Isolation of DNA Sequences with Homology to a Degenerate FaRP Oligonucleotide from the Crayfish *Procambarus clarkii*

Amy E. Peaire

Department of Biological Sciences

(Submitted in partial fulfilment of the requirements for the degree of Master of Science)

Brock University

St.Catharines, Ontario, CANADA

December, 1995

ABSTRACT

The neuropeptide FMRFamide with the sequence Phe-Met-Arg-Phe-amide was originally isolated in the clam *Macrocallista nimbosa* (Price and Greenberg, 1977). Since its discovery, a large family of FMRFamide-related peptides termed FaRPs have been found to be present in all major animal phyla with functions ranging from modulation of neuronal activity to alteration of muscular contractions. However, little is known about the genetics encoding these peptides, especially in invertebrates.

As FaRP-encoding genes have yet to be investigated in the invertebrate *Malacostracean* subphylum, the isolation and characterization of FaRP-encoding DNA and mRNA was pursued in this project. The immediate aims of this thesis were: (1) to amplify mRNA sequences of *Procambarus clarkii* using a degenerate oligonucleotide primer deduced from the common amino acid sequence of isolated *Procambarus* FaRPS, (2) to determine if these amplification products encode FaRP gene sequences, and (3) to create a selective cDNA library of sequences recognized by the degenerate oligonucleotide primer. The polymerase chain reaction - rapid amplification of cDNA ends (PCR-RACE) is a procedure in which a single gene-specific primer is used in conjunction with a generalized 3' or 5' primer to amplify copies of the region between a single point in the transcript and the 3' or 5' end of cDNA of interest (Frohman et al., 1988). PCR-RACE reactions were optimized with respect to primers used, buffer composition, cycle number, nature of genetic substrate to be amplified, annealing, extension and denaturation temperatures and times, and use of reamplification procedures. Amplification products were cloned into plasmid vectors and recombinant products were isolated, as were the recombinant plaques formed in the selective cDNA library. Labeled amplification products were hybridized to recombinant

bacteriophage to determine ligated amplification product presence. When sequenced, the five isolated PCR-RACE amplification products were determined not to possess FaRP-encoding sequences. The 200bp, 450bp, and 1500bp sequences showed homology to the *Caenorhabditis elegans* cosmid K09A11, which encodes for cytochrome P450; transfer-RNA; transposase; and tRNA-Tyr, while the 500bp and 750bp sequences showed homology with the complete genome of the *Vaccinia* virus. Under the employed amplification conditions the degenerate oligonucleotide primer was observed to bind to and to amplify sequences with either 9 or 10bp of 17bp identity.

The selective cDNA library was observed to be of extremely low titre. When library titre was increased, white plaques were isolated. Amplification analysis of eight isolated λgt11 sequences from these plaques indicated an absence of an insertion sequence.

The degenerate 17 base oligonucleotide primer synthesized from the common amino acid sequence of isolated *Procambarus* FaRPs was thus determined to be non-specific in its binding under the conditions required for its use, and to be insufficient for the isolation and identification of FaRP-encoding sequences. A more specific primer of longer sequence, lower degeneracy, and higher melting temperature (T_m) is recommended for further investigation into the FaRP-encoding genes of *Procambarus clarkii*.

ACKNOWLEDGEMENTS

I am extremely grateful to Dr. A.J. Castle and Dr. A.J.Mercier, and wish to thank them for their great supervision, helpfulness, and enthusiasm for this project. I appreciate the helpful advice of my committee members Dr. Y.Haj-Ahmad and Dr.A.Bown.

I would also like to thank my husband Rene and my parents James and Wanda for their constant support during this project. Without them this thesis would not have been completed. Finally, I wish to thank Helen for the many long hours spent helping me with the scans.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APS	ammonium persulfate
CCK	cholecystokinin
cDNA	complementary deoxyribonucleic acid
CIP	calf intestinal alkaline phosphatase
CNS	central nervous system
DEPC	diethyl pyrocarbonate
DF_2	DRNFLRFamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	dinucleotide triphosphate
dscDNA	double stranded complementary deoxyribonucleic acid
DSK	drosulfakinin
EDTA	ethylenediaminetetraacetic acid
FaRP	FMRFamide related peptide
FLI	FMRFamide-like immunoreactivity
HPLC	high performance liquid chromatography
LSK	leukosulfakinin
mRNA	messenger ribonucleic acid
NF_1	NRNFLRFamide
O.D.	optical density
PCR	polymerase chain reaction
PCR-RACE	polymerase chain reaction - rapid amplification of cDNA ends

PEG	polyethylene glycol
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
SCPb	small cardioactive peptide b
SDS	sodium dodecyl sulfate
SSC	saline - sodium citrate
sscDNA	single stranded complementary deoxyribonucleic acid
SSPE	saline - sodium phosphate - EDTA
TAE	tris - acetate - EDTA
TBE	tris - borate - EDTA
TCA	trichloroacetic acid
TE	tris - EDTA
TEMED	N,N,N',N', - tetramethylethylenediamine
T_m	melting temperature
tRNA	transport ribonucleic acid
UV	ultraviolet

INTRODUCTION

The neuropeptide FMRFamide (Phe-Met-Arg-Phe-NH_2) was originally identified in the clam *Macrocallista nimbosa* (Price and Greenberg, 1977). Immunoreactivity to FMRFamide-specific antisera has since been reported in all major animal phyla, while several FMRFamide-related peptides, or FaRPs, have been isolated from arthropods (Trimmer et al., 1987), coelenterates (Grimmelikhuijzen and Grall, 1985) and chordates (Dockray et al., 1983). FaRPs are generally defined as any peptide that can be found by employing an assay for FMRFamide (Price and Greenberg, 1989). They have been observed to act as neurohormones, neurotransmitters and neuromodulators in a wide range of tissues in various organisms, with functions ranging from regulation of the cardiovascular system to enhancement of muscular contractions, to modification of neuronal activity.

Recently, two new FaRPs known as DF_2 (DRNFLRFamide) and NF_1 (NRNFLRFamide) have been isolated from crayfish pericardial organs (Mercier et al., 1993). These peptides were found to be cardioexcitatory and to augment synaptic transmission at a neuromuscular synapse.

In some invertebrate organisms studied to date, single genes have been found to encode for FMRFamide-related peptides (Linacre et al., 1990), (Schaefer et al., 1985). This is not the case in *Drosophila melanogaster*, however, where a second gene coding for drosulfakinins, peptides with structural similarity to FaRPs, has been isolated (as reviewed by Duve et al., 1992), and the presence of a third gene encoding at least one FaRP has been postulated (Nichols, 1992). In the invertebrate FaRP genes, encoded cleavage sites are of varying similarity between species, and spacer sequences, which conform somewhat in length and frequency of occurrence, cannot be said to have any clear function or ancestral form. The lack of a representative FaRP gene from

the *Malacostracean* subphylum limits our understanding of invertebrate FaRP genes and has prevented the establishment of possible generalized invertebrate gene traits.

Recent investigations have determined length and abundance of invertebrate FaRP gene transcripts. Invertebrate FaRP-encoding mRNA has been characterized for the fruit fly *Drosophila* (Chin et al., 1990), the sea slug *Aplysia californica* (Schaefer et al., 1985), the nematode *Caenorhabditis elegans* (Rosoff et al., 1992) and the snail *Lymnaea stagnalis* (Linacre et al., 1990). In these studies, different length and abundance of mRNA transcripts were found in the different organisms. The findings of both Lehman et al., (1984) and Schaefer et al., (1985) in *Aplysia* suggest that different mRNAs are transcribed in varying amounts according to the tissue in which they are produced. As well, there is speculation that multiple RNA transcripts may arise from a single gene by a number of mechanisms including alternate initiation site recognition, alternate splicing and alternate polyadenylation, reflecting properties of the FaRP-encoding gene itself.

Determination of the genetic sequences coding for FaRPs in the *Malacostracean Procambarus clarkii* would contribute much information to the knowledge of invertebrate FaRP DNA and mRNA in general. Knowledge of the number of *Procambarus* FaRP genes could aid in the construction of phylogenetic relationships between different invertebrate species and in the establishment of genetic homology between individual peptides as well as between peptide families. Such knowledge could also provide information regarding the possible existence and structure of one or more common FaRP precursor genes. Identification of transcription sites for FaRP mRNA would contribute to the understanding of the means of communication by neuropeptides in the organism, suggesting putative sites of action, pathways of neuropeptide movement, and control mechanisms. In addition, the size, number and sites of transcription of

FaRP-encoding RNA would all contribute to the characterization of the FaRPs within the crayfish and provide much needed information concerning the genetic properties and phylogenetic relationships of FaRPs in invertebrates.

Genetic studies of crayfish FaRPs were investigated based upon the known amino acid sequences of the two identified *Procambarus* FaRP sequences, NF_1, and DF_2, as no portions of FaRP gene sequences or adjacent sequences are known that could be used to locate *Procambarus* FaRP gene sequences.

Thus, a common amino-acid sequence from the common NFLRF(Gly) portion of NF_1 and DF_2 was used to create a 16-fold degenerate 17-base oligonucleotide with binding capability to the coding strand of NF_1 and DF_2-encoding DNA. This mixed oligonucleotide was used as a primer in PCR amplification of FaRP sequences, as previous experiments using the common gene-isolating technique of low stringency genomic library screening resulted in the isolation of false positives when using this short oligonucleotide as a probe for FaRP sequences (Peaire, 1993).

The purposes of this study were three-fold. The first purpose was the optimization of PCR-RACE reactions for the amplification of *Procambarus clarkii* cDNA sequences with a degenerate oligonucleotide primer designed to recognize the two sequenced *Procambarus* FaRPs. This procedure involved selection and/or variation of primers used, of buffer composition, cycle number, the nature of the genetic substrate to be amplified, annealing, extension, and denaturation temperatures and times, and of the use of re-amplification procedures. The second purpose of the study was to determine if the amplification products encoded FaRP gene sequences. Amplification products were cloned into plasmid vectors for easy amplification and isolation of DNA sequences. Colony lift hybridizations were performed to determine if the cloned sequences

were actual PCR-RACE amplification products. Product sequences were also analyzed to determine stringency of primer binding and for comparison with known sequences. The final purpose of this study was to use the degenerate oligonucleotide to prime second strand double-stranded cDNA (dscDNA) synthesis. These sequences were then ligated to λgt11 vectors in an attempt to create a selective cDNA.

LITERATURE REVIEW

1.1 Neuropeptides

Neuropeptides are found in the nervous systems of all animals (Sherman et al., 1989) as well as in unicellular organisms (Scharrer, 1990). These peptides are produced by both neural and non-neural tissues. It is well established that many neuropeptides are localized in nerve-cell bodies and in nerve terminals (Kruk and Pyrock, 1979). In addition to being found in the central nervous system (CNS), several neuroactive peptides are found in the non-neural cells of the gastro-intestinal tract (Kruk and Pyrock, 1979).

The molecular diversity of neuropeptides is great, with as many as 100-200 different peptide molecules serving as neuronal messengers in one animal species (Kandel, 1983), (Joosse, 1986).

In humans, there are at least 36 neuropeptides in the nervous system (Altman, 1989). Overall, neuropeptides appear to be particularily important in regulating feelings and drives, such as pain, pleasure, hunger, thirst and sex (Alberts et al., 1989).

1.1.1 Neuropeptide functions

Neurohormones, neurotransmitters, and neuromodulators have traditionally been distinguished from each other in regard to their sites of action, their duration of effect, and the type of effect produced.

Neurotransmitters are classically defined as chemicals that travel across the synaptic cleft from a neuron to another cell, which as a result of the arrival of the neurotransmitter, responds by becoming either excited or inhibited with respect to a certain activity or process (Arms and

Camp, 1989). Generally, neurotransmitters are low-molecular-weight molecules, synthesized by nerve cells, that underlie a form of cell-cell interaction in the nervous system characterized as synaptic transmission. They function over relatively small and discrete areas for brief periods (milliseconds to seconds) (Barker, 1977).

Veca and Dreisbach (1988) state that in order for a compound to be classified as a true neurotransmitter, four criteria must be met: (1) the substance must be synthesized by the neuron, (2) it must be released by the neuron in sufficient amounts to exhibit an effect on another neuron or effector organ, (3) exogenous application in appropriate quantities must mimic the action of the endogenously released compound, and (4) a mechanism must exist to remove the neurotransmitter from the site of action. Neurotransmitter effects at synapses that use channel-linked receptors are immediate, simple, brief, and the transmitter released from one axon terminal acts only on a single postsynaptic cell (Alberts et al., 1989).

In contrast to neurotransmitters, neuromodulators released from one axon terminal may act on many cells in the neighborhood of that terminal, with the effects being slow, complex, long-lasting, and often spatially diffuse. Neuromodulators work in many different ways, such as by influencing the binding of another ligand at a specific site, or binding to a receptor whcih then stimulates cAMP synthesis via a protein. These effects are positive or negative depending on whether effector increases or decreases the protein's ligand-binding affinity (Voet and Voet, 1990).

Recent investigations have shown that neuromodulators play important roles in shaping simple behaviors. They can act at many different sites within an animal in a coordinated fashion, modulating the neural circuits for motor patterns in the central nervous system, altering

motoneuron excitability, and modulating muscle response to motoneuron input (Harris-Warrick et al., 1989).

Neurohormones can be structurally similar or even identical to neurotransmitters. Unlike neurotransmitters, however, they are released from sites near the circulation. They can act on distant target sites, they can have long time courses of action (minutes to hours), and they can exert multiple actions on a variety of nonneural and neural cellular elements. In addition, neurohormones have greater potencies than neurotransmitters (Beltz, 1988). Neurohormones are released from neurosecretory cells whose terminals end on the walls of vascular channels (Scharrer, 1977). These neurohormones may act to change activity in central pattern-generating circuits (Dickinson, 1989), in muscle fibres (Bishop et al. 1987), in subsets of motoneurons (Harris-Warwick and Kravitz, 1984), and in interneurons (Glanzman and Krasne, 1983) in such a way as to promote a certain behavioral sequence or posture. Pasztor and MacMillan (1990) have shown that individual modulators can have different effects at different sensory loci within an organism, and that species specificity in neuromodulator actions exists even in closely related species.

The distinctions between neurohormones, neuromodulators and neurotransmitters are somewhat blurred. Current evidence shows that many peptides once regarded as conventional neurohormones in both vertebrate and invertebrate systems are versatile compounds capable of acting as transmitters in one area of the nervous system, as neurohormones in another, and as neuromodulatory substances in yet another (Beltz, 1988). For example, somatostatin acts not only as a neurohormone regulating pituitary function, but also as a neurotransmitter or neuromodulator in brain sites outside the hypothalamus (Reichlin, 1981). Neuropeptides typically have the following properties: low overall concentrations within organisms, high potency, marked

specificity, and apparent low rate of *de novo* synthesis *in vitro* (Reichelt and Edminson, 1977). Extreme potency and targets activated by picomolar concentrations characterize those neuropeptides with transmitterlike or hormonal effects. In certain regions of the central nervous system, neuropeptides may be under feedback control in which systems for neuropeptide breakdown are activated (Reichelt and Edminson, 1977).

In invertebrates, the behaviour of neuropeptides differs from that of classical transmitters (Voet and Voet, 1990). Joosse (1988) reports that unlike classical transmitters, no special enzymatic machinery has been identified which effects the rapid breakdown of peptides at the sites of their release, thus causing their actions to last longer than those of classical transmitters. As well, peptides in the nervous system may diffuse to more widely distributed receptors (Kandel, 1983), and some neuropeptides are able to bind to several different receptors. Thus peptides can potentially transmit more information than classical transmitters (Schwyzer 1980). Despite the apparently greater amount of energy expended in synthesizing large neuropeptides as compared to small classical transmitter molecules, these larger neuropeptides have persisted as messengers throughout animal evolution. This observation may be explained by the finding that both their large size and possession of direct genomic coding provide the potential for greater adaptational and evolutionary control than that found in classical transmitters (Joosse, 1988).

When changes occur in amino acid composition due to DNA mutations, peptides have the advantage over classical transmitters. The small size of classical transmitters increases the possibility that any change in their structure will seriously affect the function, while a change in the structure of large peptides may be gradual. This gradual change in structure makes peptide messengers suitable candidates for a crucial role in the evolutionary development of regulatory systems (Joosse, 1988). Small changes in molecular structure may alter the rate or duration of

the process being controlled, an alteration which may provide a survival advantage for the organism in which it occurs (Joosse, 1988). Neuropeptides are encoded by DNA, while the structures of the synthesizing enzymes and not the classical transmitter itself are encoded within the genome. Thus compared to neuropeptides, classical transmitters have the possible disadvantage of changes in the genome affecting receptor binding, enzymatic breakdown and/or reabsorption of the transmitter, any of which may prove lethal to the organism (Joosse, 1988).

Another positive aspect of the direct genomic coding of neuropeptide structure is the occurrence of families of genes coding for the same neuropeptide precursor molecule, but with slightly different molecular structures for the final product (as reviewed by Joosse, 1988). From the viewpoint of population genetics this is a highly advantageous situation for the successful control of adaptive changes.

1.1.2 Neuropeptide biosynthesis

What peptide or peptides are produced and secreted by a cell depends not only on the particular mRNAs expressed in the cell but also on the production and distribution of the specific enzymes involved in fashioning the final peptide product. Intracellular routing mechanisms also act in this process by determining the fates of the newly synthesized peptides such as secretion, degradation, and intracellular function (Loh and Gainer, 1983). There is also evidence supporting external regulation of where neuropeptides may be produced. Recent findings suggest that localized extracellular signals are capable of regulating the transcriptional states of nuclei; resulting in synapse-specific gene expression (Burden, 1993).

Neuropeptides appear to be made by conventional peptide synthesis from amino acids which are linked at ribosomes according to a sequence determined by an appropriate messenger

RNA. Such synthesis appears to be confined to the cell body, since synthesis of peptides in nerve terminals has not been observed and the cell body of a typical neuron contains high numbers of ribosomes, some crowded togethor in the cytosol and others attached to the rough endoplasmic reticulum. Although the dendrites of neurons often contain some ribosomes, no ribosomes are present in the neuron axon (Alberts et al., 1989). This cell body synthesis contrasts with the synthesis of established neurotransmitters, which can also be synthesized in nerve terminals (Kruk and Pyrock, 1979).

A recent development supports the concept that glial elements also possess the ability to synthesize neuropeptides. This idea is based on the demonstration that certain neuropeptide genes are expressed in astrocytes. For example, the highest concentrations of mRNA sequences for the production of somastostatin and proenkephalins (as reviewed by Scharrer, 1990) have been recorded in the hypothalamic astroglia (Scharrer, 1990).

Nonneural secretion of a neuropeptide has been demonstrated experimentally in *Hydra* (Schaller, 1979). Normally, head activator hormone is produced primarily in nerves and stored in neurosecretory granules. However, nerve-free *Hydras* were found to contain about 8 times more head activator than did controls, with the hormone determined to be produced by myoepithelial cells (as reviewed by Greenberg and Price, 1983).

In invertebrate nervous systems, there are neurons ressembling classical neurosecretory cells which are specialized for peptide secretion. These cells contain many secretory granules within the perikarya and an abundance of rough endoplasmic reticulum. Although many of these neurosecretory cells are endocrine cells, others have axons which extend into the central neuropile where the neuropeptides are released (Golding and Pow, 1988).

The hypothesis that all neurons, including those containing conventional transmitters, may secrete peptides is supported by the finding of secretory granules in virtually all neurons (as reviewed by Hokfelt et al., 1984). More than one neuropeptide can be produced and released by the same neuron (as reviewed by Boer and Van Minnen, 1988). Both classical neurotransmitters and neuropeptides have been found in the same neuron. For example, in the snail *Lymnaea stagnalis*, some neurons contain the classical transmitters serotonin and dopamine and at least one vasotocin-like neuropeptide (as reviewed by Boer and Van Minnen, 1988).

Many neuropeptides can be found within the brain and other nervous tissues (as reviewed by Hadley, 1988). Peptides originally considered to be gastrointestinal substances were later shown to be present in brain tissue. For example, cholecystokinin (CCK) and substance P, first isolated as gastrointestinal peptides, were subsequently localized in the mammalian brain (Hadley, 1988). The ability of neurons, pituitary cells, or skin cells to produce a "peptide hormone", which is able to act as a neuropeptide, is believed to be due to the fact that all of these cells are derived from a "neuroendocrine-programmed" ectoblast. This concept attempts to provide an explanation for the observation that substances commonly considered neuropeptides are found in the central nervous system, the pituitary gland, and such peripheral organs as the gastro-intestinal tract and the skin (as reviewed by Hadley, 1988).

1.1.3 Neuropeptide precursors

All peptide neurotransmitters examined so far are synthesized first as a larger peptide, then the active form of the molecule is produced through progressive cleavings by enzymes (Bloom, 1981). Neuropeptides are encoded as parts of high molecular weight precursors. These pro-proteins are inactive forms (Scharrer, 1990) which undergo posttranslational processing including

sequential proteolytic cleavage by special converting enzymes and additional steps of modification such as glycosidation, phosphorylation, acetylation, and amidation. As well, some precursor molecules can give rise to different products depending on the site of synthesis (as reviewed by Scharrer, 1990).

Much speculation exists concerning the role of the peptide precursor molecule. A function that appears to be fulfilled by certain biosynthetic precursors is that of a prohormone that provides a means to generate multiple biologic activities from a single gene product (Habener, 1981). Thus, one precursor is able to give rise to a host of fragments of widely differing activities. Another role of precursors may be to ensure the correct formation of the secondary or tertiary structure of a polypeptide following biosynthesis (Gainer, 1977). This precursor formation is particularily useful in allowing proper formation of peptides composed of connecting subunits encoded on the same precursor molecule. Palade (1975) has also suggested that large precursor molecules may be synthesized to prevent leakage of small peptide products through the permeable rough endoplasmic reticulum membrane prior to its being packaged into a granule membrane. Another postulated function of the peptide precursor is that of intracellular signalling, by which cells can distinguish between differing classes of proteins and direct them to their specific sites of action (Habener, 1981).

1.1.4 Neuropeptide sites of action

Neuropeptides may diffuse to more widely distributed receptors than classical transmitters (Kandel, 1983), with certain peptides possessing the ability to bind to several different receptors (Schwyzer, 1980). Neuropeptides bind with high-affinity and high specificity to receptor(s) possessing saturable and reversible binding site(s) (Hadley, 1988).

The sites of action of neuropeptides determines whether we classify them as neuromodulators, neurohormones, or neurotransmitters. As neurotransmitters, neuropeptides mediate signal transmission between neurons, between neurons and muscles, or between neurons and glands (Voet and Voet, 1990). These transmitter molecules are released into the synaptic cleft and then diffuse to the postsynaptic membrane, where they can bind to specific sites (Voet and Voet, 1990). Peptide neurotransmitters use channel-linked receptors which, upon binding neurotransmitter, change their conformation so as to create an open channel for specific ions to cross the membrane, thus altering membrane permeability. Several transmitters may act simultaneously, and the effect of a given substance depends on the context of all other transmitters present. This principle has been demonstrated recently for three interacting neuropeptides in the rat hypothalamus (Albers et al., 1991). Neuromodulators act upon the same cells as do neurotransmitters, but are bound to different types of receptors. They possess non-channel-linked receptors which work by being functionally coupled to an enzyme that in the presence of neuromodulator catalyzes the production of an intracellular messenger. This messenger then causes changes in the postsynaptic cell including modifications of ion channels in its membrane (Alberts et al., 1989). Neurohormones, once released, travel through the circulation to a variety of targets remote from neuron synapses and widely distributed throughout the body.

1.2 FMRFamide related peptides (FaRPs)

The neuropeptide FMRFamide, with the amino acid sequence Phe-Met-Arg-Phe-NH_2 was originally isolated from ganglia of the clam *Macrocallista nimbosa* by Price and Greenberg (1977). Since this initial discovery, the distribution of FMRFamide-like peptides has been

reported in all the major animal phyla. A number of FMRFamide-like peptides have been isolated from coelenterates (Grimmelikhuijzen and Graff, 1985) arthropods (Boer et al., 1980), and chordates (Dockray et al., 1983). Recently FMRFamide-like immunoreactivity (FLI) has been reported in the nervous system of *Hydra* (Bode, 1992).

FMRFamide-related peptides (FaRPs) are generally defined as any peptide that can be detected using an assay for FMRFamide, and no requisite biological effects need to be involved (Price and Greenberg, 1989). More specifically, FaRPs are a heterogeneous assembly of peptides which vary from 4 to 36 amino acids in length, having a penultimate arginine and a C-terminal amide as their only common structural features (Price and Greenberg, 1989). Recently FaRPs have been classified as a large neuropeptide group present in both vertebrates and invertebrates (Boer et al., 1980; Greenberg and Price, 1983; Price and Greenberg, 1989). Among arthropods, FaRPs have been detected in at least 6 insect species and 3 crustacean species (as reviewed by Mercier et al., 1992). Studies of various organisms have indicated that the N terminus of certain FaRPs is crucial for at least some biological activities (Duve et al., 1992).

In molluscs, FaRPs have been shown to comprise the peptides FLRFamide, SPFLRFamide, and several heptapeptides with the general sequence XDFFLRFamide [where X is Glu, Gly, Ser, or Asn] (as reviewed by Taghert and Schneider, 1990). In 1987, Trimmer et al. isolated the peptides SDRNFLRFamide (termed F_2) and TNRNFLRFamide (termed F_1) from lobster tissue. Table I lists selected FaRP sequences which have been detected throughout the phyla.

TABLE I

Selected FaRP sequences throughout the phyla

Molluscs

Phe-Met-Arg-Phe-amide(Price & Greenberg,1977) *Macrocallista*
Phe-Leu-Arg-Phe-amide (Martin et al., 1981) *Octopus*
Glu-Asp-Pro-Phe-Leu-Arg-Phe-amide (Price et al., 1990) *Helix*
Gly-Asp-Pro-Phe-Leu-Arg-Phe-amide (Ebberink et al., 1987) *Lymnaea*
Ser-Asp-Pro-Phe-Leu-Arg-Phe-amide (Ebberink et al., 1987) *Lymnaea*
Asn-Asp-Pro-Phe-Leu-Arg-Phe-amide (Price et al., 1990) *Helix*
Tyr-Gly-Gly-Phe-Met-Arg-Phe-amide (Martin et al., 1981) *Octopus*

Insects

Tyr-Gly-Gly-Phe-Met-Arg-Phe-amide (Nambu et al.,1988) *Drosophila*
Asp-Pro-Lys-Gln-Asp-Phe-Met-Arg-Phe-amide (Schneider & Taghert, 1988)
Thr-Pro-Ala-Glu-Asp-Phe-Met-Arg-Phe-amide
Ser-Asp-Asn-Phe-Met-Arg-Phe-amide
Ser-Pro-Lys-Gln-Asp-Phe-Met-Arg-Phe-amide
Pro-Asp-Asn-Phe-Met-Arg-Phe-amide
Ala-Ala-Met-Asp-Arg-Tyr-amide
Met-Asp-Ser-Asn-Phe-Ile-Arg-Phe-amide
Ser-Ala-Pro-Gln-Asp-Phe-Val-Arg-Ser-amide

Crustacea

Ser-Asp-Arg-Asn-Phe-Leu-Arg-Phe-amide (Trimmer et al.,1987) *Homarus*
Thr-Asn-Arg-Asn-Phe-Leu-Arg-Phe-amide
Asp-Arg-Asn-Phe-Leu-Arg-Phe-amide (Mercier et al.,1992) *Procambarus*
Asn-Arg-Asn-Phe-Leu-Arg-Phe-amide

Coelenterates

pGlu-Gly-Arg-Phe-amide (Grimmelkhuijzen & Graff,1986) *Anthropleura*

Vertebrates

Tyr-Gly-Gly-Met-Arg-Phe (Stern et al.,1986) *Bos*
Phe-Leu-Phe-Gln-Pro-Gln-Arg-Phe-amide
Ala-Gly-Glu-Gly-Leu-Ser-Ser-Pro-Gln-Arg-Phe-amide
Phe-Phe-Trp-Ser-Leu-Ala-Ala-Pro-Gln-Arg-Phe-amide (Yang et al.,1985) *Bos*
Leu-Pro-Leu-Arg-Phe-amide (Dockray et al.,1983)*Gallus*
Tyr-Val-Met-Gly-His-Phe-Arg-Trp-Arg-Phe-amide (Williams, 1983) *Rattus*

1.2.1 FaRP functions

In general, FaRPs are capable of acting as neurotransmitters, neuromodulators, or hormones (Linacre et al., 1990) as are most other neuropeptides. Specifically, physiological studies have demonstrated that FaRPs have a variety of modulatory effects on both central and peripheral targets (Schneider and Taghert, 1990). These peptides have been demonstrated to decrease morphine-induced analgesia in mammals, (Yang et al., 1985) and to increase the amplitude and rate of induced muscle contractions in insects (Evans and Meyers, 1986).

FMRFamide itself has been found to possess anti-opiate properties in both molluscs and mammals (Greenberg et al., 1983) (Kavaliers and Hirst, 1985) (Raffa, 1988). In the sea hare *Aplysia,* FMRFamide performs multiple roles, functioning as a chemical messenger from neurons to other neurons, to muscles, and to glands (Schaefer et al., 1985). In this organism extensive research has shown FMRFamide to modulate the activity of specific central neurons, to cause several kinds of membrane conductance changes in the CNS, and to shorten the presynaptic action potential in sensory neurons. FMRFamide has also been shown to act in peripheral systems by inhibiting spontaneous activity of the gut and salivary gland, and to stimulate contractions of the buccal mass and gill (Linacre et al., 1990).

There are no general characteristics that can be assigned to the effects of FMRFamide or its analogs. Results show that FaRPs exhibit multiple actions on various tissues, reflecting the structural variation, not only of the peptides, but also of their receptors (Kobayashi and Muneoka, 1989). It is a common observation that a single FaRP may produce different effects at each of its receptive tissues (Greenberg et al., 1983).

There is evidence that FaRPs can function as both cardioregulatory hormones and transmitters: FMRFamide-like immunoreactivity is present in *Helix* blood (Price et al., 1985) as

well as in the heart itself (Lehman and Price, 1987). Immunoreactivity in the locust nervous system suggests that there is differential expression of FaRPs in different parts of the nervous system, and bioactivity testing suggests that FaRPs in the locust function both as circulating neurohormones and as locally released neuromodulators or neurotransmitters (Robb and Evans, 1990).

Different FaRPs including those with close structural similarities have been observed to elicit varying physiological responses. The tetrapeptide and heptapeptide FaRPs have distinct effects on muscle tone and posture in pulmonate snails (as reviewed by Price and Greenberg, 1989), while FMRFamide and FLRFamide modulate antagonistic gastropod buccal muscles (Kobayashi and Muneoka, 1989). The modulatory effects of different FaRPs on heart activity in invertebrates have been found to be dependent on the chemical structure of the peptide and on the invertebrate species (Painter and Greenberg, 1982). In invertebrate hearts, FaRPs elicit responses that vary from being inhibitory to biphasic to excitatory (Mercier and Russenes, 1992; Painter and Greenberg, 1982; Robb et al. 1989; Kravitz et al., 1987).

Observations that tetra- and hepta-FARPs have large differences in potencies and that in some cases tetra-FaRPs produce opposing actions to the hepta-FaRPs (Cottrell and Davies,1987), (Lehman and Greenberg, 1987) have led to the suggestion that there is more than one type of FaRP receptor in molluscan species (Cottrell et al., 1992).

1.2.2 FaRP production sites and sites of action

FaRPs are known to be produced in the central nervous systems of invertebrates. Their concentrations, however, are not uniform throughout constituent tissues, and the sites of action of FaRPs are diverse and vary between species.

Several different tissues in invertebrates have been found to display FMRFamide-like immunoreactivity (FLI). Such immunoreactivity often indicates the presence of either FMRFamide or a molecule recognizing the RFamide sequence, but cannot identify the tissue in question as either a FaRP production site or site of action (Duve and Thorpe, 1988). In the marine molluscs *Aplysia californica* and *Pleurobranchea california,* FMRFamide immunoreactivity was observed in the buccal ganglia, the cerebral ganglion, the abdominal ganglion, and the pedal and pleural ganglia (Soinila and Mpitsos, 1991), while in octopus nervous system FMRFamide-like immunoreactivity was found in a neuropil layer of the vena cava (Martin et al., 1981).

1.2.3 Crayfish FaRPs

Immunoassays to determine the presence of FaRPs in the crayfish *Procambarus clarkii* (Mercier et al., 1991) have shown FMRFamide-like immunoreactivity to be particularly concentrated in the pericardial organs, which are known to release neurosecretory products that modify cardiac and neuromuscular functions. Immunoreactivity was also found in the second thoracic roots and the connective tissue sheath surrounding the nerve cord (Mercier et al., 1991). The presence of FLI in cells associated with crayfish hindgut suggests that members of the FMRFamide-related family of peptides may be involved in coordinating feeding and digestion in arthropods (Mercier et al., 1991). Physiological experiments with isolated crayfish hearts and hindguts support the notion that crayfish FaRPs are likely to play roles in controlling circulation and defecation (Mercier et al., 1991).

Two new FaRPs in the crayfish *Procambarus clarkii* have been extracted and sequenced recently (Mercier et al., 1993). These peptides, structurally similar to the lobster peptides F_1 and

F_2, were named DF_2, which has the sequence DRNFLRFamide, and NF_1, which has the sequence NRNFLRFamide. NF_1 and DF_2 are cardioexcitatory and augment synaptic transmission at neuromuscular synapses (Mercier et al., 1993; Skerrett et al., 1995). Mercier et al. (1992) report that the peptides DF_2 and NF_1 increased the rate and amplitude of spontaneous contractions in semi-isolated crayfish hearts. Both of these peptides also increased the amplitude of excitatory postsynaptic potentials recorded in phasic abdominal extensor muscles in response to nerve stimulation. FMRFamide-like immunoreactive material from the crayfish hindgut appears to contain at least one peptide distinct from NF_1 and DF_2, based on elution on high performance liquid chromatography (Mercier et al., 1995). Partial sequencing analysis has shown this peptide to have the sequence pEXXDHVFLRFamide (where "X" indicates an unidentified amino acid)(Mercier et al., 1995).

1.2.4 FaRP-encoding DNA

FaRP-encoding genes have been characterized in species representing all of the invertebrate subphyla with the exception of *Malacostraceans* to which group the crayfish *Procambarus clarkii* belong. *Aplysia californica* and *Lymnaea stagnalis* (both molluscs), *Drosophila melanogaster* and *Drosphila virilus* (members of the arthropod family), and the annelid *Caenorhabditis elegans* have all been found to possess varying degrees of similarity in their FaRP-encoding genes. In invertebrates studied to date, several different FMRFamide-related peptides have been encoded by the same gene, mRNA and protein precursors, with the occasional occurance of multiple peptide copies (Linacre et al., 1990). For example, the FMRFamide precursor gene of *Aplysia* encodes 28 copies of FMRFamide, one copy of FLRFamide, and one copy of a peptide ending in Tyr-Leu-Arg-Phe-amide. The available evidence indicates that

several organisms, such as *Lymnaea*, *Aplysia,* and *Caenorhabditis* (Linacre et al., 1990), (Schaefer et al., 1985) (Rosoff et al., 1992) possess a single FaRP-encoding gene. However, the existance of other genes containing small regions encoding peptides with sequences related to FMRFamide cannot be excluded. In *Aplysia* a second neuropeptide gene expressed in the neuron L5 encodes a peptide that ends with the sequence Arg-Phe-NH_2 (as reviewed by Schneider and Taghert, 1988), while either two genes encoding heptapeptide FaRPs occur in the snail *Helix*, or a single gene encoding at least five heptapeptides is present (Price et al., 1990). In *Caenorhabditis elegans*, a single FaRP-encoding gene has been isolated (Rosoff et al., 1992). However, in the related nemotode *Panagrellus redivinus* two new FaRPs not found on the *C. elegans* FaRP gene (flp-1) have been discovered. As previously all *Panagrellus* FaRPs have been found in *C. elegans*, it has been suggested that at least two different FaRP genes may be present in nematodes (Maule et al., 1994). As well, the absence of methionine-containing FaRPs in the flp-1 gene of *C. elegans*, coupled with HPLC analysis indicating a peak which coelutes with authentic FMRFamide, suggests the presence of a second FaRP gene in *C. elegans* (Rosoff et al., 1992). In *Drosophila* it has been found that there are at least two types of genes with the potential for producing 10 or more FMRFamide-related peptides. The *Drosophila* FMRFamide gene itself produces five distict FMRFamides and three other peptides with variations in the FMRFamide C-terminal sequence (Schneider and Taghert, 1988). The drosulfakinin gene, encoding peptides homologous to both cholecystokinin and FMRFamide gives rise to other peptides ending C-terminally with -Met-Arg-Phe-NH_2 (as reviewed by Duve et al., 1992). Recently, a new peptide with the sequence TDVDHVFLRFamide has been isolated from *Drosophila* (Nichols,1992). As neither the FMRFamide or the drosulfakinin gene encode this

peptide, it has been postulated that another gene encoding -Arg-Phe-NH_2 peptides exists in this organism.

The size and structure of FMRFamide-encoding genes has also been investigated in invertebrates. In the fruit fly *Drosophila melanogaster,* the exons are 106 and 1352 base pairs long, while in the related species *Drosophila virilus*, they are 169 and at least 1232 base pairs long. The first exon of the *Drosophila* FMRFamide gene consists of the 5' untranslated region and is spliced to the initiator methionine codon. The second exon encodes the prohormone and the 3' untranslated region of the mRNA. The promoter region contains a TATA box 30 nucleotides upstream from a consensus transcription start site (Chin et al., 1990). In both species the intron is approximately 2500 base pairs long (Taghert and Schneider, 1990). *Lymnaea stagnalis* also possesses a single FaRP-encoding gene made up of 2 different exons. The first exon encodes for FMRFamide, and is located 3.4kb 3' to the second GDP/SDPFLRFamide exon (Saunders et al., 1991). The *Aplysia* FaRP gene possesses at least 2 exons, but differs from *Drosophila* and *Lymnaea* in that the precursor molecule is encoded by more than a single exon. Finally, the flp-1 gene of *C. elegans* differs from the other invertebrate FaRP genes in that it has six exons (Rosoff et al., 1992). Three of the separating intrusions are < 85bp, one of medium length (156bp) and one relatively large (434bp).

In general, FaRP-encoding genes are composed of highly repetitive stretches containing a FaRP sequence and its associated processing signals. In *Aplysia* many of these repeated sequences are more than 90% identical at the nucleotide level (Price et al., 1987b). In the FaRP genes of *Drosophila, Aplysia,* and *Lymnaea,* some FaRP sequences are not separated by spacer regions, but instead are arranged in tandem pairs. Spacer regions separating FaRP sequences in genes are generally short, but vary both within and between species. In *Aplysia*, spacers are

usually composed of either seven or eight amino acids and are highly acidic, neutralizing the positive charge of the basic peptide region. These spacers are less conserved than the peptide-encoding regions (Taussig and Scheller, 1986). In *Lymnaea*, spacer regions, while still short, are more variable in size and amino acid composition. Linacre et al. (1990) postulate that these regions do not possess any biological function other than to separate FaRP sequences. This lack of biological significance for spacers also seems probable in 3 other cases: (1) *Drosophila melanogaster*, which possesses only one short spacer region between its 13 encoded FaRP sequences (Schneider and Taghert, 1988), (2) the GDP/SDPFLRFamide exon of *Lymnaea*, in which 15 encoded FaRP sequences are separated by only 2 amino acids used for amidation of peptides and cleavage (Saunders et al., 1991), and (3) in the flp-1 gene of *C. elegans*, in which spacer regions are absent between most of the putative peptides, while those that are present are not highly acidic and may represent peptides of unknown classes (Rosoff et al., 1992). Conversion of a precursor molecule into its final products occurs most frequently at adjacent basic amino acid residues or at single arginyl residues (Wold, 1981; Loh et al., 1984). Within the invertebrates studied to date, some similarities in potential cleavage sites exists between organisms. In the *Aplysia* FMRFamide gene, nearly all of the FMRF sequences are flanked on the carboxy terminus by Gly-Lys residues, and on the amino terminus by Lys-Arg-, suggesting that single basic lysine residues indicate cleavage sites to processing enzymes (Schaefer et al., 1985). The FMRFamide exon of the *Lymnaea stagnalis* FaRP gene was also found to possess similar cleavage recognition sites with the exception of a Lys-Ser residue on the amino terminus of one FMRFamide (Linacre et al., 1990). The GDP/SDPFLRFamide exon of the *Lymnaea* FaRP gene has both Gly-Lys residues, as found in *Aplysia*, and Gly-Arg residues (Saunders et al., 1991). These Gly-Arg and Lys-Ser residues of the *Lymnaea* FaRP exons are also common to the

Drosophila FaRP gene which additionally contains Leu-Leu, Lys-Ser, Gly-Arg, Arg-Ser and Lys-Lys cleavage sites (Schneider and Taghert, 1988). *C. elegans* also shows limited homology to the other invertebrate cleavage sites, possessing either mono-, di-, or tribasic amino acids (Rosoff et al., 1992). Figure 1 shows schematic diagrams of the FaRP-encoding peptide precursors of *Drosophila* and *Lymnaea,* representative of the similarities and diversities of the invertebrate FaRP sequences characterized to date.

1.2.5 FaRP-encoding RNA

FaRP-encoding mRNAs have been characterized in selected invertebrate species. The number of FaRP-encoding RNAs in different species has been a subject of controversy, with different numbers of RNA transcripts being reported for a single organism.

In the snail *Lymnaea stagnalis* isolation of RNA and subsequent RNA blot hybridizations have demonstrated the presence of a single FaRP RNA transcript 1.7 kb in length (Linacre et al., 1990).

Attempts at identifying RNA transcripts of FaRP genes in Drosophila have yielded some conflicting results. Single RNA transcripts of approximately 1.7kb have also been reported in other organisms, but these results coexist with other conflicting results. Schneider and Taghert (1988) isolated two RNA transcripts from the fruit fly *Drosophila melanogaster*; one transcript was 1.7kb in length and the other was approximately 0.7kb in length. Subsequently, the same researchers reported that when the RNA hybridizations were repeated using the same cDNA probe but with its 3' untranslated end removed, hybridization to the smaller 0.7kb transcript did not occur. It was suggested that this 0.7kb RNA represents an unrelated gene that has homology to only the 3' untranslated end of the gene (Schneider and Taghert, 1990). However, the

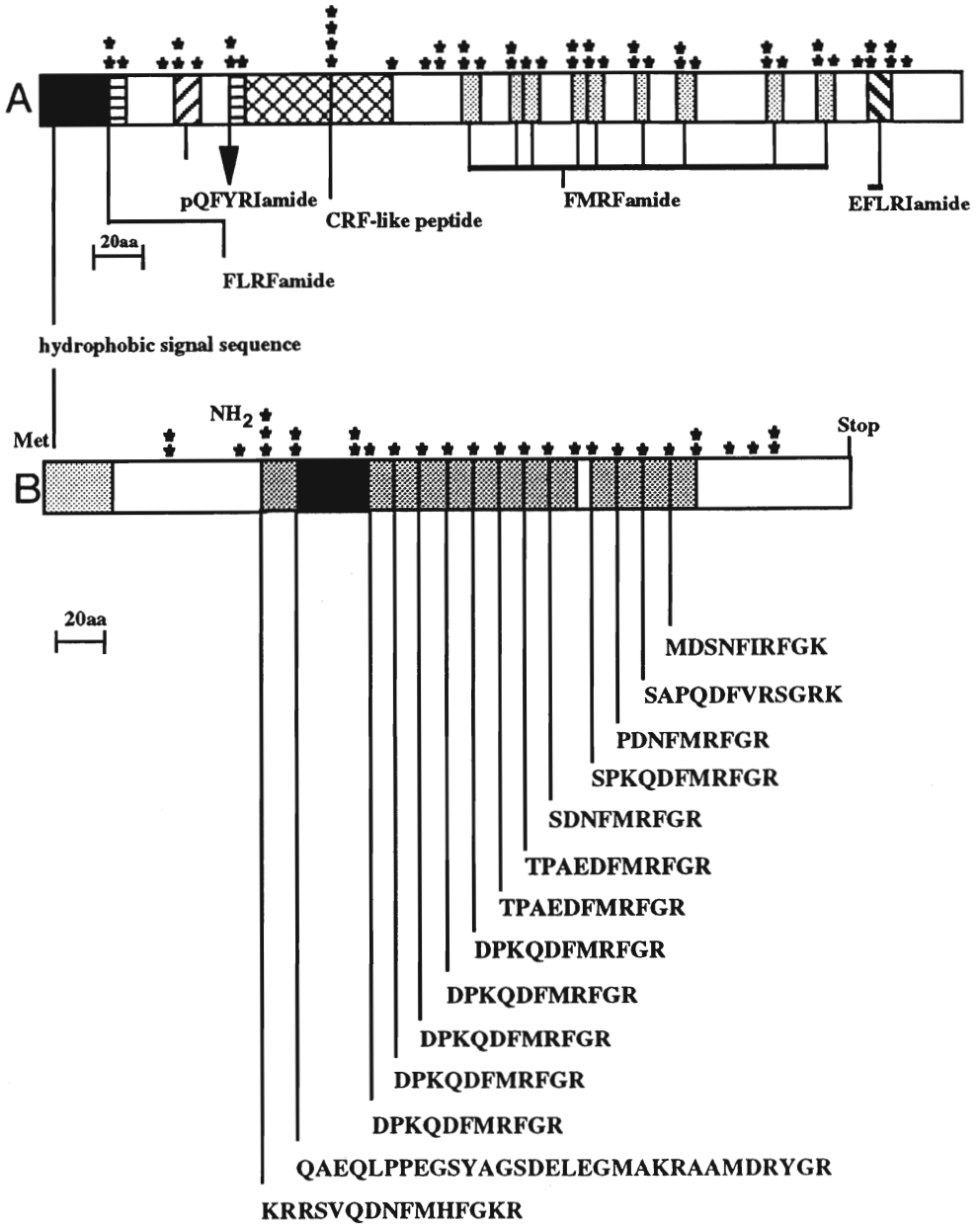

Figure 1:Schematic representations of the FMRFamide polyprotein precursor in *Lymnaea stagnalis* (A) and in *Drosophila melanogaster*(B), as adapted from Linacre et al. (1990) and Schneider and Taghert (1988). These two representative sequences demonstrate the multiple FaRPs commonly encoded on invertebrate FaRP precursors, the widespread lack of spacer regions between encoded peptides, the occurance of short spacer sequences when present, and the presence of varying numbers of basic residues between FaRP sequences. Deduced peptide cleavage products are represented as boxes within the precursors. Stars indicate the presence and number of basic residues.

Drosophila DSK gene encoding drosulfakinin peptides homologous to FMRFamide has been found by Northern hybridization to possess a single mRNA transcript approximately 0.8kb long (Nichols et al., 1988). This related mRNA may account for the varying hybridization results of Schneider and Taghert (1988, 1990).

In the sea slug *Aplysia* both the number of FaRP RNA transcripts and their respective sizes have been reported for neuron L5 (Shyamala et al., 1986) and for neurons L10, L12, and L13 (Schaefer et al., 1985). Shyamala et al. (1986) report a single 1.8kb transcript expressed mainly in the abdominal ganglion, and less intensely in the pleural ganglion, while Schaefer et al. (1985) described multiple RNA transcripts in many tissues. These latter findings included two major RNA species of 1.4kb and 3.2kb and a less abundant transcript of more than 4.0 kb. The smaller of the intense transcripts was predominantly expressed in the abdominal ganglion, while the larger transcript was expressed in all ganglia examined. Less abundant RNA transcripts were found principally in the pleural ganglia; all were smaller than the largest major transcript, while some were less than 0.5kb in length. Although both of these studies used RNA hybridization as their method of detection, the different radioactively-labelled cDNA clones used in the two experiments were found to possess no substantial homology with each other (Schaefer et al., 1985).

The FaRP encoding gene of *Lymnaea* has been shown to undergo cell-specific alternative RNA splicing. The two FaRP gene exons are spliced onto a common upstream exon encoding a hydrophobic leader sequence, while *in situ* hybridization data shows that there is mutually exclusive cytoplasmic expression of each of the neuropeptide-encoding exons. The observed cellular specificity of the splicing events suggests that cell-specific alternative splicing may be an important method of establishing neuronal diversity in this organism (Saunders et al., 1992).

The flp-1 gene of *Caenorhabditis elegans* provided the first concrete example of a distinct FMRFamide-like peptide being derived from alternative splicing (Rosoff et al., 1992). In this organism, an alternative 3' splice acceptor site between exons 3 and 4 results in the substitution of AGSDPNFLRFG for one of the copies of SADPNFLRFG found in the other translation product, suggesting a unique role for the substituted peptide in the animal (Rosoff et al., 1992). Recent investigations of a similar FaRP-encoding gene in *Caenorhabditis vulgaris* have also demonstrated the production of two transcripts by alternative 3' splicing, suggesting functional significance of the splicing procedure (Schinkmann and Li, 1994).

Observations of multiple FaRP mRNA transcripts have been linked to differences in encoded peptides. For example, FaRP-containing neurons of *Helix aspersa* appear to express either the tetra-FaRPs or the hepta-FaRPs, but not both (Cottrell et al., 1992). In these experiments, mRNA transcripts were probed by molecular hybridization using cloned cDNA probes, and FaRPs expressed by neurons were examined using radioimmunoassays (RIA) and high performance liquid chromatography (HPLC). Analyses of *Helix* cDNAs suggest that one species of mRNA encodes tetra-FaRPs, while another mRNA encodes all of the hepta-FaRPs and the base sequence for one septa-FaRP.

1.2.6 FaRP-encoding protein precursors

Invertebrate FaRP precursors studied to date are similar in that they contain diverse yet related neuropeptides, some of which are present in multiple copies, while others are present in single copies (Taghert and Schneider, 1990; Linacre et al., 1990; Schaefer et al., 1985). In the snail *Lymnaea stagnalis*, a single gene was found to encode a precursor protein containing 9 copies of FMRFamide, 2 copies of the related peptide Phe-Leu-Arg-Phe-NH_2 (FLRFamide), and

a single copy of the putative pentapeptides Gln-Phe-Tyr-Arg-Ile-NH_2 (post translationally modified to pQFYRIamide) and Glu-Phe-Leu-Arg-Ile-NH_2 (EFLRIamide) (Linacre et al., 1990). The functional significance, if any, of the 28:1 molar ratio of FMRFamide to FLRFamide found in the *Aplysia* precursor protein is not clear (Taussig and Scheller, 1986); however the resemblance to the mammalian enkephalin precursor, which contains seven copies of met-enkephalin and a single copy of Leu-enkephalin is striking (Comb et al, 1982). FMRFamide may be less stable than FLRFamide due to the potential for oxidation of the thiol group on the methionine side-chain. Alternatively, the two peptides may have different susceptibilities to degradation by proteolytic enzymes or, as is the case with the enkephalins, the molecules may have different affinities for surface receptors (Taussig and Scheller, 1986).

All secretory peptides, including the FaRPs, are encoded by their genes as a segment of a larger precursor molecule. After translation, the secretory products are processed out of the precursor by proteolytic enzymes (Greenberg and Thorndyke, 1989). The observation of differential expression of FaRPs in various tissues (Robb and Evans, 1990) suggests that different FaRP-encoding precursors may be expressed differentially in various cell types or that the same precursor is activated in all cell types, but its products are processed differently in different tissues.

1.3 FaRP Similarities to Other Peptide Families

The family of FMRFamide-related peptides displays striking but as yet unresolved similarities with several other peptide families found throughout both vertebrates and invertebrates. Assessment of actual homology between neuropeptide families occurs by

reviewing structural similarities at all genetic levels, functional similarities, and distribution both within an organism itself and throughout the animal kingdom.

Structural similarity characteristic of peptide families could arise in two ways: through either duplication of the segment of a gene encoding a particular peptide and its processing signals, or through duplication of the entire gene, followed by subsequent point mutation in one of the duplicate sequences. Both of these processes will produce related peptides.Duplication of peptide sequences is evident where very similar sequences are part of the same precursor. For example, the 28 FMRFamide and 1 FLRFamide sequences encoded in the *Aplysia* precursor suggest duplication of a FMRFamide sequence with subsequent point mutation to change the encoded Methionine to a Leucine sequence. Even where sequence similarity is low, the peptides may have similar tertiary structures, suggesting divergence from a common genetic origin. However, peptide families or new members of existing families can also occur by chance, particularly when the peptides are short. As peptides become shorter, the possibility that similar sequences are convergent increases. If the precursor organization of a short peptide is not known, convergence may not be detectable (Price and Greenberg, 1989).

Sequence homology can be statistically evaluated according to the number of matching amino acids within a given sequence. Price (1983) calculated that any given sequence of four amino acids will occur at least once in the proteins of any species, but as the sequence increases in length the odds of it occuring by chance decrease dramatically. Thus, matches of six amino acids in a row are fairly unique and indicate a likely homology. The expression of FMRFamide actually requires a sequence of at least seven amino acids: Phe, Met, Arg, and Phe (the body of the peptide); Gly (to form the amide); and at least one basic amino acid (eg. Arg)-but more usually a dipeptide (Lys-Arg, Lys-Ser, or Arg-Ser) on each end as cleavage signals. On the basis

of the above argument, a commonality of six residues (including processing signals) is assumed to be the minimum requirement for homology (Price and Greenberg, 1989).

1.3.1 The Cholecystokinin/Gastrin family

Cholecystokinin (CCK) is one of a family of peptides, including the gastrins and caerulein peptide, that is broadly distributed both phylogenetically and anatomically (Dockray, 1989). This gastrin-CCK family is defined as consisting of peptides with the common C-terminal sequence Gly-Trp-Met-Asp-Phe-NH_2 which represents the biologically active portion of the molecule. These peptides, found in both vertebrates and invertebrates, are thought to control several aspects of vertebrate gut function, and have been found to control feeding-related behaviour in mammals (Dockray, 1989) and in the lobster (Turrigiano and Selverston, 1990). CCK-like peptides are thought to function as neurotransmitters, neuromodulators, and neurohormones (Dockray, 1989).

CCK-like immunoreactivity has been found within the stomatogastric nervous system of four species of decapod crustacea, including the crayfish *Procambarus clarkii* (Turrigiano and Selverston, 1991). All four species have CCK-like immunoreactivity within the pericardial organs, thoracic segmental nerves, and in the haemolymph. The presence of detectable levels of CCK-like peptide in these locations suggests that CCK-like peptides in decapod crustacea may be utilized as neurohormones, while a neurotransmitter or neuromodulator role is also likely (Turrigiano and Selverston, 1991).

In the crayfish, CCK may affect the amylase electrophoresis pattern (Denuce, 1982). Gastrin and CCK stimulate the secretion of digestive enzymes in the gastric juice of crayfish (as reviewed by Favrel et al., 1987). A CCK-like peptide has been found to be present in the stomatogastric nervous system of the lobster *Panulirus interruptus* and can be released into the

stomatogastric ganglion (Turrigiano and Selverston). As well, the activity of the pyloric and gastric mill can be modified by bath application of CCK. This strongly suggests that an endogenous CCK-like peptide acts as a neuromodulator of the circuits in the stomatogastric ganglion and thus plays a role in the control of gut function.

Several lines of evidence indicate the FaRPs and CCK-like peptides have homologous distributions and functions. Such evidence includes the co-existence of CCK-like peptide and the FaRPs in the pericardial organs of *Procambarus* (Turrigiano and Selverston, 1991; Mercier et al., 1992), related effects of CCK-like peptide upon the gastric mill in lobster (Turrigiano and Selverston, 1989), and of FaRPs in modulation of feeding in crayfish (Mercier et al.,1990), and the effect of both CCK-like peptides and FaRPs upon invertebrate gut motility (Nachman et al., 1986; Mercier et al., 1995) suggests a homology in function and distribution.

Further evidence of relatedness between the FaRP and the CCK/gastrin peptide families comes from studies characterizing the peptide leukosulfakinin. Leukosulfakinin (Glu-Gln-Phe-Glu-Asp-Tyr(SO_3H)-Gly-His-Met-Arg-Phe-NH_2) itself is a sulfated myotropic neuropeptide, while the leukosulfakinins in general are a family of peptides possessing the C-terminus-DYGHMRF-NH_2 (Price and Greenberg, 1989). Leukosulfakinin (LSK) exhibits sequence homology with the hormonally active portion of the vertebrate hormones human gastrin II and CCK, suggesting that these peptides are evolutionarily related. Six of the eleven leukosulfakinin amino acid residues are identical to those of gastrin II. In addition, the intestinal myotropic action of leukosulfakinin is analogous to that of gastrin, suggesting that it is unlikely that their structural convergence is coincidental (Nachman et al., 1986). Leukosulfakinins have also been found to stimulate contraction of the cockroach hindgut (Holman et al., 1986); a function similar to FaRP modulation of crayfish hindgut (Mercier, et al., 1995). These peptides demonstrate a structural

ressemblance to FMRFamide in their common terminal sequence Met-Arg-Phe-NH_2 (Dockray, 1989).

Thus, leukosulfakinins, with similarities to both gastrins and FMRFamide may provide a genetic link between the two peptide families which reinforces evidence of homology between the two families themselves. The sequence similarities which suggest a possible common precursor link between the cholecystokinin /gastrin family and the FMRFamide-related peptides are listed in Table II.

Another group of neuropeptides suggested to genetically link the FMRFamide and CCK/gastrin families is that of the drosulfakinin peptides. These DSK peptides are homologous to gastrin, CCK, and LSK sequences. That they share a similar C-terminal pentapeptide which constitutes the CCK-gastrin bioactive site argues for a possible similarity in function (Nichols et al., 1988).

Genomic DNA and cDNA clones from *Drosophila* corresponding to the DSK precursor encode three putative peptides. The three peptides (DSK-O, Asn-Gln-Lys-Thr-Met-Ser-Phe-Gly; DSK-I, Phe-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-Gly; DSK-II, Gly-Gly-Asp-Asp-Gln-Phe-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-Gly) are flanked by prohormone processing sites and contain C-terminal glycocyl residues, a potential amidation site (Nichols et al., 1988). Two of the peptides, DSK-I and DSK-II, are homologous to CCK-gastrin peptides. Each of the two homologues includes a CCK/gastrin-like C-terminal pentapeptide and a conserved sequence corresponding to the sulfated tyrosine in bioactive CCK. Comparison of the amino acid sequences of the DSK peptides with CCK reveals that four of the eight C-terminal residues are identical. In addition,

TABLE II

Sulfakinin-FMRFamide sequence similarities

Sequence similarities suggesting a possible common precursor link between sulfakinins and FMRFamide related peptides (as taken from Nichols et al., (1988), Dockray (1989), De Loof and Schooes (1990) and Hadley (1992))

Cholecystokinin (Nichols et al., 1988)
Glu-Gly-Pro-Trp-Leu-Glu-Glu-Glu-Glu-Ala-Tyr-Gly-Trp-Met-Asp-Phe-NH_2

Human gastrin I (Dockray, 1989)
Lys-Ala-Pro-Ser-Gly-Arg-Val-Ser-Met-Ile-Lys -Asn-Leu-Gln-Ser-Leu-
Asp- Pro-Ser-His-Arg-Ile-Ser-Arg-Asp-Tyr-Met-Gly-Trp-Met-Asp-Phe-NH_2

Drosulfakinin O (Nichols et al., 1988)
Asn-Gln-Lys-Thr-Met-Ser-Phe-NH_2

Drosulfakinin I (Nichols et al., 1988)
Phe-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-NH_2

Drosulfakinin II (Nichols et al., 1988)
Gly-Gly-Asp-Asp-Gln-Phe-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-NH_2

Leucosulfakinin I (Nichols et al., 1988)
Glu-Gln-Phe-Glu-Asp-Tyr-Gly-His-Met-Arg-Phe-NH_2

Leucosulfakinin II (Nichols et al., 1988)
pGlu-Ser-Asp-Asp-Tyr-Gly-His-Met-Arg-Phe-NH_2

Drosophila peptide (DeLoof and Shooes, 1990)
Tyr-Gly-Gly-Phe-Met-Arg-Phe-NH_2

Octopus peptide (Hadley, 1992)
Tyr-Gly-Gly-Phe-Met-Arg-Phe-NH_2

FMRFamide (DeLoof and Schooes, 1990)
Phe-Met-Arg-Phe-NH_2

the DSK neuropeptides share many common amino acid sequences with FMRFamide-related peptides, and so are presented as another possible link between the two peptide families.

1.3.2 Enkephalins

Finally, FMRFamide-related peptides are thought to be closely related to the enkephalins, a family of pentapeptides which interact with opiate receptors and mediate analgesic responses in mammals (Rosen et al., 1984).

Met-enkephalin-Arg-Phe(YGGFMRF), at once an enkephalin and a FMRFamide analog, was isolated from chromaffin cells of the adrenal medulla and in the striatum of mammals (Stern et al., 1979). It has been proposed that both FMRFamide and the enkephalins diverged from an ancestral sequence, Tyr-Gly-Gly-Phe-Met-Arg-Phe-NH_2, which acted at a single receptor having a pair of binding sites, each complementary to one end of the peptide (Greenberg and Price, 1983). With time, two new types of bivalent receptors evolved; each type bound to both the N- and C- terminals of the heptapeptide. Whereas the effective site for one type bound the C-terminal as in present FMRFamide-like receptors, the other kind bound effectively to the N-terminal, as do enkephalin-like receptors. Subsequent evolution presumably resulted in the current groups of FMRFamide-like and enkephalin-like peptide and their respective receptors with conserved message sequences at the C- and N- terminals respectively, and with varying sequences at the opposite ends of the molecules (Greenberg et al., 1986). The distribution of enkephalin-like peptides and FaRPs in the animal kingdom is shown in Figure 2.

Research has supported this theory of a FMRFamide-enkephalin linkage. Antisera against the opioid peptide enkephalin and against the molluscan neuropeptide FMRFamide were observed to immunostain identical nerve terminals and neurosecretory granules in a neuropil layer of the

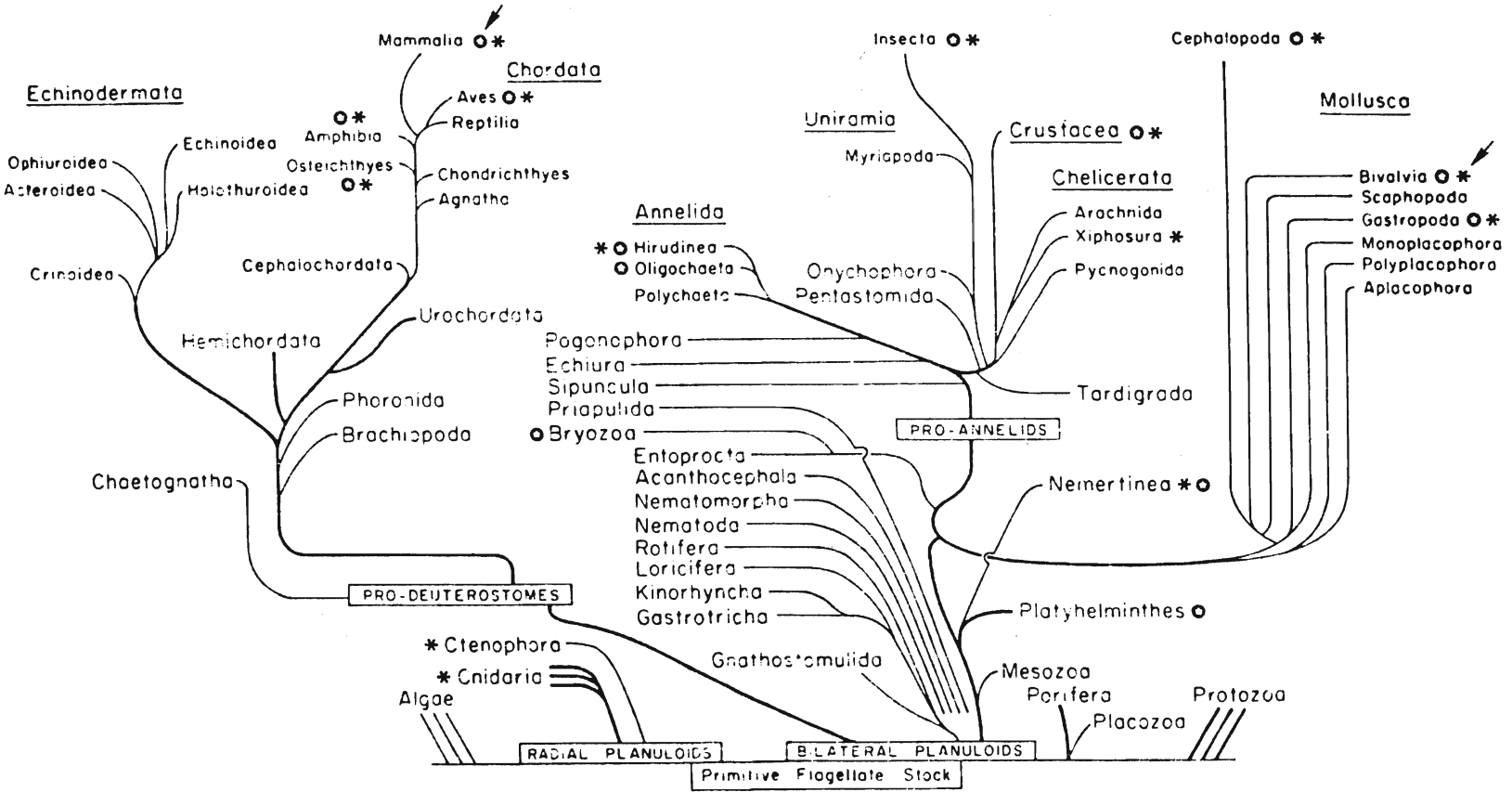

Figure 2: This figure shows all phylogenetic groups in which enkephalin-like peptides (as indicated by a circle) and FMRFamide-like peptides (as indicated by an asterisk) have been detected. The evolutionary relations of the various groups provides an indication of the origin of a possible common precursor molecule. (Figure reproduced from Greenberg et al.,1986).

octopus vena cava. This coexistance of the two different immunoreactive indicates occurence of a common enkephalin-Arg-Phe precursor (Martin et al., 1981). In the octopus a naturally occurring heptapeptide YGGFMRFamide was isolated from the vena cava neuropile (Martin et al, 1981) as well as from the brain, and was

found to possess both cardioexcitatory and opioid activity in octopus nerves (Voigt and Martin, 1986). This molecule contains both the FMRFamide sequence and the pentapeptide sequence YGGFM of the opioid enkephalin (Brownstein, 1980) and, thus, may be responsible for the coexistence of enkephalin-like and FMRFamide-like immunoreactivity (Takayanagi and Takeda, 1988). Thus, this octopus peptide seems to provide the final link between the two peptide families.

Homology between FMRFamide and opioid peptides has also been demonstrated at the genetic level by Taussig and Scheller (1986). They showed that in *Aplysia californica*, the gene coding for the FMRFamide precursor molecule contains sequences coding for peptides related to the opioid peptides.

1.3.3 Small Cardioactive Peptide b (SCPb)

In 1982, the small cardioactive peptide SCPb (Met-Asn-Tyr-Leu-Ala-Phe-Pro-Arg-Met-NH_2) was isolated from the opisthobranch gastropod, *Aplysia* (Morris et al., 1982). With the discovery of the sequence similarity between SCPb and FMRFamide, a new peptide family including both peptides was suggested, and the proposed minimum structural requirement for biological activity was Phe-a-Arg-b-NH_2 (where a and b are hydrophobic amino acids). However, SCPb has subsequently been found to be inactive on the *Busycon* radula protractor muscle and on the *Mercenaria* heart, both classic FMRFamide bioassays (Price and Greenburg, 1980).

Furthermore, SCPb did not react in a FMRFamide radioimmunoassay, and the two peptides were demonstrated to have distinct distributions within the central ganglia of *Aplysia* (Lehman et al., 1984). In conclusion, the structural similarity between SCP_b and FMRFamide appears to be fortuitous as the distribution and activity of the two peptides are different (Lehman et al., 1984) and thus no homology was found to exist.

1.4 Aims of This Thesis

The general aims of this thesis were the isolation and characterization of the FaRP-encoding gene or genes in the crayfish *Procambarus clarkii*. To this end, the objectives of this thesis were threefold. The first was to optimize the amplification of *Procambarus* cDNA using degenerate oligonucleotide primers designed to recognize the two sequenced *Procambarus* FaRPs. The second objective of this thesis was to determine the efficacy of these primers in the characterization of amplification products by sequencing reactions. Finally, the efficacy of these degenerate oligonucleotides in their priming of second strand cDNA synthesis for insertion into a selective cDNA library was studied. This study may provide a molecular basis for further FaRP or crayfish investigations.

1.4.1 Crustacean investigations

Crustaceans have been useful in scientific investigations because of their small size, ease of care and availability (Cooke and Sullivan, 1982). More importantly, their relatively small neuronal complement allows individual behaviours to be studied at the neuronal level (Beltz and Kravitz, 1986).

1.4.2 FaRP investigations

FaRP investigations in general are of interest, as they serve to determine the range and variety of FaRPs which exist, as well as indicating genetic relationships between species and between neuropeptide families. By investigation of the FaRP-encoding genes and mRNA of an organism, information is received concerning the genetic and functional structure of the organism of interest.

1.4.3 Molecular investigations

Determination of the number of FaRP encoding genes, their characterization, and investigation into the sites of this gene transcription often forms a basis for further research.

Since homology implies inheritance of common features from common ancestors, detection of homology is best achieved by comparing the genes encoding the precursor molecules. Thus, determination of the structure and organization of FaRP genes in various organisms may reveal the relationship between different species. As well, such studies may provide information with respect to the evolution of FMRFamide and related peptides in other organisms, reveal a time scale of origin of the FaRP peptides, or suggest the structure of a more general FaRP precursor.

Determination of the structures of DNA and mRNA also aids in the understanding of the mechanisms and patterns of both evolution and regulation. For example, conserved regions between different species in the putative promoter region are useful in defining DNA regions that may regulate the expression of the gene and to determine such characteristics as whether the copy number of the peptides within a precursor is critical for its functions, and the extent to which slightly divergent peptide sequences mediate different peptide functions.

Different methods of identifying and charcterizing a specific gene or segment of DNA exist. In the search for FaRP-encoding genes, several of these methods have been employed. Schneider and Taghert (1990) report that using a combination of genomic sequencing, Northern blot analysis, and primer extensions, the organization and expression of the FaRP gene in *Drosophila* was determined. cDNA synthesis, cloning and Maxam-Gilbert sequencing procedures were also used in the characterization of FaRP genes in *Lymnaea* (Linacre et al., 1990) and *Aplysia* (Schaefer et al., 1985).

1.5 Further Investigations

Characterization of FaRP-encoding genes and RNA has previously been used towards the long term goal of genetic analysis of neuropeptide expression and function (Schneider and Taghert, 1988). Sequencing of the DNA and the mRNA are needed before a complete explanation of the control of the expression of FaRPs in this species can be obtained (Robb and Evans, 1990). For example, by defining the transcription start site and promoter region of the FMRFamide gene, neuropeptide gene regulation can thus be studied.

The possibility of genetic manipulation in simple organisms does exist, and provides a complementary means of studying neuropeptide function. By the possible disturbance of cellular distribution of FaRP gene expression, alteration of the compositon of neuropeptides in their precursor form, or by the abolishment of all FaRP expression, new insights into the roles of neuropeptides on animal physiology, behaviour, and development could be gained. However, the first step to these procedures is the isolation and examination of the FaRP-encoding gene or genes.

MATERIALS AND METHODS

2.1 Animal Care

Crayfish *Procambarus clarkii* were obtained from Atchafalaya Biological Supply Co. (Raceland, LA) and were kept at 15°C in aerated freshwater tanks and fed a mixed vegetable diet prior to use. RNA was extracted from crayfish 3-5cm in length.

2.2 RNA Isolation

2.2.1 Crayfish dissection

Prior to dissection crayfish were cooled on ice for 20 minutes. During the dissection gloves were worn to prevent possible contamination with RNAses, and instruments used in the dissection were rinsed with 0.5M EDTA.

Animals were humanely sacrificed by cutting and destroying the brain with scissors. Animals then underwent dissection to isolate suboesophageal and thoracic ganglia, abdominal ganglia, deep abdominal flexor muscle, and heart.

2.2.2 RNA isolation procedure

RNA was isolated from tissue by means of the acid guanidium isothiocyanate-phenol-chloroform extraction procedure of Chomczynski and Sacchi (1987). In this procedure tissue was homogenized with 1mL of denaturing solution (5M guanidium isothiocyanate, 25mM sodium citrate, pH7, 0.5% sarcosyl, 0.1M β-mercaptoethanol). Guanidium isothiocyanate is a potent denaturing agent which readily denatures proteins (Cox, 1968), resulting in disintegration of

cellular structures and nucleoprotein dissociation from nucleic acids as the protein secondary structure is lost.

In the procedure, RNA was separated from other cellular components by the addition of acidic sodium acetate (which lowers the pH, rendering DNA insoluble in the aqueous phase), phenol, (which disrupts the tertiary structure of proteins, resulting in protein aggregation), and 49:1 chloroform:isoamyl alcohol (which permits further protein extraction, and also prevents RNA from fractioning into the phenol phase). Centrifugation at 10,000 x g was performed to separate RNA in the aqueous phase from DNA and proteins in the interphase and phenol phase. RNA was then isolated from solution by precipitation in 70% alcohol, purified by resuspension in the denaturing solution and subsequent alcohol precipitation. The pellet was solubilized for storage in 0.5% SDS, a detergent which weakly inhibits ribonuclease actions (Chomczynski and Sacchi, 1987).

In this procedure and in all procedures involving the use of RNA, all glassware, tubes, and pipette tips were treated with diethyl pyrocarbonate (DEPC), which is a strong inhibitor of RNAses (Fedorcsak and Ehronberg, 1966). Solutions used were made with 0.1% DEPC-treated H_2O. Equipment was soaked overnight in a 0.1% DEPC solution at 37°C, then autoclaved for 15 minutes to remove traces of DEPC which might otherwise have modified purine residues in RNA (Maniatis et al., 1989).

2.2.3 Formaldehyde gel electrophoresis

RNA was separated according to size by means of formaldehyde gel electrophoresis. This separation is possible because at pH 7, RNA is negatively charge and moves with a mobility dependant upon the density of the agarose gel and upon its own fragment size. The formaldehyde

gel was made up with 12.5mL formaldehyde, 7mL 10 x running buffer, and 0.56g agarose in 50.5mL DEPC-treated H_20. This mixture with the exception of formaldehyde, was heated until boiling, cooled to approximately 50°C, formaldehyde added, and poured into a casting tray with the open ends taped and a well comb in place. After gel solidification, the tape and comb were removed and the tray placed in an electrophoresis chamber with the well nearest the cathode. 550mL of running buffer consisting of 55.0mL of 10 x running buffer, 391.8mL DEPC H_20 and 103.2mL of 37% formaldehyde was poured into the chamber, covering the gel. Formaldehyde was present in both the gel and the running buffer due to its role as a hydrogen bond inhibitor which permits the RNA to remain single-stranded. To each 10-20μL sample run on the gel, an appropriate amount of 5 x formaldehyde gel-loading buffer (80% formamide, 5% glycerol, 1mM EDTA pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol FF) was added. After loading the gel, RNA fragments were separated by electrophoresis at 60V for approximately two hours.

2.2.4 Determination of concentration and purity of total RNA

The purity of total RNA was examined by denaturing agarose gel electrophoresis. In this procedure RNA is separated according to size. Low molecular weight smears represent nucleic acid degradation.

Whereas gel electrophoresis shows DNA contamination and nucleic acid degradation, UV spectrophotometry permits nucleic acid quantitation and indicates if protein and phenol contamination are present. RNA must be free of DNA contamination in order to receive accurate spectrophotometer quantitation, as DNA absorbs light at the same wavelength as does RNA.

In ultraviolet (UV) spectrophotmetry, 1μL of RNA solution is mixed with 199μl of DEPC-treated H_2O. Background absorbance for the sample is determined by scanning DEPC-

treated H_2O along a 240nm -320nm wavelength interval. The adjusted absorbance of the diluted RNA sample is read along the same interval in optical density (O.D.) units. The concentration of RNA can then be determined using the equation 1 $O.D._{260}$ = 40μg/mL (Berger, 1987). RNA purity is assessed by determination of the ratio of absorbance at 260nm/280nm. Values for RNA solutions of 1.9 to 2.0 are acceptable. Lower values are usually indicative of phenol or protein contamination (Berger, 1987). In order to render cuvettes RNAse-free, cuvettes are washed briefly in 50mM NaOH, and then rinsed with DEPC-treated H_2O (Promega technical bulletin).

2.2.5 mRNA isolation

Isolation of mRNA from total RNA was performed using the Poly ATtract mRNA Isolation System (Promega). This system uses a biotinylated oligo(dT) primer which hybridizes to the 3' poly (A) region of mRNA. Streptavidin coupled to paramagnetic particles is then added to solution, binding to the biotinylated oligo(dT)-mRNA hybrids, and separating mRNA from the total RNA in solution by magnetization of the paramagnetic particles. The solution is removed, the paramagnetic particles are washed to remove any contaminating tRNA or rRNA, and the purified mRNA is eluted from the solid particles by addition of ribonuclease-free deionized water, which breaks existing hydrogen bonds between the mRNA poly (A) tail and the oligo(dT) primer.

2.2.6 Determination of mRNA concentration and purity

UV spectrophotmetry can be used to determine mRNA concentration and purity in the same manner as for total RNA. However, when mRNA has been isolated from less than 1mg of total RNA, absorbance must be read directly without sample dilution. 40μg/mL of mRNA has

an absorbance of 1 at 260nm, while pure mRNA has an A_{260}/A_{280} absorbance ratio of 2 or more (Promega).

2.3 cDNA Synthesis

2.3.1 First strand cDNA synthesis

First strand cDNA was synthesized from mRNA according to the procedure of Maniatis et al. (1989). In this procedure, oligo(dT) binds to the poly(A) tail of mRNA, and reverse transcriptase incorporates the appropriate dNTPs into the growing cDNA strand. RNasin is present to prevent RNA degradation, while 2.5mM $MgCl_2$ optimizes oligo(dT) primer binding and enzyme activity (Innis, 1990). When working with single-stranded cDNA, it is essential that all tubes and pipette tips used are siliconized using dichlorodimethylsilane to avoid sscDNA adherance to these surfaces.

2.3.2 Second strand cDNA synthesis

Second strand cDNA for selected FaRP sequences was synthesized according to the procedure of Maniatis et al. (1982). In this procedure the specific mixed oligonucleotide was used to prime synthesis of selected second strands of cDNA with DNA polymerase. Second strands were completed with reverse transcriptase (Maniatis et al., 1982).

Mung Bean Nuclease was used to digest remaining single stranded DNA. Varying amounts of Mung Bean Nuclease (0U, 0.5U, 1U, 2U) were used in reaction aliquots so as to optimize the reaction; sufficient amounts of enzyme must be used so that maximal degradation of single stranded DNA occurs, with minimal degradation of the double-stranded cDNA

molecule. After the Mung Bean Nuclease reaction was performed, a subsequent fill-in reaction was performed so as to render full-length any double-stranded cDNA molecules which may have been somewhat degraded by excessive Mung Bean Nuclease activity (Maniatis et al., 1982). The resulting double-stranded cDNA was then methylated to protect internal restriction enzyme sites (Maniatis et al., 1982) and was ligated to radioactively phosphorylated linkers (Maniatis et al., 1982), duplex molecules which are blunt at both ends and which possess specific restriction enzyme sites which can be ligated to compatible termini in a vector, and so greatly increase cloning efficiency (Wu et al., 1987). Linker-dscDNA molecules were then digested with the EcoRI restriction endonuclease in order to cleave duplex linkers and create cohesive ends. Excess linkers, contaminating dNTPs, mRNAs and single stranded cDNAs were removed from the linker-dscDNA molecules by product purification through A5M-columns. In this procedure, samples are run through an A5M-column, and aliquots of column effluent are taken. The radioactivity in cpm units of these aliquots is taken and those samples which comprise part of a single inital rise in radioactivity are taken and ethanol precipitation is performed on them. These are the samples of interest as they contain the ^{32}P-labelled linker-dscDNA molecules, and precede the unincorporated ^{32}P movement, observed as a final large increase in radioactivity, through the A5M-column.

2.4 Polymerase Chain Reaction

The polymerase chain reaction (PCR) is a technique often used to overcome the limitations of sequence analysis due to low product amounts or the presence of extraneous material, by selectively amplifying of specific DNA sequences by a factor of 10^6 (Saiki et al., 1985; Mullis and Faloona, 1987). The PCR procedure uses two oligonucleotide primers that

flank the DNA segment to be amplified in continuing cycles of heat denaturation of DNA strands, annealing of primers to their complementary DNA strands, and extension of the primed DNA strands. The oligonucleotide primers bind to opposite DNA strands, and are oriented such that DNA synthesis proceeds across the region between the two primers, doubling the amount of DNA of interest. Additionally, since the PCR products themselves are capable of binding to the primers, each successive cycle of PCR effectively doubles the amount of DNA synthesized in the previous cycle (as reviewed by Saiki et al., 1988).

2.5 PCR-Rapid Amplification of cDNA ends (PCR-RACE)

Conventional PCR is limited in its use as it requires sequence information on both sides of the region of interest. Thus, to permit amplification in which only a single short stretch of sequence within an mRNA is known, Frohman et al. (1988) introduced the PCR-rapid amplification of cDNA ends (PCR-RACE) procedure. In this technique, cDNAs are generated using the PCR technique to amplify copies of the region between a single point in the transcript and the 3' or 5' end of mRNA. For amplification of the mRNA 3' end, a single specific primer which binds within the mRNA, and a general non-specific-type primer which binds to the 3' poly (A) tail of the mRNA are used, as illustrated in Figure 3.

For this thesis, PCR-RACE amplification of the 3' ends of mRNA was the method of choice. No information exists as to the the nature of sequences within or flanking FaRP-encoding gene or genes in the crayfish *Procambarus clarkii*. Thus, FaRP sequences were deduced from the common NFLRF(Gly) sequence of the two known crayfish FaRPs, NF_1 and DF_2, and the codon bias found in the *Drosophila* FaRP gene (Schneider and Taghert, 1988; Akashi, 1994) was employed to generate a 16-fold degenerate oligonucleotide primer 17 bases in length (see Table

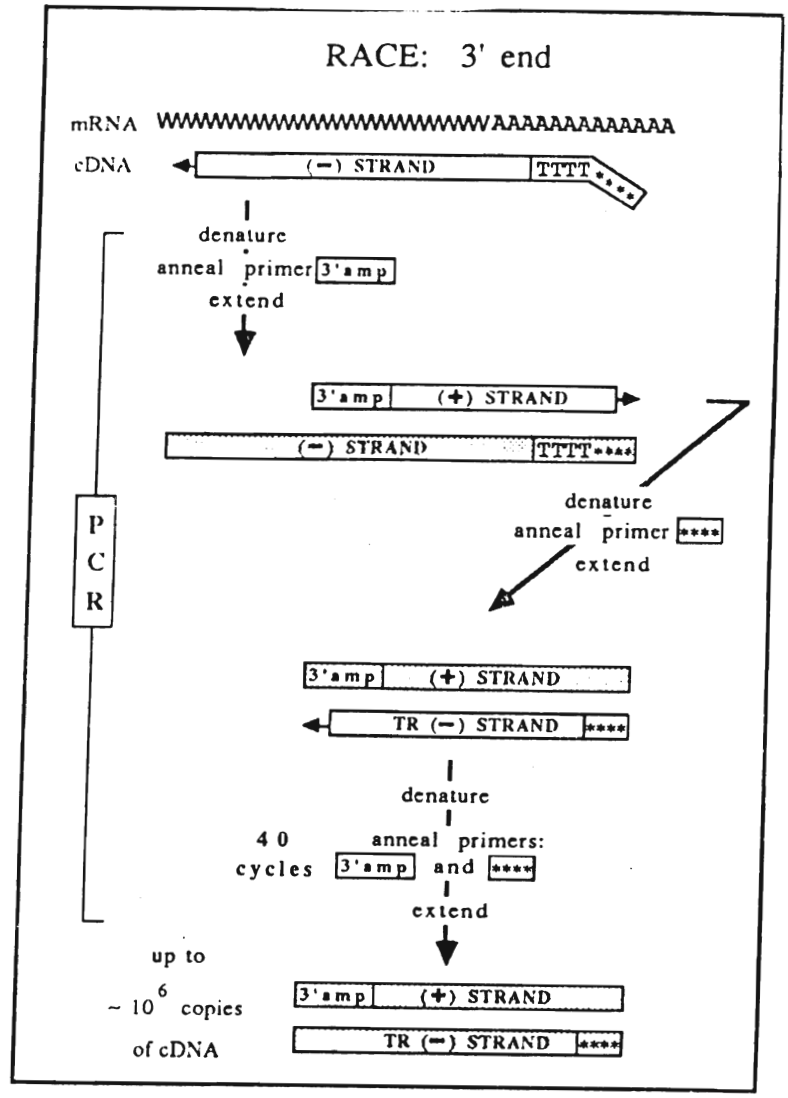

Figure 3: The general (dT)17adaptor-primer is representaed by TTTT****, while the cDNA primer, specific to the gene of interest and complementary to the (-) strand, is labelled 3' amp. Open rectangles represent DNA strands actively being synthesized, while shaded rectangles represent DNA previously synthesized. A (-) or (+) strand is referred to as "truncated" (TR) when that strand is shorter than its repective original (-) or(+) strand. This diagram has been simplified at each step to illustrate only how the strand synthesized in the previous step is used (reproduced from Frohman et al., 1988).

III). The non-specific primer chosen was the NOT1 adaptor primer (Promega) due to low annealing temperatures required with oligo(dT) use (Frohman et al., 1988). The PCR-RACE procedure has some inherent disadvantages and difficulties which may be overcome by adjustment of reaction parameters. For some mRNAs, not all of the 3'end cDNAs generated are full length (Frohman et al., 1988). This problem can be overcome by either increasing extension time, or if short strands are the result of dT primer binding to an A-rich region in the coding sequence upstream of the poly (A) tail, then primer concentration can be decreased. (Frohman et al., 1988). It has also been observed that some of the PCR-RACE reaction products are single-stranded cDNAs. This has been suggested to occur because use of a gene-specific primer and a primer that binds to every cDNA present may lead to greater production of one of the strands, or to mismatched pairing of some strands after the final round of PCR (Frohman et al., 1988). This problem can be overcome by altering buffer conditions. Due to cloning and amplification biases in favour of short molecules, if more than one sequence is amplified then the shorter cDNA species are overrepresented. Another problem inherent in the PCR-RACE procedure is the occurrence of non-specific amplification, where sequences other than those of interest are amplified due to primer redundancy or mismatches (Frohman et al., 1988). In order to optimize PCR reactions, some adjustments can be made to some of the reaction parameters. These include alteration of reaction buffer composition, particularly $MgCl_2$ concentraton, adjustment of primer concentration, dNTP concentration, enzyme concentration, enzyme choice, annealing time and temperature, and extension time and temperature (Saiki, 1989).

2.5.1 Specific primer selection

As neither positions of the FaRP-encoding gene(s) in *Procambarus*, nor adjacent sequences are known, the selection of a specific primer for the PCR-RACE reaction was made

TABLE III

Degenerate oligonucleotide primer sequences

Amino acid sequence of degenerate oligonucleotide primers coding for the amino acid sequence NFLRF(G) used as specific 5' end primer in PCR-RACE 3' end amplification, and as primers for the synthesis of second strands of cDNA in the construction of a selective cDNA library.

N F L R F G

1. AAT TTC AAG AGG TTC GG
2. AAT TTC AAG AGA TTC GG
3. AAT TTC AAG AGT TTC GG
4. AAT TTC AAG AGC TTC GG
5. AAT TTC AAG CGG TTC GG
6. AAT TTC AAG CGA TTC GG
7. AAT TTC AAG CGT TTC GG
8. AAT TTC AAG CGC TTC GG
9. AAC TTC AAG AGG TTC GG
10. AAC TTC AAG AGA TTC GG
11. AAC TTC AAG AGT TTC GG
12. AAC TTC AAG AGC TTC GG
13. AAC TTC AAG CGG TTC GG
14. AAC TTC AAG CGA TTC GG
15. AAC TTC AAG CGT TTC GG
16. AAC TTC AAG CGC TTC GG

from the common known amino acid sequence of the endogenous *Procambarus* FaRPs NF_1 and DF_2. Although these peptides possess a common RNFLRFamide sequence, the oligonucleotide sequence was based upon the NFLRFamide sequence so as to reduce both cost and primer degeneracy, and so as not to preclude primer binding to other possible FaRP sequences. The specific primer synthesized was, as according to the directives of Cooper and Isola (1993) "a "specific" deoxyoliogonucleotide of limited degeneracy that represents the most likely codon combinations following reverse translation of a previously determined C-terminal amino sequence". The codon usage was assumed to be that found in the *Drosophila* FaRP gene (Schneider and Taghert, 1988). As suggested, the specific primer utilizes only the most prevalent codons of an amino acid, uses a singular 3' most codon or partial codon, and does not exceed 16-fold degeneracy. (Cooper and Isola, 1993). As in this case the final amino acid is Glycine (subsequently cleaved to form the NH_2 group) which has a GGX (where X represents T,C,A or G) codon usage, only the initial GG sequence was used in the degenerate primer. This partial codon usage follows the directives of Cooper and Isola (1993) which state that in primer design "terminating one nucleotide from the most highly degenerate third is a great help, ensuring that the last two nucleotides are homologous". Cooper and Isola (1993) also suggest designing of primers such that the amino acid sequence is encoded by the lowest possible number of codons from available protein sequence information. As the Arginine (R) residue is encoded by eight possible codons with similar frequencies of occurence, the N terminal arginine was omitted from the primer sequence; reducing primer degeneracy from 128-fold to 16-fold. This use of a restricted number of codons is an effective means of controlling the loss of amplification product specificity characteristic of degenerate probes (Cooper and Isola, 1993), while PCR is generally

assumed to be acceptably efficient when using primers with 15-20% bp mismatches with the template (Rychlik, 1993).

2.5.2 Determination of oligonucleotide concentrations

The concentration of dissolved oligonucleotide primers was measured by means of UV spectrophotometry. These oligonucleotides, which absorb ultraviolet wavelengths between 250nm and 270nm due to the spectral characteristics of their four composite bases, have specific molar extinction coefficients for these bases which allow oligonucleotide concentrations to be determined. The molar extinction coefficients of oligonucleotides differ from those of longer DNA molecules due to the "hypochromic effect" in the DNA, in which the intensity of absorption is reduced due to the interaction between absorbing units placed in an orderly array (Berger, 1987).

In the determination of oligonucleotide concentration, the molar extinction coefficient for the particular sequence must be calculated, as outlined by Wallace and Miyada, (1987). This value is found by summing the contribution of each base. The molar extinction coefficients at 260nm for the bases are: G, 12,010; A, 15,200; T, 8400; and C, 7050. The concentration of a solution of the oligonucleotide can then be determined by measurement of the absorbance of the solution at 260nm.

2.5.3 Annealing temperature selection

In the selection of a gene-specific primer for the amplification of cDNA 3' ends, the calculated melting temperature (T_m) of a perfect hybrid is determined. However, when using a pool of oligonucleotides which differ greatly in their G+C composition as the gene-specific

primer, it is impossible to determine a representative T_m. Since it is unknown which oligonucleotide sequence will perfectly match the genomic sequence of interest, conditions must be used that allow that oligonucleotide with the lowest G+C content to anneal efficiently. Optimal annealing temperatures are generally 2-4°C less than the predicted T_m of the most A/T rich member of the oligonucleotide pool (as reviewed by Maniatis et al., 1989). The T_m of the most A/T rich oligonucleotide is calculated by means of the equation:

(# of residues of A+T)2°C + (# of residues of G+C)4°C = T_m

Thus, for the sequence AAT TTC AAG AGT TTC GG, the most A/T rich oligonucleotide, the T_m is: (11)2°C + (6) 4°C = 46°C.

The annealing temperature, chosen to be 2°C less than the T_m of the most A/T rich member of the oligonucleotide pool is 44°C. Maniatis et al (1989) note, however, that the use of such "lowest common denominator" conditions can lead to false positives being received due to mismatched oligonucleotides of greater G+C content possibly being more stable than a perfectly matched hybrid formed by the correct oligonucleotide.

2.5.4 Selection of generalized 3' end mRNA primer

In the selection of an adaptor primer for amplification of cDNA 3' end, it must be noted that although long stretches of residues anneal perfectly to the poly (A) tail of mRNA, they do not base pair well at the temperatures used to prevent mismatches of the specific primers (Frohman et al., 1988). The calculated T_m of oligo$(dT)_{17}$ is: (17)2°C = 34°C, 12°C below the T_m of the most A/T rich member of the gene-specific oligonucleotide pool primer. Thus, the NOT1 adaptor primer with the sequence 5'-d[AATTCGCGGCCGCTTTTTTTTTTTTTT]-3' and a T_m of 74°C is used. Use of this primer allows annealing temperature to be optimized according to

the requirements of the gene-specific oligonucleotide pool primer. As Rychlik (1993) states "the optimal annealing temperature range is unusually broad when primers exhibiting low 3'-terminal stability (as evidenced by high T occurrence) are used. Thus, the NOT1 primer, with a long 3' poly(T) sequence, should not experience annealing problems at the much lower temperatures conducive to the degenerate oligonucleotide primers.

The NOT1 primer was chosen as it possessed the qualities desired in an adaptor oligonucleotide: "a closely matched G+C content, ... and a lack of sequence homology with the unique oligonucleotide" (Cooper and Isola, 1993), as well as a specific restriction endonuclease site which aids in detection of specific cDNA clones (Kuehl and Battey, 1993).

Adaptor primers in general do not bind to mRNA strands at their GC-rich end in the initial annealing cycle of PCR-RACE, but in subsequent amplification cycles, bind completely to transcribed cDNA molecules at a defined site. This avoids the production of cDNAs of varying length found in oligo(dT) primer binding to different sites along the poly(A) tail of mRNAs and cDNAs (Moqadam and Siebert, 1994). The increased length and binding specificity of these molecules, as well as their use at a higher annealing temperature as compared to oligo(dT), make these primers much more useful in the amplification of large amounts of specific 3'end cDNA sequences.

2.5.5 Determination of optimal PCR-RACE temperatures

PCR-RACE involves incubation of samples at the three temperatures needed for strand denaturation, primer annealing and extension to occur. Double-stranded DNA is typically denatured by briefly heating the sample to 90-95°C; annealing occurs generally at 40-60 C according to the T_m of the most A/T rich primer sequence ,while extension take place at 72°C

(Saiki, 1989). The temperatures in the PCR-RACE procedure used were 95°C denaturing temperature, 44°C annealing temperature, and 72°C extension temperature.

2.5.6 Determination of optimal denaturation, annealing and extension times

The length of time that the sample is exposed to the specific temperature of each cycle is determined by several factors. Saiki (1989) states that "extensive denaturation is unnecessary, and limited exposure to elevated temperatures helps maintain maximum polymerase activity throughout the reaction". As well, long annealing times are unnecessary as hybridization is almost instantaneous due to the large molar excess of primer present in the reaction mix. The length of time for primer extension can vary according to the length of the sequence being amplified. The average rate of primer extension is 1000 bp per minute (Saiki, 1989). Thus, sufficient time must be given for the complete sequence to be created. Conversely, excessive incubation time will permit non-specific amplification products to be produced in large numbers (Saiki et al., 1988).

As the length of the mRNA sequence of interest was unknown, but estimated to be maximally 1.7kb in length, as observed in the FaRP genes of other invertebrates, an extension time of 2 minutes per cycle was chosen. A relatively long annealing time of 1 minute was chosen as this process may take longer for the pool of oligonucleotides which may be annealing at temperatures well below their T_m. The length of time of denaturation was also chosen to be 1 minute.

At the commencement of the PCR-RACE procedure, a 5 minute denaturation was performed in order to ensure that complete denaturation occured and that any possible contaminating enzyme activity was stopped. A final 10 minute extension was performed so as to permit full length final products to be produced.

2.5.7 Determination of optimal number of PCR-RACE cycles

The number of PCR amplification cycles used is determined by the amount of product needed, with an increasing number of cycles resulting in increasingly larger amounts of product. However, Chelly et al., (1988) found the "the rate of amplification was exponential for 20 to 30 cycles, after which it decreased drastically and reached a plateau", dependant upon the number of starting molecules. Thus, 30 cycles of 95°C extension for 1 minute, 44 C annealing for 1 minute, and 72°C extension for 2 minutes were performed.

2.5.8 cDNA enrichment

As a mixed oligonucleotide primer was being used to amplify a sequence or sequences which may be in low abundance, a process of "cDNA enrichment" was performed. In this process, three initial PCR-RACE cycles were performed at lowered annealing and extension temperatures, permitting sequences with varying degrees of homology to the different oligonucleotides to be amplified. This results in an "enriched" source of likely sequences for amplification at higher, more stringent temperatures (Delort et al., 1989).

2.5.9 Determination of optimal cDNA, enzyme and primer concentrations

Mullis (1989) states that "the success of the PCR reaction depends on the kinetic advantage that high concentrations of primers have over relatively low concentrations of product strands, which at equilibrium would rehybridize with each other and displace the primers". Thus, primer concentrations between 0.1μM and 1μM, and cDNA concentrations of 200ng/mL are generally used (Saiki, 1989). Enzyme concentrations must also be comparatively high, as sufficient enzyme must be available for primer extension immediately after primer hybridization

to acheive high-gain amplification. Thus, enzyme concentrations of 5-15 units(U)/100μL are commonly used (Mullis, 1989). Consequently, primers used in these reactions were at 0.3uM concentrations, while enzyme was at a 5U/100μL concentration.

2.5.10 Determination of optimal dNTP concentrations

The concentration of the four dNTPs used (dATP, dCTP, dGTP and dTTP) was 0.2mM each. Saiki (1989) states that "of each dNTP, 200μM is sufficient substrate to synthesize over 20 mg of DNA in a 100μL reaction and generally offers a good compromise between yield and fidelity."

2.5.11 Determination of optimal buffer composition

Composition of polymerase buffer is an important factor in determining the success of PCR amplifications. The working concentration of KCl is generally 50mM (Saiki, 1989), a concentration which facilitates primer annealing without the accompanying inhibition of polymerase activity that occurs at higher KCl concentrations (Innis, 1990). Tris.Cl (between pH 7.3 and 8.8) is generally used at concentrations of 10 - 50mM (Innis, 1990). Saiki (1989) reports that "optimal levels (of NaCl) will vary depending on the sequence being amplified and the nature of the primers. In most cases, best results with genomic targets are obtained in reaction with 0.5 - 1.0 mM free magnesium. (Because deoxynucleotide triphosphates quantitatively bind Mg^{2+}, the amount of free Mg^{2+} available for Taq polymerase is the concentration of $MgCl_2$ less the total

concentration of dNTPs)." Thus, with each dNTP at a working concentration of 0.2mM, a working concentration of 1.5mM $MgCl_2$ is added, resulting in a free Mg^{2+} concentration of: 1.5mM $MgCl_2$ - 4(0.2mM dNTPs) = 0.6mM free Mg^{2+}.

Gelatin, which acts as an enzyme stabilizer in PCR reactions (Innis et al., 1988) was present at a recommended working concentration of 100mg/mL (Innis, 1990).

Addition of 10% glycerol to PCR buffer was reported to maximize efficient specific amplification (Li and Negre, 1993), and thus the addition of 10% glycerol was assayed to determine if it increased product yield.

2.5.12 PCR - RACE Procedure Used

The final PCR-RACE procedure determined was the addition of, in the order given:

31.5 µL sterile H_2O

5 µL 10 x PCR buffer

1 µL each dNTP (@ 10mM)

0.5 µL Taq polymerase (@ 2.5U/mL)

3 µL NOT1 primer adaptor (@ 5mM)

2 µL mixed oligonucleotide primer (@ 5mM)

3 µL cDNA (@ 3.3 ng /µL)

Sterile mineral oil was thinly layered over this reaction mixture to prevent evaporation of the reaction mixture, and the 500 µL PCR tube smeared with Type 25 Silicone Compound (GC Electronics) so as to maximize heat conductance from the thermal cycler to the tube. A Hybaid^R thermal cycler was used for PCR reactions, with an amplification program of:

1 cycle of 95°C (5 min) 27°C (1 min) 45°C (2 min)

2 cycles of 95°C (1 min) 27°C (1 min) 72°C (2 min)

30 cycles of 95°C (1 min) 44°C (1 min) 72°C (2 min)

1 cycle of 44°C (1 min) 72°C (10 min)

Immediately after completion of the PCR-RACE reaction, samples were placed at -20°C.

2.5.13 Second amplication of PCR-RACE products

In the second amplification or reamplification of PCR-RACE products (Cooper and Isola, 1993), identical PCR-RACE reaction components and procedures were followed, except that the 3 μL of sample cDNA used in the first PCR-RACE procedure was replaced with 3μL of the initial PCR-RACE product.

2.6 PCR-RACE Product Recovery

2.6.1 Agarose gel electrophoresis

An appropriate volume of 6 x sample buffer (20% glycerol, 2% sodium dodecyl sulfate (SDS), 0.5% Bromophenol blue in sterile H_2O) was added to 10 μL of PCR-RACE product solution. DNA was separated according to size on 1.2% agarose gel made with 0.72g agarose and 1.2 mL 50 x Tris Acetate EDTA (TAE) running buffer (2M Tris Acetate, 0.05M EDTA pH 8.0) in a total volume of 60 mL. DNA was separated alongside molecular weight standards at 60V for 3 hours. The gel was immersed in 0.1 μg/mL ethidium bromide for 20 minutes in order to stain the DNA, then immersed in distilled water to remove any unincorporated ethidium bromide. The stained gel was then placed under UV light and photographed. If DNA bands were to be recovered from the gel, low-melt agarose was substituted for regular agarose.

2.6.2 Recovery of DNA from low-melt agarose gels

DNA bands of interest were cut from low-melt agarose gel with a razor blade, then isolated from the agarose by either the Glassmax DNA isolation and purification procedure (Geneclean) or the inexpensive alternative to glassmilk for DNA purification (Boyle and Lew, 1995). The underlying principles in both of these procedures are the binding capacity of silica-based matrices for DNA from neutral solutions, washing of the bound DNA to remove soluble impurities, and elution of cleaned DNA from its matrix into a solubilizing H_2O solution. The small amounts of DNA recovered were not able to be quantitated spectrophotometrically. DNA samples instead underwent fluorescence intensity comparison with standards of known DNA concentration in the ethidium bromide fluorescent quantitation of the amount of double-stranded DNA method of Maniatis et al. (1989).

2.7 Cloning of PCR-RACE Products

The study of individual pieces of cellular DNA only became possible with the advent of recombinant DNA procedures (Watson et al., 1987). DNA fragments containing sequences of interest, such as PCR-RACE products, can be replicated in bacteria or mammalian cells in a procedure termed cloning which allows sequences of interest to be amplified so that they can be obtained in sufficient quantities for nucleotide sequence analysis and other uses (Miller et al., 1983). In the cloning procedure, plasmids, which are small circular DNA molecules that replicate independently of host cell DNA, are cleaved at specific sites by specific restriction endonucleases and ligated to sequences of interest by ligase enzymes, and the recombinant plasmids are inserted into competent host bacterial cells. (Miller et al., 1983). The ability of plasmids to confer

antibiotic resistance to their host is useful for DNA cloning, since this provides a selection mechanism for obtaining the desired clones.

Selection of appropriate plasmids and host cells is essential for optimal cloning efficiency of the received products. Bacterial plasmids can range in size from 1kb to more than 200kb, and may confer phenotypes upon their bacterial hosts such as resistance to antibiotics, production of antibiotics, degradation of complex organic compounds and production of coli and enterotoxins, and restriction and modification enzymes (Maniatis, et al., 1989). The copy numbers of different plasmids in their hosts vary from 1 to 700 per cells. Different plasmids use different subsets of enzymes to duplicate the bacterial chromosome for their own replication, and so copy number can be influenced by the "compatibility" of the enzyme substrates used by the bacterial host and the chosen plasmid. (Maniatis et al., 1989). Plasmids vary in size, with smaller plasmids being preferred for many reasons. As efficiency of transformation is inversely related to plasmid size, smaller plasmids are preferred as they can accommodate larger segments of DNA before the efficiency of transformation begins to deteriorate. Larger plasmids replicate to lower copy numbers and so yield less inserted DNA. Finally, larger plasmids are more difficult to characterize by restriction mapping. Thus, the number, placement and identity of restriction endonuclease sites on the plasmid, as well as plasmid size and copy number within the host are all important determinants in choosing a particular plasmid. Host cells must also be determined with regards to growth rate, compatibility with the plasmid vector both in regards to copy number within the host and ease of detection of recombinant cells, and the host cell degree of competency (Maniatis, 1987).

It is necessary to ascertain whether the DNA sequences inserted into the plasmid vectors are those of interest. Hybridization techniques are used to ascertain that the plasmid insert

sequences are the amplified products themselves, that they recognize crayfish DNA sequences, and that they recognize sequences in RNA isolated from neuropeptide-transcribing tissues.

2.7.1 Plasmid isolation and purification

The plasmid used in the cloning of PCR-RACE amplification products was pUC 19 (represented diagramatically in Figure 4), which contains a single Hinc II restriction endonuclease site within its multiple cloning site, and for which specific forward and backward primers immediately outside the multiple cloning site are readily available.

Linearization of plasmid DNA at the Hinc II site was performed by means of incubation with Hinc II enzyme and its appropriate restriction enzyme buffer at 37°C until complete linearization of the plasmid is achieved; as judged by the size of plasmid bands on agarose gels.

As Maniatis et al. (1987) state "during ligation *in vitro*, bacteriophage T4 DNA ligase will catalyze the formation of a phosphodiester bond between adjacent nucleotides only if one nucleotide contains a 5'-phosphate group and the other contains a 3' hydroxyl group". Thus, in order to ensure that plasmid contaminants did not ligate to the plasmids, and also to prevent plasmids from religating themselves, dephosphorylation of plasmids was performed in some cases. Thus, DNA sequences of interest with 5' terminal phosphates can efficiently ligate to the dephosphorylated plasmids. The dephosphorylation procedure used was that of Maniatis et al. (1987) in which calf intestinal alkaline phosphatase (CIP) is used to remove 5' phosphate groups from linear DNAs.

RACE products were rendered blunt-ended by the Klenow fragment DNA Polymerase I reaction (Maniatis, 1989) in some cases in order to optimize blunt-end ligation to Hinc II linearized pUC 19. The Klenow fragment of *E. coli* DNA Polymerase I employs its 5' -> 3'

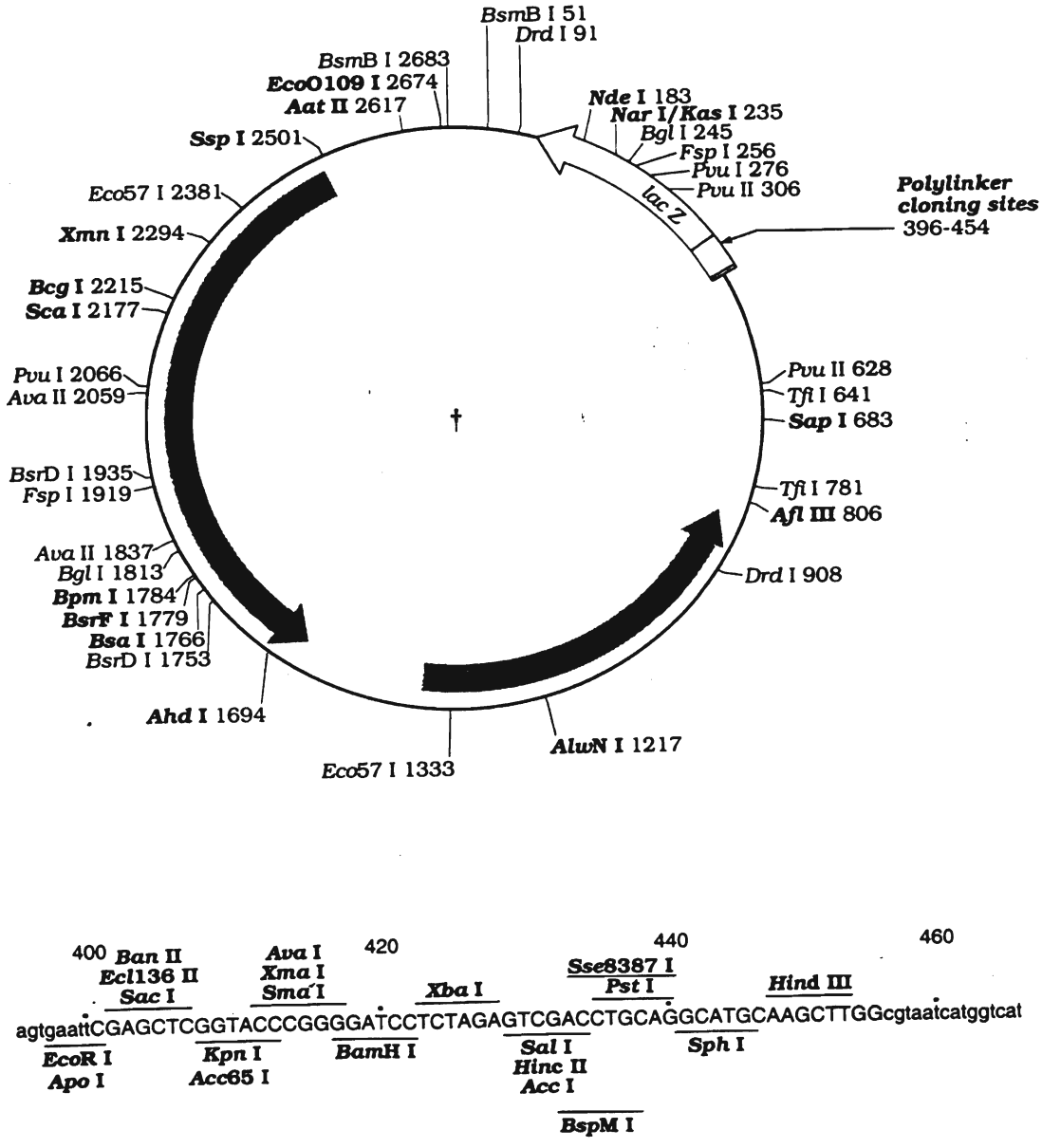

Figure 4: Figure depicting the plasmid pUC 19 used for ligation to PCR-RACE amplification products (as taken from Promega technical bulletin). This figure shows the restriction sites of those enzymes which cut the plasmid either once or twice. Unique restriction sites are depicted in bold type. The positions of the pUC 19 lac Z region (lac Z), origin of replication (ORI), and β-lactamase gene (Ap) are shown. The polylinker cloning site, in which the plasmid was cleaved and amplification products were inserted, is shown in detail at the bottom of the page, with both base sequences and restriction endonuclease sites shown.

polymerase activity to fill in recessed 3' ends of double-stranded DNA (Cobianchi and Wilson, 1987).

2.7.2 Ligation

When the Hinc II-digested pUC 19 plasmid was used in ligations, 20ng of PCR-RACE product was added to 2μL 10 x ligase buffer (250mM Tris-HCl at pH 7.5, 50mM $MgCl_2$, 25% (w/v) polyethylene glycol 8000, 5mM DTT, 4mM ATP), 200ng of Hinc II-digested pUC 19, 1μL T4 DNA ligase @ 400,000 U/mL, and sterile H_2O to bring the total reaction volume to 20μL. This mixture was incubated overnight at 16°C.

Alternate ligation procedures were also used in order to overcome the problems of blunt-end ligation caused by DNA polymerases producing PCR-RACE products with a 3'-dA overhang. The first method involves the restriction of plasmids with Sac I, generating a 3' overhang with a terminal dT which can base-pair with the 3'-dA-overhang of the PCR-RACE product. The second method involves a single-tube ligation in which digestion of plasmid DNA, rendering PCR-RACE products blunt-ended, and ligation of plasmid and PCR-RACE products are accomplished in a single step. In this procedure, Taq DNA polymerase blunt-ends PCR-RACE products by 3' exonuclease activity in the presence of dNTPs, and SmaI digests the plasmid DNA and forces the PCR-RACE product into the plasmid by continuously cutting self-ligated vector DNA (Chuang et al., 1995).

2.7.3 Preparation of competent cells

DH5α cells were grown up overnight from frozen cell stocks (Promega) in LB broth as described by Maniatis et al. (1982). Cells were rendered competent by the procedure outlined by Maniatis et al. (1987). In this procedure, bacterial cells are induced into a transient state of "competence", during which they permit the uptake of foreign DNA. Efficiency of DNA uptake is increased with exposure to combinations of DMSO and reducing agents. The means by which these agents act to render bacterial cells competent is unknown, as is the mechanism by which foreign DNA enters these cells (Maniatis et al., 1987).

2.7.4 Transformation of competent cells

100 μL preparations of competent DH5α cells were transformed with the previously ligated plasmid-insert as per the procedure of Maniatis et al. (1987), in which DNA is allowed to enter competent cells until cells are heat-shocked to end their period of "competence" and to retain the foreign DNA inside the bacterial cells.

2.7.5 Plating of transformed cells

Transformed cells were plated onto LB agar plates (10g/L bacto-tryptone, 5g/L bacto-yeast extract, 10g/L NaCl, 15g/L bacto-agar, 10mM $MgSO_4$ at pH 7.0) with 20mg/L ampicillin and 20mg/L X-Gal) in 10, 50, and 100 μL aliquots. After the solutions had been absorbed into the agar, plates were inverted and incubated overnight at 37°C.

White colonies (postulated to be competent cells transformed with plasmid containing the PCR-RACE product) were picked with a sterile toothpick, and transferred to a gridded LB Amp

X-Gal plate, and also used to innoculate LB Amp media (10g/L bactotryptone, 5g/L bacto yeast extract, 10 g/L NaCl, 20mg/L ampicillin at pH 7.0) which was then incubated overnight at 37°C for maximal colony growth.

2.7.6 Colony lifts

Selected colonies transferred to gridded agar plates were inverted and incubated overnight at 37°C to allow colony growth. Colonies were then transferred onto nitrocellulose membranes for use in colony hybridization as per the protocol of Wahl et al. (1981). This protocol transfers colonies to nitrocellulose membranes, isolates and immobilizes DNA from these colonies and allows the membranes to be stored until they can be used in hybridization experiments. In this standard protocol, a suggested modification was used that after colonies had been transferred to the nitrocellulose membrane from the agar plate, the membrane was placed onto filter paper saturated with 10% SDS for 3 minutes prior to being placed on filter paper saturated with 0.5N NaOH for 5 min. This modification aids in the isolation of plasmid DNA and SDS, a detergent, causes aggregation and binding of host cell wall proteins.

2.7.7 Isolation of plasmid DNA from colonies

The selected colonies used to innoculate 2mL of LB amp medium were grown overnight with shaking at 37°C. Recombinant plasmid DNA was then isolated from the resultant bacterial broth by the "Harvesting and lysis of bacteria - lysis by alkali" procedure of Maniatis et al. (1987). The principles underlying this procedure are that exposure of bacterial cells to EDTA, detergents, and alkali disrupts base pairing, causing the linear chromosomal DNA of the host to denature. Due to their being topologically interwined, strands of closed circular plasmid DNA are unable

to separate from one another. When conditions return to normal, plasmid DNA strands recognize their complimentary sequence, and native superhelical molecules are reformed (Maniatis et al., 1989).

2.7.8 Determination of insert presence and size

Insert presence and size were determined by restriction enzyme analysis of isolated plasmid DNA of interest. DNA from both intact and recombinant plasmids was digested to completion with a single restriction endonuclease which cuts once within the plasmid, to determine insert presence and to provide an indication of its size. The length of the linearized DNA fragments was then compared. Plasmids were also cut with a restriction enzyme that cut the plasmid once at each side of the ligation site, or with two restriction enzymes, each of which cut the plasmid once on opposite sides of the ligation site. When run on an agarose gel, band sizes were again compared to determine insert size and presence.

2.8 Selected cDNA Library Creation

2.8.1 cDNA library creation

The selected cDNA library was formed using the selected EcoRI linker-dscDNA fragments which had been purified on an A5M-column. These fragments were ligated to EcoRI digested λgt11 (Promega) (see Figure 5) as per manufacturer's direction, and the ligated DNA was used in the transformation of Y1088 bacteriophage cells using the Packagene *in vitro* packaging system (Promega).

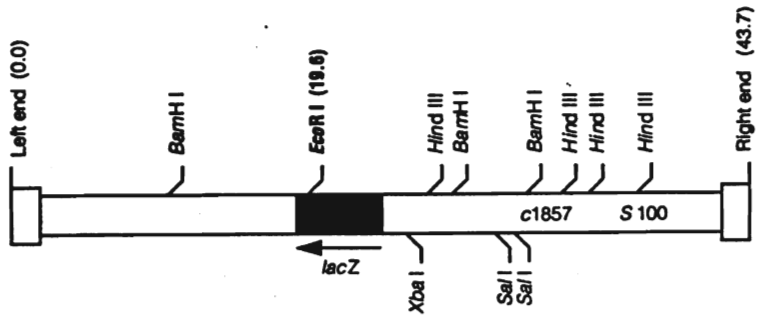

Figure 5: Figure depicting the vector λgt11 used for ligation to selected dscDNA fragments (as taken from Promega technical bulletin), and subsequently cloned in Y1090 cells in the creation of a selective cDNA library. This figure shows the EcoRI site into which the dscDNA fragments were inserted in bold type. Other restriction endonuclease sites are also shown, as is the position of the lac z region.

2.8.2 Titre increase

Phage titre, that number of plaque forming units per mL of phage solution, is increased by adding 200μL of Packagene mixture, 200μL Y1088 cells, and 1mL of LB broth with 0.2% maltose. This mixture is incubated overnight at 37°C. 50μL of chloroform is then added, the mixture is incubated for 15 minutes to break open bacterial cells, and then centrifuged briefly to precipitate cell debris.

2.8.3 Isolation of plasmid DNA from plaques

Recombinant plasmid DNA was recovered from plaques according to the method outlined by Paula Argenta (1993).

2.8.4 Determination of insert presence and size

Determination of the presence of selected dscDNA sequences within the λgt11 vector was performed by means of PCR amplification of the lac z region into which these sequences would be inserted. The primers used were ALC-2 (5' CAGGAAACAGCTATGAC 3'), a reverse primer, and ALC-4 (5' GTAAAACGACGGCCAGT 3'), a forward primer. These primers are 70bp apart from each other and are located on either side of the EcoRI insertion site. As the melting temperatures of ALC-2 and ALC-4 are 50°C and 52°C respectively, the annealing temperatrue was selected to be 48°C. An initial PCR cycle of 95°C(5 min), 48°C(1 min), 72°C(1 min) was performed, followed by 40 cycles of 95°C(1 min), 48°C(1 min), 72°C(1 min) , and a final cycle of 48°C(1 min), 72°C(10 min). Insert size and presence was calculated by determination of the received amplification product band. In order for an insert to be present, the obtained product must be greater than 102bp (the 70bp distance between the ALC-2 and ALC-4 primers + the

32bp size of the two primers themselves). The additional size of the received product band reflects the vector insert size.

2.9 Hybridization procedures

Hybridization was performed on colony lifts and DNA dot blots.

2.9.1 Preparation Of Labelled Probe

The PCR-RACE amplification products used in cloning were also used as probes to ascertain both that these sequences were successfully cloned into the selected recombinant clones, and also to determine if the amplified sequences were present in the RNA of different tissues of the crayfish *Procambarus clarkii*. Amplification products were ^{32}P labelled by Nick Translation (Promega Technical Bulletin). In this procedure *E. coli* DNA polymerase I combines the sequential addition of radioactive nucleotide residues to the 3' - hydroxyl terminus of a nick, generated by DNAse I, with the elimination of nucleotides from the adjacent 5' phosphoryl terminus (Meinkoth and Wahl, 1987).

2.9.2 Measurement of Radiolabel Incorporation

The percentage of radiolabel incorporation into the hybridization probe was determined to evaluate probe viability and to assess the amount of probe required to emit a readily detectable signal.

In this procedure, 1μL of radiolabelled probe was added to 99μL of 0.2M EDTA. In order to measure total probe radioactivity, 10μL of the diluted probe was spotted onto a glassfibre filter which was then placed in a scintillation vial. In order to measure incorporated

radioactivity, another 10μL of diluted probe was added to 100μL of 10mg/mL salmon sperm carrier DNA in a test tube on ice. 3mL of ice cold 10% trichloroacetic acid (TCA) was added to the test tube, which was left on ice for 10 minutes. The solution was filtered onto a glassfibre filter. The test tube was washed 5 x with 3mL of 10% TCA, which was subsequently added to the glassfibre filter. The filter was then washed with cold 95% ethanol, and placed into a scintillation vial. Scintillation fluid was added to both vials, and their respective radioactivities in cpm units were determined using a Beckman LS 1800 scintillation counter. Percentage radiolabel incorporation was determined by the equation:

$$\frac{\text{\# of cpm incorporated radioactivity}}{\text{\# of cpm total radioactivity}} \times 100\%$$

There must be at least 20% radiolabel incorporation for the probe to be used so as to minimize non-specific radiolabel binding and background radioactivity levels.

The amount of incorporated radiolabel per μL of probe is equal to 10 x the # of cpm incorporated radioactivity.

To obtain a signal intensity that is readily detectable, the probe must contain at least 1,000,000 cpm of incorporated radiolabel. Thus, in determining the amount of incorporated radiolabel per μL of probe, the amount of probe needed can be determined.

2.9.3 Hybridization

The hybridization procedure used was adapted from the procedure of Hybaid™. Membranes were prehybridized for 1 hr with agitation at 65°C in sealed pastic bags with 10 mL of prehybridization solution (6 x SSC [20 x SSC stock:3M NaCl, 0.3M Sodium citrate]), 0.5% sodium dodecyl sulfate (SDS), 5 x Denhardt's Solution [1% Ficoll, 1% polyvinyl prolidine, 1% bovine serum albumin] and 50μg/mL salmon sperm DNA boiled for 5 minutes, then chilled on

ice). Prehybridization procedures are used so as to preempt nonspecific nucleic acid binding on the solid supports, and so reduce background hybridization. For hybridization, probe (containing at least 1,000,000 cpm of incorporated radiolabel) was denatured by boiling for 5 minutes, then chilled on ice prior to addition to prehybridization solution. Hybridization was allowed to occur overnight with agitation at 65°C. Stringency washes were performed as per Hybaid technical bulletin, with SSC used instead of SSPE.

Once washed, membranes were placed on a solid support, covered with plastic wrap, and exposed to Kodak X-OMAT film with a Cronex intensifying screen overnight at -70°C. X-ray films were developed by an initial 5 minute submersion in Kodak GBX developer, 1 minute in distilled water, and 5 minutes in Kodak GBX fixer. The film then underwent a final 1 minute wash in distilled water.

2.10 Sequencing

Sequence determination of PCR-RACE amplification products is performed in order to determine if amplification products are FaRP-encoding sequences, to determine base pair homology with known sequences, and to assess stringency of mixed oligonucleotide primer binding. DNA sequence analysis is a molecular cloning technique which determines the nucleotide sequence of DNA. Sequencing is commonly used either to determine the site of mutation of a known gene, or to determine the entire nucleotide sequence of an unstudied DNA segment (Promega technical bulletin). The two sequencing techniques currently in use are the chemical degradation method of Maxam and Gilbert (1977) and the dideoxy method of Sanger et al. (1977).

2.10.1 The Sanger method

The Sanger method, developed in 1977, generally involves the sequencing of DNA by a procedure that involves second-strand DNA synthesis until there is incorporation of a specific radioactively labelled dideoxy nucleotide base, which prevents further chain extension. DNA sequence is found by determining the type of dideoxy base terminally incorporated in sequences of increasing length.

In this method, a template DNA molecule is initially prepared. This single-stranded DNA fragment is then annealed to homologous single-stranded primer. Primer homology is assured as both the template and the primer come from the same original double-stranded DNA fragment. The double-stranded DNA fragment is produced by site-specific cleavage of a DNA polymer. One portion of the resultant fragment sample is treated with another restriction enzyme, which produces a smaller double stranded fragment, which is dissociated to produce two small single stranded fragments, one of which is used as the primer. A second portion of the original fragment sample is used to make a template DNA strand. An exonuclease is used which cleaves single nucleotides from the 3' ends of the double stranded DNA fragment. The exonuclease is then removed, and DNA polymerase is added. With the addition of deoxynucleotides (e.g. dATP, dGTP and CTP) and a dideoxynucleotide (e.g. ddTTP), the DNA strand is re-elongated at the 3' end until the dideoxynucleotide is incorporated (in this case when a T is required to base pair with the A present on the opposite strand). With this dideoxynucleotide incorporation, the 3' end of the template DNA strand is blocked from further elongation as the lack of a 3' hydroxyl residue in dideoxy nucleotides prevents formation of a phosphodiester bond with the succeeding deoxynucleotide (Maniatis et al., 1989). This blocked template solution is heated to denature the template strands, and the solution is combined with the primer solution. At this point, the

template DNA reanneals with the homologous primer DNA, resulting in double-stranded DNA in which the shorter primer can undergo 3' end extension according to the base sequences provided by the template, but the template cannot be extended at its 3' end due to its prior incorporation of a dideoxynucleotide.

A control solution containing the entire range of extended primer lengths is subsequently prepared. In this procedure, the template-primer DNA fragment is combined with DNA polymerase and the four dNTPs, one of which is ^{32}P-labelled to generally identify newly synthesized primer extensions. As the polymerase extends the 3' end of the primer, aliquots are removed at given timed intervals and frozen to stop the elongation reaction. Recombination of these aliquots provides a control solution comprising the entire range of extended primer lengths.

Finally, four separate reactions containing the template-primer solution, three dNTPs (one of which is ^{32}P-labelled), and a given concentration of one ddNTP along with its equivalent dNTP, are run. In each of the four reactions, a different ddNTP is included. As previously, extension of the primer occurs only if the appropriate dNTP, and not its equivalent ddNTP, is incorporated at a specific base pair position. Due to the presence of both the dNTP and ddNTP equivalents, primer extension lengths vary depending on which is incorporated at each specific base pair position. The resultant extended DNA fragments of varying lengths are then hydrolyzed from the primer by treatment with a specific restriction endonuclease. Polyacrylamide gel electrophoresis of the four test solutions and the control separates the fragments according to length, and visualization of these ^{32}P labelled fragments is done using autoradiography. Fragment lengths are compared to the control autoradiograph, and the DNA sequence is determined by reading the base type of the incorporated ddNTPs in fragments of increasing length (Wood, 1984).

Recent modifications and improvements have been made to the Sanger method of sequencing. Several enzymes other than the Klenow fragment of *E. coli* DNA polymerase I can be used for primer extension in ways which can significantly affect the quantity and quality of the DNA sequence of interest (Maniatis et al., 1989). The Klenow fragment has a low processing rate, and so there is often a high background of fragments that terminate due to random disassociation of the polymerase from the template, and not by ddNTP incorporation. As well, this enzyme is not able to travel long distances along the template, thus limiting the lengths of obtained sequences to 250 - 350 nucleotides. Finally, this Klenow fragment cannot efficiently copy regions of high secondary structure in the template. As a result of these problems, other enzymes are now also used in dideoxy nucleotide sequencing. Reverse transcriptase, although not routinely used in sequencing, can be used to resolve problems caused by homopolymeric regions of A/T or G/C in the template DNA. Taq DNA polymerase is also used to determine DNA sequences that normally form extensive secondary structures, as it can work efficiently at high temperatures which preclude secondary structure formation. Finally, both Sequenase TM and Sequenase version 2.0 (which has no 3' -> 5' exonuclease activity and has a higher processivity than Sequenase TM) are used for determining long DNA sequences. Both of these have extremely high processivity, high rates of polymerization, and a wide tolerance for nucleotide used to resolve compression regions in sequencing gels (Maniatis et al. 1989).

Sanger DNA sequencing was originally carried out with [α- ^{32}P] dNTPs. However, β-particle emissions by ^{32}P caused scattering, resulting in large diffuse autoradiograph bands that were difficult to read. As well, ^{32}P decay caused radiolysis of sample DNA, limiting the storage life of the sequencing reactions to 1 or 2 days. The use of [^{35}S] dATP greatly reduced these

problems, as it has reduced scattering due to the presence of weaker β particles, and has less radiolysis. This allows sequencing reactions to be stored up to 1 week (Maniatis et al., 1989).

Both the Maxam-Gilbert and the Sanger methods of DNA sequencing are theoretically sound and functionally useful. The Maxam-Gilbert method is advantageous in that there is a redundant display of the nucleotide sequence. For example, the use of strong guanine/weak adenine display, in which there is generally enough information to distinguish both the G_2 and the A_2, is more useful than just a pure G pattern alone, as the redundant information serves as a check on the identification. (Maxam & Gilbert, 1977). However, this same redundancy makes Maxam-Gilbert's audioradiograms more difficult to read, as it is necessary to deduce the position of T residues by subtracting bands. Measuring in the C track from those in the C + T track. As well, because Maxam-Gilbert sequencing always employs ^{32}P-labelled DNA, the resultant sequencing bands are fuzzier and broader than those obtained using [^{35}S]dNTPs in the Sanger method. This lack of resolution of Maxam-Gilbert sequences limits the amount of reliable sequence in a gel to 200-250 nucleotides. (Maniatis et al., 1989). The Sanger method of sequencing is generally favoured over the Maxam-Gilbert method, as alterations and modifications to this method have been developed allowing for faster, more accurate determination of longer sequences.

2.10.2 Purification of plasmid DNA for sequencing

Prior to use in sequencing, plasmid DNA of interest was purified by one of two different procedures. DNA was isolated by either the "Purification of closed circular DNA by equilibrium centrifugation in CsCl-Ethidium Bromide gradients" procedure of Maniatis et al.(1989) followed

by the removal of ethidium bromide by 1-butanol (Maniatis et al.,1989), or by polyethylene glycol (PEG) purification.

The former procedure involves separation of plasmid and chromosomal DNA due to differences in the amount of ethidium bromide which can bind to linear and closed circular DNA molecules. Ethidium bromide intercolates between bases of DNA and binds, causing DNA helices to unwind. As a result of this unwinding compensatory superhelical turns are formed in closed circular plasmid DNA, which increases density and eventually prevents further intercalation of ethidium bromide. In contrast, linear molecules continue to bind ethidium bromide until saturation is reached, and so forms a separate upper band from the denser closed circular plasmid DNA during density-gradient centrifugation.

In the latter procedure, isolated DNA was washed with 95% ethanol, centrifuged for 10 minutes at 10,000 x g, rewashed in 80% ethanol, and the isolated DNA pellet was allowed to air dry for 10 minutes. This pellet was then resuspended in 0.5mL TE buffer containing 50μg/mL RNAse A, and incubated for 1 hr at 37°C. 0.5mL of 25:24:1 phenol:chloroform:isoamyl alcohol was then added to the DNA solution, and this mixture was centrifuged for 2 minutes at 10,000 x g to separate DNA in the aqueous phase from protein and digested RNA contaminants in the phenol phase. The upper aqueous layer was extracted, with 1mL of 95% ethanol and 100μL of 7.5M NH_4Ac added. This mixture was incubated 30 minutes at -20°C, then centrifuged 10 minutes at 10,000 x g at 4°C, and allowed to dry. The pellet was resuspended in 50μL of TE buffer, 30μL of 20% PEG in in 2.5M NaCl added, and the mixture was incubated on ice for 1-2hrs. The mixture was then centrifuged for 10 minutes at 10,000 x g at 4°C. The PEG supernatant was drawn off carefully, the pellet was washed in 80% ethanol, centrifuged 10 minutes at 10,000 x g, allowed to dry, and finally resuspended in 35μL of sterile water.

2.10.3 Denaturation of double-stranded plasmids

3-5μg of double-stranded plasmid were denatured using the alkaline denaturation method of Haltiner et al. (1985) which uses strong base in order to hydrolyze the H-bands between the two DNA strands.

2.10.4 Sequencing gel preparation

Sequencing gel plates and dividers were thoroughly washed with soap and water and were rinsed with 70% ethanol prior to sequencing apparatus assembly. Plates were washed with 0.1% sodium dodecyl sulfate (SDS), and the "upper" plate (to be removed from the gel after electrophoresis) was treated with Repel-Silane (Pharmacia Biotech) so as to prevent gel adhesion to that plate. The sequencing apparatus was assembled according to directions in the Bio-Rad Sequi Gen(R)Nucleic Acid Sequencing Cell Instruction Manual.

Sequencing gels were composed of 100mL of sequencing gel solution (420g/L urea, 220mL/L of 5 x TBE [54g/L Tris base, 27.5g/L boric acid, 0.01M EDTA, pH 8.0], 100mL/L 30% Bis-acrylamide, 110mL/L 2% sequencing acrylamide). This solution was degassed for 15 minutes, then a gel "plug" was poured by combining 10 mL of sequencing solution with 10μL of N,N,N',N'-tetramethylethylenediamine (TEMED), a polymerization initiator, and 10μL of freshly prepared 25% ammonium persulfate (APS), a polymerization catalyst (Ogden and Adams, 1987). Once this plug was polymerized, the remaining gel solution, combined with 90μL of 25% APS and 90μL TEMED, was poured into the apparatus, a plastic strip was inserted horizontally at the top of the gel to limit gel height, and the gel was allowed to polymerize.

Once polymerized, the gel apparatus was assembled into the complete sequencing apparatus with 350-500mL of 1 x TBE buffer in the lower chamber, and 1.5L of 1 x TBE buffer

in the upper chamber. The upper plastic strip was replaced with a tooth comb which created "wells" for sample insertion into the gel. The gel was run for 30-60 minutes at 1800V prior to sample application in order to acheive optimal gel temperature of 45-55°C.

2.10.5 Sequencing reactions

In the sequencing reactions, specific primers were used to sequence the Hinc II inserts of the pUC 19 plasmids and the EcoRI insert of λgt11. ALC-2 (5' CAGGAAACAGCTATGAC 3'), a reverse primer which binds from position 481 to position 465 of pUC 19, or ALC-4(5' GTAAAACGACGGCCAGT 3'), a forward primer which binds from position 379 to position 395 of pUC 19, were annealed to the denatured double-stranded templates. Second strand synthesis with radiolabel incorporation was allowed to occur, and strand termination with incorporation of specific ddNTPs was performed according to the procedures of USB technical bulletins. Either [α-^{32}P] dATP or [α-^{35}S] dATP were used in reactions.

2.10.6 Running sequencing gels

Samples were heated to 72°C 2 minutes immediately prior to loading onto the sequencing gel. 2-3μL of each sample was loaded onto the pre-warmed gel, which was then run for 2-4hrs at 1800V. When electrophoresis was completed, buffer was removed from the gel apparatus, the clamps were removed from the gel plates, and the upper gel plate was removed, exposing the sequencing gel. Whatman 3M filter paper cut to the size of the gel was placed upon the sequencing gel. The sequencing gel adherred to the filter paper and could then be removed from the "lower" sequencing plate. The gel side of the filter paper was covered with plastic wrap, and was dried in the Bio-Rad Model 58.3 gel drier for 45 minutes at a steady temperature of 80°C.

Once dried, the plastic wrap was removed from the gel, and the gel was exposed to a Kodak OMAT-AR x-ray film overnight at -70°C. X-ray films were developed as previously mentioned. Determined sequences were sent to the EMBL FASTA server in Heidelberg, Germany DE data base (E-mail: FASTA@EBI.AC.UK) for sequence comparison.

RESULTS

Procambarus clarkii total RNA was isolated from the suboesophageal and thoracic ganglia: known FaRP-transcribing tissues. RNA quality was assessed both spectrophotometrically and by gel electrophoresis. mRNA was then isolated from total RNA either by oligo-dT cellulose column purification or by use of the Poly Attract mRNA Isolation System (Promega). First strand cDNA was then synthesized for use in PCR-RACE reactions.

Prior to the use of sample cDNA in PCR-RACE reactions, optimization of PCR and PCR-RACE reactions was performed using control cDNA of known size and sequence. While no observeable products resulted from PCR reactions in which a buffer composed of 500mM KCl, 100mM Tris.Cl and 15mM $MgCl_2$ was used, PCR reactions produced observeable products of correct size when a 10 x buffer composition of 500mM KCl, 100mM Tris.Cl (pH 8.3 at room temperature), 15mM $MgCl_2$ buffer and 0.1% gelatin was used (Figure 6). Correspondingly, it has been reported that the use of gelatin improves enzyme binding (Innis et al., 1988).

Although Li and Negre (1993) report that the use of 10% glycerol maximizes efficient specific amplification, PCR-RACE amplifications of control cDNA of known amplification product band size in the presence of 10% glycerol were observed to have little positive effect or a largely negative effect upon the quantity of PCR-RACE product received (Figure 7). In the amplification of *Ustilago violacea* cDNA between the oligo (dT) type NOT1 primer and the AC2 primer (a control reaction in which products of a known size are obtained), the presence of 10% glycerol in amplification reactions resulted in a much lower yield of product than in those reactions performed without glycerol present. However, in the amplification of *Ustilago* cDNA

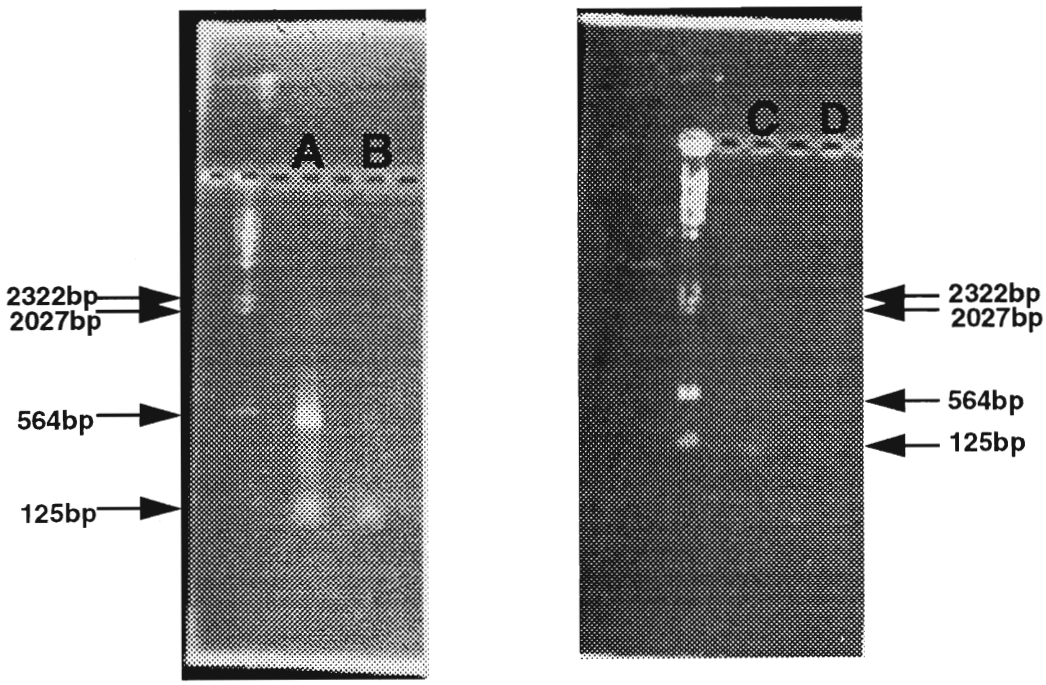

Figure 6: This figure shows the optimization of the polymerase chain reaction (PCR) procedure with regards to reaction buffer composition. Lane C depicts the amplification product of the recombinant plasmid 9-1, using the primers ALC1 and ALC2 and a 10 x buffer composition of 500mM KCl, 100mM Tris.Cl (pH 8.3 at room temperature), and 15mM $MgCl_2$. Lane A depicts the amplification products of 9-1 using the primers ALC1 and ALC2 under identical reaction conditions, with the addition of 0.01% gelatin. Lanes B and D represent negative controls which lack 9-1 DNA. D was amplified in gelatin-free conditions, while B was amplified in the presence of 0.01% gelatin.

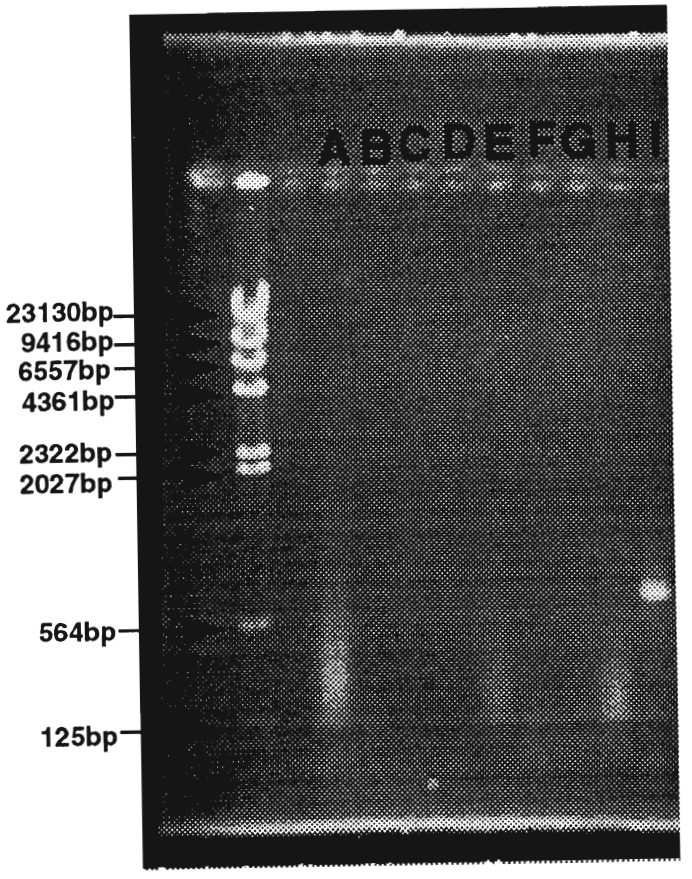

Figure 7: Figure depicting the effect of glycerol upon the yield and size of PCR-RACE amplification products. Lanes A and B both contain the amplification products of control reactions using*Ustilago* cDNA between the oligo (dT) type NOT1 primer and the AC2 primer, with amplification occuring in the presence of 10% glycerol in the lane B reaction mixture. Lane C represents a negative control which lacks cDNA, but contains both AC2 and NOT1 primers, while lane D is also a negative control; containing AC2, NOT1 and 10% glycerol. Lanes E andH both contain the amplification products of control reactions using *Ustilago* cDNA between the oligo(dT) type NOT1 Primer and the AC1 primer, with amplification occuring in the presence of 10% glycerol in the lane H reaction mixture. Lane F is a negative control which lacks cDNA, but contains both AC1 and NOT1 primers, while lane G contains AC1 , NOT1 and 10% glycerol. Lane I is a positive control of known amplification product size, used to ensure that the PCR reactions are working properly.

between the NOT1 primer and the AC1 primer, little difference was observed between products amplified in the presence of 10% glycerol. The use of "cDNA enrichment" procedures, in which three cycles of low stringency amplification precede high stringency cycles, was tested using control cDNA in PCR-RACE procedures. No observeable products were formed in the absence of cDNA enrichment, while observeable products were formed with the use of cDNA enrichment procedures (Figure 8).

With optimization of general PCR-RACE conditions using control cDNA, the PCR-RACE procedure was then optimized for the specific mixed oligonucleotide primer and generalized NOT1 adaptor-primer (Promega) of *Procambarus* FaRP sequence amplification. As optimal PCR-RACE annealing temperatures are reccommended to be 2-4°C below the lowest T_m of the primers, reactions were performed at annealing temperatures of 42°C and 44°C. At 42°C, no observeable product was formed, while at 44°C, a product band approximately 100bp in length was observed (Figure 9).

FaRP mRNA sequences of *Procambarus clarkii* are assumed to possess multiple copies of FaRPs; as deduced from known invertebrate FaRP mRNA sequences. The frequency of occurance of FaRP mRNA transcripts within *Procambarus* tissues is unknown. PCR procedures are known to preferentially amplify shorter sequences when more than one possible amplification sequence exists, and primers may have difficulty locating and binding to rarely occurring mRNAs. In order to avoid these problems, second strand cDNA synthesis of sequences recognized by the mixed oligonucleotide primer was performed. These selected sequences were ligated to EcoRI linkers and purified by passage through A5M-columns.

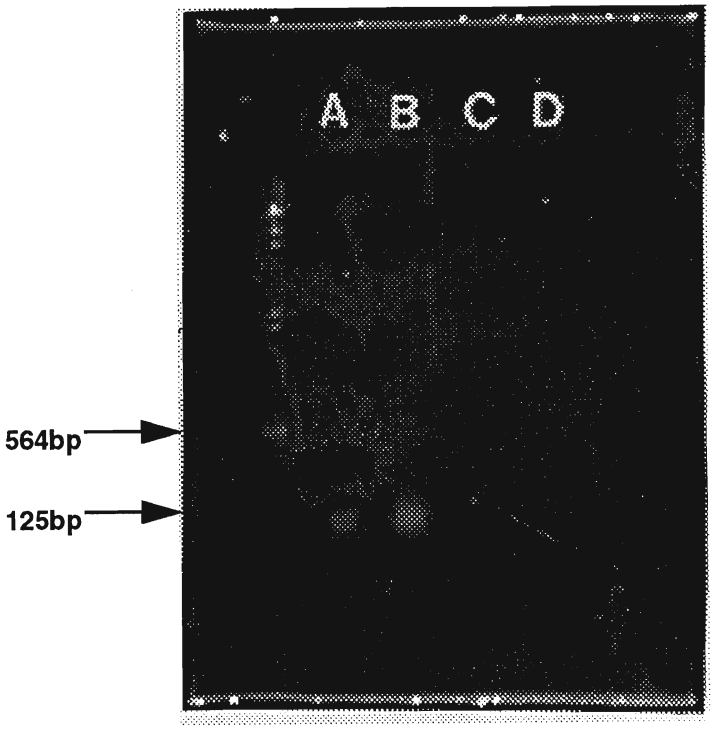

Figure 8: Effects of PCR-RACE cDNA enrichment upon product yield are shown in this figure. Lanes A and B show PCR-RACE products using three preliminary cycles of cDNA enrichment in which lowered annealing temperatures are used. In these lanes, products approximately 100bp in size are visible, in contrast to lanes C and D, in which no product is visible. Lanes C and D contain the products of identical PCR-RACE procedures, lacking the three initial cDNA enrichment cycles.

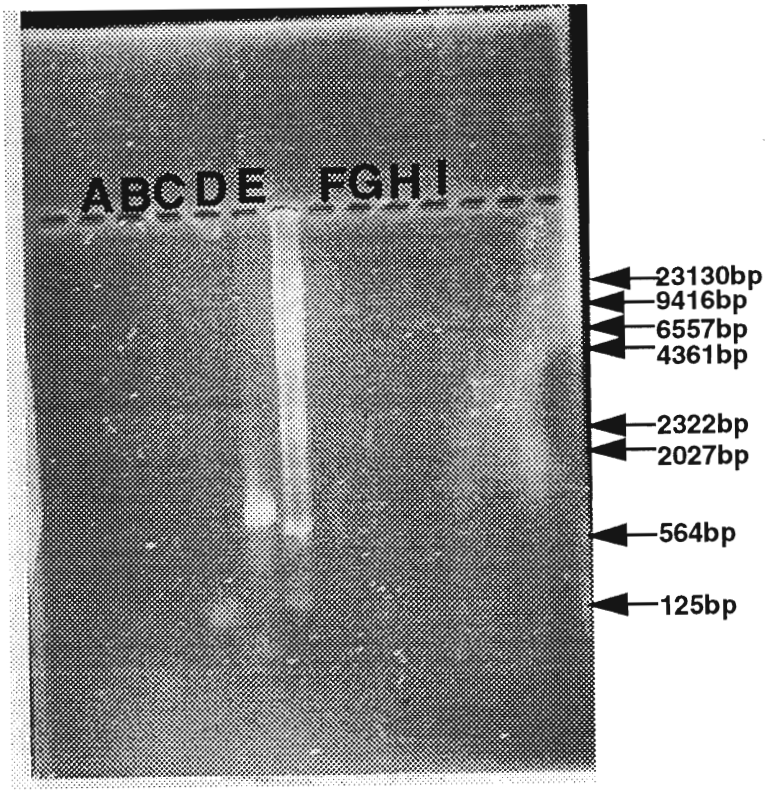

Figure 9: Figure depicting the optimization of PCR-RACE annealing temperatures. Lane D represents the approximately 100bp amplification product of *Procambarus* suboesophagel and thoracic ganglia cDNA using the mixed oligonucleotide NFLRF(Gly) primer and the NOT1 primer at a 44°C annealing temperature. The lack of observeable product in lane F is a result of identical reaction mixture being esposed to a 42°C annealing temperature. Lanes A, B, and C were negative controls performed at 44°C, while lanes G, H, and I were negative controls performed at 42°C. Lane E was a positive control performed at 44°C.

Packagene cDNA library formation was attempted using ligation of EcoRI digested λgt11 arms to 10 μL of synthesized 2U, 1U, and 0.5U Mung Bean Nuclease digested dscDNA. of plaque formation was performed. This suggests the presence of low concentrations of dscDNA. Blue-white selection was performed upon plaques grown on LB Amp X-Gal plates, and white plaques were selected for DNA isolation and investigation of insert presence and size. These white plaques were replated to ensure the consistency of formation of white plaques. The eight isolated white plaques were unable to undergo plaque hybridization procedures as the degenerate oligonucleotide primer used for synthesis of these sequences was observed to hybridize to false positives in the prior screening of the *Procambarus* genomic library (Peaire, 1993). Similarily, attempts to determine insert size by restriction endonuclease digestion and subsequent agarose gel electrophoresis failed to yield observeable amounts of insert bands due to the small expected insert size as compared to the large size of the λgt11 vector. Thus, presumed inserted sequences within the λgt11 vectors were preferentially amplified using PCR procedures (with 40 cycles at a 48°C annealing temperature). The specific primers ALC2 and ALC4 were used for this amplification reaction.

These specific primers recognize sequences within the lac z region of λgt11 surrounding the EcoRI ligation site, thus amplifying that sequence presumably containing the FaRP dscDNA of interest. As the two primers bind to positions on the lac z region that are 70bp apart, the size of the amplification product should be 102bp (70bp + the size of both primers) plus that size of the ligated sequence. As the observed amplification production sizes of the λgt11 recombinants was approximately 100bp in length (Figure 10), the presence of a FaRP sequence insert seems doubtful.

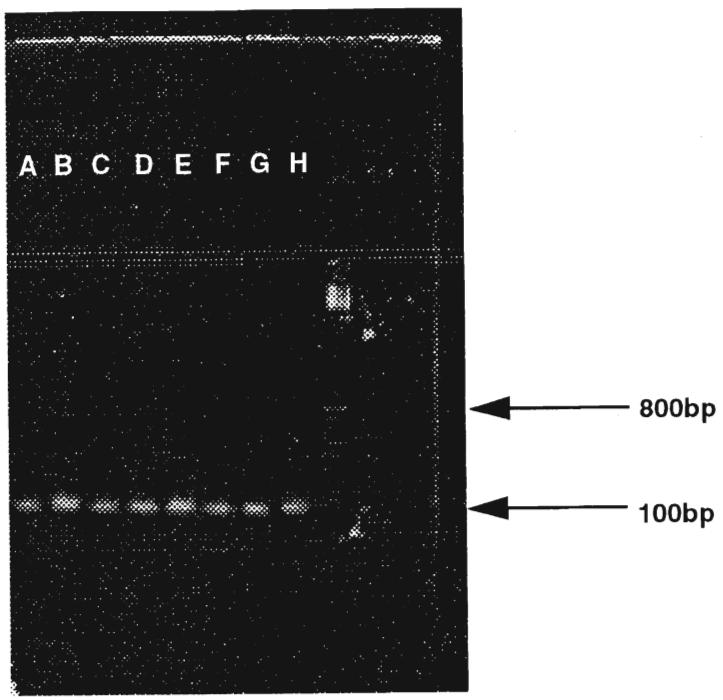

Figure 10: This figure shows the results of PCR amplification of λgt11 vectors postulated to contain inserted sequences. Amplification of the lacz region of λgt11 vectores without inserted sequences using the vectors ALC-2 and ALC-4 (as shown in lane A) produces an approximately 100bp amplification product band. Identical amplification reactions of vectors postulated to contain inserted sequences (as shown in lanes B-H) also produces bands approximately 100bp in length, indicating an absence of insertion sequences.

Purified mixed oligonucleotide primed dscDNA was also used as the substrate for PCR-RACE amplification, avoiding the difficulties of primer location of specific sequences encountered when using total cDNA. Upon amplification, although positive controls produced ample amounts of full-sized product, only small amounts of approximately 100bp product were observed in dscDNA samples (Figure 11). Re-amplification of a 3μL aliquot of the PCR-RACE reaction mixture or of isolated reaction product bands performed using the same procedure as for initial amplification. A sample of 0.5U Mung Bean Nuclease digested dscDNA was observed to produce distinct bands approximately 900bp, 750bp, and 450bp in length (Figure 12). These products were shown to be reproducible; 3μL aliquots of the reaction mixtures could be re-amplified to produce amplified bands of the same length, while 3μL aliquots of the reaction mixtures diluted as for PCR-RACE reactions did not produce observeable bands when run on the same gel (Figure 13). DNA isolated from the 750bp band, when used as the substrate in PCR-RACE procedures, produced an amplification product band approximately 750bp in length.

Amplification bands from 0.5U, 1U, and 2U Mung Bean Nuclease digested dscDNA were isolated from low-melt agarose gels, and ligated to pUC 19 plasmids for use in subsequent transformation of competent cells. Conditions were varied to maximize ligation of plasmid and insert.

The use of pUC 19 plasmid, although purified by CsCl-banding procedures (Maniatis et al., 1989), was marred by the presence of contaminating host *E. coli* DNA sequences. To prevent ligation of contaminating fragments with plasmids, pUC 19 plamid samples underwent dephosphorylation (Maniatis et al., 1987) prior to ligation.

PCR-RACE re-amplification products approximately 1500bp, 750bp, 500bp, 450bp, and 200bp in size were successfully cloned into the Hinc II site of the pUC 19 plasmid. Initial specific

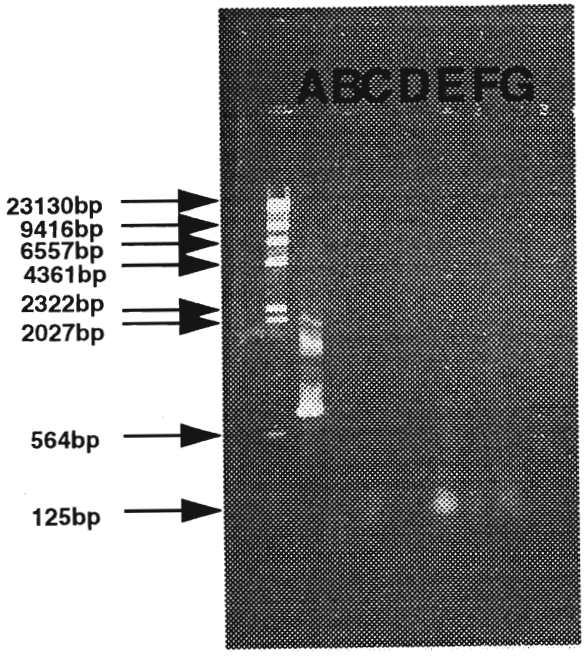

Figure 11: This figure shows the amplification products of Mung Bean Nuclease double stranded cDNAs. The degenerate NFLRF(Gly) oligonucleotide was used to prime second strand synthesis of these selected double stranded cDNAs (dscDNAs). Lane A is a positive control which produced the expected product band size of approximately 900bp, while lanes B and D are negative controls in which sample reactions lack one of the two primers used (either the degenerate NFLRF(Gly) oligonucleotide or the NOT1 primer). Lane C represents 0U Mung Bean Nuclease digested (MBN) dscDNA, while lane E is 0.5U MBN dscDNA, and lane G is 2U MBN dscDNA. All four of these lanes appear to contain varying amounts of an approximately 100bp product.

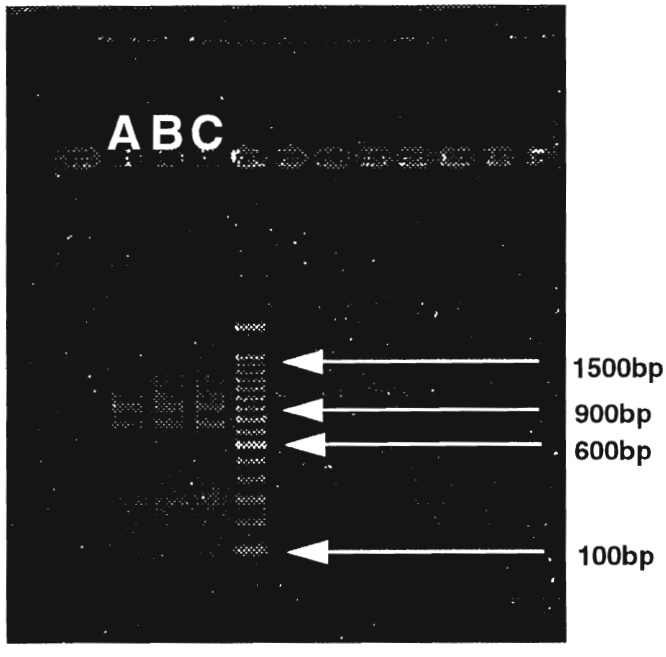

Figure 12: Figure depicting re-amplification products of 0.5U Mung Bean Nuclease digested double stranded FaRP cDNA in lanes A, B, and C. Specific bands approximately 900bp and 750bp in length are observeable.

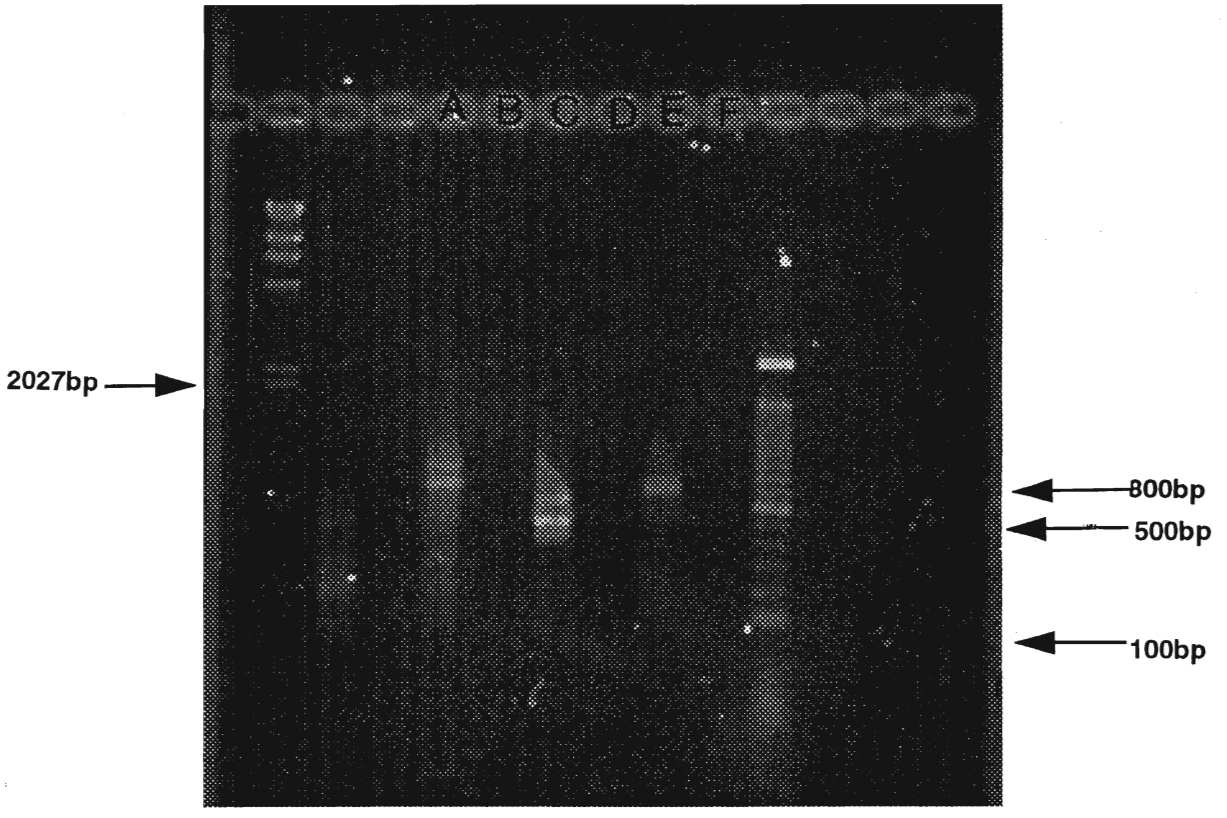

Figure 13: This figure shows reproducibility of re-amplification products. Lane A shows the PCR-RACE re-amplification product of 0.5U Mung Bean Nuclease (MBN) digested dscDNA, with distinct bands approximately 900bp, 750bp, and 200bp in length. Lane B is that amount of 0.5U MBN dscDNA first amplification product diluted as for re-amplification procedures. The lack of visible product in this lane demonstrates that the observed bands in lane A are not merely the original substrate. Identical dilutions have been performed in lanes D and F corresponding to the amplification substrates used in lanes C and E respectively. Lane C represents the PCR-RACE re-amplification product of 1U MBN dscDNA with distinct bands 600bp, 500bp, and 450bp in length. Lane E represents the PCR-RACE amplification product of an isolated band of 750bp DNA from the 0.5U MBN dscDNA. The observation of the 750bp band being formed demonstrates the reproducibility of the PCR-RACE products.

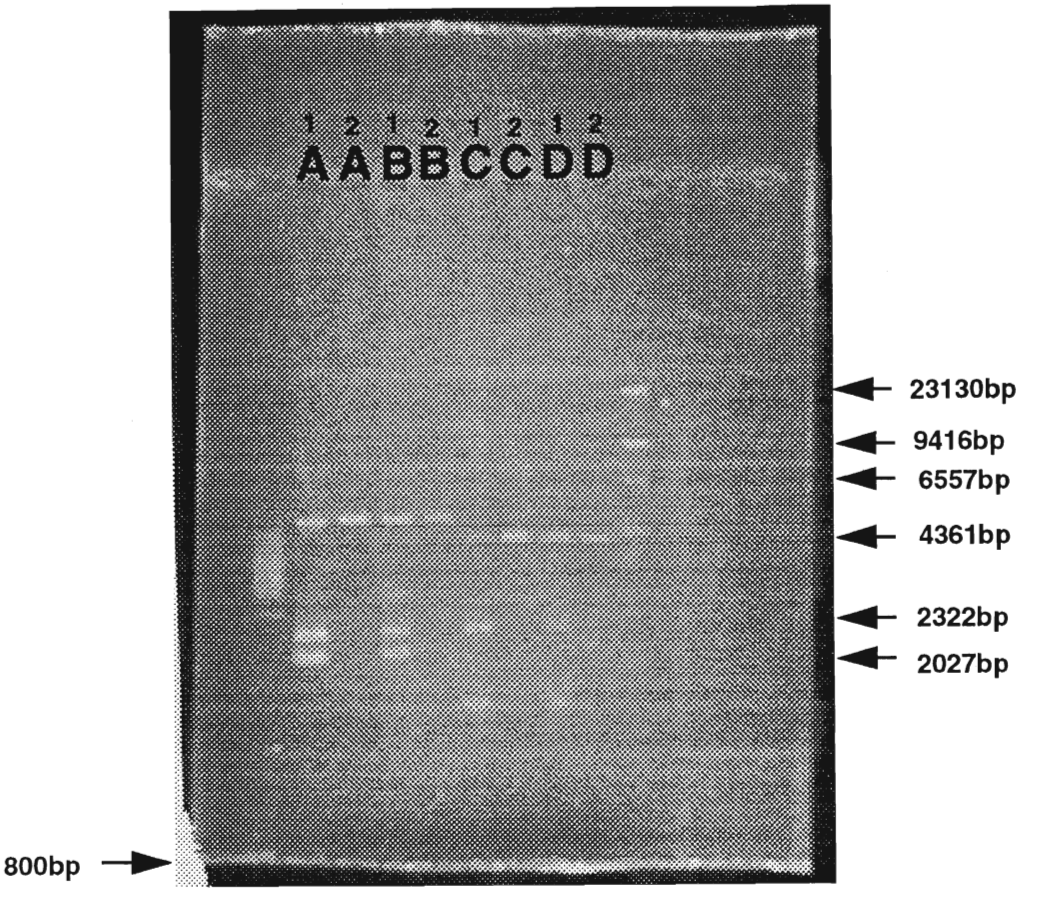

Figure 14: Figure showing insert size in recombinant pUC 19 plasmid. Recombinant plasmids were generated as a result of the ligation of Hinc II digested pUC 19 with the isolated 750bp and 200bp re-amplification product bands of 0.5U Mung Bean Nuclease digested dscDNA. The two isolated 750bp recombinant plasmids are represented by the letters A and B, while the two isolated 200bp recombinant plasmids are reresented by the letters C and D. Pvu II digestion is indicated by the number 1, while the number 2 represents EcoRI digestion. Incomplete digestion by both enzymes resulted in observeable bands being formed due to the presence of both supercoiled and linearized forms of the plasmid. Supercoiled plasmid bands are observeable at approximately 2000bp for samples A and B, and at approximately 1500bp for samples C and D. Circular plasmid bands are observeable at approximately 4900bp for samples A and B, and at approximately 4300bp for samples C and D. EcoRI digestion results in linearization of the recombinant plasmids. In lanes A2 and B2, EcoRI plasmid linearization results in an approximately 3400bp band, indicating an insert size of approximately 700bp (similar to the expected 750bp insert size). In lanes C2 and D2, EcoRI digestion results in an approximately 2900bp band being produced, indicating an insert size of approximately 200bp (corresponding to the expected 200bp insert size). Pvu II digestion, which cuts the plasmid on either side of its multiple cloning site 322bp apart, results in the formation of an approximately 2400bp sequence (pUC 19 (2686bp) - 322bp), in lanes A1, B1, C1, and D1. Incomplete Pvu II digestion, which results in the linearization of the pUC 19 plasmid, produces bands identical in size to those produced by EcoRI digestion.

endonuclease digestions of plasmid DNA isolated from colonies lacking β-galactosidase activity (and thus indicating recombination of the pUC 19 plasmid) were performed to determine insert presence and size, as shown in Figure 14. Figure 14 demonstrates the presence of an insert of approximately 700bp and 200bp in the transformation products of 750bp and 200bp PCR-RACE products respectively. In order to ensure that actual PCR-RACE products were inserted into the pUC 19 plasmid, hybridization to colony lifts were performed. Radioactively-labelled PCR-RACE product probes hybridized to recombinant colonies, indicating successful PCR-RACE product plasmid incorporation (see Figure 15 for colony hybridizations with labeled 1500bp product, Figure 16 for colony hybridizations with labeled 750bp product, Figure 17 for 500bp hybridizations, Figure 18 for 450bp hybridizations, and Figure 19 for 200bp hybridizations). Positive controls, in which radioactively - labeled PCR-RACE products of a given size were hybridized to the same unlabeled PCR-RACE product, were performed to ensure that hybridization reactions were being performed under favourable conditions.

Sequence determination of the approximately 1500bp, 450bp, and 200bp products indicated that these products possessed identical sequences at their 3' and 5' ends. The sequence amplified by the NFLRF(G)-encoding degenerate oligonucleotide primer was **AGCTTATAGATACCAGG**, with 10 of the 17 bases matching those found in the primer sequences (bases identical to those found in the degenerate oligonucleotide sequence are printed in bold type). This amplified sequence does not encode FaRP sequence.

The NOT1 adaptor-primer was observed to be similarily non-specific in its binding, with 12 of the 28 bases binding to those found in the primer sequence. The amplified sequence recognized by NOT1 was AG**AATCGCAACGTCATGTCATGTTAAGA**.

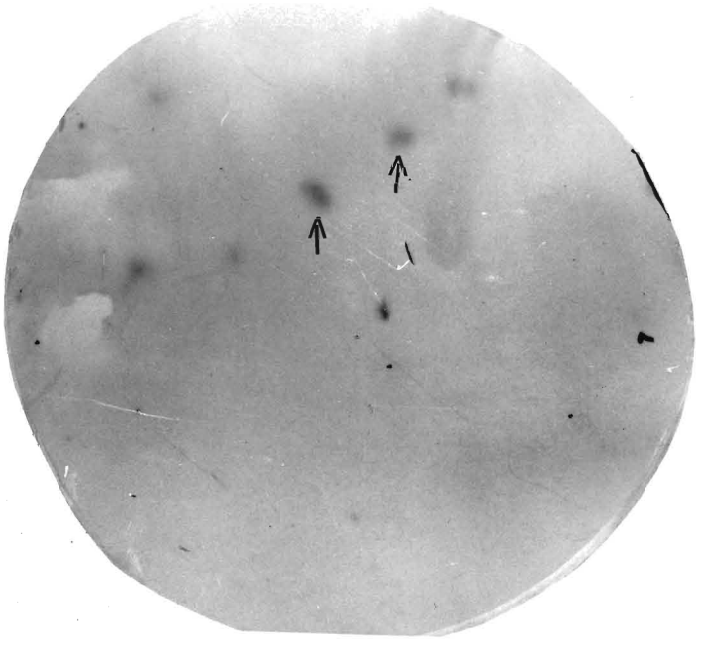

Figure 15: Figure showing hybridization of ^{32}P - labeled 1500bp PCR-RACE product to colony lifts of transformed DH5α cells containing recombinant pUC 19 plasmid. The dark spots indicated by arrows are areas of probe hybridization to DH5α colonies containing the PCR-RACE 1500bp sequences. Spots not indicated by arrows are points of false hybridization which do not correspond to locations of DH5α colonies.

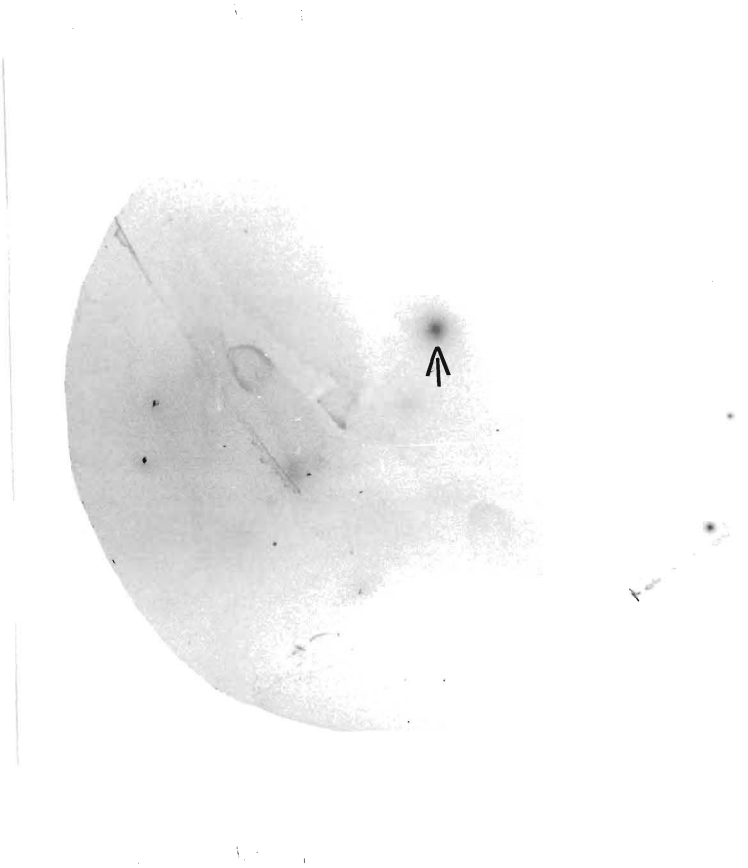

Figure 16: Figure showing hybridization of ^{32}P - labeled 750bp PCR-RACE product to colony lifts of transformed DH5α cells containing recombinant pUC 19 plasmid. The dark spots indicated by arrows are areas of probe hybridization to DH5α colonies containing the PCR-RACE 750bp product. Spots not indicated by arrows are points of false hybridization which do not correspond to locations of DH5α colonies.

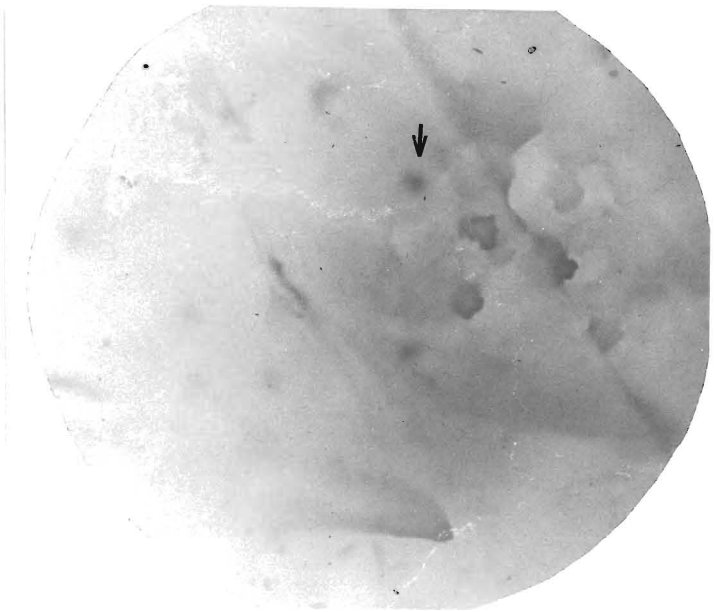

Figure 17: Figure showing hybridization of ^{32}P - labeled 500bp PCR-RACE product to colony lifts of transformed DH5α cells containing recombinant pUC 19 plasmid. The dark spots indicated by arrows are areas of probe hybridization to DH5α colonies containing the PCR-RACE 500bp product. Spots not indicated by arrows are points of false hybridization which do not correspond to locations of DH5α colonies.

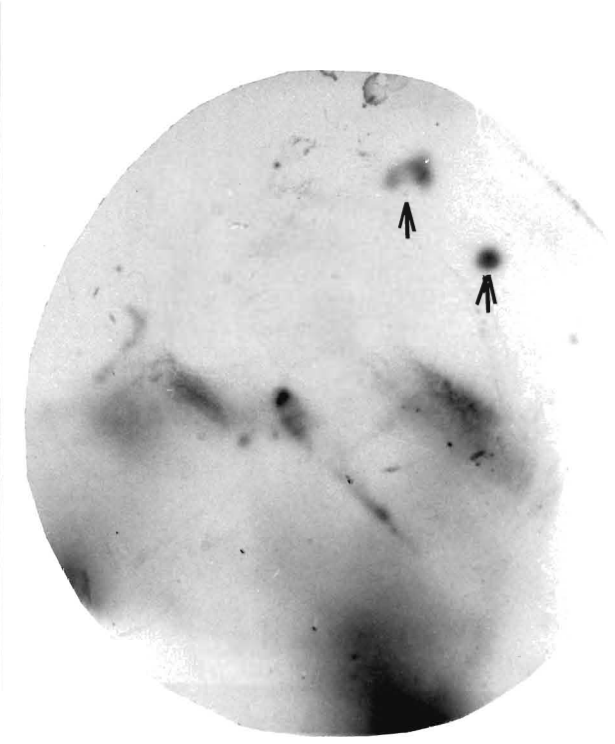

Figure 18: Figure showing hybridization of ^{32}P - labeled 450bp PCR-RACE product to colony lifts of transformed DH5α cells containing recombinant pUC 19 plasmid. The dark spots indicated by arrows are areas of probe hybridization to DH5α colonies containing the PCR-RACE 450bp product. Spots not indicated by arrows are points of false hybridization which do not correspond to locations of DH5α colonies.

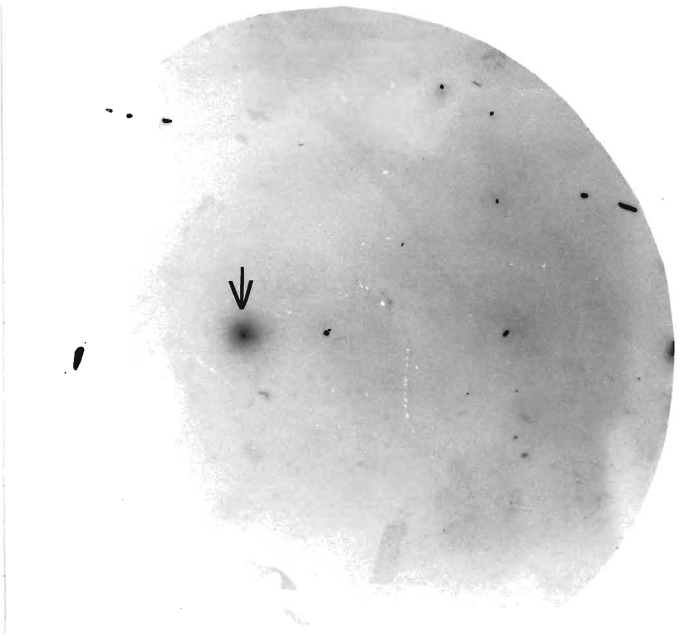

Figure 19: Figure showing hybridization of ^{32}P - labeled 200bp PCR-RACE product to colony lifts of transformed DH5α cells containing recombinant pUC 19 plasmid. The dark spots indicated by arrows are areas of probe hybridization to DH5α colonies containing the PCR-RACE 200bp product. Spots not indicated by arrows are points of false hybridization which do not correspond to locations of DH5α colonies.

A

```
DNA  sequence    153 b.p.   aatcataggacc ... tattgccttaag   linear

          |   10     |   20     |   30     |   40     |   50     |   60
    1 aatcatagga ccttggacta ctatcctact  acgagtagct actctgcttc cttcgtctct 60
   61 cttctcttgc ttctctctct acttctctct  tcttcgttat agtcttttat gcaataatgg 120
  121 gtgggtggtt gaccaaattg ttattgcctt  aag                             153
          |   10     |   20     |   30     |   40     |   50     |   60
```

B

```
DNA  sequence    166 b.p.   AGAATCGCAACG ... AGGACTAGGAAT   linear

          |   10     |   20     |   30     |   40     |   50     |   60
    1 AGAATCGCAA CGTCATGTCA TGTTAAGAGC  GAACTTACAT GTTCATGGTT AAGAAATAAC 60
   61 GAAGAAGTTA GTTTTCATGT GTCCGTTACA  TACCTAGTAA ATAGGTCAAG ATGTAGAGTC 120
  121 ATAGCGAAAT AGAATGGATG CGTCATTTGT  TATTAGGACT AGGAAT               166
          |   10     |   20     |   30     |   40     |   50     |   60
```

Figure 20: This figure depicts the 153bp 5' end mRNA sequence for the 200bp, 450bp, and 1500bp sequences (A), and the166bp 3' end mRNA sequence for the 200bp, 450bp, and 1500bp sequences (B).

Partial DNA sequences were obtained which correspond to the 5' and 3' ends of the 200bp, 450bp, and 1500bp sequences (Figure 20). DNA sequence comparisons to known sequences was performed using the Fasta Server program (Pearson and Lipman, 1988). No significant DNA sequence matches were obtained, and no homology to known FaRP-encoding sequences was detected. Instead, the 153bp partial sequence corresponding to the degenerate oligonucleotide-primed end of the sequence was found to possess 69.0% identity in a 58 bp overlap to the *Caenorhabditis elegans* cosmid K09A11 which encodes for the cytochrome P450 gene, the Transfer-RNA gene, the transposase gene and the tRNA-Tyr gene (Figure 21).

Sequence determination of the 500bp and 750bp amplification products indicated that these products possessed identical sequences at their 3' and 5' ends.

The sequence amplified by the NFLRF(G)-encoding degenerate oligonucleotide primer was **AAT**CAT**AGGACCTT**GAC, with 9 of the 17 bases matching those found in the primer sequences (bases identical to those found in the degenerate oligonucleotide sequence are marked in bold type). This amplified sequence does not encode for a FaRP sequence.

The NOT1 Adaptor-primer was also observed to bind non-specifically. The amplified sequence primed by the NOT1 primer was **ATATCCGCAGGTTTAACCACAAAATATA,** with 16 of the 28 bases matching those found in the primer sequence.

Partial DNA sequences were obtained which correspond to the 5' and 3' ends of the 500bp and 750bp sequences (Figure 22). Comparison of the amplified sequences to known gene sequences, as performed using the Fasta Server program (Pearson and Lipman, 1988) did not detect any homology to known FaRP-encoding sequences. No significant DNA sequence matches were obtained. That sequence with the highest homology to the 5' end of the mRNA

sequence was determined to be the complete genome of the *Vaccinia* virus. The two sequences possess 66.0% identity in a 50bp overlap (Figure 23).

The generated partial sequences were translated in 3 phases using the Strider program to obtain the corresponding protein sequences. The longest available sequence in each phase, as determined by stop codon position, was examined for similarity to known protein sequences using the Blitz program. No significant protein matches were obtained.

```
SCORES      Init1: 56 Initn: 175 Opt: 75
             69.0% identity in 58 bp overlap

                 20        30        40        50        60        70
2041d1 CTGACCTCTACTATCCTACTACGAGTAGCTACTCTGCTTCCTTCGTCTCTCCTTCTCTTG
                                          ||||| |  || |||| | |  ||| ||||
Cek09a CATGTAAGCGTGGACAAAATTTTCAGGTTCACTCTCC--CCATCGTAT-TTTTTCACTTG
     4470      4480      4490      4500        4510      4520

                 80        90       100       110       120       130
2041d1 CTTCTCTCTCTACTTCTTCTTCGTTATAGTCTTTTATGCAATAATAATGGGTGGGTGGTT
       |     | ||||||||||||||| || |
Cek09a C----GTAGCTACTTCTTCTTCGATAAATGTATGGTACTGAAAATCTCACTCAATTTCTA
            4530      4540      4550      4560      4570      4580
```

Figure 21: This figure shows the results of the comparison of received amplification product sequences to known sequences by means of the EMBL Fasta Server in Heidelberg, Germany. The amplified sequence submitted for comparison was a 153bp partial sequence corresponding to that region primed by the degenerate NFLRF(G) oligonucleotide primer. This 153bp partial sequence was identical for the 200bp, 450bp, and 1500bp amplification products. The known sequence with the greatest homology to this 153bp sequence is the *Caenorhabditis elegans* cosmid K09A11, which contains the genes for cytochrome P450, transfer-RNA, transposase, and tRNA-Tyr. As shown in this figure, the cosmid has 69.0% identity in a 58bp overlap with the sequence of interest.

A

```
DNA  sequence    158 b.p.   aatcataggacc ... tagaggtgctac   linear

          |   10     |   20     |   30     |   40     |   50     |   60
    1 aatcatagga ccttggacac tgcttaagca  attaacattg tgcacacaga catatatatg 60
   61 cataaagact ataagaagaa gaagtagaga  gagaaagcga agagagaaga ggaagcagaa 120
  121 gcgagatagc tactctgagt gatatgtaga  ggtgctac                        158
          |   10     |   20     |   30     |   40     |   50     |   60
```

B

```
DNA  sequence    165 b.p.   ATATCCGCAGGT ... ATAGTAGAGGAG   linear

          |   10     |   20     |   30     |   40     |   50     |   60
    1 ATATCCGCAG GTTTAACCAC AAAATATACT  TAGGCAATAA CCAATTTGGC AAAACCACCA 60
   61 TATTATGCAT AAAGCTATAC TGAAGAAGAA  GTAGAGAGAG AAGCGAAGAG AAGGAGAAGA 120
  121 CGAAGAGAAA GCAGAGATAG CTACTGCTAG  TGGATAGTAG AGGAG                165
          |   10     |   20     |   30     |   40     |   50     |   60
```

Figure 22: This figure depicts the 158bp 5' end mRNA sequence for the 500bp and 750bp sequences (A), and the165bp 3' end mRNA sequence for the 500bp and 450bp sequences (B).

```
SCORES      Init1: 58 Initn: 152 Opt: 81
             66.0% identity in 50 bp overlap

         80        90       100       110       120       130
24a004 TATAAGAAGAAGAAGTAGAGAGAGAAAGCGAAGAGAGAAGAGGAAGCAGAAGCGAGATAG
                                     | || |||| || |  |  |||||||  |
Pxvacc AAAATTATAATGAATATTATTATGATGATTATGATAGAACTGGTATGAACAGCGAGAGCG
          17953

        140       150       160
24a004 CTACTCTGAGTGATAGTAGAGGTGCTAC
        || |  |||||||| || |
Pxvacc ATAGTGAGAGTGATAATATATCAATCAAAACAGAATATGAGAATGAATATGAATTCTATG
          17959
```

Figure 23: This figure shows the results of the comparison of received amplification product sequences to known sequences by means of the EMBL Fasta Server in Heidelberg, Germany. The amplified sequence submitted for comparison was 158bp partial sequence corresponding to that region primed by the degenerate NFLRF(G) oligonucleotide primer. This 158bp partial sequence was identical for the 500bp and 750bp amplification products. The known sequence with the greatest homology to this 158bp sequence is the complete genome of the *Vaccinia* virus. As shown in this figure, the virus has 66.0% identity in a 50bp overlap with the sequence of interest.

DISCUSSION

In this study, PCR-RACE reactions were optimized for the amplification of *Procambarus clarkii* cDNA sequences with a degenerate oligonucleotide primer designed to recognize the two sequenced *Procambarus* FaRPs. Reactions conditions including buffer composition, primers used, cycle number, nature of the amplification substrate, annealing, extension, and denaturation temperatures and times, and the use of re-amplification procedures, were varied and optimized.

The lack of observeable product in PCR reactions in which buffer did not contain 0.1% gelatin may be accounted for, as the use of gelatin in PCR reactions is reported to improve enzyme binding (Innis et al., 1988). Without efficient enzyme binding, new strand synthesis is impaired, resulting in reduced amounts of amplification product.

Although the addition of 10% glycerol to PCR buffer has been reported to maximize efficient specific amplification (Li and Negre, 1993), its presence in test reactions was observed to have little positive effect or a largely negative effect upon the quantity of PCR-RACE product received. It is unknown why these observations were received, as the exact role of glycerol in PCR reactions is unclear (Li and Negre, 1993).

Although PCR annealing temperatures are generally reccommended to be 2-4^{0}C below the lowest T_m of the primers, annealing performed at 42^{0}C (4^{0}C below the lowest T_m of the primers) was found not to produce any observeable product. This lack of observeable product may be accounted for by low primer annealing stringency at lower temperatures. With lowered stringency, many different sequences with little homology to the primers may be recognized and amplified to a low degree, obstructing consistant amplification of specific sequences with a high degree of homology to the primers.

The observation of products of much larger length in PCR-RACE re-amplification reactions as compared to products in initial amplification reactions may be accounted for. PCR-RACE reactions preferentially amplify shorter sequences (Frohman et al., 1988). With re-amplification, longer, underrepresented sequences may appear due to further amplification cycles. Re-amplification was performed upon samples instead of simply increasing the number of amplification cycles in initial PCR-RACE reactions, as product formation may be impaired due to depletion of reaction constituents.

In the determination of insert presence in pUC 19 plasmid (Figure 14), the Pvu II and EcoRI restriction endonucleases showed incomplete plasmid digestion. This may be accounted for by an excess of salts in the reaction mixture, which impairs enzyme activity.

Generated PCR-RACE amplification products were studied to determine if they encoded FaRP sequences. Although the 200bp, 450bp, and 1500bp sequences were found to have homology to the *Caenorhabditis elegans* cosmid KO9A11, and the 500bp and 750bp sequences to have homology with the complete genome of the *Vaccinia* virus, in both cases these similarities were deemed to be fortuitous due to the low identity percentages (69% and 66% respectively) over short sequence overlaps (58bp and 50bp respectively). Although amplification products of five different lengths were studied, only two different sequences were recognized and bound by the degenerate oligonucleotide primer, while the NOT1 adaptor-primer was observed to recognze and bind to two sequences which were not poly (A) sequences. The 200bp, 450bp, and 1500bp sequences possess identical 3' and 5' sequences, and the 500bp and 750bp sequences also possess 3' and 5' sequences identical to each other. The absence of any known gene sequences in *Procambarus* prevents identification of these isolated sequences. They are however, postulated to be repeated sets of sequences found along *Procambarus* mRNA. The presence of multiple

copies of highly similar sequences would encourage binding of the two primers, while the tendancy of PCR to amplify shorter fragments (Frohman et al., 1988) may have contributed to the binding of the NOT1 adaptor-primer to sequences internal to the poly (A) tail. The low stringency of the NOT1 adaptor-primer binding may also be attributed to the 74°C melting temperature (T_m) of this molecule being markedly higher than the 44°C annealing temperature used in the PCR-RACE reactions.

The low stringency of the degenerate oligonucleotide primer may be due to several factors. The use of the 16 different sequences togethor as a "mixed" primer neccesitated the use of an annealing temperature below the lowest T_m of the 16 different sequences. Thus, sequences with T_ms up to 8°C above the used annealing temperature may bind with low stringency to non-FaRP sequences. If these sequences are shorter than, or in greater copy number than FaRP-encoding sequences present, they may be preferentially amplified. Additional stringency problems may be encountered with the use of cDNA enrichment procedures. Sequences recognized and amplified at lowered annealing temperatures may continue to be amplified at raised temperatures if they are similar to sequences of primers annealing well below their T_ms.

Although the determination of primer sequences based upon *Drosophila* FaRP codon usage was successfully used by Baro et al.(1993) in lobster, and this procedure was necessary due to the absence of any sequenced *Malacostracean* genes, the determination of primer sequences based upon *Drosophila* FaRP codon usage may also be a factor in the amplification of unwanted sequences. If *Procambarus* FaRP codon usage differs from that of *Drosophila*, then the synthesized primers may not recognize encoded FaRP sequences. Finally, the use of a 17bp oligonucleotide primer, which is of minimal length for a primer, reduces the specificity of binding. Although this primer was specifically designed so as to recognize both NF_1 and DF_2, as well as

any other possible NFLRF(Gly)-type sequences present, an increased primer length and sequence designed to recognize either NF_1 or DF_2 would be expected to yield amplification products of greater specificity.

The failure to receive inserts from the selected dscDNA library despite the consistently white colour of colonies during blue/white selection (indicative of vector recombination), may be explained by short adaptor insertion into the λgt11 vector, or by base pair insertion or deletion during ligation, at which time the vector is most succeptible to these reactions.

A failure to insert sequences into the λgt11 vector may be due simply to extremely low amounts of dscDNA produced by degenerate oligonucleotide priming.

Possible future investigations to determine FaRP gene sequences using oligonucleotide primers may benefit from the use of the SchistoFLRFamide (PDVDHVFLRFamide) primer/probe received courtesy of Dr. Bill Bendena (Queens University, Kingston, ON), which should recognize the encoded *Procambarus* FaRP pEXXDHVFLRFamide (where "X" indicates an unidentified amino acid). In further investigations it is recommended that a longer oligonucleotide primer of lower degeneracy be used, if possible, so as to specifically bind to FaRP-encoding sequences.

SUMMARY

1. The PCR-RACE procedure was optimized for the degenerate oligonucleotide primer.
2. Re-amplification was important in PCR-RACE procedures.
3. Five PCR-RACE amplification products were sequenced.
4. The PCR-RACE amplification products were determined not to be FaRP-encoding sequences.
5. The degenerate oligonucleotide primer was observed to bind non-specifically.
6. It is unlikely that a selected cDNA library was successfully created.
7. Alternative strategies may be more useful in isolation of FaRP-encoding gene sequences.

LITERATURE CITED

Akashi, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. Genetics 136: 927-935.

Albers, R.T., S.Y. Liou, E.G. Stopa, and R.T. Zoeller. 1991. Interaction of colocalized neuropeptides: functional significance in circadian timing system. J. Neurosci. 11: 846-851.

Alberts, B., D. Bray. J. Lewis, M. Raff, K. Roberts, J.D. Watson. 1989. Molecular Biology of the Cell 2nd ed. p. 1219. Garland Publishing Inc., New York.

Alexandrowicz, J.S. 1953. Nervous organs in the pericardial cavity of the decapod crustacea. J. Mar. Biol. Assoc. UK 31: 563-580.

Altman, J. 1989. Neuropeptide gives nerve cell its signature. New Scientist 121: 31.

Argenta, P. M. 1990. Analysis of ribonucleic acid associated with fimbriae in *Ustilago violacea*.BSc (Hon.) Thesis. Department of biological sciences, Brock University, St. Catharines, Ont.

Arms, K. and P.S. Camp. 1979. Biology 2nd ed. p.931. Saunders College Publishing, Philadelphia.

Barker, J.L. 1977. Physiological roles of peptides in the nervous system, p. 295 - 305. *In* H. Gainer, Peptides in Neurobiology. Plenum Press, New York.

Barnes, R.D. 1987. Invertebrate Zoology. 5th ed. p.893. Saunders College Publishing, Philadelphia.

Beltz, B.S. 1988. Endocrinology of Selected Invertebrate Types. p.258. Alan R. Liss, Inc. New York.

Beltz, B.S. and Kravitz, E.A., 1986. Aminergic and Peptidergic neuromodulation in crustacea. J. Exp. Biol. 124: 115-141.

Beltz, B.S. 1988. Crustacean neurohormones, p. 235 - 258. *In* H. Laufer, R.G.H. Down, Endrocinology of Selected Invertebrate Types. Alan R. Liss, Inc., New York.

Berger, S.L., 1987. Quantifying ^{32}P-labelled and unlabelled nucleic acids, p. 49-54. *In* S.L. Berger, and A.R. Kimmel, Guide to Molecular Cloning Techniques. Academic Press, Inc., Orlando.

Bishop, C. A., J. J. Wine, F. Nagy, and M. O'Shea. 1987. Physiological consequences of a peptide cotransmitter in a crayfish nerve-muscle preparation. J. Neurosci. 7: 1769-1779.

Bloom, F.E. 1981. Neuropeptides. Scient. Amer. 241: 148-168.

Bode, H.R. 1992. Continuous conversion of neuron phenotype in *Hydra*. TIG 8: 279-283.

Boer, H. H., and J. Van Minnen. 1988. Immunocytochemistry of hormonal peptides in molluscs: optical and electron microscopy and the use of monoclonal antibodies. p.41 *In* M.C. Thorndyke, and G.J. Goldworthy. Neurohormones in Invertebrates, Cambridge University Press, Cambridge.

Boer, H.H., L.P.C. Schot, J.A. Vennstra, and D. Reichelt. 1980. Immunocytochemical identification of neural elements in the central nervous systems of a snail, some insects,a fish, and a mammal with antiserum to the molluscan cardioexcitatory tetrapeptide FMRFamide. Cel. Tissue Res. 213: 21-217.

Boyle, J.S. and A.M. Lew. 1995. An inexpensive alternative to glassmilk for DNA purification. TIG 11: 20.

Brownstein, M.J. 1980. Opioid peptides: search for the precursors. Nature 287: 678-679.

Brusca, R.C., and G.J.Brusca. 1990. Invertebrates. p.568.Sinauer Associates, Inc. Sunderland, Massachusettes.

Burden, S.J. 1993. Synapse-specific gene expression. TIG 9: 12-16.

Chelley, J., J.C. Kaplan, P. Maire, S. Gautron, and A. Kahn. 1988. Transcription of the dystrophin gene in human muscle and non-muscle tissues. Nature 333: 858-860.

Chin, A.C., E.R. Reynolds, and R.H. Scheller. 1990.Organization and expression of the *Drosophila* FMRFamide-related prohormone gene. DNA and Cell Biol. 9: 263-271.

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

Chuang, S-E., K-C. Wang, and A-L. Cheng. 1995. Single-step direct cloning of PCR products. TIG 11:7.

Cobianchi, F., and S.H. Wilson. 1987. Enzymes for modifying and labelling DNA and RNA, p. 94-110. *In* S.L. Berger, and A.R. Kimmel, Guide to Molecular Cloning Techniques. Academic Press, Inc., Orlando.

Comb, M., P.H. Seeberg, J. Adelman, L. Eiden and H. Herbert. 1982. Primary structure of the human Met- and Leu-enkephalin precursor and its mRNA. Nature 295: 663-666.

Cooke, I.M., and R.E. Sullivan. 1982. Hormones and Neurosecretion. p. 205 - 290. *In* H.L. Atwood, and D.C. Sandeman, The Biology of Crustacea - Neurobiology, Structure and Function. v.3. Academic Press, Inc., Orlando.

Cotrell, G.A., D.A. Price, K.E. Doble, S. Hettle, J.Sommerville, and M. McDonald. 1992. Identified *Helix* neurons: mutually exclusive expression of the tetrapeptide and heptapeptide members of the FMRFamide family. Biol. Bull. 183: 113-122.

Cottrell, G.A., and N.W. Davies. 1987. Multiple receptors for a molluscan peptide (FMRFamide) and related peptides in *Helix*. J. Physiol. 382: 51-68.

Cox, R.A. 1968. *In* L. Grossman, and Moldave, K., eds., Methods in Enzymology. Vol.12, pp. 120-129, Academic Press, Orlando, FL.

De Loof, A., and L. Schooes. 1990. Homologies between the amino acid sequences of some vertebrate peptide hormones and peptides isolated from invertebrate sources. Comp. Biochem. Physiol. 95 B: 459-468.

Delort et al., 1989. An efficient strategy for cloning 5' extremities of rare transcripts permits isolation of multiple 5' untranslated regions of rat tryptophan hydroxylase mRNA. Nucleic Acids Res. 17: 6439-6448.

Denuce, J.M. 1982. Do crustaceans react to mammalian gastrointestinal hormone? Preliminary experiment with cholecystokinin-pancreoxymin (CCK-PZ) and secretin. Arch. Int. Physiol. Biochem. 90:820.

Dickenson, P. 1989. Modulation of simple motor patterns. Seminars Neurosci. 1: 15-24.

Dockray, G.J., K.R. Reeve, J. Shivey, R.J. Gayton, and C.S. Barnard. 1983. A novel active pentapeptide from chicken brain identified by antibodies to FMRFamide. Nature 305: 325-330.

Dockray, G.J. 1989. Gastrin, cholecystokinin (CCK), and the leukosulfakinins. Biol. Bull. 177: 195-197.

Duve, H., and A. Thorpe. 1988. Isolation and characterization of vertebrate-type peptides in insects. Pgs. 102-138. *In* Thorndyke, M. C., and G.J.Goldsworthy, eds., Neurohormones in Invertebrates. Cambridge University Press, Cambridge.

Duve, H., A.J. Johnsen, J.C. Sewell, A.G. Scott, I. Orchard, J.F. Rehfeld, and A. Thorpe. 1992. Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH_2 neuropeptides (designated calli FMRFamides) from the blowfly *Calliphora vomitora*. Proc. Natl. Acad. Sci. USA. 89: 2326-2330.

Ebberink, R.J.M., D.A.Price, H.vanLoenhout, K.E.Doble, J. Riehm, W.P.M.Geraerts, and M.J.Greenberg. 1987. The brain of *Lymnaea* contains a family of FMRFamide-like peptides. Peptides 8: 515-522.

Evans, P.D., and C.M.Meyers. 1986. The modulatory actions of FMRFamide and related peptides on locust skeletal muscle. J.Exp.Biol. 126: 403-422.

Favrel, P., A. Van-Wormhoudt, J.M. Studler and C. Bellon. 1987. Immunochemical and biochemical characterization of gastrin/cholecystokinin-like peptides in *Palaemon serratus* (crustacea decapoda): intermolt variations. Gen. Comp. Endo. 65: 363-372.

Fedorcsak, I. and L. Ehrenberg. 1966. Effects of diethyl pyrocarbonate and methyl methanesulfonate on nucleic acids and nucleases. Acta. Chem. Scand. 20: 107.

Frohman, M.A., M.K. Dush, and G.R. Martin, 1988. Rapid production of full-length cDNAs from rare transcripts:amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA. 85: 8998-9002.

Gainer, H. 1977. Biosynthesis of neuronal peptides. p. 181-214. *In* H. Gainer, Peptides in Neurobiology. Plenum Press, New York.

Geraerts, W.P.M., E. Vreugdenhil, and R.H.M. Ebberink. 1988. Bioactive peptides in molluscs p. 261-267. *In* M.C. Thorndyke, and G.J. Goldsworthy, Neurohormones in Invertebrates. Cambridge University Press, Cambridge.

Glanzman, D.L. and F.B. Krasne. 1983. Serotonin and octopamine have opposite modulatory effects on the crayfish's lateral giant escape reaction. J. Neurosci. 1: 15-24.

Golding, D.W. and D.V. Pow. 1988. The new neurobiology -ultrastructural aspects of peptide release as revealed by studies of invertebrate nervous systems. p. 18 *In* M.C. Thorndyke, and G.J. Goldsworthy, Neurohormones in Invertebrates. Cambridge University Press, Cambridge.

Greenberg, M.J., and M.C. Thorndyke. 1989. Consistency and variability in peptide families: introduction. Biol. Bull. 177: 167-171.

Greenberg, M.J., S.D. Painter, K.E. Doble, G.T. Nagle, and D.A. Price. 1983. The Molluscan neurosecretory peptide FMRFamide: comparative pharmacology and relationship to the enkephalins. Fed. Proc. 42: 82-86.

Greenberg, M.A., and D.A. Price. 1983. Invertebrate neuropeptides: native and naturalized. Ann. Rev. Physiol. 45: 271-288.

Greenberg, M.J., S.M. Lambert, H.K. Lehman, and D.A. Price. 1986. The enkephalins and FMRFamide-like peptides: the case for coevolution, p. 93-101. *In* C.B. Stefano, ed.

CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms. Vol I. CRC Press, Inc. Boca Raton.

Grimmelikhuijzen, C.J.P., and D. Graff. 1985. Arg-Phe-amide-like peptides in the primitive nervous system of coelenterates. Peptides 6: 477.

Gubler, U. 1987. Second-Strand cDNA Synthesis: Classical Method. *In* S.L. Berger, A.R. Kimmel eds. Guide to Molecular Cloning Techniques. Academic Press, Inc., Orlando.

Gubler, U. Second-Strand cDNA Synthesis: mRNA Fragments as Primers. *In* S.L. Berger, A.R. Kimmel eds. Guide to Molecular Cloning Techniques. Academic Press, Inc., Orlando.

Habener, J.F. 1981. Principles of peptide-hormone biosynthesis, p. 2134. *In* J.B. Martin, S. Reichlin, and K.L. Bick. Neurosecretion and Brain Peptides, Raven Press, New York.

Hadley, M.C. 1988. Endocrinology 2nd ed. p.608. Prentice Hall, New Jersey.

Harris-Warrick, R.M. and E.A. Kravitz, 1984. Cellular mechanisms for modulation of posture by octopamine and serotonin in the lobster. J. Neurosci. 4: 1976-1993.

Harris-Warrick, R.M., R.E. Flamm, B.R. Johnson, and P.S. Katz. 1989. Modulation of neural circuits in crustacea. Amer. Zool. 29: 1305-1320.

Hockfelt, T., O. Johansson, and M. Goldstein. 1984. Chemical anatomy of the brain. Science. 225: 1326-34.

Holman, G.M., B.J.Cook, R.J.Nachman. 1986. Isolation, primary structure and synthesis of leucomyosuppressin, an insect neuropeptide that inhibits spontaneous contractiion of the cockroach hindgut. Comp. Biochem. Physiol. 85C: 329-333.

Innis, M.A., K.B. Myambo, D.H. Gelfand, and M.A.D. Brow. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. Proc. Natl. Acad. Sci. USA. 85: 9436-9440.

Innis, M.A. 1990. PCR Protocols: A Guide to Methods and Applications. Academic Press Inc., San Diego, California.

Joosse, J. 1988. What is special about peptides as neuronal messengers?, p. 1-3. *In* Thorndyke, M.C. and G.J. Goldsworthy, Neurohormones in Invertebrates. Cambridge University Press, Cambridge.

Kandel, E.R. 1983. Neurobiology and molecular biology: the second encounter. Cold Spring Harb. Symp. Quant. Biol. 48: 891-908.

Kavaliers, M., and M. Hirst. 1985. FMRFamide, a putative endogenous opiate antagonist: evidence from suppression of defeat-induced analgesia and feeing in mice. Neuropeptides 6: 485-494.

Kobayashi, M., and Y. Muneoka. 1989. Functions, receptors, and mechanisms of the FMRFamide-related peptides. Biol. Bull 177: 206-209.

Kobierski, L.A., B.S. Beltz, B.A. Trimmer, and E.A. Kravitz. 1987. FMRFamide-like peptides of *Homarus americanus*: distribution, immunocytochemical mapping, and ultrastructural localization in terminal varicosities. J. Comp. Neurol. 266: 1-15.

Kravitz, E.A., L.Kobierski, B.A.Trimmer, and M.F.Goy. 1987. Peptide F_1: a myoactive lobster peptide related to FMRFamide. Soc. Neurosci. Abstr. 13: 1257.

Kruk, Z.L. and C. J. Pycod. 1979. Neurotransmitters and Drugs. p.141. University Park Press, Baltimore.

Kuehl, W.M. and J. Battey. 1993. Generation of a polymerase chain reaction renewable source of subtractive cDNA. p. 287-305. *In* B.A. White, PCR Protocols: Current Methods and Applications. Humana Press, New Jersey.

Laufer, H., and R.G.H Downer. 1988. Endocrinology of selected invertebrate types. p. 278. Alan R. Liss, Inc., New York.

Lehman, H.K., Price, D.A. and M.J. Greenberg. 1984. The FMRFamide-like neuropeptide of *Aplysia* is FMRFamide. Biol. Bull 167: 460-466.

Lehman, H.K., and M.J. Greenberg. 1987. The actions of FMRFamide-like peptides on visceral and somatic muscles of the snail *Helix aspersa*. J. Exp. Biol. 131: 55-68.

Lehman, H.K. and D.A. Price. 1987. Localization of FMRFamide-like peptides in the snail *Helix aspersa*. J. Exp. Biol. 131: 37-53.

Li, Y.H. and S. Negre. 1993. Use of glycerol for enhanced efficiency and specificity of PCR amplification. TIG 9:297.

Linacre, A., E. Kellett, S. Saunders, K. Bright, P.R. Benjamin, and J.F. Burke. 1990. Cardioactive neuropeptide Phe-Met-Arg-Phe-NH_2 (FMRFamide) and novel related peptides are encoded in multiple copies by a single gene in the snail *Lymnaea stagnalis*. J. Neurosci, 10: 412-419.

Loh, L., and H. Gainer. 1983. Biosynthesis and processing of neuropeptides, p. 79-113. *In* Kreiger, D.T., M.J. Brownstein, and J.B. Martin, Brain Peptides. John Wiley & Sons, New York.

Lohn, Y.P., M.J. Brownstein, and H. Gainer. 1984. Proteolysis in neuropeptide processing and other neural functions. Annu. Rev. Neurosci. 7: 189-222.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratory Press, New York.

Maniatis, T., J. Sambrook, E.F. Fritsch. 1989. Molecular Cloning: a Laboratory Manual. 2nd ed. Vol I. Cold Spring Harbor Laboratory Press, New York.

Martin, R., D. Frosch, C. Kiehling and R. H. Voigt. 1981. Molluscan neuropeptide-like and enkephalin-like material coexists in octopus nerves. Neuropeptides 2: 141-150.

Maule, A.G., C.Shaw, J.W.Bowman, D.W.Halton, D.P.Thompson, T.G.Geary and L.Thim. 1994. KSAYMRFamide: a novel FMRFamide-related heptapeptide from the free-living nematode, *Panagrellus redivius*, which is myoactive in the parasitic nematode, *Ascaris suum*. Biochem. Biophys. Res.Comm. 200: 973-980.

Maxam,. A.M. and W. Gilbert, 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74: 560-564.

McPherson, M.J., P. Quirke, G.R. Taylor, eds. 1991. PCR, a Practical Approach IRL Press at Oxford University Press, Oxford, New York.

Meinkoth, J. and G.M. Wahl. 1987. Nick translation. pp.91-93. *In* S.L. Berger, and A.R. Kimmel, eds., Guide to Molecular Cloning Techniques. Academic Press, Inc., Orlando, FL.

Mercier, A.J. I. Orchard, and V. TeBrugge, 1991. FMRFamide-like immunoreactivity in the crayfish nervous system. J. Ex. Biol. 156: 519-538.

Mercier, A.J., I. Orchard, V. Te Brugge, and M. Skerrett. 1993. Isolation of two FMRFamide-related peptides from crayfish pericardial organs. Peptides 14: 137-143.

Mercier, A.J. and R.T. Russenes. 1992. Modulation of crayfish hearts by FMRFamide-related peptides. Biol. Bull. 182: 1-8.

Mercier, A.J., I. Orchard, and A.B. Lange. 1995. Neuropeptides associated with the hindgut of crayfish (*Procambarus clarkii*).p.A5. *In* W.W. Burggen, ed., Physiological Zoology. Vol. 68, No. 4. University of Chicago Press, Chicago.

Mercier,A.J., I. Orchard, and A.B.Lange. 1995. Neuropeptides associated with the hindgut of crayfish *(Procambarus clarkii)* Phys. Zoo. 68:52.

Miller, M., J. Baxter and B. Eberhardt. 1983. In Brain Peptides. D.T. Krieger, M.J. Brownstein, J.B. Martin eds. Wiley-Interscience Publications, John Wiley & Sons, New York.

Moqadam, F., and P.D. Siebert. 1994. Rapid amplification of the 3' end of cDNAs with the 3'-AmpliFINDER™ RACE kit. CLONTECHniques 4:6.

Morris, H.R., M. Panico, A. Karplus, P.E. Lloyd, and B. Riniker. 1982. Elucidation by FAB-MS of the structure of a new cardioactive peptide from *Aplysia*. Nature 300: 643-645.

Mullis, K.B. 1989. The polymerase chain reaction: why it works. p 237-243. *In* H.A. Erlich, R.Gibbs, H.H. Kazazion, Jr., Current Communications in Molecular Biology. Cold Spring Harbor Laboratory Press, U.S.A.

Mullis, K.B. and F. A. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. Methods Enzymol. 155: 335-350.

Nachman, R.J. G.M. Holman, W.F. Haddon, and N. Long, 1986. Leucosulfakinin, a sulfated insect neuropeptide with homology to gastrin and cholecystokinin. Science 234: 71-73.

Nambu, J.R., C. Murphy-Erdosh, P.C.Andrews, G.J.Feistner, and R.H.Scheller. 1988. Isolation and characterization of a *Drosophila* neuropeptide gene. Neuron 1: 55-61.

Nichols, R. 1992. Isolation and Structural Characterization of *Drosophila* TDVDHVFLRFamide and FMRFamide-Containing Neural Peptides. J. Mol. Neurosci 3: 213-218.

Nichols, R., S.A. Schneuwly, and J.E. Dixon. 1988. Identification and characterization of a *Drosophila* homologue to the vertebrate neuropeptide cholecystokinin. J. Biol. Chem. 263: 12167-12170.

Ogden, R.C., and D. A. Adams. 1987. Electrophoresis in Agarose and Acrylamide Gels. *In* S.L. Berger and A.R. Kimmel, eds. Guide to Molecular Cloning Techniques. Academic Press, Orlando.

Painter, S.D., and M.J. Greenberg. 1982. A survey of the responses of bivalve hearts to the molluscan neuropeptide FMRFamide and to 5-hydroxytryptamine. Biol. Bull. 162: 311-332.

Palade, G. 1975. Intracellular aspects of the process of protein synthesis. Science 189: 347.

Pasztor, V.M., and D.L. MacMillan. 1990. The Actions of Proctolin, Octopamine and Serotonin on Crustacean proprioceptors show species and neurone specificity. J. Exp. Biol. 152: 485-504.

Peaire, A. 1993. Isolation and partail purification of FaRP-encoding DNA sequences from the crayfish *Procambarus clarkii*. BSc. (Hon.) Thesis. Department of Biological Sciences, Brock University, St.Catharines, Ont.

Pearson, W.R. and D.J.Lipman. 1988. Improved tools for biological sequence comparison. PNAS USA 85: 2444-2448.

Price, D.A. 1983. FMRFamide: assays and artifacts p. 184-190. *In* J. Lever and H.H. Boer, eds. Molluscan Neuroendocrinology. North Holland Publishing Co., New York.

Price, D.A., W. Lesser, T.D. Lee, K.E. Doble, and M.J. Greenberg 1990. Seven FMRFamide-related and two SCP-related cardioactive peptides from *Helix*. J. Exp. Biol. 154: 421-437.

Price, D.A., and M.J. Greenberg. 1977. Structure of a molluscan cardioexcitatory neuropeptide. Science 197: 670 - 671.

Price, D.A., C.G. Cobb, K.E. Doble, J.K. Kline and M.J. Greenberg. 1987b. Evidence for a novel FMRFamide-related heptapeptide in the pulmonate snail *Siphonaria pectinata*. Peptides 8: 533-538.

Price, D.A., G.A. Cottrell, K.E. Doble, M.J. Greenberg,.Jorenby, H.K. Lehman, and J.P. Riehm. 1985. A novel FMRFamide-related peptide in *Helix*: pQDRFLRFamide. Biol. Bull. 169: 256-266.

Price, D.A. and M. J. Greenberg. 1977. Structure of a molluscan cardioexcitatory neuropeptide. Science 197: 670 671.

Price, D.A., and M.J. Greenberg. 1989. The hunting of the FaRPs: the distribution of FMRFamide-related peptides. Biol. Bull. 177: 198-205.

Price, D.A., and M.J. Greenberg. 1980. Pharmacology of the molluscan cardioexcitatory neuropeptide FMRFamide. Gen. Pharmacol. 11: 237-241.

Raffa, R.B. 1988. The action of FMRFamide (Phe-Met-Arg-Phe-NH_2) and related peptides on mammals. Peptides 9: 915-922.

Reichelt, K. L. and P.D. Edminson. 1977. Peptides containing probable transmitter candidates in the central nervous system, p. 171-179. *In* H. Gainer, Peptides in Neurobiology. Plenum Press, New York.

Reichlin, Seymour. 1981. Systems for the study of regulation of Neuropeptide Secretions. *In* M. Martin, and S.Reichlin, eds. Neurosecretion and Brain Peptides . Bick Raven Press, New York

Robb, S., and P.D. Evans. 1990. FMRFamide-like peptides in the locust: distribution, partial characterization and bioactivity. J. Exp. Biol. 149: 335-360.

Robb, S., L.C.Packman, and P.D.Evans. 1989. Isolation, primary structure and bioactivity of SchistoFLRFamide, a FMRFamide-like neuropeptide from the locust, *Schistocerca gregaria*. Biochem. Biophys. Res.Comm. 160:850-856.

Rosen, H., J. Douglass, and E. Herbert. 1984. Isolation and characterization of the rat proenkephalin gene. J. Biol. Chem. 259: 14309-14313.

Rosoff, M.L., T.R. Burglin, and C. Li. 1992. Alternatively spliced transcripts of the flp-1 gene encode distinct FMRFamide-like peptides in *Caenorhabditis elegans*. J. Neurosci. 12: 2356-2361.

Rychlik, W. 1993. Selection of primers for polymerase chain reaction pp. 31-40. *In* B.A.White, ed. PCR Protocols: Current Method and Applications. Humana Press, New Jersey.Saiki, R.K., 1989.

Saiki, R.K. 1989. Optimization of the Polymerase Chain Reaction p.25. H.A. Erlich, R. Gibbs, and H.H. Kazazion Jr. (eds). Current Communications in Molecular Biology. Cold Spring Harbor Laboratory Press, New York.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K. B. Mullis, and H. A. Erlich, 1988. Primer-directed enzymatic Amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim, 1985. Enzymatic amplification of beta-globin genomic sequences and restriction sites analysis for analysis of sickle cell anemia. Science 230: 1350.

Sanger, F., S. Nicklen, and A.R. Coulson, 1977. DNA Sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Saunders, S.E., K. Bright, E. Kellet, P.R. Benjamin, and J.F. Burke. 1991. Neuropeptides Gly-Asp-Pro-Phe-Leu-Arg-Phe-amide (GDPFLRFamide) and Ser-Asp-Pro-Pre-Leu-Arg-Phe-amide (SDPFLRFamide) are encoded by an exon 3' to Phe-Met-Arg-Phe-NH_2(FMRFamide) in the Snail *Lymnaea stagnalis*. J. Neurosci. 11: 740-745.

Saunders, S.E., E. Kellet, K. Bright, P.R. Benjamin and J.F. Burke. 1992. Cell-specific alternative RNA splicing of a FMRFamide gene transcript in the brain. J Neurosci 12: 1033-1039.

Schaefer, M., M.R. Picciotto, T. Kreiner, R.R. Kaldany, R. Taussig, and R.H. Scheller. 1985. *Aplysia* neurons express a gene encoding multiple FMRFamide neuropeptides. Cell 41: 457-467.

Schaller, H.C. 1979. Neuropeptides in *Hydra*. Trends Neurosci. 2: 120-122.

Scharrer, B. 1977. Peptides in neurobiology: historical introduction, p. 1-6. *In* H. Gainer, Peptides in Neurobiology. Plenum Press, New York.

Scharrer, B. 1990. The neuropeptide saga. Amer. Zool. 30: 887-895.

Schinkmann, K., and C. Li. 1994. Comparison of two *Caenorhabditis* genes encoding FMRFamide (Phe-Met-Arg-Phe-NH_2)-like peptides. Molecular Brain Res. 24: 238-246.

Schneider, L.E., and P.H. Taghert. 1990. Organization and expression of the *Drosophila* Phe-Met-Arg-Phe-NH_2 neuropeptide gene. J. Biol. Chem. 265: 6890-6895.

Schneider, L.E., and P.H. Taghert. 1988. Isolation and characterization of a *Drosophila* gene that encodes multiple neuropeptides related to Phe-Met-Arg-Phe-NH_2 (FMRFamide). Proc. Natl. Acad. Sci. USA. 85: 1993-1997.

Schwyzer, R. 1980. Organization and transduction of peptide information. Trends in Pharmacol. Sci. 3: 327-331.

Sherman, T.G., H. Akil, S. J. Watson, 1989. The Molecular Biology of Neuropeptides. p. 26. Vol VI, No. 1 Elsevier Science Publishers, Netherlands.

Shyamala, M., J.M. Fisher, and R.H. Scheller. 1986. A neuropeptide precursor expressed in *Aplysia* neuron L5. DNA 5: 453-462.

Skerrett, M., A. Peaire, P. Quigley, and A.J.Mercier. 1995. Physiological effects of two FMRFamide-related peptides from the crayfish *Procambarus clarkii.* J. Exp. Biol. 198, 109-116.

Soinila, S., and G.J. Mpitsos. 1991. Immunohisto-chemistry of diverging and converging neurotransmitter systems in mollusks. Biol. Bull. 181: 484-499.

Stern, A.S., R.V. Lewis, S. Kimura, J. Rossier, L.D. Gerber, L. Brink. S. Stein, and S.Udenfriend. 1979. Isolation of the opioid heptapeptide Met-enkephalin [Arg^6, Phe^7] from bovine adrenal medullary granules and striatum. Proc. Natl. Acad. Sci. USA. 76: 6680-6683.

Taghert, P.H. and L.E. Schneider. 1990. Interspecific comparison of a *Drosophila* gene encoding FMRFamide-related neuropeptides. J. Neurosci. 10: 1929-1942.

Takayanagi, H., and N. Takeda. 1988. Coexistence of FMRFamide, met-enkephalin and serotonin in molluscan neurons. Comp. Biochem. Physiol. 91A: 613-620.

Taussig, R., and R.H. Scheller. 1986. The *Aplysia* FMRFamide gene encodes sequences related to mammalian brain peptides. DNA 5: 453-461.

Trimmer, B.A., L.A. Kobierski, E.A. Kravitz. 1987. Purification and characterization of FMRFamide-like immunoreactive substances from the lobster nervous system: isolation and sequence analysis of two closely related peptides. J. Comp. Neurol. 266: 16-26.

Turrigiano, G.G., and A.I. Selverston. 1990. A cholecystokinin-like hormone activates a feeding-related neural circuit in lobster. Nature 344: 866-868.

Turrigiano, G.G., and A.I. Selverston. 1991. Distribution of cholecystokinin-like immunoreactivity within the stomatogastric nervous systems of four species of decapod crustacea. J. Comp. Neurol. 305: 164-176.

Veca, A. and J.H. Dreisbach. 1988. Classical neurotransmitters and their significance within the nervous system. J. Chem. Ed. 65: 108-111.

Voet, D., and J.G. Voet. 1990. Biochemistry. p.1223. John Wiley & Sons, New York.

Voight, K.H. and R. Martin. 1986. Neuropeptides with cardioexcitatory and opioid activity in octopus nerves. pp. 127-138. *In* G.B. Stefano, ed., CRC Handbook of Comparative Opioid and Related Neuropeptide Mechanisms. Vol I. CRC Press, Inc., Boca Raton.

Wahl, G.M., E. Ong, J. Meinkoth, R. Franco, and M. Barinaga. 1981. Methods for the transfer of DNA, RNA, and protein to nitrocellulose and diazotized paper solid supports. Scheicher and Schuell, Inc., Keene. NH.

Wallace, R.B. and C.G. Miyada. 1987. Oligonucleotide probes for the screening of recombinant DNA libraries, p. 432-442. *In* S.L. Berger, and A.R. Kimmel, eds., Guide to Molecular Cloning Techniques. Academic Press Inc., Orlando, FL.

Watson, J.D., N.H. Hopkins, J.W. Roberts, J.A. Steitz, and A.M. Weiner. 1987. Molecular Biology of the Gene 4th ed. p. 1163. Benjamin/Cummings Publishing Company, Inc. California.

Williams, R.G. 1983. Immunohistochemical studies of FMRFamide-like immunoreactivity in rat brain. Brain Res. 276: 213-229.

Wold, F. 1981. *In vivo* chemical modifications of proteins (post-translational modification). Annu. Rev. Biochem. 50: 783-814.

Wu, R., T. Wu, and A. Ray. 1987. Adaptors, Linkers, and Methylation. pp. 343-348. *In* S.L. Berger, and A.R. Kimmel, eds., Guide to Molecular Cloning Techniques. Academic Press Inc., Orlando, FL.

Yang, H-Y.Y., W. Fratta, E.A. Majane, and E. Costa. 1985. Isolation, sequencing, synthesis, and pharmacological characterization of two brain peptides that modulate the action of morphine. PNAS USA 82:7757-7761.