



Grasshopper (*Melanoplus differentialis*) lectin genes : southern analysis and polymerase chain reaction
by Tanya Gedik

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in
Biochemistry

Montana State University

© Copyright by Tanya Gedik (1996)

Abstract:

A component of an invertebrate's innate immune response to pathogens includes lectin proteins. Lectins have the ability to discriminate self from non-self by recognizing specific carbohydrates that are present on the surface of microorganisms. Lectins bind these carbohydrates and target them for humoral or cellular defensive reactions. Hemolymph of grasshopper, *Melanoplus differentialis*, contains a lectin with two carbohydrate recognition domains (CRDs) with specificity toward galactosidic and glucosidic carbohydrates (Stebbins and Hapner 1985). The protein, GHA, is a C-type lectin in light of its dependence on calcium for sugar binding activity. GHA is known to associate with fungal blastospores and aid in their removal from the hemolymph by hemocytes (Wheeler et al. 1993). GHA protein has been isolated, as have two related grasshopper lectin cDNA clones (Hapner K.D., Rognlie M.C. and Radke J.R. Unpublished results). These clones, Clone 3 and 4, show 80% sequence identity. Partial amino acid sequence of the GHA protein revealed that it was not encoded by Clone 3 or 4. This fact suggested that the grasshopper may contain multiple C-type lectins and may have multiple lectin genes encoding these proteins.

The objectives of this study are to confirm that grasshopper genomic DNA contains multiple C-type lectin genes and to determine the intron character of genes 3 and 4 coding for Clones 3 and 4, respectively. Primary methodology includes Southern analyses, polymerase chain reaction (PCR), endonuclease restriction and random primed probe preparation.

Restricted grasshopper genomic DNA gives multiple bands on autoradiographs hybridized with ³²P-labeled grasshopper C-type lectin cDNA probes. Interpretation of the results indicates the presence of at least four C-type lectin genes in the grasshopper genome. PCR amplification was performed on grasshopper genomic DNA with primer sets that anneal to either Clone 3 or 4. Restriction analyses of the PCR products indicated gene 3 and 4 to be the amplification products. Southern analysis, with grasshopper C-type lectin cDNA probe, proved the PCR product were amplified from C-type lectin sequences. The results strongly suggested that both CRD-coding regions of gene 4, and the carboxyl CRD-coding region of gene 3, lack introns. The intronless character of the CRD-coding regions of C-type lectin genes indicates possible evolutionary relationship with intron-lacking CRDs of lectins from other organisms.

GRASSHOPPER (*MELANOPLUS DIFFERENTIALIS*) LECTIN GENES: SOUTHERN
ANALYSIS AND POLYMERASE CHAIN REACTION

by

Tanya Gedik

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 1996

N378
G2679

APPROVAL

of a thesis submitted by

Tanya Gedik

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Kenneth D. Hapner Kenneth D Hapner 11-21-96
(Signature) Date

Approved for the Department of Chemistry and Biochemistry

David M. Dooley David M Dooley 11/21/96
(Signature) Date

Approved for the College of Graduate Studies

Robert L. Brown Rob Brown 12/6/96
(Signature) Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowed only for scholarly purposes, consistent with "fair use" as described in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature J. Hedite

Date 11/21/96

ACKNOWLEDGMENTS

I wish to thank my professor, Dr. Kenneth D. Hapner, for his support and guidance throughout my studies. I also thank the other members of my Graduate Committee: Dr Martin Teintze and Dr. Patrik R. Callis. I am grateful to my laboratory colleagues for their encouragement and enthusiasm: Jay R. Radke, Don L. Wenzlick and my sister, Layla Gedik. I thank my family, and close friends, who have supported me.

TABLE OF CONTENTS

| | Page |
|---|------|
| LIST OF TABLES | vii |
| LIST OF FIGURES | viii |
| ABSTRACT | x |
| | |
| INTRODUCTION | 1 |
| Insect Immunity | 1 |
| Animal Lectins | 3 |
| Classification of C-type Lectins | 5 |
| C-type Lectin Evolution | 6 |
| Insect Lectins | 7 |
| Published GHA Work | 8 |
| Current GHA Work | 9 |
| Research Rationale and Approaches | 10 |
| Southern Analysis | 11 |
| PCR Amplification | 12 |
| Research Objectives | 13 |
| | |
| MATERIALS AND METHODS | 14 |
| Primers and Probes | 14 |
| Probe Preparation from Plasmid | 19 |
| Probe Preparation by PCR | 19 |
| Radioactive Isotope Labeling of 580bp Probe | 20 |
| Biotin Labeling of 879bp Probe | 21 |
| DNA Electrophoresis | 23 |
| Grasshopper Genomic DNA Preparation | 23 |
| Grasshopper Genomic DNA Restriction | 25 |
| Southern Analysis | 26 |
| PCR Amplification of Genomic DNA | 27 |
| PCR Optimization using 3'152 and 3'NT Primers | 27 |
| Restriction Endonuclease Cleavage of PCR Products | 28 |
| Southern Analysis of PCR Products | 29 |
| Restriction Endonuclease Enzyme Activities | 30 |
| Standards and Controls | 31 |

| | |
|---|-----------|
| DNA Size Standards for Southern Analyses | 31 |
| Southern Analysis Controls | 32 |
| PCR Controls | 32 |
| RESULTS | 33 |
| Confirmation of Restriction Endonuclease Activity | 33 |
| Activities of Enzymes used in Genomic Digests | 33 |
| Activities of Enzymes used in Restriction of PCR Products | 37 |
| Southern Analysis Standards and Controls | 39 |
| Genomic DNA Controls | 39 |
| Hybridization Control | 40 |
| Southern Standard DNA Ladder | 41 |
| Preparation of 879bp Biotin-Labeled Probe | 41 |
| Preparation of 580bp Radiolabeled Probe | 43 |
| Grasshopper Genomic DNA Preparation | 44 |
| Grasshopper Genomic DNA Restriction | 46 |
| Determination of Lectin Gene Number | 46 |
| Biotin Southern Analysis | 46 |
| Radioactive Southern Analysis | 49 |
| PCR Optimization | 53 |
| Determination of Intron Nature of Lectin Genes | 55 |
| Restriction Analysis of 4052/3'NT PCR Products | 55 |
| Southern Analysis of 4052/3'NT PCR Products | 57 |
| Restriction Analysis of 3152/3'NT PCR Products | 58 |
| Southern Analysis of 3152/3'NT PCR Products | 61 |
| DISCUSSION | 64 |
| Optimization of Experimental Methodology | 64 |
| Southern Analysis | 64 |
| Biotin- versus Radio-Labeled Probes | 64 |
| C-Type Lectins in Salmon Sperm DNA | 66 |
| Optimization of PCR | 67 |
| Grasshopper Lectin Gene Number | 69 |
| Intronic Nature of Lectin Genes | 74 |
| Lectin Classification and Evolution | 76 |
| Newly Discovered Clone 4 Sequence | 78 |
| Future Work | 79 |
| CONCLUSIONS | 81 |
| REFERENCES CITED | 82 |

LIST OF TABLES

| Table | Page |
|---|------|
| 1. Primer Td's and Sequences | 18 |
| 2. DNA and Mg ²⁺ Concentrations used to Obtain the PCR Results Shown in Figure 11 | 54 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1. Map of Recombinant PGem Plasmid | 15 |
| 2. Alignment of Clone 3 and Clone 4 with Annealing Positions of Primers | 16 |
| 3. Illustration of Primer and Probe Annealing Sites on Clone 3 and 4 | 17 |
| 4. Assay of the Restriction Endonuclease Enzymes Utilized in Genomic Southern Analyses | 34 |
| 5. Restriction Endonuclease Activity in Presence of Grasshopper Genomic DNA | 36 |
| 6. Assays of the Restriction Endonuclease Enzymes Utilized in PCR Product Restriction Analysis | 38 |
| 7. Determination of Biotinylated 879bp Probe Concentration | 42 |
| 8. Appearance of Isolated Unrestricted and Restricted Grasshopper Genomic DNA on 1% Ethidium Bromide Agarose Gel | 45 |
| 9. X-ray Film of Southern Analysis on Grasshopper Genomic DNA and Salmon Sperm Control DNA | 47 |
| 10. Autoradiograph of Southern Analysis on 15 μ g Grasshopper Genomic DNA and 15 μ g Barley Control DNA | 52 |
| 11. Optimization of PCR with Grasshopper Genomic DNA Template using 3152 and 3'NT Primers | 54 |

| | |
|--|----|
| 12. PCR Amplification of Grasshopper Genomic DNA Template using Primers 4052 and 3'NT | 56 |
| 13. PCR Amplification of Grasshopper Genomic DNA using Primers 3152 and 3'NT | 59 |
| 14. Illustration of Intronless Nature of Genes Encoding Grasshopper Clones 3 and 4 cDNA | 62 |
| 15. The 5' Terminal Sequence of the Coding Region of Clone 4 | 79 |

ABSTRACT

A component of an invertebrate's innate immune response to pathogens includes lectin proteins. Lectins have the ability to discriminate self from non-self by recognizing specific carbohydrates that are present on the surface of microorganisms. Lectins bind these carbohydrates and target them for humoral or cellular defensive reactions. Hemolymph of grasshopper, *Melanoplus differentialis*, contains a lectin with two carbohydrate recognition domains (CRDs) with specificity toward galactosidic and glucosidic carbohydrates (Stebbins and Hapner 1985). The protein, GHA, is a C-type lectin in light of its dependence on calcium for sugar binding activity. GHA is known to associate with fungal blastospores and aid in their removal from the hemolymph by hemocytes (Wheeler et al. 1993). GHA protein has been isolated, as have two related grasshopper lectin cDNA clones (Hapner K.D., Rognlie M.C. and Radke J.R. Unpublished results). These clones, Clone 3 and 4, show 80% sequence identity. Partial amino acid sequence of the GHA protein revealed that it was not encoded by Clone 3 or 4. This fact suggested that the grasshopper may contain multiple C-type lectins and may have multiple lectin genes encoding these proteins.

The objectives of this study are to confirm that grasshopper genomic DNA contains multiple C-type lectin genes and to determine the intron character of genes 3 and 4 coding for Clones 3 and 4, respectively. Primary methodology includes Southern analyses, polymerase chain reaction (PCR), endonuclease restriction and random primed probe preparation.

Restricted grasshopper genomic DNA gives multiple bands on autoradiographs hybridized with ³²P-labeled grasshopper C-type lectin cDNA probes. Interpretation of the results indicates the presence of at least four C-type lectin genes in the grasshopper genome. PCR amplification was performed on grasshopper genomic DNA with primer sets that anneal to either Clone 3 or 4. Restriction analyses of the PCR products indicated gene 3 and 4 to be the amplification products. Southern analysis, with grasshopper C-type lectin cDNA probe, proved the PCR products were amplified from C-type lectin sequences. The results strongly suggested that both CRD-coding regions of gene 4, and the carboxyl CRD-coding region of gene 3, lack introns. The intronless character of the CRD-coding regions of C-type lectin genes indicates possible evolutionary relationship with intron-lacking CRDs of lectins from other organisms.

INTRODUCTION

Insect Immunity

Insects have been remarkably successful in evolution. Current estimates are that they make up 90% of all extant animal species and colonise all terrestrial ecological niches (Hoffmann 1995). Consequently, they are confronted by an extremely large variety of potentially harmful microorganisms. Insects are able to build up an efficient defense system that has both a physical and an innate facet. The hard external skeleton functions as a physical barrier to pathogen invasion. A current view (Hoffmann et al. 1996) describes the innate response of insects as three interconnected reactions. The first is the induction of proteolytic cascades by wounding, even when potentially harmful microorganisms are absent. The proteolytic coagulation cascade leads to localized blood clotting that may immobilize the foreign invader and allow other processes to destroy the pathogen, as well as restricting blood loss (Muta and Iwanaga 1996). The prophenoloxidase cascade leading to melanization of large invaders is another example of a proteolytic cascade. Potentially cytotoxic quinoid intermediates of melanin generated in the prophenoloxidase cascade are thought to have bactericidal and fungicidal activity (Vass and Nappi 1996). The second innate response includes a variety of cellular defense reactions, that consist predominantly of phagocytosis or encapsulation of invading

microorganisms. Phagocytosis involves endocytosis of pathogens, mainly by plasmatocytes and granular cells, with lysosomal breakdown. Encapsulation is a multicellular process in which foreign objects too large for phagocytosis are surrounded by hemocytes recruited from the circulation (Ratcliffe 1993). The cells lyse and flatten, forming a layer of cells around the foreign organism. Melanotic compounds may be deposited in the inner layers. This capsule may stop the growth and development of the invader or kill it directly. The third innate response is the induction of the transient synthesis of a battery of peptides by the fat body that are secreted into the hemolymph. Close to 100 antimicrobial peptides and proteins have now been characterized. They include defensins, magainins, cecropins, proline-rich and glycine-rich polypeptides. Understanding the mode of action of these peptides remains unsatisfactory due to their only recent discovery, although it has been proposed that cecropins could act as detergents thereby causing lysis of bacterial cells through the disintegration of their cytoplasmic membranes (Hoffmann et al. 1996). Another strongly held idea is that an insect's innate immune response includes a fourth component. This component involves lectin proteins that are thought to protect the insect from parasitic invasions by having the ability to discriminate self from non-self (Arason 1996). Lectins bind avidly and reversibly to carbohydrates. Carbohydrates are present on cell surfaces and carry, per unit weight, more information than can amino acids or proteins (Sharon and Liz 1995). Lectins can detect subtle differences in carbohydrate structures, a characteristic useful and important in biological recognition and differentiation.

Animal Lectins

Lectins are ubiquitous proteins that function in fertilization, development, leukocyte migration and self/non-self distinction (Arason 1996). The latter role originates from their ability to discriminate, through hydrogen bonding and hydrophobic interactions, between endogenous carbohydrates or those that are presented by microbial invaders. Animal lectins have enormous structural diversity but carbohydrate binding activity can often be ascribed to a limited polypeptide segment of each lectin, designated the carbohydrate-recognition domain (CRD) (Drickamer 1993). Several types of CRD have been discerned, each of which shares a pattern of invariant and highly conserved residues over a 115-140 amino acid region. Three major groups of animal lectins; P, S and C-types, contain CRDs with distinct sequence motifs. Proteins of the major lectin groups share properties beyond similarity of primary structure (Drickamer and Taylor 1993). For example, S-type lectins often are dependent on reducing agents, such as thiols, for full activity and they all bind β -galactosides. P-type CRDs bind mannose-6-phosphate as their primary ligand. The animal C-type lectins are characterized by a dependence on calcium for sugar binding activity (Drickamer 1994). They occur in serum, extracellular matrix, and membranes (Drickamer and Taylor 1993).

The C-type lectin family includes among others the hepatic asialoglycoprotein receptor (Lodish 1991), macrophage mannose receptor (Sharon and Liz 1995), selectins (Lasky 1992), and soluble collectins (Hoppe and Reid 1994). The hepatic asialoglycoprotein receptor is a membrane-bound lectin found on the surface of

hepatocytes. This receptor binds certain glycoproteins that have lost terminal sialic acid residues, and the receptor-ligand complexes are then internalized. The macrophage mannose receptor may participate in antimicrobial defense by mediating phagocytosis of infectious organisms that expose mannose-containing glycans on their surface (Sharon and Liz 1995). The collectins include the rat mannose binding protein (MBP) that mediates humoral defense either via complement fixation or by direct opsonization of potential pathogens (Drickamer 1993). The three dimensional structure of the CRD of rat MBP has been determined by X-ray crystallography (Weis et al. 1991). The structure appears divided by two transverse β -strands that separate a compact scaffold of two helices and two β -sheets from an extended loop. The loop creates a pocket for two calcium ions and a binding site for the carbohydrate ligand. The carbohydrate binding site is thus exposed at the surface of the CRD, allowing for binding to sugars contained within complex oligosaccharide chains. The selectin family members play a crucial role in leucocyte trafficking to sites of inflammation, and in the migration of lymphocytes to specific lymphoid organs (Lasky 1992). The X-ray crystal structure of E-selectin provided a second example of a C-type lectin CRD (Graves et al. 1994). The three dimensional structures of the CRDs of rat MBP and E-selectin are very similar, although loop regions flanking the carbohydrate binding site differ significantly. The difference leads to altered directionality of carbohydrate-binding residues as well as the complete lack of a pocket around the Ca^{2+} site. The changes in structure enable MBP to bind mannose while E-selectin recognizes a sialic acid analogue.

Classification of C-type Lectins

Most C-type lectins have alternative functional domains in addition to their CRDs. These additional functional regions can be classified into several groups (Bezouska 1991). Group I C-type lectins have hyaluronic acid-binding regions while group II lectins are joined to N-terminal membrane anchors to form type II transmembrane proteins. All of the CRDs of collectins are associated with collagenous domains and are classed as group III C-type lectins. Selectins, group IV, have epidermal growth factor-like domains, while group V proteins consist of a type I transmembrane domain. Group VI consists of merely one protein, the macrophage mannose receptor. The protein contains a fibronectin type-II repeat domain but, unlike the C-type lectins mentioned thus far, has multiple CRD domains. Some C-type lectins consist simply of isolated CRDs and form group VII. Such proteins occur in snake venom (Hirabayashi et al. 1991) and in some invertebrate body fluids such as BRA-2 and BRA-3 lectins from acorn barnacle (Takamatsu et al. 1994 and Takamatsu et al. 1993). Insect soluble C-type lectins also belong in group VII.

A particularly striking observation was made by Drickamer et al. (1991) when the gene structures of members of these C-type lectin groups were compared. The CRDs, from group I or II lectins, are encoded by three exons. The two introns within the CRD-coding regions of group II genes are found at exactly corresponding positions. Similarly, the CRDs in group I are found at nearly these same positions. Collectins and selectins, groups III and IV respectively, lack introns within their CRD-coding regions. Group VI,

the macrophage mannose receptor, is one C-type lectin that does not fall into the gene structure classification. Introns are found in all of its eight CRDs but their number varies. Evolutionary relationships of C-type lectins have been considered based on occurrence of introns in the CRD-coding regions (Bezouska et al. 1991).

C-type Lectin Evolution

Drickamer and collaborators (Bezouska et al. 1991) have proposed an order of events that may have occurred during the evolution of lectins containing the C-type CRDs. Divergence of intron-containing and intron-lacking CRDs preceded shuffling events in which the other functional domains were associated with the CRDs. For example, during evolution a CRD-encoding gene segment became juxtaposed to a collagenous domain and all of the group III C-type lectins derive from this single precursor. Similar arguments are made for the group I, II and IV proteins. Therefore, from the long-term evolutionary point of view, it appears useful to classify C-type lectins on the basis of their genetic organization rather than domain shuffling (Arason 1996).

It is not known how group VI, the macrophage mannose receptor, evolved (Drickamer 1993) but it is thought that duplication of CRDs that led to its generation must have been an early event, occurring at roughly the same period as the duplications that led to the progenitor CRDs for each of the other groups of C-type lectins.

Insect Lectins

Lectins from several insects have been isolated and characterized and have been proposed as defense molecules. It is thought that the *Sarcophaga* lectin, produced by the flesh fly *Sarcophaga peregrina*, has dual functions in defense and in development (Natori 1990). During ontogenesis of developing *S. peregrina* only certain cells proliferate to form body structures, while unwanted cells are eliminated. *Sarcophaga* lectin is essential in removal of the unnecessary cells and foreign pathogens by mediating cell lysis. Recently, a C-type lectin was found in *Drosophila melanogaster* that has similar functions in defense and development as does *Sarcophaga* lectin (Haq et al. 1996), although the two lectins are assumed not to be structurally related. In the silkworm, *Bombyx mori*, it was reported that the hemagglutinating activity increased significantly in the hemolymph after infection with cytoplasmic polyhedrosis virus (Mori et al. 1992). The *Bombyx* lectin protein is induced concomitantly with infection thereby suggesting the lectin's involvement in the silkworm's defense system. The beet armyworm, *Spodoptera exigua*, also has a lectin that has been characterized as a defense molecule (Pendland et al. 1988). A galactose-specific agglutinin purified from *S. exigua* sera opsonizes fungal cells having exposed galactose residues. These fungal cells are rapidly cleared from the *S. exigua* hemolymph in *in vivo* studies. The hemolymph of the American cockroach, *Periplaneta americana*, contains *Periplaneta* lectin that acts as an opsonin to facilitate phagocytosis of injected bacteria by hemocytes (Kawasaki et al. 1993). The fat body of the cockroach has recently been shown to contain a family of lectins with similar

sequence to *Periplaneta* lectin (Kawasaki et al. 1993). Kawasaki et al. claims their find is the first published demonstration of the presence of a lectin-related protein family in an insect. Unconfirmed data has indicated grasshopper, *Melanoplus differentialis*, as having a family of hemolymph lectin proteins. One of these proteins, named GHA for 'grasshopper hemagglutinin', is thought to have a role in defense through its pathogen agglutinating activity (Wheeler et al. 1993). Research on these lectins is the focus of work in Dr. Hapner's laboratory.

Published GHA Work

GHA, a C-type lectin found in the hemolymph of grasshopper *M. differentialis*, was purified by affinity chromatography on a column of Sepharose-galactose followed by elution with EDTA. The agglutinin has binding specificity toward galactosidic and glucosidic carbohydrates (Stebbins and Hapner 1985). Hemagglutination activity was destroyed by treatment of the hemoagglutinin with heat, trypsin or EDTA. The mature GHA, is a glycoprotein and was measured to be approximately 70kDa by non-reducing electrophoresis. The protein was shown to contain two disulfide-linked polypeptide chains. The hemagglutinin is released from fat body, ovary and testes tissues as demonstrated by metabolic incorporation of ³⁵S-methionine into the relevant tissue cultures (Stiles et al. 1988). The lectin does not opsonize asialo human erythrocytes, *Bacillus thuringiensis* bacteria nor spores of *Nosema locustae* (Bradley et al. 1989). The lectin does associate with blastospores from *Beauveria bassiana* (Wheeler et al. 1993).

Insects injected with *B. bassiana* blastospores treated with agglutinin have the fungal cells cleared more than twice as rapidly as those not treated. It is suggested that the grasshopper hemagglutinin has a role in immune recognition of this fungus and functions in its removal from the hemolymph.

Current GHA Work

Two clones, Clones 3 and 4, have been isolated and sequenced from a grasshopper fat body cDNA library (Radke J.R. Unpublished work). These clones are 80% identical with only one segment where they show significant differences. Clone 3 is 1221bp and includes a 972bp open reading frame (ORF) coding for 324 amino acids. The initiating codon, stop codon and polyA tail are represented. There is no sequence available to complete the 5' end of Clone 4's ORF. A 5' Rapid Amplification of cDNA Ends (RACE) procedure is underway to obtain the putative 120bp of missing 5' ORF. The first 19 amino acid residues of the ORF of Clone 3 are uncharged and mostly hydrophobic. These residues most probably represent a signal peptide. The coded amino acid sequence includes two glycosylation consensus sequences, at least one of which is glycosylated (Wenzlick D.L. Unpublished work). The amino acid sequence also includes two C-type lectin CRDs that are approximately 30% identical to one another and to other invertebrate C-type lectins. GHA has been HPLC-purified and subsequently undergone amino acid analysis, molecular mass determination and cyanogen bromide fragmentation with Edman sequencing (Hapner K.D. and Wenzlick D.L. Unpublished work). Accurate

molecular mass determination with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI/TOF MS) has demonstrated that the grasshopper lectin is a disulfide-bond stabilized dimeric molecule consisting of two glycosylated monomers of identical size. The dimeric GHA molecule has been measured to approximately 72kDa while the monomers have the equivalent mass of 36.1kDa. Edman protein sequencing of two cyanogen bromide fragments has produced sequences that differ from sequences found in Clones 3 or 4. Therefore, three different C-type lectin sequences have been documented in the lab. One hypothesis from these observations is that the grasshopper contains a family of C-type lectins.

Information of three dimensional structure was gained through computer modeling of the GHA CRDs (Radke J.R. Unpublished work). X-ray crystal structures of rat MBP (Weis et al. 1991) and E-selectin (Graves et al. 1994) were used as reference proteins. The sites of GHA expression are being determined through Northern analysis and reverse transcription PCR amplification (Gedik L. Unpublished work).

Research Rationale and Approaches

Knowledge of homologous Clone 3 and 4 cDNA sequences and the amino acid sequences from fragments of GHA lectin protein has suggested the presence of a family of grasshopper lectin genes. Southern analysis can confirm that a family of genes may encode multiple grasshopper lectins. Southern analyses use either radioactive or non-radioactive isotopes for DNA detection. The techniques may have differing

sensitivities. Comparison of the two techniques may determine which is more appropriate for genomic Southern analyses. Knowledge of intron occurrence in C-type lectin genes is useful in lectin classification and evolutionary relationships (Arason 1996). Gene structure can be investigated without the availability of a grasshopper genomic library. An indirect approach using PCR amplification may be used to determine the size of the lectin gene. A gene larger than the mRNA it encodes will suggest the gene contains introns. Southern analysis and PCR amplification techniques are briefly discussed below.

Southern Analysis

Southern analysis involves the detection of a specific fragment of DNA. The DNA of interest is immobilized onto a nylon membrane. Subsequently, a 'probe' is required. A probe is a DNA fragment of complementary sequence to the immobilized DNA. The probe is modified to allow for its detection. This modification involves the incorporation of biotinylated or radioactive nucleotides into the probe DNA. When the probe is added to immobilized DNA, complementary sequences anneal, and the target bands are visualized by autoradiography or chemiluminescence. One chemiluminescent detection method involves a complex of biotin and a streptavidin alkaline phosphatase conjugate. A phenylphosphate-substituted 1,2-dioxetane substrate is cleaved by the alkaline phosphate and this triggers the decomposition of the 1,2-dioxetane with the simultaneous production of light. The light emission is detected using X-ray film.

PCR Amplification

PCR is a method for the amplification of DNA sequences *in vitro*. PCR is based on a series of incubation steps at different temperatures. One set of these steps, referred to as a PCR cycle, allows the annealing and extension of two primers, usually 17- to 20-mers, complementary to the target. The temperature is then raised to denature the DNA. The PCR process is a repetition of the cycle. The target is copied with each cycle, resulting in an exponential amplification. With PCR, DNA sequences can be amplified by at least 10^5 fold and potentially as high as 10^9 fold (Saiki et al. 1988). Reaction setup at room temperature may allow for non-specific primer annealing and extension (Chou et al. 1992). Undesirable non-specific constructs that begin this way are amplified throughout the remaining PCR cycles, resulting in misprimed products. Hot start is a technique that ensures that the polymerase enzyme is unable to function during PCR set up at room temperature. Perkin-Elmer AmpliTaq Gold™ (Roche Molecular Systems Inc., Branchburg, NJ) was one of the thermostable DNA polymerases used in the PCR amplification. The enzyme is provided in an inactive state and high temperatures are required to activate the enzyme. Using a pre-PCR heat step provides a PCR hot start, since primer extension cannot occur during PCR set up when the enzyme is inactive. Another hot start method used recombinant Taq DNA polymerase (Life Technologies, Grand Island, NY) and Mg^{2+} -free PCR buffer. Mg^{2+} was provided in a wax bead and the Mg^{2+} released only once the bead melted at higher temperatures.

Research Objectives

The overall objective of this research is to gain more understanding of the genes encoding C-type lectins in grasshoppers. The specific goals are listed below:

1. Confirm that the grasshopper genome contains a family of C-type lectin genes.

Determine this through genomic Southern analyses. Subobjectives include comparison of radioactive and non-radioactive Southern analyses.

2. Determine if genes representing Clone 3 and 4 are continuous or discontinuous and contain introns. Sub-objectives involve strategic primer design, PCR optimization and confirmatory differential endonuclease restriction analysis.

MATERIALS AND METHODS

Primers and Probes

Two probes were utilized in Southern analyses. One probe was obtained by cleaving out the 879bp grasshopper cDNA insert, from pGem 3.0 recombinant plasmid, using EcoRI and Acc I restriction enzymes (figure 1). EcoRI alone cleaves out the 879bp grasshopper insert but also generates a fragment of phage and plasmid DNA that is 920bp in length. This latter fragment may not be resolved on an agarose gel, making isolation of the grasshopper insert difficult. Thereby, the plasmid was cleaved with Acc I to cut the 920bp fragment into smaller sizes. The cDNA fragment, referred to as '879bp' probe, represents 72% of the total sequence of Clone 3. The second probe, named '580bp' probe, was PCR amplified from pGem 3.0 template with the primers 5'B and 3'D (figure 3). The regions of Clones 3 and 4 where the probes anneal are shown in figure 2.

Oligomer primers, required for PCR experimentation, were purchased from NBI (National Biosciences Inc., Plymouth, MN). Primers were required that were either specific to individual grasshopper cDNA clones or annealed to both clones. Primer design is an important part of PCR optimization. Rules in the design of efficient primers include length between 17-25bp, 50-60% GC composition, above 55°C T_d, non-complementarity at the 3' ends of primer pairs and non-complementarity to self. All these factors were considered when the primers were designed from grasshopper Clone 3 and 4

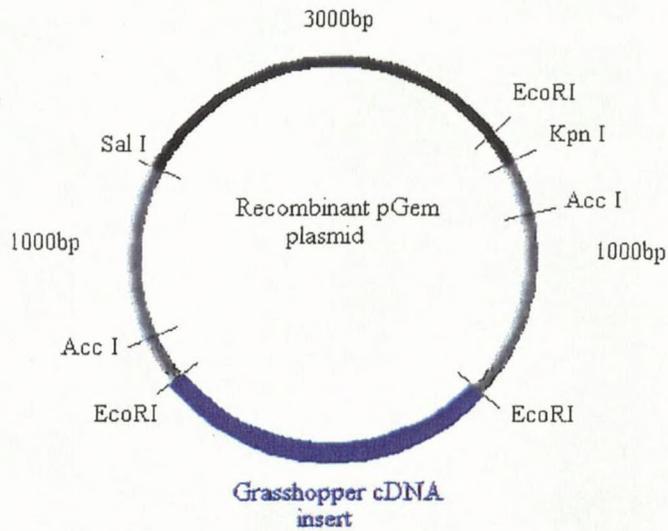


Figure 1. Map of Recombinant PGem Plasmid. Black portion indicates pGem-7Zf(+) plasmid. Grey regions indicate λ gt11 DNA. The blue region represents inserted 879bp grasshopper Clone 3 cDNA. The plasmid was utilized in probe DNA preparation, determination of restriction enzyme activities and creation of a standard ladder for Southern analyses.

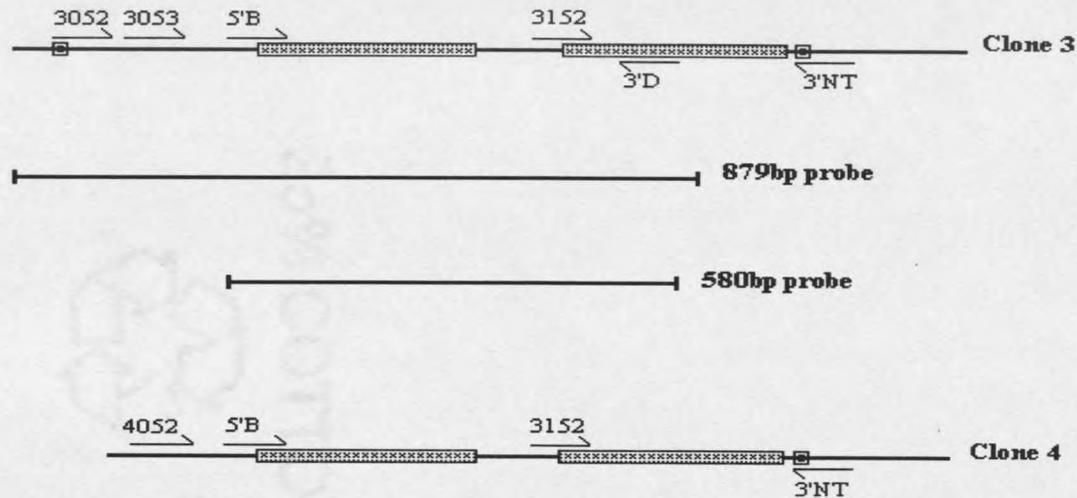


Figure 3. Illustration of Primer and Probe Annealing Sites on Clone 3 and 4. Sequences run 5' to 3'. Squares represent start and stop translation codons. Stippled boxes represent carbohydrate recognition domains. The probes were either biotinylated or radiolabeled and used in Southern analyses. The primers were utilized in PCR experiments on grasshopper genomic DNA to determine the intronic character of the genes coding Clones 3 and 4.

Table 1. Primer Td's and Sequences. Td's were generated by the 'nearest neighbour' method and calculated in the OLIGO computer program (National Biosciences Inc., Plymouth, MN). See figure 3 for primer annealing sites.

| PRIMER | PRIMER Td (°C) | PRIMER SEQUENCE (5'-3') |
|--------|----------------|-------------------------|
| 3052 | 69.1 | ATGCAGCTGG TGACGGTGTG |
| 3053 | 67.5 | CACCACAGGG ACTCGACGAC |
| 5'B | 61.5 | TCAAGCTGTA CCGCATAATG |
| 3152 | 66.7 | TCTACAAGGT GCCACGCCGA |
| 3'D | 61.8 | CGGTAACGAA GTCACCTTCC |
| 3'NT | 65.6 | GTCTGGGCCA TTCGCAGTTG |
| 4052 | 62.1 | ACAAAACGTG TCAAAAAGCC |

sequences (figure 2). The primer sequences were thoroughly examined on the computer software OLIGO (National Biosciences Inc., Plymouth, MN). Primers selected are listed in table 1.

Probe Preparation from Plasmid

The 879bp grasshopper cDNA insert was cleaved out of pGem 3.0 recombinant plasmid to create the 879bp probe as shown in figure 1. The 50 μ l reaction mixtures contained 2 μ g pGem 3.0, 1X buffer 4 (New England Biolabs, Beverly, MA) (20mM Tris-acetate, 10mM magnesium acetate, 50mM potassium acetate, 1mM DDT at pH 7.9), 1 μ l Acc I (10 Units) and 1 μ l EcoRI (12 Units). Reactions were incubated overnight at 37°C and reactions were terminated by addition of 4 μ l loading dye (5% glycerol, 0.01% bromophenol blue, 0.01% xylene cyanol, 0.6mM EDTA, 0.1% SDS). Restriction products were electrophoresed as described later. The 879bp band was excised from the gel and purified with Prep-A-Gene® (Bio-Rad Laboratories, Hercules, CA). Concentration of the purified DNA was estimated by comparative agarose electrophoresis with known λ DNA standards.

Probe Preparation by PCR

The 580bp probe was PCR amplified from pGem 3.0 using the primer set 5'B and 3'D as shown in figure 3. The amplifications were performed in 0.5ml micro-centrifuge

tubes (VWR Scientific Products, West Chester, PA). One unit recombinant Taq DNA polymerase (Life Technologies, Grand Island, NY) was added to each 50 μ l reaction (1X Mg²⁺-free buffer that contained 60mM Tris, 15mM (NH₄)₂SO₄, pH 8.5; 0.2mM each deoxyribonucleotides; 0.2 μ M 5'B primer; 0.2 μ M 3'D primer; 1ng pGem 3.0 plasmid). A HotWax™ Mg²⁺ bead (Invitrogen Corporation, San Diego, CA) was added to each tube to provide 2.5mM Mg²⁺ final concentration. 'Hot start' PCR was performed in a Perkin Elmer Thermal Cycler (Roche Molecular Systems Inc., Branchbury, NJ). The cycling parameters were:- an initial 3 minutes at 94°C; 35 cycles of 45 seconds at 94°C for denaturing, 45 seconds at 55°C for primer annealing and 2 minutes at 72°C for extension. The final extension time was 10 minutes followed by soaking at 4°C. The product was purified in a QIAquick PCR Purification Kit spin column (QIAGEN Inc., Chatsworth, CA) that separates fragments of 100bp or larger from fragments smaller than 40bp. The concentration of purified 580bp DNA was estimated by electrophoresis with λ DNA as concentration standards.

Radioactive Isotope Labeling of 580bp Probe

A 25ng (45 μ l) aliquot of 580bp probe DNA was denatured in water at 85°C for 10 minutes and cooled on ice for 10 minutes. The denatured DNA was added to a Ready-To-Go™ DNA Labeling Bead reaction tube (Pharmacia Biotech Inc., Piscataway, NJ) containing ingredients for random priming incorporation of label. To this mixture was added 5 μ l (50 μ Ci) of [α ³²P]dCTP (DuPont NEN® Research Products, Boston, MA).

The tube was incubated at 37°C for 1 hour. The reaction mixture was then spun through a Bio-spin® 30 Chromatography Column (BioRad Laboratories) at 1,100xg for 4 minutes. The DPM, disintegrations per minute, activity of the collected, purified sample was determined in an instrument detecting Cerenkov radiation (Bioscan Inc., Washington, DC). The specific activity (DPM/μg) of the probe was calculated. The calculation assumed that the amount of probe DNA doubled during the random primed incorporation of label, therefore terminating with 50ng of probe DNA. The approximate specific activity was 2×10^9 DPM/μg, with 1×10^9 DPM/μg being the recommended minimum activity (Sambrook et al. 1989). After labeling, the probe was denatured in a waterbath at 85°C for 15 minutes, transferred to ice, and added to the hybridization buffer of a Southern analysis. The probe will bind to complementary or similar sequences on the Southern membrane.

Biotin Labeling of 879bp Probe

The 879bp probe was labeled with biotin using USB™ Random Primed Images® Biotin Labeling Kit (United States Biochemical Corporation, Cleveland, OH). A 25ng (10μl) aliquot of 879bp DNA was mixed with 4μl water and 2μl reaction mixture (random hexanucleotide mixture in reaction buffer). This mixture was boiled in a water bath for 10 minutes and transferred to ice for 5 minutes. A 3μl aliquot of the nucleotide mixture (0.167mM dGTP, 0.167mM dATP, 0.167mM dTTP, 0.125mM biotin-14-dCTP, 0.042mM dCTP in TE, pH 7.5 buffer) and 1μl exonuclease-free Klenow enzyme were

added and the reaction was incubated overnight at 37°C. The reaction was terminated by addition of 2µl 0.2M EDTA, pH 8.0.

The biotinylated 879bp probe concentration was determined by dot blot analysis according to the protocol supplied by USB Gene Images® Non-Isotopic Nucleic Acid Detection Kit. The 879bp DNA and amp^r control probe were diluted in TE buffer containing 25µg/ml herring DNA (Life Technologies). Dilutions were pipetted onto a damp Hybond™-N+ positive nylon membrane (Amersham International, Cleveland, OH) that was previously soaked in 2X SSC solution (0.3M NaCl, 0.03M sodium citrate, pH 7.4). The membrane was baked at 80°C for 30 minutes followed by 15 minutes agitation in blocking buffer (0.5% casein, 50mM Tris-Cl, 100mM NaCl, 0.1% SDS, pH 10) at room temperature. A 1:5000 volume of streptavidin alkaline phosphatase (SAAP) was added to the blocking buffer and the membrane agitated for a further 10 minutes. The post-SAAP wash involved soaking the membrane in post-SAAP wash buffer (0.05M Tris, pH 10, 0.10M NaCl, 0.1% SDS) for 2 minutes followed by rinsing in water for 30 seconds. The Post-SAAP wash was repeated three times. Finally, the blot was washed in 200ml post-SAAP wash buffer that contained no SDS. Next, the chemiluminescent LumiPhos® 530 (United States Biochemical Corp.) was sprayed onto the membrane and the membrane sealed in a Micro-Seal® plastic bag (Dazey Corporation, Industrial Airport, KS). The blot was placed in the dark for 12 hours and was subsequently exposed to Kodak X-OMAT™ film (Eastman Kodak Company, Rochester, NY) for approximately 12 hours and developed. The intensities of the 879bp and control probe signals on the film were compared to estimate the concentration of biotinylated 879bp probe. The

average concentration was 10ng/ μ l of biotinylated probe, a concentration sufficient for use in Southern analyses.

DNA Electrophoresis

Agarose gel electrophoresis was utilized in both Southern and PCR experiments. The electrophoresis used a 1.0% agarose gel containing 0.1 μ g/ml ethidium bromide. The DNA samples were electrophoresed in 0.5X TBE running buffer (45mM Tris, 45mM boric acid, 1.2mM EDTA, pH 8) at 90V for approximately 90 minutes. Genomic Southern electrophoresis was performed in a 12.5cm x 19cm flat bed submarine electrophoresis apparatus and the samples run for 7 hours at 60V. Gel bands were visualized with a 312nm UV light box (Spectronics Corporation, Westburg, NY) and photographed on a videographic printer (Ultra Lum Inc., Carson, CA).

Grasshopper Genomic DNA Preparation

Adult *Melanoplus differentialis* grasshoppers were provided by the USDA Rangeland Insect Laboratory (Montana State University, Bozeman, MT). The insects were insectary housed and maintained on a diet of bran and lettuce. Isolation of genomic DNA was according to Sambrook et al. (1989). Insects were washed for 2 minutes each in:- soapy water, 1% bleach, and water. The insects were pinned and dissected via a ventral incision from anus to head. The head and gut were removed. The carcass was

ground in liquid nitrogen and then placed in a 50ml Fisherbrand® sterile, polypropylene centrifuge tube (Fisher Scientific, Pittsburgh, PA). Digestion buffer (100mM NaCl, 10mM Tris HCl, 25mM EDTA, 0.5% SDS, 0.1mg/ml proteinase K) was added at a concentration of 1.2ml per 100mg carcass weight. The tube was agitated overnight at 50°C. Nucleic acids were extracted with phenol/chloroform/isoamyl alcohol (24:24:1). This involved adding an equal volume of phenol/chloroform/isoamyl alcohol to the tube and then spinning the tube at 1700xg for 10 minutes. The top aqueous layer, containing the nucleic acids, was pipetted into a new 50ml polypropylene centrifuge tube and the bottom organic phase was discarded. The phenol/chloroform/isoamyl alcohol extraction was repeated and the resulting aqueous layer transferred to a 15ml Fisher sterile polypropylene centrifuge tube. The nucleic acids were recovered with ethanol and ammonium acetate. One half volume of 7.5M NH₄Ac and two volumes of ice cold 100% ethanol were added to the recovered aqueous layer. The mixture was incubated at -20°C for 30 minutes to precipitate the DNA. The precipitated DNA was pelleted by spinning the tube at 12000xg for 3 minutes. The supernatant was discarded, then 1ml 70% ethanol was added to the DNA and the tube flicked to wash the pellet. The tube was then spun at 12,000xg for 5 minutes. The supernatant was discarded and the DNA pellet dried for 5 minutes in a SpeedVac Concentrator (Savant Instruments Inc., Farmingdale, NY). The dried pellet was resuspended in 1ml TE buffer. DNase-free RNase (Boehringer Mannheim Corporation, Indianapolis, IN) at 1µg/ml and 0.1% SDS were added to the redissolved DNA and incubated for 1 hour at 37°C. Nucleic acid extraction, beginning with addition of phenol/chloroform/isoamyl alcohol (24:24:1), was repeated two more

times but the final extraction was terminated after the addition of 1ml TE buffer. An aliquot of the isolated genomic DNA was electrophoresed and examined under UV light to determine the integrity of the genomic DNA. This included ensuring no DNA shearing through the lack of ethidium bromide smearing on the gel. Also, lack of RNA contamination was determined by lack of ethidium bromide fluorescence between 5-0.16 kbp. Absorbancies at OD₂₆₀ and OD₂₈₀ were measured in a Techtronic double beam UV-vis spectrometer. OD₂₆₀ readings measure the concentration of the isolated grasshopper genomic DNA and OD₂₆₀/OD₂₈₀ ratios measure its purity.

Grasshopper Genomic DNA Restriction

Southern analysis requires the isolated genomic DNA to undergo digestion with a restriction endonuclease enzyme. EcoRI digestion of genomic DNA shall be described and this procedure can be applied to other restriction enzyme digests of grasshopper genomic DNA. A 15µg aliquot of grasshopper genomic DNA was pipetted, with wide-mouthed pipette tips (Rainin Instrument Co. Inc., Woburn, MA) into 3.5µl of 1X buffer H (Promega) and the mixture was made up to 32µl with water. This mixture was incubated at 4°C for 2 hours to aid in solubilization of the genomic DNA. A 1.5µl (18 Units) aliquot of EcoRI (New England Biolabs) was added to the mixture and incubated overnight at 37°C. A second 1.5µl aliquot of EcoRI enzyme was then added to ensure continued activity. The reaction was incubated a further 12 hours at 37°C and was terminated by transferring to ice and addition of 4µl loading dye.

Southern Analysis

Restriction enzyme-cleaved genomic DNA samples were electrophoresed and the gel was subsequently soaked according to the USB Gene Images™ Non-Isotopic Nucleic Acid Detection Kit protocol. This involved soaking the gel for 10 minutes in 0.25M HCl, then for 15 minutes in 0.6M NaCl/0.4M NaOH, followed by 30 minutes in fresh 0.6M NaCl/0.4M NaOH. The gel was then agitated twice in fresh 10X SSC, each for 15 minutes. DNA in the gel was transferred, through capillary transfer, to a Hybond™-N+ positive-nylon membrane under neutral conditions. The membrane was subsequently baked at 80°C for 30 minutes to immobilize the DNA. The nylon membrane with immobilized DNA was agitated in hybridization buffer (7% SDS, 1% casein, 1mM EDTA, 0.25M Na₂HPO₄, pH 7.4) in a glass tube for 4 hours at 65°C. The denatured, labeled probe was added to the hybridization buffer and the membrane agitated for a further 24 hours during which the probe binds to complementary and homologous sequences.

The hybridization buffer was discarded and washes were undertaken. For radiolabeled probe, the membrane was washed with 1X SSC/0.1% SDS as follows:- 5 minutes at room temperature; 30 minutes at 65°C; 30 minutes at 65°C. For biotin-labeled probe, hybridization washes were performed as suggested in the USB Gene Images™ Non-Isotopic Nucleic Acid Detection Kit protocol. Hybridization washes commenced with agitation of the membrane in 200ml 2X SSC/0.5% SDS for 2 minutes then rinsing in water, followed by a wash in 2X SSC/0.5% SDS for 20 minutes. The membrane was

then rinsed in water and washed twice in 200ml 0.2X SSC/0.1% SDS at 65°C for 30 minutes each, with water rinses between. Finally, the membrane was agitated in blocking buffer for one hour, followed by addition of a 1:5000 dilution of SAAP. The membrane was soaked in this SAAP solution for 10 minutes. The following post-SAAP washes, LumiPhos[®] 530 application and chemiluminescence detection were done as described earlier. An exception from the previously described procedure was the membrane was washed for 20 minutes in 0.2X SSC/0.1% SDS rather than 3 minutes. Membranes were sealed in Micro-Seal[®] plastic and exposed to Kodak X-OMAT[™] film for approximately 24 hours.

PCR Amplification of Genomic DNA

PCR Optimization using 3152 and 3'NT Primers

Primer design, Mg²⁺ ion concentration, and template concentration are all important factors in PCR amplification reactions (Innis and Gelfand 1990). Primers were designed to be suitable in PCR. Mg²⁺ ion concentration, and amount of grasshopper genomic template utilized required optimization. These two parameters were varied as shown in the statistical 'central composite design' of Boleda et al. (1996). The 0.5ml reaction tubes (VWR Scientific Products, West Chester, PA) contained a 50µl reaction mixture with GeneAmp[®] 1X PCR Buffer (Roche Molecular Systems Inc., Branchburg, NJ) composed of 10mM Tris-HCl, 50mM KCl, 1.5mM MgCl₂, 0.001% (w/v) gelatin, pH

8.3; 0.2 mM each deoxyribonucleotides; 1.25U AmpliTaq Gold™ DNA polymerase (Roche Molecular Systems Inc., Branchburg, NJ); 0.2 μ M 3132 primer; 0.2 μ M 3'NT primer. DNA amount used was either 115ng, 195ng, 390ng, 585ng or 665ng. Mg²⁺ concentration utilized was either 1.1mM, 1.5mM, 2.5mM, 3.5mM or 3.9mM. A 25mM MgCl₂ solution (Life Technologies) was used to increase the final Mg²⁺ concentration above the initial 1.5mM present in the buffer. For Mg²⁺ ion concentrations lower than 1.5mM, Mg²⁺-free buffer (Life Technologies) was the buffer used and 25mM MgCl₂ added. A 3 μ l aliquot of loading dye was added and the PCR product electrophoresed with Rsa I-digested PUC DNA (donated by Talbot L., Montana State University, Bozeman, MT) as standard ladder. Yields of the amplified fragments were compared by noting ethidium bromide intensities. A 585ng aliquot of grasshopper genomic DNA and 3.5mM Mg²⁺ resulted in highest yield and were used for subsequent PCR amplifications. All other PCR conditions were unaltered.

Restriction Endonuclease Cleavage of PCR Products

Restriction reactions were undertaken on a number of PCR products to determine if specific genes had been amplified. The enzymes chosen were Sal I and Bgl II. Sal I cleaves Clone 4 and not Clone 3, while Bgl II cleaves only Clone 3, based on the known sequence of Clones 3 and 4. After amplification, PCR products were purified through QIAquick Spin PCR Purification Kit spin columns. Reaction mixtures for Sal I restriction involved adding 6 μ l of 10X buffer D (Promega) and 1 μ l (10 Units) of Sal I

enzyme (Promega) to the purified PCR product and making up to 60 μ l with water.

Reaction set up was identical for Bgl II except buffer M and Bgl II enzyme (Boehringer Mannheim) were utilized. Reactions were terminated by transfer to ice and addition of 7 μ l loading dye. Aliquots of 30 μ l restriction products were electrophoresed in a 4.0% polyacrylamide gel to ensure adequate resolution of low molecular weight DNA fragments. The gel was made up as described in Sambrook et al. (1989) and polymerized in a vertical Mini-PROTEAN II Electrophoresis System (Bio-Rad) apparatus.

Electrophoresis was continued for 90 minutes at 90V. The gel was soaked in 0.5 μ g/ml ethidium bromide for 20 minutes and then photographed on a videograph printer (Ultra Lum Inc.).

Southern Analysis of PCR Products

Southern hybridization experiments were performed on a number of PCR products. After amplification, 40 μ l aliquots of PCR products were electrophoresed and subsequently transferred to a HybondTM-N+ positive-nylon membrane. Hybridization, washes, and exposure to film were carried out as described previously. An autoradiograph signal, resulting from specific binding of lectin cDNA probe, would confirm that a lectin gene had been amplified.

Restriction Endonuclease Enzyme Activities

Restriction enzymes utilized in this research project were tested to ensure they were active. Lambda DNA (15 μ g) replaced genomic DNA for assays verifying the activity of restriction enzymes used in genomic grasshopper restrictions. Other conditions were kept identical to genomic restriction reactions. Active enzymes would cleave the λ DNA to produce predicted sizes upon agarose gel electrophoresis. Another type of experiment ensured the activities of Kpn I and Sac I restriction endonuclease enzymes in genomic DNA cleavage. This involved addition of 55pg of pGem 3.0 plasmid to the Kpn I/Sac I genomic DNA restriction. The restriction reaction and subsequent Southern analysis followed the procedure described previously. Active enzymes would cleave the plasmid and produce a 3kbp fragment containing the 879bp grasshopper cDNA. A 3kbp fragment visible on a Southern X-ray film would indicate the enzymes were active.

Restriction enzymes were also used in PCR-product cleavage as described later. These enzymes were shown to be active by incubating each with pGem plasmid, that contained a grasshopper cDNA insert, and had a restriction site for that enzyme. The 20 μ l reaction mixture contained 1 μ l restriction endonuclease enzyme (10-18 Units), 200ng pGem recombinant plasmid with Clone 3 or 4 cDNA insert, and 2 μ l 10X buffer recommended by the enzyme manufacturers. The reaction was incubated for 30 minutes at 37°C. The reactions were terminated by transferring to ice and addition of 3 μ l loading dye. Restriction products were electrophoresed and their size estimated by comparison to

ϕ X174/Hae III DNA standard ladder (Promega).

Standards and Controls

DNA Size Standards for Southern Analyses

A DNA ladder was required for estimation of DNA fragment size. The ladder was generated from the products of selected restriction reactions on recombinant pGem plasmid 3.0 (figure 1). The restriction reactions produce fragments containing the 879 bp grasshopper cDNA insert. These fragments bind the '879bp' and '580bp' probes and are therefore visible on autoradiographs. The restriction enzymes utilized and the subsequent fragments generated are:-Kpn I = 5825bp; Kpn I/Sac I = 2863bp; Acc I = 1757bp, EcoRI = 879bp. The reaction mixtures were 50 μ l volume and contained 4.45 μ g pGem 3.0, 2 μ l restriction endonuclease enzyme (20-24 Units), and 5 μ l 10X buffer supplied with the enzyme. The reactions were incubated at 37°C for 72 hours. Reactions were terminated by addition of 3 μ l loading dye. Restriction products were electrophoresed and the bands that contained the 879 bp insert were excised from the gel and purified with Prep-A-Gene[®]. These DNA fragments, referred to as the Southern ladder, were electrophoresed, transferred to nylon membranes and served as size standards for Southern analyses.

Biotinylated Hind III-digested λ DNA (New England Biolabs) was used as the standard ladder in biotin-labeled probe Southern analyses.

Southern Analysis Controls

Negative controls for Southern hybridization experiments included salmon sperm DNA (Life Technologies) and barley DNA (donated by Talbot L., Montana State University, Bozeman, MT) that were restricted with the same protocol as grasshopper genomic DNA. Other controls were pGem 3.0 restricted with Acc I and EcoRI restriction endonucleases. These enzymes cut the recombinant plasmid at five positions. One cleavage product is the 879bp grasshopper cDNA insert that is complementary to 879bp and 580bp probes. This restricted fragment was a positive control. Restricted plasmid fragments that do not include grasshopper insert cDNA served as negative controls.

PCR Controls

Controls for PCR amplification were setup identically to the other PCR reactions except the negative control had no template DNA and the positive control had 1ng recombinant pGem plasmid as the template.

RESULTS

Confirmation of Restriction Endonuclease Activity

Activities of Enzymes used in Genomic Digests

Restriction endonuclease enzymes were used for digestion of genomic DNA prerequisite to Southern analyses. These enzymes were assayed to ensure they were active. The restriction reaction conditions followed those used for digestion of grasshopper genomic DNA, with the exception that the DNA digested was 15 μ g λ DNA. The digested products were electrophoresed in ethidium bromide agarose gels and gave results as shown in figure 4. The observed DNA fragment sizes correlated with the sizes expected, with the exception of Kpn I (lane 4). Two fragments, at 29.9kbp and 17kbp, were anticipated in the Kpn I restriction. The 17kbp band can be visualized but the 29.9kbp fragment does not appear on the gel. Kpn I is known to be active due to its ability to restrict plasmid DNA as shown in figure 5. The reason for lack of the 29.9kbp fragment in lane 4 may be due to insufficient resolution on the gel to enable differentiation of the two restriction fragments. Overall, it appears the enzymes show complete activity toward cleavage of 15 μ g λ DNA. These enzymes and restriction conditions were deemed sufficient for digestion of 15 μ g genomic DNA in Southern analyses.

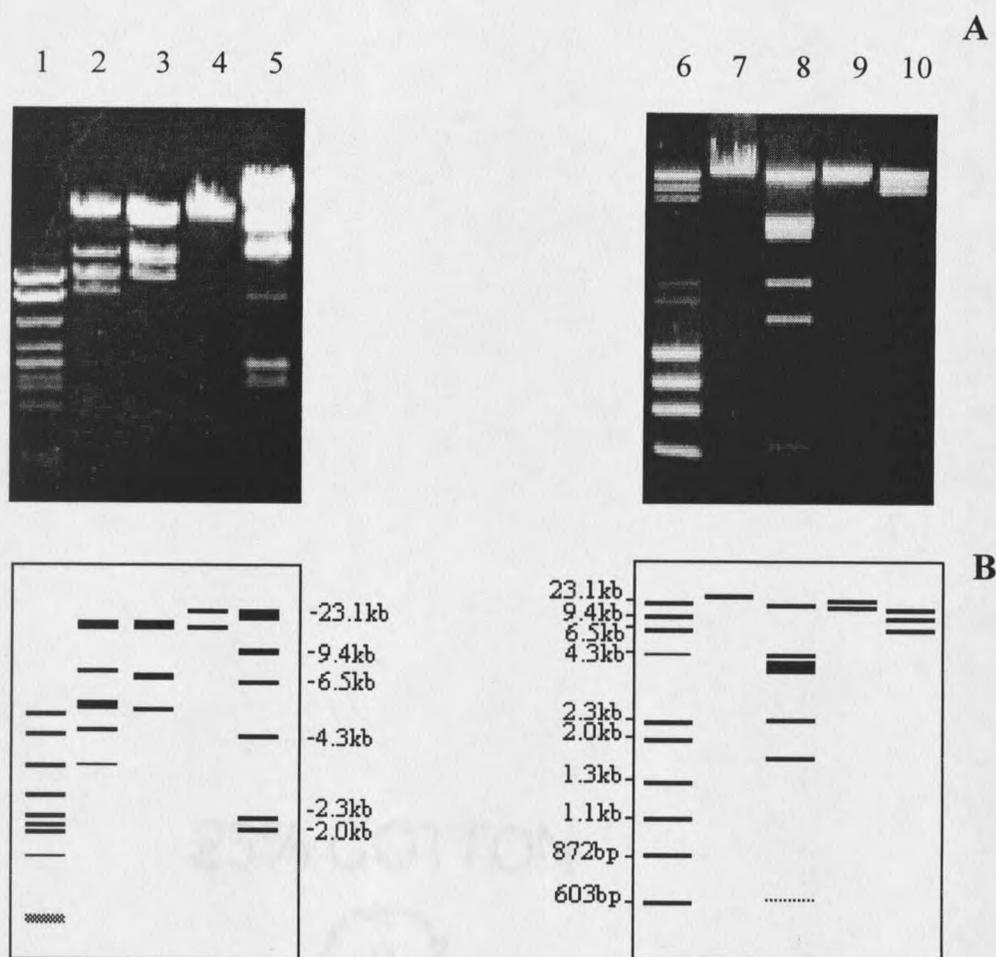


Figure 4. Assay of the Restriction Endonuclease Enzymes Utilized in Genomic Southern Analyses. Aliquots of 15 μ g λ DNA were cleaved under conditions identical to grasshopper genomic restriction in Southern analyses. A) Ethidium bromide agarose gels of restriction endonuclease-digested λ DNA. Aliquots of 375ng were electrophoresed in each lane. B) Illustrations of the expected digested DNA fragments. Overall, the enzymes cleaved λ DNA into expected fragment sizes and are therefore suitable for digestion of genomic DNA in Southern analyses.

Abbreviations: kb = kilo base pairs, bp = base pairs.

Lane 1: Pst I
 Lane 2: EcoRI
 Lane 3: BamHI
 Lane 4: Kpn I
 Lane 5: HindIII-digested λ DNA
 standard ladder

Lane 6: Hind III-digested λ DNA and Hae III-
 digested ϕ X174 DNA standard ladders
 Lane 7: Undigested λ DNA
 Lane 8: Pvu II
 Lane 9: Sal I
 Lane 10: Sma I

An experiment was performed to confirm sufficient activity of restriction enzymes in an environment containing genomic DNA. A 55pg aliquot of pGem 3.0 recombinant plasmid was added to a grasshopper genomic DNA restriction reaction with Kpn I and Sac I restriction endonuclease enzymes. If the Kpn I and Sac I restriction endonucleases are sufficiently active, the pGem 3.0 recombinant plasmid would be cleaved at its Kpn I and Sac I restriction sites (figure 1). This would yield two DNA fragments, one 3.0kbp and the other 2.8kbp. The 2.8kbp fragment represents pGem plasmid DNA while the other contains grasshopper cDNA. The restricted fragments were electrophoresed on an agarose gel and subsequently transferred onto a positive nylon membrane. The membrane was hybridized with 879bp biotinylated grasshopper cDNA probe that is complementary to, and would be expected to bind to, the grasshopper cDNA insert contained in the 2.8kbp restriction fragment. Therefore, a 2.8kbp band would be visible on the X-ray film if both enzymes had cleaved their respective sites. Insufficient restriction of either Kpn I or Sac I would yield a ~6kbp fragment on the X-ray film, representing linearized pGem 3.0 recombinant plasmid. The control for the experiment was Kpn I/Sac I digestion of 15µg grasshopper genomic DNA without the addition of pGem 3.0 plasmid DNA. The actual result, shown in figure 5 (lane 1), gave a low intensity signal at approximately 2.8kbp. No band can be seen in the control lane (lane 2). The overall result confirms Kpn I and Sac I are active under conditions used in genomic DNA digests. The enzymes may be able to also cleave the grasshopper genomic DNA in the restriction reactions.

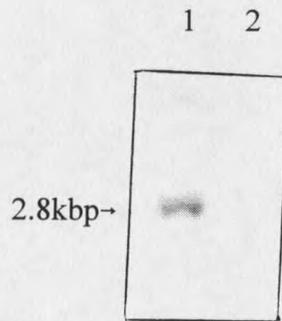


Figure 5. Restriction Endonuclease Activity in Presence of Grasshopper Genomic DNA, Biotin-Based Southern Blot. Lane 1: Kpn I/Sac I restriction enzyme cleavage of a mixture of 30 μ g of grasshopper genomic DNA and 55pg of pGem 3.0 recombinant plasmid containing a C-type lectin grasshopper cDNA insert (figure 1). Lane 2: same reaction without plasmid. Southern hybridization was undertaken with biotinylated 879bp grasshopper C-type lectin cDNA probe. The figure shows a strong signal on X-ray film at 2.8kpb (lane 1). The 2.8kpb band is the expected cleaved pGem 3.0 fragment containing the grasshopper cDNA insert. The size was determined by comparison to biotinylated Hind III-digested λ DNA standard ladder. No band is seen in the control (lane 2). The result confirms that Kpn I and Sac I enzymes cleave plasmid DNA in a plasmid/genomic DNA mixture and may also be expected to cleave the grasshopper genomic DNA in the restriction reaction.

Activities of Enzymes used in Restriction of PCR Products

Restriction analyses were performed on PCR products, as will be described in detail later. It was important to establish the activity of the enzymes as specific cleavage of the PCR products determined the identity of the amplified product. Bgl II, Aat II and Sal I restriction endonucleases were utilized in cleaving the products yielded from PCR amplifications. An experiment was performed to determine that these enzymes were active, the results of which are seen on the agarose gels in figure 6. In this experiment, each enzyme was incubated with 200ng recombinant pGem plasmid containing grasshopper cDNA insert. All the enzymes cleave the plasmid at one site, yielding a linear ~6kbp DNA fragment. The enzymes were shown to be active as they cleaved the plasmid DNA to yield DNA fragments of expected size. Thick bands are produced in Aat I-, and Sal I-, restrictions (lanes 3 and 4) but they can still be distinguished from the uncleaved plasmid DNA in lane 2. This may be due to overloading of the restriction products as aliquots of 200ng were electrophoresed. The experiment confirmed the Bgl II, Aat II and Sal I enzymes were active and may be used in PCR-product restriction experiments.

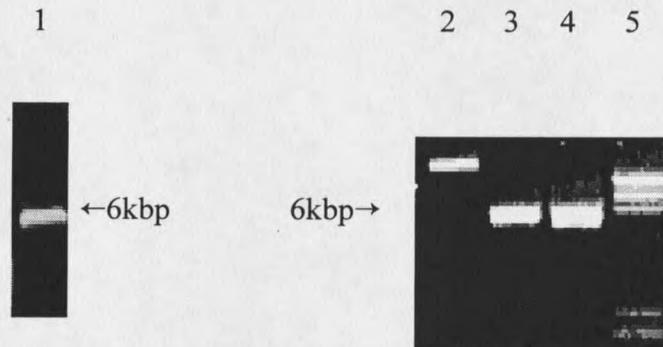


Figure 6. Assays of the Restriction Endonuclease Enzymes Utilized in PCR Product Restriction Analysis. Aliquots of 200ng pGem recombinant plasmid, containing Clone 3 or 4 insert, were cleaved with the enzymes listed below. All the enzymes cleave the plasmid at one site, therefore, the restricted plasmid DNA will migrate on an agarose gel as a single linear ~6kbp fragment. The figure represents ethidium bromide agarose gels of 200ng recombinant pGem plasmid. Lanes 3 and 4 appear overloaded but can be distinguished from the uncleaved plasmid DNA control in lane 2. The enzymes cleave the plasmid DNA and so are active and may be used in restriction of PCR-products.

Lane 1: Bgl II-restricted

Lane 2: Unrestricted recombinant plasmid

Lane 3: Aat II-restricted

Lane 4: Sal I-restricted

Lane 5: Hind III-restricted λ DNA ladder

Southern Analysis Standards and Controls

Genomic DNA Controls

Aliquots of 15 μ g salmon sperm DNA, and 15 μ g barley DNA, were digested under conditions identical to grasshopper genomic cleavage. The salmon sperm and barley were serving as negative controls as they were thought to lack genes homologous to the grasshopper C-type lectin cDNA probes used in Southern analyses. BamHI-, and Pst I-, restricted salmon sperm DNA yielded bands on the Southern X-ray film of figure 9 (lanes 5 and 10, page 47) indicating the presence of DNA homologous to the probe utilized. The hybridization probe was biotinylated 879bp grasshopper C-type lectin cDNA. Binding to the DNA of salmon sperm may be due to the occurrence of C-type lectin genes in this organism. A C-type lectin was found in unfertilized eggs from salmon *Oncorhynchus kisutch* (Yousif et al. 1995). The bands seen in figure 9 may be a gene encoding this lectin protein. Further experiments employed barley DNA as the negative control.

Restricted barley DNA did not produce signals visible on the autoradiograph in figure 10 (lanes 9-10, page 52). The probe utilized in this experiment was radiolabeled 580bp grasshopper C-type lectin cDNA. It appears that no sequences homologous to the grasshopper cDNA probe exist in barley. Unlike salmon sperm, barley is a true genomic DNA negative control and was used in further Southern analyses.

Hybridization Control

EcoRI/Acc I-digested pGem 3.0 recombinant plasmid serves both as negative and positive controls in Southern analyses. An 879bp restriction product represents the grasshopper C-type lectin cDNA insert cleaved out of the pGem 3.0 plasmid (figure 1). This fragment is complementary to the grasshopper cDNA probes utilized in Southern blots and, therefore, serves as a positive control. Negative controls are the 3.8kbp, 640bp, 280bp and 240bp plasmid fragments produced in the EcoRI/Acc I cleavage. The sensitivity and stringency of the Southern analyses will be judged according to the signals produced from the control plasmid fragments on a Southern X-ray film.

A high intensity band appears in lane 11 of the Southern X-ray film (figure 9, page 47) with biotin-labeled 879bp grasshopper cDNA probe. This band is the 879bp positive control DNA fragment. A low intensity band is seen for the 3.8kbp negative control in lane 11. This band contains approximately 130pg of plasmid DNA. It appears the biotinylated 879bp probe is not sufficiently washed off 130pg non-specific DNA in hybridization washes. But, it is insignificant compared to the very intense band seen for the 879bp positive control fragment that represents 29pg of DNA complementary to the probe. Therefore, the restricted pGem control serves both to indicate positive hybridization and potential non-specific binding of the probe.

The pGem control was used in a Southern analysis with radiolabeled 580bp grasshopper cDNA probe. The Southern autoradiograph is seen in figure 10 (page 52). The band in lane 11 is the positive control 879bp grasshopper cDNA fragment from the

EcoRI/Acc I digestion described above. The negative control plasmid DNA fragments are not seen on the autoradiograph. Therefore, the Southern analysis is sufficiently stringent to eliminate non-specific binding to the probe.

Southern Standard DNA Ladder

A DNA ladder was required for radioactive Southern analyses to determine the size of signals produced on autoradiographs. The ladder was generated from restriction of pGem 3.0 recombinant plasmid (figure 1). A selection of restriction endonuclease enzymes produced the following DNA fragments that contained the 879bp grasshopper cDNA insert: 5825bp, 2863bp, 1757bp and 879bp. These fragments all hybridize with the radiolabeled Southern probes 879bp and 580bp grasshopper C-type lectin cDNA and should be visible on autoradiographs. Figure 10 (page 52) is an autoradiograph of a Southern with the radiolabeled 580bp cDNA probe. The ladder, referred to as the 'Southern ladder', is seen in lane 1. All bands in this lane are of high intensity. The ladder is useful in estimating the sizes of the signals on autoradiographs from restricted genomic DNA.

Preparation of 879bp Biotin-Labeled Probe

Cleavage of pGem 3.0 recombinant plasmid with Acc I and EcoRI restriction endonuclease enzymes yielded five separate fragments on agarose gel electrophoresis

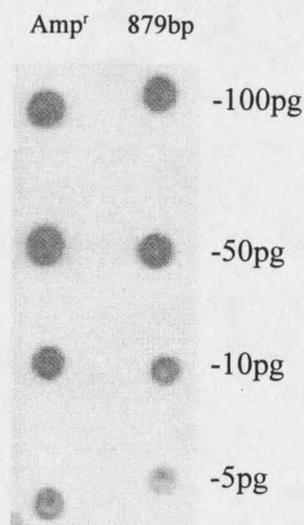


Figure 7. Determination of Biotinylated 879bp Probe Concentration. The figure represents a dot blot on an X-ray film of biotinylated 879bp probe and known concentrations of biotinylated amp^r control probe. A HybondTM-N+ positive nylon membrane was dotted with the biotinylated probes. Streptavidin alkaline phosphatase (SAAP) conjugate bound to the biotin in the probes. SAAP-cleavage of the lumiphore, LumiPhos[®] 530, produced a chemiluminescent signal. The higher the concentrations of biotin incorporated into the probe, the more intense the chemiluminescent signal. Concentration of biotinylated 879bp probe was estimated through comparison with signals from the amp^r probe of known biotin concentration. The biotinylated 879bp probe concentration was estimated as 10ng/ μ l and was used in biotinylated Southern analyses.

(results not shown). One gel band corresponded to the 879bp grasshopper cDNA insert. This band was isolated and purified and its concentration estimated as 2.5ng/ μ l by comparative agarose gel electrophoresis with known λ DNA standards. This was an adequate concentration for subsequent biotin labeling.

Concentration of biotinylated 879bp probe was estimated by the comparisons of chemiluminescent signals on X-ray film. Serial dilutions of 879bp probe were compared with known concentrations of biotinylated amp^r control probe (figure 7): Both probes yielded similar dot blot intensities and must have therefore been of similar concentrations. The biotinylated 879bp probe's concentration was estimated as 10ng/ μ l, a useful concentration range for subsequent Southern analyses. The probe was used in the biotinylated Southern analysis in figure 9 (page 47).

Preparation of 580bp Radiolabeled Probe

The 580bp probe was PCR amplified from pGem 3.0 recombinant plasmid using the primer set 5'B/3'D (figure 3). Concentration of the purified 580bp fragment was estimated, with reference to λ DNA standards, as 5ng/ μ l. A 25ng aliquot was used for probe radiolabeling. Incorporation of [α ³²P]dCTP was followed by spin purification through a Bio-spin 30[®] chromatography column. Non-incorporated radioactive nucleotides are visibly green. The top portion of the column was green in color after centrifugation. It was therefore assumed that all the non-incorporated nucleotides were bound in the column and were not contaminating the probe. The larger fragments of

DNA were collected and their specific activities calculated as approximately 2×10^9 DPM/ μ g. This activity was suitable for Southern hybridizations and used in radiolabeled Southern analyses.

Grasshopper Genomic DNA Preparation

Genomic DNA was required for Southern and PCR analyses to investigate the number and structure of C-type lectin genes in grasshopper. Isolated, precipitated genomic DNA was difficult to resolubilize in TE buffer. Heating and flicking the samples encouraged DNA solubilization. Preparations yielded an intense, high molecular weight band with insignificant smearing at lower molecular weight, upon agarose gel electrophoresis with ethidium bromide as shown in lane 1 of figure 8. This indicated that the DNA had not been extensively sheared during precipitation. Contaminating RNA is likely to be degraded and visualized as a low molecular weight smear on agarose gels. Low molecular weight smearing was not observed on the agarose gel in figure 8 indicating probable lack of RNA contamination. A portion of the DNA sample had not migrated out of the gel well. Low solubility of the genomic DNA or DNA bound-protein contamination may have prevented the DNA from entering the gel. The DNA concentration was calculated from OD_{260} as approximately 700 μ g/ml with the assumption that 1.0 OD is equivalent to 50ng/ml of double stranded DNA (Sambrook et al. 1989). OD_{260}/OD_{280} ratios were above 1.7, showing the DNA was of sufficient purity for subsequent experimentation.

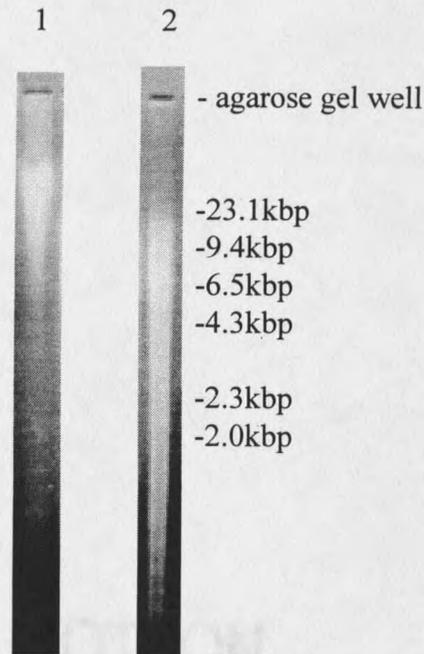


Figure 8. Appearance of Isolated Unrestricted (Lane 1) and Restricted (Lane 2) Grasshopper Genomic DNA on 1% Ethidium Bromide Agarose Gel. Lane 1: 15 μ g grasshopper genomic DNA. Lane 2: 15 μ g grasshopper genomic DNA after EcoRI restriction endonuclease digestion. There is insignificant smearing at lower molecular weight in lane 1. This indicated that the DNA had not been extensively sheared during preparation. The smear in lane 2 indicated that the DNA had been cleaved by the restriction enzyme. The lane 2 digest was subsequently transferred to a positive nylon membrane and hybridized with a grasshopper C-type lectin cDNA probe in a Southern analysis. The sizes shown to the right of the lanes represent a Hind III-digested λ DNA ladder (gel not shown).

Grasshopper Genomic DNA Restriction

All restriction endonuclease enzymes utilized in genomic DNA digests produced smearing upon ethidium bromide agarose gel electrophoresis (results not shown). This indicated the DNA had been extensively cleaved by the restriction enzyme. An example of EcoRI-digested grasshopper genomic DNA is shown in the agarose gel in lane 2 of figure 8. Faint, distinct bands are visible over the background DNA smearing on the original gel. These bands result from restriction site repeats in genomic DNA and are characteristic of the restriction enzyme used (Kroczek 1993). Their appearance confirms an adequate enzymatic digestion of the genomic DNA as well as sufficient separation during gel electrophoresis. A small portion of unmigrated DNA appeared in the wells, perhaps due to DNA-bound protein contaminants or incompletely solubilized genomic DNA.

Determination of Lectin Gene Number

Biotin Southern Analysis

A Southern analysis was performed to gain knowledge of the number of C-type lectin genes in grasshopper. Aliquots of 15 μ g grasshopper genomic DNA were digested with different restriction endonuclease enzymes and subsequently hybridized with biotin-labeled 879bp grasshopper cDNA probe. The resultant Southern membrane was exposed

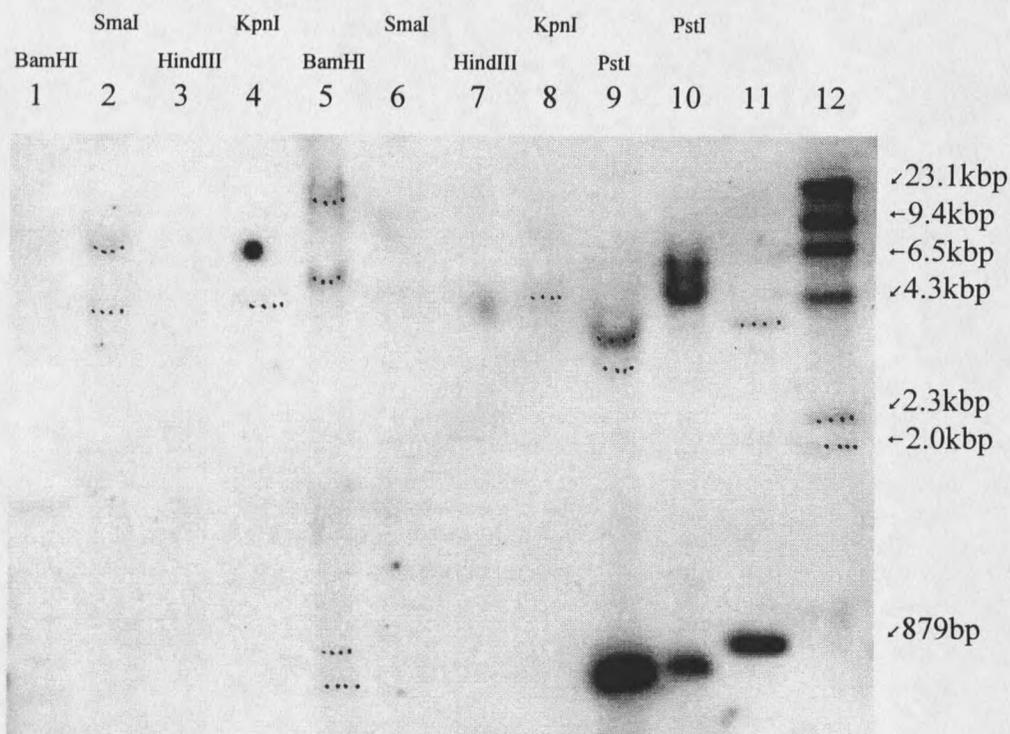


Figure 9. X-ray Film of Southern Analysis on Grasshopper Genomic DNA and Salmon Sperm Control DNA. Aliquots of 15µg genomic DNA were digested with the enzymes shown above the lanes. Southern analysis was performed at 65°C, with biotinylated 879bp grasshopper cDNA probe (figure 3). LumiPhos™ 530 was sprayed onto the membrane. Lumi-Phos™ signal was developed for 2 days after which the membrane was exposed to the film for 3 hours. Light high molecular weight bands are visible in grasshopper genomic digests in lanes 2, 4 and 9. This indicates the presence of lectin genes in the grasshopper genome. Dots have been added where low resolution of the scanned X-ray film does not allow for adequate visualization of faint bands.

Lanes 1-4, 9: 15µg digested grasshopper genomic DNA

Lanes 5-8, 10: 15µg digested salmon sperm control DNA

Lane 11: 200pg EcoRI/Acc I-digested pGem 3.0 control DNA; the fragments are: 3800bp, 640bp, 280bp, 240bp plasmid DNA; 879bp grasshopper cDNA complementary to the 879bp probe

Lane 12: 10ng biotinylated Hind III-digested λ DNA.

to the X-ray film for three hours and is shown in figure 9. The Southern analysis controls (lanes 5-8 and 10-12), are discussed in previous and following sections.

Two faint bands are visible with Sma I-restricted grasshopper DNA in lane 2 at approximately 8kbp and 4kbp. Kpn I restriction (lane 4) produces a very light band at around 4.3kbp. Pst I restriction of grasshopper DNA (lane 9) produces a low intensity signal at ~3.5kbp and two high intensity bands at approximately 4.0kbp and 755bp. The 755bp band was later discovered, by dot blot analysis, to be an unknown contaminant in Pst I buffer, and is to be ignored (results not shown). The Pst I buffer was not used in subsequent experiments.

No bands are visible in BamHI and Hind III restrictions (lanes 1 and 3, figure 9). This is an unexpected result as C-type lectin cDNA has been isolated from grasshopper hemolymph (Stebbins and Hapner 1985), therefore C-type lectin genes exist in the grasshopper genome. An explanation for the lack of bands may have been that the hybridization temperature of 65°C was too 'stringent' and caused the probe not to anneal to homologous sequences.

The X-ray film in figure 9 was exposed to the Southern chemiluminescent membrane for three hours. To increase band intensities on the film, another film was exposed to the Southern chemiluminescent membrane for three days. Background 'noise' increased while the band intensities did not increase substantially (results not shown). Exposure of the film to the Southern chemiluminescent membrane for a few hours appears optimal. The Southern chemiluminescent membrane was resprayed with LumiPhos™ 530 and an X-ray film exposed for 24 hours. The developed X-ray film had

a dark background that made it difficult to differentiate bands. The faint 4.3kbp band visible in Kpn I restriction (lane 4, figure 9) appeared as a more intense band on the X-ray film exposed to the resprayed membrane (results not shown). This confirmed that the light signal in Kpn I restriction in figure 9 is a valid band.

Interpretation of the results from the biotinylated Southern analysis (figure 9) is difficult. Kpn I-restricted grasshopper DNA (lane 4) gives one band, indicating a single C-type lectin gene homologous to the 879bp C-type lectin grasshopper cDNA probe. Pst I and Sma I restrictions (lanes 9 and 2) gave two bands, suggesting the presence of more than one C-type lectin gene. Some bands may have gone unobserved due to the biotinylated probe being unable to produce a strong enough signal to be visible on the X-ray film. Kroczek (1993) claimed that low sensitivity is characteristic with biotin-labeled probes while a Southern analysis with radiolabeled probe is a more sensitive technique. Results with radioactive probes are discussed below.

Radioactive Southern Analysis

Biotinylated probes may be too low in sensitivity to allow detection of genes in Southern analyses. Therefore, radiolabeled probes were used to achieve a more accurate estimate number of lectin genes. A Southern analysis was performed on grasshopper genomic DNA and hybridized with radiolabeled 580bp grasshopper cDNA probe (figure 3). The 580bp probe sequence represents 66% of the 879bp probe utilized in the biotinylated Southern analysis. The radiolabeled probe was expected to bind to the same

target sites as the biotinylated probe. The resulting Southern autoradiograph is shown in Figure 10. It is clear that the radiolabeled probe gives a dramatic increase in multiplicity of bands. Lanes 2 and 3 contain 10 μ g and 15 μ g digested grasshopper DNA, respectively (figure 10). The bands are more intense for the 15 μ g DNA suggesting this amount of DNA is required for the signals to be optimally visible. All digested DNA bands in figure 10 are of lower intensity and resolution than they appear on the actual autoradiograph. The computer scanning program used to copy the figures was unable to produce high resolution pictures. Dashed lines were drawn to represent bands seen on the original autoradiograph. They appear more defined on the actual autoradiograph film.

The probe utilized in the Southern analysis in figure 10 was a portion of a C-type lectin cDNA clone. The probe was expected to anneal to C-type lectin genes of homologous sequence to the probe. Pst I-restricted grasshopper genomic DNA (lane 3) shows five moderately intense bands ranging from ~8kbp to ~2.4kbp, and three smaller very faint bands of approximately 1.6kbp, 1.5kbp and 1.4kbp. BamHI and Sma I (lanes 4 and 5) give four bands, while both Sal I and Pvu II (lanes 7 and 8) show five bands. The EcoRI digest (lane 6) yields the most intense bands of the genomic digests. A reason for the wide range of band intensities shown in the genomic digests may be due to some C-type lectin gene sequences in grasshopper genome having low homology to the 580bp probe. Low homology may cause the probe to bind weakly and be partially washed off in the Southern hybridization washes. Another explanation for the low intensities may be due to some C-type lectin genes containing restriction sites recognized by the restriction enzymes used in figure 10. If these sites are present in regions of 580bp probe annealing

then the C-type lectin gene would be cleaved leaving possibly only short genomic fragments that hybridize to the probe. These short fragments would hybridize to the probe less strongly, thereby allowing the probe to be partially washed off in the hybridization washes with consequent decrease in sensitivity.

The region of Clones 3 and 4 where the probe anneals does not contain restriction sites for enzymes used in the experiment in figure 10. Therefore, if genes 3 and 4 do not contain introns within the regions of probe annealing then the genes will not be fragmented by the enzymes. It has been shown that the region of gene 4 where the 580bp probe anneals is intron free (figure 12, page 56) and so does not contain 'unknown' restriction endonuclease cleavage sites. Gene 3 may also be intronless, but only 37% of the region complementary to the 580bp probe has been proven to be intronless (figure 13, page 59). With these ideas, it can be confirmed that one band from each restriction in figure 10 represents gene 4. It can also be presumed that another band represents gene 3. Additional bands likely correspond to genes or gene fragments additionally present.

Some background smearing appears generally in the Southern ladder (lane 1) and genomic digests (lane 2-8). An explanation for this background is unclear. It may be due to the probe binding to areas of high DNA concentration but this is unlikely as the probe does not bind to 15 μ g unrestricted genomic DNA (results not shown). Also, no DNA is present between the major signals in the Southern ladder in lane 1 yet a background smear is still evident.

It can be strongly suggested, from the Southern analysis with radiolabeled probe in figure 10, that total grasshopper DNA contains multiple C-type lectin genes. The exact

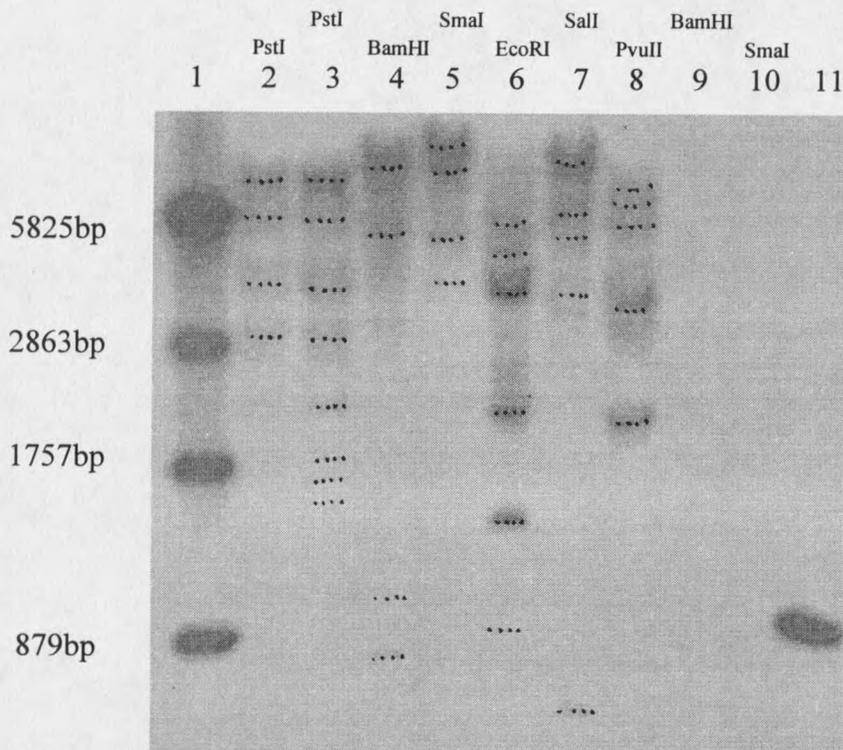


Figure 10. Autoradiograph of Southern Analysis on 15 μ g Grasshopper Genomic DNA and 15 μ g Barley Control DNA. The genomic DNAs were restricted with the enzymes shown above the lanes. Southern hybridization was performed at 65°C with radiolabeled 580bp grasshopper cDNA probe. Multiple high molecular weight bands can be seen for digested genomic DNA in lanes 2-8. This suggests the presence of multiple C-type lectin genes in grasshopper. Dots are added where low resolution of the scanned image does not allow for adequate resolution of bands.

Lane 1: Southern ladder - the DNA fragments contained grasshopper cDNA complementary to the 580bp probe

Lane 2: 10 μ g grasshopper genomic DNA

Lanes 3-8: 15 μ g grasshopper genomic DNA

Lanes 9-10: 15 μ g barley control DNA

Lane 11: 50pg EcoRI/Acc I-digested pGem 3.0 control DNA; the fragments are: 3800bp, 640bp, 280bp, 240bp plasmid DNA; 879bp grasshopper cDNA complementary to the 580bp probe

number cannot be extrapolated from the results, but there appears to be between three and eight C-type lectin genes to which the 580bp probe binds. This result could be expected in view of *Periplaneta americana* cockroach in which a lectin-related protein family exists (Kawasaki et al. 1996). Also, three distinct lectin sequences have been observed in this laboratory, in the form of two cDNA clones and one purified lectin protein (data unpublished).

PCR Optimization

PCR optimization was performed on grasshopper genomic DNA with 3152 and 3'NT primers (figure 3). DNA and Mg^{2+} ion concentrations were varied (table 2) and the PCR results shown in Figure 11. Low DNA concentrations, in conjunction with low Mg^{2+} ion concentrations, produced no visible bands on the agarose gel (lanes 2, 4, 9-10 of figure 11). This suggested either lack of PCR amplification or insufficient amplification for visualization on the gel. Relatively high intensity bands of ~1.6kbp, ~1.3kbp and ~410bp occurred when 585ng of genomic DNA was amplified with 3.5mM Mg^{2+} , shown in lane 7. These conditions were judged to be optimal and were used in subsequent experiments.

The results shown in figure 11 confirm that optimization is critical in genomic PCR amplification. A more detailed description of the bands obtained in the PCR experiment will be discussed later.

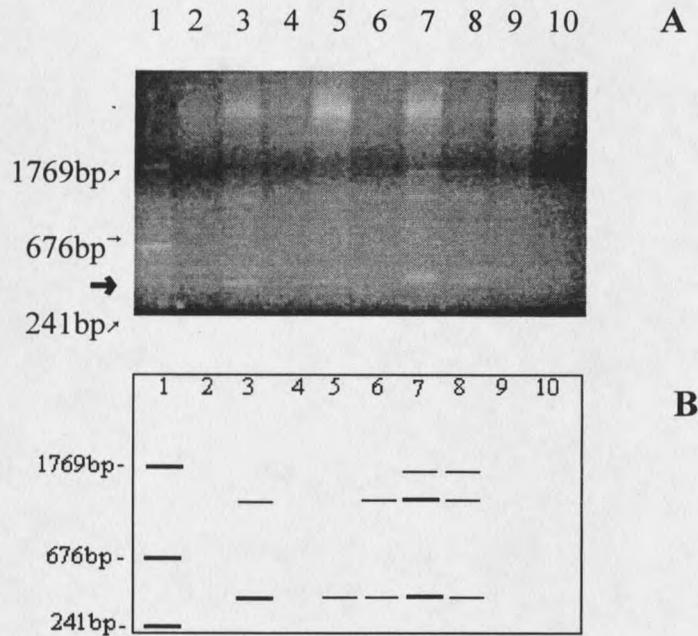


Figure 11. Optimization of PCR with Grasshopper Genomic DNA Template using 3152 and 3'NT Primers (Figure 3). DNA and Mg^{2+} concentrations were varied as shown in Table 2. Lane 1 is Ras I-digested PUC DNA standard ladder A) Ethidium bromide agarose gel of the PCR-amplified products. Large arrow indicates amplified ~ 410bp fragment. B) Illustration of the results in A. Lane 7 gave the highest yield of PCR products and, therefore, is optimal for amplification from genomic DNA. DNA and Mg^{2+} amounts of 585ng and 3.5mM, respectively, were used in subsequent PCR experiments.

Table 2. DNA and Mg^{2+} Concentrations used to Obtain the PCR Results Shown in Figure 11. Explanation of the results are given in figure 11.

| Gel Lane | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Grasshopper DNA (ng) | 390 | 390 | 390 | 665 | 115 | 585 | 195 | 585 | 195 |
| Mg^{2+} conc. (mM) | 2.5 | 3.9 | 1.1 | 2.5 | 2.5 | 3.5 | 3.5 | 1.5 | 1.5 |
| Result: number of bands on gel | — | 2 | — | 1 | 2 | 3 | 3 | — | — |

Determination of Intron Nature of Lectin Genes

Restriction Analysis of 4052/3'NT PCR Products

Knowledge of the structure of lectin CRD-coding regions within the gene gives insight into the lectin protein's evolution and relationship with lectins from other organisms (Drickamer 1993). PCR amplification was performed on grasshopper genomic DNA template to determine the intronic character of gene 4. Primers 3'NT and 4052 were utilized (figure 3). Primer 3'NT binds to both Clone 3 and 4, at a position 58 nucleotides downstream from the translation termination codon. Primer 4052 anneals to Clone 4, but not Clone 3, near the 5' end of the ORF and would amplify a 885bp fragment of Clone 4 when paired with 3'NT. Lack of cross-binding to Clone 3 was proved by using 4052 and a 3' terminal Clone 3 primer to amplify pGem 3.0 recombinant plasmid template. No amplification occurred (results not shown) confirming that primer 4052 was unable to bind, or amplify, Clone 3. Therefore, genomic PCR with 4052/3'NT primers would not amplify gene 3, but would be expected to amplify a region of gene 4.

PCR analysis of grasshopper genomic DNA with 3'NT and 4052 primers yielded a ~870bp band visible after polyacrylamide gel electrophoresis as shown in lane 1 of figure 12A. This ~870bp band was thought to represent an amplification product of gene 4. An 885bp fragment would be produced if gene 4 were intronless between the 3'NT/4052 primers, as calculated from the known cDNA sequence (figure 2).

To confirm that the ~870bp band in lane 1 was indeed amplified from gene 4, the

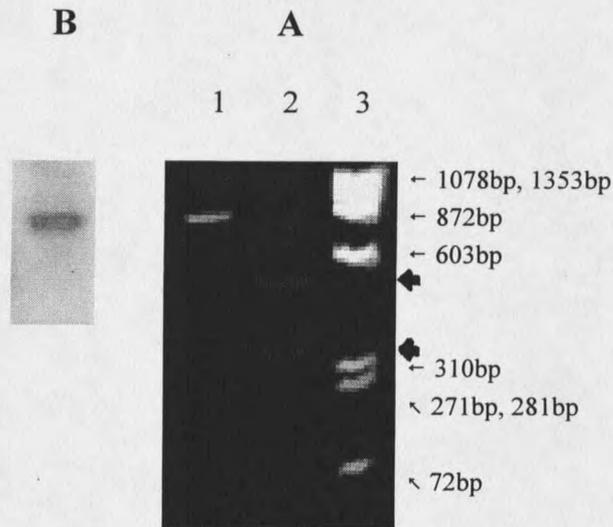


Figure 12. PCR Amplification of Grasshopper Genomic DNA Template using Primers 4052 and 3'NT (Figure 3). A) Polyacrylamide gel of the PCR product (lane 1) and after restriction with Sal I restriction endonuclease (lane 2). Arrows indicate the restriction band fragments. Lane 3 is Hind III-digested λ DNA standard ladder. B) Southern autoradiograph band, hybridized with radiolabeled 580bp probe, representing the unrestricted PCR product in lane 1 of A. The results suggest gene 4 is the amplification product.

PCR product was cleaved with Sal I restriction endonuclease. No Sal I restriction sites are in Clone 3, while there is a single cleavage site in Clone 4. Sal I restriction, of the 885bp Clone 4 cDNA sequence between 3'NT/4052 primers, produces two fragments of 532bp and 353bp. These bands are visible after Sal I cleavage of the genomic PCR product from 3'NT/4052 primers (lane 2, figure 12A). This result strongly suggests that gene 4 is the amplified product of the 3'NT/4052 PCR reaction and confirms it is not Clone 3.

The PCR results and restriction analysis (figure 12A) verify that 85% of the gene 4 ORF is continuous and lacking intron sequences. This includes the two CRD-coding regions (figure 3).

Southern Analysis of 4052/3'NT PCR Products

As previously described, restriction analysis of the genomic PCR product from 4052/3'NT primers, indicated gene 4 was the amplified product (figure 12A). A Southern analysis was performed to confirm that the PCR product was amplified from a C-type lectin gene. PCR amplification products from grasshopper genomic DNA with 4052 and 3'NT as the primer pair were electrophoresed on an ethidium bromide agarose gel. The band was transferred to a positive-nylon membrane and a subsequent Southern analysis was undertaken with radiolabeled 580bp grasshopper C-type lectin cDNA probe. The result, seen in figure 12B, shows an intense band at around 870bp. An 885bp fragment is produced when the 4052 and 3'NT primers anneal to grasshopper cDNA (figure 2). The

~870bp product in figure 12B is of the expected size and is consistent with a lack of introns in gene 4 between the 4052/3'NT primers. This region includes both CRDs.

The restriction analysis of the 4052/3'NT PCR product described previously (figure 12A) strongly suggests that gene 4 is the amplification product from 4052/3'NT primers. This conclusion is supported by the Southern analysis in figure 12B. Both experiments confirm a lack of introns in gene 4 between the 4052/3'NT primers. This distance represents 85% of the entire ORF. Gene 4 has two CRD-coding domains that are amplified from 4052/3'NT primers (figure 3). These domains are intronless and may have evolutionary relationships with other intronless C-type lectin proteins, as will be described later.

Restriction Analysis of 3152/3'NT PCR Products

PCR amplification was utilized to determine the gene makeup of a portion of gene 3. Knowledge of the intronic character of the CRD-coding region of gene 3 will allow classification of the C-type lectin protein encoded by gene 3. The classification scheme will be described later. Primers 3152 and 3'NT (figure 3) were used to PCR amplify grasshopper genomic DNA. Primers 3152 and 3'NT are complementary to Clone 3 and 4 (figure 2). Therefore, it was hypothesized that they amplify both gene 3 and 4 in a PCR reaction with grasshopper genomic DNA as template. The genomic PCR amplification from 3152/3'NT primers gave three bands after agarose gel electrophoresis as seen in lane 7 of figure 11. The ~1300bp and ~1600bp bands were determined to be non-specific

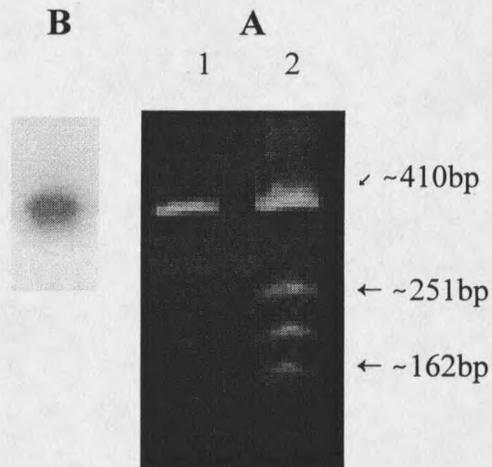


Figure 13. PCR Amplification of Grasshopper Genomic DNA using Primers 3152 and 3'NT (Figure 3). A) Polyacrylamide gel of the PCR product (lane 1) and after restriction with Aat II restriction endonuclease (lane 2). Arrows indicate the approximate size of the PCR product and the fragments produced after cleavage. Sizes of bands were estimated from comparison to Hae III digested ϕ X174 DNA (not shown) B) Southern autoradiograph band, hybridized with radiolabeled 580bp probe, representing the unrestricted PCR product in lane 1 of A. Interpretation of the results presented in the figure suggests gene 3 is the amplification product and is intronless between the primers 3152 and 3'NT.

PCR products, as described later, and so will be ignored. A band at around 200bp is visible in lane 2 (figure 13) and can also be seen as a weak band in lane 1. This was Southern negative, therefore was a PCR artifact and was not further considered. If gene 3 and 4 are intronless in the 3' portion of their ORFs then products from the amplification would be 413bp and 411bp, respectively, as calculated from cDNA sequences (figure 2). The band in lane 7 of figure 11 is in this region, although a 2bp difference could not be resolved. This result suggests gene 3 and/or gene 4 have been amplified and are of a size consistent with a lack of introns between the 3152/3'NT primers.

The PCR-amplified product from 3152/3'NT primers and genomic DNA template was restricted with Aat II restriction endonuclease. Aat II does not cleave Clone 4 but is known to restrict Clone 3, within a region that the two primers amplify, to produce 251bp and 162bp fragments. Lane 2 of figure 13A shows the cleaved DNA fragments of the Aat II restriction of genomic PCR products using 3152/3'NT primers. It appears that the PCR product is a portion of gene 3 as Aat II is known to cleave the cDNA sequence encoded by gene 3. The bands seen in lane 2 are ~251bp and ~162bp in length indicating gene 3 was the amplification product. The overall results in figure 13 suggest that gene 3 is the amplification product and does not contain introns in the 3' end of its ORF between primers 3152 and 3'NT.

The ~410bp band that is present in lane 1 of figure 13A is visible after Aat II restriction (lane 2). This fragment is probably uncleaved, amplified gene 4 because 3152 and 3'NT primers are also complementary to Clone 4 and are likely to amplify gene 4 in addition to gene 3. Assuming the uncleaved fragment is gene 4 then the result would

indicate that gene 4 is intronless in the region between 3152/3'NT primers. This interpretation is consistent with data from 4052/3'NT amplification of gene 4 (figure 12, page 56).

Southern Analysis of 3152/3'NT PCR Products

PCR amplification with 3152 and 3'NT primers on grasshopper genomic DNA yielded three bands of ~1600bp, ~1300bp and ~410bp on an ethidium bromide agarose gel (lane 7, figure 11). The bands were transferred by capillary action, and immobilized, onto a HybondTM-N+ nylon membrane. The immobilized DNA was hybridized with radiolabeled 580bp cDNA probe. The ~1300bp and ~1600bp PCR products did not produce signals on the autoradiograph (result not shown). This result suggests that these two bands are not homologous to the grasshopper C-type lectin cDNA probe and are, therefore, products of non-specific annealing and amplification. A portion of the Southern autoradiograph is shown in figure 13B. The signal represents the ~410bp band seen in lane 1 of figure 13A. The 580bp probe is homologous to C-type lectin-coding sequences and so is expected to bind to products derived from C-type lectin genes. The Southern analysis confirmed a C-type lectin gene was amplified in the PCR reaction. The size of the band is ~410bp, similar to the 413bp distance between the 3152 and 3'NT primers when they are represented on a grasshopper clone map (figure 2). The size of the band on the autoradiograph in figure 13B strongly suggests a lack of introns in the portion of gene 3 that is amplified by the 3152/3'NT primers. Specific restriction analysis

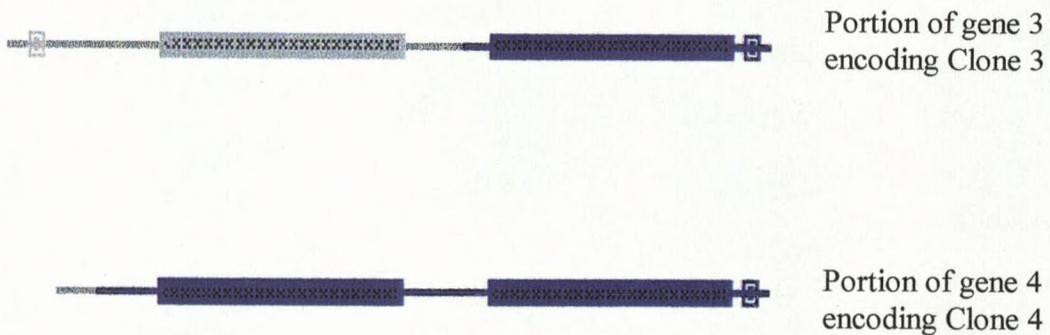


Figure 14. Illustration of Intronless Nature of Genes Encoding Grasshopper Clones 3 and 4 cDNA. Sequences run 5' to 3'. Squares represent start or stop translation codons. Stippled boxes represent CRD-coding regions. Blue areas show the portions of the genes that are intron-free. The areas of Clones 3 and 4 that have not had their gene intronic character determined are shown in grey. In summary, both CRDs in Clone 4 and the C-terminal CRD in Clone 3 have been shown to be intron-free.

in figure 13A indicates the ~410bp band from the 3152/3'NT PCR reaction is a gene 3 product. Therefore, it appears there are no introns between the 3152/3'NT primers that cover the 3' end of gene 3 including the carboxyl CRD-encoding region (figure 3).

In summary, it appears gene 4 is intronless, including both its CRD-coding domains. The carboxyl CRD-coding region of gene 3 has been shown to lack introns. The amplification of the 5' CRD-coding region of gene 3 was unsuccessful, probably the result of inadequate primers. An illustration of the intronless nature of genes 3 and 4 is shown in figure 14. Continuous CRD-coding regions in genes 3 and 4 may indicate possible evolutionary relationship to intronless C-type lectin genes from other organisms. This relationship will be discussed in the next section.

DISCUSSION

The main objective of this thesis has been achieved in that the presence of multiple lectin genes in the grasshopper has been documented. In addition, two genes, corresponding to cDNA Clones 3 and 4, have been shown to be without introns in the CRD domains. Completion of this project required use of procedures in the field of molecular biology, some of which required modification to obtain reproducible data. The techniques included endonuclease restriction, Southern analysis and PCR amplification.

Optimization of Experimental Methodology

Southern Analysis

Biotin- versus Radio-Labeled Probes. Biotinylated and radiolabeled grasshopper C-type lectin cDNA probes were utilized in Southern analyses of grasshopper genomic DNA. The autoradiograph (figure 10) contains signals from genomic digests that do not appear on the biotin Southern (figure 9). A reason for these extra bands from the radiolabeled Southern analysis may be that the autoradiograph is the result of a more sensitive technique. This idea comes with the assumption that the probe does not bind non-specifically to areas of high DNA concentration. This is a valid assumption for three

reasons. Firstly, the probe was shown not to bind to a 15 μ g band of non-restricted grasshopper DNA (results not shown). Also, the radiolabeled probe does not anneal to 25pg of 3.8kbp negative control DNA in lane 11. Finally, the hybridization temperature and the post-hybridization washes performed on the Southern membrane were stringent. High stringency included hybridizing at 65°C; while washing involved low salt concentrations of 1X SSC with added SDS detergent. KroczeK (1993) claims that lower sensitivity is obtained with non-radioactive labeling methods and low sensitivity does not easily allow a routine detection of single copy genes on Southern blots. C-type lectin genes in grasshopper may be difficult to detect in a Southern analysis with a biotin-labeled probe.

The advantages of biotin Southern analyses are they utilize probes that can be stored for prolonged periods and are not subject to radiation-related degradation. Chemiluminescent detection is safe and the biotin system, unlike systems based on a color reaction that include BCIP and NBT substrates (Kerkhof 1992), is readily detected using standard X-ray film to produce a non-fading, permanent experimental record. The weakness of the biotin system includes higher background problems, a longer experimental procedure and, as explained previously, apparent lower sensitivity than radioactive Southern analyses.

The random-primed labeling used for production of Southern probes was based on methodology developed by Feinberg and Vogelstein (1983). An alternative labeling procedure, known as nick-labeling, involves nicking one strand of double-stranded DNA and replacing the nucleotides downstream from the nick with radioactive nucleotides by

means of DNA polymerase (Sambrook et al. 1989). Random-primed labeling has advantages over nick-labeling as the former produces probes with higher specific activity due to both the input DNA not being degraded during the reaction and label being incorporated equally along the entire length of the input DNA. However, the resultant random-primed probe is statistically shorter than nick-labeled probes (Feinberg and Vogelstein 1983).

C-Type Lectins in Salmon Sperm DNA. Restricted salmon sperm DNA yields bands on the Southern X-ray film in figure 9. This suggests that salmon sperm contains lectin genes homologous to the grasshopper cDNA probe. In fact, proteins with homologous sequences to lectins have been reported in many fish species including sea raven, *Hemitripterus americanus* (Ng and Hew 1992), smelt, *Osmerus mordax* (Ewart et al. 1992) and in the ova of coho salmon, *Oncorhynchus kisutch* (Yousif 1994). It has been shown that a coho salmon C-type lectin binds to specific bacterial cells and may have a function in the defense system of the fish. Interestingly, this is a role suggested for the grasshopper hemagglutinin. Sequence alignments have shown that the proteins in sea raven and smelt have C-type lectin CRDs but sea raven has lost its Ca^{2+} binding capacity while smelt has retained just one Ca^{2+} binding site. These proteins are fish antifreeze proteins (AFPs) and their CRDs may have the ability to bind to an ice crystal lattice (Ewart et al. 1992). C-type lectin genes in salmon sperm may be of sufficient homology to bind the grasshopper lectin probe. No bands were observed for digested barley DNA on Southern autoradiographs, as shown in lanes 9-10 of figure 10. The overall conclusion

is that salmon sperm DNA contains C-type lectin-like sequences while barley DNA does not. Salmon sperm was discontinued as the negative control in Southern analyses and was replaced by barley DNA.

Optimization of PCR

Some of the encountered problems with PCR amplification included: no detectable product or a low yield of the desired product, the presence of non-specific background bands due to mispriming or misextension of the primers, and formation of primer dimers. Optimal conditions were established for PCR amplification.

Deoxynucleotide triphosphate concentrations, primer concentrations and amplification cycle number used were within ranges suggested by Innis and Gelfand (1996) and shown to be adequate for PCR amplification carried out by L. Gedik and J.R. Radke in this lab (unpublished work). Innis and Gelfand (1990) claim that the most likely cause for failure of a PCR is incomplete denaturation of the target template. Initial PCR reactions performed on grasshopper genomic DNA amplified products with Taq DNA polymerase (Life Technologies) gave no amplified products (results not shown). In one approach, the genomic DNA template fragments were decreased in length to enable the template to be more efficiently denatured. Fragmentation of the template DNA included cleaving the genomic DNA with restriction endonucleases. The enzymes were Not I, that has an 8bp recognition sequence, or BamHI that recognizes a 6bp sequence. Statistically, BamHI cuts the genomic DNA into smaller fragments than does NotI due to

BamHI's shorter recognition sequence. Other experiments involved shearing the genomic DNA by either sonication, or vortexing, for 90 seconds. [$\alpha^{35}\text{S}$]dATP was added to the PCR reaction mixture and, following PCR thermocycling, polyacrylamide electrophoresis, and gel drying the PCR products were visualized on an autoradiograph. The autoradiograph showed no bands (results not shown). Subsequently, Taq DNA polymerase was replaced with AmpliTaq GoldTM DNA polymerase (Roche Molecular Systems Inc.). Amplification fragments were produced when AmpliTaq GoldTM DNA polymerase was used on genomic DNA template. It is probable that the initial 10 minutes at 94°C required to activate the enzyme is also beneficial in adequate denaturation of the template and therefore promotes subsequent extension and amplification. The use of AmpliTaq GoldTM was the key to resolving the genomic PCR portion of the work.

A relatively long primer extension time of 2 minutes was used in the PCR reactions. This length of time was chosen to allow complete extension of targeted genes containing intronic DNA. Primer annealing temperatures are usually 5°C below the Tds of the amplification primers (Innis and Gelfand 1990). A primer set should have a Td difference of 5°C or less and the longer the amplification product, the closer the Tds. Amplification could not be obtained from grasshopper genomic DNA using the 5'NT primer in conjunction with either 3'NT or 3132 primers (results not shown). The primer pairs had Td differences over 8°C (table 1). It was concluded that this temperature difference was too dissimilar for amplification to occur. Alternative primers, 3052 and 3053 (figure 3), were subsequently designed to anneal to the 5' region of Clone 3 and had

Tds more compatible with primers 3'NT and 3132. Although the 3052 and 3053 primers amplified from pGem 3.0 plasmid, they were unsuccessful in amplification of gene 3 (results not shown).

Genomic DNA and Mg^{2+} conditions chosen were 585ng and 3.5mM, respectively. These were shown to be optimal in the PCR optimization experiment shown in figure 11. The optimal DNA concentration was within the 50ng to 1 μ g range typically used for single copy loci (Saiki 1990). The relatively high Mg^{2+} concentration of 3.5mM produces relatively high PCR yields but also increases non-specific products.

Grasshopper Lectin Gene Number

An aim of this research was to determine the number of C-type lectin genes in grasshopper. The presence of at least three C-type lectin genes was implied from lectin cDNA and protein data, available in the laboratory. Two grasshopper C-type lectin cDNA clones, Clones 3 and 4 (figure 2), have been isolated and sequenced and are 80% homologous (Radke J.R. Unpublished results). A cyanogen bromide-cleaved fragment of isolated grasshopper C-type lectin hemagglutinin protein (GHA) has a different sequence from those encoded by Clones 3 and 4 (Hapner K.D. Unpublished results). Assuming the clones and the isolated protein are all encoded by separate genes then it appears at least three C-type lectin genes exist in the grasshopper's genome.

Southern analyses were performed on grasshopper genomic DNA to confirm the presence of multiple C-type lectin genes in grasshopper. A resultant Southern X-ray film

is shown in figure 9. The probe utilized was biotinylated 879bp C-type lectin cDNA probe (figure 3). The results on the X-ray film in figure 9 are difficult to interpret. BamHI and Hind III digests (lanes 1 and 3, respectively) of grasshopper genomic DNA give no bands suggesting there are no genes present in grasshopper that are homologous to the 879bp probe. This result is unlikely as the 897bp probe was generated from a lectin clone isolated from grasshopper. The grasshopper is therefore expected to contain a gene coding the clone sequence. Sma I digestion of grasshopper genomic DNA (lane 2) gives two bands at approximately 8kbp and 4kbp, while Pst I restriction (lane 9) produces signals at approximately 4.0kbp, 3.5kbp and 755bp. The latter band is a contaminant so can be ignored. Restriction of grasshopper DNA with Kpn I (lane 4) produces one signal at ~4.3kbp. Sma I, Pst I and Kpn I digests indicate the presence of one or more C-type lectin genes in grasshopper. The suggestion of the existence of one grasshopper lectin gene was proven to be incorrect after interpretation of Southern analyses with radiolabeled C-type lectin probe.

The biotylated probe, used in figure 9, was replaced with a radiolabeled probe in order to further investigate lectin gene number in grasshopper. Southern signals obtained from digested grasshopper genomic DNA in figure 9 are of low intensity. Southern analyses with radiolabeled probes have been shown to produce higher intensity signals than probes modified for chemiluminescent detection (Kroczeck 1993). The signals produced in the autoradiograph (figure 10) are higher in intensity than the chemiluminescent signals seen in the Southern X-ray film (figure 9), indicating that radiolabeled probes produce higher sensitivity blots. But, a few bands in the

autoradiograph (figure 10) are still very faint. Reasons for these low intensities include the following. First, a C-type lectin target gene may be endonuclease restricted within the region of probe binding. This would produce two fragments, both unable to completely bind the probe and therefore the probe is more likely to be washed off in stringent hybridization washes. Second, the low intensity bands may represent C-type lectin genes that are of low homology to the probe causing the probe to detach in the hybridization washes. Finally, not all the target DNA migrated through the gel, as shown by ethidium bromide fluorescence in gel wells after electrophoresis (lane 2, figure 8). The reason for non-migration of the DNA was thought to be low solubility of the genomic DNA or DNA-bound protein contaminants.

Genomic Southern analyses were undertaken with radiolabeled probe as this technique appears more sensitive than chemiluminescent detection and would possibly give a more accurate determination of lectin gene number in grasshopper. Results, with radiolabeled 580bp lectin cDNA probe, are shown on the autoradiograph in figure 10. An initial observation is that more bands appear in grasshopper genomic digests, some of which are of higher intensity, than when Southern are hybridized with radiolabeled probes than when biotinylated probes are utilized (figure 9). When examined in more detail it appears Pst I digestion of grasshopper genomic DNA (lane 3) gives eight bands, three of which are ~1700bp and are very faint. Sal I and Pvu II (lanes 7 and 8, respectively) show five bands while BamHI and Sma I (lanes 4 and 5, respectively) give four bands. The presence of multiple bands with grasshopper genomic DNA strongly suggests the existence of multiple lectin genes in grasshopper. This conclusion is

acceptable in view of *Periplaneta americana* cockroach that contains a lectin-related protein family (Kawasaki et al. 1996). Lectins are proposed to function in invertebrate immune defense (Drickamer 1993). Multiple C-type lectin recognition molecules may have evolved to regulate the grasshopper's response to infection.

The precise number and size of the bands produced for digested grasshopper genomic DNA in the Southern autoradiograph (figure 10) gives inexact indication of the number of lectin genes present. Without prior knowledge of sequences and intronic character of all lectins in grasshopper, bands shown on the autoradiograph cannot be assigned to a particular gene. For instance, certain lectin gene sequences may contain restriction sites, both in the coding region and possible intronic regions, that are recognized and cleaved by the enzymes used in the Southern digestion. If the genes are cleaved in the region where the Southern probe binds then two bands may appear on the autoradiograph representing the single gene. The restriction enzymes used in the Southern analyses in figures 9 and 10 do not cut the coding sequences of genes 4 or 3 within areas of probe binding. Gene 4 has been shown (figure 12) to contain no introns in areas where the Southern probes hybridize. Therefore, gene 4 does not have intronic DNA that may contain 'unknown' endonuclease restriction sites and so is not fragmented by endonuclease restriction. Gene 4 should be represented by one band in the Southern autoradiograph (figure 10).

Recently, Kawasaki et al. (1996) subjected a cDNA library of cockroach, *Periplaneta americana*, fat body to PCR amplification. Eight degenerate primers were used for PCR amplifications. The primers corresponded to partial amino acid sequences

of *Periplaneta* lectin. Analysis revealed many similar, but not identical, *Periplaneta* lectin-related cDNAs. Some *Periplaneta* lectin-like cDNAs were cloned, followed by deduction of the amino acid sequences of proteins encoded by these cDNAs. The sequences revealed that the proteins constitute a discrete family. This result is the first demonstration of the presence of a lectin-related protein family in an insect. Multiple lectin-related proteins in the cockroach implies its genome contains multiple lectin genes. The research in this thesis has indicated that multiple lectin genes exist in the grasshopper, *Melanoplus differentialis*, genome. The grasshopper may be another example of an insect containing a family of lectin-related proteins.

Genomic Southern analysis was performed on *Sarcophaga peregrina* (Takahashi et al. 1985). The probe used was a 780bp fragment from the coding region of *Sarcophaga* lectin cDNA. The Southern autoradiograph showed a single band with two different restriction endonuclease enzymes. Therefore, it is likely that *Sarcophaga peregrina* has a single *Sarcophaga* lectin gene.

Recently, a C-type lectin has been discovered in *Drosophila melanogaster* (Haq et al. 1996). A Southern analysis was performed on *D. melanogaster* total DNA and hybridized with ³²P-labeled *Drosophila* lectin cDNA probe. Digests of *Drosophila* genomic DNA gave single bands on an autoradiograph irrespective of the restriction enzyme used. The Southern result indicated that *Drosophila melanogaster* contains a single C-type lectin gene.

In summary, a limited amount of research has focused on C-type lectin gene number in insects. A single C-type lectin gene is present in flies *Sarcophaga peregrina*

and *Drosophila melanogaster*. There appear to be multiple genes encoding C-type lectins in cockroach *Periplaneta americana* and, from this work, grasshopper *Melanoplus differentialis*.

Intronic Nature of Lectin Genes

Knowledge of the intronic character of the CRD-coding region of a C-type lectin can enable the lectin protein to be classified into a specific lectin group and indicate the protein's possible evolutionary path (Bezouska et al. 1991). PCR was used to amplify regions of genes coding for grasshopper C-type lectins. Examination of the length of the amplified gene fragment gives insight into the intronic makeup of the gene. Specific primers were used to amplify portions of gene 3 and 4, the genes that encode Clones 3 and 4, respectively. The size of the amplified products were compared with the distance between the primers when annealed to Clone 3 and 4 cDNA sequences (figure 2). Genomic amplification products longer than corresponding products from grasshopper cDNA template would suggest the presence of intronic DNA between the two primers.

Primers 4052/3'NT PCR-amplified a large portion of gene 4 (figure 12A, lane 1). The 4052/3'NT PCR product was successfully cut with Sal I confirming the presence of a Sal I restriction site, uniquely to Clone 4. This result indicated gene 4 was the amplification product (figure 12A, lane 2). Southern analysis of the 4052/3'NT PCR product, hybridized with 580bp grasshopper C-type lectin probe, verified that a C-type lectin had been amplified (figure 12B). The PCR product, ~870bp, approximated the

expected 885bp fragment from 4052/3'NT amplification of Clone 4 cDNA. These restriction and Southern analyses of the PCR product strongly suggests that no introns occur between 4052 and 3'NT primers in lectin gene 4 in the grasshopper genome. The amplification product represents 85% of the ORF of gene 4. The entire coding region of gene 4 may be uninterrupted as are 17% of all known insect genes (Lewin 1994).

The 3' end of the ORF of gene 3, that contained a CRD-coding region (figure 3), was also shown to lack introns. Primers 3152 and 3'NT (figure 3) PCR-amplified a DNA fragment from grasshopper genomic DNA template (figure 13A). This product was cleaved with a restriction enzyme known to have a restriction site in Clone 3 (figure 13A, lane 2). The restriction enzyme cleaved the PCR product, indicating that gene 3 may be the amplification product seen in lane 1 of figure 13A. The 3152/3'NT PCR product gave a signal in a Southern analysis when hybridized with 580bp grasshopper C-type lectin cDNA probe (figure 13B). The Southern blot indicated that a C-type lectin sequence had been amplified in the 3152/3'NT genomic PCR reaction. The PCR product ~410bp is a size approximating the 413bp fragment expected from 3152/3'NT amplification of Clone 3 cDNA (figure 2). Therefore, restriction and Southern analyses confirmed that gene 3 was the amplification product and its size indicated the lack of introns between the 3152/3'NT primers. The amplified region constitutes 37% of the ORF of gene 3 and includes the carboxyl CRD-coding region.

Attempts at amplifying the 5' end of gene 3, using either 3052 or 3053 primer in conjunction with 3'NT (figure 3), were unsuccessful. These primer combinations produced multiple bands on agarose gels but no signals appeared on Southern blots with

radiolabeled 580bp C-type lectin cDNA probe. The Southern blot suggested that authentic C-type lectin genes had not been amplified (results not shown). The primers may be homologous to non-C-type lectin genes and may have amplified sequences unrelated to the probe. The 3052 primer may have been a particularly non-specific primer as it bound to the initiating ATG codon and the putative signal sequence that is similar in many secreted proteins. L. Gedik, in this laboratory, also found 3052 to anneal to, and amplify, unwanted sequences in RT-PCR work (unpublished work). Why the 3052 primer did not amplify any gene 3 DNA is unexplained.

One can not speculate with certainty, on the gene structure of the 5' end of the gene 3 coding region as it need not be identical to gene 4. An example of differences in the gene organization of two similar proteins is seen in the invertebrate acorn barnacle, *Megabalanus rosa* (Takamatsu et al. 1993 and Takamatsu et al. 1994). One of the barnacle's proteins, BRA-2, is a C-type lectin encoded by a gene that is entirely lacking introns. The other barnacle protein, BRA-3, has its CRD-coding region interrupted by three exons. In the case of the grasshopper genes, there is a duplicate domain within a single polypeptide chain, not two proteins encoded by two genes. One can speculate that the 3' CRD-coding region of gene 3 is intronless, like its carboxyl CRD-coding region, since they are the same in gene 4.

Lectin Classification and Evolution

C-type lectins are classified into groups I-VII depending on the functional

domains they have in addition to their CRD regions. It is unclear how to classify the C-type lectin encoded by gene 4. The CRD-coding regions of gene 4 are intronless, therefore should fall into group III or IV C-type lectins that also lack introns in their CRD-coding regions. These groups include pulmonary surfactant apoprotein, bovine conglutinin and rat MBP (Arason 1996). Groups III and IV have additional functional domains associated with them. Collectins, group III, have collagenous domains (Hoppe and Reid 1994) while selectins, group IV, consist of an epidermal growth factor-like domain (Drickamer 1993). Unlike groups III and IV, the proteins encoded by gene 3 and 4 have no known additional functional domains although a significant fraction of the polypeptide chain, 70 out of 304 amino acids, is situated amino terminal to the two CRDs. A lack of additional functional domains is a characteristic of group VII C-type lectins. But genes 4 and 3 do not neatly fit into group VII because the proteins encoded by genes 3 and 4 contain an additional CRD-coding domain. If one views the grasshopper genes as having a single duplicated domain perhaps they could be categorized as group VII. Alternatively, like the macrophage mannose receptor in group VI (Drickamer 1993), the grasshopper lectins may need to be placed in a novel C-type lectin group as they lack precise characteristics required for classification into groups I-VII.

Evolution of C-type CRDs is thought to have involved divergence of intron-containing and intron-lacking CRDs, followed by shuffling events that associated CRDs with other functional domains (Bezouska et al. 1991). The protein encoded by gene 4 may have formed from intron-lacking progenitors. The same evolutionary steps can be

suggested for gene 3, although the intron character of the N-terminal CRD-coding region is unknown. Known C-type lectin proteins from invertebrates that are without introns in their CRD-coding regions are acorn barnacle lectin BRA-2 (Takamatsu 1994) and, present data, the protein encoded by gene 4 from grasshopper. These lectins, as well as the vertebrate collectins and selectins, may have originated from a common progenitor protein and, therefore, have an evolutionary relationship.

Newly Discovered Clone 4 Sequence

A cDNA fragment representing the 5' region of gene 4 has very recently been cloned and sequenced (Radke J.R. Unpublished results). The coding region (ORF) of Clone 4 is 978bp, in comparison to 972bp for Clone 3, corresponding to 326 and 324 amino acids, respectively.

The 5' terminal sequence of Clone 4 contains an EcoRI recognition site (figure 15) and there is also an EcoRI site further downstream. Consequently, EcoRI digestion of Clone 4 cDNA produces a 767bp fragment. Gene 4 is known to be intron free from primer 4052 to 3'NT (figure 3). Assuming that gene 4 does not contain introns 78bp upstream from the 4052 primer, then EcoRI can restrict gene 4 to produce a fragment 767bp in length. This is important in light of the EcoRI digestion of grasshopper genomic DNA in the Southern analysis in figure 10. The autoradiograph should produce a 767bp fragment with EcoRI digestion if gene 4 is intronless between its two EcoRI recognition sites. A very faint band is seen ~880bp in the EcoRI digestion (lane 6). This

may represent the 767bp band but migrated slower in the agarose gel than expected.

The additional 5' sequence of the coding region of Clone 4 provides more potential for primer design that may be subsequently used for PCR amplification and determination of the intronic character of this region.

5' ATGGC CTGCC CCCTT ATTAT TATTT TGAGA CCAGA CAACT GTGTT
AATGG TAGGG ACCGG GGCTC AGCAG AATTC CGGCG 3'

Figure 15. The 5' Terminal Sequence of the Coding Region of Clone 4. The final 3' residues, CGGCG, represent the first 5' nucleotides of Clone 4 in figure 2. Underlined are the translation initiation site (ATG) and the EcoRI recognition sequence (GAATTC). This 5' sequence completes the ORF sequence of Clone 4. Data from J.R. Radke (unpublished work).

Future Work

Further characterization of the genes encoding C-type lectins in *Melanoplus differentialis* is necessary to determine their intronic character. Immediate possibilities for future work include:

1. Sequence the amplified regions of genes 3 and 4 to unambiguously determine their intronic nature. Sequencing is required as some introns are less than 35 nucleotides in length (Warson et al. 1992) and may be too small to be resolved on agarose or polyacrylamide gels.
2. PCR-amplify the 5' portion of the coding region of gene 3 that includes the 5' terminal

CRD-coding domain. Confirm that the entire coding region is intronless as is gene 4.

3. Completion of nucleic acid sequence within the extensive 5' untranslated regions of

Clone 3 and 4.

CONCLUSIONS

The main objectives of this research thesis were to confirm that multiple C-type lectin genes exist in the grasshopper and to indicate the intronic character of the genes representing Clones 3 and 4. These goals have been met. Major milestones achieved during the work include:

1. Confirmation that grasshopper, *Melanoplus differentialis*, contains multiple C-type lectin genes. The exact number cannot be confirmed, however the presence of multiple signals on Southern blots is clear.
2. Determination that the gene representing Clone 4 is intronless over 85% of its ORF, including both CRD-encoding regions. The Clone 3-encoding gene is continuous over 37% of its 3' coding region that includes the carboxyl terminal CRD-coding domain.
3. Proteins encoded by genes 3 and 4 may have evolved from a progenitor protein that was also an ancestor of collectins and selectins, since all the genes encoding these proteins lack introns in their CRD-coding regions.

REFERENCES CITED

- Arason G.J. (1996). Lectins as defense molecules in vertebrates and invertebrates. *Fish and Shellfish Immunology* **6**: 277-289.
- Bezouska K., Crichlow G.V., Rose J.M., Taylor M.E. and Drickamer K. (1991), Evolutionary conservation of intron position in a subfamily of genes encoding carbohydrate-recognition domains. *Journal of Biological Chemistry* **266**: 11604-11609.
- Boleda M.D., Briones P., Farres J., Tyfield L. and Pi R. (1996). Experimental design: a useful tool for PCR optimization. *BioTechniques* **21**: 134-140.
- Bradley R., Stuart G.S., Stiles B. and Hapner K.D. (1989). Grasshopper haemagglutinin: immunochemical localization in haemocytes and investigation of opsonic properties. *Journal of Insect Physiology* **35**: 353-361.
- Chou Q., Russel M., Birch D.E., Raymond J. And Bloch W. (1992). Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic Acid Research* **20**: 1717-1723.
- Drickamer K. (1993). Evolution of Ca²⁺-dependent animal lectins. *Progress in Nucleic Acid Research and Molecular Biology* **45**: 207-232.
- Drickamer K. (1994). *Molecular Glycobiology* (Fukuda M., ed.). Oxford University Press, Oxford.
- Drickamer K. and Taylor M.E. (1993). Biology of animal lectins. *Annual Review of Cell Biology* **9**: 237-264.
- Ewart K.V., Rubinsky B. and Fletcher G.L. (1992). Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biochemical and Biophysical Research Communications* **185**: 335-340.
- Feinberg A.P. and Vogelstein B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**: 6-13.
- Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S., Huang K.S., Presky D.H., Familletti P.C., Wolitzky B.A. and Burns D.K. (1994). Insight into E-

selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature* **367**: 532-538.

Haq S., Kubo T., Kurata S., Kobayashi A. and Natori S. (1996). Purification, characterization, and cDNA cloning of a galactose-specific C-type lectin from *Drosophila melanogaster*. *Journal of Biological Chemistry* **271**: 20213-20218.

Hirabayashi H., Kusunoki T. and Kasai K. (1991). Complete primary structure of a galactose-specific lectin from the venom of the rattlesnake *Crotalus atrox*. *Journal of Biological Chemistry* **266**: 2320-2326.

Hoffmann J.A. (1995). Innate immunity of insects. *Curr. Op. Immuno.* **7**: 4-10.

Hoffmann J.A., Reichhart J.M. and Hetru C. (1996). Innate immunity in higher insects. *Current Opinion in Immunology* **8**: 7-13.

Hoppe H.J. and Reid K.B.M. (1994). Collectins - soluble proteins containing collagenous regions and lectin domains - and their role in innate immunity. *Protein Science* **3**: 1143-1158.

Kawasaki N., Itoh N. and Kawasaki T. (1994). Gene organization and 5'-flanking region sequence of conglutinin: A C-type mammalian lectin containing a collagen-like domain. *Biochemical and Biophysical Research Communications* **198**: 597-604.

Kawasaki K., Kubo T. and Natori S. (1993). A novel role of *Periplaneta* lectin as an opsonin to recognize 2-keto-3-deoxy octonate residues of bacterial lipopolysaccharides. *Comparative Biochemistry and Physiology; Biochemistry and Molecular Biology* **106**: 675-680.

Kawasaki K., Kubo T. and Natori S. (1996). Presence of the *Periplaneta* lectin-related protein family in the American cockroach *Periplaneta americana*. *Insect Biochemistry and Molecular Biology* **26**: 355-364.

Kennedy J.F., Palva P.M.G., Corella M.T.S., Cavalcanti M.S.M. and Coelho L.C.B.B. (1995). Lectins, versatile proteins of recognition. *Carbohydrate Polymers* **26**: 219-230.

Kerkhof L. (1992). A comparison of substrates for quantifying the signal from a nonradiolabeled DNA probe. *Analytical Biochemistry* **205**: 359-364.

Kroczek R.A. (1993). Southern and Northern analysis. *Journal of Chromatography* **618**: 133-145.

Lasky L.A. (1992). Selectins: interpreters of cell-specific carbohydrate information

during inflammation. *Science* **258**: 964-969.

Lewin B. (1994). *Genes V* (Lewin B., ed.). Oxford University Press, Oxford.

Lodish H.F. (1991). Recognition of complex oligosaccharides by the multisubunit asialoglycoprotein receptor. *Trends in Biological Sciences* **16**: 374-377.

Mori H., Iwamoto S.I., Kotani E., Sumida M., Matsumoto T. and Matsubara F. (1992). Isolation of cDNA clones coding for humoral lectin of silkworm, *Bombyx mori*, larvae. *Journal of Invertebrate Pathology* **59**: 40-45.

Muta T. and Iwanaga S. (1996). The role of hemolymph coagulation in innate immunity. *Current Opinion in Immunology* **8**: 41-47.

Natori S. (1990). Dual functions of insect immunity in defense and development. *Research in Immunology* **141**: 938-939.

Newfeld S.J. and Gelbart W.M. (1995). Identification of two *Drosophila* TGF- β family members in the grasshopper *Schistocerca americana*. *Journal of Molecular Evolution* **41**: 155-160.

Ng N.F.L. and Hew C.L. (1992). Structure of an antifreeze polypeptide from the sea raven. *Journal of Biological Chemistry* **267**: 16069-16075.

Pendland J.C., Heath M.A. and Boucias D.G. (1988). Function of a galactose-binding lectin from *Spodoptera exigua* larval haemolymph: opsonization of blastospore from entomogenous hyphomycetes. *Journal of Insect Physiology* **34**: 533-540.

Ratcliffe N.A. (1993) Cellular defense responses of insects: unresolved problems Parasites and Pathogens of Insects (Beckage N.E., Thompson S.N. and Federici B.A., eds.). Oxford University Press, Oxford. Vol. 1: 267-304.

Saiki R.K. (1990). Amplification of Genomic DNA. PCR Potocols (Innis M.A., Gelfand D.H. Sninsky J.J. and White T.J., eds.). Academic Press Inc., San Diego.

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

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

Sharon N. and Lis H. (1995). Lectins - proteins with a sweet tooth: functions in cell

recognition. *Essays in Biochemistry* **30**: 59-75.

Stebbins M.R. and Hapner K.D. (1985). Preparation and properties of haemagglutinin from haemolymph of acrididae (grasshoppers). *Insect Biochemistry* **15**: 451-462.

Stiles B., Bradley R.S., Stuart G.S. and Hapner K.D. (1988) Site of synthesis of the haemolymph agglutinin of *Melanoplus differentialis* (Acrididae: Orthoptera). *Journal of Insect Physiology* **34**: 1077-1085.

Takahashi H., Komano H., Kawaguchi N., Kitamura N., Nakanishi S. and Natori S. (1985). Cloning and sequencing of cDNA of *Sarcophaga peregrina* humoral lectin induced on injury of the body wall. *Journal of Biological Chemistry* **260**: 12228-12233.

Takamatsu, N., Heishi M., Muramoto K., Kamiya H., and Shiba T. (1994). Cloning and analysis of the gene encoding lectin from the acorn barnacle *Megabalanus rosa*. *Gene* **150**: 407-408.

Takamatsu N., Takeda T., Kojima M., Heishi M., Muramoto K. and Kamiya H. (1993). Acorn barnacle *Megabalanus rosa* lectin (BRA-3): cDNA cloning, gene structure and seasonal changes of mRNA and protein levels. *Gene* **128**: 251-255.

Vass E. and Nappi A. (1996). The fruit fly providing insight into mechanisms of non-self recognition. *Lab Animal* :33-36.

Watson J.D., Gilman M., Witkowski J. and Zoller M. (1992). Recombinant DNA: 2nd edition (Watson J.D., Gilman M., Witkowski J. and Zoller M., eds.). Scientific American Books, New York.

Weis W.I., Kahn R., Fourme R., Drickamer K. and Hendrickson W.A. (1991). Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science* **254**: 1608-1615.

Wheeler M.B., Stuart G.S. and Hapner K.D. (1993). Agglutinin mediated opsonization of fungal blastospores in *Melanoplus differentialis* (Insecta). *Journal of Insect Physiology* **39**: 477-483.

Yousif A.N., Albright L.J. and Evelyn T.P.T. (1994). Purification and characterization of a galactose-specific lectin from the eggs of coho salmon *Oncorhynchus kisutch*, and its interaction with bacterial fish pathogens. *Diseases of Aquatic Organisms* **20**: 127-136.

Yousif A.N., Albright L.J. and Evelyn T.P.T. (1995). Interaction of coho salmon *Oncorhynchus kisutch* egg lectin with the fish pathogen *Aeromonas salmonicida*. *Diseases of Aquatic Organisms* **21**: 193-199.

MONTANA STATE UNIVERSITY LIBRARIES



3 1762 10309115 1