This is to certify that the

dissertation entitled

Molecular analysis of the structure and function of two chitin-binding proteins: barley and stinging nettle lectins.

presented by

David Ross Lerner

has been accepted towards fulfillment
of the requirements for

Doctoral degree in Botany and Plant Pathology

Natasha Raikhel
Major professor

Date Feb. 2, 92

MOLECULAR ANALYSIS OF THE STRUCTURE AND FUNCTION OF TWO CHITIN-BINDING PROTEINS: BARLEY AND STINGING NETTLE LECTINS

By

David Ross Lerner

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Botany and Plant Pathology

1992

ABSTRACT

MOLECULAR ANALYSIS OF THE STRUCTURE AND FUNCTION OF TWO CHITIN-BINDING PROTEINS: BARLEY AND STINGING NETTLE LECTINS

BY

David Ross Lerner

Plants possess a family of chitin-binding proteins containing one or more conserved domains, 40 to 43 amino acids in length, responsible for binding to chitin. Stinging nettle and barley lectins consist of two and four repeated chitin-binding domains, respectively. Besides binding chitin, both of these lectins can inhibit the growth of insect larvae *in vitro*. Nettle lectin, however, can also inhibit the growth of some chitin-containing fungi. The objectives of this dissertation were to isolate and characterize barley and nettle lectin cDNA clones and to investigate the expression and structure of these lectins as they relate to possible functions in plant defense.

The barley lectin cDNA clone, isolated using a partial wheat germ agglutinin (WGA) cDNA as a heterologous probe, was shown to be over 90% identical to

WGA. Expression of barley lectin mRNA was localized to the adult root tips and the embryonic root cap and coleorhiza with RNA gel blot analysis and *in situ* hybridizations. Barley lectin was also shown to be synthesized with a glycosylated carboxyl-terminal propeptide not found in the mature protein.

A synthetic gene for nettle lectin was prepared and used as the probe for cloning the nettle lectin cDNA. Although nettle lectin is only 86 amino acids long, the cDNA encoded a 372 amino acid polypeptide containing a putative signal sequence, the lectin, and a carboxyl-terminal domain with chitinase activity and similarity to other plant chitinases. Southern and northern blot analyses were used to determine that nettle lectin was encoded by a small multigene family expressed in rhizomes and inflorescence.

To initiate investigations into the relationship between structure and antifungal activity of chitin-binding proteins, modified barley lectin (mBL) genes encoding approximately two chitin-binding domains, the size of nettle lectin, were constructed using oligonucleotide directed mutagenesis. Analysis of mBL expressed in either *E. coli* or tobacco revealed that the mBL retained chitin-binding activity.

The results of these investigations are consistent with the hypothesis that plant chitin-binding proteins have defense-related functions. The potential for modifying chitin-binding proteins for investigating basic biological problems and practical applications has also been demonstrated.

To my parents who, through their love and example,

have shown me that anything is possible.

ACKNOWLEDGEMENTS

I would like to thank the members of my committee for their helpful discussion throughout my work on this dissertation. I would also like to thank Dr. John Olhrogge for help with the design and synthesis of the synthetic gene for stinging nettle lectin. Everyone in Natasha's laboratory, past and present, deserve thanks for their constant help with techniques, ideas, encouragement and forebearance of numerous questions.

Natasha deserves a special acknowledgement for her constant encouragement and support. She has often gone far beyond what was required or expected to ensure my development as a productive researcher.

I am also indebted to my dear friends Neera Agrwal and David Rhoads kept me amused, distracted and in shape. Finally, I must thank Yolanda for being with me even when she was miles away.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Blc3	Barley lectin cDNA clone #3
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
cDNA	Complementary deoxyribonucleic acid
CTPP	Carboxyl-terminal propeptide
dATP	Deoxy-adenosine triphosphate
dGTP	Deoxy-guanosine triphosphate
DPA	Days post-anthesis
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethylene glycol-bis(ß-aminoethyl ether)-N,N,N',N'-tetraacetic acid
Endo H	Endoglycosidase H
FW	Fresh weight
GlcNAc	N-acetylglucosamine
HPLC	High pressure liquid chromatography
HRGP	Hydroxyproline rich glycoprotein
IPTG	Isopropyl-ß-D-thiogalactopyranoside
kb	Kilobases
kDa	Kilodaltons

mBL	Modified barley lectin
MBP	Maltose binding protein
mRNA	Messenger ribonucleic acid
nt	nucleotide
oligo	oligodeoxyribonucleotide
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pfu	plaque forming unit
PMSF	Phenylmethyl sulfonyl fluoride
PVP	Polyvinyl pyrrolidone
RBV	Remizol brilliant violet
rGTP	Riboguanosine triphosphate
RT	Room temperature
SDS	Sodium dodecyl sulfate
SSC	Standard sodium citrate
TE	Tris (10 mM), EDTA (0.1 mM)
TMV	Tobacco mosaic virus
Tris	Tris(hydroxymethyl)amino methane
UDA	*Urtica dioica* agglutinin
WGA	Wheat germ agglutinin
win1	Wound inducible gene 1 (potato)
win2	Wound inducible gene 2 (potato)

CHAPTER 1

INTRODUCTION:

Chitin-Binding Proteins and Evidence for Their Role in Plant Defense

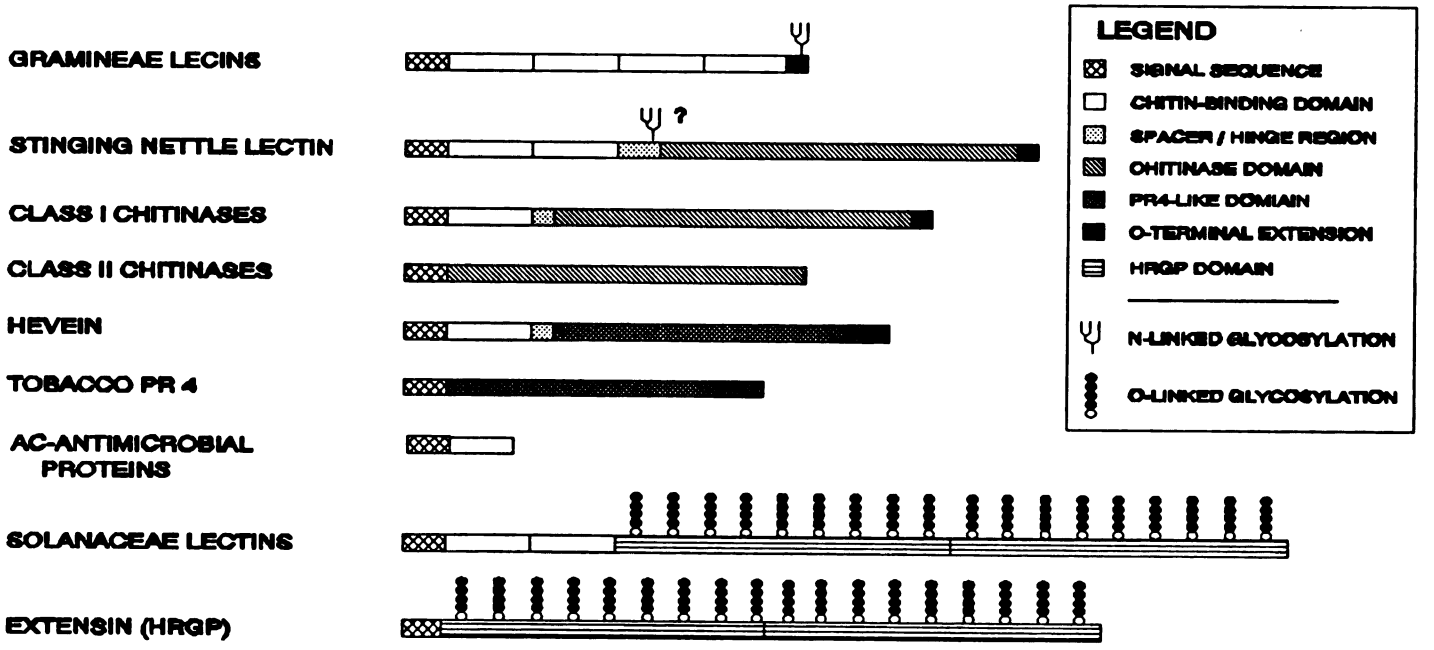
INTRODUCTION

A family of chitin-binding proteins in plants can be identified by the presence of a distinct and conserved chitin-binding domain (figure 1, open boxes). This domain is characteristically 40-43 amino acids in length and is rich in Gly and Cys residues. The structure of the chitin-binding domain has been elucidated from X-ray crystallography of wheat germ agglutinin (WGA) (Wright, 1977). WGA has four tandem repeats of the distinct chitin-binding domain. Each chitin-binding domain of WGA contains 8 Cys residues making up four disulfide bonds. This maintains the tight globular structure of each domain since there are no distinct regions of α-helix or ß-sheet in the chitin-binding domains (for review see Wright *et al.*,1991). Further studies of WGA crystallized with bound polysaccharides revealed the specific amino acid residues directly involved in carbohydrate binding (Wright, 1980). These residues and the Cys residues involved in disulfide bridges are highly conserved in all chitin-binding proteins examined (Wright *et al.*, 1991).

Chitin-binding proteins can be divided between proteins with multiple chitin-binding domains repeated in tandem and proteins containing a single chitin-binding domain. In both cases the chitin-binding domain is often fused to the amino-terminal of a structurally unrelated domain (figure 1). Only proteins with multiple chitin-binding domains (ie. Gramineae lectins) are able to agglutinate red blood cells indicating the presence of multiple N-acetylglucosamine (GlcNAc), the chitin monomer, binding sites.

Figure 1. The domain structure of chitin-binding and related proteins in plants. Boxes representing domains of each class of protein are shown (not to scale). Domain structures have been interpreted mostly from the deduced amino acid sequence of cloned cDNAs. Ac-antimicrobial proteins and Solanaceae lectin domain structure has been deduced solely from biochemical data since genes for these proteins have not been cloned. Stinging nettle lectin contains a potential N-linked glycosylation site , denoted by the question mark (?), but no data is available on its utilization.

THE CHITIN-BINDING FAMILY AND RELATED PROTEINS



The wide distribution of both classes of chitin-binding proteins in monocots and dicots indicates conserved and essential functions for these proteins. In spite of this, the endogenous functions have yet to be conclusively demonstrated. There are many indications, however, that these proteins are involved in plant defense against microbial pathogens or insects (Chrispeels and Raikhel, 1991). In this review I will describe the chitin-binding proteins, their processing, regulation and *in vitro* activities (Table 1) with particular emphasis on properties related to possible plant defense functions.

PROTEINS CONTAINING ONLY CHITIN-BINDING DOMAINS

Proteins comprised solely of chitin-binding domains are of special interest since their properties and characteristics are not influenced by the presence of an unrelated domain. In all cases these proteins are highly resistant to heat and proteases, and retain chitin-binding activity after heating to 70°C for 15 min. These proteins range in size from 3 kDa monomeric polypeptides to 36 kDa homodimers containing 8 chitin-binding domains.

Gramineae Lectins

The Gramineae lectins are, perhaps, the most studied proteins containing chitin-binding domains. These lectins have been isolated from wheat, barley, rye, rice and a number of other Gramineae species (for review see Raikhel and Lerner,

Table 1. Chitin-binding proteins properties and *in vitro* activities

PROTEIN	MOL. WT.: MATURE/ PROPROTEIN	# CHITIN-BINDING DOMAINS	CARBOXYL-TERMINAL DOMAIN	*IN VITRO* ACTIVITIES
Gramineae lectins	18 kDa/23 kDa 36 kDa	4/monomer 8/dimer	None	Agglutination Insecticidal
Stinging nettle lectin	8.5 kDa/ (38 kDa)[1]	2	Chitinase	Agglutination Fungicidal Insecticidal
Class I chitinase	30 kDa	1	Chitinase	Fungicidal
Hevein	4.5 kDa/ 20 kDa	1	PR4-like	Fungicidal
Solanaceae lectins	30 to 68 kDa	2[2]/monomer	HRGP-like	Agglutination
Ac-AMP	3 kDa	$^2/_{3^{rd}}$	?	Fungicidal

[1]Molecular weight shown in parenthesis is deduced from sequence data.
[2]Based on biochemical data only.

1991). In addition, cDNA clones have been isolated for the three WGA isolectins (Raikhel and Wilkins, 1987; Smith and Raikhel, 1989), barley lectin (Lerner and Raikhel, 1989), and rice lectin (Wilkins and Raikhel, 1989). Mature Gramineae lectins are 36 kDa homodimers with four chitin-binding domains in each monomer. Biochemical studies, including subunit exchange experiments (Peumans *et al.*, 1982), show that the Gramineae lectins, and in particular the cereal lectins, are very similar. Carbohydrate binding sites of these lectins have been localized to the cleft between the subunits with both monomers contributing residues which interact directly with the ligand. Although dimer formation is apparently essential for the Gramineae lectin ligand binding, Wright and coworkers (1991) suggest that dimer formation is not required for saccharide binding in other chitin-binding proteins.

In all cases studied, the Gramineae lectins are expressed in specific cell layers in both embryos and root tips of adult plants (for review see Raikhel and Lerner, 1991). Immunolocalization (Mishkind *et al.*, 1983) and *in situ* hybridization to mRNA (Lerner and Raikhel, 1989; Wilkins and Raikhel, 1989; Raikhel *et al.*, 1988) allowed the determination that root tip expression is restricted to the root cap cells and outer layer of cells at the root tip. These studies also show that Gramineae lectins are expressed in tissues which protect the embryonic leaves, the coleoptile, and roots, the coleorhiza, in the seed and while germinating. Expression of the Gramineae lectins in the coleorhiza and progenitor root cap cell of the embryo was found in all species studied (Mishkind *et al.*, 1983). Gramineae lectin expression in the coleoptiles, however, varies from species to species

(Mishkind *et al.*, 1983). In rice, the lectin is expressed throughout the coleoptile. Rye contains lectin in the single outermost and innermost cell layers, while in wheat, WGA is present only in the outermost cell layer of the coleoptile. Barley lectin is not found in the coleoptile at all (Mishkind *et al.*, 1983 and Lerner and Raikhel, 1989). Interestingly, root cap cells, the coleoptile and the coleorhiza are all considered to be protective tissues and are the first to contact soil microbes as the roots and germinating seedlings grow through the soil.

Stinging Nettle Lectin

Stinging nettle lectin (*Urtica dioica* agglutinin, UDA) is a 8.5 kDa chitin-binding protein which, like the Gramineae lectins, can agglutinate red blood cells (Peumans *et al.*, 1984). Unlike the Gramineae lectins however, stinging nettle lectin does not effectively bind the chitin monomer, GlcNAc, which may account for its lower agglutination specific activity (Shibuya *et al.*, 1986). The amino acid sequence of stinging nettle lectin, as determined by analysis of tryptic fragments (Chapot *et al.*, 1986) and the cloned cDNA (Lerner and Raikhel, 1992), shows two chitin-binding domains in tandem. Proteins with multiple chitin-binding domains therefore are not restricted to the Gramineae. This may be important evolutionarily especially if the domain duplication occurred prior to divergence into monocots and dicots as suggested by Wright, *et al.* (1991). Also, like the Gramineae lectins, stinging nettle lectin is found in the seeds (Lerner and Raikhel, 1992) and in underground tissues, in this case the roots and rhizomes (Peumans, *et al.*, 1984).

Localization of stinging nettle lectin to cell layers within these tissues has not been determined.

Hevein

Hevein is perhaps the archetypical chitin-binding protein because it consists of exactly one 43 amino acid chitin-binding domain. This domain, when found in other plant proteins, has been referred to as the 'hevein domain'. Hevein accumulates in the vacuole-like lutoid bodies of specialized cells, laticifers, of rubber trees (*Hevea brasiliensis*) (Walujono *et al.*, 1975). These cells form an anastomosing system through the trunk, stems and leaves of rubber trees and are responsible for latex synthesis (Esau, 1960).

***Amaranthus caudatus* Antimicrobial Peptides (Ac-AMP1 and Ac-AMP2)**

Two very small, 30 and 31 amino acid, chitin-binding proteins have been isolated from the seeds of amaranth (*Amaranthus caudatus*) (Broekaert *et al.*, 1992). Peptide sequencing data indicates Ac-AMP1 and 2 are identical except for an additional Arg residue at the carboxyl-terminus of Ac-AMP2. Thus Ac-AMP1 and Ac-AMP2 may be derived from a single precursor polypeptide by differential post-translational processing or degradation during the isolation procedure. These chitin-binding proteins appear to be truncated variants of the chitin-binding domain. Alignment of the Ac-AMP peptide sequences and chitin-binding domains from other proteins shows that approximately 50% of the residues are identical with the insertion of a 4 residue gap in the Ac-AMP sequence. Wright and

coworkers (1991) refer to residues 3 to 30 of a consensus chitin-binding domain as the "inner conserved core" and define it as responsible for ligand binding. Ac-AMPs amino acid sequence corresponds to this "inner conserved core". It is not surprising therefore that the shortened chitin-binding domains of Ac-AMPs do bind chitin.

PROTEINS CONTAINING CHITIN-BINDING DOMAINS FUSED TO AN UNRELATED DOMAIN

Class I Chitinases

The class I chitinases contain a single 40 to 43 amino acid chitin-binding domain linked by a short hinge region to the chitinase catalytic domain (for review see Boller, 1988). The chitin-binding domain generally gives these proteins a high isoelectric point. Because of this, these chitinases are often referred to as the basic chitinases. Proteins containing a homologous chitinase catalytic domain by itself, the acidic or class II chitinases are expressed from genes which do not encode a chitin-binding domain. Thus the class II chitinases are not derived from the class I chitinases by removal of the chitin-binding domain. The class I chitinases are expressed in many plant tissues including leaves, flowers, seeds, and roots (Shinshi *et al.*, 1987; Lotan, 1989). These chitinases have been identified in virtually all plants examined (for reviews see Linthorst, 1991; Bol *et al.*, 1990).

Prostinging Nettle Lectin

The stinging nettle lectin mRNA encodes a polypeptide extending 174 amino acids beyond the 86 amino acids encoding the lectin's two chitin-binding domains (Lerner and Raikhel, 1992). This carboxyl-terminal domain has approximately 45% homology with class I and class II chitinase catalytic domains. The identity of prostinging nettle lectin's carboxyl-terminal domain with these other chitinase catalytic domains, however, is considerably less than the identity between the class I and class II chitinases. In spite of this, chitinase activity is observed when the 37 kDa prostinging nettle lectin or the 28 kDa carboxyl-terminal domain alone is expressed in *E. coli* (Lerner and Raikhel, 1992). It is not known, however, whether prostinging nettle lectin exists as a stable protein or is just a transient precursor of stinging nettle lectin in the plant.

Prohevein and Potato Wound Inducible Genes, winI and winII

Like stinging nettle lectin, the hevein mRNA encodes a polypeptide extending far beyond (141 amino acids) the single chitin-binding domain (43 amino acids) which makes up hevein (Broekaert *et al.*, 1990). Highly homologous polypeptides are also encoded by two wound inducible genes from potato, winI and winII (Stanford *et al.*, 1989). The 20 kDa prohevein, 4 kDa hevein and 16 kDa hevein carboxyl-domain are all found as major protein constituents of rubber tree latex (Lee *et al.*, 1991) although these polypeptides have not been identified in potato. Proteins with homology to the hevein carboxyl-domain alone have now been identified as the pathogenesis-related protein from tobacco, PR-4, and

tomato, PR-P2 (Linthorst *et al.*, 1991). Cloning the PR-4 and PR-P2 cDNAs reveals that they are encoded by mRNAs which do not encode the chitin-binding domains (Linthorst *et al.*, 1991).

Solanaceae Lectins

Chitin-binding proteins from potato tubers, tomato fruits, and thorn-apple seeds have been identified which contain a Cys and Gly rich, chitin-binding, domain linked to a highly glycosylated hydroxyproline-rich domain (Broekaert *et al.*, 1987; Allen, *et al.*, 1978; Nachbar, *et al.*, 1980). Although little protein sequence data are available, the amino acid composition and chitin-binding activity of these proteins suggests they contain chitin-binding domains homologous to other chitin-binding proteins. In addition, the agglutinating activity exhibited by these lectins suggests they contain multiple chitin-binding domains. Amino acid and carbohydrate compositions of the hydroxyproline-rich domain indicate that this region is very similar to other hydroxyproline-rich glycoproteins (HRGPs) such as extensin. These unique lectins are the only chitin-binding proteins known to contain a high amount of carbohydrate.

PROTEIN PROCESSING AND TARGETING

Chitin-binding proteins are secretory proteins which are synthesized on the ER and processed through the Golgi apparatus before accumulation in the

vacuole or cell wall (Chrispeels and Raikhel, 1991; Meins *et al.*, 1992). The analysis of cDNA clones of chitin-binding proteins indicates they all encode a putative amino-terminal signal sequence which would target the nascent polypeptides to the ER.

Carboxyl-Terminal Extensions and Vacuolar Targeting

In addition to the signal sequence, the cDNAs of several chitin-binding proteins encode a short extension beyond the carboxyl-terminus of the mature protein. The Gramineae lectins, for example, have well characterized carboxyl-terminal propeptides (CTPP's). The WGA isolectins and barley lectin have 15 amino acid CTPP's which are almost identical in sequence (for review see Raikhel and Lerner, 1991). Rice lectin contains a CTPP which differs in length (26 amino acids) and amino acid sequence from the cereal lectin CTPPs (Wilkins and Raikhel, 1989). A common feature of the Gramineae lectin CTPPs is an asparagine-linked glycosylation site which is utilized to yield a glycosylated precursor (Mansfield *et al.*, 1988). This glycosylated precursor is processed into the smaller, unglycosylated, mature polypeptide. The tobacco class I chitinase is also synthesized with a carboxyl-terminal extension which is not found in the mature protein (J.-M. Neuhaus, Botanisches Institut der Universitat Basel, personal communication). The chitinase extension is only 6 amino acids long and does not contain a glycosylation site. Comparison of cloned class I chitinases indicates that they all encode similar regions 6 to 10 amino acids long.

The CTPP of barley lectin and the carboxyl-terminal extension of the tobacco class I chitinase have recently been shown to be the vacuolar targeting signals for these proteins. Bednarek *et al.* (1990) demonstrated that barley lectin is secreted when expressed in tobacco plants without the CTPP while the wild type barley lectin is faithfully targeted to the vacuole. The CTPP of barley lectin (Bednarek and Raikhel, 1991) and the carboxyl-terminal extension of the tobacco class I chitinase (Neuhaus *et al.*, 1991b) are sufficient for redirecting a normally secreted protein, cucumber class III chitinase, to the vacuole when expressed in tobacco. Interestingly, the secreted class II chitinases, which do not possess the amino-terminal chitin-binding domain, also do not have the carboxyl-terminal extension encoded in their cDNA's. Mauch and Staehelin (1989) suggest that the vacuolar localization of the class I chitinase and ß-1,3-glucanase in bean leaves is directly related to their role as a last line of the plant's defense against invading pathogens.

Two other chitin-binding proteins, prostinging nettle lectin and prohevein, have carboxyl domains longer than homologous proteins which do not contain the chitin-binding domain. Like the class I chitinases, the stinging nettle lectin chitinase domain is 10 amino acids longer than the class II chitinases. The intra cellular localization and potential processing of this extension have not yet been investigated. As stated earlier, the carboxyl-domain of prohevein is homologous to tobacco PR-4 which does not contain a chitin-binding domain. Similar to the chitinases, the vacuolar prohevein is 14 amino acids longer than the secreted PR-4. The actual carboxyl-terminus of prohevein, however, is still unknown. Thus, a

pattern is emerging wherein chitin-binding proteins containing an extensive carboxyl-domain also have a short carboxyl-terminal extension which may be responsible for their targeting to the vacuole. Homologous proteins lacking the chitin-binding domain and carboxyl-terminal extension are found in the apoplastic space outside the plasma membrane. It is important to note that chitin-binding domains by themselves are not sufficient for targeting or redirecting proteins to the vacuole (Bednarek and Raikhel, 1991; Neuhaus *et al.*, 1991b).

Additional Processing Events

Several chitin-binding proteins undergo internal proteolytic cleavage events. Approximately 85% of the rice lectin monomer (18 kDa) is cleaved into 8 and 10 kDa polypeptides (Stinissen *et al.*, 1983). Rice lectin also contains an additional pair of Cys residues (Wilkins and Raikhel, 1989) not found in the other Gramineae lectins, which form a disulfide bridge between the 8 and 10 kDa polypeptides (Wright *et al.*, 1991).

Prohevein undergoes an internal cleavage to release hevein from the carboxyl-domain (Lee *et al.*, 1991). During this processing a 6 amino acid 'spacer' region between the two domains is lost. It is interesting to note that, similar to rice lectin, apparently not all of the prohevein is cleaved and all three polypeptides (prohevein, hevein and the carboxyl-domain) are found in the rubber tree latex (Lee *et al.*, 1991).

Presumably, prostinging nettle lectin is also processed to release the 8.5 kDa lectin from the carboxyl-terminal chitinase domain. This processing event,

however, has not been demonstrated experimentally. In light of the accumulation of the 8.5 kDa lectin to high levels in the rhizomes and the processing events demonstrated for prohevein, it is likely that prostinging nettle lectin undergoes a specific proteolytic cleavage as well.

REGULATION OF GENE EXPRESSION

Developmental Regulation

Many chitin-binding proteins are expressed in specific tissues. As a result, their expression is, in part, developmentally regulated. The Gramineae lectins, for example, are expressed in developing embryos and accumulate as the embryos mature. Raikhel *et al.* (1988) found WGA mRNA is first detectable at 10 days post anthesis (DPA) and reaches maximum levels at 30 to 40 DPA. The WGA mRNA level later decreases as the seeds desiccate and increases slightly as the seeds germinate, due to increased lectin expression in the growing root tips (Raikhel *et al.*, 1988). Stinging nettle lectin is also found in seeds and not in leaves or stems (Lerner and Raikhel, 1992) indicating developmental control over its accumulation. Interestingly, stinging nettle lectin mRNA and protein are not found in aerial stems but accumulate to high levels in stems which have been buried to generate rhizomes (Lerner and Raikhel, 1992). The developmental conversion of these stems to rhizomes must, therefore, induce expression of the lectin. The class I chitinases have been found in many different plant tissues and developmental

regulation of expression has been particularly noted by higher mRNA levels in older leaves (Shinshi *et al.*, 1987) and floral tissues (Lotan *et al.*, 1989).

Hormonal Regulation

Exogenously applied plant growth regulators, abscisic acid (ABA) and ethylene, have been shown to enhance the levels of message and protein for several chitin-binding proteins. Gramineae lectin accumulation in embryos and root tips of adult plants is increased by ABA (Raikhel *et al.*, 1986; Triplett and Quatrano, 1982). For example, WGA accumulation in root-tips is increased 2 to 3 fold by exogenous application of 10 mM ABA (Raikhel *et al.*, 1986). Endogenous ABA also accumulates in developing wheat embryos beginning just before the developmental increase in WGA accumulation (Raikhel *et al.*, 1987). This is consistent with the hypothesis of ABA control over WGA expression. Exogenously applied ABA (50 uM) also causes a several-fold increase in Hevein message in rubber tree latex (Broekaert *et al.*, 1990). Hevein and chitinase have been shown to accumulate upon application of ethephon which converts to ethylene. Hevein mRNA in rubber tree leaves, stems, and latex increased several-fold in the presence of 0.1% ethephon (Broekaert *et al.*, 1990). Broglie *et al.* (1986) found a 30-fold increase in chitinase activity of bean seedlings following exposure to 1mg/ml ethephon.

Stress Induced Expression

Environmental stresses often result in the increased production of ABA or ethylene by the plant. It is not surprising, therefore, that osmotic stress and wounding can also cause the accumulation of chitin-binding proteins. The effects of osmotic stress on expression of chitin-binding proteins have been examined only for the Gramineae lectins. WGA and ABA accumulate when seedlings are stressed by air drying or chemically with mannitol (0.5 M) or polyethylene glycol (180 g/l) (Cammue *et al.*, 1989). Under these conditions WGA levels increase up to 10-fold, but only in tissues at the root tips where the protein is normally found (Cammue *et al.*, 1989). Similar results have been found for expression of chitin-binding proteins after wounding. The proteins accumulate to higher than normal levels in tissues where they are otherwise expressed. Within 24 hours of wounding, dramatic increases in hevein message in rubber tree leaves is seen. A slightly lesser increase in hevein mRNA accumulation is found in rubber tree stems and latex (Broekaert *et al.*, 1989). The results for latex expression are somewhat difficult to interpret since latex is collected by 'wounding' the rubber tree. Because of this, the high levels of expression seen in 'unwounded' latex may also be due to wounding. As implied by their names, the hevein homologues in potato, wound inducible mRNAs 1 and 2 (win1 and win2) accumulate in wounded potato plants (Stanford *et al.*, 1989). Interestingly, these two genes seem to have arisen through gene duplication and show differential expression in potato tissues after wounding (Stanford *et al.*, 1989). Specifically, win2 will accumulate in wounded tubers while win1 does not. Both mRNAs accumulate in wounded leaves

and stems but not in wounded roots. Class I chitinases have also been shown to accumulate in leaf tissue in response to wounding (Hedrick *et al.*, 1988; Parsons *et al.*, 1989). Wounded stinging nettle leaves do not, however, accumulate the stinging nettle lectin mRNA which also encodes a chitinase (Lerner and Raikhel, unpublished results).

Induction by Plant Pathogens

Although plants contain no chitin, insect pests and many fungal and bacterial pathogens do. Plant responses to these organisms by the accumulation of chitin-binding proteins are strongly suggestive of the role chitin-binding proteins may play in plant defense. Tobacco mosaic virus infection causes the systemic induction of the classical acidic pathogenesis-related (PR) proteins, and the class I chitinases (Legrand *et al.*, 1987). The levels of tobacco class I chitinase, for example, increase 20-fold after 3 days of infection by tobacco mosaic virus (TMV) (Vögeli-Lange *et al.*, 1988). In addition, the tobacco class I chitinase mRNA increases 6-fold after infection by the bacterial pathogen *Pseudomonas tabaci* and increases 200-fold after infection by the fungal pathogen *Phytophthora parasitica* var *nicotianae* (Meins and Ahl, 1989). In bean (*Phaseolus vulgaris*), infection of hypocotyls with the fungus *Colletotrichum lindemuthianum* causes an increase in class I mRNA which peaks at approximately 7 days after inoculation (Hedrick *et al.*, 1988). A fungal cell wall elicitor causes an increase of the class I chitinases in suspension cultured cells of tobacco and rice (Hedrick *et al.*, 1988; Zhu and Lamb, 1991 respectively). WGA accumulation is also induced 2- to 5-fold in root-tips by

infection with several different fungi and fungal cell wall elicitors (Cammue *et al.*, 1990).

Many insect pests cause wounding-like injuries as they eat the plants. It is likely, therefore, that hevein, the potato win genes, and the class I chitinases will accumulate in response to insects since they are induced in wounding experiments.

IN VITRO ACTIVITIES

Activities due to the Chitin-Binding Domains

Many chitin-binding proteins are made up solely of chitin-binding domains and activities observed for these proteins can be attributed to these domains. Besides binding chitin, a number of these proteins agglutinate red blood cells and are therefore termed lectins. Agglutination of red blood cells requires the proteins to crosslink the cells by binding to extracellular oligosaccharides. Two binding sites are required for crosslinking, although a single chitin-binding domain will potentially bind the same oligosaccharides. Stinging nettle lectin and the Gramineae lectins, which have 2 and 4 chitin-binding domains respectively, are able to agglutinate red blood cells. Potato lectin, which also agglutinates blood cells, should have multiple chitin-binding domains, although no direct evidence for this is available. All chitin-binding proteins tested which contain a single chitin-binding domain, however, do not show agglutination activity.

Perhaps more significant is the finding that several proteins containing only chitin-binding domains can inhibit the growth of fungi *in vitro*. Stinging nettle lectin, hevein, and the Ac-AMPs have been shown to retard the growth of a number of chitin-containing fungi (Broekaert *et al.*, 1989; Van Parijs *et al.*, 1991; and Broekaert *et al.*, 1992 respectively). A striking feature of all these proteins is their small size. Stinging nettle lectin, the largest of these proteins, is only 8.5 kDa while hevein is 4 kDa and the Ac-AMPs are approximately 3 kDa. The Gramineae lectins, on the other hand, form dimers of 36 kDa and do not show antifungal activity (Schlumbaum *et al.*, 1986). It is tempting to speculate that the small size of the chitin-binding proteins with antifungal activity allows them to penetrate into, and interact with, the fungal wall where the Gramineae lectins cannot. An estimate of fungal wall pore size supports this hypothesis (Money, 1990). This study predicts that proteins larger than 15 to 20 kDa will not be able to pass through the fungal cell wall.

Several chitin-binding proteins have also been tested for their effect on insects. WGA, rice lectin and stinging nettle lectin are all able to retard the growth of cowpea weevil larvae *in vitro* (Murdock *et al.*, 1990; Huesing, *et al.*, 1991b). WGA and rice lectin are similar in activity and two to four times more effective per mole than stinging nettle lectin (Huesing *et al.*, 1991a). In addition, WGA has been shown to be deleterious to two major corn pests, the European corn borer and the Southern corn borer when added to their diet (Czapla *et al.*, 1991). While low doses of these lectins don't kill the insects, they delay larval development time by approximately 1.5 days per 0.1% (w/w) increase in lectin concentration (Huesing

et al., 1991b). This may be significant since plants and their insect pests often exist in an equilibrium where a delay in insect development may shift the balance in favor of the plant.

Activities due to the Carboxyl-Terminal Domains

The chitinase activities observed for the class I chitinases and prostinging nettle lectin are the most obvious *in vitro* activities which can be attributed to the carboxyl-domain of chitin-binding proteins. In addition, the antifungal activity exhibited by the class I chitinases differs from that observed for stinging nettle lectin, hevein or Ac-AMPs. Microscopic examination reveals that the fungal hyphal tips burst in the presence of chitinase while the other chitin-binding proteins slow the growth of the hyphae and cause a characteristic swelling just behind the hyphal tip (Broekaert *et al.*, 1989; Broekaert *et al.*, 1992; Van Parijs *et al.*, 1991).

No function or activities have been ascribed to the carboxyl-terminal domain of hevein, win1 or win2. This domain, however, is highly conserved in plants as divergent as rubber tree, potato, tobacco and tomato suggesting an important function within the plant.

The Solanaceous lectins have a domain reminiscent of other hydroxyproline rich glycoproteins such as extensin. These HRGPs are thought to be involved in strengthening the plant cell wall, especially in response to pathogens (for review see Bowles, 1990). It is likely that the Solanaceous lectins have similar roles (Casalongué and Pont Lezica, 1985).

Synergistic Effects

A synergistic enhancement of antifungal activity between stinging nettle lectin and chitinase is one of the most interesting findings from *in vitro* activity experiments of chitin-binding proteins. Broekaert *et al.* (1989) found that although stinging nettle lectin and chitinase would inhibit fungal growth individually, combinations of the two acted cooperatively. The potential for enhancing plant resistance against fungal pathogens by constitutive expression of class I chitinases has been addressed by two groups. Broglie and coworkers (1991) found that tobacco (*Nicotiana tabacum* cv. Xanthi) and canola (*Brassica napus* cv. Westar) expressing a bean chitinase gene under control of the cauliflower mosaic virus 35S promoter had an increased ability to survive in soils infected with *Rhizoctonia solani* and showed delays in disease symptom development. These transgenic tobacco plants however were only slightly resistant to *Cercospora nicotiana.* The other group (Neuhaus *et al.*, 1991a) found that high levels of a tobacco (*N. tabacum*) class I chitinase gene expressed in *Nicotiana sylvestris* also did not enhance resistance to *C. nicotianae.* Thus, these studies are inconclusive as to the effectiveness of engineering fungal resistance in plants by overexpressing chitinase alone. Since stinging nettle lectin acts synergistically with chitinase to inhibit fungal growth *in vitro,* the possibility exists that overexpression of prostinging nettle lectin, with its lectin and chitinase domains, will enhance resistance to fungal pathogens not seen with the chitinase alone. Another possibility is that stinging nettle lectin will work cooperatively with endogenous chitinases in the transgenic plant.

Ac-AMPs do not act synergistically with either chitinase or stinging nettle lectin in inhibiting fungal hyphae (Broekaert *et al.*, 1992). Unlike stinging nettle lectin, however, the Ac-AMPs antifungal activity is strongly inhibited by divalent cations indicating that these proteins may act in different ways. The chitinases have also been found to act synergistically with glucanases for fungal inhibition (Mauch *et al.*, 1988). In this case it is likely that the two enzymes act on different polysaccharides in the fungal cell wall to give the enhanced antifungal effect.

CONCLUSIONS

The biological function of chitin-binding proteins in plants have yet to be conclusively demonstrated. There is, however, a wealth of circumstantial evidence for functions in plant defense. First there is the lack of chitin in plants and its presence in many plant pathogens and pests. Second, the conservation of amino acid sequence and chitin-binding activity within this family of proteins indicates an essential role in the plant's interactions with organisms containing chitin. Especially noteworthy is the distinct pattern of Cys residues and conservation of residues directly involved in chitin-binding. Fusion of the chitin-binding domain to a number of structurally unrelated domains suggests this domain may be genetically recombined with other polypeptides by the plant to create new antimicrobial proteins. The finding that the carboxyl-terminal domain of prohevein is

homologous to tobacco and tomato pathogenesis-related proteins suggests additional activities may be found for chitin-binding proteins.

Localization of the chitin-binding proteins in the plants also suggests they may interact with plant pathogens. The Gramineae lectins are expressed in the tissues which first expand into new areas of soil as the embryo germinates and the root system grows. The intracellular localization of the Gramineae lectins and the class I chitinases in the vacuole suggests that these proteins act after the invasion of the plant tissue which releases the contents of the cell. Stinging nettle lectin is also localized in underground tissues and seeds. Ac-AMPs are also found in seeds. Similar tissue specific localization of similar proteins in divergent plant species suggests important roles for these proteins.

Induction of chitin-binding protein expression by wounding, microbial infection, and the plant 'stress' hormones also indicates these proteins act as part of the plant's defense mechanisms. Chitinases, Gramineae lectins and hevein are all induced under these conditions. Finally, and perhaps most importantly, are the *in vitro* antifungal and insecticidal activities shown for chitin-binding proteins. Every chitin-binding protein tested exhibits insecticidal, antifungal or both activities.

REFERENCES

Allen, A.K., Desai, N.N., and Neuberger, A. (1978). Properties of potato lectin and the nature of its glycoprotein linkages. Biochem. J. **171**:665-674.

Bednarek, S.Y. and Raikhel, N.V. (1991). The barley lectin carboxyl-terminal propeptide is a vacuolar protein sorting determinant in plants. Plant Cell **3**:1195-1206.

Bednarek, S.Y., Wilkins, T.A., Dombrowski, J.E., and Raikhel, N.V. (1990). A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. Plant Cell **2**:1145-1155.

Bol, J.F., Linthorst, H.J.M. and Cornelissen, B.J.C. (1990). Plant pathogenesis-related proteins induced by virus infection. Annu. Rev. Phytopathol. **28**:113-138.

Boller, T. (1988). Ethylene and the regulation of antifungal hydrolases in plants. Oxford Surveys of Plant Mol. And Cell. Biology **5**:145-174.

Bowles, D.J. (1990). Defense-related proteins in higher plants. Annu. Rev. Biochem. **59**:873-907.

Broekaert, W.F., Allen, A.K., and Peumans, W.J. (1987). Separation and partial characterization of isolectins with different subunit compositions from *Datura stramonium*. FEBS Lett. **220**:116-120.

Broekaert, W.F., Lee, H.-I., Kush, A., Chua, N.-H., and Raikhel, N.V. (1990). Wound-induced accumulation of mRNA containing a hevein sequence in laticifers of rubber tree (*Hevea brasiliensis*). Proc. Natl. Acad. Sci. USA **87**:7633-7637.

Broekaert, W.F., Mariën, W., Terras, F.R.G., De Bolle, M.F.C., Proost, P., Van Damme, J., Dillen, L., Claeys, M., Rees, S.B., Vanderleyden, J., and Cammue, B.P.A. (1992). Antimicrobial peptides from *Amaranthus caudatus* seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins. Proc. Natl. Acad. Sci. USA, submitted.

Broekaert, W.F., Van Parijs, J., Leyns, F., Joos, H., and Peumans, W.J. (1989). A chitin-binding lectin from stinging nettle rhizomes with antifungal properties. Science **245**:1100-1102.

Broglie, K., Chet, I., Holliday, M., Cressman, R., Biddle, P., Knowlton, S., Mauvais, C.J., and Broglie, R. (1991). Transgenic plants with enhanced resistance to the fungal pathogen *Rhizoctonia solani*. Science **254**:1194-1197.

Broglie, K.E., Gaynor, J.J., and Broglie, R.M. (1986). Ethylene-regulated gene expression: Molecular cloning of the genes encoding an endochitinase from *Phaseolus vulgaris*. Proc. Natl. Acad. Sci. USA **83**:6820-6824.

Cammue, B.P.A., Broekaert, W.F., Kellens, J.T.C., Raikhel, N.V., and Peumans, W.J. (1989). Stress-induced accumulation of wheat germ agglutinin and abscisic acid in roots of wheat seedlings. Plant Physiol. **91**:1432-1435.

Cammue, B.P.A., Broekaert, W.F., and Peumans, W.J. (1990). Wheat germ agglutinin in wheat seedling roots: induction by elicitors and fungi. Plant Cell Reports **9**:264-267.

Casalongué, C. and Pont Lezica, R. (1985). Potato lectin: A cell-wall glycoprotein. Plant Cell Physiol. **26**:1533-1539.

Chapot, M.P., Peumans, W.J., and Strosberg, A.D. (1986). Extensive homologies between lectins from non-leguminous plants. FEBS Lett. **195**:231-234.

Chrispeels, M.J. and Raikhel, N.V. (1991). Lectins, lectin genes, and their role in plant defense. Plant Cell **3**:1-9.

Czapla, T.H. and Lang, B.A. (1991). Effects of plant lectins on the larval development of European corn borer (Lepidoptera: Pyralidae) and Southern corn rootworm (Coleoptera: Chrysomelidae). J. Econ. Entomol. **83**:2480-2485.

Esau, K. (1960). *Anatomy of seed plants*. John Wiley and Sons, Inc., New York, NY.

Hedrick, S.A., Bell, J.N., Boller, T., and Lamb, C.J. (1988). Chitinase cDNA cloning and mRNA induction by fungal elicitor, wounding, and infection. Plant Physiol. **86**:182-186.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991a). Effects of wheat germ isolectins on development of cowpea weevils. Phytochemistry **30**:785-788.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991b). Rice and stinging nettle lectins: Insecticidal activity similar to wheat germ agglutinin. Phytochemistry **30**:3565-3568.

Lee, H.-I., Broekaert, W.F., Raikhel, N.V. (1991). Co-and post-translational processing of the hevein preproprotein of latex of the rubber tree (*Hevea brasiliensis*). J. Biol. Chem. **266**:15944-15948.

Legrand, M., Kauffmann, S., Geoffroy, P., and Fritig, B. (1987). Biological function of pathogenesis-related proteins: Four tobacco pathogenesis-related proteins are chitinases. Proc. Natl. Acad. Sci. USA **84**:6750-6754.

Lerner, D.R. and Raikhel, N.V. (1989). Cloning and characterization of root-specific barley lectin. Plant Physiol. **91**:124-129.

Lerner, D.R. and Raikhel, N.V. (1992). The gene for stinging nettle lectin (*Urtica dioica* agglutinin) encodes both a lectin and a chitinase. J. Biol. Chem. submitted.

Linthorst, H.J.M. (1991). Pathogenesis-related proteins of plants. Crit. Rev. Plant Sci. **10**:123-150.

Linthorst, H.J.M., Danhash, N., Brederode, F.T., Van Kan, J.A.L., De Wit, P.J.G.M., and Bol, J.F. (1991). Tobacco and tomato PR proteins homologous to *win* and pro-hevein lack the "hevein" domain. Mol. Plant-Microbe Interact. **4**:586-592.

Lotan, T., Ori, N., and Fluhr, R. (1989). Pathogenesis related proteins are developmentally regulated in tobacco flowers. Plant Cell **1**:881-887.

Mauch, F., Mauch-Mani, B., and Boller, T. (1988). Antifungal hydrolases in pea tissues: II. Inhibition of fungal growth by combinations of chitinase and ß-1,3-glucanase. Plant Physiol. **88**:936-942.

Mauch, F. and Staehelin, A. (1989). Functional implications of the subcellular localization of ethylene-induced chitinase and ß-1,3-glucanase in bean leaves. Plant Cell **1**:447-457.

Mansfield, M.A., Peumans, W.J. and Raikhel, N.V. (1988). Wheat-germ agglutinin is synthesized as a glycosylated precursor. Planta **173**:482-489.

Meins, F.Jr. and Ahl, P. (1989). Induction of chitinase and ß-1,3-glucanase in tobacco plants infected with *Pseudomonas tabaci* and *Phytophthora parasitica* var. *nicotianae*. Plant Science **61**:155-161.

Meins, F.Jr., Neuhaus, J.-M., Sperisen, C., and Ryals, J. (1992). The primary structure of plant pathogenesis-related glucanohydrolases and their genes. In: *Genes Involved in Plant Defense* (Boller, T. and Meins, F.Jr., eds). Springer Verlag, Wien-New York. In press.

Mishkind, M.L., Palevitz, B.A., Raikhel, N.V., Keegstra, K. (1983). Localization of wheat germ agglutinin-like lectins in various species of the gramineae. Science **220**:1290-1292.

Money, N. (1990). Measurement of pore size in the hyphal cell wall of *Achlya bisexualis*. Exp. Mycol. **14**:234-242.

Murdock, L.L., Huesing, J.E., Nielsen, S.S., Pratt, R.C., and Shade, R.E. (1990). Biological effects of plant lectins on the cowpea weevil. Phytochemistry **29**:85-89.

Nachbar, M.S., Oppenheim, J.D., and Thomas, J.O. (1980). Lectins in the U.S. diet. Isolation and characterization of a lectin from tomato (*Lycopersicon esculentum*). J. Biol. Chem. **225**:2056-2061.

Neuhaus, J.-M., Ahl-Goy, P., Hinz, U., Flores, S., and Meins, F.Jr. (1991a). High-level expression of a tobacco chitinase gene in *Nicotiana sylvestris*. Susceptibility of transgenic plants to *Cercospora nicotiana* infection. Plant Mol. Biol. **16**:141-151.

Neuhaus, J.-M., Sticher, L., Meins, F.Jr., and Boller, T. (1991b). A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. Proc. Natl. Acad. Sci. USA **88**:10362-10366.

Parsons, T.J., Bradshaw, H.D.Jr., and Gordon, M.P. (1989). Systemic accumulation of specific mRNAs in response to wounding in poplar trees. Proc. Natl. Acad. Sci. USA **86**:7895-7899.

Peumans, W.J., Stinissen, H.M., and Carlier, A.R. (1982). Subunit exchange between lectins from different cereal species. Planta **154**:568-572.

Peumans, W.J., De Ley, M., and Broekaert, W.F. (1984). An unusual lectin from stinging nettle (*Urtica dioica*) rhizomes. FEBS Lett. **177**:99-103.

Raikhel, N.V., Bednarek, S.Y., and Wilkins, T.A. (1988). Cell-type-specific expression of a wheat-germ agglutinin gene in embryos and young seedlings of *Triticum aestivum*. Planta **176**:406-414.

Raikhel, N.V., Hughes, D.W. and Galau, G.A. (1987). An enzyme-immunoassay for quantitative analysis of abscisic acid in wheat. In: *Molecular Biology of Plant Growth Control* (Fox, J.E. and Jacobs, M. eds.), Alan R. Liss, New York, pp 197-207.

Raikhel, N.V., Palevitz, B.A., and Haigler, C.H. (1986). Abscisic acid control of lectin accumulation in wheat seedlings and callus cultures. Plant Physiol. **80**:167-171.

Raikhel, N.V. and Lerner, D.R. (1991). Expression and regulation of lectin genes in cereals and rice. Develop. Genetics. **12**:255-260.

Ra

S

Sl

Sh

Sm

Star

Raikhel, N.V. and Wilkins, T.A. (1987). Isolation and characterization of a cDNA clone encoding wheat germ agglutinin. Proc. Natl. Acad. Sci. USA. **84**:6745-6749.

Schlumbaum, A., Mauch, F., Vögeli, U., and Boller, T. (1986). Plant chitinases are potent inhibitors of fungal growth. Nature **324**:365-367.

Shibuya, N., Goldstein, I.J., Shafer, J.A., Peumans, W.F., and Broekaert, W.F. (1986). Carbohydrate-binding properties of the stinging nettle (*Urtica dioica*) rhizome lectin. Arch. Biochem. Biophys. **249**:215-224.

Shinshi, H., Mohnen, D., and Meins, F.Jr. (1987). Regulation of a plant pathogenesis-related enzyme: Inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. Proc. Natl. Acad. Sci. USA **84**:89-93.

Smith, J.J. and Raikhel, N.V. (1989). Nucleotide sequences of cDNA clones encoding wheat germ agglutinin isolectins A and D. Plant Mol. Bol. **13**:601-603.

Stanford, A., Beven, M., and Northcote, D. (1989). Differential expression within a family of novel wound-induced genes in potato. Mol. Gen. Genet. **215**:200-208.

Stinissen, H.M., Peumans, W.J., and Carlier, A.R. (1983). Two-step processing of *in vivo* synthesized rice lectin. Plant Mol. Biol. **2**:33-40.

Triplett, B.A. and Quatrano, R.S. (1982). Timing, localization, and control of wheat germ agglutinin synthesis in developing wheat embryos. Develop. Biol. **91**:491-496.

Van Parijs, J., Broekaert, W.F, Goldstein, I.J., and Peumans, W.J. (1991). Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. Planta **183**:258-264.

Vögeli-Lange, R., Hansen-Gehri, A., Boller, T., and Meins, F.Jr. (1988). Induction of the defense-related glucanohydrolases, ß-1,3-glucanase and chitinase, by tobacco mosaic virus infection of tobacco leaves. Plant Science **54**:171-176.

Walujono, K., Scholma, R.A., Beintema, J.J., Mariono, A., and Hahn, A.M. (1975). In *Proceedings of the international rubber conference* Vol. 2, pp. 518-531. Rubber Research Institute Malaysia, KuaLa Lumpur.

Wilkins, T.A. and Raikhel, N.V. (1989). Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos. Plant Cell **1**:541-549.

Wright, C.S. (1977). The crystal structure of wheat germ agglutinin at 2.2Å resolution. J. Mol. Biol. **111**:439-457.

Wright, C.S. (1980). Crystallographic elucidation of the saccharide-binding mode on wheat germ agglutinin and its biological significance. J. Mol. Biol. **141**:267-291.

Wright, H.T., Sandrasegaram, G., and Wright, C.S. (1991). Evolution of a family of N-acetylglucosamine binding proteins containing the disulfide rich domain of wheat germ agglutinin. J. Mol. Evol. **33**:283-294.

Zhu, Q. and Lamb, C.J. (1991). Isolation and characterization of a rice gene encoding a basic chitinase. Mol. Gen. Genet. **226**:289-296.

CHAPTER 2

Cloning and Characterization of Root-Specific Barley Lectin

[This chapter is reprinted with permission from:

Lerner, D.R. and Raikhel, N.V. (1989).

Plant Physiology **91**:124-129]

ABSTRACT

Cereal lectins are a class of biochemically and antigenically related proteins localized in a tissue-specific manner in embryos and adult plants. To study the specificity of lectin expression, a barley (*Hordeum vulgare* L.) embryo cDNA library was constructed and a clone (BLc3) for barley lectin was isolated. BLc3 is 972 nucleotides long and includes an open reading frame of 212 amino acids. The deduced amino acid sequence contains a putative signal peptide of 26 amino acid residues followed by a 186 amino acid polypeptide. This polypeptide has 95% sequence identity to the antigenically indistinguishable wheat germ agglutinin isolectin-B (WGA-B) suggesting that BLc3 encodes barley lectin. Further evidence that BLc3 encodes barley lectin was obtained by immunoprecipitation of the *in vitro* translation products of BLc3 RNA transcripts and barley embryo poly A^+ RNA. *In situ* hybridizations with BLc3 showed that barley lectin gene expression is confined to the outermost cell layers of both embryonic and adult roots tips. On Northern blots, BLc3 hybridizes to a 1.0 kilobase mRNA in poly A^+ RNA from both embryos and root tips. We suggest, on the basis of immunoblot experiments, that barley lectin is synthesized as a glycosylated precursor and processed by removal of a portion of the carboxyl terminus including the single N-linked glycosylation site.

INTRODUCTION

Lectins are a class of proteins with very specific carbohydrate binding properties. Many of the plant lectins are well characterized in their sugar binding specificities and, in some cases, the crystalline structure of the protein is known (Etzler, 1985). In spite of this, the biological significance of plant lectins remains elusive (Etzler, 1985). The Gramineae lectins all specifically bind N-acetylglucosamine and are closely related antigenically and biochemically (Stinissen *et al.*, 1983). These lectins are especially interesting because of their unique patterns of expression in specific cell layers of embryonic organs and in the root tips of adult plants (Mishkind *et al.*, 1983). We are investigating how these specific patterns of lectin accumulation are regulated and what this may indicate about possible roles of these proteins within the plant.

Lectins from wheat, barley and rye (cereal lectins) are all dimers with 18 kDa subunits which are synthesized as 23 kDa precursors (Stinissen *et al.*, 1985). Wheat germ agglutinin (WGA) is the best characterized cereal lectin. Wheat (*Triticum aestivum* L.), however, is a hexaploid with each diploid genome contributing an antigenically indistinguishable isolectin (isolectins A, B and D) (Stinissen *et al.*, 1983). Functional dimers of WGA isolectins form *in vivo* by random association of the isolectin monomers (Peumans *et al.*, 1982). Direct sequencing of all three isolectin genes has revealed greater than 90% sequence identity between them (Smith and Raikhel, 1989a). These features of the WGA

system have made molecular and cellular studies of individual isolectin expression particularly difficult.

In this study, we have circumvented the difficulties of the WGA system by studying barley lectin. Barley, a diploid, contains a lectin shown to be antigenically indistinguishable from WGA (Stinissen *et al.*, 1983). The lectins are so similar that active heterodimers containing wheat and barley lectin subunits can be formed *in vitro* (Peumans *et al.*, 1982). Barley lectin accumulates in the embryonic and adult root tips, but unlike in wheat, rye and rice; no lectin is found in the coleoptile (Mishkind *et al.*, 1983). By studying lectin expression in barley, we avoid the possible complications of discerning coleoptile-specific versus root-specific regulatory elements and differential expression of isolectins. A barley lectin cDNA clone, BLc3, was isolated from a barley embryo lambda gt10 library. Using this clone as an *in situ* hybridization probe we localized lectin mRNA in the embryonic and adult root tips. We also present evidence that the barley lectin precursor is glycosylated and undergoes carboxyl terminal processing to produce the mature polypeptide.

MATERIALS AND METHODS

Plant Material

Barley (*Hordeum vulgare* L. var. Betzes) was grown in soil under growth chamber conditions with a 13 h light cycle (440 $\mu E/m^2/s$) or under greenhouse

conditions. Developing grains were harvested at 15 to 25 days post-anthesis (dpa) and stored at -70°C for RNA isolation or used directly for *in situ* hybridization and protein extraction. For isolation of root tips and coleoptiles, barley grains were surface sterilized with 10% commercial bleach for 20 min, rinsed with sterile distilled water and germinated on Whatman #1 filter paper over 0.7% agar for 3 d.

RNA Isolation and Northern Blot Analysis

Total RNA was isolated from developing embryos and root tips (3 to 5 mm) or whole coleoptiles of 3-d-old seedlings by the method of Finkelstein and Crouch (Finkelstein and Crouch, 1986). Polyadenylated RNA (poly A^+ RNA) was purified by oligo-deoxythymidine (oligo-dT) cellulose affinity chromatography using the method of Silflow et al. (Silflow *et al.*, 1979) except that the poly A^+ RNA was eluted at room temperature. For Northern analysis, poly A^+ RNAs were separated electrophoretically, 2 ug per lane, on 2% agarose gels containing 6% formaldehyde and transferred to Immobilon N (Millipore, Bedford, MA) as previously described (Raikhel and Wilkins, 1987). The pre-hybridizations and hybridizations were performed as previously described (Raikhel and Wilkins, 1987) except that the amounts of SDS and salmon sperm DNA were increased to 0.1% and 250 ug/ml respectively. The blots were probed with WGA-B cDNA or BLc3 labeled with α-^{32}P-ATP by the random primers method (Feinberg and Vogelstein, 1983).

Cloning and Sequencing a cDNA for Barley Lectin

Poly A^+ RNA used as a template for cDNA was prepared as described (Mansfield *et al.*, 1988). The poly A^+ RNA was examined for lectin mRNA by Northern analysis using a partial cDNA clone for WGA-B (Raikhel and Wilkins, 1987) as the probe. The presence of full-length, translatable barley lectin mRNA was demonstrated by *in vitro* translation followed by immunoprecipitation with anti-WGA antiserum (Mansfield *et al.*, 1988). The cDNA synthesis reaction was primed with oligo-dT, and the second strand was synthesized using a modification of the Gubler and Hoffman (1983) method with the Bethesda Research Laboratories (Gaithersburg, MD) cDNA Synthesis System. The cDNA was ligated into lambda gt10 (Stratagene, San Diego, CA) with EcoRI linkers (New England Biolabs, Beverly, MA) and packaged *in vitro* using Gigapack Gold (Stratagene). Plaques (5×10^5) were screened with ^{32}P-labeled WGA-B cDNA at low stringency hybridization conditions (Wilkins and Raikhel, 1989) and positive plaques were purified at high stringency (Raikhel *et al.*, 1988) with the same probe. Inserts from purified plaques were subcloned into the EcoRI site of pUC 119 (Vieira and Messing, 1987) and sequenced by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using α-^{35}S-dATP in place of α-^{32}P-dATP and 7-deaza-dGTP in place of dGTP (Mizusawa *et al.*, 1986). The complete sequence of both strands of one clone, designated BLc3, was obtained by sequencing overlapping deletions generated by T4 DNA polymerase (Dale and Arrow, 1987). Sequence analysis was performed with Microgenie (Beckman, Fullerton, CA) and Editbase software (courtesy of N. Nielsen, Purdue Univ., West Lafayette, IN).

In vitro Translations and Immunoprecipitation of BLc3 RNA Transcripts

To generate RNA transcripts, BLc3 was subcloned into the EcoRI site of Bluescript KS^+ (Stratagene). The construct, designated pBsBLc3, was linearized with XhoI or XbaI for sense or antisense RNA transcripts, respectively. For increased efficiency of translation, 'capped' transcripts were generated using an RNA Transcription Kit (Stratagene) according to the manufacturer's protocol with the modifications described below. Capping analog, 0.5 mM m^7GpppG (Pharmacia, Piscataway, NJ), and 0.05 mM rGTP were initially used and 2 aliquots of rGTP were added at 10 min intervals to concentrations of 0.30 mM and 0.55 mM rGTP, respectively.

Two μg of 'capped' sense transcripts or 10 μg barley embryo poly A^+ RNAs were translated in a rabbit reticulocyte lysate (Promega, Madison, WI) using 50 μCi ^{35}S-methionine (Tran^{35}S-label; ICN Biomedicals, Irvine, CA) per reaction. The *in vitro* translation products were immunoprecipitated (Hondred *et al.*, 1987) using anti-WGA antiserum (Mansfield *et al.*, 1988). Samples were carboxyamidated with 2.4 M iodoacetamide at 37°C for 30 min to optimize resolution of the lectins (Raikhel *et al.*, 1984). Translation products were analyzed by SDS-PAGE on 12.5% acrylamide gels and visualized by fluorography.

Analysis of Barley Lectin Synthesized *In Vivo*

Barley embryos (300), 15 to 25 dpa, were isolated onto moistened 3MM paper. Embryos were then incubated in 0.1 mM ABA for 4 h at room temperature to enhance lectin synthesis (Triplett and Quatrano, 1984). Acid soluble protein was

extracted and affinity-purified on immobilized GlcNAc as previously described (Mansfield *et al.*, 1988). Affinity purified lectin, from 100 embryos, was digested with 10 mUnits Endo-β-N-acetylglucaminidase H (Endo H, Calbiochem, San Diego, CA) at 37° C for 18 h. Samples were lyophilized, carboxyamidated, separated on SDS-PAGE, as above, and electroblotted onto nitrocellulose (Towbin *et al.*, 1979). Lectin was detected immunologically with anti-WGA antiserum or anti-WGA-B 172-186, an antiserum specific for the 15 amino acid pro-peptide at the carboxyl terminus of pro-WGA (Smith and Raikhel, 1989b).

In Situ Hybridization

For use as *in situ* hybridization probes, ^{35}S-UTP-labeled sense and antisense RNA transcripts were produced from linearized pBsBLc3. Labeled transcripts were partially hydrolyzed with alkali to an average size of 150 nucleotides for increased efficiency of hybridization to mRNA in the tissue sections. Barley embryos (15 to 25 dpa) and 3-d-old root tips from growing seedlings were cryosectioned to 8 μm and processed as previously described (Raikhel *et al.*, 1988).

RESULTS

Isolation and characterization of barley cDNA clone BLc3.

Eight putative barley lectin clones were isolated from the unamplified barley embryo cDNA library. The 972 nucleotide sequence for one of these clones, designated BLc3 (Figure 1), was determined from overlapping sequential deletions. BLc3 contains a start codon at nucleotides 16-18 initiating a 212 amino acid open reading frame (calculated mol wt = 21,208 D). Amino acid residues -26 to -1 make up a putative signal sequence (figure 1, broken underline). The cleavage site for the signal sequence predicted by the method of von Heijne (1986) matches the amino terminus predicted by sequence identity to mature WGA-B. This putative signal sequence is followed by a 186 amino acid protein with high percentages of Cys (17%) and Gly (22%) and low percentages of His (0.5%), Met (1%), Arg, Ile, Phe, Trp, and Val (1.5% each). A single potential site for Asn-linked glycosylation, Asn-Ser-Thr, is found at residues 206 through 208 (figure 1, marked with asterisks). The deduced amino acid sequence of BLc3 is 95% identical to that of WGA-B. Table 1 lists the amino acid differences between BLc3 and WGA-B. The coding region is followed by two consecutive TGA termination codons (marked with squares) and a 321 nucleotide 3' untranslated region. Four putative polyadenylation signals (Figure 1, underlined) are located at positions 688 and 754 (AATAAT), and at positions 832 and 946 (AATATA). Since an extensive poly A^+ tail is not found however, the exact 3' end of the barley lectin mRNA is unknown.

Figure 1. Nucleotide and deduced amino acid sequence of barley lectin cDNA clone BLc3. The deduced amino acid sequence is from the first methionine residue and numbered along the right margin. The putative signal sequence (broken underline) and carboxyl-terminal extension (double underline) are presumably no present in the mature protein. The single potential asparagine-linked glycosylation site is designated with asterisks. Two stop codons at the end of the coding region are indicated with squares. The four potential polyadenylation signals are underlined. An extensive poly(A^+) tail is not present, so the actual site of polyadenylation is unknown.

```
                    M  K  M  M  S  T  R  A  L  A  L  G  A  A  A  V  L  A  F  A   -7
  1 CAGAAAACAAGAAGGATGAAGATGATGAGCACCAGGGCCCTCGCTCTCGGCGCGGCCGCCGTCCTCGCCTTCGCG

    A  A  T  A  H  A  Q  R  C  G  E  Q  G  S  N  M  E  C  P  N  N  L  C  C  S   19
 76 GCGGCGACCGCGCACGCCCAGAGGTGCGGCGAGCAGGGCAGCAACATGGAGTGCCCCAACAACCTCTGCTGCAGC

    Q  Y  G  Y  C  G  M  G  G  D  Y  C  G  K  G  C  Q  N  G  A  C  Y  T  S  K   44
151 CAGTACGGGTACTGCGGCATGGGCGGCGACTACTGCGGCAAGGGCTGCCAGAACGGCGCCTGCTACACCAGCAAG

    R  C  G  T  Q  A  G  G  K  T  C  P  N  N  H  C  C  S  Q  W  G  Y  C  G  F   69
226 CGCTGCGGCACTCAGGCCGGCGGCAAGACATGCCCTAACAACCACTGCTGCAGCCAGTGGGGTTACTGCGGCTTC

    G  A  E  Y  C  G  A  G  C  Q  G  G  P  C  R  A  D  I  K  C  G  S  Q  A  G   94
301 GGCGCCGAGTACTGCGGCGCCGGCTGCCAGGGCGGCCCCTGCCGCGCCGACATCAAGTGCGGCAGCCAGGCCGGC

    G  K  L  C  P  N  N  L  C  C  S  Q  W  G  Y  C  G  L  G  S  E  F  C  G  E  119
376 GGCAAGCTTTGCCCCAACAACCTCTGCTGCAGCCAGTGGGGTTACTGCGGCCTCGGCTCCGAGTTCTGCGGCGAG

    G  C  Q  G  G  A  C  S  T  D  K  P  C  G  K  A  A  G  G  K  V  C  T  N  N  144
451 GGCTGCCAGGGCGGTGCTTGCAGCACCGACAAGCCGTGCGGCAAGGCCGCCGGCGGCAAAGTTTGCACCAACAAC

    Y  C  C  S  K  W  G  S  C  G  I  G  P  G  Y  C  G  A  G  C  Q  S  G  G  C  169
526 TACTGCTGCAGCAAGTGGGGATCCTGTGGCATCGGCCCGGGCTACTGCGGCGCAGGTTGCCAGAGCGGCGGCTGC
                                  *  *  *
    D  G  V  F  A  E  A  I  A  A  N  S  T  L  V  A  E  ■  ■
601 GACGGTGTCTTCGCCGAGGCCATCGCCGCCAACTCCACTCTTGTCGCAGAATGATGATCTTGCTAATGGCAGTAT

676 TATTGCAACGACGAATAATCCGTGGCAGTTTTGTTGCCACGTACGGTCTCCCTTCACTTACTTTTAGCACTAGTC

751 CTTAATAATTCTCCAGCCTTGCAATATGACGTGCAGGTTGCTACATGCATGGACATATTGCAGTGAGAAGTACTG

826 TGTGGCAATATAGGGTGTACTATTGTTGCCACAAATTTAGTTCTTTCTTGTTACGTACGTACAGTTGTCAGGATG

901 CATGCATCCCCGTTGTAATGTTGGAGTACTCCATGATTTCGTTGCAATATATATTGCCATGAGTCTAAAG
```

Table 1. Differences in deduced amino acid sequence between barley lectin and WGA-B

AMINO ACID POSITION	BLc3	WGA-B
	Conservative substitutions	
41	Tyr	Trp
48	Thr	Ser
64	Trp	Tyr
139	Lys	Arg
179	Ala	Thr
184	Val	Leu
	Non-conservative substitutions	
9	Asn	Gly
66	Tyr	His
123	Gly	Asn
135	Ala	Asp

To verify that BLc3 encodes barley lectin, BLc3 RNA transcripts and barley embryo poly A^+ RNA were each translated *in vitro*. The products were then immunoprecipitated with anti-WGA antiserum and resolved on SDS-PAGE. *In vitro* translation of BLc3 RNA transcripts produced a protein of M_r 21 kDa (figure 2, lane 1). A M_r 21 kDa polypeptide was also specifically immunoprecipitated from *in vitro* translation products of embryo poly A^+ RNA (figure 2, lane 2). These M_r's agree well with the mol wt of 21.2 kDa calculated from the deduced amino acid sequence.

Post-translational Modifications of Barley Lectin

To investigate the *in vivo* synthesis of barley lectin, Western blots of affinity purified lectin from developing barley embryos were probed with anti-WGA antiserum. Affinity-purified barley lectin contained two polypeptides of M_r 18 kDa and M_r 23 kDa (figure 3, lane 1). Mature barley lectin has the same mobility as purified WGA (M_r 18 kDa; figure 3, lane 2). The M_r 23 kDa protein is most likely the barley lectin precursor (Stinissen *et al.*, 1985). *In vivo* labeling studies with barley embryos also show a M_r 23 kDa band after immunoprecipitation with anti-WGA antiserum (data not shown). In addition, pulse labeling studies in wheat have shown that WGA is also synthesized as a M_r 23 kDa precursor (Mansfield *et al.*, 1988). Based on the 95% amino acid sequence identity with WGA-B and the evidence presented above, we will refer to the M_r 23 kDa form as the barley lectin precursor. The barley lectin precursor migrates more slowly, M_r 23 kDa, on SDS-PAGE than predicted from the deduced amino acid sequence alone (21.2 kDa).

FIGURE 2. *In vitro* translation and immunoprecipitation analysis of poly A^+ RNA and BLc3 transcripts. Poly A^+ RNA isolated from 15 to 25 dpa developing embryos (lane 1) and BLc3 RNA transcripts (lane 2) were translated *in vitro* using rabbit reticulocyte lysate and ^{35}S-methionine. Translation products were immunoprecipitated with anti-WGA antiserum, separated on SDS-PAGE and visualized with fluorography. A single product with M_r 21 kDa was immunoprecipitated in each case indicating BLc3 encodes the barley lectin.

21 kD -

Since the polypeptide deduced from the clone BLc3 includes the only potential glycosylation site at the carboxyl-terminus, we investigated whether this glycosylation site was utilized. Affinity-purified protein from developing barley embryos was treated with Endo-ß-N-acetylglucaminidase H. Endo H will specifically cleave high mannose oligosaccharides linked to Asn residues. The smaller size of a protein after Endo H digestion would confirm the presence a high mannose, N-linked glycan. In this experiment we used an antiserum specific for the carboxyl-terminal portion of pro-WGA, anti-WGA-B 172-186 (Smith and Raikhel, 1989b). Binding of anti-WGA-B 172-186 to pro-barley lectin was expected since there are only 2 conservative amino acid differences between the pro-peptide of WGA-B and the last 15 residues encoded by BLc3 (Table I). Anti-WGA-B 172-186 detected the M_r 23 kDa precursor band but failed to recognize mature barley lectin in the sample (figure 3, lane 3). This provides further evidence that the M_r 23 kDa band represents pro-barley lectin. Endo H digestion of affinity-purified barley lectin reduced the size of pro-barley lectin by M_r 3 kDa (fig. 3, lane 4), indicating the presence of a high-mannose oligosaccharide.

Cellular Localization and Temporal Expression of Barley Lectin

The spatial distribution of barley lectin mRNA was determined by *in situ* hybridization with BLc3 antisense RNA transcripts. Barley lectin mRNA was localized to the coleorhiza, outer cell layers of the radicles, and the root caps of the developing embryo (figure 4a and b). Lectin mRNA was also found in the root tip and root cap of 3-d-old seedlings (figure 4c and d). Lectin mRNA was not

FIGURE 3. Western blot analysis of native and Endo H treated barley lectin. Isolated barley embryos, 15 to 25 dpa, were treated with 0.1 mM ABA (4 h) to enhance lectin expression. Barley lectin was affinity purified from acid extracted protein and resolved on SDS-PAGE prior to transfer onto nitrocellulose. Western blots were probed with either anti-WGA antiserum, lanes 1 and 2; or anti-WGA-B 172-186, lanes 3, 4 and 5. Anti-WGA-B 172-186 is an antiserum specific for the pro-peptide of pro-WGA. Barley lectin has a M_r 23 kDa putative precursor and a M_r 18 kDa mature form, lane 1. Commercial WGA, lane 2, contains only the mature lectin, M_r 18 kDa. Anti-WGA-B 172-186 detects only the M_r 23 kDa pro-barley lectin, lane 3. Treatment for 18 h at 37 °C with Endo H changes the M_r of pro-barley lectin to M_r 20 kDa, lane 4. Anti-WGA-B 172-186 does not detect commercial WGA since no pro-WGA is present, lane 5.

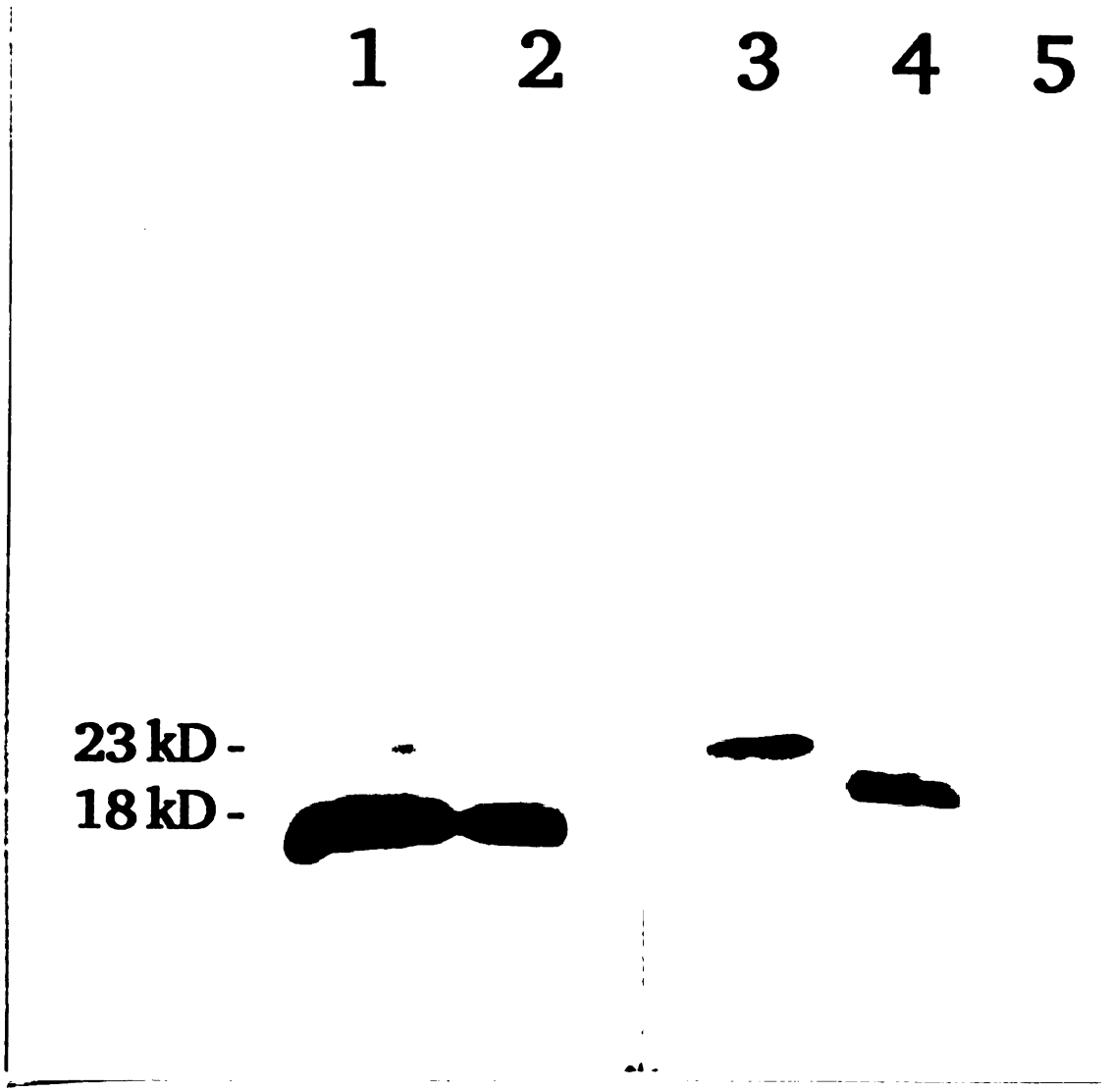
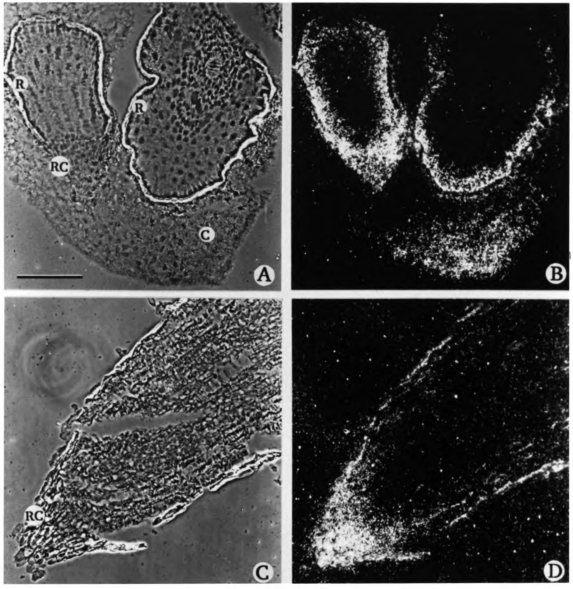

1
2
3
4
5
23 kD -
18 kD -

FIGURE 4. Localization of barley lectin mRNA by *in situ* hybridization. Barley embryos, 15 to 25 dpa, and root tips from 3-d-old seedlings were cryosectioned to 8 μm and probed with BLc3 antisense RNA transcripts. Silver grains developed in the autoradiographic emulsion appear as bright areas with darkfield optics. Phase contrast micrograph of developing embryo, panel A, shows the coleorhiza (C), radicles (R), and embryonic root cap (RC). Darkfield micrograph of the same section, panel B, localizes barley lectin mRNA in the cells of the coleorhiza, the outer cell layer of the radicle and the root cap. Phase contrast, panel C, and Darkfield, panel D, micrographs of root tips from germinating seedlings show specific hybridization of the probe to the root tip and particularly the root cap (RC). Scale bar, 50 μm. Magnification, 400X.


R
R
RC
C
A
B
RC
C
D

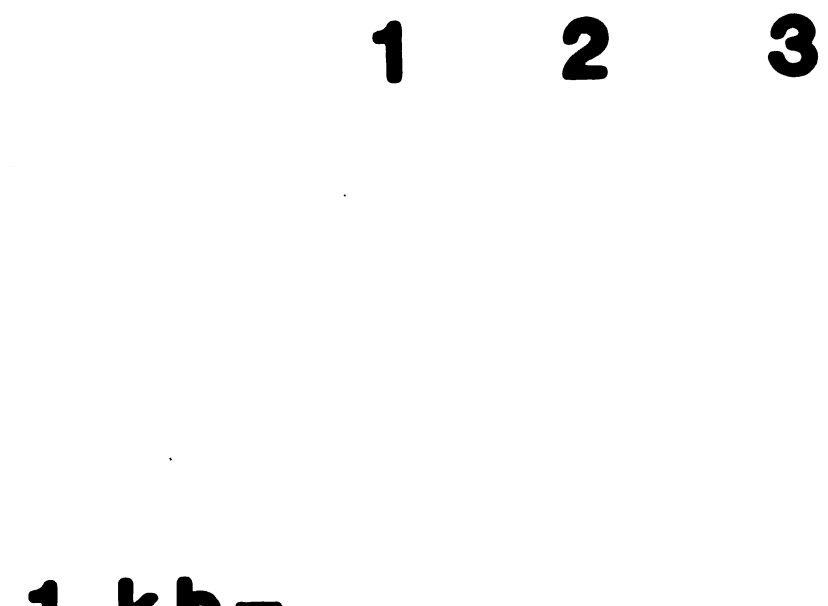
detected in the primordial leaves, coleoptile or scutellum of the embryo (data not shown). Sense BLc3 RNA transcripts, used to monitor non-specific binding of labeled nucleic acids to the sections, did not bind significantly to any tissue (data not shown).

To determine if barley lectin mRNAs from embryos and adult roots were the same size, Northern blot analysis was performed (figure 5). A 1.0 kb mRNA was detected in poly A^+ RNA from both tissues (figure 5, lanes 1 and 2). No detectable lectin mRNA was found in coleoptiles of 3-d-old seedlings (figure 5, lane 3).

DISCUSSION

Our goal is to gain an understanding of the mechanisms controlling the specificity of expression observed in the cereal lectins. Previous work shows that *de novo* synthesis of both lectin mRNA (Raikhel *et al.*, 1988) and protein (Raikhel *et al.*, 1984) is responsible, at least in part, for the pattern of accumulation seen for WGA expression. These data suggest that transcriptional control accounts for some of the observed specificity. The pattern of lectin expression found in cereals is species specific. However, only barley lectin is expressed solely in the adult and embryonic roots. As an initial step in understanding this root-specific expression, a cDNA clone for barley lectin, BLc3, was isolated and characterized.

FIGURE 5. Northern analysis of poly A^+ mRNA from root tips and coleoptiles of 3-d-old seedlings and 15 to 25 dpa barley embryos. Poly A^+ RNA was separated on a formaldehyde/agarose denaturing gel, immobilized on nitrocellulose and hybridized at high stringency with ^{32}P-labeled cDNA clone BLc3. BLc3 hybridizes to a 1.0 kb mRNA from both embryos, lane 1, and root tips, lane 2. No hybridization to coleoptile poly A^+ RNA, lane 3, was observed.


1 2 3
1 kb-

Complementary DNA Clone BLc3 Encodes Barley Lectin

BLc3 was shown to encode barley lectin by *in vitro* translation experiments followed by immunoprecipitation of the products. As shown previously, barley lectin and WGA are immunologically indistinguishable (Stinissen *et al.*, 1983). Thus, anti-WGA antiserum should immunoprecipitate *in vitro* translation products of BLc3. Our results show a M_r 21 kDa polypeptide was immunoprecipitated by anti-WGA antiserum. These data were supported by *in vitro* translation and immunoprecipitation of barley embryo poly A^+ RNA. Here, a single M_r 21 kDa band was also immunoprecipitated by anti-WGA antiserum. The identical M_r of the immunoprecipitated products from both sources indicates that BLc3 probably contains the entire coding region of barley lectin.

Analysis of the amino acid sequence encoded by BLc3 provides further evidence that BLc3 encodes barley lectin. The amino acid composition, rich in Gly and Cys while poor in several other amino acids, is characteristic of the cereal lectins (Peumans *et al.*, 1982). In addition, there were only 10 differences (95% sequence identity) between WGA-B and the deduced amino acid sequence of barley lectin (Table I). Six of these differences were conservative substitutions (Microgenie, Beckman); making the structural similarity even greater. The striking sequence identity found between BLc3 and WGA-B explains the immunological similarity (Stinissen *et al.*, 1983) and the agglutinating activity of WGA/barley lectin heterodimers (Peumans *et al.*, 1982).

The translated sequence of BLc3 is given from the first methionine codon. It is unknown, however, which of the initial methionine residues (-26, -24 or -23) is

used to initiate translation *in vivo*. The coding region of BLc3 begins with a typical tripartite signal sequence (residues -26 to -1) characteristic of secretory proteins. This signal sequence was expected in a full length clone since previous studies have localized cereal and rice lectins to the vacuoles/protein bodies (Mansfield *et al.*, 1988; Mishkind *et al.*, 1982; Stinissen *et al.*, 1985). The predicted cleavage site (von Heijne, 1986) for the signal sequence corresponds exactly to the amino terminus of mature WGA. These data support the hypothesis that Gln #1 is the amino-terminus of the mature barley lectin although the actual terminus is unknown.

Glycosylation and Cleavage of a Propeptide from Probarley Lectin

The results presented in this paper showed that the precursor for barley lectin (M, 23 kDa) is larger than predicted from the cDNA sequence (mol. wt. 21.2 kDa). Pro-barley lectin was found to be Endo H sensitive and therefore glycosylated with a high mannose glycan. This glycan would account for only part of the additional size of the precursor. In these experiments a polyclonal antiserum, anti-WGA-B 172-186, specific for pro-WGA was used (Smith and Raikhel, 1989b). Anti-WGA-B 172-186 specifically recognized pro-barley lectin and deglycosylated pro-barley lectin, but did not bind mature barley lectin. This makes anti-WGA-B 172-186 an especially powerful tool for investigating modifications of the carboxyl-terminal end of barley lectin. The results of these experiments allow us to tentatively assign the carboxyl-terminus of mature barley lectin as Gly #171, although the actual terminal residue is unknown. Furthermore, mature barley lectin

has the same M_r as WGA on SDS-PAGE and the region surrounding the carboxyl-terminus of mature WGA is identical in barley lectin (Table I). Thus, based on our results with anti-WGA-B 172-186 and sequence identity with WGA-B, the carboxyl terminal portion of the barley lectin precursor (double underlined in figure 1) is probable absent in mature barley lectin. WGA (Mansfield *et al.*, 1988), rice lectin (Wilkins and Raikhel, 1989) and β-glucanase (Shinshi *et al.*, 1988), have also been shown to be synthesized as glycosylated precursors and undergo carboxyl-terminal processing of the polypeptide.

Temporal and Cellular Localization of Barley Lectin

In situ hybridization experiments show barley lectin mRNA is localized to the root tip of the adult plant and the analogous structures in the embryo. As might be expected, this pattern of expression coincides with that for lectin accumulation (Mishkind *et al.*, 1983). WGA-B mRNA shows a similar pattern of expression (Raikhel *et al.*, 1988), however, recent data showing greater than 90% identity between wheat isolectin mRNAs (Smith and Raikhel, 1989a) will make precise analysis of individual isolectin expression difficult. Furthermore, the complicated pattern of WGA accumulation in different genotypes of wheat remains unexplained (Raikhel *et al.*, 1986). The barley lectin system, devoid of isolectin complications, is therefore more amenable to the study of root tip-specific protein expression. The cDNA for barley lectin presented in this paper will provide a valuable tool for the isolation of gene promotor sequences for barley lectin and characterization of the cis-elements involved in root-tip-specific expression.

REFERENCES

Dale, R.M.K. and Arrow, A. (1987). A rapid single-stranded cloning, sequencing, insertion, and deletion strategy. Methods Enzymol. **155**:204-214.

Etzler, M.E. (1985). Plant lectins: molecular and biological aspects. Annu. Rev. Plant Physiol. **36**:209-234.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6-13.

Finkelstein, R.R. and Crouch, M.L. (1986). Rapeseed embryo development in culture on high osmoticum is similar to that in seeds. Plant Physiol. **81**:907-912.

Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. Gene **25**: 263-269.

Hondred, D., Wadle, D.-M., Titus, D.E. and Becker, W.M. (1987). Light stimulated accumulation of the peroxisomal enzymes hydroxypyruvate reductase and serine: glyoxylate aminotransferase and their translatable mRNAs in cotyledons of cucumber seedlings. Plant Mol. Biol. **9**:259-275.

Mansfield, M.A., Peumans, W.J. and Raikhel, N.V. (1988). Wheat-germ agglutinin is synthesized as a glycosylated precursor. Planta **173**:482-489.

Mishkind, M.L., Raikhel, N.V., Palevitz, B.A. and Keegstra, K. (1982). Immunocytochemical localization of wheat germ agglutinin in wheat. J. Cell Biol. **92**:753-764.

Mishkind, M.L., Palevitz, B.A., Raikhel, N.V. and Keegstra, K. (1983). Localization of wheat germ agglutinin-like lectins in various species of the gramineae. Science **220**:1290-1292.

Mizusawa, S., Nishimura, S. and Seela, F. (1986). Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucl. Acids Res. **14**:1319-1324.

Peumans, W.J., Stinissen, H.M., Carlier, A.R. (1982). Isolation and partial characterization of wheat-germ-agglutinin-like lectins from rye (Secale cereale) and barley (*Hordeum vulgare*) embryos. Biochem. J. **203**:239-243.

Peumans, W.J., Stinissen, H.M. and Carlier, A.R. (1982a). A genetic basis for the origin of six different isolectins in hexaploid wheat. Planta **154**:562-567.

Peumans, W.J., Stinissen, H.M. and Carlier AR (1982b). Subunit exchange between lectins from different cereal species. Planta **154**:568-572.

Raikhel, N.V., Bednarek, S.Y. and Lerner, D.R. (1988). *In situ* RNA hybridization in plant tissues. In: *Plant Molecular Biology Manual* (Gelvin, S.B. and Schilperoort, R.A., eds), Section B9, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 1-32.

Raikhel, N.V., Bednarek, S.Y. and Wilkins, T.A. (1988). Cell-type-specific expression of a wheat-germ agglutinin gene in embryos and young seedlings of Triticum aestivum. Planta **176**:406-414.

Raikhel, N.V., Mishkind, M.L. and Palevitz, B.A. (1984). Characterization of a wheat germ agglutinin-like lectin from adult wheat plants. Planta **162**:55-61.

Raikhel, N.V., Palevitz, B.A. and Quatrano, R.S. (1986). Pattern of wheat germ agglutinin accumulation in different genotypes of wheat. In: *Lectins* (Bog-Hansen, T.C. and van Driessche, E. eds). vol V, Walter de Gruyter & Co., Berlin, pp 75-81.

Raikhel, N.V. and Wilkins, T.A. (1987). Isolation and characterization of a cDNA clone encoding wheat germ agglutinin. Proc. Natl. Acad. Sci. USA **84**:6745-6749.

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

Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. and Meins, F. (1988). Evidence for N- and C- terminal processing of a plant defence-related enzyme: Primary structure of tobacco prepro-β-1,3-glucanase. Proc. Natl. Acad. Sci. USA **85**:5541-5545.

Silflow, C.D., Hammett, J.R. and Key, J.L. (1979). Sequence complexity of polyadenylated ribonucleic acid from soybean suspension cultured cells. Biochem. **18**:2725-2731.

Smith, J.J. and Raikhel, N.V. (1989a). Tools to study the post-translational modification of WGA. Plant Physiol. **89S**:102.

Smith J.J. and Raikhel, N.V. (1989b). Production of an antibody specific for the propeptide of wheat germ agglutinin. Plant Physiol. **91**:473-476.

Stinissen, H.M., Chrispeels, M.J. and Peumans, W.J. (1985). Biosynthesis of lectin in roots of germinating and adult cereal plants. Planta **164**:278-286.

Stinissen, H.M., Peumans, W.J., Carlier, A.R. (1983). Occurrence and immunological relationships of lectins in gramineous species. Planta **159**:105-111.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350-4354.

Triplett, B.A. and Quatrano, R.S. (1984). Timing, localization and control of wheat germ agglutinin synthesis in developing wheat embryos. Develop. Biol. **91**:491-496.

Vieira, J. and Messing, J. (1987). Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3-11.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucl. Acids. Res. **14**:4683-4690.

Wilkins, T.A. and Raikhel, N.V. (1989). Expression of rice lectin is governed by two temporally and spatially regulated mRNAs in developing embryos and young seedlings. Plant Cell **1**:541-549.

CHAPTER 3

The Gene for Stinging Nettle Lectin (*Urtica dioica* Agglutinin) Encodes both a Lectin and a Chitinase.

[The text and figures in this chapter have been submitted for publication to the *Journal of Biological Chemistry*]

ABSTRACT

Chitin-binding proteins are present in a wide range of plant species, including both monocots and dicots, even though these plants contain no chitin. To investigate the relationship between *in vitro* antifungal and insecticidal activities of chitin-binding proteins and their unknown endogenous functions, the stinging nettle lectin (*Urtica dioica* agglutinin, UDA) cDNA was cloned using a synthetic gene as the probe. The nettle lectin cDNA clone contained an open reading frame encoding 374 amino acids. Analysis of the deduced amino acid sequence revealed a 21 amino acid putative signal sequence and the 86 amino acids encoding the two chitin-binding domains of nettle lectin. These domains were fused to a 263 amino acid carboxyl extension with partial identity to a chitinase catalytic domain. The authenticity of the cDNA clone was confirmed by deduced amino acid sequence identity with sequence data obtained from tryptic digests of the stinging nettle lectin, RNA gel blot and PCR analyses. RNA gel blot analysis also showed the nettle lectin message was present primarily in rhizomes and inflorescence (with immature seeds) but not in leaves or stems. Chitinase enzymatic activity was found when the chitinase-like domain alone or the chitinase-like domain with the chitin-binding domains were expressed in *E. coli*. This is the first example of a chitin-binding protein with both a duplication of the 43 amino acid chitin-binding domain and a fusion of the chitin-binding domains to a structurally unrelated domain, the chitinase domain.

INTRODUCTION

Chitin, a polymer of ß-1,4-N-acetylglucosamine, is found in the cell wall of many fungi, the exoskeleton and digestive tract of some insects and in some nematodes. It is curious then that plants, which contain no chitin, express a family of chitin-binding proteins with a conserved chitin-binding domain. These plants include both monocots and dicots ranging from wheat and barley to tobacco and rubber trees. Chitin-binding proteins are secretory proteins which may be involved in plant defense (see Chrispeels and Raikhel, 1991 for review). Several genes and proteins from this family have been isolated and characterized. Some of them, for example the Gramineae lectins, have been shown to have insecticidal activity by *in vitro* experiments (Huesing *et al.*, 1991a). Others, such as the class I chitinases and hevein, possess antifungal activity in *in vitro* experiments (Van Parijs *et al.*, 1991). Recently, a chitin-binding lectin isolated from the rhizomes of stinging nettle (*Urtica dioica*), *Urtica dioica* agglutinin (UDA), has been shown to possess both antifungal and insecticidal activities (Broekaert *et al.*, 1989 and Huesing *et al.*, 1991b, respectively).

The rhizomes of stinging nettle serve as underground storage tissues and as a source of vegetative meristems for regeneration of shoots. For both these functions rhizomes would require mechanisms for protection against soil pathogens. Nettle lectin has been found to accumulate to high levels (1 gm/kg) in the rhizomes (Peumans *et al.*, 1984a). This small, 8.5 kDa protein is made up of two 43 amino acid glycine and cysteine rich domains, possesses several

isoforms (Van Damme et al., 1988) and has been shown to have homology to the chitin-binding domains of other proteins (Chapot *et al.*, 1986).

In this paper we report the isolation of a cDNA clone encoding stinging nettle lectin using a synthetic gene as the probe. The cDNA clone was found to encode a putative signal sequence adjacent the two nettle lectin chitin-binding domains in the amino-terminal portion of the deduced amino acid sequence. In addition, a long and unexpected carboxyl-terminal domain with partial identity to a chitinase catalytic domain was observed. The validity of the clone was addressed by several approaches and the predicted properties of both domains were analyzed by expression in *E. coli*.

MATERIALS AND METHODS

Plant Material

Urtica dioica plants were grown under standard greenhouse conditions. *U. dioica* seed was from Dr. W.J. Peumans (Katholieke Univ. Leuven, Belgium) and Egessa (FRG). Rhizomes were generated by cutting stems off at the base of the plant, burying these stems under 1 to 2 cm of soil, and growing them under standard greenhouse conditions. These rhizomes were tested for the presence of stinging nettle lectin by agglutination of trypsin-treated rabbit erythrocytes (Peumans *et al.*, 1984b).

General Molecular Biology Methods

Unless stated otherwise, all general molecular biology methods were done as per *Molecular Cloning: A Laboratory Manual* (Maniatis *et al.* 1982). Chemicals were from Sigma (St. Louis, MO) and restriction enzymes were from Boehringer Mannheim (Indianapolis, IN).

RNA Isolations and RNA Gel Blot Analysis

Rhizome and root tissue was washed free of soil, blotted dry, frozen in liquid N_2 and stored at -80° C. Frozen tissue was ground in liquid N_2, mixed with 2.6 ml/gm extraction buffer [0.2 M Tris-HCl pH 9.0, 0.4 M NaCl, 0.025 M EGTA, 1% SDS, 0.5% PVP-40) and ground further with a polytron for 4 min. Proteinase K (0.5 mg/ml) was added and incubated at 37° C with constant mixing for 60 min. The extract was then mixed with BaCl (0.075 M) and KCl (0.035 M), incubated on ice 15 min, centrifuged at 16,000xg for 10 min at 4° C and the supernatant filtered through Miracloth (Calbiochem, San Diego, CA). RNA in the supernatant was precipitated with 2 M LiCl at -20° C overnight and collected by centrifugation at 5,000xg for 20 min at 4° C. The pellet was resuspended in TE (10 mM Tris-HCl pH 8, 0.1 mM EDTA), the RNA reprecipitated with sodium acetate and ethanol and stored at -20° C. Poly(A)$^+$RNA was isolated from total RNA with oligo-dT sepharose (Silflow *et al.*, 1979). Total RNA from all other tissues was isolated by the method of Nagy *et al.* (1988).

RNA gel blot analysis was performed by separating 20 ug total RNA per lane and 5 ul RNA ladder (BRL, Gaithersburg, MD) as size standards on a 2%

agarose/formaldehyde gel as previously described (Raikhel and Wilkins, 1987). ^{32}P-labelled probes were prepared with random primers (Feinberg and Vogelstein, 1983), the unincorporated nucleotides were removed with Nuctrap columns (Stratagene, San Diego CA). Medium stringency washes were done with 2x SSC with 0.1% SDS at 65° C. High stringency washes were done with 0.2x SSC with 0.1% SDS at 65° C.

Design of the Synthetic Gene for Stinging Nettle Lectin

The amino acid sequence of stinging nettle lectin as determined by Chapot et al. (1985) and Steven Michnick (at Harvard University, personal communication) was used as a template for designing the synthetic gene. Since the synthetic gene was intended for expression in transgenic tomato, codon usage was derived primarily from a consensus of tomato genes (Genebank, version 59, 3/89). Codons from highly conserved amino acid residues found in other chitin-binding proteins were also used. In addition, where four G residues in a row were unavoidable, such as tryptophan (TGG) next to glycine (GGX), the guanidine residues were split between two adjacent oligonucleotides since these sequences are difficult to synthesize accurately (Gait *et al.*, 1980). Complementary pairs of oligonucleotides were designed to contain one 5' and one 3' eight base overhang. This greatly reduced the potential for mispaired ligations compared to pairs with two 5' or 3' overhangs. To facilitate cloning, restriction endonuclease recognition sites were included in the synthetic gene flanking the coding region for stinging nettle lectin. In addition, three extraneous base pairs were added to each end of

the synthetic gene to prevent exonuclease activity during construction from obliterating restriction sites at the extreme ends of the gene.

Synthetic Gene Construction

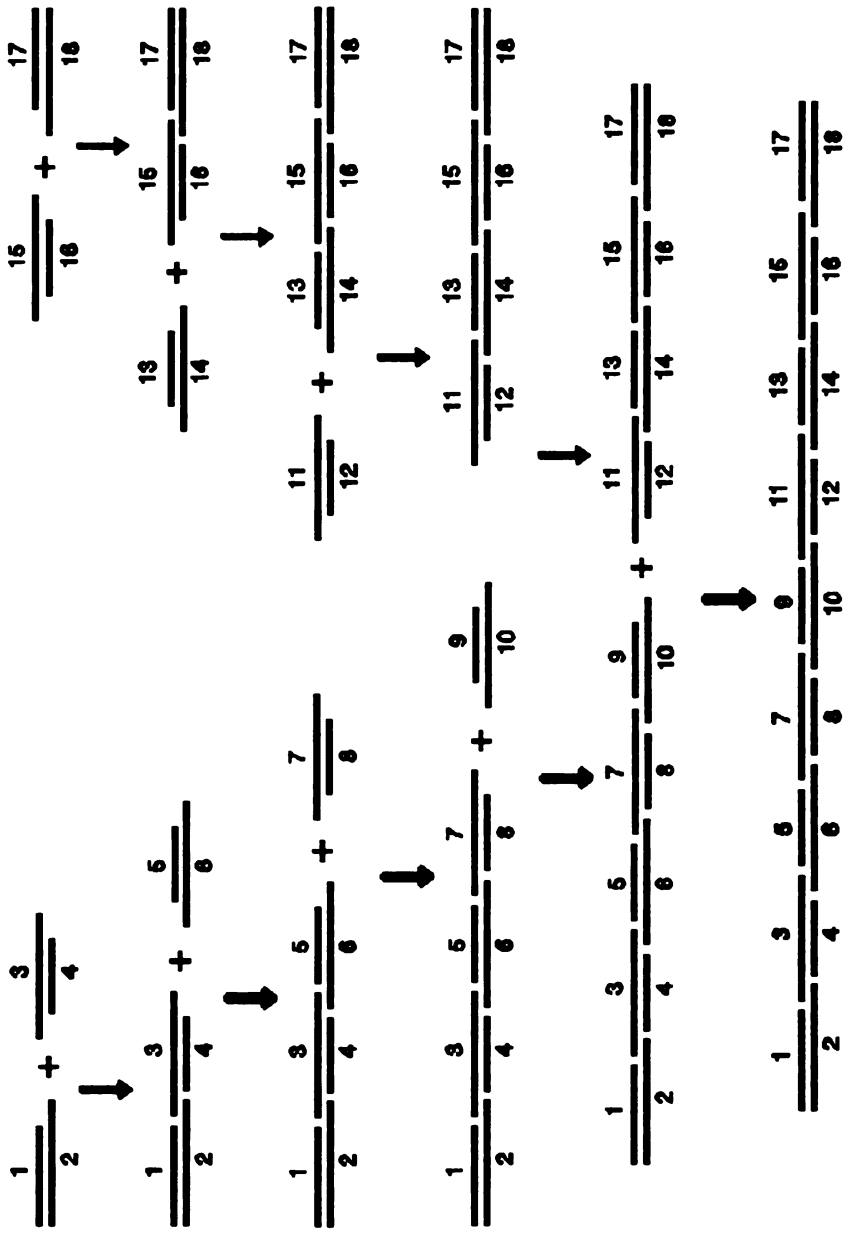
Oligonucleotides were synthesized using β-cyanoethyl chemistry on an Applied Biosystems DNA synthesizer (model 380B, Foster City, CA) by the Macromolecular Structure Facility (Michigan State Univ., East Lansing, MI). Purification of the oligonucleotides was by gel electrophoresis (Sambrook, *et al.* 1989) or HPLC using a Zorbax Bioseries oligo column (Dupont, Wilmington, DE). Purified oligonucleotides were desalted with SEPAK C-18 columns (Waters Associates, Bedford, MA) as recommended by the manufacturer.

One nmol of each purified oligonucleotide, except those at the terminal 5' ends of the synthetic gene, were phosphorylated with T4 polynucleotide kinase (New England Biolabs, Beverly, MA) (Sambrook *et al.*, 1989) and separated from free nucleotides with Nuctrap push columns (Stratagene, La Jolla, CA), using 10 mM Tris, pH 8.0.

Complementary pairs of oligonucleotides, 200 pmol each, were annealed by heating to 80° C then cooling slowly to room temperature. The annealed pairs, 100 pmol each, were then ligated in sequential reactions (Figure 1). Initially, the outermost pairs of annealed oligonucleotides were mixed with the adjacent pairs in separate reactions which included 60 units T_4 DNA ligase (New England Biolabs), 200 mM Tris-HCl (pH 7.6), 50 mM $MgCl_2$, 50 mM dithiothreitol (DTT) and 500 ug/ml bovine serum albumin (BSA) and then incubated for 30 min at 15° C.

Figure 1. Schematic diagram of the sequential ligation reactions used to produce the synthetic stinging nettle lectin gene from 18 synthetic oligonucleotides. Complementary pairs of purified phosphorylated oligonucleotides (ie. #1 + #2, #3 + #4, etc.) were annealed in separate reactions. Annealed pairs of oligonucleotides were sequentially ligated from the ends of the synthetic gene as shown. The full length synthetic gene was restricted at EcoRI sites positioned near the ends of the gene and ligated into pUC119 for sequence analysis.

SYNTHETIC STINGING NETTLE LECTIN GENE LIGATION SCHEME



For subsequent ligation reactions the next adjacent pairs of annealed oligonucleotides were added to the previous reaction along with 30 units T_4 DNA ligase. Tris, $MgCl_2$, DTT, and BSA were also added to bring concentrations to initial values and the reactions incubated for 30 min at 15° C. Finally, the two halves of the synthetic gene were ligated together with an additional 60 units of T_4 DNA ligase for 2.5 h at 15° C. The ligation reactions were stopped by heating to 70° C and the products precipitated with sodium acetate and ethanol. To assess the reaction products 10% of the reaction mix was removed at the end of each ligation step, end labeled with γ-^{32}P-ATP, resolved on a 12.5% non-denaturing polyacrylamide gel, and visualized with autoradiography (data not shown). The synthetic gene was phosphorylated as described above for ligation to the cloning vector, pUC119.

Restriction digestion by EcoRI of the synthetic gene and pUC119 vector were performed using Boehringer Mannheim enzymes following the manufacturers instructions. Vector dephosphorylation reactions were performed with calf intestinal phosphatase (Sigma, St. Louis, MO) for 30 min at 55° C. The synthetic stinging nettle lectin gene (approximately 2.5 pmol) and vector (0.25 pmol) were ligated as above with 100 units ligase for 12 h at 15° C. The ligation products were cloned by standard methods and the sequence of the synthetic gene confirmed by dideoxy chain termination sequencing (Sanger *et al.*, 1977).

Complementary DNA Library Construction and Screening

A stinging nettle rhizome lambda gt10 cDNA library was constructed with *U. dioica* rhizome poly(A)$^+$RNA and EcoRI adapters (Promega, Madison WI) as previously described (Lerner and Raikhel, 1989). 5 x 10^5 plaque forming units (pfu) of the amplified library (1.5 x 10^5 original pfu) were screened at medium stringency (2X SSC + 0.1% SDS at 65 °C) with random-primer labelled synthetic stinging nettle lectin gene. Positive plaques were purified to homogeneity and the lambda DNA isolated. Due to loss of the EcoRI adapter sites during construction of the library, insert DNA was isolated by PCR with primers just flanking the EcoRI site of lambda gt10 (5'-AGCAAGTTCAGCCTGGTTAA-3' and 5'-TTATGAGTTATTTCTTCCAGG-3'). PCR products were phosphorylated with T4 polynucleotide kinase and ligated into the SmaI site of pUC119 for cloning and sequence analysis. The complete sequence of cDNA clones was obtained with dideoxy chain termination sequencing of deletions generated by the method of Dale and Arrow (1987).

Independent isolates of nettle lectin cDNAs were obtained from a separate and unamplified nettle rhizome cDNA library prepared as described above. A random primer ^{32}P-labeled uda1 fragment (nt #267 to 1317) was used to probe approximately 5 x 10^5 plaques. Plaques were purified and insert cDNAs partially sequenced as described above.

Polymerase Chain Reaction DNA Amplification

Amplification reactions were carried out with deoxynucleotides, buffers and enzyme concentrations as recommended by the enzyme manufacturer. Taq polymerase (Amplitaq, Perkin Elmer/Cetus, Norwalk, CT) was used for amplifying inserts from the lambda gt10 library. Replinase (NEN/Dupont, Wilmington, DE) was used for amplification of reverse transcribed RNA (see below). Reactions were carried out on a Perkin-Elmer thermocycler with an initial denaturation step of 94° C for 4 min followed by 30 cycles of 94° C for 1 min, 65° C for 2 min, 72° C for 3 min. A final polymerization step of 72° C for 7 min was added after the 30 cycles. Reaction products were purified from unincorporated nucleotides with PCR purification columns (Qiagen, Studio City, CA). For subcloning, the reaction products were size fractionated on 1% low melting point agarose gels (SeaKem LE, FMC, Rockland, ME).

Reverse Transcription with PCR

First strand cDNA was generated from 1 ug *U. dioica* total rhizome RNA using 10 ng oligo-dT primer and the cDNA synthesis kit from BRL (Gaithersburg, MD) following the manufacturers instructions in a 20 ul reaction. First strand cDNA was then used as the template for the PCR with 5 ul reverse transcription reaction per 25 ul PCR reaction. Primers for the PCR were identical or complementary to the stinging nettle lectin cDNA clone as specified: #1: 5'-TCTGCCGTAGTGAT-CATG-3' (nt #40 to 57), #2: 5'-AGCGGTACTGGCATTTGC-3' (nt #348 to 329),

#3: 5'-ATGGTAGCTGTAGAAGC-3' (nt #495 to 479), #4: 5'-GTCGCAGTACCTCT-TGTA-3' (nt #1044 to 1027).

Southern Blot Analysis

Genomic DNA was isolated by the method of Dellaporta *et al.* (1983) and 40 ug were cleaved with BamHI (16, 80 or 160 units), HindIII (160 units), or XbaI (160 units) using standard reaction conditions for 20 hours. The reaction products were extracted with phenol, phenol/$CHCl_3$/Isoamyl alcohol (25:24:1), $CHCl_3$/isoamyl alcohol (24:1), precipitated with sodium acetate and ethanol (Sambrook *et al.*, 1989) and resuspended in 10mM Tris, pH 8.0 with 0.1 mM EDTA. This digested DNA was redigested, extracted and precipitated as above then size separated on a 0.8% agarose gel and capillary blotted onto transfer membrane (Nitroplus, Micron Separations Inc., Westboro, MA). ^{32}P-labelled uda1 was prepared and used to probe blots as described for RNA gel blots. Blots were washed at high stringency (0.2x SSC + 0.1% SDS at 65° C) and exposed to autoradiographic film for visualization of bound probe.

Sequence Analysis and Comparisons

Sequence analysis and comparisons were performed with MICROGENIE software (Beckman, Fullerton, CA) and EDITBASE software (courtesy of N. Nielsen, Purdue University, West Lafayette, IN). Amino acid sequence of the chitinase genes was deduced from the nucleotide sequence in Genebank. Amino acid sequence alignments were done manually.

Expression of the uda1 Encoded Domains in E. coli

Oligonucleotide directed mutagenesis (Dale and Arrow, 1987) was used to insert restriction sites into the cDNA for convenient cloning into the T7 RNA polymerase *E. coli* expression vector pET3a (Studier *et al.*, 1990). One oligonucleotide (#1: 5'-GGTCTAGTGTCGGCATATGCAGAGGTGCGGAAGC-3') was used to insert an NdeI site at the beginning of the region encoding the mature stinging nettle lectin. A second oligonucleotide (#2: 5'-GCCAGTACCGC-TGCTAACATATGATCGGCAACGTCGTCG-3') was used to insert a stop codon immediately following the sequence encoding mature stinging nettle lectin along with an NdeI site at the beginning of the chitinase-like region. A third oligonucleotide (#3: 5'-TGTTGCGGCGTAAAACATATGCTAGTCCTCCCCAAG-3') was used to place an NdeI site just following the stop codon for the open reading frame. Pairs of these oligonucleotides were then used in separate mutagenesis reactions to construct NdeI bound inserts encoding the mature stinging nettle lectin (#1 and #2), the chitinase-like region (#2 and #3), and the entire open reading frame without the putative signal sequence (#1 and #3). Mutagenized nettle lectin cDNAs were sequenced to confirm mutagenesis and check for errors introduced during mutagenesis. These inserts were ligated into the NdeI site of pET3a for expression of the polypeptides in *E. coli*. *E. coli*, BL21(DE3) (Studier, et al., 1990), cultures containing the constructs were grown to OD_{595} 0.5, induced with 100 uM isopropyl ß-D-thiogalactoside (IPTG) and incubated for 5 h. Cells were collected by centrifugation and resuspended in 1 ml extraction buffer (50 mM Tris (pH 8.0), 1 mM EDTA, and 100 mM NaCl) per gm cells. Crude lysate was

prepared by freezing the cell suspension at -20° C, thawing at 37° C for 20 min., and removing the cell debris by centrifugation at 25000xg for 15 min. (Studier, et al., 1990).

Chitinase Assay

The colorimetric chitinase assay was performed by the method of Wirth and Wolf (1990). Briefly, two parts crude extract, diluted to 1 ug/ul protein with extraction buffer, was mixed with one part 0.2 M sodium acetate (pH 3.6) and one part chitin dye substrate (carboxymethyl-chitin-Remazol Brilliant Violet 5R (CM-chitin-RBV) from Loewe Biochemica GmbH, Otterfing, FRG) and then incubated at RT for 12 h. Undegraded substrate was precipitated with one part 1 M HCl for 10 min on ice and removed by centrifugation at 2500xg for 10 min. Absorbance at 550 nm was used to measure soluble RBV released by the chitinase activity in the extracts. A blank was prepared with extraction buffer in place of the *E. coli* extract.

RESULTS

To further our studies on chitin-binding proteins and their endogenous functions the stinging nettle lectin cDNA was cloned. An initial attempt to use the barley lectin cDNA (Lerner and Raikhel, 1989), another chitin-binding protein, as a heterologous probe was unsuccessful due to insufficient hybridization specificity

as determined by RNA gel blot analysis (data not shown). Additional attempts to use degenerate oligonucleotides as a probe or generate a specific probe by PCR were also unsuccessful. Finally, we decided to synthesize a gene encoding the reported stinging nettle lectin amino acid sequence.

Assembly and Cloning of the Synthetic Stinging Nettle Lectin Gene

The synthetic gene was assembled by sequential ligation reactions of annealed oligonucleotide pairs in two separate reactions starting at each end of the gene. The two half-genes were then ligated together (Figure 1). This strategy was used to limit the number of potential mismatched ligations. Denaturing polyacrylamide gel electrophoresis indicated that both complete and partial ligation products were present at each step. The final ligation reaction products were restricted with EcoRI and ligated into appropriately restricted and dephosphorylated pUC119. Sequence analysis showed that several synthetic gene clones were missing single base pairs but others were accurately synthesized and assembled. One of these correctly synthesized clones was used for further experiments.

Isolation of a cDNA Clone for Stinging Nettle Lectin

The synthetic gene was used as the probe to screen an amplified stinging nettle rhizome cDNA library. Eight positive plaques were purified through successive rounds of screening and their inserts isolated via PCR due to loss of the EcoRI restriction sites during library construction. Isolated inserts were

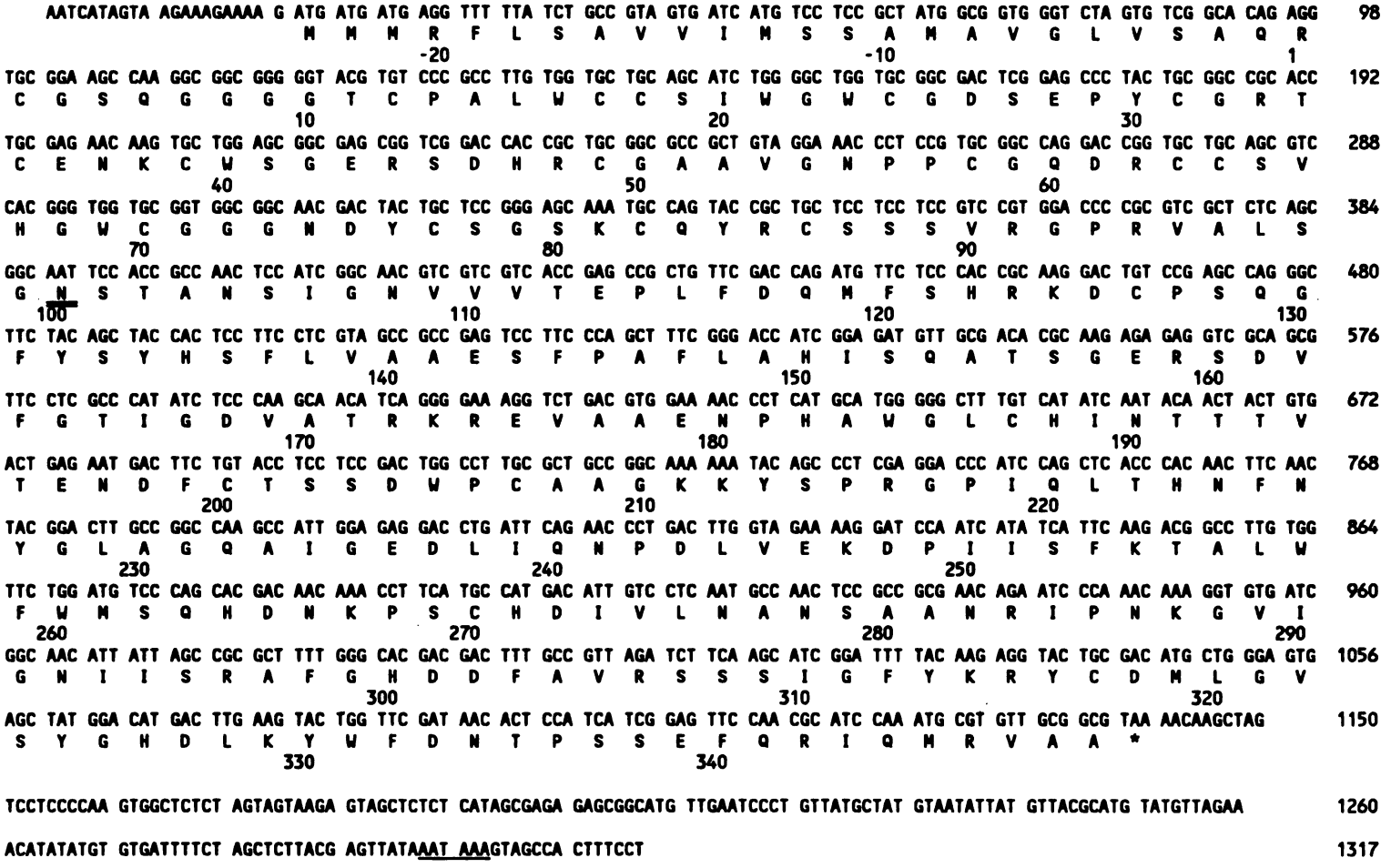
blunt-end ligated into the SmaI site of pUC119 and sequenced. Only one clone, uda1, encoded nettle lectin amino acid sequence and was completely sequenced in both directions. Uda1 was approximately 72% identical to the synthetic gene used as the probe. Although the deduced amino acid sequence for stinging nettle lectin was only 86 amino acids long, uda1 encoded an open reading frame of 372 amino acids. The deduced amino acid sequence included a putative signal peptide (23 amino acids), the two chitin-binding domains of nettle lectin (86 amino acids), a 'spacer' domain (19 amino acids), and a 244 amino acid carboxyl-terminal domain (Figure 2).

Determining the Authenticity of uda1

The unexpected structure of uda1 prompted us to use several methods to verify that the isolated cDNA corresponded to the actual message for stinging nettle lectin. First, RNA gel blot analysis of stinging nettle rhizome total RNA was performed to analyze the hybridization pattern with cDNA clone fragments corresponding to the stinging nettle lectin encoding domain or the carboxyl-terminal domain. A single band of approximately 1.3 kb in size was detected with each probe (data not shown). These results also indicated that a near full length cDNA clone had been isolated since the hybridization was to a message of approximately the same size as the cDNA.

Next, PCR was used to analyze whether the hybridization seen in the RNA gel blot analysis was in fact due to a single mRNA species and not different messages of approximately the same size. Reverse transcribed total rhizome RNA

Figure 2. Nucleotide and deduced amino acid sequence of the stinging nettle lectin cDNA clone uda1. The translation starts at the first methionine of the open reading frame and the potential polyadenylation signal is underlined. Deduced amino acids are listed below the nucleotide sequence and numbered respective to the start of mature stinging nettle lectin. The single potential asparagine-linked glycosylation site in the deduced amino acid sequence (#100) is marked by a double underline.



AATCATAGTA AGAAAGAAAA G ATG ATG ATG AGG TTT TTA TCT GCC GTA GTG ATC ATG TCC TCC GCT ATG GCG GTG GGT CTA GTG TCG GCA CAG AGG 98
M M M R F L S A V V I M S S A M A V G L V S A Q R
-20 -10 1

TGC GGA AGC CAA GGC GGC GGG GGT ACG TGT CCC GCC TTG TGG TGC TGC AGC ATC TGG GGC TGG TGC GGC GAC TCG GAG CCC TAC TGC GGC CGC ACC 192
C G S Q G G G G T C P A L W C C S I W G W C G D S E P Y C G R T
10 20 30

TGC GAG AAC AAG TGC TGG AGC GGC GAG CGG TCG GAC CAC CGC TGC GGC GCC GCT GTA GGA AAC CCT CCG TGC GGC CAG GAC CGG TGC TGC AGC GTC 288
C E N K C W S G E R S D H R C G A A V G N P P C G Q D R C C S V
40 50 60

CAC GGG TGG TGC GGT GGC GGC AAC GAC TAC TGC TCC GGG AGC AAA TGC CAG TAC CGC TGC TCC TCC TCC GTC CGT GGA CCC CGC GTC GCT CTC AGC 384
H G W C G G G N D Y C S G S K C Q Y R C S S S V R G P R V A L S
70 80 90

GGC AAT TCC ACC GCC AAC TCC ATC GGC AAC GTC GTC GTC ACC GAG CCG CTG TTC GAC CAG ATG TTC TCC CAC CGC AAG GAC TGT CCG AGC CAG GGC 480
G N S T A N S I G N V V V T E P L F D Q M F S H R K D C P S Q G
100 110 120 130

TTC TAC AGC TAC CAC TCC TTC CTC GTA GCC GCC GAG TCC TTC CCA GCT TTC GGG ACC ATC GGA GAT GTT GCG ACA CGC AAG AGA GAG GTC GCA GCG 576
F Y S Y H S F L V A A E S F P A F L A H I S Q A T S G E R S D V
140 150 160

TTC CTC GCC CAT ATC TCC CAA GCA ACA TCA GGG GAA AGG TCT GAC GTG GAA AAC CCT CAT GCA TGG GGG CTT TGT CAT ATC AAT ACA ACT ACT GTG 672
F G T I G D V A T R K R E V A A E N P H A W G L C H I N T T T V
170 180 190

ACT GAG AAT GAC TTC TGT ACC TCC TCC GAC TGG CCT TGC GCT GCC GGC AAA AAA TAC AGC CCT CGA GGA CCC ATC CAG CTC ACC CAC AAC TTC AAC 768
T E N D F C T S S D W P C A A G K K Y S P R G P I Q L T H N F N
200 210 220

TAC GGA CTT GCC GGC CAA GCC ATT GGA GAG GAC CTG ATT CAG AAC CCT GAC TTG GTA GAA AAG GAT CCA ATC ATA TCA TTC AAG ACG GCC TTG TGG 864
Y G L A G Q A I G E D L I Q N P D L V E K D P I I S F K T A L W
230 240 250

TTC TGG ATG TCC CAG CAC GAC AAC AAA CCT TCA TGC CAT GAC ATT GTC CTC AAT GCC AAC TCC GCC GCG AAC AGA ATC CCA AAC AAA GGT GTG ATC 960
F W M S Q H D N K P S C H D I V L N A N S A A N R I P N K G V I
260 270 280 290

GGC AAC ATT ATT AGC CGC GCT TTT GGG CAC GAC GAC TTT GCC GTT AGA TCT TCA AGC ATC GGA TTT TAC AAG AGG TAC TGC GAC ATG CTG GGA GTG 1056
G N I I S R A F G H D D F A V R S S S I G F Y K R Y C D M L G V
300 310 320

AGC TAT GGA CAT GAC TTG AAG TAC TGG TTC GAT AAC ACT CCA TCA TCG GAG TTC CAA CGC ATC CAA ATG CGT GTT GCG GCG TAA AACAAGCTAG 1150
S Y G H D L K Y W F D N T P S S E F Q R I Q M R V A A *
330 340

TCCTCCCCAA GTGGCTCTCT AGTAGTAAGA GTAGCTCTCT CATAGCGAGA GAGCGGCATG TTGAATCCCT GTTATGCTAT GTAATATTAT GTTACGCATG TATGTTAGAA 1260

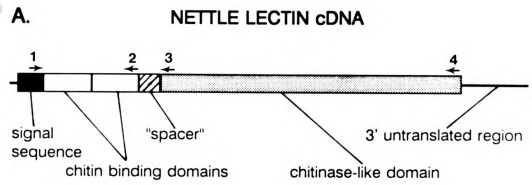
ACATATATGT GTGATTTTCT AGCTCTTACG AGTTATAAAT AAAGTAGCCA CTTTCCT 1317

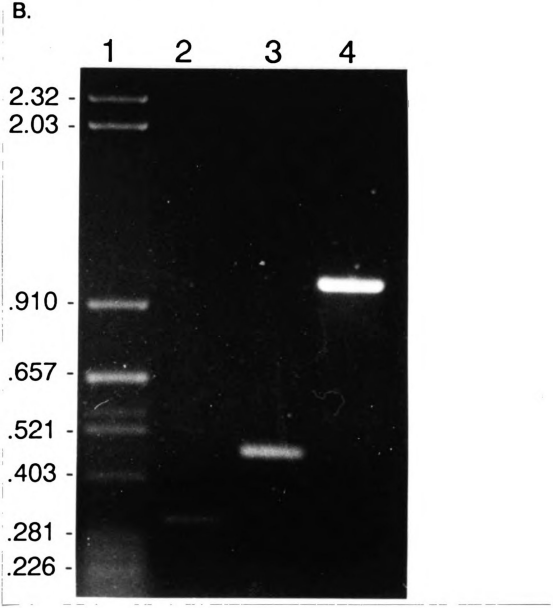
was used as the template for the PCR. Primers for the reactions corresponded to: the beginning of the open reading frame (Figure 3, #1), the end of the stinging nettle lectin encoding domain (Figure 3, #2), the beginning of the carboxyl-terminal domain (Figure 3, #3), and the end of the carboxyl-terminal domain (Figure 3, #4). An elevated annealing temperature (65° C) was used for the reactions to obtain highly specific products. Figure 3 shows that each set of primers produced a single predominant PCR product as detected by an ethidium bromide stained agarose gel. The size of each product was exactly as predicted from the nettle lectin cDNA clone. Primers #1 and #2 generated a 303 bp product, primers #1 and #3 a 451 bp product and primers #1 and #4 a 995 bp product. The PCR products from each reaction were then subcloned into pUC119 and partially sequenced. In all cases the sequence matched that of uda1. In addition, three partial cDNA clones were independently isolated from an unamplified *U. dioica* rhizome cDNA library using uda1 as a probe. Sequence obtained from these clones matched uda1 sequence (data not shown). The amino acid sequence, RNA gel blot, PCR, and clone sequence data all indicated that an authentic nettle lectin cDNA was isolated.

Comparison of uda1 Deduced Amino Acids to Published Amino Acid Sequence of Stinging Nettle Lectin

The deduced amino acid sequence of stinging nettle lectin in uda1 matched 68 of the 72 residues (94%) reported by Chapot *et al.* (1986)(Figure 4). Three of the differences (Ala (52) for Gln, Pro (57) for Thr, and Asp (61) for Leu) were within

Figure 3. PCR analysis using reverse transcribed total RNA from *U. dioica* as the template with primers from uda1. A) Diagram showing positions of the primers on the cDNA clone. B) PCR products were separated by electrophoresis on a 1% agarose gel containing 0.5 ug/ml ethidium bromide and visualized with UV light. Lanes 2, 3, and 4 are total PCR products generated with the primers indicated. HindIII cut lambda DNA with AluI cut pBR322 are the size standards (lane 1).


A.
NETTLE LECTIN cDNA
1
2
3
4
signal sequence
chitin binding domains
"spacer"
chitinase-like domain
3' untranslated region


B.
1
2
3
4
2.32
2.03
.910
.657
.521
.403
.281
.226

```
                    1         10         20         30         40
uda1 (deduced):     QR/CGSQGGGGTCPALW CCSIWGWCGDSEPYCGR/TCENK/CWSGER/
Nettle lectin:      QR/CGSQGGGGTCPALR/CCSIWGWCGDSEPYCG-/      CWSGER/
(tryptic fragments)

                         50         60         70         80
uda1 (deduced):     SDHR/CGAAVGNPPCGQDR/CCSVHGWCGGGNDYCSGSK/CQYR/C
Nettle lectin:           CGAQVGNPTCGQLR/CCSVHGWCGGGNDYC-/   CQYR/
(tryptic fragments)
```

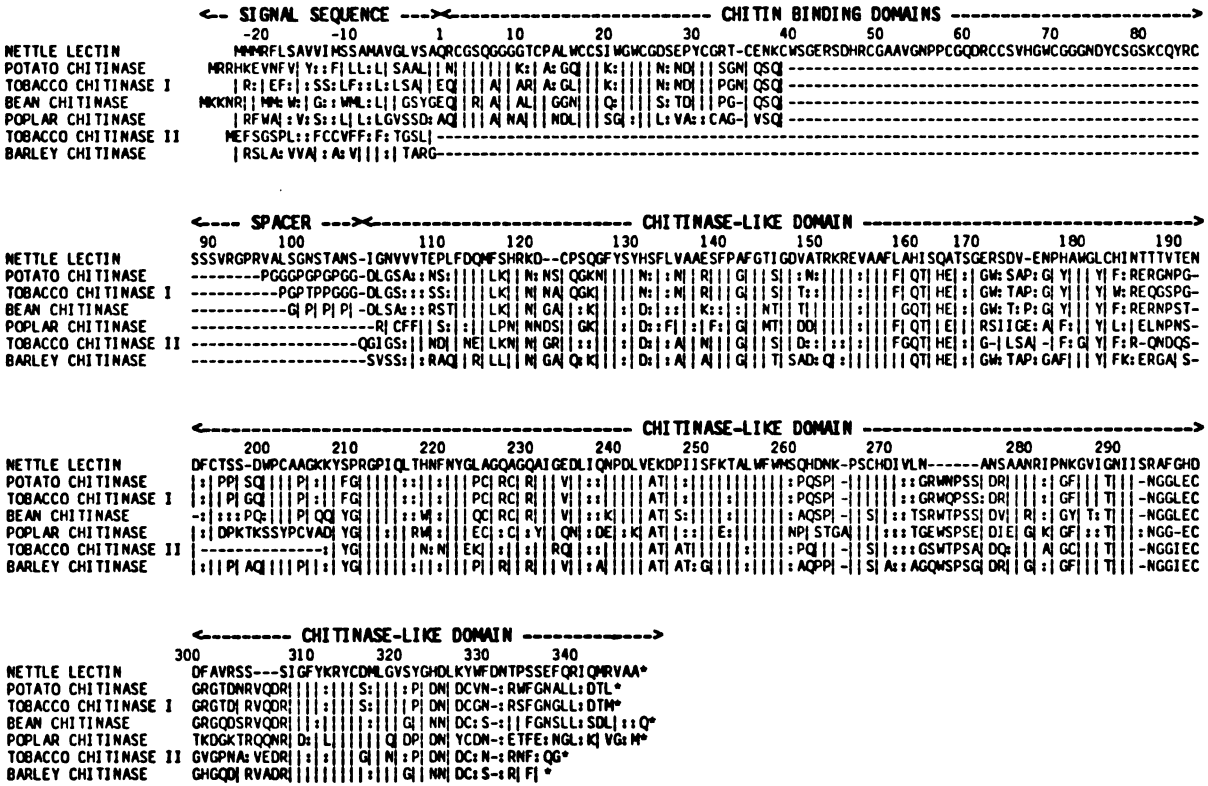
Figure 4. Comparison of stinging nettle lectin amino acid sequences. Alignment of amino acid sequence from Edman degradation of stinging nettle lectin tryptic digests (Chapot *et al.*, 1986) with the deduced amino acid sequence of uda1. Trypsin cleavage sites are indicated with a slash (/). Non-identical amino acids are double underlined. Incompletely sequenced fragments are indicated with a dash (-) at the carboxyl end.

the sequence of a single tryptic fragment. The fourth unmatched amino acid was Trp (16) for Arg in the N-terminal sequence. In addition, 14 residues not identified by peptide sequencing are present in the deduced amino acid sequence encoding stinging nettle lectin. Nine of these residues are predicted to be completely contained within two tryptic fragments which were not sequenced by Chapot *et al.* (1986). The remaining five residues would be present at the carboxyl-terminal ends of tryptic fragments only partially sequenced by Chapot *et al.* (1986).

Comparison of the uda1 Deduced Amino Acid Sequence with Cloned Chitinases

Further analysis of the deduced amino acid sequence from the carboxyl-terminal domain revealed extensive similarity with the deduced amino acid sequence of cloned chitinases (figure 5). This similarity extends beyond the chitin-binding domains to include the chitinase catalytic domain. As expected, stinging nettle lectin contains two chitin-binding domains (86 amino acids) whereas the class I chitinases and hevein possess only a single chitin-binding domain (approximately 43 amino acids). The carboxyl-terminal domain encoded by uda1 has 40-46% identity (approximately 60% with conservative substitutions) with the catalytic domain of the cloned chitinases (Figure 5). The 'spacer' domain between the chitin-binding and chitinase-like domains has no similarity to previously published sequence for stinging nettle lectin or to the deduced amino acid sequence of the cloned class I chitinases. This 'spacer' region does, however, contain the only potential asparagine(N)-linked oligosaccharide recognition site (N-

Figure 5. Comparison of uda1 deduced amino acid sequence with the deduced amino acid sequences of various chitinases. Putative signal sequences, chitin-binding domains, 'spacer' regions and chitinase catalytic domains are indicated. Sequences were aligned manually with gaps to preserve the highest degree of similarity. Amino acids identical to uda1 are denoted with a vertical line (|), conservative amino acid substitutions are denoted with a colon (:) and gaps in the aligned sequences are denoted with a dash (-). Amino acid sequences were derived from nucleotide sequence in the GeneBank database [accession numbers: Potato chitinase, X08011 (Gaynor, 1988); Tobacco chitinase I, M15173 (Shinshi *et al.*, 1987); Bean chitinase, M13968 (Broglie *et al.*, 1986); Poplar chitinase, M25337 (Parsons *et al.*, 1989); Tobacco chitinase II, M29869 (Payne *et al.*, 1990); Barley chitinase, M36989 (Leah, *et al.*, 1991)].


<-- SIGNAL SEQUENCE ---><---------- CHITIN BINDING DOMAINS ---------->
NETTLE LECTIN
POTATO CHITINASE
TOBACCO CHITINASE I
BEAN CHITINASE
POPLAR CHITINASE
TOBACCO CHITINASE II
BARLEY CHITINASE
<---- SPACER ---><---------- CHITINASE-LIKE DOMAIN ---------->
<---------- CHITINASE-LIKE DOMAIN ---------->
<--------- CHITINASE-LIKE DOMAIN --------->

S-T) at deduced amino acid residue #100. It remains to be determined whether the stinging nettle lectin precursor is glycosylated or not.

Tissue Distribution of Stinging Nettle Lectin in U. dioica

Total RNA isolated from rhizomes, lower and upper stem sections, inflorescence with immature seeds, and leaves were examined for the presence of stinging nettle lectin message. Relatively high steady state levels of the 1.3 kb nettle lectin mRNA were detected in the rhizomes and the inflorescence containing immature seeds (Figure 6, lanes 1 and 4). In addition, longer exposures of the autoradiograms revealed much lower levels of stinging nettle lectin message in the RNA isolated from the upper portion of the stem. No hybridization was detected with the RNA samples from the lower stem portion or leaf.

Southern Blot Analysis

Previous studies have reported that stinging nettle lectin in rhizomes of stinging nettle is a complex mixture of isolectins (Van Damme *et al.*, 1988). Southern blot analysis of *U.dioica* genomic DNA using uda1 as a probe was performed to analyze whether the reported isolectins could be due to multiple nettle lectin genes. Figure 7 shows several hybridizing bands of genomic DNA cleaved with BamHI, HindIII and XbaI. Lanes 1, 2 and 3 contain DNA digested with increasing amounts (16, 80 and 160 units) of BamHI. The identical pattern of hybridizing bands indicates that the digestion in lane 1 was complete and that lane 3 contains at least a 10-fold excess of enzyme. The DNA in lanes 4 and 5 were

Figure 6. RNA gel blot analysis of total RNA from *U. dioica* tissues. 20 ug of total RNA (lanes 2-6) was electrophoresed on a 2% agarose/formaldehyde denaturing gel transferred to nitrocellulose, hybridized with ^{32}P-labeled uda1, washed at high stringency (0.2X SSC + 0.1% SDS at 65° C), and visualized with autoradiography. Lanes 2-6, rhizomes, lower stem sections, upper stem sections, inflorescence (with immature seeds) and leaves, respectively.

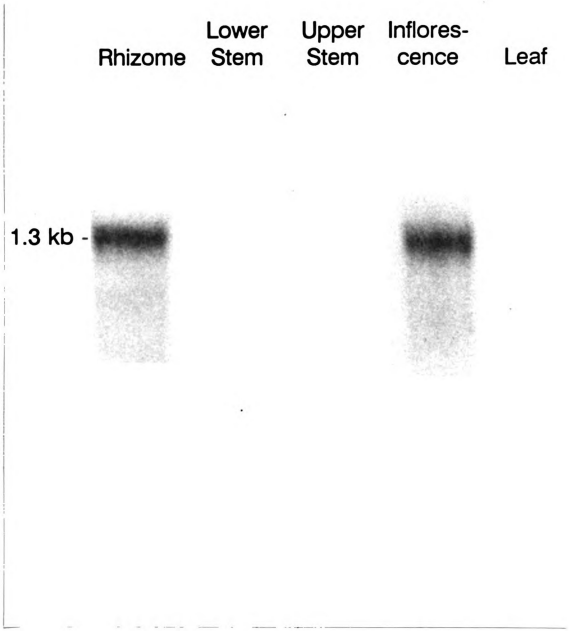
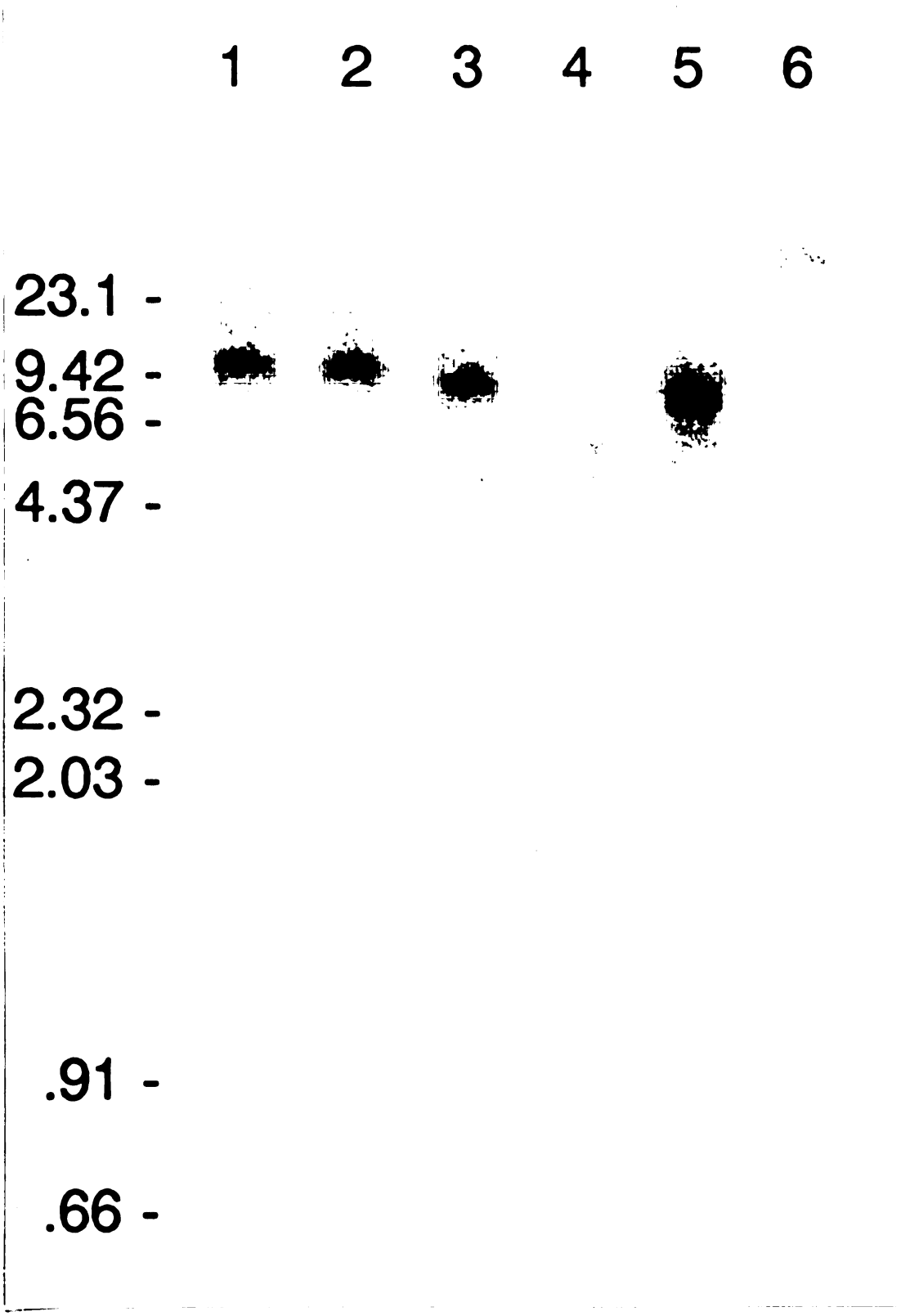

Rhizome
Lower Stem
Upper Stem
Inflores-cence
Leaf
1.3 kb

Figure 7. Southern blot analysis of *U. dioica* genomic DNA. 40 ug genomic DNA was digested two times 20 h with 16 U BamHI (lane 1), 80 U BamHI (lane 2), 160 U BamHI (lane 3), 160 U HindIII (lane 4) or 160 U XbaI (lane 5). 5 ug per lane of digested DNA (lanes 1-5) or undigested DNA (lane 6) was electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose, hybridized with ^{32}P labeled uda1, washed at high stringency (0.2X SSC + 0.1% SDS at 65° C), and visualized with autoradiography.

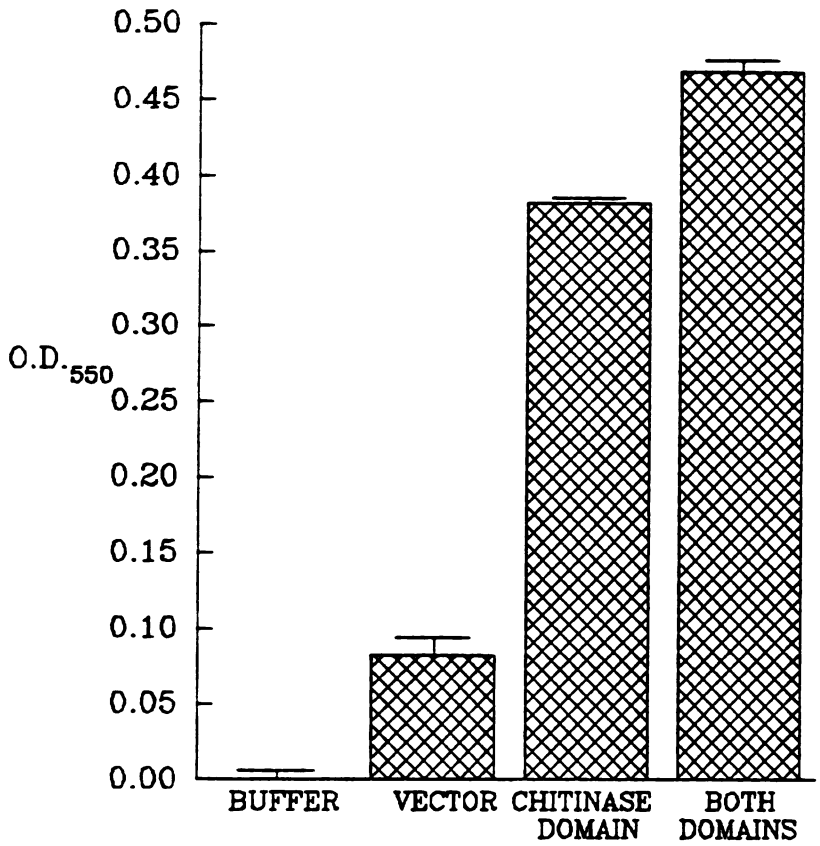

1
2
3
4
5
6
23.1 -
9.42 -
6.56 -
4.37 -
2.32 -
2.03 -
.91 -
.66 -

digested with 160 units of HindIII and XbaI respectively. Since uda1 contains one BamHI restriction site but has no sites for HindIII or XbaI and multiple hybridizing bands were seen in each digest, these data suggests that nettle lectin is encoded by a small multigene family. Thus, some of the nettle lectin isolectins found in the rhizomes may be encoded by independent genes rather than the result of posttranslational processing.

Expression of the uda1 Encoded Domains in E. coli

To assess the properties of the domains encoded by uda1, constructs were prepared encoding the chitin-binding domain alone, the chitinase-like domain alone, or the chitin-binding domain with the chitinase-like domain. These constructs were placed under the control of the T7 RNA polymerase which, in turn was controlled by the lacUV5 promotor and induced by the presence of IPTG (Studier, *et al.*, 1990). Expression of the chitinase-like domain or the full coding region, minus the putative signal sequence, in *E. coli* results in chitinase activity in crude cell extracts. Figure 8 shows a comparison of the levels of chitinase activity in the crude cell extracts measured by the colorimetric assay of Wirth and Wolf (1990). Both the chitinase domain alone (Figure 8, #3) and the chitinase domain with the chitin-binding domain (Figure 8, #4) have considerably higher activity per ug extract protein than the vector alone (Figure 8, #2).

Figure 8. Chitinase assay of crude cell extracts from *E. coli* cells expressing uda1 encoded domains. Colorimetric chitinase assays were done with crude cell extracts from IPTG induced *E. coli* cells with the vector pET3a (vector), pET3a containing the chitinase-like encoding domain of uda1 (chitinase domain), or pET3a containing the chitin-binding and the chitinase-like encoding domains of uda1. Optical density readings were normalized to reactions containing extraction buffer instead of the cell extracts. All values are the average of three independent reactions with the error bars representing standard deviations.


0.50
0.45
0.40
0.35
0.30
O.D.550
0.25
0.20
0.15
0.10
0.05
0.00
BUFFER
VECTOR
CHITINASE DOMAIN
BOTH DOMAINS

DISCUSSION

In this paper we have presented the nucleotide and deduced amino acid sequence of a cDNA clone for stinging nettle lectin. Initially, barley lectin cDNA was used as a heterologous probe for stinging nettle rhizome RNA gel blots. Results from these experiments showed that the heterologous probe would not specifically hybridize to a nettle rhizome message. Next, PCR was used in an attempt to generate a specific probe for the stinging nettle lectin gene. This strategy involved synthesizing redundant oligonucleotides encoding all possible codons for several regions of the amino acid sequence. Specific PCR products of the predicted size were obtained from some of these reactions. Subcloning and sequencing these PCR products, however, revealed that with all primer combinations, only primer/primer products had been made. In retrospect it is clear that these results were due to the highly G/C rich nature of the stinging nettle lectin gene. In addition to the base composition of the primers, it is obvious from the deduced amino acid sequence of uda1 that a tryptic fragment, to which some of the primers were made, had been previously arranged incorrectly (Dr. Stephen Michnick, Harvard University). Thus, PCR primers to this region would lead to DNA synthesis away from the opposing primers and result in no amplification.

Stinging nettle lectin is a small, 86 amino acid protein with 72 amino acid residues known from tryptic fragment sequencing (Chapot *et al.*, 1986). Although the complete amino acid sequence was not available, homology between stinging nettle lectin and other chitin-binding proteins and positioning of the tryptic

fragments (Steven Michnick, Harvard University, personal communication) permitted designing a synthetic gene for use as a hybridization probe. This synthetic gene showed a high level of hybridization specificity by RNA gel blot analysis (data not shown) which allowed for screening the cDNA library.

The cDNA clone isolated had an unexpected structure. Along with the anticipated region encoding the two chitin-binding domains of stinging nettle lectin and the putative signal sequence, the open reading frame possessed an additional 263 amino acids. Four lines of evidence show that the cDNA isolated was the authentic stinging nettle lectin gene. First, the deduced amino acids of the stinging nettle lectin encoding domain of uda1 match 68 of 72 amino acids sequenced from stinging nettle lectin tryptic fragments (Chapot *et al.* 1986) (Figure 4). Second, RNA gel blot analysis showed that both the stinging nettle lectin and the chitinase encoding regions hybridized to a single mRNA species similar to the size of uda1 (1300 bp). Third, PCR analysis of reverse transcribed total rhizome RNA shows that the stinging nettle lectin encoding domain and the carboxyl-terminal domain exist on the same mRNA species (Figure 3). Finally, independently isolated stinging nettle lectin cDNA clones from an unamplified library matched the sequence of uda1.

Although the extended structure of the uda1 open reading frame was unexpected, it was not entirely surprising. All of the previously cloned genes for chitin-binding proteins seem to be the result of either chitin-binding domain duplications, as in the Gramineae lectins, or fusions of a single chitin-binding

domain with an unrelated domain, as in hevein and the class I chitinases (see Chrispeels and Raikhel, 1991 for review). The stinging nettle lectin is the first example of a chitin-binding protein gene resulting from both a domain duplication and fusion with an unrelated domain. Differences between the deduced amino acid sequence of uda1 and previous amino acid sequence of stinging nettle lectin tryptic fragments (Chapot *et al.*, 1986) are most probably due to variations between isolectins. Stinging nettle lectin has been shown to have up to seven isoforms (Van Damme *et al.*, 1988). It is also possible that mistakes arose in amino acid sequencing since the majority of the differences lie on a single tryptic fragment.

The chitinase domain of uda1 remains enigmatic both as to its endogenous function and processing. The deduced amino acids for the carboxyl-terminal domain possess approximately 60% similarity with the catalytic domain of other cloned chitinases. These chitinases, however, are over 90% similar to each other. Where studied, it is known that the class I chitinases do not undergo a processing event to release the chitin-binding domain from the chitinase catalytic domain. Presumably, this type of processing event does occur with the uda1 encoded protein since stinging nettle lectin, when isolated from the plant, is a very small protein (8.5 kDa, 86 amino acids) containing only the two chitin-binding domains. Lee *et al.* (1991) have shown that the hevein preproprotein is processed *in vivo* to yield the chitin-binding protein (hevein, 5 kDa) and a polypeptide of unknown function (16 kDa). The Gramineae lectins have also been shown to undergo post-translational processing of the carboxyl-terminal propeptide with 15 to 27 amino

acids removed (see Raikhel and Lerner, 1991 for review). Differences between the uda1 encoded 'spacer' region and the 'hinge' region of the cloned class I chitinases suggests that there may be differences in their functions. The chitinase 'hinge' region tends to be short, 8 to 11 residues, and rich in proline and glycine while the 'spacer' region of the uda1 deduced amino acid sequence was longer, 19 residues, and contained only one proline. Since the class I chitinases have never been shown to undergo a cleavage to remove the chitin-binding domain it is likely that this 'spacer' region in the stinging nettle lectin proprotein contains the recognition site for proteolytic cleavage. Future studies will address processing of the stinging nettle lectin proprotein. Future studies will also include intracellular localization of the stinging nettle lectin and the chitinase domain in stinging nettle (*U. dioica*) and transgenic plants. The possibility of separate targeting signals for the two domains is intriguing. Vacuolar targeting domains have recently been identified for barley lectin and the tobacco class I chitinase (Bednarek *et al.*, 1990, and Neuhaus, *et al.*, 1991b, respectively). In both cases a short carboxyl terminal propeptide (CTPP) has been shown to be necessary and sufficient for targeting secretory proteins to the vacuole (Bednarek and Raikhel, 1991, and Neuhaus, *et al.*, 1991b). Comparison of the 'spacer' region encoded in the nettle lectin cDNA shows no sequence homology with the CTPP of barley lectin except for the presence of an N-linked glycosylation site. The glycosylation site, however, is not required for correct targeting of barley lectin (Wilkins, *et al.*, 1990). In addition, the chitinase domain encoded by uda1 extends beyond the carboxyl-terminus of the secreted class II chitinases. Again this is similar to the CTPP of the vacuolar,

class I, chitinases but no distinct sequence similarity can be found. Expression of uda1 in transgenic plants will provide an experimental system to study targeting of the nettle lectin domains.

Expression of three uda1 constructs in *E. coli* using a T7 RNA polymerase expression system showed that uda1 encodes a protein with chitinase enzymatic activity. Expression of constructs encoding the chitinase domain alone or the chitin-binding domain with the chitinase domain both show marked chitinase activity.

The potential plant defense functions of the chitin-binding proteins presents an intriguing opportunity for enhancement of plant disease resistance by overexpression of these proteins in transgenic plants. It has been shown that increasing the chitinase activity in transgenic plants by overexpression of a chitinase in tobacco had no discernable effect on fungal resistance (Neuhaus *et al.*, 1991a). However, Broekaert *et al.* (1989) has shown that stinging nettle lectin and chitinase have different modes of action for inhibition of fungal growth *in vitro*. In addition, stinging nettle lectin acts synergistically with chitinase for inhibiting fungal growth *in vitro*. These data suggest that expression of stinging nettle lectin in transgenic plants may enhance resistance to fungal pathogens where expression of chitinase did not. In addition, this is an exciting opportunity to analyze whether the *in vitro* synergistic antifungal effect between stinging nettle lectin and chitinase is observed in the transgenic plants.

REFERENCES

Bednarek, S.Y., Wilkins, T.A., Dombrowski, J.E., and Raikhel, N.V. (1990). A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. Plant Cell **2**:1145-1155.

Bednarek, S.Y. and Raikhel, N.V. (1991). The barley lectin carboxyl-terminal propeptide is a vacuolar protein sorting determinant in plants. Plant Cell **3**:1195-1206.

Broekaert, W.F., Lee, H.-I., Kush, A., Chua, N.-H., and Raikhel, N.V. (1990). Wound-induced accumulation of mRNA containing a hevein sequence in laticifers of rubber tree (*Hevea brasiliensis*). Proc. Natl. Acad. Sci. USA **87**:7633-7637.

Broekaert, W.F., Van Parijs, J., Leyns, F., Joos, H., and Peumans, W.J. (1989). A chitin-binding lectin from stinging nettle rhizomes with antifungal properties. Science **245**:1100-1102.

Broglie, K.E., Gaynor, J.J., and Broglie, R.M. (1986). Ethylene-regulated gene expression: Molecular cloning of the genes encoding an endochitinase from *Phaseolus vulgaris*. Proc. Natl. Acad. Sci. USA **83**:6820-6824.

Cammue, B.P.A., Broekaert, W.F., Kellens, J.T.C., Raikhel, N.V., and Peumans, W.J. (1989). Stress-induced accumulation of wheat germ agglutinin and abscisic acid in roots of wheat seedlings. Plant Physiol. **91**:1432-1435.

Chapot, M.P., Peumans, W.J., and Strosberg, A.D. (1986). Extensive homologies between lectins from non-leguminous plants. FEBS Lett. **195**:231-234.

Chrispeels, M.J. and Raikhel, N.V. (1991). Lectins, lectin genes, and their role in plant defense. Plant Cell **3**:1-9.

Dale, R.M.K., and Arrow, A. (1987). A rapid single-stranded cloning, sequencing, insertion, and deletion strategy. Methods Enzymol. **155**:204-214.

Dellaporta, S.L., Wood, J., and Hicks, J.B. (1983). A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. **1**:19-21.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6-13.

Gait, M.J., Popov, S.G., Singh, M., and Titmas, R.C. (1980). Rapid synthesis of oligodeoxyribonucleotides V. Further studies in solid phase synthesis of oligodeoxyribonucleotides through phosphotriester intermediates. Nucl. Acids Symp. Ser. **7**:243-257.

Gaynor, J.J. (1988). Primary structure of an endochitinase mRNA from *Solanum tuberosum*. Nucl. Acids Res. **16**:5210.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991a). Effects of wheat germ isolectins on development of cowpea weevils. Phytochemistry **30**:785-788.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991b). Rice and stinging nettle lectins: Insecticidal activity similar to wheat germ agglutinin. Phytochemistry **30**:3565-3568.

Leah, R., Tommerup, H., Svendsen, I., and Mundy, J. (1991). Biochemical and molecular characterization of three barley seed proteins with antifungal properties. J. Biol. Chem. **266**:1564-1573.

Lee, H.-I., Broekaert, W.F., Raikhel, N.V. (1991). Co-and post-translational processing of the hevein preproprotein of latex of the rubber tree (*Hevea brasiliensis*). J. Biol. Chem. **266**:15944-15948.

Lerner, D.R. and Raikhel, N.V. (1989). Cloning and characterization of root-specific barley lectin. Plant Physiol. **91**:124-129

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 1st Edition. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY.

Mishkind, M.L., Palevitz, B.A., Raikhel, N.V., Keegstra, K. (1983). Localization of wheat germ agglutinin-like lectins in various species of the gramineae. Science **220**:1290-1292.

Nagy, F., Kay, S.A., and Chua, N.-H. (1988). Analysis of gene expression in transgenic plants. Plant Mol. Biol. Manual **B4**:1-29.

Neuhaus, J.-M., Ahl-Goy, P., Hinz, U., Flores, S., and Meins, F.Jr. (1991a). High-level expression of a tobacco chitinase gene in *Nicotiana sylvestris*. Susceptibility of transgenic plants to *Cercospora nicotiana* infection. Plant Mol. Biol. **16**:141-151.

Neuhaus, J.-M., Sticher, L., Meins, F.Jr., and Boller, T. (1991b). A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. Proc. Natl. Acad. Sci. USA **88**:10362-10366.

Parsons, T.J., Bradshaw, H.D.Jr., and Gordon, M.P. (1989). Systemic accumulation of specific mRNAs in response to wounding in poplar trees. Proc. Natl. Acad. Sci. USA **86**:7895-7899.

Payne, G., Ahl, P., Moyer, M., Harper, A., Beck, J., Meins, F., and Ryals, J. (1990). Isolation of complementary DNA clones encoding pathogenesis-related proteins P and Q, two acidic chitinases from tobacco. Proc. Natl. Acad. Sci. USA **87**:98-102.

Peumans, W.J., De Ley, M., and Broekaert, W.F. (1984a). An unusual lectin from stinging nettle (*Urtica dioica*) rhizomes. FEBS Lett. **177**:99-103.

Peumans, W.J., Nsimba-Lubaki, M., Carlier, A.R., and Van Driessche, E. (1984b). A lectin from Bryonia dioica root stocks. Planta **160**:222-228.

Raikhel, N.V. and Lerner, D.R. (1991). Expression and regulation of lectin genes in cereals and rice. Develop. Genetics. **12**:255-260.

Raikhel, N.V. and Wilkins, T.A. (1987). Isolation and characterization of a cDNA clone encoding wheat germ agglutinin. Proc. Natl. Acad. Sci. USA. **84**:6745-6749.

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

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

Shinshi, H., Mohnen, D., and Meins, F.Jr. (1987). Regulation of a plant pathogenesis-related enzyme: Inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. Proc. Natl. Acad. Sci. USA **84**:89-93.

Silflow, C.D., Hammett, J.R. and Key, J.L. (1979). Sequence complexity of polyadenylated ribonucleic acid from soybean suspension cultured cells. Biochem. **18**:2725-2731.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). The use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60-89.

Van Damme, E.J.M., Broekaert, W.F., and Peumans, W.J. (1988). The Urtica dioica agglutinin is a complex mixture of isolectins. Plant Physiol. **86**:598-601.

Van Parijs, J., Broekaert, W.F, Goldstein, I.J., and Peumans, W.J. (1991). Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. Planta **183**:258-264.

Walujono, K., Scholma, R.A., Beintema, J.J., Mariono, A., and Hahn, A.M. (1975) In Proceedings of the International Rubber Conference, Vol. 2. Rubber Research Institute Malaysia, KuaLa Lumpur, pp. 518-531.

Wilkins, T.A., Bednarek, S.Y., and Raikhel, N.V. (1990). Role of propeptide glycan in post-translational processing and transport of barley lectin to vacuoles in transgenic tobacco. Plant Cell **2**:301-313.

Wirth, S.J., Wolf, G.A. (1990). Dye-labelled substrates for the assay and detection of chitinase and lysozyme activity. J. Microbiol. Methods **12**:197-205

CHAPTER 4

Size, Structure and Function Relationships of Chitin-Binding Proteins: Preliminary Studies on Modifying Barley Lectin.

INTRODUCTION

Chitin-binding proteins provide a unique opportunity to examine the structural features responsible for their activity. These proteins have common structural elements (figure 1) and chitin-binding activity but differ in their ability to inhibit the growth of fungi and insect larvae *in vitro* (Table 1). Chitin-binding proteins are related by the presence of a conserved chitin-binding domain. This domain has been duplicated in tandem in the Gramineae lectins and in stinging nettle lectin. The mature Gramineae lectins are dimers with each subunit consisting of four copies of the 43 amino acid chitin-binding domain. Stinging nettle lectin is approximately half the size of the Gramineae lectin monomer consisting of two tandem repeats of the chitin-binding domain. Other chitin-binding proteins, prohevein and the class I chitinases for example, contain only one copy of the chitin-binding domain which has been fused to a structurally unrelated domain.

The Gramineae lectins, especially wheat germ agglutinin (WGA), are the most studied of the chitin-binding proteins. Biochemically and immunologically these proteins are very similar (Peumans *et al.*, 1982a, Mishkind *et al.*, 1983). Subunit exchange experiments, for example, have shown that WGA and barley lectin monomers are interchangeable in the active dimer (Peumans *et al.*, 1982b). Wheat is a hexaploid plant with three distinct diploid genomes. Each of these genomes contributes a separate isolectin which differs in only a few amino acids (Wright and Raikhel, 1989). Barley lectin is very similar to the WGA isolectins also

Figure 1. Gramineae lectins and homologous proteins. A graphical representation of the domain structure of the Gramineae lectins and some related proteins. Open boxes represent the 43 amino acid, cysteine and glycine rich chitin-binding domain found throughout this family of proteins. Grey boxes represent the unrelated domain fused to the chitin-binding domain in hevein and the potato wound inducible mRNAs, win1 and win2. The black boxes represent the chitinase catalytic domain which is fused to the chitin-binding domain in the class 1, basic, chitinases.

GRAMINEAE LECTINS AND HOMOLOGOUS PROTEINS
WGA
BARLEY LECTIN
RICE LECTIN
NETTLE LECTIN
HEVEIN
POTATO WIN1
POTATO WIN2
BEAN CHITINASE
TOBACCO CHITINASE
POPLAR WIN8

Table 1. *In vitro* activities of chitin-binding proteins.

Protein	Antifungal Activity	Insecticidal Activity	Reference
Gramineae lectins	-	+++	Huesing *et al.*, 1991a,b Schlumbaum *et al.*, 1986
Nettle lectin	+++	++	Broekaert *et al.*, 1989 Huesing *et al.*, 1991b
Hevein	++	NA	Van Parijs *et al.*, 1991
Class I chitinases	+++	NA	Schlumbaum *et al.*, 1986

Relative specific activity for inhibition of growth is represented by +++, highly active; ++, moderately active; +, some activity; -, no activity; NA, data not available.

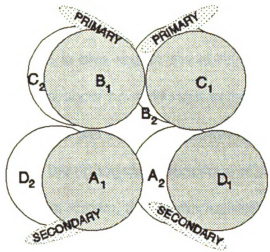
differing by only a few amino acids (Lerner and Raikhel, 1989). Because of the close similarity between the Gramineae lectins, in particular WGA and barley, structural information obtained for WGA is likely to be applicable throughout the Gramineae lectins.

These Gramineae lectins bind N-acetylglucosamine (GlcNAc), GlcNAc oligomers, and chitin, the GlcNAc polymer (for review see Goldstein and Hayes, 1978). The active ligand-binding protein is a dimer with two primary and two secondary binding sites in the cleft between the two subunits (figure 2). Elucidation of the amino acid residues directly involved in binding chitin has been achieved through the crystallization of WGA in the presence of the disaccharide GlcNAc-GlcNAc (Wright, 1980; for review see, Reeke and Becker, 1988). Four amino acids from one subunit (Ser62, Tyr64, His66, Tyr73) and one amino acid (Glu115) from the other subunit interact directly with the disaccharide (Wright, 1980).

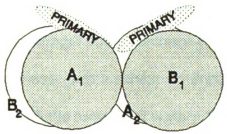
Stinging nettle lectin is similar to the Gramineae lectins in its ability to agglutinate red blood cells indicating the presence of two ligand binding sites. Like the Gramineae lectins, mature stinging nettle lectin consists only of chitin-binding domains. Stinging nettle lectin has been reported to contain up to 85 amino acids when the amino acid composition is examined (Peumans *et al.*, 1984; Van Damme *et al.*, 1988). Only 72 amino acids, however, have been sequenced from tryptic fragments and subsequently aligned (Chapot *et al.*, 1986, and Dr. Steve Michnick, personal communication). Amino acid composition and SDS-PAGE analyses indicate that stinging nettle lectin is approximately 8.5 kDa in size

Figure 2. Graphical representation of mature barley lectin and stinging nettle lectin. The three dimensional structure of barley lectin is known from crystallographic data of WGA (Wright, 1980). The chitin-binding domains are designated A, B, C, D beginning at the amino-terminal of the protein. The domains of the two monomers are designated with a 1 or 2 respectively. The active, carbohydrate binding protein, is a homodimer with the binding sites in the cleft between the two monomers. 'P' and 'S' represent the primary and secondary carbohydrate binding sites, respectively. The structure for stinging nettle lectin is inferred from that of WGA (model 1) and biochemical evidence suggesting that the active stinging nettle lectin is a monomer (model 2)(Peumans *et al.*, 1984).

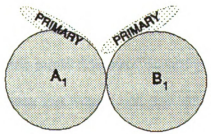
BARLEY LECTIN AND STINGING NETTLE LECTIN 3-DIMENSIONAL STRUCTURE



BARLEY LECTIN



STINGING NETTLE LECTIN
Model 1



STINGING NETTLE LECTIN
Model 2

(Peumans *et al.*, 1984). Alignment of the stinging nettle lectin amino acid sequence with other chitin-binding domains shows that the sequenced amino acids contain the essential elements of two chitin-binding domains in tandem. These include fourteen cysteine residues and amino acids homologous to those involved in binding chitin in the Gramineae lectins. Furthermore, both the Gramineae lectins and stinging nettle lectin can retard the growth of insect larvae *in vitro* (Huesing *et al.*, 1991a and b). Stinging nettle lectin and the Gramineae lectins differ both in size and in the ability to inhibit the growth of fungi. The Gramineae lectins show no antifungal activity (Schlumbaum *et al.*, 1986) while stinging nettle lectin can inhibit the growth of a variety of fungi (Broekaert *et al.*, 1989). Although the endogenous functions of the Gramineae lectins and stinging nettle lectin are not known, the *in vitro* activities suggest that these proteins play a role in plant defense.

The goals of this study were to examine the relationships between protein size and *in vitro* antifungal activity of these chitin-binding proteins. One hypothesis is that the Gramineae lectins exhibit no antifungal activity because they are too large to interact with the fungal cell walls in the same way as stinging nettle lectin. Measurements of fungal cell wall porosity (Money, 1990) indicate that Gramineae lectins, especially as active 36 kDa dimers, will not be able to move through the wall. Stinging nettle lectin however, at half the size of the Gramineae lectins, may be able to permeate the fungal cell wall where it could interact with the chitin to inhibit fungal growth. To test this hypothesis, *in vitro* mutagenesis of the barley lectin cDNA was used to construct a modified barley lectin (mBL) gene encoding

a polypeptide resembling stinging nettle lectin in size. This mBL was shown to retain the ability to bind chitin when expressed in *E. coli* or tobacco.

MATERIALS AND METHODS

Plant Material

Tobacco plants (*Nicotiana tabacum*) were maintained in sterile conditions by bimonthly nodal transfers on Mirashige and Skoog salt mix (BRL/Gibco, Gaithersburg, MD) with 1.5% sucrose and 0.7% phytagar (BRL/Gibco). Plants were grown in GA7 vessels (Magenta Corp., Chicago, IL) in a growth chamber at 28° C with a 16 h light cycle. For generation of seeds from transgenic plants, the GA7 lids were set slightly open to allow to plants to acclimatize to humidity conditions of the air for several days. Plants were then transferred to soil, covered with plastic wrap and grown at RT in the laboratory with a 16 h light cycle for several days before transfer to the greenhouse.

General Molecular Techniques

Except where explicitly stated all molecular biology techniques were done according to Molecular Cloning: A Laboratory Manual 1st (Maniatis *et al.*, 1982) or 2nd (Sambrook *et al.*, 1989) editions. Chemicals were from Sigma Corp. (St. Louis, MO), and restriction enzymes from Boehringer Mannheim Biochemicals (Indianapolis, IN).

Mutagenesis Strategy

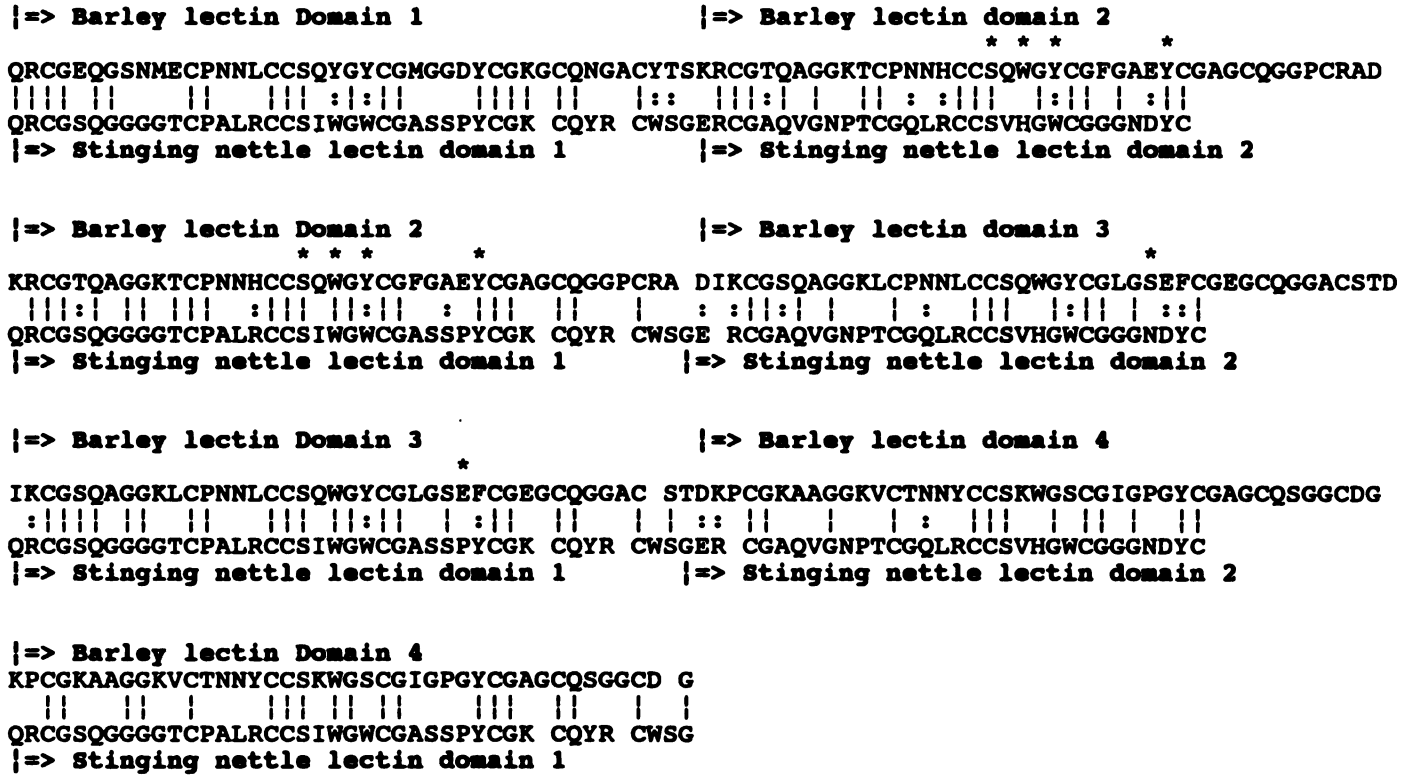
To design the mBL the amino acid sequence of barley lectin was aligned with the known amino acid sequence for stinging nettle lectin (Chapot *et al.*, 1986) (figure 3). Additional information was provided by Dr. Steven Michnick (personal communication) to arrange the tryptic fragment sequences of Chapot. According to the above sources, mature stinging nettle lectin was between 72 and 85 amino acids in length with 72 amino acids of known sequence. This corresponded to approximately two chitin-binding domains of barley lectin. Alignment of the known stinging nettle lectin with the barley lectin amino acid sequence showed the presence of one full and one truncated chitin-binding domains with the truncated domain containing all five amino acids required for chitin-binding. Therefore, mBL was design to contain 72 amino acids corresponding to domain B and approximately two thirds of domain C. These domains were chosen because they contain the primary carbohydrate binding sites of the mature barley lectin (figure 3).

Oligonucleotide Directed Mutagenesis

Barley lectin cDNA clone blc3 (Lerner and Raikhel, 1989), in pUC119, which had previously been altered to add XbaI sites external to the EcoRI sites (Wilkins *et al.*, 1990) was used as the template for the mutagenesis. Mutagenesis was performed using the "MUTA-GENE"™ phagemid *in vitro* mutagenesis kit (BioRad laboratories, Richmond, CA). Four oligonucleotides were used to form single

Figure 3. Alignment of stinging nettle lectin amino acid sequence with the deduced amino acid sequence of barley lectin domain pairs. Identical amino acids are denoted by the | symbol. Conservative amino acid substitutions are designated by a colon. Barley lectin amino acids directly involved in carbohydrate binding are denoted by an asterisk.

COMPARISON OF BARLEY LECTIN AND STINGING NETTLE LECTIN CHITIN-BINDING DOMAINS



deletion mutants of the barley lectin cDNA: 1) 5'-GACCGCGCACGCCCAGCGCTG-CGGCACTCAGG-3' to delete domain A (Δnt #97-225); 2) 5'-GTCGACTCTAGAATT-CCAGCGCTGCGGCACTCAGG-3', to delete the signal sequence and domain A (Δnt #24 to 225); 3) 5'-CGGCTCCGAGTTCTGCGGTGTCTTCGCCGAGG-3', to delete the last third of domain C and domain D (Δnt #445 to 603); and 4) 5'-CGG-CTCCGAGTTCTGCTGATGATCTTGCTAATGGC-3', to delete the last third of domain C, domain D, and the carboxyl terminal propeptide (Δnt #445-651).

Barley Lectin Gene Double Deletion Mutants

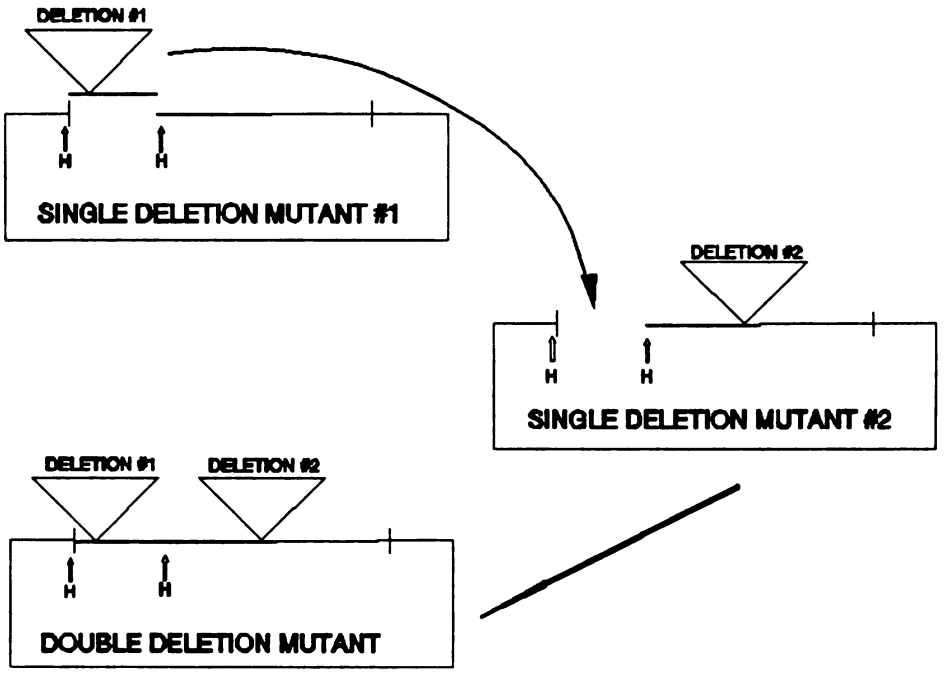
To generate double deletion constructs of barley lectin, HindIII restriction sites within the pUC119 polylinker and at nt #379 were utilized as shown in figure 4. The HindIII fragment from a single deletion construct, made with oligonucleotide #1 or #2 above, was used to replace the unmutagenized HindIII fragment from a single deletion construct made with oligonucleotide #3 or #4. In this way three modified barley lectin (mBL) genes were assembled. For expression in *E. coli*, the mBL gene encoded just domain B and two thirds of domain C (DL111, figure 4). The second mBL gene, for expression in transgenic plants, encoded the barley lectin signal sequence, domain B, two thirds of domain C and the carboxyl terminal extension (DL108, figure 4).

Overexpression in *E. coli*

The mBL gene, DL111, was ligated into the EcoRI site of the maltose binding protein (MBP) expression vector pIH821 (New England Biolabs, Beverly,

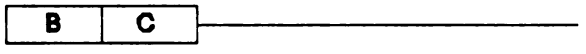
Figure 4. Barley lectin double deletion construction. Single deletion barley lectin constructs were generated by oligonucleotide directed mutagenesis. The HindIII (H) fragment containing a deletion was used to replace the corresponding fragment in the second deletion mutant. Using this general method two mBL genes were constructed. The first, for expression in *E. coli* (DL111) contained the region encoding domain B, two thirds of domain C followed by a stop codon and the 3' untranslated region. The second construct (DL108) was designed for expression in transgenic tobacco. DL108 contained the 5' untranslated region and the region encoding the signal sequence immediately followed by the regions encoding domain B, two thirds of domain C and the 18 amino acid carboxyl-terminal propeptide along with the 3' untranslated region.

BARLEY LECTIN DOUBLE DELETION CONSTRUCTION

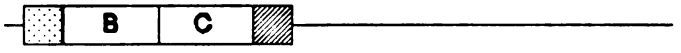


BARLEY LECTIN: UNMODIFIED

DL111: MODIFIED FOR BACTERIAL EXPRESSION



DL108: MODIFIED FOR PLANT EXPRESSION



MA) and cloned into the *E. coli* strain MV1193. Induction, amylose column purification and Factor Xa cleavage of *E. coli* expressed proteins was performed according to New England Biolabs protocols. Briefly, one liter of LB, supplemented with 100 ug/ml ampicillin to select for pIH821 derivatives, was inoculated with 10 ml of a saturated overnight culture. The culture was incubated at 37° C to an optical density at 595 nm ($O.D._{595}$) of 0.3 to 0.4, induced with 100 uM IPTG and incubated at 37° C for 2 h. Cells were collected by centrifugation, resuspended in 50 ml lysis buffer (10 mM sodium phosphate pH 7.2, 30 mM NaCl, 10 mM EDTA, 10 mM EGTA, 0.25% Tween 20), frozen and/or stored at -20° C, thawed at 37° C and sonicated to release protein from the cells. The soluble protein extract was cleared by centrifugation at 9,000 xg for 30 min, mixed 1:1 with amylose column buffer (10 mM sodium phosphate pH 7.2, 0.5 M NaCl, 10 mM ß-mercaptoethanol, and 1 mM EGTA) with 0.25% Tween 20.

Amylose Column Purification of MBP-Fusion Protein

Soluble protein was mixed with approximately 3 ml rehydrated amylose column matrix (New England Biolabs) and incubated 10 to 16 h at 4° C. Unbound protein was removed and saved (effluent fraction). The amylose beads were collected in a 5 ml Quick-Sep column (Isolab, OH), washed with 15 ml column buffer with 0.25% Tween 20 followed by 25 ml column buffer alone. Bound proteins were eluted from the amylose column with 10 mM maltose in 0.8 to 1 ml fractions. Fractions containing protein, as shown by $O.D._{280}$ above baseline, were pooled (eluant fraction), diluted to <200 ug/ml protein with dialysis buffer (100 mM

NaCl, 10 mM Tris-HCl pH 8.0, and 1 mM EGTA) and dialyzed extensively at 4° C against the dialysis buffer. Eluant and Effluent fraction were concentrated in Centricon 30 microconcentrators and stored at 4° C. For release of mBL from the maltose binding protein, fusion protein was mixed with 1.5% (w/w) Factor Xa (New England Biolabs) in 100 mM NaCl, 20 mM Tris-HCl pH 8.0, and 2 mM $CaCl_2$ and incubated 16 h at RT (Factor Xa cleavage experiments were done by Antje Heese, Michigan State University).

Chitin Column Purification of mBL Constructs

Affinity purification on chitin was performed according to Peumans *et al.* (1984). Chitin column matrix, provided by Dr. Willem Broekaert, was resuspended in phosphate buffered saline (PBS). Swollen chitin was washed extensively with PBS, regenerated with 6 M guanidine-HCl with 0.1 M NaOH, and again washed extensively with PBS. Concentrated amylose column eluate containing 50 ug protein was diluted to 1 ml with PBS and applied to a 5 ml column with 250 chitin bed volume. The column was sealed at both ends and rocked 2 h at 4° C. Column effluent containing unbound protein was collected, the column washed with 10 ml PBS and bound protein eluted with 6 M guanidine-HCl with 0.1 M NaOH. Fractions were dialyzed (3500 molecular weight cut-off dialysis membrane) extensively against H_2O, lyophilized and resuspended in SDS-PAGE sample buffer (Chitin-column purification of mBL expressed in *E. coli* was done by Antje Heese, Michigan State University).

Denaturing Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blotting

SDS-PAGE was performed essentially as described (Laemmli, 1970) using 15% acrylamide resolving gels. Samples were boiled 2 min immediately prior to electrophoresis. Semi-dry electrophoretic transfer of protein from the gel onto Immobilon P (Millipore, Bedford, MA) membrane was accomplished with the semi-dry electroblotter (Millipore) according to the manufacturer's recommended procedures. Current density was increased from 10% to the maximum over four steps (10, 30, 60, 90%) 15 min each to enhance transfer of the small proteins (mBL is approximately 8.5 kDa). After transfer, remaining protein binding sites on the transfer membrane were blocked with saturation buffer (3% gelatin and 1% BSA in tris buffered saline, 150 mM NaCl with 20 mM Tris-HCl pH 8) 1 h at RT and 8 h at 4°C. Bound mBL was detected with anti-WGA antiserum and visualized with alkaline phosphatase-linked secondary antibody using nitroblue tetrazolium as the substrate as described (Blake *et al.*, 1984; Mansfield *et al.*, 1988).

Microsequencing mBL Expressed in *E. coli*

For sequencing small quantities of mBL bound to transfer membrane different SDS-PAGE and electroblotting conditions were required. Electrophoresis was done as described above except specially pure, recrystallized SDS (from Dr. J. Leykem, MSU Macromolecular Structure Facility) was used and time allowed for polymerization of the acrylamide gels was increased to ensure its complete polymerization. The resolving gel was polymerized for 10 to 12 h at 4°C and the stacking gel was polymerized for at least 4 h at RT. A CAPS/methanol buffer

system (Matsudaira, 1987) was used in a tank electroblotter (Biorad) for the transfer of protein from the polyacrylamide gel to the Immobilon P transfer membrane. This buffer system was required since residues from the Tris-glycine/methanol normally used interfere with accurate determination of the amino acid sequence. After transfer to immobilon P, the outermost lanes were removed and the mBL in these lanes detected with anti-WGA antiserum as described above. The stained mBL in these lanes was used as a guide for the precise location of the mBL in the unstained lanes. The mBL in the unstained lanes was cut from the remaining membrane and used for sequencing. The amino acid sequence of mBL was determined by Dr. J. Leykem (MSU-Macromolecular Structure Facility) using an Applied Biosystems gas phase sequenator and reverse phase HPLC as described by Hewick et al. (1981) (SDS-PAGE and electroblotting of mBL for microsequencing was done by Antje Heese, Michigan State University).

Expression of mBL in Transgenic Tobacco

Modified barley lectin construct DL108 was ligated into the Xbal site of the plant expression vector pGA643 and transformed into *E. coli* MV1193. This put the mBL gene under control of the CaMV 35S promoter allowing for constitutive expression in the transgenic plant. The plasmid was mobilized to *Agrobacterium tumafaciens* by triparental mating and leaf disk transformations of tobacco were performed as previously described (Wilkins *et al.*, 1990). Transformed regenerated plants were selected for resistance to kanamycin and mBL expressed in the transgenic tobacco was detected by western blot analysis of leaf extract

affinity purified on a chitin column. One gm leaf tissue was ground to a fine powder in liquid N_2, mixed with 2 ml extraction buffer (PBS with 2 mM PMSF and 1% PVP-40) and centrifuged at 15,000 xg 10 min at 4°C. The supernatant was filtered through Miracloth (Calbiochem, La Jolla, CA) and stored on ice. Chitin column affinity chromatography and western blot analysis were performed as described above.

RESULTS

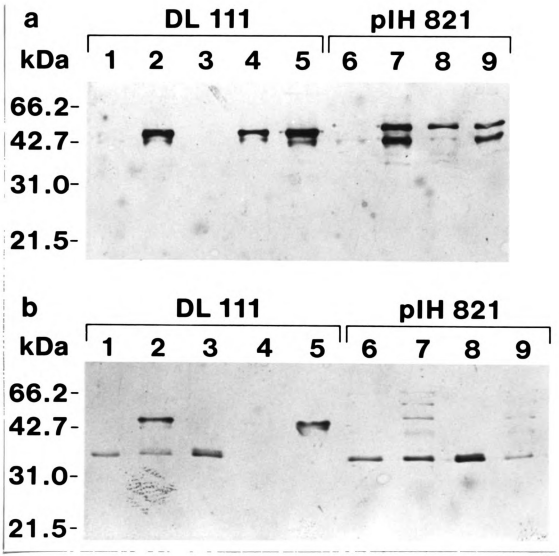
To investigate the relationship between size and *in vitro* activities of chitin-binding proteins, modified barley lectin (mBL) genes were constructed. Mature barley lectin monomers are 18 kDa in size, consisting of four 43 amino acid chitin-binding domains. Stinging nettle lectin consists of approximately two chitin-binding domains containing up to 85 amino acids. The amino acid sequence for only 72 residues of stinging nettle lectin was known at the beginning of this study. Modified barley lectin genes were constructed to encode polypeptides approximating the size of the stinging nettle lectin (figure 2). For expression in *E. coli* the mBL gene DL111 encoded the 72 amino acids from the start of domain B (figure 3). For expression in transgenic tobacco the second mBL gene, DL108, contained the same 72 amino acids along with the BLc3 5' untranslated region for correct initiation of translation and the regions encoding the signal sequence and

carboxyl-terminal propeptide for correct targeting of the mBL to the vacuole (figure 3).

Expression of mBL in E. coli

The mBL gene designed for expression in *E. coli*, DL111, was ligated into the EcoRI site of the maltose binding fusion protein expression vector pIH821. Induction with IPTG of the maltose binding protein-fusion proteins, MBP-lacZα (47 kDa, vector without insert) or MBP-mBL (47 kDa), was detected by western blot analysis of identical SDS-PAGE gels using anti-MBP (figure 5, panel a) or anti-WGA antibodies (figure 5, panel b). As expected anti-MBP recognized both 47 kDa fusion proteins, MBP-mBL (panel a, lane 2) and MBP-lacZα control (panel a, lane 7). Anti-WGA specifically recognized MBP-mBL (panel b, lane 2) but did not bind MBP-lacZα (panel b, lane 7). The 47 kDa fusion proteins were not detected in uninduced cells (panels a and b, lanes 1 and 6). The MBP-mBL and MBP-lacZα fusion proteins were mostly present in the soluble fraction (panels a and b, lanes 4 and 9) although due to technical errors little MBP-mBL was detected with the anti-WGA antiserum on this blot (panel b, lane 4). Some MBP-lacZα was detected in the insoluble fraction (panel a, lane 8) but no insoluble MBP-mBL could be detected (panels a and b, lane 3). The 47 kDa fusion proteins specifically bound to amylose and were eluted from an amylose affinity column with 20 mM maltose (panels a and b, lane 5; MBP-lacZα data not shown). Native MBP (40 kDa) was also detected using anti-MBP antiserum (panel a, lane 5). A cross-reactive *E. coli* protein of approximately 35 kDa along with several other less abundant proteins

Figure 5. Expression of fusion protein in *E. coli*. MBP-mBL (lanes 1-5) or the vector without the mBL, pIH821 (lanes 6-9) was expressed in *E. coli* following induction with IPTG. Immunoblot analysis of identical gels was performed with anti-MBP antiserum (panel a) or anti-WGA antiserum (panel b) and visualized by alkaline phosphatase linked secondary antibodies. Uninduced cells (lanes 1 and 6) show cross reactive *E. coli* proteins. Induced cells (lanes 2 and 7) show the MBP-mBL fusion protein specifically reacts with each antiserum. Lanes 3 and 8 contain the insoluble protein fraction. Lanes 4 and 9 contain the soluble protein fraction which was used for amylose column affinity purification. Affinity purified MBP-mBL (1 ug) eluted from the amylose column with 20 mM maltose is shown in lane 5. All lanes, except #5, contain protein extract from $5x10^6$ cells (figure courtesy of Antje Heese).


a
DL 111
pIH 821
kDa
1 2 3 4 5 6 7 8 9
66.2-
42.7-
31.0-
21.5-
b
DL 111
pIH 821
kDa
1 2 3 4 5 6 7 8 9
66.2-
42.7-
31.0-
21.5-

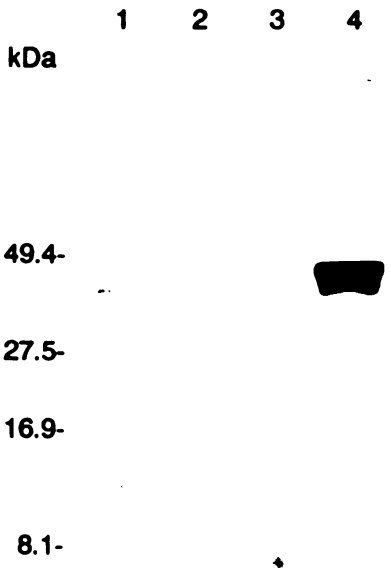
were often detected with anti-WGA but were most insoluble (panel b, lanes 3 and 8) and did not bind to the amylose column (panel b, lane 5).

The 8 kDa mBL could be released from the amylose column purified fusion protein, MBP-mBL, by cleavage with Factor Xa (figure 6). Factor Xa cleaved (lanes 1-3) proteins were affinity purified on chitin columns. Proteins in the column effluent (lane 1), wash (lane 2) and eluate (lane 3) were detected by western blot analysis with anti-WGA antibodies. The released mBL and the MBP-mBL fusion protein were both capable of binding to chitin (data for fusion protein not shown). Chitin column eluate of Factor Xa cleaved protein contains the 8 kDa mBL released from the fusion protein (lane 3). The efficiency of cleavage by Factor Xa, however, was too low to allow for the production of large quantities of mBL. Storage of the fusion protein at 4° C for an extended period (ie. several months) also caused the degradation of the MBP moiety but not the mBL moiety of the fusion protein (data not shown).

To confirm that the 8 kDa protein detected after cleavage with factor Xa was the intact mBL, microsequencing of the amino terminal residues of this protein was performed by the MSU macromolecular structure facility. Five cycles of Edman degradation detected Gln-Arg-Xxx-Gly-Glu at the amino terminal end of the protein. The four identified amino acids confirmed the sequence of the mBL. This technique cannot identify Cys residues which further confirms the mBL amino acid sequence since mBL has Cys as the third residue.

Figure 6. Chitin column affinity purified mBL expressed in *E. coli*. Amylose column purified protein from cells containing the vector (pIH821) with the mBL gene (pDL111) was visualized with anti-WGA antiserum. Protein fractions were digested with Factor Xa before affinity chromatography on chitin columns. The chitin column effluent (lane 1) and wash (lane 2) do not contain the MBP-mBL fusion protein or the mBL released with Factor Xa. The chitin-binding MBP-mBL (47 kDa) and mBL (8 kDa) eluted from the column with 6 M guanidine + 0.1 N NaOH (lane 3). Lane 4 contains 0.75 ug MBP-mBL fusion protein for size comparison.

IMMUNOBLOT ANALYSIS OF THE BARLEY LECTIN CONSTRUCT EXPRESSED IN *E. coli*



Expression of mBL in Tobacco Plants

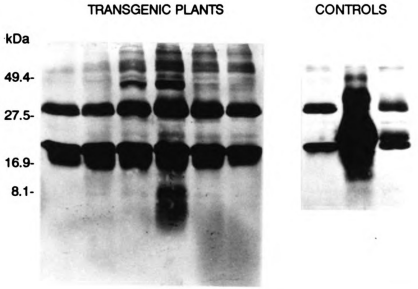
For expression in tobacco plants the mBL gene (DL108) was introduced into pGA643 under control of the constitutive CaMV 35S promotor. This vector also contains the selectable marker nptII which imparts kanamycin resistance to the transformed plants. Six kanamycin resistant plants were screened for the presence of active mBL by immunoblot detection of the 8 kDa protein in chitin column eluate (figure 7). Although a number of abundant cross-reactive chitin-binding proteins are detected by anti-WGA antiserum, the 8 kDa mBL could be detected in several transgenic plants, especially DL108-5 (lane 4). This protein was not a native tobacco protein since it could not be detected in chitin column eluate of leaf extract from an untransformed plant (lane 7) or a plant transformed with the vector pGA643 alone (lane 9). Protein from a transgenic plant expressing the full length barley lectin was affinity purified on chitin and provided a positive control for the western blot (lane 8). The levels of expression of the mBL in transformed tobacco plants, however, was very low. Plants produced by selfing the transformants also failed to produce high levels of mBL (data not shown).

DISCUSSION

A family of chitin-binding proteins is found in a wide variety of plants. These proteins all contain a Cysteine and Glycine rich domain of approximately 43 amino acids. The chitin-binding domain has been fused to an unrelated domain in

Figure 7. Expression of mBL in transgenic tobacco. Immunoblot detection with anti-WGA antiserum of leaf extract proteins which bound to a chitin column. Lanes 1-6 are extracts from separate kanamycin resistant plants transformed with DL108. The 8 kDa chitin-binding mBL can be seen in several lanes and was most abundantly present in a single plant (lane 4). The control lanes contain extracts from an untransformed plant (lane 7), a plant expressing large quantities of the full length barley lectin gene (lane 8), and a plant transformed with the vector, pGA643, without any insert (lane 9).

IMMUNOBLOT ANALYSIS OF THE BARLEY LECTIN CONSTRUCT EXPRESSED IN TRANSGENIC TOBACCO



prohevein and the class I chitinases. Other proteins in this family are made up of the chitin-binding domains alone. The very small protein hevein, 4 kDa, is made up of a single chitin-binding domain and can be isolated from rubber tree latex by its affinity to chitin. Because of this, chitin-binding domains are often referred to as 'hevein' domains. Stinging nettle lectin has approximately two chitin-binding domains repeated in tandem while the Gramineae lectins, such as barley lectin, consist of four chitin-binding domains in tandem. The chitin-binding proteins are thought to play roles in plant defense since plants contain no chitin whereas many plant pathogens and pests do (Chrispeels and Raikhel, 1991). Recently several groups have shown that several proteins which consist solely of the chitin-binding domains (hevein, stinging nettle lectin, and the Gramineae lectins) exhibit *in vitro* antifungal and/or insecticidal activities (Table I). The Gramineae lectins can inhibit the growth of insect larvae (Huesing *et al.*, 1991a) but do not show any antifungal activity (Schlumbaum *et al.*, 1986). Stinging nettle lectin, which at 8.5 kDa is approximately half the size of the Gramineae lectin monomers, inhibits the growth of both insect larvae (Huesing *et al.*, 1991b) and some fungi (Broekaert *et al.*, 1989). Hevein can also inhibit the growth of some fungal hyphae (Van Parijs *et al.*, 1991) but has not been tested for insecticidal activity.

To examine the relationship between protein size and these *in vitro* activities, the barley lectin gene was modified to encode a protein approximately the size of the nettle lectin. At the onset of this study only partial sequence data for the stinging nettle lectin was known. Chapot *et al.* (1986) determined the location of 72 stinging nettle lectin amino acids residues on 6 tryptic fragments. The

sequence of these residues was further confirmed by Dr. Steven Michnick who also determined the order of the tryptic fragments (personal communication). This amino acid sequence data was aligned with the barley lectin cDNA deduced amino acid sequence (figure 4). When aligned with the barley lectin domains containing the primary ligand-binding sites (B and C), the 72 amino acids of stinging nettle lectin sequence contains residues homologous to those directly involved in Gramineae lectin ligand binding. Thus the mBL gene was designed to encode the 72 amino acids from the start of domain B. For expression in transgenic plants the mBL gene also encoded the barley lectin signal sequence, for translocation into the ER, and carboxyl-terminal propeptide, for targeting the mBL to the vacuole.

The mBL was expressed in *E. coli* to quickly determine if an active, chitin-binding protein would be produced. In addition, large quantities of mBL could be obtained simply by increasing the size of the bacterial culture. By using the maltose binding fusion protein expression system the fusion protein could be affinity purified on an amylose column (figure 5). This was an important design consideration since it was thought the mBL would very likely be inactive when produced in *E. coli* and require a denaturation/renaturation procedure to isolate a chitin-binding protein. Once the fusion protein was collected the mBL could be isolated by cleavage with the protease Factor Xa and the MBP removed by its affinity to amylose.

It is striking that both the MBP-mBL fusion protein and mBL alone could bind to the chitin column. Several important conclusions can be drawn from these

data. First, it is possible to modify the barley lectin, eliminating half of the protein, while retaining its ability to bind chitin (Figure 6). Second, the mBL expressed in bacteria does not require denaturation/renaturation to be active. Disulfide bonds of eukaryotic proteins are not always assembled correctly when expressed in *E. coli* (Tsuji *et al.*, 1987). The mBL should have seven disulfide bonds which, presumably, are required for correct conformation into the chitin-binding protein. Finally, the finding that a carboxyl-linked mBL, in the MBP-mBL fusion protein, can impart chitin-binding activity to the MBP polypeptide (data not shown) is important for the potential of adding chitin-binding activity to other proteins. All known proteins containing the chitin-binding domain and an unrelated domain, ie hevein and the class I chitinases, have the chitin-binding domain at the amino-terminal end of the protein. Hevein linked to the carboxyl-terminal of MBP has also been shown to bind chitin (Lee *et al.*, 1991) indicating this versatility may be a general property of the chitin-binding domains. The potential for future protein engineering to provide chitin-binding activity to other unrelated proteins is enhanced by this property since placement of the chitin-binding domains could be at either the amino or carboxyl ends of the polypeptide.

The bacterial expression vector used to express the mBL is designed to encode a Factor Xa cleavage site between the MBP and the protein of interest. Amino-terminal sequencing of the mBL released by Factor Xa was used to confirm that protease cleavage had occurred correctly. The quantity of Factor Xa required, however, was prohibitively expensive for collecting large amounts of the mBL. Because of this restraint, antifungal assays of the *E. coli* expressed mBL were not

continued. Preliminary antifungal assays performed with the MBP-mBL fusion protein did not show significant antifungal activity (data not shown).

Expression and analysis of mBL in transgenic plants was the ultimate goal of this investigation. According to the WGA model for chitin-binding the mBL would have to be correctly synthesized, processed, and assembled into the dimer in order to possess chitin-binding activity. As shown in figure 7 several of the kanamycin resistant, transgenic tobacco plants expressed the 8 kDa chitin-binding mBL. Again these data demonstrate that, working from the known biochemical, crystallographic, and sequence data of the chitin-binding proteins, it is possible to 'engineer' a protein only 1/2 the size of the native barley lectin but retaining chitin-binding activity when expressed in tobacco. The active, chitin-binding mBL in the transgenic plants was detectable, however, only with the highly sensitive western blot analysis. This level of expression was too low for use in the *in vitro* antifungal assay. In addition, before enhanced fungal resistance in the transgenic plants would be expected, mBL should accumulate to the level required for stinging nettle's *in vitro* antifungal activity (50 ug/ml or approximately 50 ug/gm FW). Since the amount of the mBL in one gm leaf tissue was barely detectable with the western blot analysis it was clear that enhanced fungal resistance due to the mBL could not be expected.

Future studies will include enhancing the expression of the mBL gene in transgenic tobacco by use of a stronger promotor/enhancer system. In addition, recent cloning of the stinging nettle cDNA (Lerner and Raikhel, 1992) has provided more accurate and complete sequence information for this protein. These data will

be used to construct a second generation of mBL genes which will be expressed in transgenic plants. These studies will allow us to determine if the antifungal activity shown for nettle lectin but not for the Gramineae lectins, including barley lectin, is primarily due to the size of the active protein. Although the goal of determining whether the mBL exhibits antifungal activity was not achieved, this study has demonstrated the feasibility of substantially modifying the barley lectin while retaining the chitin-binding activity.

REFERENCES

Blake, M.S., Johnston, K.H., Russel-Jones, G.J., and Gotshclich, E.C. (1984). A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on western blots. Anal. Biochem. **136**:175-179.

Broekaert, W.F., Van Parijs, J., Leyns, F., Joos, H., and Peumans, W.J. (1989). A chitin-binding lectin from stinging nettle rhizomes with antifungal properties. Science **245**:1100-1102.

Chapot, M.P., Peumans, W.J., and Strosberg, A.D. (1986). extensive homologies between lectins from non-leguminous plants. FEBS Lett. **195**:231-234.

Chrispeels, M.J. and Raikhel, N.V. (1991). Lectins, lectin genes, and their role in plant defense. Plant Cell **3**:1-9.

Goldstein, I.J. and Hayes, C.E. (1978). The lectins: Carbohydrate-binding proteins of plants and animals. Adv. Carbohydr. Chem. Biochem. **35**:120-340.

Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981). A gas-liquid solid phase peptide an protein sequenator. J. Biol. Chem. **256**:7990-7997.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991a). Effect of wheat germ isolectins on development of cowpea weevil. Phytochemistry **30**:785-788.

Huesing, J.E., Murdock, L.L., and Shade, R.E. (1991b). Rice and stinging nettle lectins: Insecticidal activity similar to wheat germ agglutinin. Phytochemistry **30**:3565-3568.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

Lee, H.-I., Broekaert, W.F., and Raikhel, N.V. (1991). Co- and post-translational processing of the hevein preproprotein of latex of the rubber tree (*Hevea brasiliensis*). J. Biol. Chem. **266**:15944-15948.

Lerner, D.R. and Raikhel, N.V. (1989). Cloning and characterization of root-specific barley lectin. Plant Physiol. **91**:124-129.

Lerner, D.R. and Raikhel, N.V. (1992). The gene for stinging nettle lectin (*Urtica dioica* agglutinin) encodes both a lectin and a chitinase. J. Biol. Chem., in press.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning: A laboratory manual, 1st edition. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY.

Mansfield, M.A., Peumans, W.J., and Raikhel, N.V. (1988). Wheat germ agglutinin is synthesized as a glycosylated precursor. Planta **173**:482-489.

Matsudaira, P. (1987). Sequence from picomole quantities of protein electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035-10038.

Mishkind, M.L., Palevitz, B.A., and Raikhel, N.V. (1983). Localization of wheat germ agglutinin-like lectins in various species of the gramineae. Science **220**:1290-1292.

Money, N.P. (1990). Measurement of pore size in the hyphal cell wall of *Achlya bisexualis*. Experimental Mycology **14**:234-242.

Peumans, W.J., De Ley, M., and Broekaert, W.F. (1984). An unusual lectin from stinging nettle (*Urtica dioica*). FEBS Lett. **177**:99-103.

Peumans, W.J., Stinissen, H.M., and Carlier, A.R. (1982a). Isolation and partial characterization of WGA-like lectins from Rye (*Secale cereale*) and barley (*Hordeum vulgare*) embryos. Biochem. J. **203**:239-243.

Peumans, W.J., Stinissen, H.M., and Carlier, A.R. (1982b). Subunit exchange between lectins from different cereal species. Planta **154**:568-572.

Reeke, G.N. Jr. and Becker, J.W. (1988). Carbohydrate-binding sites of plant lectins. Curr. Top. in Microbiol. and Immunol. **139**:35-58.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: A laboratory manual, 2nd edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

Schlumbaum, A., Mauch, F., Vögeli, U., and Boller, T. (1986). Plant chitinases are potent inhibitors of fungal growth. Nature **324**:365-367.

Tsuji, T., Nakagawa, R., Sigimoto, N., and Fukugara, K. (1987). Characterization of disulfide bonds in recombinant proteins: Reduced human interleukin 2 in inclusion bodies and its oxidative refolding. Biochemistry **26**:3129-3134.

Van Parijs, J., Broekaert, W.F., Goldstein, I.J., and Peumans, W.J. (1991). Hevein: An antifungal protein from rubber tree (*Hevea brasiliensis*) latex. Planta **183**:258-264.

Van Damme, E.J.M., Broekaert, W.F., and Peumans, W.J. (1988). The *Urtica dioica* agglutinin is a complex mixture of isolectins. Plant Physiol. **86**:598-601.

Wilkins, T.A., Bednarek, S.Y., and Raikhel, N.V. (1990). Role of propeptide glycan in post-translational processing and transport of barley lectin to vacuoles in transgenic tobacco. Plant Cell **2**:301-313.

Wright, C.S. (1980). Crystallographic elucidation of the saccharide-binding mode on wheat germ agglutinin and its biological significance. J. Mol. Biol. **141**:267-291.

Wright, C.S., and Raikhel, N.V. (1989). Sequence variability in three wheat germ agglutinin isolectins: Products of multiple genes in polyploid wheat. J. Mol. Evol. **28**:327-336.

CHAPTER 5

SUMMARY AND FUTURE PROSPECTS

The data presented in this dissertation support the hypothesis that Barley lectin and stinging nettle lectin are involved in plant defense. Barley lectin expression was localized to the root-tips and the coleorhiza of the embryo by *in situ* hybridizations. This indicates that its function may relate to interactions with chitin containing soil microbes. Expression of stinging nettle lectin mRNA in the seeds and rhizomes, which also contact the soil, was reminiscent of barley lectin expression. The finding that the stinging nettle lectin mRNA encoded an unexpected chitinase domain was also suggestive of a plant defense function for this protein. This is particularly the case since stinging nettle lectin and chitinase act synergistically to inhibit fungal growth during *in vitro* experiments (Broekaert, 1989). Data was also presented which showed that barley lectin can be substantially modified and still bind chitin when expressed in either *E. coli* or tobacco. This was the case after deletion of over half the protein and fusion to the carboxyl-terminus of a maltose binding protein indicating that the chitin-binding proteins are quite amenable to molecular manipulations. Design and construction of artificial genes encoding chitin-binding domains alone or chitin-binding domains fused to unrelated domains is now a reality.

The unexpected structure of the protein encoded by the stinging nettle lectin mRNA leads us to a number of questions. First the processing of the prostinging nettle lectin polypeptide needs to be addressed. Production of prostinging nettle lectin domain specific antibodies will provide the tools necessary for determining the fate of the individual prostinging nettle lectin domains. There

are many interesting questions concerning these processing events. Is the prolectin processing tissue specific or developmentally regulated? It will be particularly interesting to determine if expression or processing of the stinging nettle lectin is influenced by ABA, ethylene, wounding or pathogens. Also, do the 38 kDa prostinging nettle lectin or the 28 kDa chitinase domain accumulate in the plant? If so, then it will be important to determine the tissue specificity, intracellular localization and whether the chitinase domain is active. Answering these questions will help in understanding the endogenous function of the stinging nettle lectin.

Stinging nettle lectin may be useful in engineering plants with enhanced fungal or insect resistance. It will be important to determine if processing of prostinging nettle lectin in transgenic plants is the same as in the stinging nettle plant. Perhaps the most probing question is whether stinging nettle lectin, or the proprotein, can enhance the plant's resistance to microbial pathogens or insect pests. The possibility exists that stinging nettle lectin will act in concert with endogenous chitinases in the plant. With the expression of large quantities of stinging nettle lectin and barley lectin, with various modifications, in transgenic plants it should be possible to determine why stinging nettle lectin inhibits fungal growth but barley lectin does not.

Chitin-binding proteins in plants are a widespread and varied family of proteins with potential functions in plant defense. Work presented in this dissertation has broadened the variety of protein structures we can expect within this family. In addition, these investigations have opened up new avenues of

research by showing the variability possible through molecular manipulations of chitin-binding protein genes. These are, perhaps, the first steps toward a general understanding of the endogenous functions of chitin-binding proteins and their potential uses in genetically engineered crops.

APPENDIX A

SYNTHETIC STINGING NETTLE LECTIN GENE

SEQUENCE AND OLIGONUCLEOTIDE STRUCTURE

SYNTHETIC UDA OLIGONUCLEOTIDE STRUCTURE