2004 BELTWIDE COTTON CONFERENCES, SAN ANTONIO, TX – JANUARY 5-9

EXPRESSION OF TWO ISOFORMS OF THE ANTIFUNGAL PROTEIN OSMOTIN FROM A COTTON GENE CLUSTER

Jeffery R. Wilkinson, Kimberly D. Spradling, David W. Yoder, Irma L. Pirtle, and Robert M. Pirtle
Department of Biological Sciences
University of North Texas
Denton, TX

Abstract

Fungal diseases in cotton, such as various types of root rot and wilt, cause destruction of a large percentage of the cotton crop and large economic loss throughout the cottonbelt annually. Almost all plants naturally produce their own defensive proteins, called pathogenesis-related (PR) proteins, directed against pathogens such as fungi and bacteria. PR proteins called osmotins are made in response to fungal pathogen stress or osmotic stress. The class of PR proteins called osmotins cause osmotic rupture and destruction of pathogenic fungal cells by an unknown mechanism. Since osmotins have antifungal activity against most fungal species, it may be possible to use the osmotin gene as a novel defense gene effective against numerous pathogenic fungi. To isolate prospective cotton osmotin genes to study osmotin gene structure, organization, and expression, genomic libraries in lambda phage were screened with a tobacco osmotin gene probe. Three overlapping genomic clones were found to encompass a 29.0-kb cotton DNA segment encompassing a cluster of two genes and two pseudogenes. The two genes have an identity of 92%, with open reading frames of 729 basepairs without introns, and encode putative preproteins of 242 amino acids. Two partial cDNA clones corresponding to the two genes were isolated from a cotton cDNA library, indicating that these genes are truly expressed in cotton. The two presumptive cotton osmotin preproteins can clearly be identified as members of the class of PR5 proteins due to their identities with the deduced amino acid sequences of other PR5 preproteins. The two cotton osmotin preproteins would have N-terminal signal sequences of 24 amino acids, and the mature forms of the proteins would likely be targeted for extracellular secretion as a neutral isoforms. In addition to basal TATA and CAAT promoter elements, other prospective promoter elements, such as two ethylene response elements with the canonical sequence AGCCGCC, implicated as being positive regulatory elements in the expression of a number of PR proteins, occur in the 5'-flanking sequences of the two genes. The two pseudogenes are probably nonfunctional, since they have internal stop codons in their coding regions. When treated with ethephon and hydrogen peroxide, cotton plants apparently are induced to express the osmotin proteins, as detected by Western blot analysis with a polyclonal anti-osmotin antibody preparation.

Introduction

Plants naturally produce several classes of defense proteins, called pathogenesis-related (PR) proteins, in response to a wide range of environmental stresses (Stintzi et al., 1993; Kitajima and Sato, 1999; Velazhahan et al., 1999; Datta et al., 1999; Buchanan et al., 2000). These proteins are called pathogenesis-related (PR) proteins, and are made constitutively by plants in low amounts, but can be induced to accumulate after infection by fungal and bacterial pathogens. The PR protein superfamily includes 14 subfamilies (van Loon and van Strien, 1999; Kitajima and Sato, 1999; Velazhahan et al., 1999). The PR proteins are synthesized as either basic vacuolar-targeted or neutral/acidic extracellular-targeted preproteins (Stintzi et al., 1993). The neutral/acidic extracellular forms have N-terminal signal sequences for targeting into the secretory pathway, whereas the basic vacuolar forms have C-terminal signal sequences for vacuolar targeting (Melchers et al., 1993).

PR5 proteins have been shown to have antifungal activity against many fungal pathogens (Malehorn et al., 1994; Hu and Reddy, 1997; Koiwa et al., 1997). A PR5 protein called osmotin by Bressan and coworkers (Singh et al., 1987; Singh et al., 1989; Nelson et al., 1992) is produced during osmotic stress due to water deprivation or salt exposure in tobacco cells. The tobacco osmotin has antifungal activity, causing the release of intracellular materials and hyphal rupture of many types of fungal cells (Abad et al., 1996). Transgenic potato plants that constitutively overproduce either potato or tobacco osmotins delay development of disease symptoms after infection with the fungus *Phytophthora infestans* (Liu et al., 1994; Zhu et al., 1996). Osmotins probably do not directly and non-specifically enter fungal membranes, and may interact by forming complexes with membrane receptors such as yeast PIR proteins or STE proteins (Yun et al, 1997, 1998). From the results of Ibeas et al. (2000), it would seem that fungal cell wall phosphomannans are necessary for binding osmotins and also appear to facilitate the toxic activity of osmotins. In addition, tobacco osmotin has been shown to induce apoptosis in yeast through suppression of transcription of stress-responsive genes, and its toxicity results from concerted interaction of various unidentified signal transduction pathways and cell wall components (Narasimhan et al., 2001). The amino acid sequences of a number of PR5 proteins from plants, such as tobacco (Nelson et al., 1992), maize (Malehorn et al., 1994), and *Arabidopsis* (Hu and Reddy, 1997), have been deduced from their gene or cDNA coding regions. For example, the intronless tobacco osmotin gene has an open reading frame encoding a polypeptide of 246 amino acids with a predicted N-terminal signal sequence of 21 amino acids (Nelson et al., 1992). The *Arabidopsis* osmotin-like cDNA clone contains an open reading frame encoding a polypeptide of 245 amino acids, with a predicted N-terminal signal peptide of 20 amino acids (Hu and Reddy, 1997).

The 5'-flanking regions of several PR genes have been shown to have regulatory elements for inducible gene expression (Guilfoyle, 1997). The tobacco osmotin gene has been found to be up-regulated by an array of hormonal and environmental signals, such as fungal infection, drought, salinity, wounding, cold, ethylene, and abscisic acid (Nelson et al., 1992; Raghothama et al., 1993; Raghothama et al., 1997). The tobacco gene has several *cis*-acting elements, including two ethylene-responsive elements (GCC boxes), and an abscisic acid-responsive element containing a G-box motif, that increase osmotin expression when the tobacco plant is exposed to fungal pathogens or osmotic stress. Similar promoter elements that activate gene expression when threatened with fungal pathogens have also been found in osmotin genes from potato (Zhu et al., 1995). These promoter elements also occur in the flanking regions of other PR-protein genes, such as ethylene-responsive elements in the promoter region of the tobacco PR3 chitinase gene (Shinshi et al., 1995) and the tobacco PR-1b gene (Sessa et al., 1995). However, even though a number of promoter elements have been identified, the coordination of expression of the PR protein gene superfamily is still not well understood.

Our laboratory is studying the gene organization, chromosomal arrangement, and regulation of expression of the PR5 gene family in cotton plants, with the goal of genetically engineering the cotton osmotin genes to routinely overproduce osmotin in transgenic cotton plants. This potentially may render cotton plants more resistant to fungal pathogens and possibly increase the yield of the cotton crop across the cottonbelt. This research project falls in the priorities of the "Cotton Breeding and Genetics Initiative" of Cotton Incorporated. In this report, we describe our analyses of an osmotin gene cluster containing two genes and two pseudogenes, two cDNAs for the two genes, and detection of the osmotin proteins in cotton plant extracts by Western blot analysis using an anti-osmotin antibody preparation. The cotton osmotins are the first PR5 proteins described in cotton, and the osmotin gene cluster is apparently the first such cluster of PR genes detected in plants to date.

Materials and Methods

Isolation of Cotton Osmotin cDNA and Genomic Clones

A tobacco osmotin gene subcloned into a pGEM vector (Promega) was provided by Dr. Ray Bressan of Purdue University, West Lafayette, IN (Nelson et al., 1992). This DNA was electroporated into *E. coli* strain DH5α, amplified as described in Sambrook and Russell (2001), and further purified by HPLC. An 884-bp *Hin*dIII/*Sal*I fragment was purified as template to generate radioactively-labeled random-primed DNA fragments (Feinberg and Vogelstein, 1983) for use as hybridization probe to screen cotton cDNA and genomic libraries.

A cotton cDNA library, generated from mRNA of 48-hour dark-grown cotyledons (*Gossypium hirsutum* cv. Delta Pine 62) and harbored in the Stratagene UniZAP lambda vector, was kindly provided by Dr. R.N. Trelease of Arizona State University (Ni and Trelease, 1991). The cDNA library was screened by plaque hybridization to identify positive clones encoding presumptive osmotin proteins. Prehybridization was done for 4 hours at 55°C in a solution containing 6X SSC (1X SSC is 150 mM sodium chloride and 15 mM sodium citrate, pH 7.0), 5X Denhardt's reagent, 0.5% SDS, and denatured sheared salmon sperm DNA (100 µg/ml). Hybridization was done overnight at 55°C in a solution containing 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 20 mM Tris-HCl (pH 8.0), 2 mM EDTA (pH 7.5), 2.5 mM sodium pyrophosphate (pH 8.0), denatured sheared salmon sperm DNA (100 µg/ml), and ^{32}P-labeled probe. After hybridization, the membranes were rinsed once at room temperature with 2X SSC for 5 min., once with 2X SSC, 0.1%SDS at 55°C for 30 min., and twice with 1X SSC, 0.1% SDS at 55°C for 30 min. The cloned cDNA fragments were biologically excised into the Bluescript SK(-) phagemid vector *in vivo* with proteins from ExAssist Interference-Resistant helper phage and *E. coli* SOLR strain (Stratagene). Two cDNA clones designated pCcOSM47B and pCcOSM52A were selected for DNA sequence analysis.

To isolate genomic clones encompassing putative osmotin genes, one cotton (*Gossypium hirsutum*, cv. Acala SJ-5) genomic library in the lambda vector EMBL3, generously provided by Dr. David M. Anderson of Phytogen Seeds, Placentia CA (Grula et al., 1995) and another cotton (*Gossypium hirsutum* cv. Acala SJ-2) genomic library contained in the lambda FIXII vector (Stratagene), generously provided by Dr. Thea Wilkins (University of California at Davis), were screened by plaque hybridization. The genomic libraries were screened by hybridization with probes derived from the tobacco gene or the cotton gene, using hybridization conditions described for screening the cDNA library. Three genomic clones, designated LCgOSM16B (from the EMBL3 library) and LCgOSM12A and LCgOSM7B (from the FIXII library), that gave intense positive signals were selected for analysis. The three genomic clones were shown by preliminary cycle sequencing analysis (Fan et al., 1996), using oligonucleotide primers designed for the osmotin cDNA coding regions, to contain osmotin genes identical to the coding regions of the cDNA clones pCcOSM47B and pCcOSM52A. The DNAs from the genomic clones were then digested with restriction endonucleases for physical mapping by agarose gel electrophoresis and alkaline blot hybridization. Based upon the physical map, four presumptive osmotin genes were localized. A 7.2-kb *Sal*I/*Bam*HI fragment was found to encompass one cotton osmotin gene (designated *OSMI*) in the cotton DNA segment of the genomic clone LCgOSM16B, and a 4.0-kb *Eco*RI fragment was determined to encompass the entire coding region and flanking regions of a second putative osmotin gene (designated *OSMII*) in the cotton DNA segment of the genomic clone LCgOSM12A. A 3.4-kb *Hin*dIII fragment and a 3.9-kb *Eco*RI fragment from the cotton DNA insert in the genomic clone LCgOSM7B were deduced to encompass the entire coding and flanking regions of two additional putative osmotin genes designated *OSMIII* and *OSMIV*, respectively.

DNA Sequence Analysis

Both DNA strands of the 1,052-bp cotton cDNA insert in the plasmid pCcOSM47, the 878-bp cDNA insert in the plasmid pCcOSM52, and a total of 15,007 bp of cotton genomic DNA encompassed in the four genomic subclones (from the genomic clones LCgOSM16B, LCgOSM12A, and LCgOSM7B) were sequenced, using a primer-based approach and terminator cycle sequencing (Fan et al., 1996). The genomic DNA inserts were subjected to semi-automated DNA sequencing on an Applied Biosystems Model 377XL DNA Sequencer by Lone Star Labs (Houston, TX). The sequencing reactions were done by terminator cycle sequencing with BigDye™, using 10% DMSO as denaturant. Analyses of the DNA and amino acid sequences were done with DNASIS software (Hitachi). The genomic sequences have been deposited in GenBank with Accession Numbers AF304007 and AY303690 and the cDNA sequences have been assigned GenBank Accession Numbers AF192271 and AY301283.

Genomic Blot Hybridization

Cotton genomic DNA (cv. Paymaster HS26) was purified using the procedure of Paterson et al. (1993). The genomic DNA was digested with several restriction endonucleases and the fragments resolved by electrophoresis on 0.8% agarose gels. The genomic DNA fragments were analyzed by alkaline blot hybridization with a homologous cotton DNA probe derived from an 884-bp *Eco*RV fragment containing the *OSMI* gene. The prehybridization and hybridization conditions with the homologous cotton probe were the same as described above for the tobacco gene probe, except that the stringency of hybridization was increased by raising the temperature to 60°C.

Plant Material

Cotton plants (*Gossypium hirsutum* L., cv. Paymaster HS26 and cv. Acala SJ-5) were grown in a greenhouse as previously described (Chapman and Sprinkle, 1996). Plants were germinated by soaking the seeds for 5 min. in 20% bleach, washing the seeds five times with sterile H_2O, incubating for 4 hours (with aeration at 30°C in total darkness), then rolling in pre-wetted filter paper and placing in 600 ml beakers containing 200 ml of H_2O, and incubating at 30°C in total darkness for 48 hours. Seedlings were planted in a 1:1 potting soil/sand mixture and placed in a 28°C growth chamber. When plants were three weeks old, they were sprayed with either H_2O, 100 mM H_2O_2, or 1 mM ethephon, and placed in Ziploc bags until the appropriate collection time. Plants were immediately frozen with liquid nitrogen and stored at –80°C. The cotton cv. Paymaster HS26 seeds were a gift from Dr. John J. Burke (USDA-ARS, Lubbock, TX). The cotton cv. Acala SJ-5 seeds were a gift from the California Planting Cotton Seed Distributors, Bakersfield, CA.

Preparation of a Polyclonal Anti-Osmotin Antibody

The deduced amino acid sequences derived from the osmotin cDNA sequences and the osmotin genes designated *OSMI* and *OSMII* were examined using the software packages DNASIS version 2.1 (Hitachi) and Antheprot V5.0 (http://pbil.icp.fr/ANTHEPROT) to determine an appropriate polypeptide sequence to be used as an antigen to generate a polyclonal anti-osmotin antibody in rabbits. A peptide sequence of the last 18 amino acids comprising the C-terminal end of the OSMI polypeptide (i.e. CPRGSPHIEMVGSKSQEK) was selected for synthesis of an antigenic peptide. The 18-mer oligopeptide was synthesized and the polyclonal antibody preparation in rabbits was commercially prepared by Biosynthesis Inc. (Lewisville, Texas).

Western Blot Analysis of Protein Extracts of Cotton Plants

Protein cell extracts were prepared from control and ethephon-treated cotton plants (cv. Acala SJ-5) using a modified protocol from Chlan and Bourgeois (2001). Cotton plants were sprayed with either distilled H_2O or 1 mM ethephon (Sigma). The whole plants were ground in liquid nitrogen using a mortar and pestle and the samples transferred to centrifuge tubes and placed on ice. After the liquid nitrogen evaporated, one ml of 0.1 M sodium citrate buffer (pH 5.0) per gram of tissue was added. The slurries were centrifuged for 10 min at 11,000 xg and 4°C, and the resulting supernatants filtered through Miracloth (Calbiochem). The total protein extracts were mixed with 6X sample buffer (350 mM Tris-HCl (pH 6.8), 30% glycerol, 10% SDS, 9.3% DTT, and 0.012% bromophenol blue dye) and resolved on a denaturing SDS-polyacrylamide gel (5% stacking gel and 15% running gel, Laemmli, 1970). Proteins were transferred to a nitrocellulose membrane by electroblotting. The membranes were blocked with 5% dry milk-TBS (20 mM Tris, 150 mM NaCl) for 2 hours at room temperature and incubated with the anti-osmotin antibody preparation at a 1:1,000 dilution in 5% dry milk-TBS-1% Tween-20 for one hour at room temperature. The bound antibodies were visualized using horseradish peroxidase-linked anti-rabbit secondary antibody followed by detection using ECL™ Western blotting detection reagents (Amersham Pharmacia Biotech).

Total Cotton Plant RNA Isolation for RT-PCR

To extract total RNA for use in reverse transcriptase-polymerase chain reactions (RT-PCR), 4 to 6 week-old cotton plants were sprayed with distilled H_2O, 100 mM H_2O_2, or 1mM ethephon. The leaves, stem, and root tissues were collected using a modified protocol (Chang et al., 1993; McKenzie et al., 1997) and extracted using the RNeasy Mini Kit (QIAGEN). The root, stem, and leave tissues were weighed, frozen in liquid nitrogen, and ground to a fine powder using a mortar and pestle. The ground frozen tissue samples were extracted at 65°C in 100 mM Tris-HCl (pH 8.0), 25 mM disodium EDTA, 2 M NaCl , 2% CTAB (hexadecyltrimethyl-ammonium bromide), 2% polyvinylpyrrolidone, spermidine (0.5 g/l) (Sigma), and 2% β-mercaptoethanol. The mixtures were maintained at 65°C, mixed thoroughly with chloroform, and centrifuged. The extrac-

tion procedure was repeated once for stem and root tissues and twice for leaf tissue. To precipitate the total RNA, the samples were made 2 M in lithium chloride, stored overnight at 4°C, and centrifuged. The samples were purified on RNeasy™ mini-columns (QIAGEN) and the columns washed following the protocol in the RNeasy™ manual (QIAGEN). A total of 1.5 µg of RNA per sample was run in each lane on a 1% formaldehyde denaturing gel, transferred to Hybond N^+ nylon membrane (Amersham), and analyzed using the homologous cotton osmotin probe. Total RNA from the treated cotton plants was used as templates for RT-PCR using primers designed for the *OSMI* and *OSMII* genes. The forward and reverse amplimers were: 5'-ACAAATCCCAAAGTAAGAGCTAACC-3' and 5'-CAAGGCGCAGATTAACTAGACC-3' for the *OSMI* gene and 5'-CAAATCACCAAGTAAAAACCAACC-3' and 5'-CCAAATGCAAATCAACTACTCC-3' for the *OSMII* gene. The RNA samples were amplified by RT-PCR using the Invitrogen SuperScript™ One-Step RT-PCR with Platinum® *Taq* Kit (Invitrogen) and the gene-specific primers. The PCR products were then electrophoresed on a 1.0% agarose gel to confirm the presence of the unique PCR fragments derived from the two osmotin genes (788 basepairs (bp) for the *OSMI* gene PCR product and 789 bp for the *OSMII* gene PCR product).

Vector Construction for Expression of Recombinant Osmotins in a Bacterial System

Expression of recombinant proteins in bacterial or yeast systems is much easier than in intact plants (Ausubel et al., 1997). Two plasmid constructs of two osmotin genes in the PinPoint™ Xa-3 plasmid vector system (Promega) have been generated for expression of the recombinant osmotin proteins in an *E. coli* bacterial expression system. Recently, biotinylated fusion osmotin-Factor Xa polypeptides were produced in bacterial cells transformed with the fusion constructs. Using this system, fusion proteins can be purified under mild conditions, and the purification system can theoretically yield up to 5 mg of recombinant protein per liter of bacterial culture. Also, fusion proteins produced in this expression system have a Factor Xa protease cleavage site (Factor Xa is one of the blood clotting factors), for removal of the N-terminal fusion region of Factor Xa by its specific protease to generate unfused, recombinant protein products. We have produced biotinylated fusion osmotin-Factor Xa polypeptides in bacterial cells transformed with the fusion constructs, and began the purification of small quantities of the two osmotin fusion polypeptides by affinity chromatography on an avidin resin (Promega), specific for binding the biotinylated recombinant proteins in this system. We are trying to isolate sufficient amounts of the unfused recombinant osmotins in this bacterial system, in order to bioassay their effectiveness against fungal and bacterial pathogens by filter-disk assay procedures.

Construction of Plasmid Vectors for Transformation of *Arabidopsis* and Cotton Plants

The transformation of *Arabidopsis* plants has recently been started with binary vector constructs containing the two osmotin genes, using co-cultivation with *Agrobacterium tumefacians* for transfer and integration of the constructs into the plant chromosomal DNA, in order to determine if the prospective transgenic model plants express the cotton osmotins. The binary vector is called pCAMBIA 2301 (from the Center for Application of Molecular Biology to International Agriculture (CAMBIA), Canberra, Australia). We generated two pCAMBIA constructs designed to routinely produce the cotton osmotins in the model *Arabidopsis* plants. For functional expression of the osmotins, we inserted the coding regions in-frame into CaMV 35S promoter cassettes of the binary vector. PCR amplification was used to generate DNA fragments with compatible restriction sites for proper alignment of the osmotin inserts upstream from the CaMV promoter cassettes of the pCAMBIA vector for introduction into *Agrobacterium* LBA4404 cells. We verified that the two osmotin protein-coding regions and promoter cassettes are identical in sequence to the two osmotin genes and pCAMBIA promoter elements.

Results and Discussion

Three genomic clones encompassing prospective osmotin genes were isolated from two cotton genomic libraries harbored in the EMBL3 (Acala SJ-5) and FIXII (Acala SJ-2) lambda vectors. Using either a tobacco osmotin gene probe or a cotton osmotin gene probe for alkaline blot hybridization of the immobilized DNAs, the DNAs from the genomic clones designated LCgOSM16B (in the EMBL3 vector), LCgOSM12A (in the FIXII vector), and LCgOSM7B (in the FIXII vector) were digested with various restriction endonucleases and the fragments resolved by agarose gel electrophoresis for physical mapping. The cotton genomic segments contained within the three clones were found to encompass four potential osmotin genes, later shown by DNA sequence analysis to be two osmotin PR5 genes (called *OSMI* and *OSMII*) and two osmotin pseudogenes (called *OSMIII* and *OSMIV)*. The cotton genomic inserts in the LCgOSM16B, LCgOSM12A, and LCgOSM7B clones were determined to be 16.0 kb, 15.5 kb, and 16.5 kb, respectively. The three overlapping lambda clones encompass a total of 29.0 kb of cotton genomic DNA. Four genomic subclones were generated from the three lambda clones to encompass the osmotin gene cluster. Both strands of the four subcloned cotton genomic DNA segments were sequenced using an oligonucleotide primer-based approach, primarily by semi-automated DNA sequencing. A segment of 15,007 bp of cotton genomic DNA sequence was determined from the four overlapping subclones. A total of 7,127 bp (GenBank Accession AY303690) from the cotton insert in the clone LCgOSM16B was sequenced, and a total of 8,914 bp (GenBank Accession AF304007) from the cotton inserts in the clones LCgOSM12A and LCgOSM7B were determined.

The cotton genomic segments encompass two actual osmotin genes that lack introns, and the genes are identical with the coding regions in the osmotin cDNA clones designated pCcOSM47B (*OSMI*) and pCcOSM52A (*OSMII*), respectively. The entire cDNA inserts of pCcOSM47B and pCcOSM52A were determined to be 1,052 bp (GenBank Accession AF192271) and 878 bp (GenBank Accession AY301283) in size, respectively, excluding the poly(A) tails. Based upon the genomic se-

quences, the cDNA insert of pCcOSM47B was inferred to lack 10 bp at the 5'-end of the open reading frame (ORF), and the cDNA insert of pCcOSM52A was deduced to lack 101 bp at the 5-end of its open reading frame. A prospective cap site for the 5'-end of the mature osmotin mRNAs is located 37 nucleotides (nt) before the ATG initiation codon for the *OSMI* gene and 36 nt from the start codon of the *OSMII* gene. The 3'-flanking regions of the cotton cDNA insert and osmotin genes have canonical near-upstream poly(A) polymerase elements (Hunt and Messing, 1998) of 5'-AATAAA-3' at 35 nt upstream from the poly(A) cleavage/polyadenylation site of the *OSMI* gene, which is 297 nt downstream from the stop codon, and 10 nt upstream from the poly A site of the *OSMII* gene, which is 138 nt downstream from the stop codon, inferred from comparison with the cDNA sequences.

The two osmotin genes have a wide variety of potential promoter elements in their 5'-flanking regions which could bind various transcription factors for activation of their expression in different environmental situations, perhaps even through synergistic effects (Mitchell and Tjian, 1989; Young, 1991; Thomas, 1993; Guilfoyle, 1997). Putative TATA and CAAT basal promoter elements (Mitchell and Tjian, 1989; Young, 1991) are located at –49 nt and -103 nt upstream from the ATG start codon of the *OSMI* gene, and -49 nt and -96 nt from the initiation codon of the *OSMII* gene. Two presumptive ethylene response elements (GCC boxes) with the sequence 5'-AGCCGCC-3' occur -127 nt and -63 nt from the initiation codon for the *OSMI* gene and -63 nt and -122 nt for the *OSMII* gene. Ethylene-response elements occur in promoters of numerous PR-protein genes, and many of these genes have two or more GCC boxes (Guilfoyle, 1997). Kitajima et al. (1998) showed that the GCC elements of a tobacco osmotin PR5 protein are constitutive in roots and cultured cells, and are inducible by ethylene in tobacco leaves. The Bressan group (Nelson et al., 1992; Raghothama et al., 1993; Raghothama et al., 1997) has extensively studied the promoter region of a tobacco osmotin PR5 protein, which has two GCC boxes in a 140-bp minimum promoter region. These elements are required for ethylene responsiveness but are not sufficient for maximal ethylene-induced gene activation (Raghothama et al., 1997). Three potential basic-region helix-loop-helix (bHLH) or E box motifs with the consensus sequence CANNTG occur at 486 nt, 223 nt, and 92 nt upstream from the ATG start codon of the *OSMI* gene and 445 nt, 182 nt, and 85 nt for the *OSMII* gene. The E box motif has been shown to be a seed-specific positive element for expression of the French bean β-phaseolin gene (Kawagoe et al., 1994). An H-box or AC-element (5'-CCTACC(N_7)CT(N_4)A-3') occurs in reverse orientation at 831 nt upstream from the initiation codon of the *OSMI* gene. H-box motifs occur in the promoter region of the bean phenylalanine ammonia-lyase 2 (*PAL2*) gene (Hatton et al., 1995) and may be important in developmental-specific gene activation. There are four prospective GATA motifs with the sequence 5'-GATAA-3' at 737 nt, 607 nt, 411 nt, and 307 nt upstream from the initiation codon of *OSMI* gene and 856 nt, 1100 nt, 1252 nt, 2136 nt upstream from the start codon of the *OSMII* gene. GATA elements occur in the promoter regions of many light-regulated genes (reviewed in Guilfoyle, 1997) as well as some promoters, like the CaMV35S RNA promoter, that are light-unresponsive (Lam and Chua, 1989).

The coding regions of a number of osmotin protein cDNAs/genes were compared at the nucleotide sequence level. The highest degree of identity (75%) with the cotton osmotin coding regions is with the *Fagus sylvatica* (beech tree) osmotin coding region (AJ298303). As would be expected, tobacco osmotin coding sequences have the highest identities with each other (greater than 90%), and the grape osmotin sequences also have high identities of more than 60% with each other. In contrast, the osmotin coding regions of several species have low sequence identities, such as between *Arabidopsis* and *Atriplex,* with identities of 32% and 37%. From comparison of the alignments of the deduced amino acid sequences for the putative cotton osmotin preproteins and other putative plant osmotin preproteins, a 59% identity occurs between the cotton and grape osmotin proteins, a 48% identity occurs between the cotton and tobacco preproteins, about 35% identities occur between the cotton and the thaumatin and maize preproteins, and only a 17% identity occur between the cotton and *Arabidopsis* preproteins.

The cotton osmotin preproteins of 242 amino acids, deduced from the cDNA sequences of pCcOSM47B and pCcOSM52A and the gene sequences of pCgOSM16B and pCgOSM12A were analyzed using the software package PSORT (http://psort.ims.u-tokyo.ac.jp) to predict putative targeting sequences and the preprotein signal sequence cleavage sites, and by Antheprot V5.0 (http://pbil.ibcp.fr/ANTHEPROT) to predict isoelectric points (pI) and preprotein cleavage sites. The cotton osmotin preproteins were predicted to have N-terminal preprotein signal sequences that direct the preproteins into the extracellular secretory pathway (Neuhaus and Rogers, 1998), resulting in the extracelluar secretion of the mature isoforms. The predicted preprotein C-terminal regions apparently lack the necessary signal sequences for vacuolar targeting. Both preprotein signal sequences are predicted to be cleaved between amino acids 24 and 25, resulting in alanine being the N-terminal amino acid in both isoforms. Both proteins have 16 invariant cysteine residues, crucial in the formation of eight disulfide bonds of the two osmotins (de Vos et al., 1985; Ogata et al., 1992; Koiwa et al., 1994; Batalia et al., 1996; Koiwa et al., 1997). Such invariant cysteine residues are also found in identical members and locations in all other PR5 proteins, such as those from *Vitis raparia* (AF178653), *N. tabacum* (S40046), *T. danielli* (J01209), and *A. thaliana* (U83490) (in the deduced primary structures of all six PR5 proteins). de Vos et al. (1985) observed that the turn and loop motif maintained by the disulfide bonds occurs in a number of other proteins, such as toxins, agglutinins, and cytotoxins, that interact with membrane-bound receptors, and speculated that this structural feature could be important for binding to fungal membrane receptors. Due to the 8 disulfide bonds, the cotton osmotins have similar β-strand structures that occur in thaumatin, zeamatin, and tobacco protein Pr-5d, and hence could potentially adopt the β-sandwich three-dimensional structure for interacting with fungal membrane receptor proteins (Kitajima et al., 1999; Batalia et al., 1996; de Vos et al., 1985).

A 24 kilodalton (kDa) protein was detected using the anti-osmotin antibody preparation in Western blot analyses of crude protein extracts from 4 to 6 week-old cotton plants (cv. Acala SJ-5) treated with various chemical inducers (1 mM ethephon, 550 mM NaCl, or 100 mM hydrogen peroxide). The protein was not observed in Western blots of crude cotton extracts treated with water as control. The molecular weights were estimated by comparison with protein molecular weight standards (Amersham). Cotton tissue protein extracts from 4 to 6 week-old cotton plants (cv. Acala SJ-5) were treated with either 1 mM ethephon or water, and the tissues were collected at 0, 4, 8, 12, 16, 20, 24, 48, and 96 hr. Using the anti-osmotin antibody for Western blot analysis, a 24.0 kDa protein was found to be present in the ethephon-treated samples, but it was not observed in the control sample extracts. Two smaller molecular protein bands can also be observed in the crude homogenates from both the ethephon-treated and the control samples, and were determined to be 20.0 kDa and 15.0 kDa in size. It is possible that these two bands may correspond to two predicted polypeptides of 19.5 and 14.0 kDa that would result from possible internal translation start codons in both the *OSMI* and *OSMII* genes.

To date, we have been unsuccessful in generating sufficient amounts of the recombinant cotton osmotins in this bacterial expression system for bioassaying their effectiveness against fungal pathogens. It is possible that the recombinant osmotins, which would have eight disulfide bonds, may be forming aggregates of inclusion bodies in the bacterial cells. Inclusion bodies of numerous recombinant proteins with improperly formed disulfide bonds are notorious for causing insoluble aggregates in bacterial systems, due to faulty three-dimensional folding of the recombinant proteins (Marston, 1987; Burgess, 1996). For many recombinant proteins, inclusion bodies can be very difficult to deal with, since they often prove to be intractable to dissolve. We are currently trying to determine if the recombinant osmotins are aggregating as inclusion bodies in the bacterial extracts. If this is the case, there are various experimental procedures for slowly solubilizing and renaturing proteins from inclusion bodies, and we will try several methods for dissolving them (Sambrook and Russell, 2001).

The *Eco*RV fragment encompassing the *OSMI* gene from the clone LCgOSM16B was used as hybridization probe for alkaline blot hybridization of cotton genomic DNA, in order to confirm that DNA fragments derived from the cloned osmotin genes truly occur in genomic DNA, and also to estimate the number of similar osmotin genes in the cotton genome. Double digestion of cotton genomic DNA with *Xho*I and *Bam*HI produces three hybridizing fragments of 16.9 kb, 7.6 kb, and 5.5 kb. Digestion of cotton genomic DNA with *Eco*RI generates four hybridizing fragments of 10.5 kb, 4.2 kb, 4.0 kb, and 3.2 kb. Cleavage of genomic DNA with *Eco*RV yields three hybridizing fragments of 5.2 kb, 3.3 kb, and 0.9 kb. The 5.5-kb *Xho*I/*Bam*HI and 0.9-kb *Eco*RV genomic fragments directly correspond to the 5.5-kb *Xho*I/*Bam*HI and 0.9-kb *Eco*RV DNA fragments generated from the cloned cotton DNA segment in the genomic clones LCgOSM16B and LCgOSM12A. The 3.2-kb *Eco*RI genomic fragment directly corresponds to a 3.2-kb *Eco*RI restriction sites found near the *OSMI* gene, while the 4.2-kb genomic fragment corresponds to the *OSMII* gene. The other hybridizing fragments in the *Xho*I/*Bam*HI, *Eco*RI, and *Eco*RV digests of cotton DNA must be derived from several similar osmotin genes, since the hybridization was under relatively stringent conditions. Thus, there seem to be at least three or more actual osmotin genes in the allotetraploid cotton genome.

Reverse transcriptase-polymerase chain reactions (RT-PCR) analyses were done to determine if both *OSMI* and *OSMII* transcripts were present in the total RNA extracts from roots, stems and leaves of cotton plants that had been treated with 1 mM ethephon, 100 mM hydrogen peroxide, or water (control). Unique oligonucleotide primers were designed for the 5' and 3'-untranslated regions (UTRs) of the osmotin gene sequences to amplify a 788-bp product from the *OSMI* gene and a 789-bp product for the *OSMII* gene. Both strands of the 788-bp and 789-bp *OSMI* and *OSMII* RT-PCR products were sequenced to confirm their identities. Both osmotin mRNAs apparently are present in all these tissues when treated with water, ethephon, or hydrogen peroxide.

We have begun the transformation of model *Arabidopsis* plants with binary vector constructs of the two osmotin genes, using co-cultivation with *Agrobacterium tumefacians*, in order to determine if the prospective transgenic model plants express the cotton osmotins, by a floral dip procedure. The dipped plants are grown until seeds became mature, and then the dried seeds are harvested. We are currently screening the harvested seeds from the *Arabidopsis* plants for potential transformants. Once true transformants have been identified, they will be transplanted to soil and grown to maturity. The seeds from these transformants will be harvested and used as seed stock to generate lines of transformed *Arabidopsis* plants. The transformed plants will then be assayed in order to determine if the prospective transgenic plants express the cotton osmotins, and to determine their extent of resistance against various fungal and bacterial pathogens. We will begin transformation of cotton plants with these constructs when the results from the model *Arabidopsis* plants are deemed satisfactory.

Acknowledgements

This research was supported by Texas State Support Program Agreement 01-897TX, overseen by Dr. Roy G. Cantrell, Cotton Incorporated, and by University of North Texas Organized Research Funds.

References

Abad, L.R., M.P. D'Urzo, D. Liu, M.L. Narasimhan, M. Reuveni, J.K. Zhu, X. Niu, N.K. Singh, P.M. Hasegawa, and R. A. Bressan. 1996. Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization, Plant Science 118: 11-23.

Ausubel, F.M., R, Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, 1997. Current Protocols in Molecular Biology, John Wiley & Sons, New York

Batalia, M.A., A.F. Monzingo, S. Ernst, W. Roberts, and J.D. Robertus. 1996. The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family. Nat. Struct. Biol. 3: 19-23.

Burgess, R.R. 1996. Purification of overproduced *Escherichia coli* RNA polymerase sigma factors by solubilizing inclusion bodies and refolding from Sarkosyl. Methods Enzymol. 273:145-149.

Chapman, K.D. and W.B. Sprinkle. 1996. Development, tissue-specific, and environmental factors regulate the biosynthesis of N-acylphosphatidyl-ethanolamine in cotton (*Gossypium hirsutum* L.) J. Plant Physiol. 149: 277-284.

Chlan, C.A., and R.P. Bourgeois. 2001. Class I chitinases in cotton (*Gossypium hirsutum*): characterization, expression and purification. Plant Science 161: 143-154.

Chang, S., J. Puryear, and J. Cairney. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter 11: 113-116.

Datta, K., S. Muthukrishnan, and S.K. Datta. 1999. Expression and function of PR-protein genes in transgenic plants. In: Pathogenesis-related Proteins in Plants, Datta, S.K. and Muthukrishnan, S. (Eds), CRS Press, New York, NY, pp. 261-277.

de Vos, A.M., M. Hatada, H. van der Wel, H. Krabbendam, A.F. Peerdeman, and S.-H. Kim. 1985. Three dimensional structure of thaumatin I, an intensely sweet protein. Proc. Natl. Acad. Sci. USA 82:1406-1409.

Fan, J., R.S. Ranu, C. Smith, C. Ruan, and C.W. Fuller. 1996. DNA sequencing with [α-^{33}P]-labeled ddNTP terminators: a new approach to DNA sequencing with Thermo Sequenase DNA polymerase. BioTechniques 21: 1132-1137.

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

Guilfoyle, T.J. 1997. The structure of plant gene promoters, in J.K. Setlow (Ed.), Genetic Engineering, Vol. 19, Plenum Press, New York, pp. 15-47.

Grula, J.W., R.L. Hudspeth, S.L. Hobbs, and D.M. Anderson. 1995. Organization, inheritance and expression of acetohydroxyacid synthase genes in the cotton allotetraploid *Gossypium hirsutum*. Plant Mol. Biol. 28: 837-846.

Hatton, D., R. Sablowski, M.H. Yung, C. Smith, W. Schuch, and M. Bevan. 1995. Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. Plant J. 7: 859-876.

Hu, X., and A.S.N. Reddy. 1997. Cloning and expression of a PR5-like protein from *Arabidopsis*: inhibition of fungal growth by bacterially expressed protein. Plant Mol. Biol. 34: 949-959.

Hunt, A.G. and J. Messing. 1998. mRNA polyadenylation in plants. In: A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants, J. Bailey-Serres and D.R. Gallie (Eds.), American Society of Plant Physiologists, Rockville, MD, pp. 29-39

Ibeas, J.I., H. Lee, B. Damsz, D.T. Prasad, J.M. Pardo, P.M. Hasegawa, R.A. Bressan,, and M.L. Narasimhan. 2000. Fungal cell wall phosphomannans facilitate the toxic activity of a plant PR-5 protein. Plant J. 23: 375-383.

Kawagoe, Y., B.R. Campbell, and N. Murai. 1994. Synergism between GACGTG (G-box) and CACCTG cis elements is required for activation of the bean seed storage protein beta-phaseolin gene. Plant J. 5: 885-890.

Kitajima, S., T. Koyama, Y. Yamada, and F. Sato. 1998. Constitutive expression of the neutral PR-5 (OLP, PR-5d) gene in roots and cultured cells of tobacco is mediated by ethylene-responsive cis-element AGCCGCC sequences. Plant Cell Reports 18: 173-179.

Kitajima, S. and F. Sato. 1999. Plant pathogenesis-related proteins: molecular mechanisms of gene expression and protein function. J. Biochem. 125: 1-8.

Koiwa, H., F, Sato, and Y, Tamada. 1994. Characterization of accumulation of tobacco PR-5 proteins by IEF-immunoblotting analysis. Plant Cell Physiol. 35: 821-828.

Koiwa, H. H. Kato, T. Nakatsu, J. Oda, Y. Yamada, and F. Sato. 1997. Purification and characterization of tobacco pathogenesis-related protein PR-5d, an antifungal thaumatin-like protein. Plant Cell Physiol. 38: 783-791.

E. Lam, and N.H. Chua. 1989. ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in Cab promoters. Plant Cell 12: 1147-1156.

Liu, D., K.G. Raghothama, P.M. Hasegawa, and R. A. Bressan. 1994. Osmotin overexpression in potato delays development of disease symptoms. Proc. Natl. Acad. Sci. USA 91: 1888-1892.

Malehorn, D.E., J.R. Borgmeyer, C.E. Smith, and D.M. Shah. 1994. Characterization and Expression of an Antifungal Zeamatin-like Protein (*Zlp*) Gene from *Zea mays*. Plant Physiol. 106: 1471-1481.

Marston, F.A. 1987. The purification of eukaryotic polypeptides expressed in *Escherichia coli*. In: DNA Cloning: A Practical Approach , Vol. 3 (D.M. Glover, ed.) IRL Press, Oxford, UK, pp. 59-88.

McKenzie, D.J., M.A. McLean, S. Mukerji, and M. Green. 1997. Improved RNA extraction from woody plants for the detection of viral pathogens by reverse transcription-polymerase chain reaction. Plant Disease 81: 222-226.

Melchers, L.S., M.B. Sela-Buurlage, S.A. Vloemans, C.P. Woloshuk, J.S.C. Van Roekel, J. Pen, P.J.M. van den Elzen, and B.J.C. Cornelissen. 1993. Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and β-1,3-glucanase in transgenic plants. Plant Mol. Biol. 21: 583-593.

Mitchell, P. and R. Tjian, R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins, Science 245: 371-378.

Narasimhan, M.L., B. Damsz, M.A. Coca, J.I. Ibeas, D.-J. Yun, J.M. Pardo, P.M. Hasegawa., and R.A. Bressan,. 2001. A plant defense response effector induces microbial apoptosis. Molecular Cell 8:921-930.

Nelson, D.E., K.G. Raghothama, N.K. Singh, P.M. Hasegawa, and R.A. Bressan. 1992. Analysis of structure and transcriptional activation of an osmotin gene. Plant Mol. Biol. 19: 577-588.

Neuhas, J., and J.C. Rogers. 1998. Sorting of proteins to vacuoles in plant cells. Plant Mol. Biol. 38: 127-144.

Ni, W., and R.N. Trelease. 1991. Post-transcriptional regulation of catalase isozyme expression in cotton seeds. Plant Cell 3: 737-744.

Ogata, C.M., P.F. Gordon, A.M. de Vos, and S.-H. Kim. 1992. Crystal structure of a sweet tasting protein, thaumatin I, at 1.65A^{o} resolution. J. Mol. Biol. 228: 893-908.

Paterson, A.H., C.L. Brubaker, and J.F. Wendel. 1993. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Reporter, 11: 122-127.

Raghothama, K.G., D. Liu, D.E. Nelson, P.M. Hasegawa, and R.A. Bressan. 1993. Analysis of an osmotically regulated pathogenesis-related osmotin gene promoter. Plant Mol. Biol. 23: 1117-1128.

Raghothama, K.G., A. Maggio, M.L. Narasimhan, A.K. Kononowicz, G. Wang, M.P. D'Urzo, P.M. Hasegawa, and R.A. Bressan. 1997. Tissue-specific activation of the osmotin gene by ABA, C_2H_4 and NaCl involves the same promoter region. Plant Mol. Biol. 34: 393-402.

Sambrook, J., and D.W. Russell. 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shinshi, H., S. Usami, and M. Ohme-Takagi. 1995. Identification of an ethylene-responsive region in the promoter of a tobacco class I chitinase gene. Plant Mol. Biol. 27: 923-932.

Sessa, G., Y. Meller, and R. Fluhr. 1995. A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene. Plant Mol. Biol. 28: 145-153.

Singh, N.K., C.A. Bracker, P.M. Hasegawa, A.K. Handa, S. Buckel, M.A. Hermodson, E. Pfankoch, F.E. Regnier, and R.A. Bressan. 1987. Characterization of osmotin: a thaumatin-like protein associated with osmotic adaptation in plant cells. Plant Physiol. 85: 529-536.

Singh, N.K., D.E. Nelson, D. Kuhn, P.M. Hasegawa, and R.A. Bressan. 1989. Molecular cloning of osmotin and regulation of its expression by ABA and adaptation to low water potential. Plant Physiol. 90: 1096-1101.

Stintzi, A, T. Heitz, V. Prasad, S. Wiedemann-Merdinoglu, S. Kauffmann, P. Geoffroy, M. Legrand, and B. Fritig. 1993. Plant "pathogenesis-related" proteins and their role in defense against pathogens. Biochimie 75: 687-706.

Thomas, T.L. 1993. Gene expression during plant embryogenesis and germination: an overview. Plant Cell 5: 1401-1410.

van Loon, L.C., and E.A. van Strien. 1999. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol. Mol. Plant Pathol. 55: 85-97.

Velazhahan, R., S.K. Datta, and S. Muthukrishnan. 1999. The PR-5 family: thaumatin-like proteins. In: Pathogenesis-related Proteins in Plants, Datta, S.K. and Muthukrishnan, S. (Eds), CRS Press, New York, NY, pp. 107-129.

Young, R.A. 1991. RNA polymerase II, Annu. Rev. Biochem. 60: 689-715.

Yun, D.-J., Y. Zhao, J.M. Pardo, M.L. Narasimhan, B. Damsz, H. Lee, L.R. Abad, M.P. D'Urzo, P.M. Hasegawa, and R. A. Bressan. 1997. Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. Proc. Natl. Acad. Sci. USA 94: 7082-7087.

Yun, D.-J., J.I. Ibeas, H. Lee, M.A. Coca, M.L. Narasimhan, Y. Uesono, P.M. Hasegawa, J.M. Pardo, and R.A. Bressan. 1998. Osmotin, a plant antifungal protein, subverts signal transduction to enhance fungal cell susceptibility. Mol. Cell 1: 807-817.

Zhu, B., T.H.H. Chen, and P.H. Li. 1995. Activation of two osmotin-like protein genes by abiotic stimuli and fungal pathogen in transgenic potato plants. Plant Physiol. 108: 929-937.

Zhu, B., T.H.H. Chen, and P.H. Li. 1996. Analysis of late-blight disease resistance and freezing tolerance in transgenic potato plants expressing sense and antisense genes for an osmotin-like protein. Planta 198: 70-77.