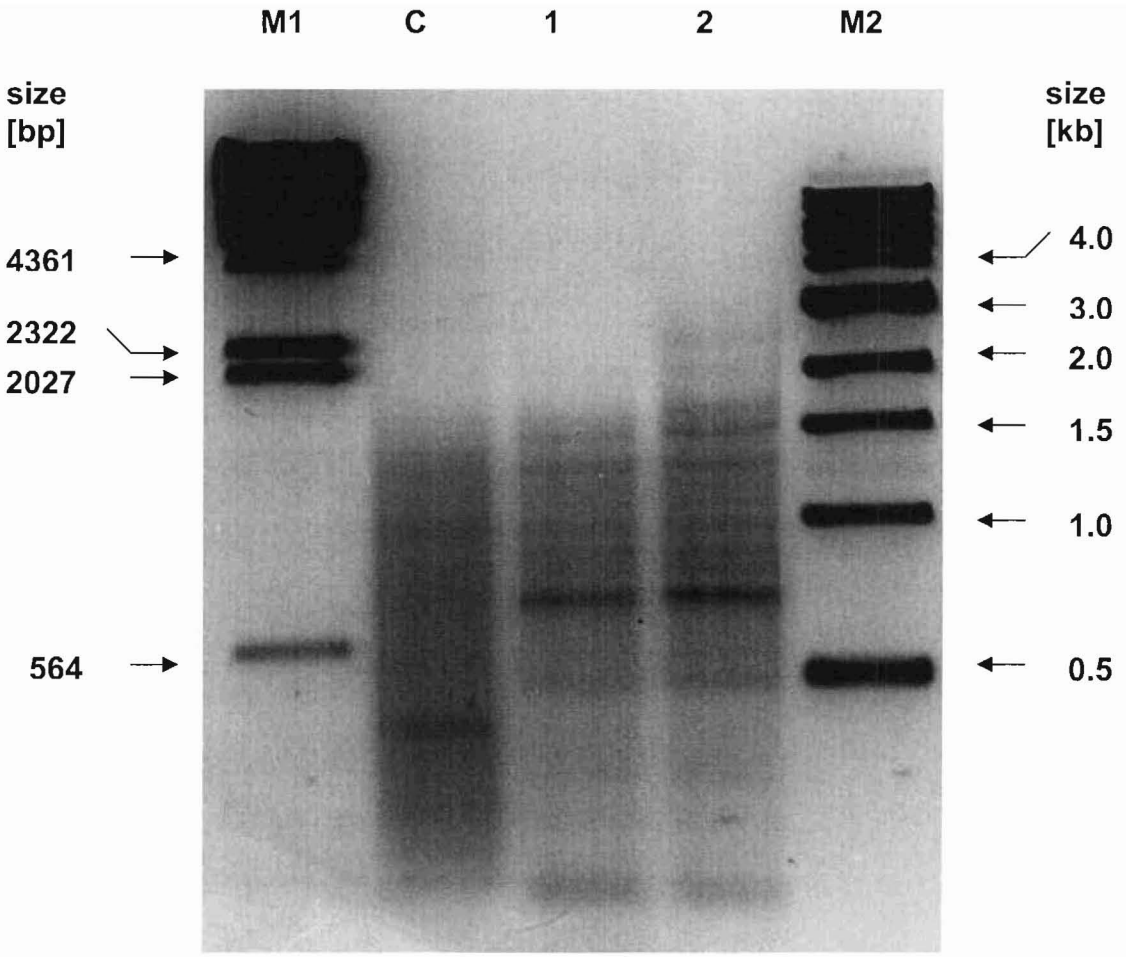


Figure 2-7 Analysis of the double-stranded cDNA

Double stranded cDNA was synthesized from total RNA extracted from locust flight muscle (Lane 1 and 2) and from human placenta Poly A+ RNA (control). The final product were analyzed on a 1.2 % agarose gel. Lane 1&2, two independent preparations of locust muscle cDNA; Lane C, control; Lane M1, λDNA/*Hin*d III fragments; M2, 1 kb DNA ladder.

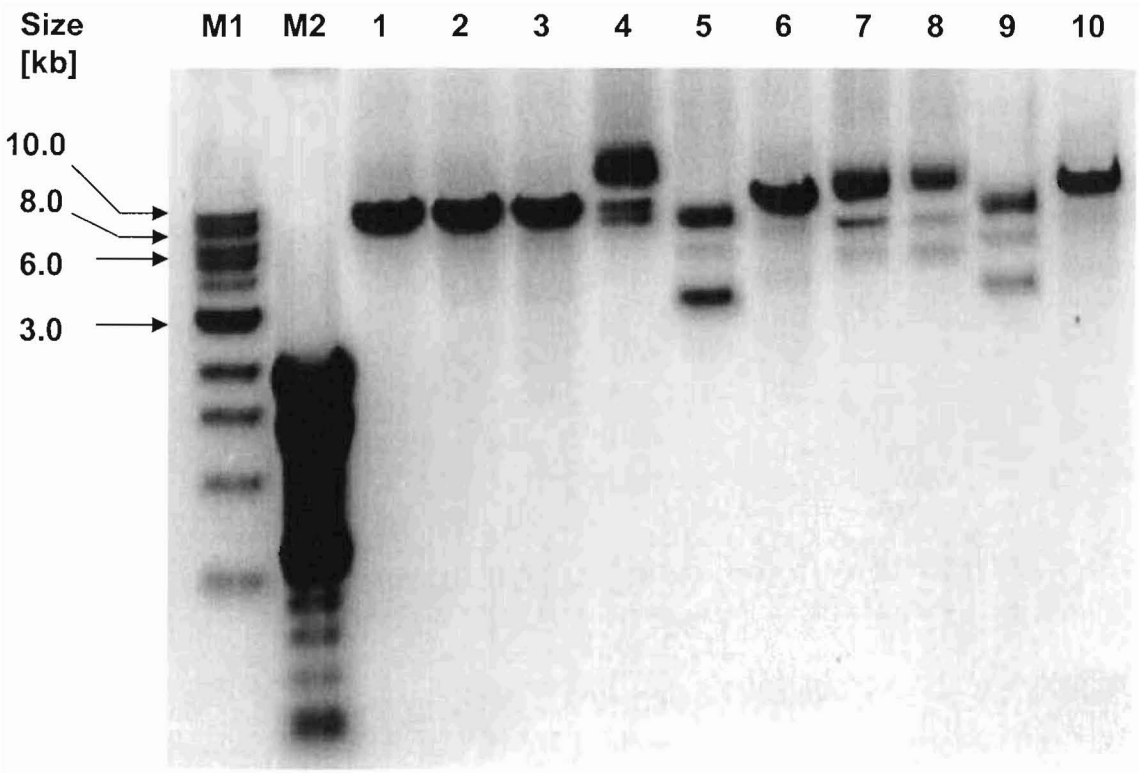


Figure 2-8 Analysis of FARE bait-transformed bacterial colonies

Plasmids prepared from 10 transformant colonies were digested with *Eco*R I and *Mlu* I, and the products were analyzed on a 1% agarose gel. Lane 1-10, digested plasmid DNA from 10 individual colonies; M1, 1 kb DNA ladder; M2, 100 bp DNA ladder.

site. Four out of five PCR products showed bands at the expected size of 330 bp (Figure 2-9), indicating the correct size of the insertion. The sequencing results confirmed the success of the FARE bait construction (Figure 2-10).

2.3.3 Optimization of *HIS3*-based FARE one-hybrid screens

The plasmid vectors, pGBT9, pHIS2, p53HIS2 and pFARE-HIS, were transformed into Y187 yeast cells. pGBT9 is the positive control vector for the transformation protocol. pHIS2 is the one-hybrid reporter vector without insertion of the target element. p53HIS2 is the one-hybrid positive control reporter vector containing tandem copies of consensus DNA binding site for p53 protein. pFARE-HIS2 is the bait vector containing FARE. After the selection plates were incubated at 30 °C for 3 to 7 days, colonies appearing were counted. The lowest concentration of 3-AT needed to allow only small colonies to grow is the optimal 3-AT concentration for one-hybrid cDNA library screening. The transformation efficiency, expressed as colony forming units (cfu) per µg DNA, can be calculated according to the following formula:

$$\frac{\text{cfu X dilution factor X total suspension volume [}\mu\text{l]}}{\text{volume plated [}\mu\text{l] X amount DNA used [}\mu\text{g]}} = \text{cfu/}\mu\text{g DNA}$$

Table 2-1 shows the number of colonies growing on each plate; 0.25 μg of pGBT9 and 0.5 μg of the other three plasmids were used for transformation. For the transformation positive control, the transformants were plated on SC/-Trp

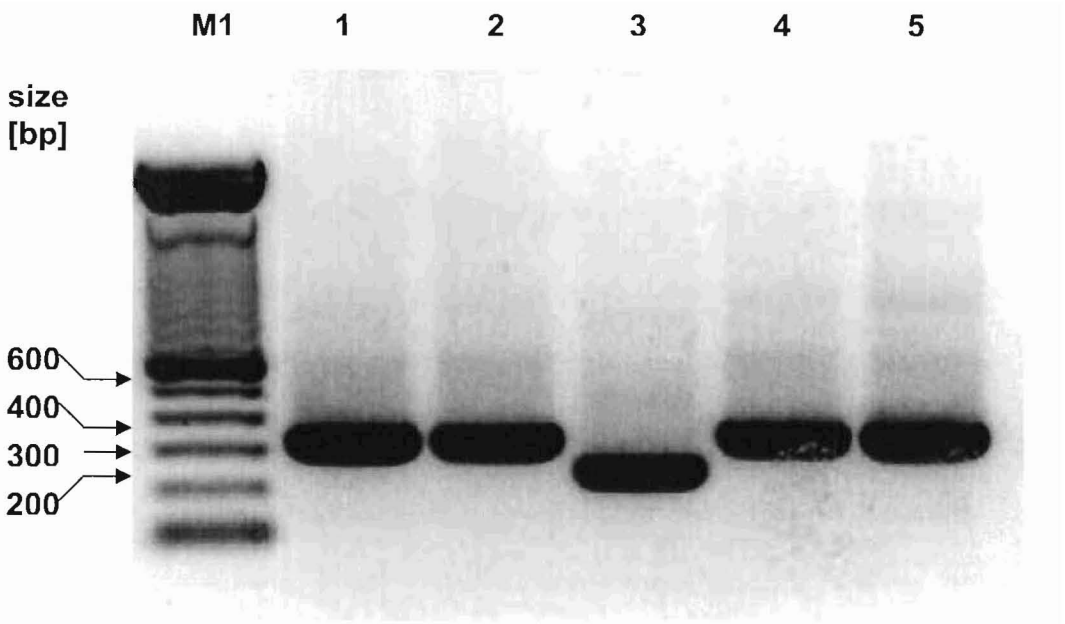


Figure 2-9 PCR of plasmids prepared from transformants of the FARE bait vector

Plasmids of five candidate transformant colonies decided by double digestion were used as templates for PCR with primers across the junctions of insertion. The PCR products were analyzed on a 1% agarose gel (Lane 1-5). M1 is the 100 bp DNA ladder.

Figure 2-10 Sequencing confirmation of the FARE bait construct

Table 2-1 Background HIS3 titration for vectors of pHIS2, p53HIS2 and pFARE-HIS2

Suspension cultures were diluted 1:10 or 1:100, as indicated, and spread on the selective media plates containing varying concentrations of 3-AT. The numbers of colonies appearing on each plate are listed. TMTC stands for too many to count.

	pGBT9		pHIS2		p53HIS2		pFARE-HIS2	
	1:10	1:100	1:10	1:100	1:10	1:100	1:10	1:100
SC/-Trp	TMTC	41	146	12	—	—	178	18
SC/-Trp/-His+3-AT								
0 mM	—	—	145	15	405	51	189	17
5 mM	—	—	—	—	47	4	—	—
10 mM	—	—	—	—	12	2(small)	121	9
15 mM	—	—	—	—	10	1(small)	—	—
20 mM	—	—	99	10	2(small)	1(small)	35	4
30mM	—	—	—	—	2(small)	0	19	2(small)
40 mM	—	—	32	2(small)	—	—	4	0
50mM	—	—	—	—	—	—	2	0
60mM	—	—	—	—	—	—	1	0
70mM	—	—	—	—	—	—	0	0

selection medium because the vector pGBT9 contains no HIS3 reporter gene but only TRP1 nutritional marker. The transformation efficiency for pGBT9 was 1.64 $x10^5$ cfu/μg DNA. The optimal 3-AT concentrations for pHIS2, p53HIS2 and pFARE-HIS2 vector were about 40 mM, 10 mM and 30 mM, respectively, which was concluded based on the colonies appearing on the 1:100 selection plates. The transformation efficiency for these three vectors were about 3 $x10^4$ cfu/μg DNA, 1 $x10^5$ cfu/μg DNA and 4 $x10^4$ cfu/μg DNA, respectively.

2.3.4 Yeast one-hybrid controls

The one-hybrid assay of one-hybrid positive and negative controls provides an insight on how the library screening works, and serves as a pilot experiment to confirm the whole procedure would work. Before screening the cDNA library, both one-hybrid controls were performed by co-transforming yeast with either one-hybrid positive vectors (pGAD-Rec2-53 prey vector and p53HIS2 bait vector) or one-hybrid negative vectors (pGAD-Rec2-53 prey vector and pHIS2 reporter vector). Colonies appearing on each plate were counted (Table 2-2). Colonies appeared on SC/-Trp plates and SC/-Leu plates due to the transformation of p53HIS and pGAD-Rec2-53 respectively, and the transformation efficiencies were about 2.5 $x10^4$ cfu/µg DNA and 9 x 10^3 cfu/µg DNA. Out of 530 clones from colonies grown on SC/-Trp/-Leu plates screened, 400 clones had been detected to have positive one-hybrid interaction which showed growth on the SC/-Trp/-His/-Leu + 3-AT plate. The number of detected positive clones remained stable with increasing 3-AT concentration, indicating that the one-hybrid interaction in positive clones is very strong and stable.

Table 2-2 One-hybrid controls

Suspension cultures of p53HIS2/ pGAD-Rec2-53 (positive control) and pHIS2/ pGAD-Rec2-53 (negative control) co-transformants were diluted 1:1 or 1:10, as indicated, and streaked on nutrient media plates containing varying concentrations of 3-AT. The numbers of colonies appearing on each plate are listed. TMTC stands for too many to count.

	p53HIS2 + pGAD-Rec2-53		**pHIS2 + pGAD-Rec2-53**	
	1:1	1:10	1:1	1:10
SD/-Trp	TMTC	127	—	—
SD/-Leu	TMTC	46	—	—
SD/-Trp/-Leu	53	2	—	—
SD/-Trp/-His/- Leu + 3-AT				
0mM	40	2	112	6
10mM	39	4	4(small)	0
20mM	40	4	—	—

After the yeast plasmids were isolated from positive clones and transformed into DH5α competent cells, the bacteria transfornants containing pGAD-Rec2-53 were selected on LB/Amp plates. The *Hin*d III digestion of pGAD-Rec2-53 isolated out of the selected transformants showed two bands as expected at the sizes of approximately 1.7 kb and 7 kb (Figure 2-11), which indicates the presence of prey vector in the yeast positive clones.

2.4 Discussion

The synthesis of a representative cDNA library depends primarily on the quality of the mRNA used, and it is essential to prevent degradation or contamination with DNA or protein. In my experiments, total RNA was first extracted with the TRIzol reagent method, subsequently purified with the RNeasy mini kit, and finally concentrated by sodium acetate precipitation. The resulting RNA preparation was of very high quality, as evident from various indicators.

The ration of A_{260}/A_{280} was 1.8, a value that should be expected for pure RNA. Two predominant bands on a denaturing agarose gel with an approximate size of 2.2 kb and 1.9 kb (Figure 2-6) represent the large and small subunit of the risbosomal RNA (rRNA). Unlike mammalian 28S and 18S, which correspond to sizes of approximately 4.5 kb and 1.9 kb, the ribosomal subunits in locust are less well characterized. However, the sizes of the ribosomal subunits have been found to vary from different sources. The extracted total RNA was used to synthesize the cDNA library (Figure 2-7) with BD SMART technology, which appeared to be of high quality as well. The BD SMART III and CDSIII anchors of the synthesized cDNA library are crucial for the homologous-recombination-

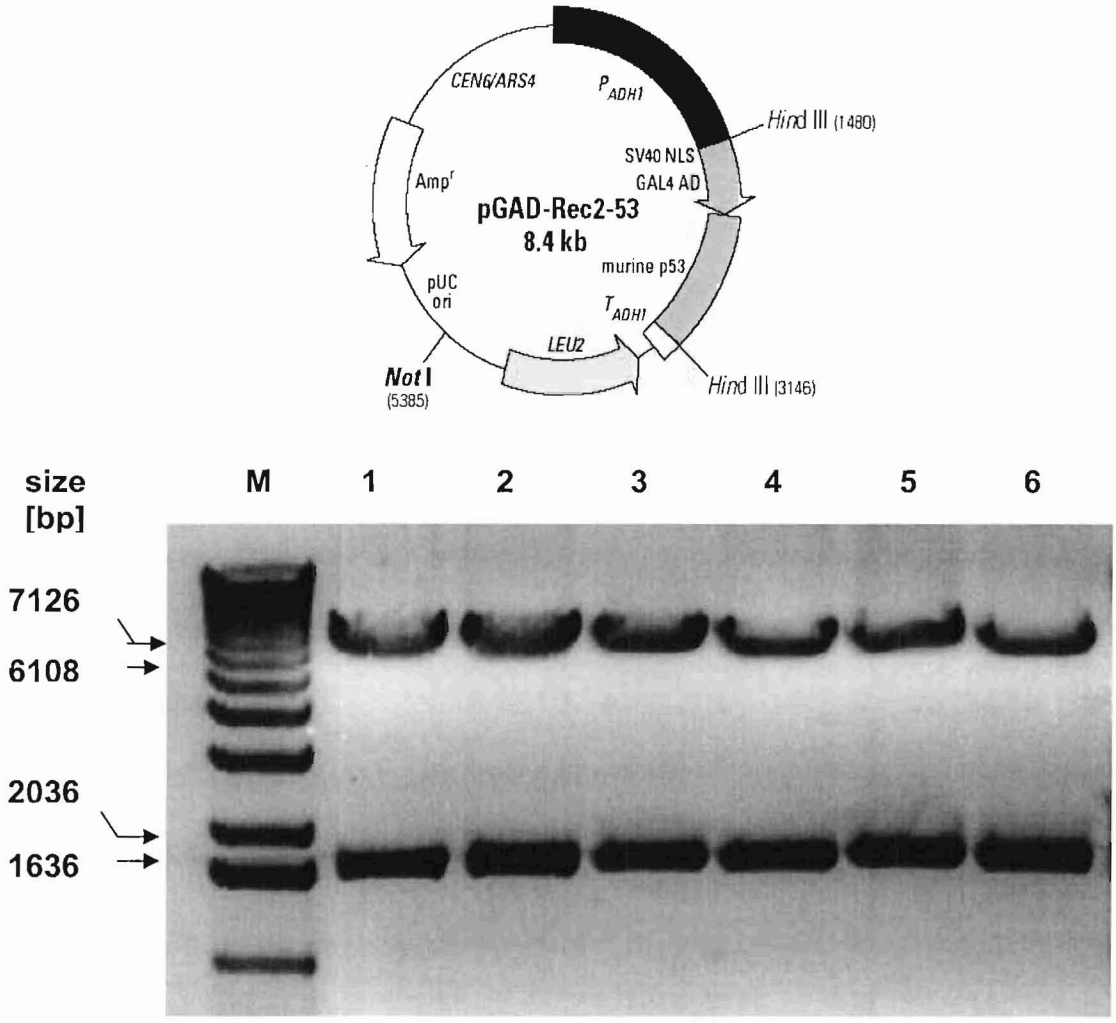


Figure 2-11 Map of pGAD-Rec2-53 AD vector and its digestion with *Hind* III

Plasmids isolated from six selected bacteria transformants were digested (Lane 1-6). M, 1 kb DNA ladder

Map of pGAD-Rec2-53 AD control vector

mediated cloning of cDNA into the AD-expressing vector and the construction a cDNA expression library, which can be screened in yeast for FARE-binding proteins in the presence of FARE-bait vector.

The FARE-bait vector contains three tandem copies of FARE upstream $P_{minHIS3}$ and the HIS3 nutritional reporter gene. The success of its construction was confirmed by sequencing (Figure 2-10). The bait vectors express low level of HIS3 in yeast due to the presence of T_C, so they must be transformed into yeast and tested for leaky HIS3 expression with 3-AT prior to the cDNA library screen. The leaky HIS3 titration is semi-quantitative, so the optimal 3-AT concentration may fluctuate and not always be fully reproducible. The transformation efficiency for the transformation positive control or for the bait vectors was considerably lower than the value suggested by the manufacturer (10^6 cfu/μg DNA).

Thus, it was important to test the one-hybrid methodology with one-hybrid controls before attempting to screen a cDNA library (Table 2-2). For the negative control, the growth of yeast transformants was suppressed with 10 mM 3-AT. For the positive controls, the growth was very stable with increasing 3-AT concentration, due to the strong interaction of p53 with its target DNA element. Approximately 75% of the screened positive clones were detected to have positive one-hybrid interaction, thus confirming that the methodology works, at least for very strong interactions as seen in p53 and its target element. However, the interactions between the FARE and its receptor may be much weaker, which could prove problematic for a library screening if high 3 AT concentrations suppress all cell growth. If the determination of the optimal 3-AT concentration is

not accurate enough, more false positives will be expected. It will be difficult to distinguish between leaky expression and real positives interaction. Moreover, the transformation efficiencies were relatively low (less than 10^6 cfu/μg DNA), which would cause additional problems in the screening of a cDNA library. Due to the low efficiency, considerably more colonies need to be screened before positive transformants are encountered. Given the difficulty to distinguish between true interactions and false positives, it may be very challenging to identify true interactions in such a library screening.

As discussed in Chapter 1, only two nuclear receptors, the ecdysone receptor (EcR) and the retinoic X receptor (RXR), have so far been found in locust species. EcR is a well-characterized receptor for the molting hormone ecdysone that binds to a steroid hormone and thus is unlikely to interact with fatty acids. Yet, it can be anticipated that a member of the nuclear receptor superfamily binds fatty acids and initiates the up-regulation of FABP expression. Moreover, in a heterodimeric complex, RXR will most likely be present. Extensive PCR screening with degenerate primers for nuclear receptors carried out in our laboratory failed to identify in locust muscle nuclear receptors other than RXR. Due to all the above considerations, it was decided to focus directly on the possible interaction between RXR and FARE.

CHAPTER 3:
CLONING OF *SCHISTOCERCA* RXR

3.1 Introduction

Since one-hybrid screening of a flight muscle cDNA library was deemed unlikely to succeed in identifying the receptor(s) binding to the fatty acid response element, it was decided to investigate the interaction between this element and a candidate receptor, RXR. RXR is known to be involved in many ligand activated gene regulation processes, binding various response elements either as homo- or as heterodimer. Prior to this thesis, RXR had never been detected in flight muscle of insects, nor had this factor been identified anywhere in the desert locust, *Schistocerca gregaria*. RXR, which is homologous to *Drosophila* ultraspiracle (USP), has been cloned, however, from a closely related locust species, the migratory locust *Locusta migratoria* (Hayward et al., 1999). *Drosophila* USP has been found to regulate the cellular response to the molting hormone ecdysone by the formation of a heterodimer with the ecdysone receptor (EcR) (Yao et al., 1992; Yao et al., 1993). The study of cellular responses to ecdysone in *Locusta migratoria* led to the discovery of a longer isoform of *Locusta* RXR, denoted as LmRXR-L (Genbank accession number AY348873). The shorter RXR isoform found earlier is referred to as LmRXR-S (Genbank accession number AF136372) (Hayward et al., 2003).

These two isoforms of LmRXR originate from a single gene, and the two transcripts of 4677 bp and 4752 bp, respectively, are the result of alternative

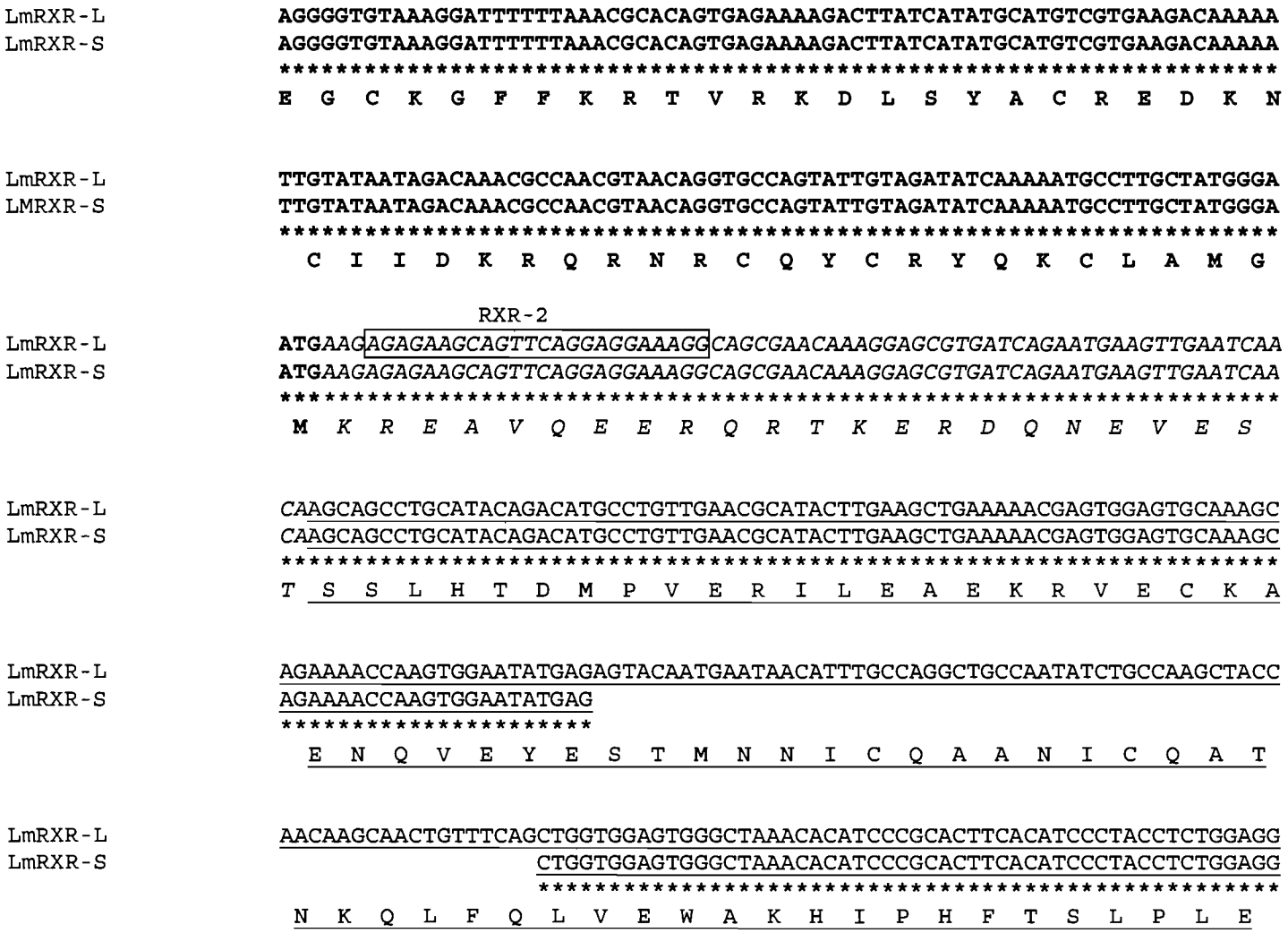
splicing (Hayward et al., 2003). The protein coding sequences are located between bp 107 and 1276 (1170 bp in size) or 107-1342 bp (1236 bp in size) of the transcripts, respectively. Gene alignment of the coding regions for the two isoforms shows essentially identical nucleotide sequences (two nucleotide difference at bp 48 and 1005 of LmRXR-L) except for a 66 bp deletion in the shorter isoform (between bp 619- 685 of LmRXR-L) (Figure 3-1). Their encoded proteins have the general nuclear receptor structure, defined by a variable domain (A/B), a DNA binding domain (DBD), a hinge region and a ligand binding domain (LBD). The only difference in the protein sequences is a 22 amino acids deletion in LmRXR-S after the first 30 amino acids of the LBD (Figure 3-1).

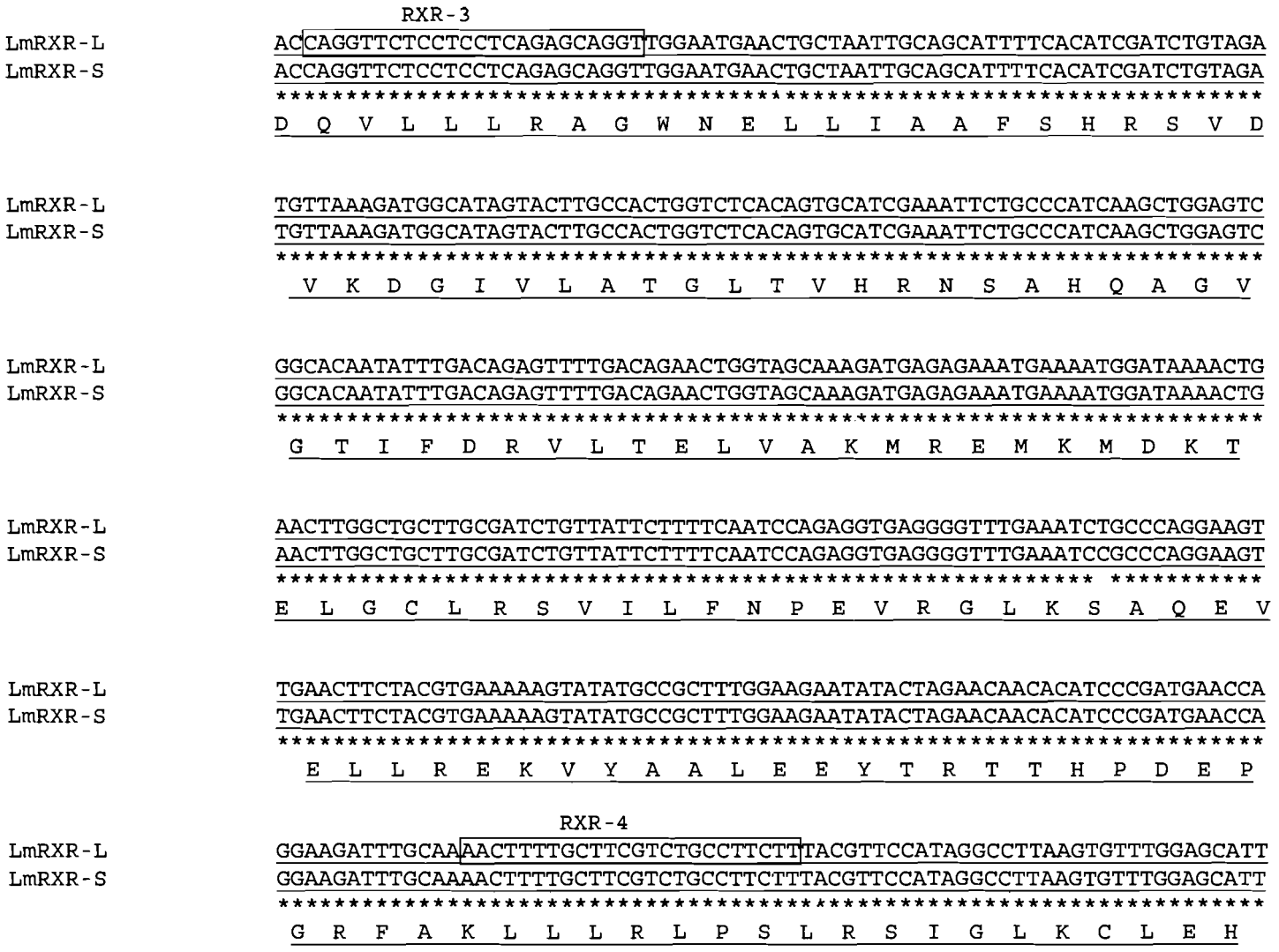
In order to study the proposed interaction between locust RXR and FARE in a one-hybrid assay, the expression of RXR in the flight muscle of *S. gregaria* must be confirmed and its cDNA sequence determined. Then, *S. gregaria* RXR (SgRXR) needs to be cloned into the AD-expressing vector and expressed in yeast cells as a GAL4 fusion protein. Because of the close phylogenic relationship of the two locust species and the highly conserved structure of RXR, RT-PCR with *S. gregaria* flight muscle RNA should be possible with nested primers designed based on the knowledge of LmRXRs. The full-length mRNA sequences of the two isoforms found in *L. migratoria* contain long 3' untranslated regions (3.4 kb), but for the purpose of this study, it is sufficient to clone mRNA up to the stop codon only. In order to clone SgRXR into the AD-expressing vector by homologous recombination, amplification of its cDNA requires the addition of

Figure 3-1 Sequence alignment of both LmRXRs

The beginning of the sequence is the variable domain; the bold is DBD; the italic is the hinge and the underlined is LBD. Primers are identified by a box.

```
                    aagcagtggtatcaacgcagagtggccattatggatggaaggaagtgaaagaggaataa
LmRXR-L                    Forward anchor primer     ATGGAAGGAAGTGAAAGAGGAATAAGTTTAGAGAAC
LmRXR-S                                              ATGGAAGGAAGTGAAAGAGGAATAAGTTTAGAGAAC
                                                     ************************************
                                                      M  E  G  S  E  R  G  I  S  L  E  N

LmRXR-L              AATCTTTCGATTAGCTCGATGGGTCCACAATCACCGTTGGATATGAAACCAGATACAGCTAGCCTAATCA
LmRXR-S              AATCTTTCGATAAGCTCGATGGGTCCACAATCACCGTTGGATATGAAACCAGATACAGCTAGCCTAATCA
                     *********** ************************************************************
                      N  L  S  I  S  S  M  G  P  Q  S  P  L  D  M  K  P  D  T  A  S  L  I

LmRXR-L              GCTCTGGCAGCTTTAGTCCCACTGGAGGACCTAACAGCCCAGGCTCTTTTACAATTGGACACAGTAGTCT
LmRXR-S              GCTCTGGCAGCTTTAGTCCCACTGGAGGACCTAACAGCCCAGGCTCTTTTACAATTGGACACAGTAGTCT
                     ************************************************************************
                     S  S  G  S  F  S  P  T  G  G  P  N  S  P  G  S  F  T  I  G  H  S  S  L

                                                                          RXR-1
LmRXR-L              CTTGAACAACTCCTCAAGCAACCAGGCAAAAGGCTCCTCATCACAGTATCCACCAAATCATCCACTGAGT
LmRXR-S              CTTGAACAACTCCTCAAGCAACCAGGCAAAAGGCTCCTCATCACAGTATCCACCAAATCATCCACTGAGT
                     ************************************************************************
                       L  N  N  S  S  S  N  Q  A  K  G  S  S  S  Q  Y  P  P  N  H  P  L  S

LmRXR-L              GGCTCTAAACATCTTTGTTCTATTTGTGGAGATCGAGCCAGTGGAAAGCACTATGGTGTTTACAGCTGTG
LmRXR-S              GGCTCTAAACATCTTTGTTCTATTTGTGGAGATCGAGCCAGTGGAAAGCACTATGGTGTTTACAGCTGTG
                     ************************************************************************
                      G  S  K  H  L  C  S  I  C  G  D  R  A  S  G  K  H  Y  G  V  Y  S  C
```


LmRXR-L
LmRXR-S
AGGGGTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTATCATATGCATGTCGTGAAGACAAAAA
AGGGGTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTATCATATGCATGTCGTGAAGACAAAAA
****
E G C K G F F K R T V R K D L S Y A C R E D K N
LmRXR-L
LMRXR-S
TTGTATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTTGCTATGGGA
TTGTATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTTGCTATGGGA
****
C I I D K R Q R N R C Q Y C R Y Q K C L A M G
RXR-2
LmRXR-L
LmRXR-S
ATGAAGAGAGAAGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAGCGTGATCAGAATGAAGTTGAATCAA
ATGAAGAGAGAAGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAGCGTGATCAGAATGAAGTTGAATCAA
****
M K R E A V Q E E R Q R T K E R D Q N E V E S
LmRXR-L
LmRXR-S
CAAGCAGCCTGCATACAGACATGCCTGTTGAACGCATACTTGAAGCTGAAAAACGAGTGGAGTGCAAAGC
CAAGCAGCCTGCATACAGACATGCCTGTTGAACGCATACTTGAAGCTGAAAAACGAGTGGAGTGCAAAGC
****
T S S L H T D M P V E R I L E A E K R V E C K A
LmRXR-L
LmRXR-S
AGAAAACCAAGTGGAATATGAGAGTACAATGAATAACATTTGCCAGGCTGCCAATATCTGCCAAGCTACC
AGAAAACCAAGTGGAATATGAG
****
E N Q V E Y E S T M N N I C Q A A N I C Q A T
LmRXR-L
LmRXR-S
AACAAGCAACTGTTTCAGCTGGTGGAGTGGGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGG
CTGGTGGAGTGGGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGG
****
N K Q L F Q L V E W A K H I P H F T S L P L E


RXR-3
LmRXR-L
LmRXR-S
ACCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAACTGCTAATTGCAGCATTTTCACATCGATCTGTAGA
ACCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAACTGCTAATTGCAGCATTTTCACATCGATCTGTAGA
D Q V L L L R A G W N E L L I A A F S H R S V D
TGTTAAAGATGGCATAGTACTTGCCACTGGTCTCACAGTGCATCGAAATTCTGCCCATCAAGCTGGAGTC
TGTTAAAGATGGCATAGTACTTGCCACTGGTCTCACAGTGCATCGAAATTCTGCCCATCAAGCTGGAGTC
V K D G I V L A T G L T V H R N S A H Q A G V
GGCACAATATTTGACAGAGTTTTGACAGAACTGGTAGCAAAGATGAGAGAAATGAAAATGGATAAAACTG
GGCACAATATTTGACAGAGTTTTGACAGAACTGGTAGCAAAGATGAGAGAAATGAAAATGGATAAAACTG
G T I F D R V L T E L V A K M R E M K M D K T
AACTTGGCTGCTTGCGATCTGTTATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCTGCCCAGGAAGT
AACTTGGCTGCTTGCGATCTGTTATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCCGCCCAGGAAGT
E L G C L R S V I L F N P E V R G L K S A Q E V
TGAACTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATACTAGAACAACACATCCCGATGAACCA
TGAACTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATACTAGAACAACACATCCCGATGAACCA
E L L R E K V Y A A L E E Y T R T T H P D E P
RXR-4
GGAAGATTTGCAAAACTTTTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATT
GGAAGATTTGCAAAACTTTTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATT
G R F A K L L L R L P S L R S I G L K C L E H

```
                                                                    gcttgaatcaccttc
LmRXR-L         TGTTTTTCTTTCGCCTTATTGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTC
LmRXR-S         TGTTTTTCTTTCGCCTTATTGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTC
                ************************************************************************
                L  F  F  F  R  L  I  G  D  V  P  I  D  T  F  L  M  E  M  L  E  S  P  S

                tgattcataacatgtcggccgcctcggcctctaga
LmRXR-L         TGATTCATAA   Reverse anchor primer
LmRXR-S         TGATTCATAA
                **********
                  D  S
```

a BD SMART III anchor and a CDSIII anchor for each dsDNA. Using gene specific primers that include the SMART III anchor in the forward primer and the CDS III anchor in the reverse primer (Figure 3-1), ds SgRXR can be first cloned into TA cloning vector for the purpose of gene sequencing and clone maintenance, and subsequently cloned into the AD-expressing vector in yeast cells.

The termini of the *Sma* I-linearized AD-expressing vector pGADT7-rec-2 are homologous to the BD SMART III and CDS III anchors, and the PCR-synthesized SgRXR amplicon has both anchors as well. Because cellular recombinases in yeast use dsDNA to repair the gap in pGADT7-rec-2, the homologous recombination-mediated cloning of SgRXR into AD-expressing vector occurs after the PCR amplicon and linearized pGADT7-rec-2 are co-transformed into yeast cells (Figure 3-2). As a yeast vector, pGADT7-rec-2 can be propagated and selected in both yeast and bacteria. The presence of *CEN6/ARS4* cassette ensures the stable and low-copy propagation of this YCp vector in yeast cells, so fewer false positive will be generated during the detection of one-hybrid interaction. The vector also contains a PUC origin of replication (PUC ori) for proper propagation in *E. coli*. For the vector selection in the host cells, the *LEU2* nutritional marker is used in yeast while the ampicillin resistance gene (Amp^r) is used in *E. coli* (Figure 3-3).

In vivo cloning of SgRXR into the pGADT7-rec-2 vector results in the expression of RXR fused with an upstream GAL4 activation domain. The high-

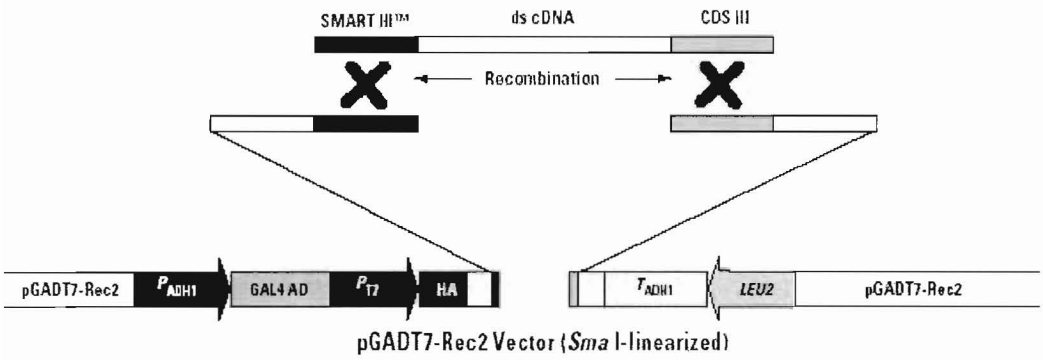


Figure 3-2 Cloning of cDNA into pGADT7-Rec2 by homologous recombination

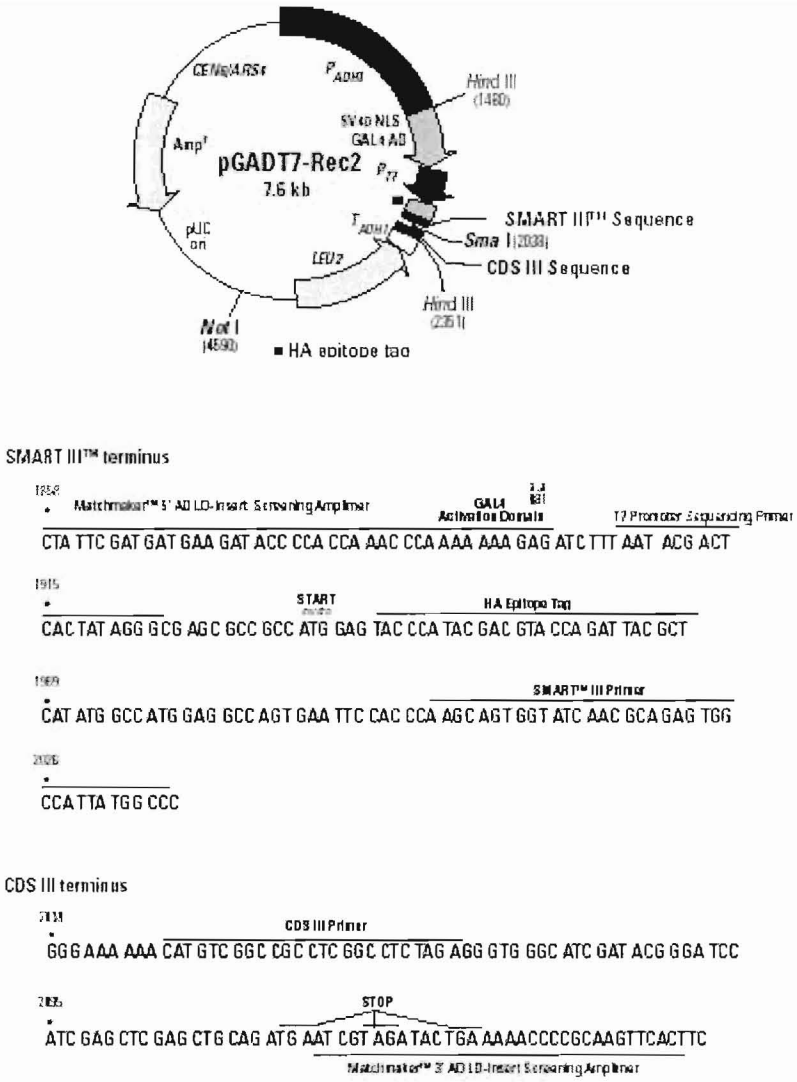


Figure 3-3 Map of pGADT7-Rec2 vector

level expression of the fusion protein is controlled by the constitutive full-length yeast ADH1 (alcohol dehydrogenase) promoter (P_{ADH1}) and ADH1 termination signal (T_{ADH1}) (Figure 3-3). The proteins expressed from insert sequences are downstream of the GAL4 AD polypeptide containing the incorporated nuclear localization signal (NLS) from simian virus 40 large tumor antigen (SV40 T), which functions to facilitate translocation of the fusion to the yeast nucleus (Chien et al., 1991). The fusion proteins also have a hemagglutinin (HA) epitope tag incorporated into the *GAL4* AD coding sequence, which allows the detection of the fusion proteins in yeast with antibodies to HA. Successful construction of the SgRXR prey vector and the correct expression of the SgRXR fusion protein are important prerequisites for the detection of interactions with the bait FARE sequences.

To verify specific interaction with FARE, bait vectors containing the original FARE sequence and partially altered sequences must be constructed and tested for their interaction with SgRXR. The design of the mutated bait sequences (MBSs) is based on its inverted repeat structure, composed of two imperfectly palindromic hexanucleotide half-sites interspaced by 3 nucleotides (IR-3). Four MBSs were designed by changing three nucleotides of within one of the half sites, both of the half sites, as well as five nucleotides on extending in both directions beyond the half sites. The sequences of the FARE and the mutated baits are illustrated in Table 3-1.

Table 3-1 Sequences of baits and nucleotides for bait construction

Bait vector	Sequences of Baits	Sequences of Target DNA Element
FARE	GG AGT GGT agt TCC CAT CC	5′-AATTC**GGAGTGGTagtTCCCATCC**GGAGTGGTagtTCCCATCC**GGAGTGGTagtTCCCATCC**A-3′ 3′-G**CCTCACCAtcaAGGGTAGG**CCTCACCAtcaAGGGTAGG**CCTCACCAtcaAGGGTAGG**TGCGC-5′
MB-1	GG AGT CAA agt TCC CAT CC	5′-AATTC**GGAGTCAAagtTCCCATCC**GGAGTCAAagtTCCCATCC**GGAGTCAAagtTCCCATCC**A-3′ 3′-G**CCTCAGTTtcaAGGGTAGG**CCTCAGTTtcaAGGGTAGG**CCTCAGTTtcaAGGGTAGG**TGCGC-5′
MB-2	GG AGT GGT agt AAG CAT C	5′-AATTC**GGAGTGGTagtAAGCATCC**GGAGTGGTagtAAGCATCC**GGAGTGGTagtAAGCATCC**A-3′ 3′-G**CCTCACCAtcaTTCGTAGG**CCTCACCAtcaTTCGTAGG**CCTCACCAtcaTTCGTAGG**TGCGC-5′
MB-3	GG AGT CAA agt AAG CAT CC	5′-AATTC**GGAGTCAAagtAAGCATCC**GGAGTCAAagtAAGCATCC**GGAGTCAAagtAAGCATCC**A-3′ 3′-G**CCTCAGTTtcaTTCGTAGG**CCTCAGTTtcaTTCGTAGG**CCTCAGTTtcaTTCGTAGG**TGCGC-5′
MB-4	AA TTT GGT agt TCC CTA TT	5′-AATTC**AATTTGGTagtTCCCTATT**AATTTGGTagtTCCCTATT**AATTTGGTagtTCCCTATT**G-3′ 3′-G**TTAAACCAtcaAGGGATAA**TTAAACCAtcaAGGGATAA**TTAAACCAtcaAGGGATAA**TGCGC-5′

3.2 Materials and methods

3.2.1 Identification of RXR and FABP in desert locust flight muscle

Four primers for the detection of SgRXR were designed according to the coding sequence of LmRXRs, and the sequences of four primers specific to SgFABP were determined based on the published sequence (Wu et al., 2001) (Figure 3-1and Table 3-2). Four possible primer combinations were used for RT-PCR reactions with total RNA extracted from the desert locust flight muscle. The reactions were performed with Ready-To-Go™ RT-PCR Beads (Amersham Biosciences, Piscataway, NJ) based on manufacturer's protocol, and PCR products were analyzed on 1% agarose gel.

3.2.2 Cloning of SgRXRs

The first-strand cDNA was synthesized from total RNA with an anchored-oligo $(dT)_{18}$ primer according to the protocol of Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Sciences, Laval, PQ), and then used directly for subsequent PCR with forward anchor primer (AAGCAGTG GTATCAACGCAGA GTGGCCATTATGGATGGAAGGAAGTGAAA GAGGAATAA) and reverse anchor primer (TCTAGAGGCCGAGGCGGCCGACATGTTATGAATCAGAAGG TGATTCAAGC). The PCR product was analyzed on a 1% agarose gel, gel-purified, and cloned into the linearized TA cloning vector pCR2.1-TOPO (Figure 3-4). The vector construct pCR2.1-SgRXR was transformed into DH5α, and grown on LB/Amp plates. Plasmid DNA was isolated from individual colonies and used for PCR with primers RXR-2/RXR-3 to confirm the presence of two

Table 3-2 Sequence of primers specific to locust RXR and FABP

Primers	Specific to RXR	Specific to FABP
Forward	RXR-1	FABP-1
	CACAGTATCCACCAAATCATCCAC	AGCTCGACTCGCAGACCAATTTTGAGGAATA
Reverse	RXR-4	FABP-4
	AAGAAGGCAGACGAAGCAAAAGTT	TATTCTCGTTGCCACCAGGTCG
Nested Forward	RXR-2	FABP-2
	AGAGAAGCAGTTCAGGAGGAAAGG	TCGAGCGGAAGGCAGGTC
Nested Reverse	RXR-3	FABP-3
	ACCTGCTCTGAGGAGGAGAACCTG	GATGATGGTGGGGTGGTC

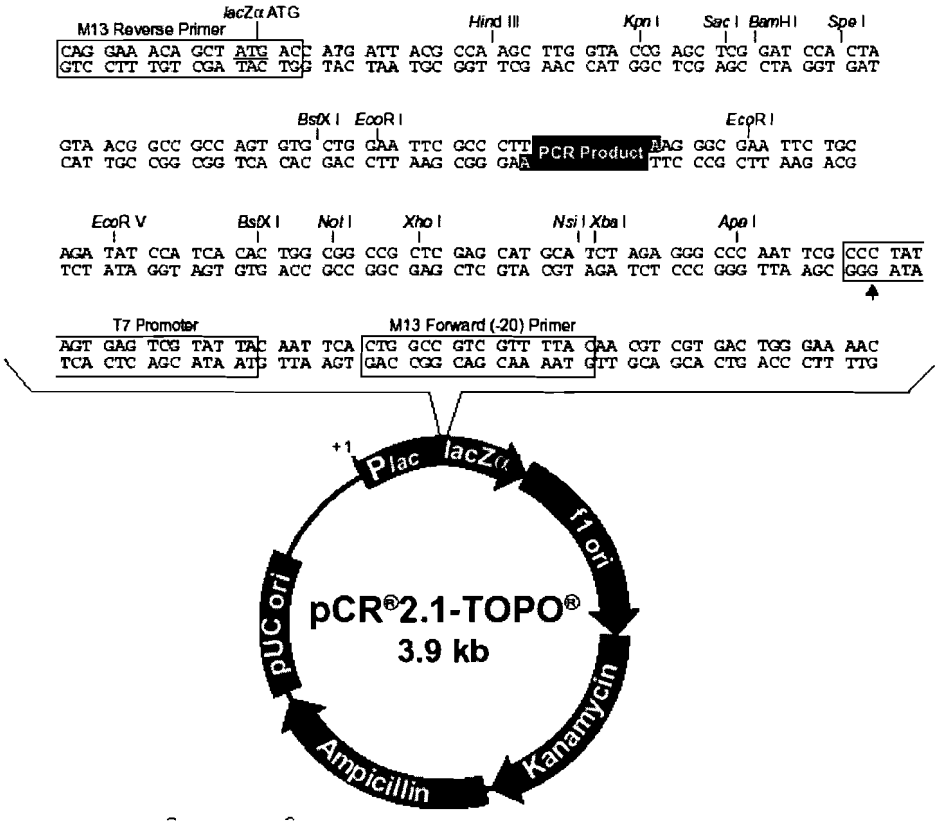


Figure 3-4 Map of pCR2.1-TOPO

isoforms, and sequenced with primers of M13R and T7 promoter to ascertain the correct amplification of the two isoforms of SgRXR.

3.2.3 RXR prey vector construction

The vector construct pCR2.1-SgRXR for either isomer was used for PCR with the forward anchor primer and reverse anchor primer. The PCR product was transformed into Y187 yeast strain together with *Sma* I-linearized AD-expressing vector pGADT7-rec-2. The yeast transformants containing SgRXR prey vector were selected on SC/-Leu plate. The plasmid vector was isolated and transformed into DH5α. The SgRXR prey vectors were digested with *Hind* III or *Not* I/*Nsi* I to check the orientation of the insert, and sequenced to confirm the correct construction of the SgRXR prey vector.

3.2.4 Construction of mutated bait vectors

The method used to construct the mutated bait vectors was the same as that used in FARE bait vector construction (see Chapter 2), involving the synthesis of mutated target elements and their cloning into the digested pHIS2 reporter vector. Two complementary oligonucleotides were designed for each of the four mutated target elements (Table 3-1), and annealed together to form ds mutated target element which was then ligated into *Eco*R I/*Mlu* I digested pHIS2 vector. The ligation products were transformed into DH5α competent cells, and the transformants containing the mutated bait vectors were selected on the LB/Kan agar medium. The mutated bait vectors were prepared from the isolated colonies, analyzed by double digestion with *Eco*R I/*Mlu* I, and amplified by PCR

with primers GGCGAAA GGGGG ATGTGCGT and TAGGGCTTTCTGCTCTGT CA, as well as sequenced to ascertain the correct construct of the mutated bait vectors.

3.3 Results

In RT-PCR with each primer combination for FABP, one discrete band of the expected size was observed: about 290 bp with FABP-1/FABP-3, 360 bp with FABP-1/FABP-4, 230 bp with FABP-2/FABP-3 and 300 bp with FABP-2/FABP-4 (Figure 3-5). In contrast, neither primer combination for RXR resulted in a PCR product at the optimal annealing temperatures. However, re-amplification of the RXR-1/RXR-4 RT-PCR product with the nested primers RXR-2 and RXR-3 showed two separate bands of approximately 230 bp and 300 bp (Figure 3-6). The two bands were eluted and sequenced, showing out three nucleotide difference compared to the sequence of corresponding region of LmRXRs.

The PCR products obtained with anchor primers showed two bands on agarose gel of around 1.2 kb (Figure 3-7). After transforming bacteria with these products, six colonies were picked from LB/Amp plates. PCR of isolated plasmids with primers LRXR-2/LRXR-3 showed that two of the transformants contained SgRXR-L sequence, three SgRXR-S sequence, and one both isoforms (Figure 3-8). Exclude the 66bp nucleotide difference, the coding sequences between two isoforms of SgRXR showed 99.66% nucleotide identity (99.83% between LmRXRs). The coding sequences of SgRXRs are highly similar to those of the long and short isoforms of LmRXR both at the RNA level (94.98% and 94.70% nucleotide identity, respectively) and at the protein level, where only a single

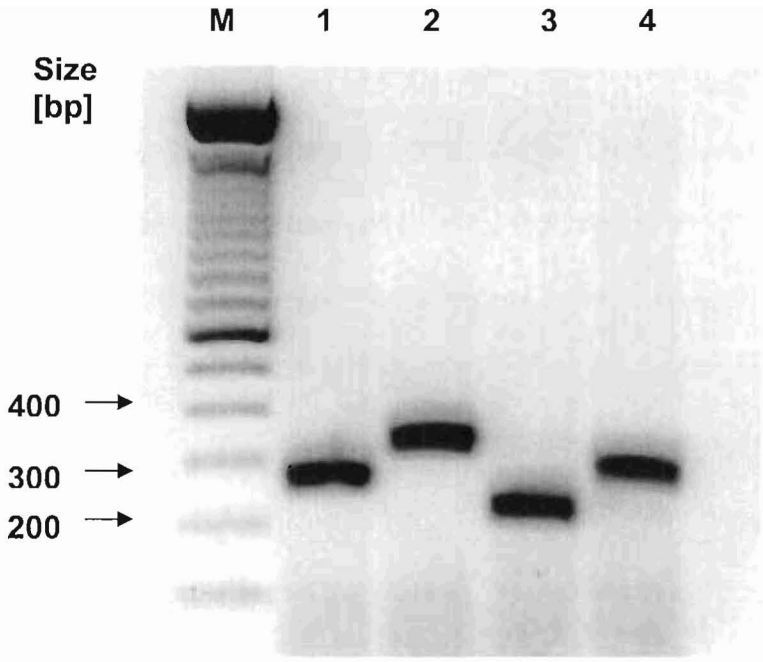


Figure 3-5 RT-PCR with primers specific to FABP

Total RNA from locust flight muscle was used for RT-PCR with primers FABP-1/FABP-3 (Lane 1), FABP-1/FABP-4 (Lane 2), FABP-2/FABP-3 (Lane 3) and FABP-2/FABP-4 (Lane 4). The PCR products were analyzed on 1% agarose gel. M, 100 bp DNA ladder

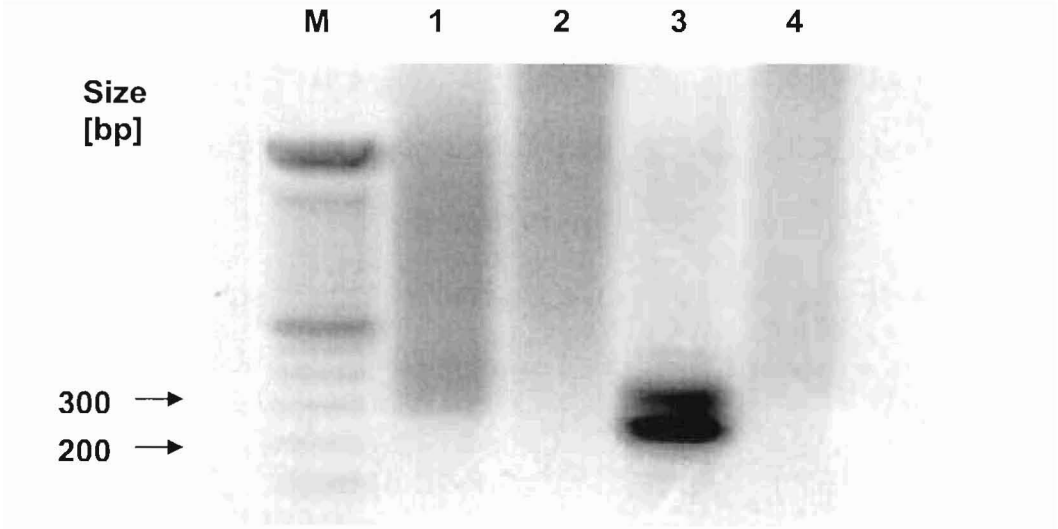


Figure 3-6 Re-amplification with nested primers specific to RXR

Re-amplifications with nested primers RXR-2/RXR-3 were performed from RT-PCR products of four primer combinations (RXR 1/RXR-3 for lane 1, RXR-2/RXR-3 for lane 2, RXR-1/RXR-4 for lane 3 and RXR-2/RXR-4 for lane 4). M, 100 bp DNA ladder.

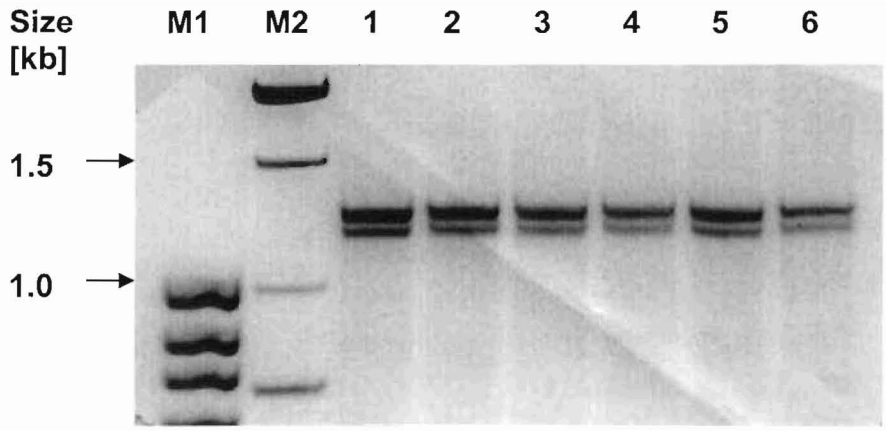


Figure 3-7 PCR product of SgRXR coding sequence

PCR products with primers containing sequences of both anchors were analyzed on 1.2% agarose gel (Lane 1-6). M1, 100 bp DNA ladder; M2, 1 kb DNA ladder.

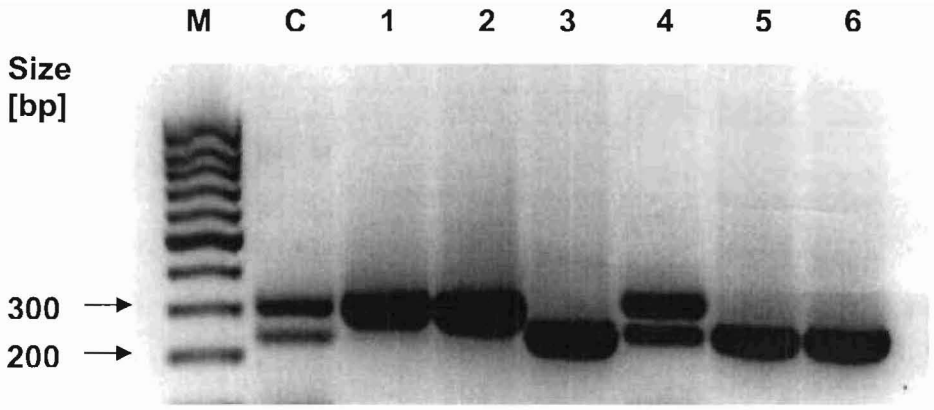


Figure 3-8 PCRs of isolated plasmids with primers RXR-2/RXR-3

Plasmids isolated from six colonies of bacterial transformants were used for PCR (Lane 1-6). PCR of the first-strand cDNA with primers RXR-2/RXR-3 serves as positive control (Lane C). M, 100 bp DNA ladder.

amino acid deletion (serine) in both isoforms and a single substitution is observed (leucine replaced by proline) in the short isoform compared to LmRXRs (Figure 3-9).

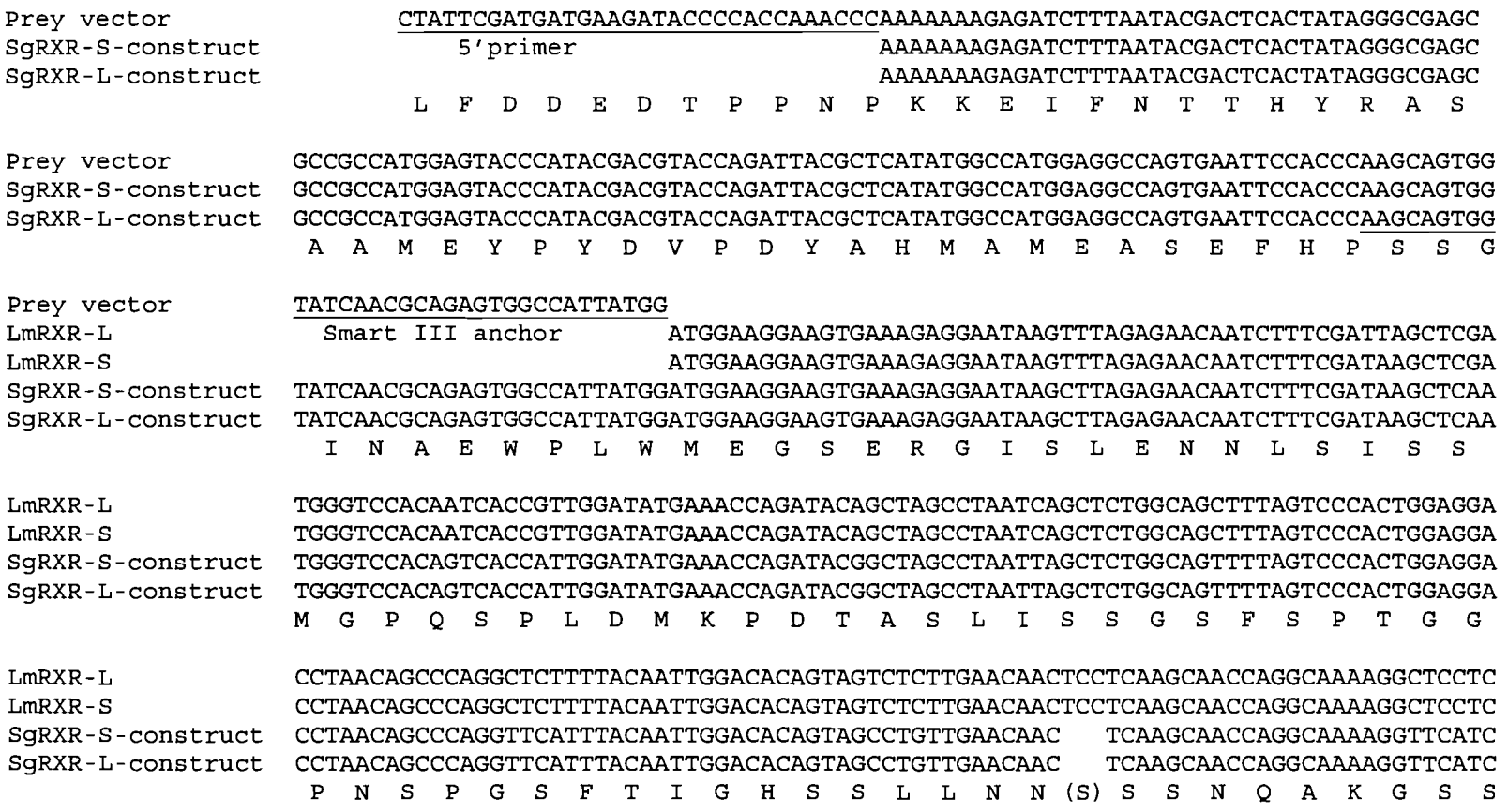
The yeast transformants growing on SC/-Leu medium contain the cloned SgRXR prey vector. The *Hin*d III digest of the SgRXR prey vector showed three bands of around 7.0 kb, 1.5 kb and 0.6 kb, respectively. The *Not* I/*Nsi* I double-digest showed two bands at the sizes of about 3.5 kb and 5.5 kb (Figure 3-10). The sequencing of the prey vector confirmed the correct insertion of SgRXRs into linearized pGADT7-rec-2 by in vivo cloning. Another single substitution of amino acid (glycine replaced aspartic acid) was observed in the short isoform of SgRXR (Figure 3-9).

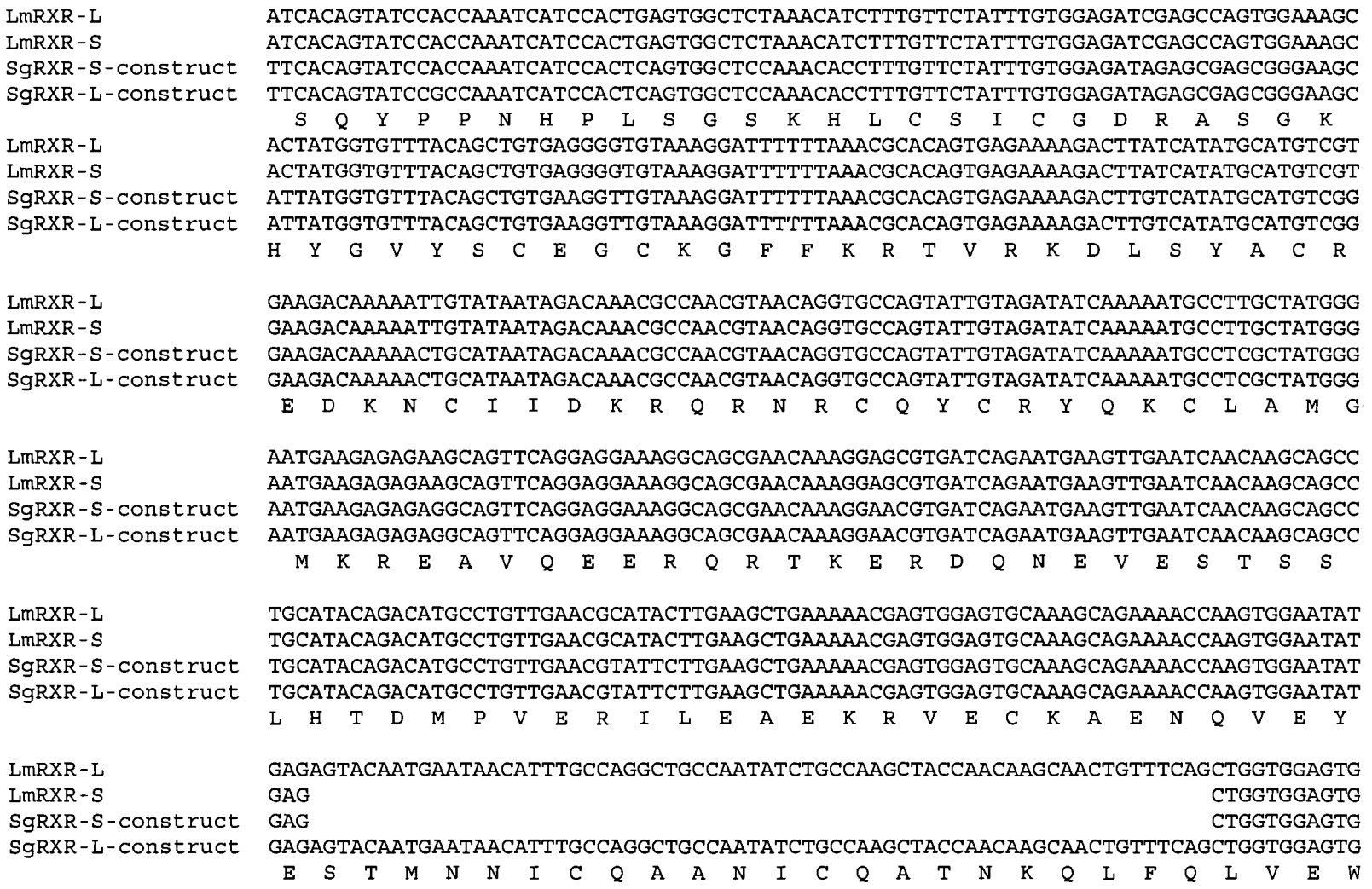
Double digestion of each mutated bait vector showed a single band of around 7 kb (Figure 3-11). Their PCR products with primers annealing upstream and downstream of the multiple cloning site showed bands of 330 bp (Figure 3-12), identical in size to the PCR product of the wild-type FARE bait vector. Sequence alignment of the different constructs confirms the successful construction of each mutated bait vector.

3.4 Discussion

If one or both isoforms of RXRs involve in FABP gene regulation as we proposed, the protein(s) must be expressed in locust flight muscle. For the first time, we were able to demonstrate the presence of RXR in flight muscle of an insect species. RT-PCR gave two distinct amplicons that are virtually identical to

Figure 3-9 Sequence alignment of SgRXR inserts in the prey vectors




LmRXR-L ATCACAGTATCCACCAAATCATCCACTGAGTGGCTCTAAACATCTTTGTTCTATTTGTGGAGATCGAGCCAGTGGAAAGC
LmRXR-S ATCACAGTATCCACCAAATCATCCACTGAGTGGCTCTAAACATCTTTGTTCTATTTGTGGAGATCGAGCCAGTGGAAAGC
SgRXR-S-construct TTCACAGTATCCACCAAATCATCCACTCAGTGGCTCCAAACACCTTTGTTCTATTTGTGGAGATAGAGCGAGCGGGAAGC
SgRXR-L-construct TTCACAGTATCCGCCAAATCATCCACTCAGTGGCTCCAAACACCTTTGTTCTATTTGTGGAGATAGAGCGAGCGGGAAGC
S Q Y P P N H P L S G S K H L C S I C G D R A S G K
LmRXR-L ACTATGGTGTTTACAGCTGTGAGGGGTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTATCATATGCATGTCGT
LmRXR-S ACTATGGTGTTTACAGCTGTGAGGGGTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTATCATATGCATGTCGT
SgRXR-S-construct ATTATGGTGTTTACAGCTGTGAAGGTTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTGTCATATGCATGTCGG
SgRXR-L-construct ATTATGGTGTTTACAGCTGTGAAGGTTGTAAAGGATTTTTTAAACGCACAGTGAGAAAAGACTTGTCATATGCATGTCGG
H Y G V Y S C E G C K G F F K R T V R K D L S Y A C R
LmRXR-L GAAGACAAAAATTGTATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTTGCTATGGG
LmRXR-S GAAGACAAAAATTGTATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTTGCTATGGG
SgRXR-S-construct GAAGACAAAAACTGCATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTCGCTATGGG
SgRXR-L-construct GAAGACAAAAACTGCATAATAGACAAACGCCAACGTAACAGGTGCCAGTATTGTAGATATCAAAAATGCCTCGCTATGGG
E D K N C I I D K R Q R N R C Q Y C R Y Q K C L A M G
LmRXR-L AATGAAGAGAGAAGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAGCGTGATCAGAATGAAGTTGAATCAACAAGCAGCC
LmRXR-S AATGAAGAGAGAAGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAGCGTGATCAGAATGAAGTTGAATCAACAAGCAGCC
SgRXR-S-construct AATGAAGAGAGAGGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAACGTGATCAGAATGAAGTTGAATCAACAAGCAGCC
SgRXR-L-construct AATGAAGAGAGAGGCAGTTCAGGAGGAAAGGCAGCGAACAAAGGAACGTGATCAGAATGAAGTTGAATCAACAAGCAGCC
M K R E A V Q E E R Q R T K E R D Q N E V E S T S S
LmRXR-L TGCATACAGACATGCCTGTTGAACGCATACTTGAAGCTGAAAAACGAGTGGAGTGCAAAGCAGAAAACCAAGTGGAATAT
LmRXR-S TGCATACAGACATGCCTGTTGAACGCATACTTGAAGCTGAAAAACGAGTGGAGTGCAAAGCAGAAAACCAAGTGGAATAT
SgRXR-S-construct TGCATACAGACATGCCTGTTGAACGTATTCTTGAAGCTGAAAAACGAGTGGAGTGCAAAGCAGAAAACCAAGTGGAATAT
SgRXR-L-construct TGCATACAGACATGCCTGTTGAACGTATTCTTGAAGCTGAAAAACGAGTGGAGTGCAAAGCAGAAAACCAAGTGGAATAT
L H T D M P V E R I L E A E K R V E C K A E N Q V E Y
LmRXR-L GAGAGTACAATGAATAACATTTGCCAGGCTGCCAATATCTGCCAAGCTACCAACAAGCAACTGTTTCAGCTGGTGGAGTG
LmRXR-S GAG CTGGTGGAGTG
SgRXR-S-construct GAG CTGGTGGAGTG
SgRXR-L-construct GAGAGTACAATGAATAACATTTGCCAGGCTGCCAATATCTGCCAAGCTACCAACAAGCAACTGTTTCAGCTGGTGGAGTG
E S T M N N I C Q A A N I C Q A T N K Q L F Q L V E W

```
LmRXR-L              GGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGGACCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAACTGC
LmRXR-S              GGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGGACCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAACTGC
SgRXR-S-construct    GGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGGGCCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAGCTGC
SgRXR-L-construct    GGCTAAACACATCCCGCACTTCACATCCCTACCTCTGGAGGACCAGGTTCTCCTCCTCAGAGCAGGTTGGAATGAGCTGC
                       A  K  H  I  P  H  F  T  S  L  P  L  E  D  Q  V  L  L  L  R  A  G  W  N  E  L
                                                                (G)-s

LmRXR-L              TAATTGCAGCATTTTCACATCGATCTGTAGATGTTAAAGATGGCATAGTACTTGCCACTGGTCTCACAGTGCATCGAAAT
LmRXR-S              TAATTGCAGCATTTTCACATCGATCTGTAGATGTTAAAGATGGCATAGTACTTGCCACTGGTCTCACAGTGCATCGAAAT
SgRXR-S-construct    TAATTGCAGCATTCTCACATCGATCTGTAGATGTTAAAGATGGTATAGTACCTGCCACTGGACTCACAGTGCATCGAAAT
SgRXR-L-construct    TAATTGCAGCATTCTCACATCGATCTGTAGATGTTAAAGATGGTATAGTACTTGCCACTGGACTCACAGTGCATCGAAAT
                     L  I  A  A  F  S  H  R  S  V  D  V  K  D  G  I  V  L  A  T  G  L  T  V  H  R  N
                                                                        (P)-s
LmRXR-L              TCTGCCCATCAAGCTGGAGTCGGCACAATATTTGACAGAGTTTTGACAGAACTGGTAGCAAAGATGAGAGAAATGAAAAT
LmRXR-S              TCTGCCCATCAAGCTGGAGTCGGCACAATATTTGACAGAGTTTTGACAGAACTGGTAGCAAAGATGAGAGAAATGAAAAT
SgRXR-S-construct    TCTGCCCACCAAGCGGGAGTGGGCACAATATTTGACCGTGTTTTGACAGAATTGGTGGCAAAGATGAGAGAAATGAAAAT
SgRXR-L-construct    TCTGCCCATCAAGCGGGAGTGGGCACAATATTTGACCGTGTTTTGACAGAATTGGTGGCAAAGATGAGAGAAATGAAAAT
                      S  A  H  Q  A  G  V  G  T  I  F  D  R  V  L  T  E  L  V  A  K  M  R  E  M  K  M

LmRXR-L              GGATAAAACTGAACTTGGCTGCTTGCGATCTGTTATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCTGCCCAGGAAG
LmRXR-S              GGATAAAACTGAACTTGGCTGCTTGCGATCTGTTATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCCGCCCAGGAAG
SgRXR-S-construct    GGATAAAACAGAACTTGGCTGCTTGAGATCTGTAATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCTGCCCAGGAGG
SgRXR-L-construct    GGATAAAACAGAACTTGGCTGCTTGAGATCTGTAATTCTTTTCAATCCAGAGGTGAGGGGTTTGAAATCTGCCCAGGAGG
                       D  K  T  E  L  G  C  L  R  S  V  I  L  F  N  P  E  V  R  G  L  K  S  A  Q  E

LmRXR-L              TTGAACTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATACTAGAACAACACATCCCGATGAACCAGGAAGATTT
LmRXR-S              TTGAACTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATACTAGAACAACACATCCCGATGAACCAGGAAGATTT
SgRXR-S-construct    TGGAACTTCTTCGTGAAAAAGTATATGCAGCTTTGGAAGAATATACTAGAACAACACACCCTGACGAACCAGGACGATTT
SgRXR-L-construct    TGGAACTTCTTCGTGAAAAAGTATATGCAGCTTTGGAAGAATATACTAGAACAACACACCCTGACGAACCAGGACGATTT
                     V  E  L  L  R  E  K  V  Y  A  A  L  E  E  Y  T  R  T  T  H  P  D  E  P  G  R  F
```

```
LmRXR-L              GCAAAACTTTTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATTTGTTTTTCTTTCGCCTTAT
LmRXR-S              GCAAAACTTTTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATTTGTTTTTCTTTCGCCTTAT
SgRXR-S-construct    GCGAAACTTCTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATTTGTTTTTCTTTCGCCTTAT
SgRXR-L-construct    GCGAAACTTCTGCTTCGTCTGCCTTCTTTACGTTCCATAGGCCTTAAGTGTTTGGAGCATTTGTTTTTCTTTCGCCTTAT
                      A  K  L  L  L  R  L  P  S  L  R  S  I  G  L  K  C  L  E  H  L  F  F  F  R  L  I

Prey vector                                                                                CATGTCGGCCGCCTCGGCC
LmRXR-L              TGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTCTGATTCATAA
LmRXR-S              TGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTCTGATTCATAA
SgRXR-S-construct    TGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTCTGATTCATAACATGTCGGCCGCCTCGGCC
SgRXR-L-construct    TGGAGATGTTCCAATTGATACGTTCCTGATGGAGATGCTTGAATCACCTTCTGATTCATAACATGTCGGCCGCCTCGGCC
                       G  D  V  P  I  D  T  F  L  M  E  M  L  E  S  P  S  D  S        CD III Anchor

Prey vector          TCTAGAGGGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTCAC
SgRXR-S-construct    TCTAGAGGGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGA                3' primer
SgRXR-L-construct    TCTAGAGGGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGAT
```

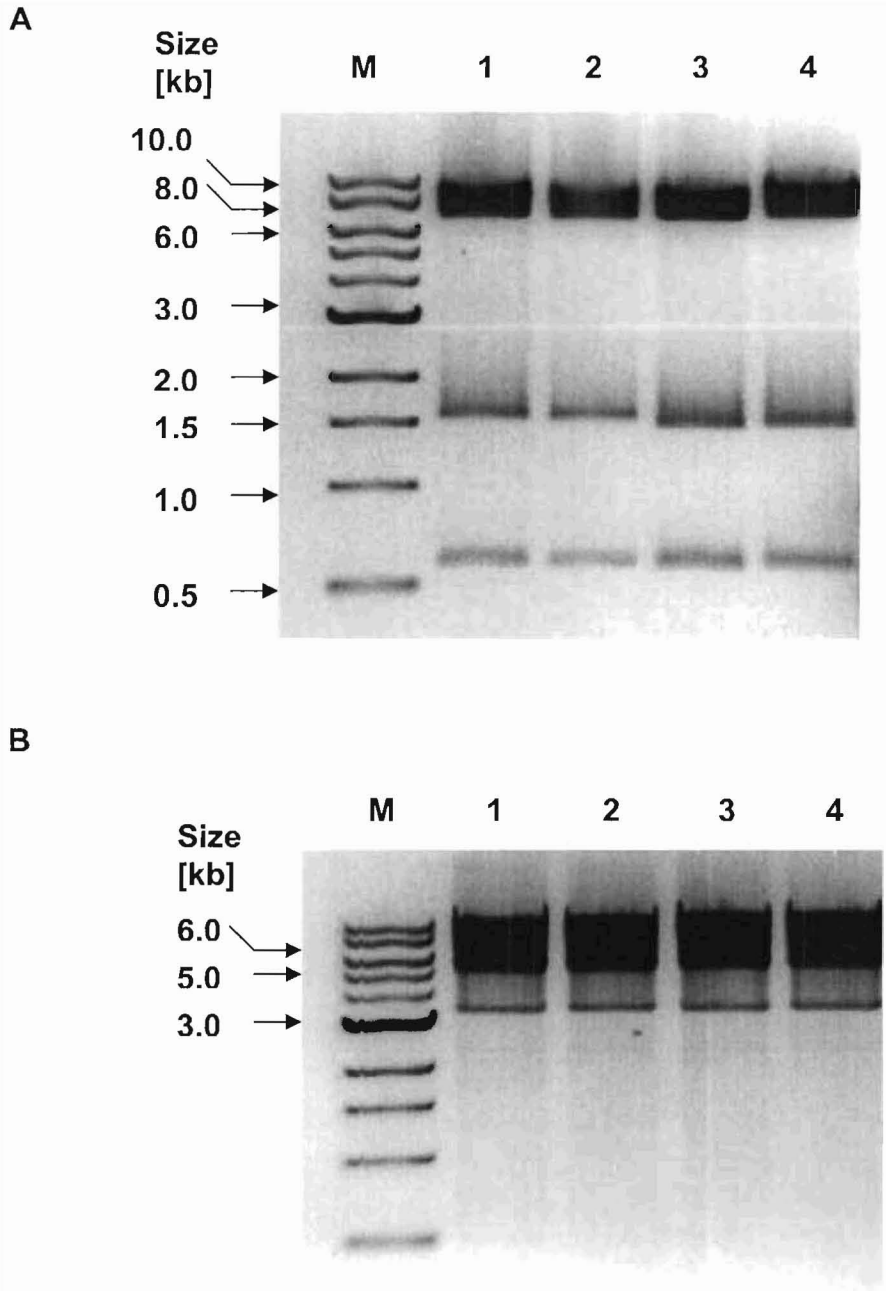


Figure 3-10 Digest of SgRXR prey vector with *Hind* III or *Not* I/*Nsi* I

Plasmids were isolated from two transformants of SgRXR-L (Lane 1&2) and SgRXR-S (Lane 3&4), respectively, then digested with *Hind* III (A) or *Not* I/*Nsi* I (B). M, 1 kb DNA ladder.

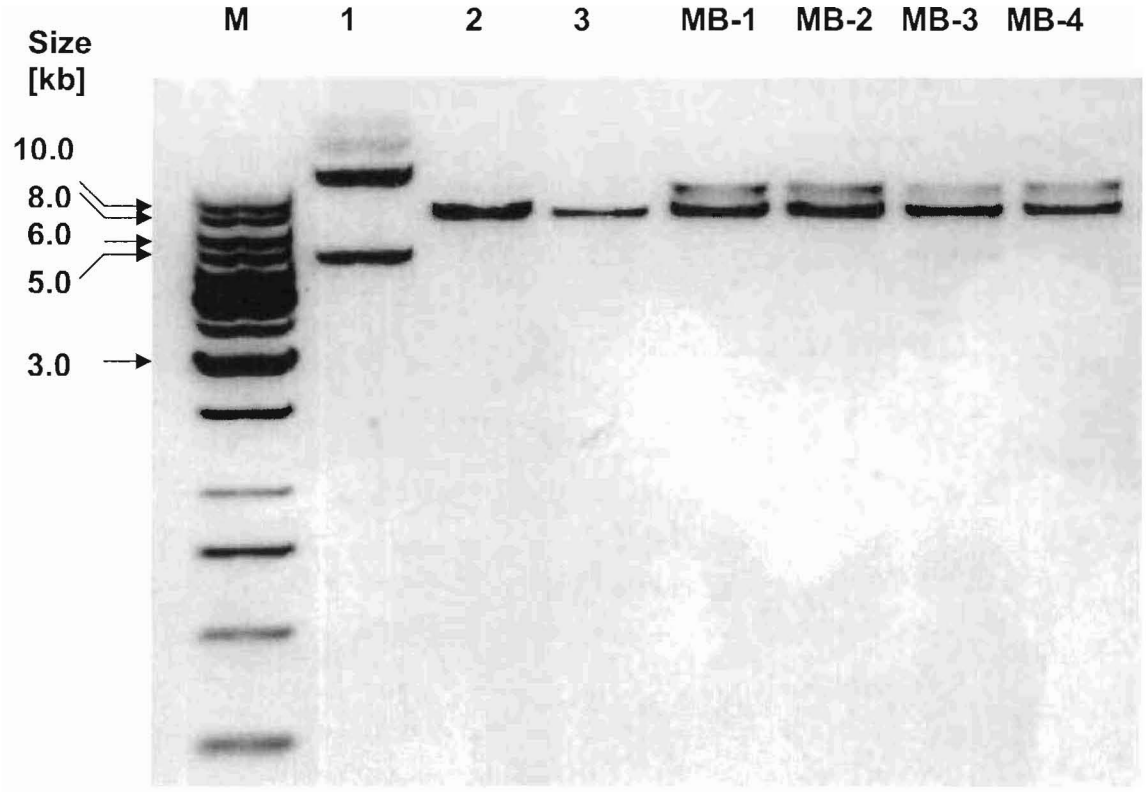


Figure 3-11 Double digestion of mutated bait vectors with *Eco*R I/*Mlu* I

The products of undigested pHIS2 reporter vector (Lane 1), the digested pHIS2 (Lane 2), FARE bait vector (Lane 3) as well and the mutated bait vectors (MB-1, MB-2, MB-3 and MB-4) were analyzed on 1% agarose gel. M, 1kb DNA ladder.

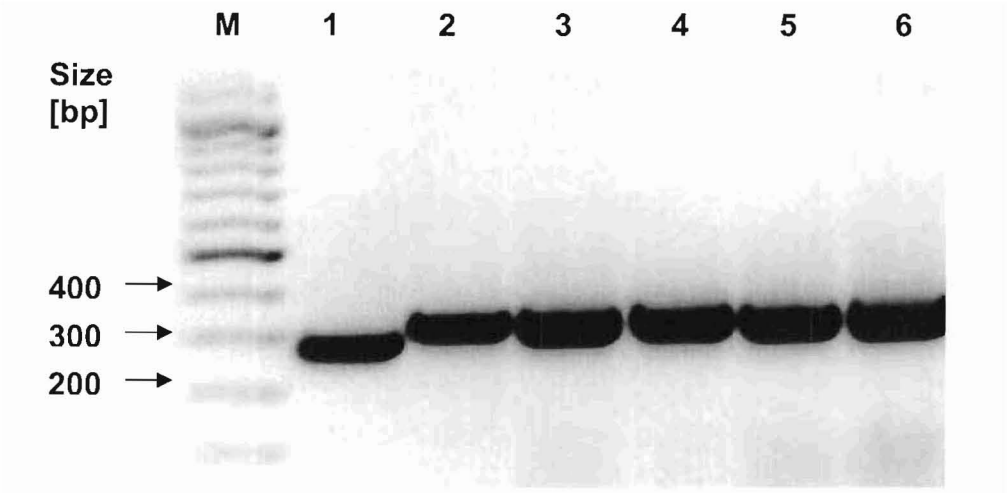


Figure 3-12 PCR of plasmids containing mutated bait sequences

PCR products of pHIS2 vector (Lane 1), FARE bait vector (Lane 2) as and the mutated bait vectors MB-1(Lane 3), MB-2(Lane 4), MB-3(Lane 5), MB-4(Lane 5) were analyzed on 1% agarose Gel. M, 100 bp DNA ladder.

the two RXR isoforms detected in the embryo or fat body of Locusta migratorya (Hayward et al. 1999; Hayward et al. 2003). The fact that nested PCR had to be performed, because no primer combination was sufficient to amplify RXR directly from total flight muscle RNA, indicates that these nuclear receptors are expressed at low levels only. In contrast, PCR with primers for the abundant muscle FABP resulted in strong bands in a single PCR reaction (Figure 3-5).

After it was confirmed that both isoforms of RXR are present in locust flight muscle, their full-length coding sequences were obtained from PCR with their anchor primers. The sizes of the PCR fragments amplified with RXR anchor primers correspond to the two isoforms of SgRXR (Figure 3-7). To obtain the complete coding DNA sequences for the two isoforms of SgRXR, the PCR fragments were cloned into the TA vector. As expected for these conserved nuclear receptors, their encoded proteins are identical to RXRs from *Locusta migratoria* except for a deletion of serine and a substitution of leucine by proline in the short isoform (Figure 3-9).

The successful recombination of SgRXR into the linearized pGADT7-rec-2 resulted in a circular and fully functional expression vector, which confers the Leu+ phenotype to Y187 leu auxotroph. Thus, the growth of the yeast transformants on SC/-Leu medium indicated the presence of the SgRXR prey vector. Restriction digestions of the prey vectors (Figure 3-10) confirmed the correct orientation of the inserts, and sequencing of the prey vector indicated the successful construction of the SgRXR prey vectors (Figure 3-9). Compared to the sequencing result of the SgRXR in the TA cloning vector, the addition of

another single difference in amino acid (glycine instead of aspartic acid) for SgRXR-S may due to the error in PCR. Further sequencing confirmed this change. Based on the successful construction of the wild type bait vector, a number of modified bait vectors were also constructed. The purpose of these mutated baits was to investigate the specificity of the FARE-RXR interaction, and the effects resulting from the change in the nucleotide sequence of the proposed target element. Because the four mutated target elements are the same in size, the bands with the same sizes were obtained from digestions (Figure 3-11) and PCRs (Figure 3-12). Only the gene sequencing of each individual mutated bait vector distinguished the nucleotide differences among them and the success in the construction of each.

In Y187 yeast strains, the two SgRXRs prey vectors should allow for the expression of both isoforms as GAL4 AD fusion proteins containing HA epitope tag. If SgRXR interacts with the target DNA sequence in the bait vector, the fused GAL4 AD protein functions as an activation domain and activates the expression of the downstream *HIS3* gene. The level of *HIS3* protein is enhanced and higher than the basal level, so the positive one-hybrid clones with true one-hybrid interaction will grow on the histidine deficient medium containing optimal concentration of 3-AT (functions to suppress the basal level expression of HIS3 gene). The higher than optimal concentration of 3-AT is needed to suppress the growth of positive one-hybrid clones. Thus, interactions of these SgRXR fusion proteins with the FARE bait or with mutated bait sequences can be investigated in Y187 yeast strain with one-hybrid system.

CHAPTER 4:
ONE–HYBRID ANALYSIS OF INTERACTION BETWEEN SGRXR AND FARE

4.1 Introduction

The analysis of one-hybrid interaction between the retinoic X receptor from *Schistocerca gregaria* (SgRXR) and the fatty acid response element (FARE) target sequence requires that yeast cells contain both the FARE bait vector and the SgRXR prey vector. The co-existence of both vectors in yeast cells can be accomplished by simultaneous or sequential LiAc-mediated yeast transformation. Sequential transformation involves the transformation of yeast cells with a single type of plasmid, and the subsequent re-transformation of the cells with the second plasmid. Simultaneous transformation is the co-transformation of yeast cells with two plasmids containing different selection markers; co-transformants will grow as colonies on SC medium that selects for both markers. While simultaneous transformation is time-saving, the probability to obtain co-transformants with both plasmids is lower than with the sequential method, and hence the co-transformation efficiency is lower than that of sequential transformation. Moreover, sequential transformation may be more convenient when several different co-transformants for a given plasmid need to be prepared.

In the current study, some other co-transformants are necessary besides the co-transformant containing SgRXR prey vector and FARE-bait vector, in order to compare and verify the specificity of the one-hybrid interactions. Such

co-transformants include different combinations of both p53 prey vector, pGADT7-rec-2 vector or each isoform of SgRXR prey vector and p53 bait vector, pHIS2 vector, FARE-bait vector or each of the mutated bait vector (Table 4-1). While co-transformants can be constructed by simultaneously transforming yeast cells with each combination of both prey and bait vectors, a more convenient way is to transform yeast cells with seven bait vectors separately, and sequentially transform them with four prey vectors respectively, or *vice versa.*

In each yeast co-transformant (Table 4-1), the prey vector expresses GAL4 AD fusion protein under the control of the constitutive full-length yeast alcohol dehydrogenase (AHD1) promoter (P_{ADH1}) and the ADH1 termination signal. The full-length AHD1 promoter has been proven to initiate high level expression of downstream proteins (Ammerer 1983). The p53 prey vector pGAD-Rec2-53 encodes murine p53 protein fused with upstream GAL4 AD polypeptide containing SV40 NLS (Figure 2-11), while the circular, empty prey vector pGADT7-rec-2 encodes SV40 NLS-containing GAL4 AD polypeptide followed by an HA tag (Figure 3-3). Our SgRXR prey vectors express either SgRXR-L or SgRXR-S downstream of the GAL4 AD containing SV40 NLS and the HA tag, so the expressed SgRXR fusion proteins can be detected with antibodies to HA by Western blotting.

The detection of the one-hybrid interaction between the fusion prey proteins and their target DNA elements is dependent on the promoter region of the *HIS3* reporter gene in the bait vector. The expression level of the *HIS3* gene

Table 4-1 List of required yeast co-transformants

The cotransformants are denoted as Y187 [bait/ prey].

	P53HIS2	pHIS2	pFARE	pMB-1	pMB-2	pMB-3	pMB-4
pGAD-SgRXR-L	Y187[p53HIS2 + pGAD-SgRXR-L]	Y187 [pHIS2 + pGAD-SgRXR-L]	Y187 [pFARE + pGAD-SgRXR-L]	Y187 [pMB-1 + pGAD-SgRXR-L]	Y187 [pMB-2 + pGAD-SgRXR-L]	Y187 [pMB-3 + pGAD-SgRXR-L]	Y187 [pMB-4 + pGAD-SgRXR-L]
pGAD-SgRXR-S	Y187[p53HIS2 + pGAD-SgRXR-S]	Y187 [pHIS2 + pGAD-SgRXR-S]	Y187 [pFARE + pGAD-SgRXR-S]	Y187 [pMB-1 + pGAD-SgRXR-S]	Y187 [pMB-2 + pGAD-SgRXR-S]	Y187 [pMB-3 + pGAD-SgRXR-S]	Y187 [pMB-4 + pGAD-SgRXR-S]
pGAD-Rec2-53	Y187[p53HIS2 + pGAD-Rec2-53]	Y187 [pHIS2 + pGAD-Rec2-53]	Y187 [pFARE + pGAD-Rec2-53]	Y187 [pMB-1 + pGAD-Rec2-53]	Y187 [pMB-2 + pGAD-Rec2-53]	Y187 [pMB-3 + pGAD-Rec2-53]	Y187 [pMB-4 + pGAD-Rec2-53]
pGADT7-Rec2	Y187[p53HIS2 + pGADT7-Rec2]	Y187 [pHIS2 + pGADT7-Rec2]	Y187 [pFARE + pGADT7-Rec2]	Y187 [pMB-1 + pGADT7-Rec2]	Y187 [pMB-2 + pGADT7-Rec2]	Y187 [pMB-3 + pGADT7-Rec2]	Y187 [pMB-4 + pGADT7-Rec2]

is regulated by two elements in the promoter region: the HIS3 minimal promoter ($P_{minHIS3}$) and an upstream activating sequence (UAS). The latter sequence serves as DNA-binding site for transcriptional regulatory proteins and facilitates significant increase in transcription upon regulatory stimulation (Guarente 1984). PminHIS3 is composed of a regulated TATA box (T_R, TATAAA) and a constitutive TATA box (T_C, TATACA) (Struhl 1986). In our study, three tandem copies of the target DNA sequences were inserted into the bait vector to act as UAS, while the DNA-binding proteins fused with the GAL4 AD serve as upstream regulatory activators. As discussed in Chapter 2, the GAL4 AD fusion proteins, as potential eukaryotic transcription factors, are composed of the DNA binding domains from suspected nuclear receptors (the tested proteins) fused with the GAL4 activation domain. When the tested proteins bind to the target response elements, the GAL4 AD, as an activation region of the fusion protein, up-regulates HIS3 transcription by interacting with the T_R-binding protein (Figure 4-1). The low level of constitutive HIS3 transcription due to T_C is often referred to as background or "leaky" HIS3 expression (Struhl 1987).

In yeast co-transformants, GAL4 AD fusion proteins are expressed and translocated into the nucleus because of a SV40 nuclear localization signal (NLS) (Chien et al., 1991). Without one-hybrid interaction, the HIS3 gene is expressed constitutively at the low basal level under the control of T_C, but the growth of the co-transformants on the histidine deficient medium can be suppressed by experimentally determined, relatively low concentrations of the inhibitor 3-AT (optimal 3-AT concentration). If a fusion protein has true one-hybrid

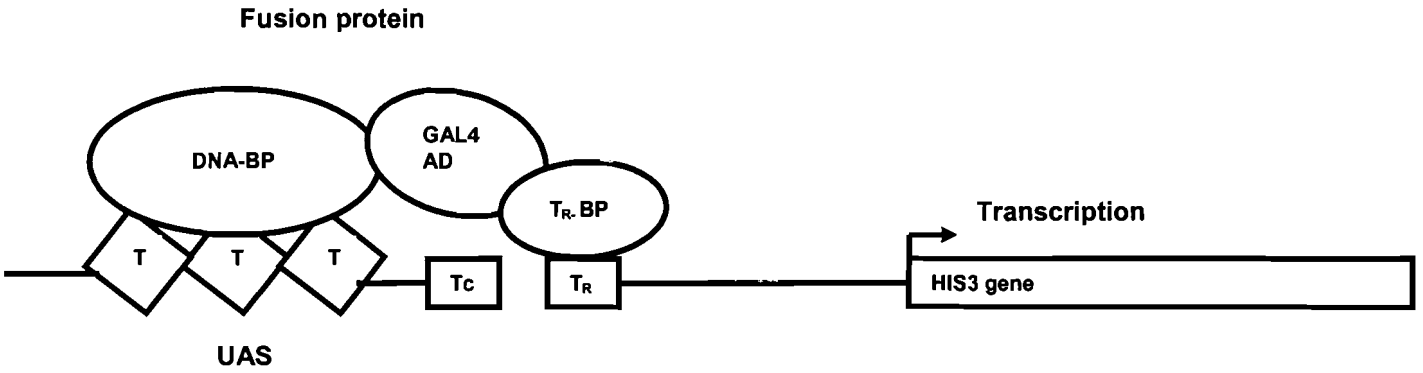


Figure 4-1 Schematic representation of *HIS3* gene regulation in yeast one-hybrid assay

interaction with the target DNA bait sequence, its activation domain is brought in contact with the T_R-binding protein, thus stimulating regulated HIS3 transcription (Struhl 1987); these co-transformants will grow well on the histidine deficient medium containing the optimal 3-AT concentration, and much higher concentrations are needed to suppress their growth. The murine p53 protein has been proven to interact with the p53 response element in a p53HIS2 vector (Thukral et al., 1994), so the co-transformant Y187 [p53HIS2 + pGAD-Rec2-53] can serve as one-hybrid positive control.

The analysis of protein-DNA interactions is based on the ability of the nuclear receptor/GAL4 AD fusion protein to stimulate the transcription of the HIS3 gene. The level of the stimulation can be assessed by the level of the expressed His3p, which in turn can be scored as cell growth on histidine deficient media containing various concentration of 3-AT (Horecka and Sprague 2000). Cells of yeast co-transformants can be applied to the selection media plates by direct spreading or streaking.

4.2 Materials and methods

4.2.1 One-hybrid analysis by direct spreading

The TA cloning vector construct pCR2.1-SgRXR for either isoform was used as template for PCR with RXR forward and reverse anchor primers. After purification, the amplicons for both isoforms were mixed and co-transformed into yeast cells together with S*ma* I-linearized pGADT7-rec-2 and FARE bait vector according to the library-scale transformation procedure in the manual of

Matchmaker library construction and screening kits (Clontech, Mountain View, CA). The co-transformants were resuspended in 0.9% NaCl solution and further diluted 1:10 or 1:100, then spread onto the selection media plates SC/-Trp, Sc/-Leu, SC/-Trp/-Leu and SC/-Trp/-Leu/-His+3-AT(0 mM, 10 mM, 20 mM and 30 mM). After 3-5 days incubation at 30 °C, colonies were counted and further verified by colony PCR with primers RXR-2 and RXR-3.

4.2.2 Construction of test co-transformants

Y187 yeast strains were first transformed with one of the bait vectors p53HIS2, pHIS2, pFARE, pMB-1, pMB-2, pMB-3 and pMB-4, respectively. The Y187 [Bait vector] transformants were selected on SC/-Trp agar medium and incubated at 30 °C for 3 days. Several isolated colonies on each transformation plate were picked and re-streaked to single colonies on fresh SC/-Trp agar plates. A single colony from each isolate was then grown as a patch to make a patch master plate for glycerol stock preparation, background HIS3 titration, and the construction of the test co-transformants.

The Y187 [Bait vector] transformants were further transformed with one of the prey vectors pGAD-SgRXR-L, pGAD-SgRXR-S, pGAD-Rec2-53 and pGADT7-Rec2, respectively. The co-transformants were plated on SC/-Trp/-Leu agar medium and incubated. After the isolated colonies for each transformation were re-streaked, patch master plates were made for glycerol stock preparation and future one-hybrid analysis.

4.2.3 Background HIS3 titration by streaking

Each medium plate containing SC/-Trp/-His+3-AT was divided into four sections. A small amount of each Y187 [bait vector] transformants was picked with a toothpick and streaked into single colony on each corresponding section. The plates were incubated at 30 °C for 3-5 days and examined for cell growth.

4.2.4 One-hybrid analysis by streaking

The test co-transformants were streaked on the SC/-Trp/-Leu/-His+3-AT agar medium plates divided into a grid or circular sections. Sixteen grids were divided equally at the back of each plate contain SC/-Trp/-Leu/-His+3-AT agar medium (Table 4-2). A pinch of each test co-transformant was picked from the patch master plate with toothpick and streaked as a horizontal line onto the corresponding grid. Each plate containing SC/-Trp/-Leu/-His+3-AT agar medium was divided equally into four sections (Figure 4-2). Each test co-transformant was picked and streaked into single colonies in the corresponding section. All the streaked plates were incubated at 30 °C for 3-5 days and evaluated for cell growth.

4.3 Results

4.3.1 One-hybrid analysis by direct spreading

After 3-day incubation, all plates were full of colonies (too many to count). However, on SC/ Trp/-Leu/-His+ 3-AT plates, the number of colonies appeared to decrease with increasing 3-AT concentration (Table 4-3). The transformation efficiency and number of clones screened were calculated as follows:

Table 4-2 Grid layout for one-hybrid analysis by streaking in grid

Y187[pbait + pGAD-SgRXR-L]

TF1-1 **Y187[p53HIS2 + pGAD-SgRXR-L]**	**TF2-2** **Y187 [pHIS2 + pGAD-SgRXR-S]**	**TF2-1** **Y187 [pHIS2 + pGAD-SgRXR-L]**	**TF1-3** **Y187[p53HIS2 + pGAD-Rec2-53]**
TF4-4 Y187 [pMB-1 + pGADT7-Rec2]	**TF3-4** Y187 [pFARE + pGADT7-Rec2]	**TF2-4** Y187 [pHIS2 + pGADT7-Rec2]	**TF1-2** Y187[p53HIS2 + pGAD-SgRXR-S]
TF3-1 Y187 [pFARE + pGAD-SgRXR-L	**TF7-3** Y187 [pMB-4 + pGAD-Rec2-53]	**TF6-4** Y187 [pMB-3 + pGADT7-Rec2	**TF5-4** Y187 [pMB-2 + pGADT7-Rec2]
TF7-1 Y187 [pMB-4 + pGAD-SgRXR-L]	**TF6-1** Y187 [pMB-3 + pGAD-SgRXR-L	**TF5-1** Y187 [pMB-2 + pGAD-SgRXR-L]	**TF4-1** Y187 [pMB-1 + pGAD-SgRXR-L]

Y187[pbait + pGAD-SgRXR-S]

TF1-1 **Y187[p53HIS2 + pGAD-SgRXR-L]**	**TF2-2** **Y187 [pHIS2 + pGAD-SgRXR-S]**	**TF2-1** **Y187 [pHIS2 + pGAD-SgRXR-L]**	**TF1-3** **Y187[p53HIS2 + pGAD-Rec2-53]**
TF4-4 Y187 [pMB-1 + pGADT7-Rec2]	**TF3-4** Y187 [pFARE + pGADT7-Rec2]	**TF2-4** Y187 [pHIS2 + pGADT7-Rec2]	**TF1-2** Y187[p53HIS2 + pGAD-SgRXR-S]
TF3-2 Y187 [pFARE + pGAD-SgRXR-S]	**TF7-3** Y187 [pMB-4 + pGAD-Rec2-53]	**TF6-4** Y187 [pMB-3 + pGADT7-Rec2	**TF5-4** Y187 [pMB-2 + pGADT7-Rec2]
TF7-2 Y187 [pMB-4 + pGAD-SgRXR-S]	**TF6-2** Y187 [pMB-3 + pGAD-SgRXR-S]	**TF5-2** Y187 [pMB-2 + pGAD-SgRXR-S]	**TF4-2** Y187 [pMB-1 + pGAD-SgRXR-S]

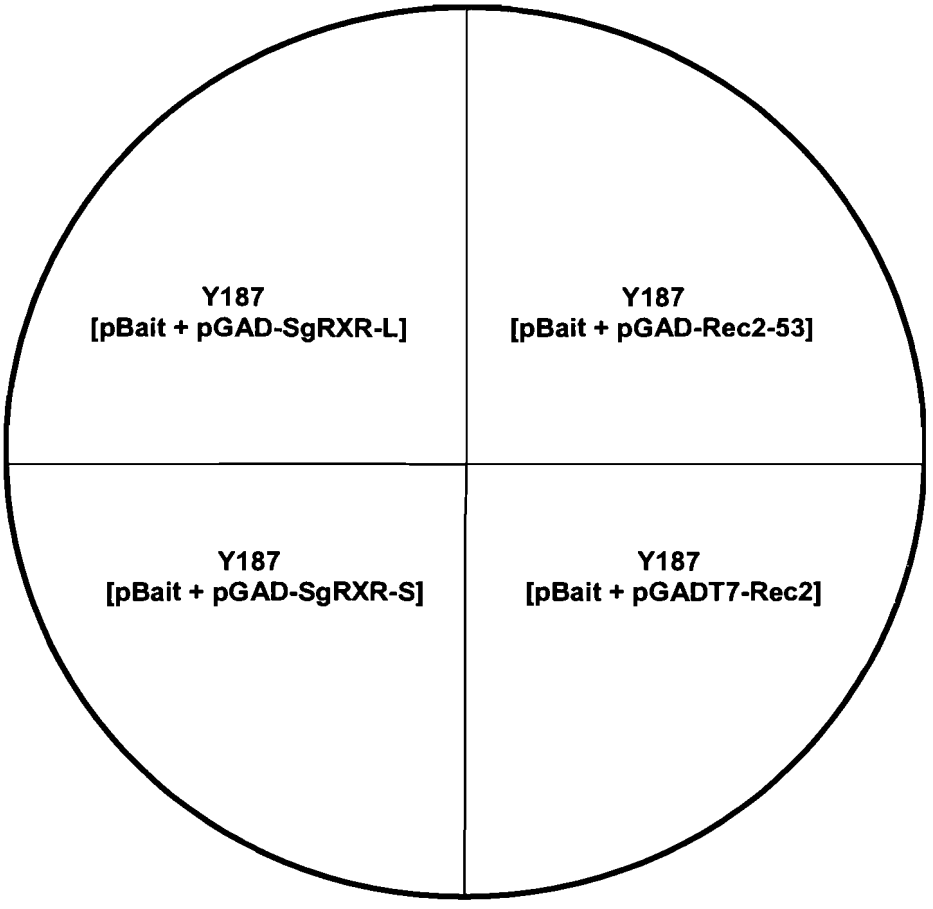


Figure 4-2 Plate layout for one-hybrid analysis by streaking in section

Table 4-3 One-hybrid analysis by direct spreading

Suspension of cotransformants was diluted, as indicated, and spread on selective media plates containing varying concentrations of 3-AT. The numbers of colonies appearing on each plate are listed. TMTC stands for too many to count.

	Y187 [pFARE + pGAD-SgRXR-L / pGAD-SgRXR-S]		
	1:1	1:10	1:100
SC/-Trp	TMTC	TMTC	620 (Briefly)
SC/-Leu	TMTC	TMTC	470 (Briefly)
SC/-Trp/-Leu	TMTC	TMTC	295 (Briefly)
SC/-Trp/-Leu/-His+ 3-AT			
0mM	TMTC	TMTC	
10mM	TMTC	TMTC	
20mM	TMTC	TMTC	
30mM (optimal concentration)	TMTC	TMTC	

Less colonies observed ↓

a. Transformation efficiency (expressed as number of transformants per 3μg pGADT7-rec-2) :

$$\frac{\text{Colonies on SC/-Leu X dilution factor X total suspension volume [ml]}}{\text{Volume plated [ml]}}$$

b. Number of colonies screened:

$$\frac{\text{Colonies on SC/-Trp/-Leu X dilution factor X total suspension volume [ml]}}{\text{Volume plated [ml]}}$$

Based on the above formulas, the transformation efficiency was calculated to be about 2.8×10^6 transformants/3 μg pGADT7-rec-2 (expected higher than 1×10^6), and about 1.77×10^6 clones was screened(expected higher than 3×10^5). Both isoforms of SgRXR were present in all six colonies picked from histidine deficient medium containing 30 mM 3-AT (Figure 4-3). However, upon re-streaking on SC/ Trp/-Leu/-His+ 30mM 3-AT plates, all colonies appeared to contain one of the two isoforms only.

4.3.2 Background HIS3 titration by streaking

Without 3-AT, all the transformants containing bait vector grew well; however, their growth decreased with increasing concentration of 3-AT (Figure 4-4). There was no growth for Y187 [p53HIS2] and Y187 [pMB-4] transformants

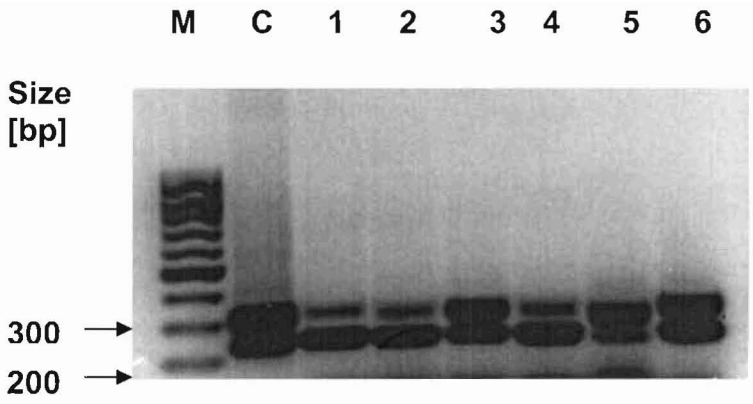


Figure 4-3 Presence of both isoforms of SgRXR in cotransformants

Six colonies were picked from histidine deficient medium containg 30mM 3-AT, and used for colony PCR with primers RXR-2 and RXR-3 (Lane 1-6). C, positive control; M, 100 bp DNA ladder.

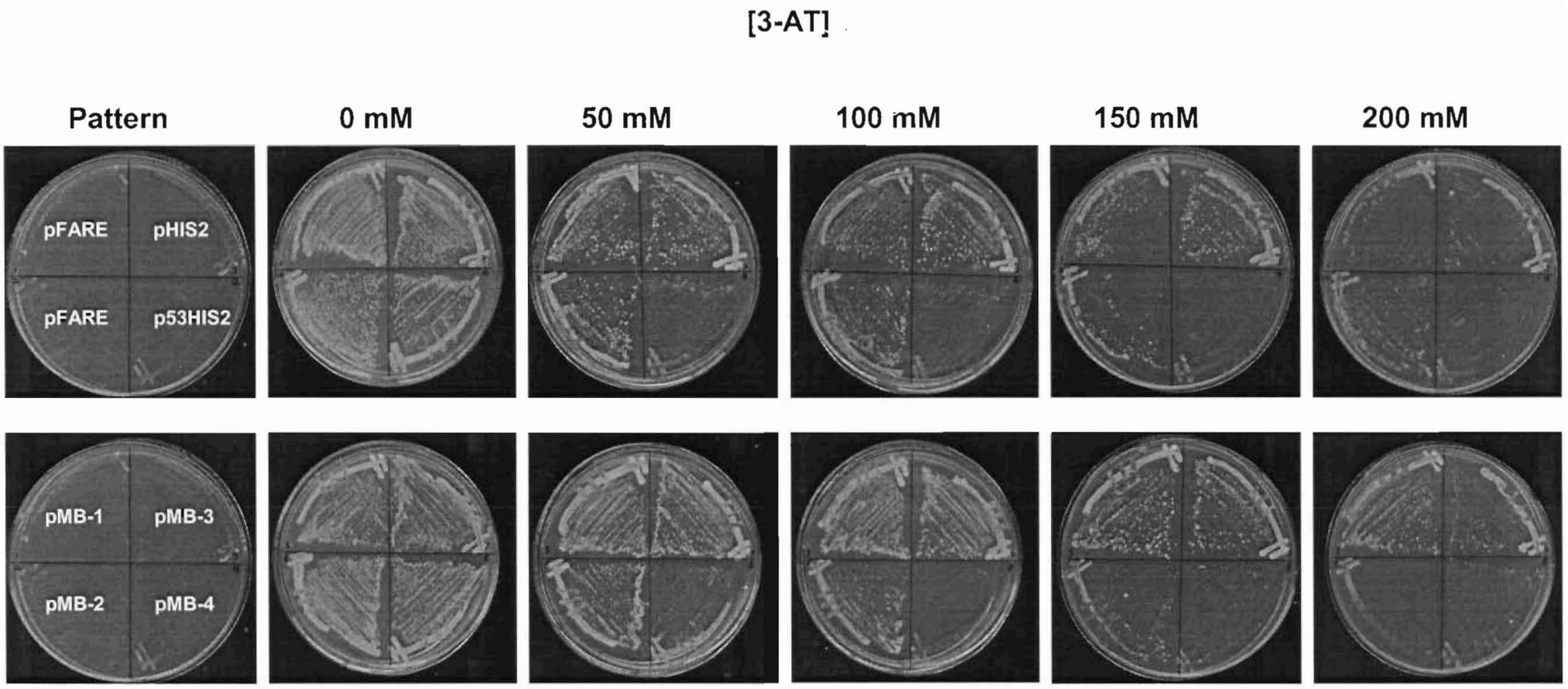


Figure 4-4 Background HIS3 titration of bait vectors

on the selection plates containing 50 mM 3-AT. The growth of Y187 [pMB-2] was inhibited on plates containing 100 mM 3-AT. The growth of Y187 [pFARE] stopped when the 3-AT concentration was 150 mM. When the 3-AT concentration reached 200 mM, Y187 [pHIS2] and Y187 [pMB-3] transformants stopped growing; there was still growth for Y187 [pMB-1], but the growth was inhibited and far less than that at 0 mM 3-AT. Two independent titration experiments showed the same results.

4.3.3 One-hybrid analysis by streaking in grid

The growth of all the test co-transformants was strong without 3-AT, but decreased with increasing 3-AT concentration, except for the positive control Y187[p53HIS2+pGAD-Rec2-53] which still grew well at 300 mM 3-AT (Figure 4-5). The growth of Y187 [p53HIS2+pGAD-SgRXR-L/-S)] was largely inhibited by 50 mM 3-AT. The co-transformants Y187 [pMB-2+pGAD-SgRXR-L/-S] and Y187 [pMB-4+pGAD-SgRXR-L/-S] showed somewhat stronger growth at 150 mM 3-AT than Y187 [pMB-2+ pGADT7-Rec2] and Y187 [pMB-4+pGADT7-Rec2]; growth of the latter two was greatly reduced at 150 mM and completely inhibited at 200 mM 3-AT. Y187 [pFARE + pGAD-SgRXR-L/-S] grew stronger at 200 mM 3-AT than Y187 [pFARE + pGADT7-Rec2], but both strains were completely inhibited at 300 mM 3-AT. There was no detectable difference in growth between Y187 [pMB-1+pGAD-SgRXR-L/-S], Y187 [pMB-3+pGAD-SgRXR-L/-S] or Y187 [pHIS2+pGAD-SgRXR-L/-S] and Y187 [pMB-1+pGADT7-Rec2], Y187 [pMB-3+pGADT7-Rec2] or Y187 [pHIS2+pGADT7-Rec2], respectively, even when the 3-AT concentration reached 300 mM.

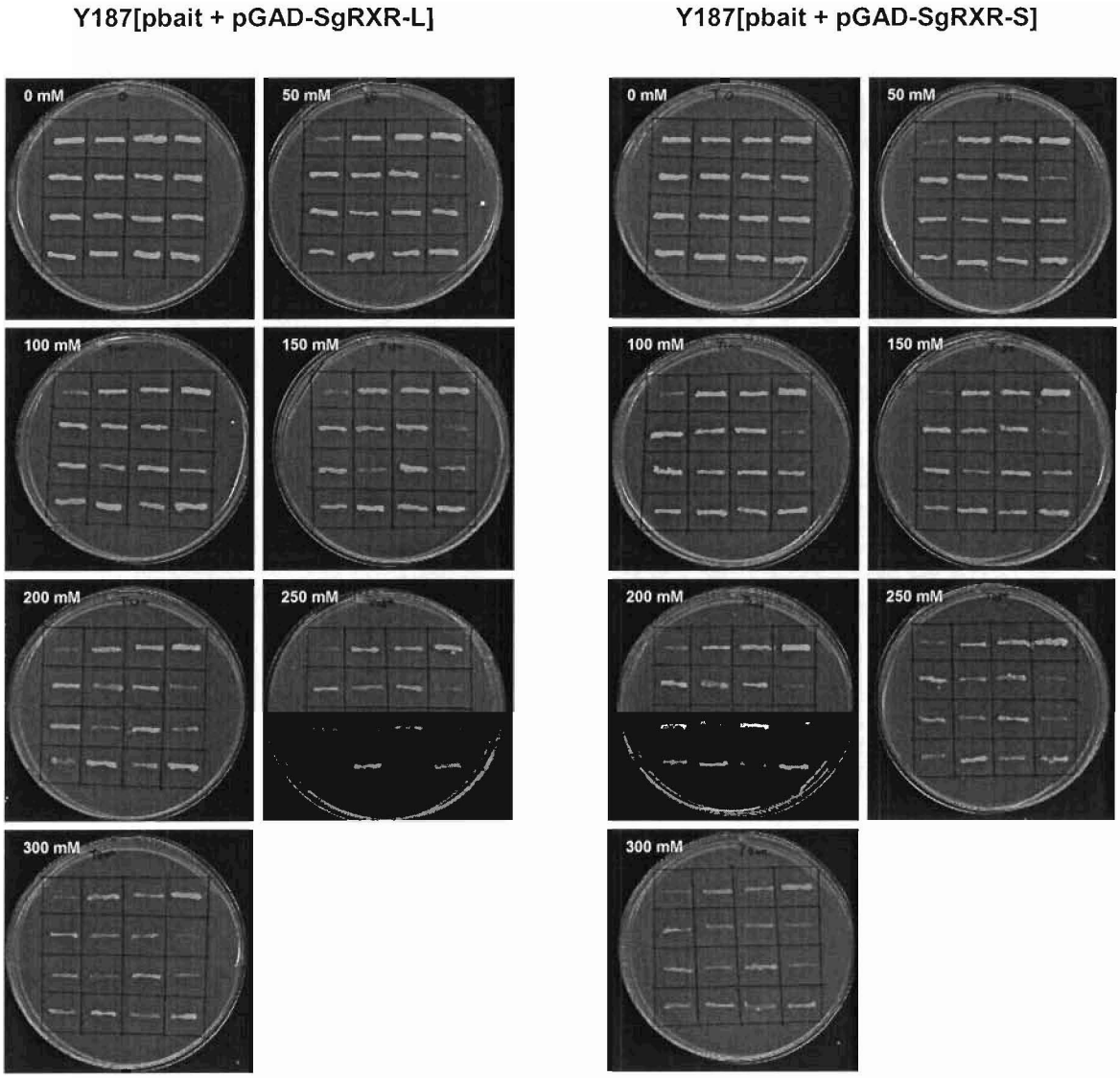


Figure 4-5 Growth evaluation of one-hybrid analysis by streaking co-transformants in grid

4.3.4 One-hybrid analysis by streaking in section

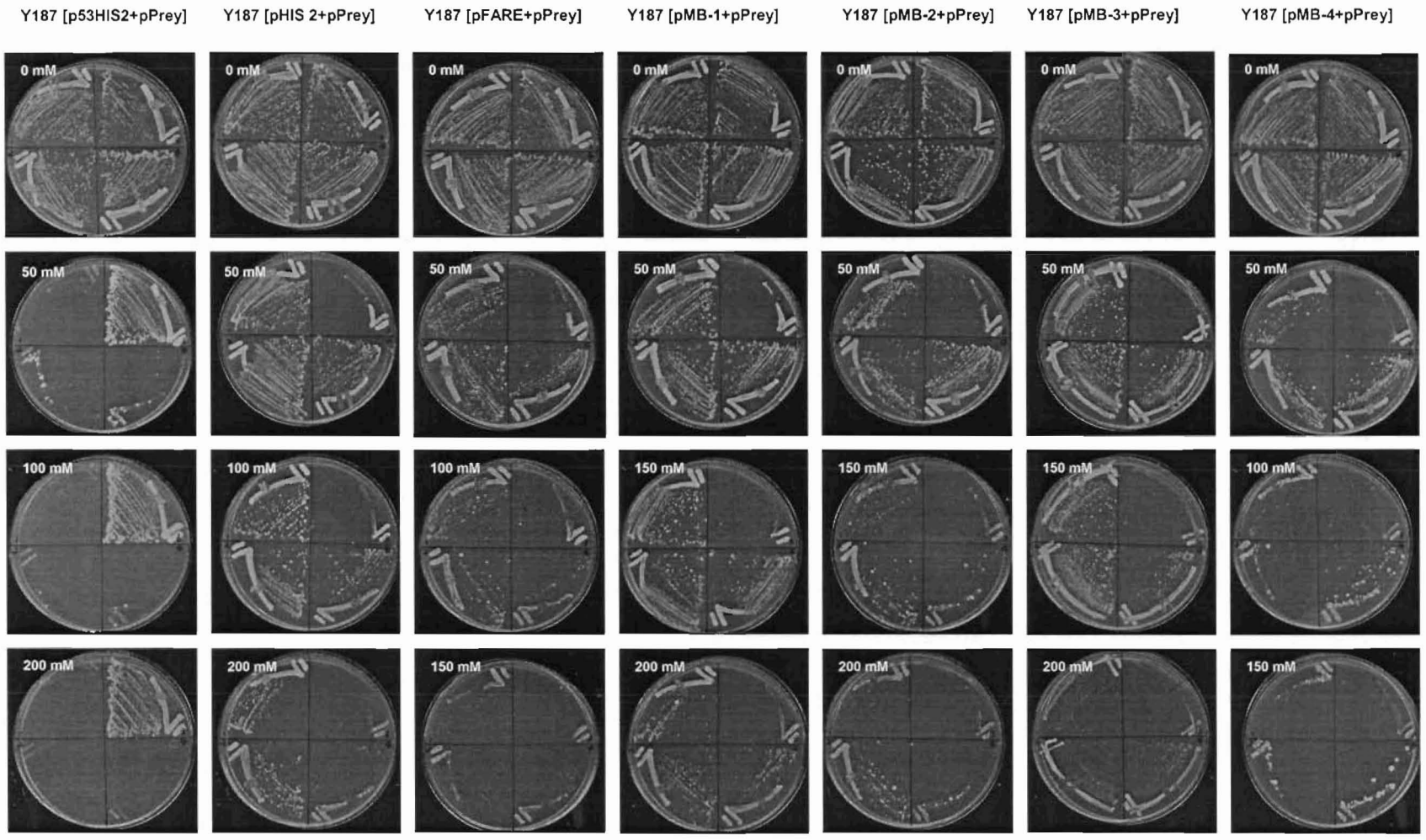
Figure 4-6 shows the growth of each test co-transformant on SC/-Trp/-Leu/-His+3-AT plate. As a positive control, Y187 [p53HIS2+pGAD-Rec2-53] grew well even at 200 mM 3-AT, while the other three p53HIS2 containing co-transformants died out at as low as 50 mM 3-AT. The growth of the co-transformants containing p53 prey vector and any bait (except p53HIS2) vector was suppressed by 50 mM 3-AT. Y187 [pHIS2+pGAD-SgRXR-L/-S] showed stronger growth than Y187 [pHIS2+pGADT7-Rec2] even at 200 mM 3-AT, but its growth was not as strong as the positive control. Y187 [pFARE + pGAD-SgRXR-L/-S] had stronger growth than Y187 [pFARE + pGADT7-Rec2] at 50 mM, but their growth was reduced at 100 mM and absent at 150 mM 3-AT.

The growth of Y187 [pMB-1+pGAD-SgRXR-L/-S] was stronger than that of Y187 [pMB-1+pGADT7-Rec2] at 150 mM 3-AT, but inhibited at 200 mM 3-AT. Y187 [pMB-2+pGAD-SgRXR-L/-S] had slightly stronger growth than Y187 [pMB-2+pGADT7-Rec2]. Y187 [pMB-3+pGAD-SgRXR-L/-S] grew stronger than Y187 [pMB-3+pGADT7-Rec2] at 50mM to 150 mM 3-AT, but also stopped growth at 200 mM 3-AT. The growth of both Y187 [pMB-4+pGAD-SgRXR-L/-S] and Y187 [pMB-4+pGADT7-Rec2] was inhibited as already at 50 mM 3-AT.

4.4 Discussion

In this study, yeast one-hybrid interactions were determined after the co-transformants were constructed with sequential transformation methods. This

Figure 4-6 Growth evaluation of one-hybrid analysis by streaking co-transformants in section



method was chosen because the each of the seven bait vectors needed to be co-presence with four different prey vectors, respectively (Table 4-1).

The six other bait vectors were constructed by restriction digestion of the pHIS2 reporter vector and the insertion of the different target DNA sequence into it. It is noteworthy that the subtle sequence differences in the promoter region of the *HIS3* gene can affect not only the interaction with its binding proteins which will change the level of induced HIS3 expression, but also the level of background *HIS3* expression due to nonspecific binding to endogenous yeast transcription factors.

Thus, the optimal 3-AT concentration needed to be separately determined for each transformant. However, the optimal 3-AT concentration needed to suppress the HIS3 background expression depends also on the amount and mode by which the bait-containing transformants are applied to the plate. As more cells were applied to the plate, higher concentration of 3-AT were needed to suppress the background growth. The optimal 3-AT concentration determined by spreading was different from that tested by streaking. With all these variables, the optimal 3-AT concentration should be treated as an approximate value. The optimal 3-AT concentrations determined by the spreading method were about 40 mM, 10 mM and 30 mM, for pHIS2, p53HIS2 and pFARE-HIS2 vector, respectively (Table 2-1), while the streaking method resulted in far higher values (200 mM, 50 mM, and 150 mM, respectively. With the mutated response elements, background expression was suppressed at concentrations varying between 50 and 200 mM (Figure 4-4).

For one-hybrid analysis by the direct spreading method, the large number of colonies observed on SC/ Trp/-Leu/-His+ 30 mM 3-AT plates may indicate specific interactions between SgRXR and the FARE. To examine the possibility that a heterodimer between the long and short isoform of RXR is needed for binding of the element, cells were co-transfected with both isoforms. Initially, both isoforms were present in the positive clones, but upon re-streaking, one of the isoform was always lost. The continued growth of the cells could indicate that either isoform can interact with the element and thus activate HIS3 expression. However, the interactions appear to be not as strong as those found for the P53 positive control, and hence one has to be careful in over-interpreting the results. It cannot be excluded that too many cells were spread on the selection medium, or that the 3-AT concentration used was too low, resulting in false positive clones. Therefore, the direct spreading method appears to be not optimally suited for our study. Furthermore, because different bait vectors had different optimal 3-AT concentrations, comparisons of the strength of interaction between the prey proteins with FARE or with other mutant baits is difficult to achieve. To test whether a specific protein interacts with its target element, the streaking method is more direct and accurate.

Cells of each co-transformant were streaked as a line in a grid (grid streaking method) or into single colonies in a section (section streaking method) on the selection medium with various concentrations of 3-AT. The stronger the one-hybrid interaction, the higher the activated expression level of HIS3 should be, thus requiring a higher the 3-AT concentration to suppress the growth.

Because more cells are applied in the initial application, grid streaking method seems to be less accurate than the section streaking method, in which cells are diluted to single colonies; nevertheless, both methods gave an indication for the binding specificity between a prey protein and its target sequence.

Not surprisingly, both streaking methods showed that p53 protein had strong interaction with its target response element, and thus growth of Y187 [p53HIS2+pGAD-Rec2-53] was still obvious at 200 or 300 mM 3-AT, while background growth was totally suppressed by 50 mM 3-AT. In contrast, the SgRXR prey proteins and protein expressed by pGADT7-Rec2 did not interact with the p53 response element, and the p53 protein did not bind to any test bait sequences except for its target p53 bait sequence.

It is not entire clear why Y187 [bait vector+pGADT7-Rec-2] seemed to grow stronger than Y187 [bait vector+pGAD-Rec2-53]; perhaps the protein expressed by pGADT7-Rec2 can interact weakly with the T_R-binding protein thus activate HIS3 expression, but the p53 prey protein which is larger in size can not bind unless there is a specific target sequence that brings the GAL4 AD close to T_R-binding protein.

pGAD-Rec2-53 does not contain the T7 RNA polymerase promoter or the HA tag, nor does it share any homology with the BD SMART III or CDS III oligonucleotides. Those four features will be expressed as amino acids in RXR prey vector and pGADT7-Rec2. Because of these differences, Y187 [bait vector+pGAD-Rec2-53] may not be a perfect negative control, and Y187 [bait vector+pGADT7-Rec2] is not a perfect negative control, either, due to size

differences. pHIS2-containing co-transformants are not good negative controls, either, because they contains DNA sequence between *Eco*R I and *Mlu* I (Figure 2-2) that is replaced by FARE or mutated bait sequences to make pFARE or pMB-1 to pMB-4. However, if the growth of a co-transformant is stronger than that of a co-transformant containing the same bait vector and pGADT7-Rec-2, it may indicate the one-hybrid interaction between the test prey protein and the test bait.

With the above caveats, we can make the following conclusions regarding the binding of SgRXR to FARE and the mutated baits, based on the results from both streaking methods. The grid streaking method showed very weak binding of SgRXR to MB-2 or MB-4 mutated bait sequences, and the section streaking method showed no binding at all between SgRXR and MB-4, making it unlikely that SgRXR specifically interacts with either sequence. In contrast, interactions of SgRXR with the wild-type FARE were much stronger, but not nearly as strong as the binding of p53 protein to its response element. The grid method did not allow any conclusions for the interactions between SgRXR and the bait sequence in pMB-1, pMB-3 or the empty pHIS2 vector, because the minimal 3-AT concentration to remove background expression were extremely high for those bait vectors (> 300 mM). However, the more accurate section method showed that SgRXR interacted more specifically with MB-1, MB-3 or the bait sequence in pHIS2. Interestingly, it appears that SgRXR had stronger binding to MB-3 than to the native FARE or mutant MB-1, while at best very weak binding to MB-2 and no binding to MB-4.

The streaking method is one of the possible tools in detecting protein-DNA binding, but there are some features which will hinder the success in detection. The cell growth is affected by the amount of cells applied on the selection plate, the viability of the co-transformants, the consistency of the cell application and possible non-specific binding of the endogenous yeast transcription factors, the correct expression of prey proteins, and other influences. Thus, the one-hybrid analysis performed here may not reflect real interactions, and some other methods should be used to confirm these conclusions.

CHAPTER 5:
GENERAL DISCUSSION

5.1 Yeast one-hybrid strategy

Our initial goal was to screen a cDNA library for a potential FAAR which binds FARE located in the promoter region of the FABP gene thus regulates the gene expression. The yeast one-hybrid assay is a method to detect protein-DNA interaction, or to screen DNA-binding proteins from a cDNA library. The detection of DNA-binding proteins requires the co-existence of a bait vector and a prey vector in yeast cells. The bait vector can be constructed by cloning tandem copies of target DNA sequence into a reporter vector containing the *HIS3* nutritional reporter gene. The prey vector can be constructed by *in vivo* cloning cDNA into a AD cloning vector, so that the potential DNA-binding proteins are expressed as fusion proteins with the GAL4 activation domain (AD). If a DNA-binding protein interacts with the target DNA sequence, the transcription of the *HIS3* gene will be activated, and the histidine auxotroph yeast strain will grow on minimal media lacking histidine.

All cells in this system express His3 protein at very low background level, resulting in some growth even without the specific activation. It is possible to reduce or eliminate background expression by the enzyme inhibitor 3-AT, but this will also reduce histidine production in positive colonies. Thus, the effective use of this technique requires that the interaction between the DNA and binding protein is strong, so that the resulting expression of the *HIS3* gene is much

stronger than the low-level background expression of the gene. If the candidate FARE-binding proteins are nuclear receptors as proposed, they should be present in very low amounts in cDNA library, which will make library screening very difficult, as many colonies need to be screened before a positive clone can be found. If the regulation of FABP gene expression requires the binding of a heterodimer to FARE, there will be even less chance to screen them out. During the course of this thesis it became apparent that limitations of the experimental system, especially with respect to reducing the background expression, were severe, and that it would be difficult to identify weak interactions. Hence, although a good quality cDNA library was synthesized, we decided to forgo screening of the library, and focus on using this system to investigate the suspected interactions between RXR and FARE. This required the cloning of RXRs into the AD cloning vector to express GAL4 fusion proteins in yeast cells.

5.1.1 cDNA library construction from locust flight muscle

In this study, high quality total RNA was isolated from locust flight muscle and used successfully for library construction. Two predominant bands for total RNA with the sizes of 2.2 kb and 1.9 kb were seen on denaturing formaldehyde agarose gels, representing the large and small subunits of ribosomal RNA. The quality of the cDNA library synthesized from total RNA was confirmed by agarose gel electrophoresis. While the library was not used for screening by the yeast one-hybrid method, it served as source for the identification and cloning of RXR.

5.1.2 Cloning of *Schistocerca* RXRs

Two RXR isoforms (SgRXR-L and SgRXR-S) were detected in the flight muscle of desert locust, and their coding sequences are highly homologous with those identified in the migratory locust (LmRXR-L and LmRXR-S). At the RNA level, the coding sequences of SgRXR-L and SgRXR-S exhibit 94.98% and 94.70% nucleotide identity to those of the LmRXRs. At the protein level, their amino acid sequences are almost identical to their homologues of the migratory locust except a single amino acid deletion in both isoforms and a single amino acid substitution in SgRXR-S compared to LmRXR-S.

Because of the identical sequence of both isoforms, it can be concluded that they are encoded by a single gene and the result of alternative splicing of the primary transcript. The difference between the short and long isoforms of SgRXR is a 66 bp deletion, resulting in the loss of 22 amino acids after the first 30 amino acids of the ligand binding domain. When aligned to the amino acid sequences of vertebrate RXRs, it becomes apparent that the deleted portion of the short isoform of locust RXRs may be important for ligand binding and the ligand-binding-induced conformational change. Thus, the short isoform may not have ligand-dependent functions while LmRXR-L or SgRXR-L could be ligand-modulated nuclear receptors (Hayward et al., 2003).

5.1.3 Interaction of SgRXRs and FARE

The full-length coding sequence for each SgRXR isoform was *in vivo* cloned in frame into the AD cloning vector in yeast, and the prey vectors expressed SgRXR fused with the GAL4 activation domain. The interaction of the

fusion protein with wild type FARE and with mutated baits was tested by the one-hybrid assay. Four mutated bait sequences were prepared by partially altering the wild type fatty acid response element (5' GG AGT GGT agt TCC CAT CC 3'). The replacement of GGT by CAA within the inner part of 5' half site resulted in mutated bait sequence 1(MB-1), whereas the replacement of TCC with AAG within the inner part of 3' half site gave rise to MB-2; and the same nucleotide changes as above within the inner parts of both half sites resulted in MB-3. The replacement of GG AGT by AA TTT at the outer part of the 5' half site and CAT CC by CTA TT at the outer part of the 3' half site resulted MB-4, which is again a palindromic sequence, however without resemblance to the consensus sequence.

All one-hybrid assays were carried out for both isoforms. In most cases, similar results were seen for each SgRXR isoforms. If indeed SgRXR-L is a ligand-modulated nuclear receptor and SgRXR-S is ligand-independent, as proposed for LmRXRs (Hayward et al., 2003), it is not surprising that both isoforms have similar behaviours without the presence of the ligand. Yet, subtle differences between the long and short isoforms were observed. It appears that SgRXR has only very weak interaction with MB-2, and no interaction with MB-4. SgRXRs interact with the wild type FARE and MB-1, but not as strong as the binding of p53 with its response element. Interestingly, the interaction of SgRXRs with MB-3 is stronger than that with the native FARE or mutant MB-1, perhaps providing an insight into the optimal consensus sequence for this element.

On the other hand, high-affinity DNA binding of nuclear receptors and the subsequent transcriptional activation may require the presence of a ligand and the ligand-induced homo- or heterodimer formation (FAAR and RXR, according to our proposed model). Stronger one-hybrid interaction may need the presence of fatty acids. Retinoic acid, on the other hand, is unlikely to be involved, as it is generally assumed that this molecule is not present in insects. Nevertheless, we can make some assumption based on the experimental results: SgRXRs interacts with FARE and MB-1 but not MB-2 or MB-4, which may indicate that the 3' half site of FARE is important for the ligand independent binding of SgRXR, while the 5' half site may be the binding site for FA-induced nuclear receptor. However, the nucleotide changes at the both site in MB-3 makes the both sites symmetric and may favor the binding of SgRXRs to both sites. It is possible that the potential FAAR is SgRXR-L, which might be activated by its ligand fatty acids, and the induced conformational change favors its binding with one of the SgRXR isoform to 5' half site of FARE thus regulate the FABP gene expression. Without the induction of its ligand, the conformation of SgRXR-L may not optimal for binding to the 5' half site, and therefore only weak interactions between SgRXRs and MB-2 were observed.

The current model for receptor-FARE interactions assumes that long chain fatty acids are the natural ligand for the receptor. As discussed before, long chain fatty acids have been shown to bind to vertebrate RXR (Harmon et al., 1995), and thus they may well be a natural ligand for locust RXR. It is important to note, however, that fatty acid mediated gene activation should occur only at high fatty

acid concentrations. Unlike hormone receptors, which bind scarce ligands with high affinity, FAAR can be expected to bind to its ligand more weakly, so that activation occurs only when fatty acid is abundant. Indeed, although free fatty acid should be present in all yeast cells, its concentration may not be high enough to induce the binding of FAAR to FARE. The nucleotide changes in MB-3 may favor the binding of both isoforms in a ligand-independent fashion, and therefore stronger interactions were observed.

5.2 Future research

The results of the present study can help to better understand the molecular mechanism by which the fatty acids regulate FABP gene expression, but further research is needed to completely elucidate the mechanism at the molecular level.

The ligand-binding of a nuclear receptor often induces a conformational change which initiates the formation of a homo- or heterodimer and favors its binding to the response element of the target gene. Therefore, the addition of ligand to the growth media of the yeast co-transformants may provide greater clarity of the ligand-dependend interactions of the isoforms of RXR.

A more direct approach to study the interactions of FARE and RXR requires the expressed proteins. In the yeast one-hybrid analysis, SgRXRs are expressed as fusions with a GAL4 AD and HA tag. Both isoforms of SgRXR can be extracted from yeast, and, if necessary, purified with anti-HA antibodies. Then, the SgRXR-FARE interaction detected by one-hybrid analysis can be

confirmed *in vitro* with wild type FARE and mutated bait sequences by electrophoretic mobility-shift assay (EMSA), in the presence or absence of fatty acids. Alternatively, protein binding studies can be employed to investigate the interaction of SgRXR with radio-labelled FARE. One caveat is that our yeast cells express a fusion protein, and the additional GAL4 activation domain and HA tag might alter the conformation of RXR and block the ligand or DNA binding. Thus, *in vitro* transcription and translation of both isofoms of SgRXR is desirable to yield native RXR for the above outlined binding studies.

An advantageous approach could be to express RXR in insect cells, such as *Spodoptera* sf9 cells. Because our experimental model is an insect species, the proteins expressed in the insect cell line are more likely to be in their native conformations. and the results may represent biologically significant interactions. sf9 cells could also be cotransfected with an SgRXR expression vector and a reporter vector containing target DNA sequence. The interaction of SgRXR with the target DNA sequence can activate the transcription of the reporter gene, such as GFP, that can be easily detected and quantified.

If no interactions between FARE and RXR can be verified by these experiments, it may seem necessary to screen the cDNA library of the locust flight muscle established here for potential FAARs other than RXRs. However, the screening may not be representative because regulatory proteins are normally present only at very low levels (Ammerer 1983), and weak but relevant interactions may not be detectable. An alternative approach could be to

specifically isolate the receptors by affinity techniques based on the binding to the response element.

Streptavidin-coupled paramagnetic beads (Dynabeads) provide a rapid and sensitive method to isolate and purify sequence-specific DNA-binding proteins. The biotin-labeled target DNA sequence, i.e., the FARE, is coupled to the beads and functions as "fishing rod" to pull the DNA-binding proteins out of the nuclear extract. A strong magnet can be used to separate the beads from the extract. After washed off the beads, the isolated potential DNA-binding proteins can be re-tested for their DNA-binding ability by EMSA, or further purified and analyzed. Bands excised from SDS-PAGE gels may be used to positively identify the proteins by sensitive proteomic techniques, such as MALDI (matrix-assisted laser desorption/ionization) mass spectrometry.

5.3 Conclusion

In this study, a cDNA library of locust flight muscle was constructed. The full- length coding sequences of two RXR isoforms were successfully isolated from the flight muscle of desert locust *Schistocerca gregaria,* and cloned into the AD cloning vector to construct prey vectors through homologous recombination in yeast cells. The bait vectors containing the wild-type FARE sequence and partially altered sequences were constructed by the method of restriction digesting and ligation. At last, the presence and the specificity of the interaction between SgRXRs and FARE were verified by yeast one-hybrid analysis. This research project provides insights for future studies on revealing the molecular mechanism of FABP gene regulation by fatty acids.

REFERENCE LIST

Ammerer, G. 1983. Expression of genes in yeast using the ADCl promoter. Methods in Enzymology 101, 192-201.

Becker, D.M. & Guarente, L. 1991. High-efficiency transformation of yeast by electroporation. Methods in Enzymology 194, 182-185.

Beenakkers, A.M., van der Horst, D.J., & van Marrewilk, W.J.A. 1985. Insect lipids and lipoproteins, and their role in physiological processs. Progress in Lipid Research 24, 19-67.

Bouton, D., Escriva, H., de Mendonca, R.L., Glineur, C., Bertin, B., Noel, C., Robinson-Rechavi, M., de Groot, A., Cornette, J., Laudet, V. & Pierce, R.J. 2005. A conserved retinoid X receptor (RXR) from the mollusk Biomphalaria glabrata transactivates transcription in the presence of retinoids. Journal of Molecular Endocrinology 34, 567-582.

Braus, G., Luger, K., Paravicini, G., Schmidheini, T., Kirschner, K. & Hutter, R. 1988. The role of the TRP1 gene in yeast tryptophan biosynthesis. Journal of Biological Chemistry 263, 7868-7875.

Burke, D., Dawson, D. & Stearns, T. 2000. Methods in yeast genetics, Cold Spring Harbor Laboratory Press, New York.

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

Cheskis, B. & Freedman, L.P. 1994. Ligand modulates the conversion of DNA-bound vitamin D3 receptor (VDR) homodimers into VDR-retinoid X receptor heterodimers. Molecular and Cellular Biology 14, 3329-3338.

Chien, C., Bartel, P., Sternglanz, R. & Fields, S. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proceedings of the National Academy of Sciences 88, 9578-9582.

Clarke, L. & Carbon, J. 1980. Isolation of a yeast centromere and constructoin of functoinal small circular chromosomes. Nature 287, 504-509.

Collingwood, T.N., Butler, A., Tone, Y., Clifton-Bligh, R.J., Parker, M.G. & Chatterjee, V.K.K. 1997. Thyroid hormone-mediated enhancement of heterodimer formation between thyroid hormone receptor beta and retinoid X receptor. Journal of Biological Chemistry 272, 13060-13065.

Crabtree, B. & Newsholme, E.A. 1975. Comparative aspects of fuel utilization and metabolism by muscle. Insect Muscle, ed. P.N.R. Usherwood, Academic Press, London, 405-491.

Das, D.K., Barua, P.K. & Jones, R.M. 1991. Release of fatty acid-binding protein from ischemic-reperfused rat heart and its prevention by mepacrine. Biochimica et Biophysica Acta 1168, 307-314.

Desvergne, B. & Wahli, W. 1999. Peroxisome proliferator-activated receptors: nuclear control of metabolism. Endocrine Reviews 20, 649-688.

DiRenzo, J., Soderstrom, M., Kurokawa, R., Ogliastro, M., Ricote, M., Ingrey, S., Horlein, A., Rosenfeld, M. & Glass, C. 1997. Peroxisome proliferator-activated receptors and retinoic acid receptors differentially control the interactions of retinoid X receptor heterodimers with ligands, coactivators, and corepressors. Molecular and Cellular Biology 17, 2166-2176.

Duplus, E., & Forest, C. 2002. Is there a single mechanism for fatty acid regulation of gene transcription. Biochemical Pharmacology 64: 893–901.

Duplus, E., Glorian, M. & Forest, C. 2000. Fatty acid regulation of gene transcription. Journal of Biological Chemistry 275, 30749-30752.

Fink, G.R. 1964. Gene-enzyme relations in histidine biosynthesis in yeast. Science 146, 525-527.

Ghosh, S., Selby, M.J. & Peterlin, B.M. 1993. Synergism between Tat and VP16 in trans-activation of HIV-1 LTR. Journal of Molecular Biology 234, 610-619.

Giguere, V. 1999. Orphan nuclear receptors: from gene to function. Endocrine Reviews 20, 689-725.

Gilbert, L.I. & Chino, H. 1974. Transport of lipids in insects. Journal of Lipid Research 15, 439-456.

Glass, C. 1994. Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocrine Reviews 15, 391-407.

Glatz, J.F.C. & van der Vusse,G. J. 1996. Cellular fatty acid-binding proteins: Their function and physiological significance. Progress in Lipid Research 35, 243-282.

Guarente, L. 1984. Yeast promoters: Positive and negative elements. Cell 36, 799-800.

Gunge, N. 1983. Yeast DNA plasmids. Annual Review of Microbiology 37, 253-276.

Harmon, M.A., Boehm, M.F., Heyman, R.A. & Mangelsdorf, D.J. 1995. Activation of mammalian retinoid X receptors by the insect growth regulator methoprene. Proceedings of the National Academy of Sciences 92, 6157-6160.

Haunerland, N.H. 1994. Fatty acid binding proteins in locust and mammalian muscle: Comparison of structure, function and regulation. Comparative Biochemistry and Physiology 109B, 199-208.

Haunerland, N.H. 1997. Transport and utilization of lipids in insect flight muscles. Comparative Biochemistry and Physiology 117B, 475-482.

Haunerland, N.H., Spener, F. 2004a. Properties and physiological significance of fatty acid binding proteins. Advances in Molecular and Cell Biology 33, 99-123(article in press).

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

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

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

Hinnen, A., Hicks, J.B. & Fink, G.R. 1978. Transformation of yeast. Proceedings of the National Academy of Sciences 75, 1929-1933.

Horecka, J. & Sprague, G.F.J. 2000. Use of imidazoleglycerolphosphate dehydratase(His3) as a biological reporter in yeast. Methods in Enzymology 326, 107-119.

Hwang, S.T., Urizar, N.L., Moore, D.D. & Henning, S.J. 2002. Bile acids regulate the ontogenic expression of ileal bile acid binding protein in the rat via the farnesoid X receptor. Gastroenterology 122, 1483-1492.

Issemann, I., Prince, R., Tugwood, J. & Green, S. 1992. A role for fatty acids and liver fatty acid binding protein in peroxisome proliferation. Biochemical Society Transactions 20, 824-827.

Ito, H., Fukuda, Y., Murata, K. & Kimura, A. 1983. Transformation of intact yeast cells treated with alkali cations. Journal of Bacteriology 153, 163-168.

Jump, D.B. & Clarke, S.D. 1999. Regulation of gene expression by dietary fat. Annual Review of Nutrition 19, 63.

Kakizawa, T., Miyamoto, T., Kaneko, A., Yajima, H., Ichikawa, K. & Hashizume, K. 1997. Ligand-dependent heterodimerization of thyroid hormone receptor and retinoid X receptor. Journal of Biological Chemistry 272, 23799-23804.

Kammer, A.E. & Heinrich, B. 1978. Insect flight metabolism. Advances in Insect Physiology 13, 133-228.

Laudet, V. & Gronemeyer, H. 2002. The nuclear receptor factsbook. Acdemic Press, London, UK.

Laudet, V. 1997. Evolution of the nuclear receptor superfamily: Early diversification from an ancestral orphan receptor. Journal of Molecular Endocrinology 19, 207-226.

Mahadevan, S. & Struhl, K. 1990. Tc, an unusual promoter element required for constitutive transcription of the yeast HIS3 gene. Molecular and Cellular Biology 10, 4447-4455.

Mangelsdorf, D.J. & Evans, R.M. 1995. The RXR heterodimers and orphan receptors. Cell 83, 841-850.

Murray, A. & Szostak, J. 1983. Pedigree analysis of plasmid segregation in yeast. Cell 34, 961-970.

Noy, N. 2000. Retinoid-binding proteins: Mediators of retinoid action. Biochemical Journal 348, 481-495.

Ockner, R.K., Manning, J.A., Poppenhausen, R.B. & Ho, W.K.L. 1972. A binding protein for fatty acids in cytosol of intestinal musosa, liver, myocardium, and other tissues. Science 177, 56-58.

Owen, G.I. & Zelent, A. 2000. Origins and evolutionary diversification of the nuclear receptor superfamily. Cellular and Molecular Life Sciences 5, 809-827.

Panzeri, L. & Philippsen, P. 1982. Centromeric DNA from chromosome VI in *Saccharomyces cerevisiae* strains. EMBO Journal 1, 1605–1611.

Perera, S.C., Zheng, S., Feng, Q.L., Krell, P.J., Retnakaran, A. & Palli, S.R. 2005. Heterodimerization of ecdysone receptor and ultraspiracle on symmetric and asymmetric response elements. Archives of Insect Biochemistry and Physiology 60, 55-70.

Poirer, H., Niot, I., Monnot, M.C., Braissant, O., Meunier-Durmort, C., Costet, P., Pineau, T., Wahl, W., Wilson, T.M. & Besnard, P. 2001. Differential involvement of peroxisome-proliferator-activated receptors alpha and delta in fibrate and fatty-acid-mediated induction of the gene encoding liver fatty-acid-binding protein in the liver and the small intestine", Biochemical Journal 355, 481-488.

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

Rastinejad, F. 2001. Retinoid X receptor and its partners in the nuclear receptor family. Current Opinion in Structural Biology 11, 33-38.

Renaud, J.P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H. & Moras, D. 1995. Crystal structure of the RAR-γ ligand-bindind domain bound to all-trans retinoic acid. Nature 378, 681-689.

Robinson-Rechavi, M., Garcia, H.E. & Laudet, V. 2003. The nuclear receptor superfamily. Journal of Cell Science 116, 585-586.

Rose, M.D. & Broach, J.R. 1991. Cloning genes by complementation in yeast. Methods in Enzymology 194, 195-230.

Ryan, R., Prasad, S., Henriksen, E., Wells, M. & Law, J. 1986. Lipoprotein interconversions in an insect, *Manduca sexta*. Evidence for a lipid transfer factor in the hemolymph. Journal of Biological Chemistry 261, 563-568.

Schaap, F.G., Van der Vusse, G.J. & Glatz, J.F.C. 1998. Fatty acid-binding proteins in the heart. Molecular and Cellular Biochemistry 180, 43-51.

Schachtrup, C., Emmler, T., Bleck, B., Sandqvist, A. & Spener, F. 2004. Functional analysis of peroxisome-proliferator-responsive element motifs in genes of fatty acid-binding proteins. Biochemical Journal 382, 239-245.

Sikorski, R.S. & Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics 122, 19-27.

Stargell, L. & Struhl, K. 1996. Mechanisms of transcriptional activation in vivo: Two steps forward. Trends in Genetics 12, 311-315.

Struhl, K. 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. Molecular and Cellular Biology 6, 3847-3853.

Struhl, K. 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell 49, 295-297.

Thornton, J.W. 2003. Nonmammalian nuclear receptors: Evolution and endocrine disruption. Pure and Applied Chemistry 75, 1827-1839.

Thukral, S. K., Blain, G. C., Chang, K. K. & Fields, S. 1994. Distinct residues of human p53 implicated in binding to DNA, simian virus 40 large T antigen, 53BP1, and 53BP2. Molecular and Cell Biology 14, 8315-8321.

Tontonoz, P., Hu, E., Graves, R.A., Budavari, A.I. & Spiegelman, B.M. 1994. mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. Genes and Development 8, 1224-1234.

van der Horst, D.J. 1990. Lipid transport function of lipoproteins in flying insects. Biochemica et Biophysica Acta 1047, 195-211.

van der Horst, D.J., van Marrewijk, W.J.A. & Diederen, J.H.B. 2001. Adipokinetic hormones of insect: release, signal transduction, and responses. International Review of Cytology 211, 179-240.

van der Vusse,G.J., van Blisen, M., Glatz, J.F.C., Hasselbaink, D.M. & Luiken, J.J. 2002. Critical steps in cellular fatty acid uptake and utilization. Molecular and Cellular Biochemistry 239, 9-15.

van der Vusse, G.J., Glatz, J.F., Stam, H.C. & Reneman, R.S. 1992. Fatty acid homeostasis in the normoxic and ischemic heart. Physiological Reviews 72, 881-940.

van der Vusse,G. J., van Bilsen, M. & Glatz, J.F.C. 2000. Cardiac fatty acid uptake and transport in health and disease. Cardiovascular Research 45, 279-293.

Veerkarnp, J.H. & Maatman, R.G.H.J. 1995. Cytoplasmic fatty acid-binding proteins: Their structure and genes. Progress in Lipid Research 34, 17-52.

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

Weisiger, R.A. 1996. Cytoplasmic transport of lipids: role of binding proteins. Comparative Biochernistry and Physiology 115B, 319-331.

Wu, Q., Andolfatto, P. & Haunerland, N. H. 2001. Cloning and sequence of the gene encoding the muscle fatty acid binding protein from the desert locust, *Schistocerca gregaria*. Insect Biochemistry and Molecular Biology 31, 553-562.

Wu, Q., Chang, W., Rickers-Haunerland, J., Higo, T., & Haunerland, N.H. 2002. Characterization of a new fatty acid response element that controls the expression of the locust muscle FABP gene. Molecular and Cellular Biochemistry 239, 173-180.

Wu, Q. & Haunerland, N.H. 2001. A novel fatty acid response element controls the expression of the flight muscle FABP gene of the desert locust, *Schistocerca gregaria*. European Journal of Biochemistry 268, 5894-5900.

Yamomoto, Y. & Niki, E. 1990. Role of antioxidants in lipid peroxidation, CRC Press, Boca Raton.

Yao, T., Forman, B.M., Jiang, Z., Cherbas , L., Chen, J., McKeown, M., Cherbas, P. & Evans, R.M. 1993. Functional ecdysone receptor is the product of EcR and Ultraspiracle genes. Nature 366, 476-479.

Yao, T., Segraves, W., Oro, A., McKeown, M. & Evans, R. 1992. Drosophila ultraspiracle modulates ecdysone receptor function via heterodimer formation. Cell 71, 63-72.

Zhang, J. & Haunerland, N.H. 1998. Transcriptional regulation of FABP expression in flight muscle of the desert locust, *Schistocerca gregaria*. Insect Biochemistry and Molecular Biology 28, 683-691.

Zimmerman, A.W. & Veerkamp, J.H. 2002. New insights into the structure and function of fatty acid-binding proteins. Cellular and Molecular Life Science 59, 1096-1116.