Characterization of the multigene family encoding endopolygalacturonase in the basidiomycete *Chondrostereum purpureum*

by

Holly L. Williams

B.Sc., Simon Fraser University, 1996

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Department

of

Biology

UNIVERSITY OF VICTORIA

January 2001

APPROVAL

Name: **Holly Williams**

Degree: **Master of Science**

Title of Thesis: **Characterization of the multigene family encoding endopolygalacturonase in the basidiomycete *Chondrostereum purpureum***

We accept this thesis as conforming to the required standard

[redacted]

Dr. W. E. Hintz, Supervisor (Department of Biology)

[redacted]

Dr. P. von Aderkas, Departmental Member (Department of Biology)

[redacted]

Dr. S. F. Shamoun, Departmental Member (Department of Biology)

[redacted]

Dr. S. Misra, Outside Member (Department of Biochemistry)

[redacted]

Dr. W. W. Kay, External Examiner (Department of Biochemistry)

Date Approved Jan. 24, 2001

ABSTRACT

Chondrostereum purpureum is a white-rot basidiomycete fungus which is being developed as a biocontrol agent of hardwoods for use in reforestation sites and hydroelectric rights-of-ways. It produces several plant cell wall-degrading enzymes that are suspected virulence factors, including endopolygalacturonase (endoPG). To search for new endoPG genes in *C. purpureum*, degenerate oligonucleotide primers were designed based on conserved regions of seventeen ascomycete endoPG genes and two published endoPG sequences from *C. purpureum*. These primers were used to amplify endoPG gene fragments from *C. purpureum* 2128u genomic DNA. Fifty-six of the amplification products were analyzed by sequence determination or restriction digestion to reveal segments of three previously uncharacterized endoPG genes, as well as *epgA*, an endoPG gene sequenced from *C. purpureum* 2128u. The three new genes were named *epgB*, *epgC*, and *epgD*, and the full length sequences of two copies of *epgB* (*epgB1* and *epgB2*) were retrieved from a genomic library of *C. purpureum* 2128u. The identities between the deduced polypeptide sequences for *epgA*, *epgB1*, *epgC* and *epgD* ranged from 61.8% to 80.0%. The two *epgB* genes shared 97.6% nucleotide identity and 97.1% amino acid identity, with the majority of differences between them near the 3' ends of the genes. Phylogenetic analysis of these five basidiomycete endoPGs, seventeen ascomycete endoPGs, four ascomycete exoPGs, two bacterial endoPGs, and two bacterial exoPGs separated the *C. purpureum* endoPGs into their own monophyletic clade; whereas the endoPGs from ascomycete species having multiple endoPGs were divided between different monophyletic groups within the ascomycete endoPG clade. This split between the ascomycete and basidiomycete endoPGs suggests the two fungal phyla diverged before the duplication of existing endoPG genes. Mechanisms regulating the expression of the endoPG genes in *C.*

purpureum were examined by northern analyses using mycelia grown in media with different carbohydrate sources. The *epgA* gene was expressed at high levels when glucose, sucrose or pectin was used as the carbon source; whereas both *epgB* genes, *epgC*, and *epgD* were expressed at low or non-detectable levels on all media types. The expression of *epgA* appeared to be constitutive and not repressed by simple sugars, suggesting EPGA may be the most important of the five endoPGs in the plant cell wall degradation caused by *C. purpureum*.

Examiners:

Dr. W. E. Hintz, Supervisor (Department of Biology)

Dr. P. von Aderkas, Departmental Member (Department of Biology)

Dr. S. F. Shamoun, Departmental Member (Department of Biology)

Dr. S. Misra, Outside Member (Department of Biochemistry)

Dr. W. W. Kay, External Examiner (Department of Biochemistry)

ACKNOWLEDGMENTS

I would like to thank Dr. William Hintz for his guidance, support and enthusiasm. He has been an excellent supervisor, and is a tremendous source of knowledge. My lab mates Josh Eades, Brad Temple, Elisa Becker, Louise Hahn, Paul de la Bastide, Kirk Leifso, and Mike Pinchback were always happy to offer advice, and they made the Hintz lab a fun and interesting place to work. Dr. Ben Koop generously provided advice for the phylogenetic analysis, and his technicians sequenced countless DNA samples. Roderick Haesevoets was very accommodating in providing sequence data for the last endoPG gene. My family and friends gave me encouragement and support throughout my studies. I gratefully acknowledge the funding provided by grants from B.C. Hydro, the Natural Sciences and Engineering Research Council of Canada, and the Science Council of British Columbia, as well as financial support from the M.C. Melburn Award, Amelia Leith Memorial Fellowship and University of Victoria Graduate Teaching and Research Fellowships.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3'	three prime
5'	five prime
bp	base pair
C-terminal	carboxyl-terminal
C. purpureum	*Chondrostereum purpureum*
CWDE	cell wall-degrading enzyme
cDNA	copy deoxyribonucleic acid
dCTP	deoxycytosine triphosphate
DEP-C	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
endoPG	endopolygalacturonase
GPD	glyceraldehyde-3-phosphate dehydrogenase
kb	kilobase pair
M	molar
ml	millilitre
mM	millimolar
N-terminal	amino-terminal
ng	nanogram
PCR	polymerase chain reaction
PG	polygalacturonase
PGIP	polygalacturonase-inhibiting protein
PL	pectate lyase
pI	isoelectric point
RAPD	randomly amplified polymorphic DNA
rDNA	ribosomal DNA
RNA	ribonucleic acid
SCAR	sequence-characterized amplified region
SDS	sodium dodecyl sulphate
TAE	tris acetate EDTA
Taq	*Thermus aquaticus*
UV	ultraviolet

INTRODUCTION

1.1 Vegetation management

Fast-growing hardwoods, such as alder (*Alnus spp.*), black cottonwood (*Populus nigra*), trembling aspen (*Populus tremuloides*), birch (*Betula sp.*), and maple (*Acer spp.*), often outcompete industrially valuable conifers for light, nutrients, water and space (Wall et al., 1992). These hardwoods are considered to be weeds in the context of conifer reforestation, and are also problematic to utility companies. Traditionally, growth of these deciduous species has been controlled by manual cutting followed by application of chemical herbicides to prevent resprouting from cut stumps. This is generally effective for suppression of unwanted vegetation, but chemical agents often affect non-target species, and may accumulate in the soil and water table with long term use. Increasing public opposition to the use of chemical herbicides has helped to drive research on alternative approaches to vegetation management, including the development of biological control agents (Dorworth, 1990; Wall et al., 1992).

1.2 Biological control

Biological control is defined as the deliberate use of natural enemies to suppress the growth or reduce the population of a weed species (Watson, 1989). This involves the direct application of living agents or their metabolites to the target species to reduce its population. The classical or inoculative approach of biological control requires the release of a pathogen into a small portion of the total weed infestation. If conditions are favorable, the pathogen spreads to establish a weed-suppressing epidemic (Charudattan, 1988). Both the weed and pathogen are often from a foreign locale where they have co-evolved. This method is

particularly useful for controlling foreign pests (Templeton et al., 1979; Charudattan, 1988). A more recent innovation is the inundative approach, whereby large quantities of a local endemic pathogen are applied to the unwanted species, in the same general manner as a chemical agent. The endemic pathogen is applied in massive inoculations in order to temporarily overcome constraints to its epidemic development in nature (Templeton et al., 1979). The inundative method has a lower risk to the environment than the classical method of biocontrol, because the native pathogen generally does not survive over time in sufficient numbers to initiate a fresh epidemic on new weed infestations (Charudattan, 1988). Thus, there is usually a need for repeated application of the inundative agent to control weed populations (Charudattan, 1988). Inundative control has been applied to agricultural weeds with considerable success, and is now being developed as a strategy for use in forest vegetation management (Markin and Gardner, 1993).

Several criteria must be met in order for an organism to be considered a candidate for development as a biocontrol agent. It must grow rapidly on culture media, readily infect its target host under a variety of environmental conditions, and suppress growth or induce mortality of the target weed species in a relatively short time (Wall et al., 1992). Since many diseases of forest weeds are caused by fungal pathogens, fungi which naturally infect unwanted hardwoods were investigated as potential bioherbicides for use on conifer plantations and hydroelectric rights-of-ways.

1.3 *Chondrostereum purpureum* as a biological control agent

1.3.1 Chondrostereum purpureum as a phytopathogen

The phytopathogenic fungus *Chondrostereum purpureum* (Pers. ex. Fr.) Pouzar was first recognized as a candidate biocontrol agent in Holland (Scheepens, 1980) for the control of the American black cherry (*Prunus serotina*). These trees were introduced from North America around 1920 to improve understory and litter in planted pine stands, but soon became forest weeds, colonizing cleared areas and suppressing the growth of native species. *Chondrostereum purpureum* is currently marketed as a "wood decay promoter" (BioChon™) in the Netherlands (de Jong, 2000).

The potential of using *C. purpureum* to control Canadian hardwood weeds was first recognized by Dr. R. E. Wall of the Canadian Forest Service (Pacific Division). This basidiomycete is able to colonize many deciduous species, but does not affect conifers, making it an attractive candidate for control of woody species in reforestation sites and rights-of-ways (Wall, 1986). *Chondrostereum purpureum* is widespread throughout North America, and collections have also been made in Europe, Japan and New Zealand. It may potentially be used as a native bioherbicide in many areas of the world.

Chondrostereum purpureum is more commonly recognized as the causal agent of silver-leaf disease, which affects a wide range of fruit trees in New Zealand, South Africa, Europe, Canada, the United States (Brooks and Moore, 1926) and Japan (Miyairi et al., 1977). The silvering of the foliage in response to fungal infection is caused by the partial physical separation of palisade cells from the epidermal layer, as well as separation of the cuticle from the epidermis (Brooks and Moore, 1926). The resulting air pockets reflect incident light and give the leaves a silvery sheen.

Generally, *C. purpureum* invades its hosts though fresh pruning wounds, cut stumps or stem lesions. It can debilitate target plants relatively quickly by disrupting the vascular tissues, severing the physical connection between foliage and roots. If the fungus invades the cambium of the lower stem or the root crown, it effectively prevents growth of adventitious shoots (Brooks and Moore, 1926; Miyairi, 1988; Wall 1990). Because *C. purpureum* is wound invasive, it does not generally spread quickly, and inundation of a wounded host is necessary for it to be effective as a biological control agent (de Jong et al., 1990; Wall, 1991). This characteristic is highly favorable because it allows application to target trees with little danger of infecting adjacent trees. Nonetheless, concerns have been raised that use of *C. purpureum* as a biocontrol agent for forestry may increase the incidence of silver-leaf disease in orchard crops near areas of application. Risk analysis studies addressing these concerns indicate that the relatively short lived basidiospores produced by the fungus are unlikely to pose a hazard to non-target trees. Simulated basidiospore dispersal with a Gaussian plume model estimated 160 basidiospores/m^3 at a distance of 5000 meters (de Jong et al., 1990). A 500 meter buffer zone between treated sites and orchards was recommended, since this would imply a risk of the same magnitude as naturally occurring spores. Field trials in which two different isolates of *C. purpureum* were applied to Sitka alder (*Alnus sinuata*) and trembling aspen suggested that the risk to nontarget trees was very low, as isolate-specific genetic markers showed that biocontrol isolates were recovered only from stumps to which they were applied, and were absent from control stumps (Becker et al., 1999a). Of 220 stumps cut adjacent to experimental sites in Quebec where two native strains of *C. purpureum* were deployed, 15% became infected with *C. purpureum*, but at least 85% of the infections were not attributable to deployed strains of the fungus (Gosselin et al., 1999b). Naturally occurring spores were thus the main sources of

inoculum, implying minimal risk associated with use of *C. purpureum* as a biological control agent in the forest (Gosselin et al., 1999b).

1.3.2 Efficacy of Chondrostereum purpureum in infecting and suppressing target species

New Brunswick and Ontario field trials testing the frequency of successful stump infection by *C. purpureum* four months post-treatment found an average of 84% speckled alder (*Alnus rugosa*) stems and 54% trembling aspen stems infected by the applied fungus (Becker et al., 1999b). In British Columbia, 90% of biocontrol-treated Sitka alder (*Alnus viridis*) stems and 40% of treated aspen stumps were infected with *C. purpureum* four months after application (Becker et al., 1999a). Field efficacy was tested two years following treatment in British Columbia, at which point close to a 90% Sitka alder clump mortality was observed. This mortality level was not significantly different from that using triclopyr herbicide (Harper et al., 1999). For trembling aspen, *C. purpureum* treatment significantly increased stem mortality and reduced stump sprout activity; however, neither *C. purpureum* nor triclopyr herbicide were considered successful in control of aspen due to high densities of stump sprouts and root suckers (Harper et al., 1999).

1.3.3 Evaluation of the genetic structure of Chondrostereum purpureum populations

Several studies have been done to assess the population structure of *C. purpureum* in order to evaluate the extent of the risks associated with deploying a single isolate of the fungus across large geographical distances. The extent of genetic variability was estimated by variation in protein profiles (Ekramoddoullah et al., 1993), isozymes (Shamoun and Wall, 1996), the internal transcribed spacer (Shamoun et al., 1991) and the large non-transcribed

spacer (Ramsfield et al., 1996) regions of the ribosomal DNA (rDNA), random amplified polymorphic DNA (RAPDs) (Gosselin et al., 1995; 1999a; Spiers et al., 2000), sequence-characterized amplified regions (SCARs) in mitochondrial DNA (Ramsfield et al., 1999), and restriction polymorphisms in total mitochondrial DNA (Ramsfield et al., 1999). None of these studies identified any host or geographical specializations. Mitochondria-specific SCAR analysis (Ramsfield et al., 1999) and RAPD analysis (Gosselin et al., 1999a) indicated that variation within Canadian subpopulations was greater than variation among subpopulations, providing evidence that extensive gene flow has historically tended to homogenize the populations. This suggested very little risk in using a single natural isolate as a biological control agent across the whole of Canada (Ramsfield et al., 1996; Gosselin et al., 1999a).

1.4 Plant cell wall structure and cell wall-degrading enzymes

The plant cell wall can be considered to be a polymeric mesh consisting primarily of cellulose, hemicellulose, pectic substances, and proteins (McNeil et al., 1984). Cellulose and hemicellulose are integral components of the cell wall, and the pectin matrix helps provide structural organization to these and other cell wall components (Carpita and Gibeaut, 1993). Pectin is one of the primary constituents of outer wall regions within the middle lamella, and constitutes a substantial proportion (~35%) of primary walls of dicots (Darvill et al., 1980). Because pectic substances are more exposed than other cell wall components, they are especially susceptible to enzymatic degradation (Collmer and Keen, 1986) and the action of pectin-degrading enzymes may expose underlying structures to other cell wall-degrading enzymes (Bateman and Basham, 1976; Talmadge et al., 1973). When the pathogens

Colletotrichum lindemuthianum (English et al., 1971), *Fusarium oxysporum* f. sp. *lycopersici* (Jones et al., 1972), *Fusarium roseum* (Mullen and Bateman, 1976), and *Verticillium albo-atrum* (Cooper, 1977) were grown on isolated plant cell walls, pectic enzymes were the first cell-wall degrading enzymes to be synthesized, while cellulose-degrading enzymes were among the last to be induced.

The pectin matrix of plant cell walls is composed primarily of rhamnogalacturonan, comprised of chains of α-1,4- linked galacturonic acid residues interspersed with 1,2-linked rhamnose (Cooper et al., 1978) (Fig. 1). The carboxyl groups of the galacturonate residues may be methoxylated (McNeil et al., 1984), or may be cross-linked by calcium, which confers rigidity to cell walls (Cooper, 1984). The neutral sugar polymers arabinan, xylan and/or arabinogalactan are attached as side chains covalently linked to rhamnogalacturonan (McNeil et al., 1984), and may serve as a bridge between the rhamnogalacturonan and the hemicellulosic wall components (Talmadge et al., 1973).

To breach the first line of plant defense, most plant pathogens have evolved enzymes capable of degrading the various constituents of plant cell walls. Rhamnogalacturonan, the main component of the pectin matrix, is degraded by hydrolytic polygalacturonases (PG) or by pectin or pectate lyases (PNL, PL) which attack the α-1,4- bonds linking the galacturonic acid residues in the pectic molecule (Fig. 2). These enzymes may attack the chain internally (*endo*-), releasing oligomers, or terminally (*exo*-), releasing monomeric products, although some enzymes combine the two modes of action (Cooper et al., 1978). Certain lyases and polygalacturonases preferentially attack partially methylated pectin, while others prefer the non-esterified pectate. The methoxy groups on pectinic acids are removed by pectin methylesterase (PME), which can thereby decrease steric hindrance and facilitate the action of

FIGURE 1. Possible structure of pectic polymers, composed of rhamnogalacturonan and covalently linked neutral polysaccharides. GalA, galacturonic acid; Rha, rhamnose; Gal, galactose; Ara, arabinose (Albersheim, 1976; Darvill et al., 1980)

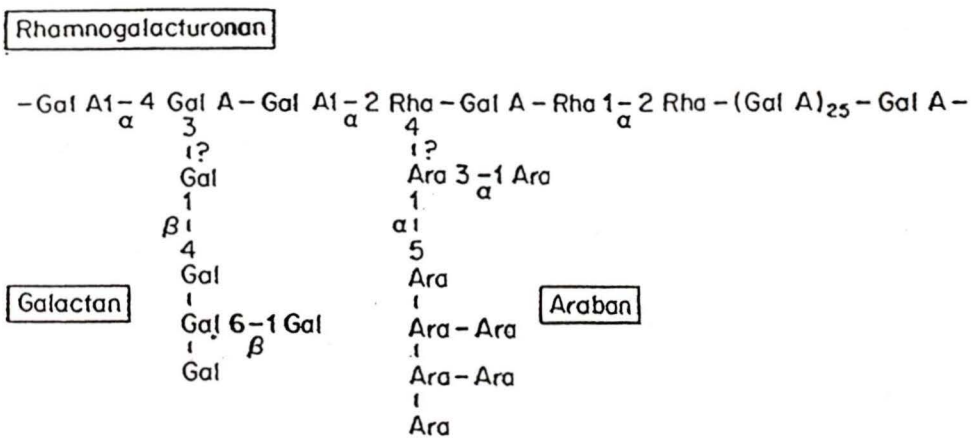
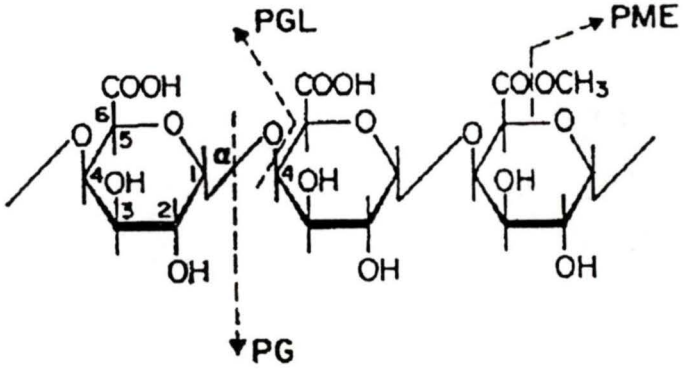

Rhamnogalacturonan
-Gal A1-4 Gal A-Gal A1-2 Rha-Gal A-Rha 1-2 Rha-(Gal A)25-Gal A-
α
α
3
?
Gal
1
β
4
Gal
Gal 6-1 Gal
β
Gal
4
?
Ara 3-1 Ara
α
1
α
5
Ara
Ara-Ara
Ara-Ara
Ara
Galactan
Araban

FIGURE 2. Cleavage of an α-1,4-linked polygalacturonide chain by endopolygalacturonase (PG) and endopolygalacturonide lyase (PGL), and demethoxylation by pectin methylesterase (PME). Note the numbering of carbons on the galacturonosyl units, the α configuration of the glycosyl bond, and the methylated or free carboxyls (Cooper, 1983).


PGL
PME
COOH
COOH
COOCH3
OH
OH
OH
OH
OH
OH
PG

depolymerases specific for non-esterified substrate (Keon et al., 1987; Rexová-Benková and Markovic, 1976). Rhamnogalacturonase (RHG) cleaves the bond between the alternating galactose and rhamnose residues in rhamnogalacturonan (Suykerbuyk et al., 1995). The neutral polysaccharides linked to rhamnogalacturonan and other cell wall polymers may be cleaved by *endo-* and *exo-* glycanases and glycosidases (Cooper et al., 1978), xylanases, arabinases and arabinogalactanases, which may act in concert to increase access to the main polymers (Annis and Goodwin, 1997).

The plant cell wall presents both a barrier to infection and a source of metabolizable substrates to plant pathogens. Furthermore, the oligogalacturonates released from the plant cell wall by endopolygalacturonase can provide nourishment for the pathogen, but may also act as an elicitor of the plant defense response. As a result endopolygalacturonase may function in host-pathogen interactions as both a pathogenicity factor and a potential avirulence factor (Cervone et al., 1989). The oligogalacturonates found to most strongly induce host defense reactions were those with a degree of polymerization between 10 and 13; whereas smaller fragments were less markedly biologically active (Hahn et al., 1981; Nothnagel et al., 1983; Jin and West, 1984, Davis et al., 1986a; 1986b; Cervone et al., 1987; McNeil et al., 1984). It has been suggested that the accumulation of elicitor-active oligogalacturonides may be facilitated by plant polygalacturonase-inhibiting proteins (PGIPs), which bind to endoPGs and retard polygalacturonase-mediated hydrolysis of polypectate (Cervone et al., 1989). Plant PGIPs are constitutively produced at low levels and quickly induced by fungal infection (Bergmann et al., 1994; Cervone et al., 1993), suggesting they play a role in resistance when they are overexpressed. A high level of PGIP in bark tissue of Chinese chestnut was suggested to explain the resistance of this host to the chestnut blight fungus *Cryphonectria parasitica*,

which readily infects American chestnut (Gao and Shain, 1995). Some PGIPs interact with a broad spectrum of fungal PGs (Cervone et al., 1990), while others have limited specificity (Desiderio et al., 1997). Multiple PGIPs produced by the same plant may also have differing specificities (Desiderio et al., 1997).

The ability of host plants to recognize pathogenicity factors or their degradation products may have played an important role in the progression of fungi from saprophytes to facultative pathogens and finally to biotrophic pathogens. In living plants, it may be sufficient or even desirable for the pathogen to secrete low levels of pectinolytic enzymes, to avoid the cytotoxic effects associated with these enzymes (Davis et al., 1984). An endoPG purified from the saprophytic fungus *Aspergillus niger* exhibited a specific activity significantly higher than those of purified polygalacturonases from phytopathogenic fungi (Cervone et al., 1987). In the obligate biotrophic rust fungus *Uromyces viciae-fabae*, pectinase and cellulase production is strictly controlled by the differentiation of infection structures (Frittrang et al, 1992; Heiler et al., 1993), whereas in necrotrophs (e.g. *Botrytis allii, Rhizoctonia solani, Sclerotinia sclerotiorum*, etc.), production of pectic enzymes is not related to their differentiation and may precede infection (Alghisi and Favaron, 1995). Low PG activity levels were detected *in planta* and in axenic culture of the largely biotrophic fungus *Claviceps purpurea* (Tenberge et al., 1996), and mychorrhizal fungi also seem to produce smaller quantities of PG than saprophytic fungi (Peretto et al., 1993).

1.5 Evidence for the role of endopolygalacturonase as a pathogenicity factor

1.5.1 Physiological correlations between endoPG and virulence

Several factors are important to the establishment of fungal pathogens during the infection process. These include recognition and attachment to the plant surface, formation of infection structures, penetration of the host, and colonization of the host tissue (Schafer, 1994). For *C. purpureum*, penetration of the host and colonization of the host tissue may be especially important to establishment of disease because this fungus requires a pre-existing wound surface on the tree for its entry. Cell wall-degrading enzymes, in particular polygalacturonases (PGs), have been implicated as virulence factors in numerous phytopathogens. For example, studies with purified endoPGs and PLs from many plant pathogenic bacteria and fungi clearly demonstrated their unique ability to cause cell maceration and separation, while exo-acting PGs and other cell wall-degrading enzymes were relatively ineffective (Bateman and Basham, 1976). Polygalacturonase was the first detectable enzyme secreted by *Fusarium oxysporum* (Jones et al., 1972) and *Colletrotrichum lindemuthianum* (English et al., 1971) when grown *in vitro* on plant cell walls. PG activity was detected in sunflower plants inoculated with virulent isolates of *Sclerotium bataticola*, whereas no PG activity was detected in plants inoculated with avirulent isolates (Chan and Sackston, 1970). High concentrations of endoPGs were visualized around pectin-rich areas of bean tissue being degraded by the fungus *C. lindemuthianum* (Benhamou et al., 1991) and in areas where extensive pectin breakdown was occurring in soybean tissue treated with an endoPG of *Sclerotinia sclerotiorum* (Favaron et al., 1993). Considerably more polygalacturonase activity was detected in tissues of American chestnut (*Castanea dentata*) susceptible to *Cryphonectria parasitica* than in tissues of resistant Chinese chestnut (*Castanea mollissima*) (Gao and Shain,

1995). A strain of *Aspergillus flavus* that constitutively produced an endoPG caused more severe disease symptoms in developing cotton bolls than low-virulence strains that lacked a major endoPG (Brown et al., 1992; Cleveland and Cotty, 1991). EndoPG purified from *Aspergillus niger* and applied to cowpea (*Vigna unguiculata*) pods elicited a necrotic response (Cervone et al., 1987), while endoPG purified from *Chondrostereum purpureum* and injected into apple seedlings induced silver-leaf disease symptoms (Miyairi, 1988).

1.5.2 Genetic manipulations to definitively establish the role of endoPG in pathogenicity

Recombinant DNA technology offers the most refined approach to determine whether endopolygalacturonase is required for the pathogenicity of a fungal species. If the genes encoding endoPGs important in the pathogenic process were to be disabled, the transformed fungus would be expected to be significantly less virulent. Genetic modifications of endoPG genes in phytopathogens have so far yielded conflicting results. The PG of *Agrobacterium tumefaciens* biovar 3 was disrupted and found to be necessary to the development of root decay in grape seedlings (Rodriquez-Palenzuela et al., 1991). Disruption of the endoPG gene in the bacterium *Pseudomonas solanacearum* resulted in greatly reduced disease symptoms on tomato (*Lycopersicon esculentum*) (Schell et al., 1988) and on eggplants (*Solanum melongena*) (Huang and Allen, 1997). Disruption of one of two exoPGs led to a lesser yet still significant drop in virulence. Double disruptants, lacking both the endoPG and exoPG, were the least virulent of all (Huang and Allen, 1997). Disruption of one of the six identified endoPG genes in the fungus *Botrytis cinerea* significantly reduced the aggressiveness of the fungus on all three host plant tissues tested (ten Have et al., 1998). Similarly, disruption of an endoPG gene in aggressive isolates of *Aspergillus flavus* caused a significant drop in the ability

of the fungus to damage and spread within its host. When this same gene was introduced into isolates of *A. flavus* lacking the gene, the transformants became significantly more aggressive (Shieh et al., 1997).

There is not always such a clear correlation between endoPG disruption and reduction in pathogenicity. For example, the virulence of the chestnut blight fungus *Cryphonectria parasitica* did not change with disruption of the *enpg-1* gene. It should, however, be considered that while this endoPG was the major form produced *in vitro*, it was found to contribute less than 5% of the total PG activity in canker tissue. As two new endoPGs were identified that predominated *in vivo*, this particular disruption did not allow the role of endoPG in this pathogen to be resolved (Gao et al., 1996). Three isolates of the tomato vascular wilt fungus *Fusarium oxysporum* deficient in the major *in vitro* endoPG (PG-1) were transformed with a vector containing the *pg1* gene and promoter, but no difference was detected between the virulence level of the transformants and their wild-type counterparts. Although this study provided strong evidence that PG-1 was not essential for virulence of *F. oxysporum* (Di Pietro and Roncero, 1998), the importance of other endoPGs in this species was not examined. Disruption of both the endo- and exoPG genes in the fungal maize pathogen *Cochliobolus carbonum* did not result in a reduction in pathogenicity (Scott-Craig et al., 1990; 1998).

1.6 Significance of multiplicity in genes encoding endopolygalacturonase

There appears to be a great deal of variation in the number of endopolygalacturonase genes within the fungi. Multigene families of seven endoPG genes were identified in both *Aspergillus niger* (Parenicová et al., 1998) and *Sclerotinia sclerotiorum* (Reymond et al., 1994). Six endoPG genes were sequence characterized in *Botrytis cinerea* (Wubben et al.,

1999). Three endoPGs were identified in *Cryphonectria parasitica* (Gao et al., 1996) and *Aspergillus flavus* (Whitehead et al., 1995), and two endoPG genes were sequenced in each of *Colletotrichum lindemuthianum* (Centis et al., 1996; 1997), *Claviceps purpurea* (Tenberge et al., 1996), and *Fusarium oxysporum* (Di Pietro and Roncero, 1998; Accession No. AF078156). It appears only one endoPG gene is present in each of *Cochliobolus carbonum* (Scott-Craig et al., 1990) and *Fusarium moniliforme* (Caprari et al., 1993).

There are advantages to having multiple endoPG genes. For pathogens, multiplicity of cell wall-degrading enzymes may confer a greater adaptive capacity and flexibility (Keon et al., 1987). For example, in the phytopathogenic fungus *S. sclerotiorum*, the large range in the net charges of multiple endoPG forms may allow them to react differentially according to various pectic substrates and environments (Martel et al., 1996). Successive replacement of isoforms by new enzymes may allow endoPGs essential in the early stages of cell wall degradation to be produced first, and then replaced by enzymes successively better able to degrade smaller oligomers. Two isoforms present after two days of culturing *S. sclerotiorum* in media containing polygalacturonic acid were not detected after 10 days, and the activity of another isoform dropped from 45% of the total activity of all purified isoforms in 4-day old cultures to only 2% in 10-day old cultures. Another PG, not detected after 2 days and representing 0.5% of the PG activity after 4 days, was the main fraction after 10 days of culture, accounting for 67% of the activity of the purified enzymes (Martel et al., 1998). In *B. cinerea*, expression of a constitutively produced endoPG appeared to induce the expression of several other endoPG genes (Leone and van den Heuvel, 1987). The constitutively produced endoPG was the only isozyme secreted in the outer region of plant lesions (Kamoen and van der Cruyssen, 1996), and disruption of the gene showed that the enzyme was required for full virulence of the

fungus (ten Have et al., 1998). Further support for a coordinated degradation of pectin by multiple endoPGs was provided by the finding that the seven characterized endoPG enzymes in *Aspergillus niger* differed in regulation of expression and in processive behaviour on polygalacturonate and a series of oligogalacturonates of various lengths (Benen et al., 1999; Parenicová et al., 1998; 2000a; 2000b).

It has been suggested that multigene families of cell wall-degrading enzymes may allow pathogens to infect a wide range of host species (Beaulieu and van Gijsegem, 1992; Fraissinet-Tachet et al., 1995). The finding that full virulence of the bacterium *Erwinia chrysanthemi* depended on different pectate lyases to infect various hosts provided support for this theory (Beaulieu and van Gijsegem, 1992), however, in considering the endopolygalacturonase gene families in fungi, as many examples contradict this theory as those which support it. For instance, *B. cinerea* infects over 200 different plant species (Jarvis, 1977) and has a multigene family of six identified endoPG genes; but Southern hybridization analyses indicated that homologs of these six genes were also present in four other *Botrytis* species, each of which only infects a single host species (Wubben et al., 1999). PG forms with differences in both pI values and in capacity to macerate fruit tissues were extracted from different plant species infected by *Monilinia fructigena* (Byrde, 1979), whereas the PG isozymes produced by *S. sclerotiorum* on apple fruits and sunflower stems had similar properties (Marciano et al., 1982). The single-host chestnut pathogen *C. parasitica* produces at least three endoPG isozymes (Gao et al., 1996), whereas the cereal pathogen *F. moniliforme* has only one endoPG gene but infects a wide variety of plants (Caprari et al., 1993).

1.7 Regulation of endoPG expression

Gene disruption could well be the most powerful tool for studying the role of putative virulence factors in phytopathogenic fungi, but the success of any study also depends on understanding how expression of the gene of interest is regulated. Before beginning gene disruption studies, all of the genes encoding the enzyme of interest should be characterized, and factors which affect their expression levels *in vitro* and *in planta* should be determined. Disruption of a gene which is weakly or not expressed during infection would provide little insight into the role of its corresponding enzyme in pathogenicity.

The mechanisms involved in regulating fungal cell wall-degrading enzymes have not been well studied. Most cell wall-degrading enzymes are primarily regulated at the level of transcription, with production induced by specific substrate components and repressed by readily metabolised carbohydrates such as glucose (Cooper, 1983). Many CWDE enzymes are subject to self-catabolite repression, whereby gene expression is halted when products are released from the substrate at a rate exceeding the utilization ability of the cell. Some inducible enzymes may be synthesized at a basal level, and a minority of CWDE are constitutively produced (Cooper, 1983).

The major *in vitro* extracellular endopolygalacturonase of *Fusarium oxysporum* showed the most common *in vitro* regulation pattern for fungal endoPGs, being induced by citrus pectin, polygalacturonic acid, host plant vascular tissue, and 0.1% galacturonic acid, and repressed by 1% galacturonic acid and glucose (Di Pietro and Roncero, 1998). Similarly, the endo- and exoPGs in *Cochliobolus carbonum* were induced by the presence of pectin and suppressed by 2% sucrose (Scott-Craig et al., 1998). In contrast, a medium containing potato extracts and 2% glucose was highly favorable for endoPG production in *Chondrostereum*

purpureum (Miyairi et al., 1977). In *Colletotrichum lindemuthianum*, *CLPG1* and *CLPG2* were both transiently induced on pectin, with *CLPG2* only briefly detected. *CLPG1* was highly expressed *in planta* during the necrotrophic phase of parasitism of the fungus, while at this stage *CLPG2* was not expressed. Two of the seven recently characterized endoPGs from the saprophytic fungus *Aspergillus niger*, encoded by *pgaA* and *pgaB*, were unusual in that they preferred partially methylated substrates and were constitutively produced. These properties along with their different pH optima led to the speculation that they may act as 'scouting' enzymes which allow rapid adaptation of the fungus to pectic substrates (Parenicová et al., 2000a). In *Botrytis cinerea*, the six endoPGs were differentially expressed over time both in glucose and in polygalacturonic acid media (Wubben et al., 1999), and one of the enzymes was constitutively produced (Leone and van den Heuvel, 1987; van der Cruyssen et al., 1994). In *Aspergillus flavus* one member of a multigene family of endoPGs appeared to be constitutively produced, and disruption of this gene revealed it was an important pathogenicity factor (Shieh et al., 1997). Likewise, disruption of the distinct constitutively produced endoPG in *B. cinerea* demonstrated that its expression was essential for full virulence of the fungus (ten Have et al., 1998). The ability of some organisms to produce endoPG in the absence of an inducer may be thus be a key element in their pathogenicity.

1.8 Research objectives

The objectives of this study were to clone and characterize the endopolygalacturonase genes in *Chondrostereum purpureum* and to determine factors controlling the regulation of their expression. This information will allow the genetic manipulation of these genes in order to definitively determine the role endoPG enzymes play in the pathogenicity of this species.

The *C. purpureum* endoPGs characterized in this study are the first sequenced polygalacturonase genes in a basidiomycete fungus. The presence of multiple endoPGs in this species and in most ascomycete fungi raised the question of whether the multigene families of endoPG were present in ancestral fungi prior to the divergence of the ascomycetes and basidiomycetes, or whether the endoPG gene duplications occurred independently in the two lineages. The evolution of fungal endoPG genes was examined using phylogenetic analysis of endoPGs from *C. purpureum* and several ascomycete species.

MATERIALS AND METHODS

2.1 *Chondrostereum purpureum* culture conditions and DNA extraction

The *C. purpureum* homokaryotic isolate 2128u was derived from a germinated basidiospore of the heterokaryotic isolate 2128 isolated from bigleaf maple (*Acer macrophyllum*) at Mesachie Lake, British Columbia, Canada. Mycelium for DNA extraction was grown at 23°C in liquid culture of 1.5 % malt extract (Difco) for 2 weeks prior to harvesting by filtration. Total genomic DNA was extracted from freeze-dried mycelium using the method of Ramsfield et al. (1996).

2.2 Design of degenerate PCR primers for detection of new endoPG genes

To search for endoPG genes in *Chondrostereum purpureum*, degenerate oligonucleotide primers were designed according to conserved regions of 17 endoPG gene sequences from 8 species of ascomycete fungi, *epgA* from *C. purpureum* 2128u (Tang, 1996), and a newly published cDNA sequence for a second endoPG from *C. purpureum* isolate ASP-4B (Miyairi *et al.*, 1997). These sequences were aligned using the ClustalW 1.8 algorithm (Thompson *et al.*, 1994), and the two regions of the aligned nucleotide sequences having the highest degree of sequence conservation were selected for primer design. The forward primer (Cpepgs-For) 5'T A/G T/G/C T G G G A T/C C G G A/C A/T A/T/G G G A/T/G/C A/T/G G/C A/T/C A A C corresponded to nucleotide positions 553 through 576 of *epgA*. The reverse primer (Cpepgs-Rev) 5'A/G G A A/G A T/G/C A/T C C A/G T G A/G/C C C A/T/G C C A/T/G G A A/G C A corresponded to the reverse complement of nucleotide positions 989 through 1012 of *epgA*. The amplification product was expected to be approximately 460 bp

long. PCR amplifications were performed using ~30 ng total genomic DNA from *C. purpureum* 2128u as template, 2.5 μM each of primers Cpepgs-For and Cpepgs-Rev, 2 μl 10X *Taq* DNA polymerase buffer (Pharmacia), 200 μM dNTPs, and 1 unit *Taq* DNA polymerase (Pharmacia) in a final volume of 20 μl. The first 5 cycles of PCR amplification were 94°C, 60 s; 53°C, 90 s; 72°C, 120 s. The annealing temperature was increased to 58°C for the next 25 cycles. The program was followed by an extension at 72°C for 10 min. The PCR products were size separated on 1% low melting temperature agarose gels (Sigma) and visualized by staining in ethidium bromide. DNA fragments of the expected size were excised, purified by the Wizard PCR DNA purification system (Promega), and cloned into the pGEM-T vector (Promega). The nucleotide sequences of the cloned PCR products were identified by sequencing or by restriction analysis. The sequenced inserts were checked for similarity to other endoPGs using the National Center for Biotechnology information BLAST WWW server.

2.3 DNA sequencing

All DNA sequences in this study were obtained by automated fluorescence sequencing performed on a ABI 373A automated sequencer (Applied Biosystems, Foster CA) using the protocol specified in the ABI PRISM dye primer cycle sequencing ready reaction manual. For each sample, four reactions were set up (A,C,G,T) for PCR amplification, each containing deoxynucleoside triphosphates (dNTP), fluorescent linked universal M13 primer, pyrophosphatase, AmpliTaq DNA polymerase FS, one of four "termination" dideoxynucleoside triphosphates (ddNTP), and 250 ng of plasmid DNA. Thermal cycle

sequencing was performed using the suggested parameters (ABI protocol manual). On completion, the A, C, G and T sequencing mixtures were pooled, phenol/chloroform treated, precipitated, and resuspended in 4.1 µl formamide. The samples were denatured at 90°C for two minutes and loaded onto a 5% acrylamide-urea sequencing gel. The samples were run on the automated sequencer overnight and primary sequence data was produced by the ABI-373A.

2.4 Genomic library construction, screening and lambda DNA purification

Total genomic DNA was digested with *Mbo*I such that the average fragment size was between 10 to 20 kilobase pairs as determined by agarose gel electrophoresis. The partially digested DNA was ligated to *Bam*HI-predigested EMBL-3 λ DNA replacement vector (Promega Scientific, Mississauga, Ont.) and the ligated DNA was packaged *in vitro* using commercial Lambda phage particle extracts (Gigapack, Stratagene, San Diego, CA). The recombinant phage DNA library of *C. purpureum* was plated using the host *E. coli* LE392 and plaques were lifted onto nylon membranes (Genescreen, Dupont) for library screening. Radio-labelled endoPG gene probes were synthesized by the incorporation of ^{32}P-dCTP during PCR amplification using the newly cloned endoPG gene fragments as target DNA. Plaque hybridization was done overnight according to Sambrook *et al.* (1989), and the blots were washed 4 times at room temperature in 2X SSC, 0.1% (w/v) SDS and twice at 55° C in 0.1X SSC, 0.1% SDS. Following 12 to 48 hours exposure to Kodak X-OMATAR film, positive recombinant λ phage were identified, picked, eluted, replated, and rescreened using plaque hybridization until well-isolated, positively hybridizing phage clones were obtained. Lambda phage were picked and eluted into 100 ml LB (10 g NaCl, 10 g tryptone, 5 g yeast extract in 1

L dH_2O, pH 7.0 with NaOH) + 10 mM $MgSO_4$ and grown for 15 hours at 37°C on a rotary shaker at 150 rpm. The cell debris was pelleted at 10,000 rpm for 15 min and 1 μg/ml each of DNase and RNase were added to the supernatant. Following a 2 hour incubation at 37° C, an equal volume of PEG solution (50% w/v PEG 4000, 0.01 M Tris-HCl pH 7.5, 2 M NaCl, 0.10 M $MgSO_4 \cdot 7H_2O$) was added, and the mixture was chilled for 2 hours on ice. Phage particles were pelleted at 8000 rpm for 20 minutes, resuspended in 2 ml phage dilution buffer (5.8 g NaCl, 2.0 g $MgSO_4 \cdot 7H_2O$, 50 ml 1 M Tris-HCl pH 7.5, 5.0 ml 2% w/v gelatin, in 1 L dH_2O), and extracted with an equal volume of chloroform until no white interface remained. Following addition of 25 μl 10 % SDS and 25 μl 500 mM EDTA pH 8.0, the solution was incubated for 15 minutes at 65°C, extracted once with phenol and twice with chloroform, and precipitated with an equal volume of isopropanol. The λ DNA pellet was washed with 70 % cold ethanol, dried, and resuspended in 50 μl TE (10 mM Tris-Cl, 1 mM EDTA, pH 7.6).

2.5 Subcloning endoPG genes and compiling their sequences

Purified recombinant λ phage clones were digested with various restriction endonucleases having a 6-bp specificity and the digestion products were size separated on a 0.8 % agarose gel. The digested DNA were Southern blotted as described below and probed with the appropriate ^{32}P-dCTP-labeled PCR endoPG fragment. Fragments large enough to contain the entire gene and flanking DNA, but small enough to be easily cloned into plasmid vector (4 to 8 kb), were subcloned into one of the bacterial plasmids Bluescript II K/S + or pUC18 using the following shotgun cloning strategy: λ DNA digested with the appropriate enzyme was ligated into plasmid DNA digested with the same enzyme, and the ligation mix was used to transform *E. coli* αDH10 competent cells (GibcoBRL). Several ampicillin-

resistant colonies were picked for plasmid DNA extraction using Wizard minipreps (Promega), and the plasmids were digested with the chosen enzyme to identify a plasmid containing the appropriate DNA fragment from the λ clone. A detailed restriction map was constructed for each plasmid to facilitate subcloning and resectioning. Subsectioned plasmids were prepared for automated sequencing using the QIAprep spin miniprep kit (Qiagen). Multiple alignments of the sequences were created using the Generunner program (Hastings Software) and used to compile the full length gene sequences. The five gene and deduced polypeptide sequences were aligned using ClustalW 1.8 (Thompson et al., 1994) and adjusted manually to correct minor flaws in the alignment near the 5' ends of the genes.

2.6 Determination of gene copy number

Total genomic DNA from isolate 2128u was digested with *Bgl*II and *Eco*RI (Pharmacia), separated on four identical 0.8% 1 X TAE (50X: 242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA) agarose gels, and transferred to Genescreen Plus (NEN Life Science Products) according to the methods of Sambrook et al. (1989). The blots were hybridized with one of each ^{32}P-dCTP-labeled PCR fragment of *epgA*, *epgB*, *epgC*, or *epgD*, derived from amplification of the appropriate ~460 bp endoPG PCR fragment. These fragments were obtained by gel purification following PCR amplification of the λ EMBL-3 clone containing the appropriate endoPG gene with the Cpepgs-For and Cpepgs-Rev primers. The same procedure was used for washing and exposing the blots as for recombinant lambda DNA hybridization. To confirm the copy number of the *epgB* genes, an additional 2128u *Sac*I (Pharmacia) digest was blotted and probed with *epgB*.

2.7 Design of PCR primers for the amplification of *epgB* gene fragments

To determine whether the *cppg1* gene (Accession number AB043873) from isolate ASP-4B highly similar to the 2128u *epgB1* gene corresponded to the second *epgB* gene in isolate 2128u, a primer pair was developed to specifically amplify a ~670 bp product from within the *epgB* genes. The sequences of *epgA*, *epgB1*, and *cppg1* were aligned and two segments well conserved between *epgB1* and *cppg1* but significantly different from *epgA* were chosen as templates for primer design. The forward primer, B-For, corresponded to positions 652 through 677 in *epgB1*, and the reverse primer, B-Rev, corresponded to the reverse complement of positions 1287 through 1320 in *epgB1*. Differences between the *Bam*HI, *Bcl*I, *Sal*I, *Hin*fI, and *Hae*III restriction patterns of these fragments in *cppg1* and *epgB1* were identified using the Generunner program, and these enzymes (Pharmacia) were used to digest gel-purified (Wizard PCR preps, Promega) PCR products amplified from 2128u genomic DNA.

2.8 Phylogenetic, nucleotide, and polypeptide sequence analyses

A database search using the deduced peptide sequence for *C. purpureum epgA* was performed using the National Center for Biotechnology information BLAST WWW server. The polypeptide sequences of the five *C. purpureum* endoPGs, all of the endoPGs from ascomycete species having more than two sequenced endoPGs (*Aspergillus niger, Botrytis cinerea,* and *Sclerotinia sclerotiorum*), the four published fungal exoPG sequences, an endoPG from each of the phytopathogenic bacteria *Erwinia carotovora* and *Pseudomonas solanacearum*, and an exoPG from *P. solanacearum* were included in the phylogenetic

analysis. (See Figure 7 caption for sequence references.) As the N- and C- termini of the PGs were not well conserved in length or sequence composition, they were omitted and a conservative core of the protein sequences was used in order to obtain an unequivocal alignment for the phylogenetic analysis. The portion used in the analysis of the PGs extended from the relatively conservative sequence GXDITV (X = any amino acid) roughly 100 residues from the N-termini through the conservative sequence CSXWTW near the 3' end of the polypeptides. These sequences were aligned using the Multiple Alignment Program (Huang, 1994), using the BLOSUM50 weight matrix, a gap opening penalty of 12, gap extension penalty of 1, and constant gap penalty of 20. For each insertion/deletion, a single gap character was introduced (irrespective of the length of the gap) and positioned to uniquely identify different gaps. Maximum parsimony phylogenetic trees were created using a heuristic search in PAUP 4.0b4a (Swofford, 2000) with simple, step-wise sequence addition and tree bisection reconnection branch swapping (TBR), and saving all optimal trees for subsequent branch swapping steps (MULPARS). The number of changes was determined as the minimum number of nucleotide changes required to convert one amino acid sequence to another. No outgroup taxa were specified, and clade stability was assessed by 1000 bootstrap replications.

Percent nucleotide and amino acid identities between the *C. purpureum* endoPGs were calculated as the percentage of nucleotides or amino acids shared between sequence pairs, based on the alignment in Fig. 3. Percent identities between all other polygalacturonase sequences were obtained from pairwise alignments created in the ClustalW multiple alignment program provided by the EMBL European Bioinformatics Institute website (http://www.ebi.ac.uk/clustalw/).

Sequence motifs were identified and isoelectric points were predicted using the program Generunner (Hastings Software Inc.). Signal peptidase sites were predicted using PSORT II for animal and yeast sequences (http://psort.nibb.ac.jp:8800/). Codon usage of the endoPG sequences was determined using the Virtual Genome Center DNA sequence translation website (http://alces.med.umn.edu/webtrans.html).

2.9 RNA extraction

Mycelia for RNA extraction were grown for 2 weeks as for DNA extraction, filtered using miracloth (Calbiotech), rinsed several times with sterile water, fragmented in a sterile Waring blender for 10 seconds, and transferred into one of six different media. All six media were prepared using 3.0 g D-L malic acid, 2.1 g $NaNO_3$, 1 g KH_2PO_4, 0.5 g yeast extract, and 0.1 g $MgSO_4 \cdot 7H_2O$ in 1 L dH_2O, and were supplemented with one of: 1) no carbon source (minimal media), 2) 2% glucose, 3) 2% sucrose 4) 1% citric pectin (Sigma), 5) 1% citric pectin plus 2% glucose, or 6) 1% citric pectin plus 2% sucrose, and adjusted to pH 5.0 with 5 M NaOH. These cultures were grown at 23°C and harvested for RNA extraction after 2 days, 4 days and 8 days. The mycelia were filtered using miracloth, rinsed with dH_2O, thoroughly blotted on miracloth and absorbent paper towels, and the fresh weights determined.

Approximately equal weights of blotted mycelia from the six media types were homogenized in 2 ml TRIZOL reagent (GibcoBRL) for 30 seconds. 500 µl chloroform was added, thoroughly mixed and allowed to incubate at room temperature with the homogenized mycelia for 5 minutes. The mixture was centrifuged at 10,000g for 10 minutes, and the supernatant added to 1 ml isopropanol. After a 5 minute incubation, the RNA was pelleted at

10,000g for 3 minutes. The pellet was washed twice with 75% ethanol in DEP-C treated H_2O, dried at room temperature, suspended in 50 µl DEP-C treated dH_2O, and dissolved at 65°C for 10 minutes.

2.10 Northern hybridization

Total RNA isolated from cultures of *C. purpureum* was mixed with 50 µl of freshly prepared loading buffer (0.75 ml deionized formamide, 0.15 ml 10X MOPS [0.2 M 3-(N-morpholino) propanesulfonic acid, 50mN sodium acetate, 10 mM EDTA, pH 7.0 and autoclaved], 0.24 ml formaldehyde, 0.1 ml RNase-free water, 0.5 ml glycerol, 0.08 ml 10% bromophenol blue) and 1 µl ethidium bromide, and divided equally between four wells to create four identical Northern blots. The gels were composed of 1% agarose, 1X MOPS buffer dissolved in DEP-C treated water, and were supplemented with 5% v/v 37% formaldehyde. Electrophoresis was in 1X MOPS buffer at 50V. Approximately equal loading of the RNA in each lane was determined by ethidium bromide staining. Gels were photographed and soaked 5 minutes in 0.05% NaOH made up in 1X SSC, soaked in 10X SSC for 40 minutes, and blotted to nylon membrane (GeneScreen Plus, DuPont) for 12 hours. Northern blots were prehybridized 2 hours and hybridized 12 hours at 42° C in 50% (v/v) deionized formamide, 0.47X Denhardt's solution, 4.7X SSPE, 0.1% (w/v) SDS, 10% dextran sulfate, then washed 3 times at room temperature in 1X SSC, 0.1% (w/v) SDS and twice at 55° C in 0.1X SSC, 0.1% SDS. The *epgA*, *epgB*, *epgC* and *epgD* probes were prepared as described for DNA blotting. The size of the hybridizing RNA bands were estimated based upon the position of the small and large ribosomal RNA bands in the photographs taken of the

gels. Small ribosomal RNA subunit sequences were available on GenBank for twenty-four basidiomycete species, and all were well conserved in sequence and length (1734 to 1806 bases). It is likely that the *C. purpureum* small RNA subunit is also in this size range. The *C. purpureum* endoPG mRNAs were estimated to be between 1300 and 1450 bases in length, based on the cDNA sequence (1434bp) of the *epgB* homologue in isolate ASP-4B (Miyairi et al., 1997). Single hybridizing bands slightly smaller than the small ribosomal RNA band were assumed to correspond to mRNA of the *C. purpureum* endoPG under investigation.

RESULTS AND DISCUSSION

3.1 Identification of a multigene family of endoPGs in *C. purpureum*

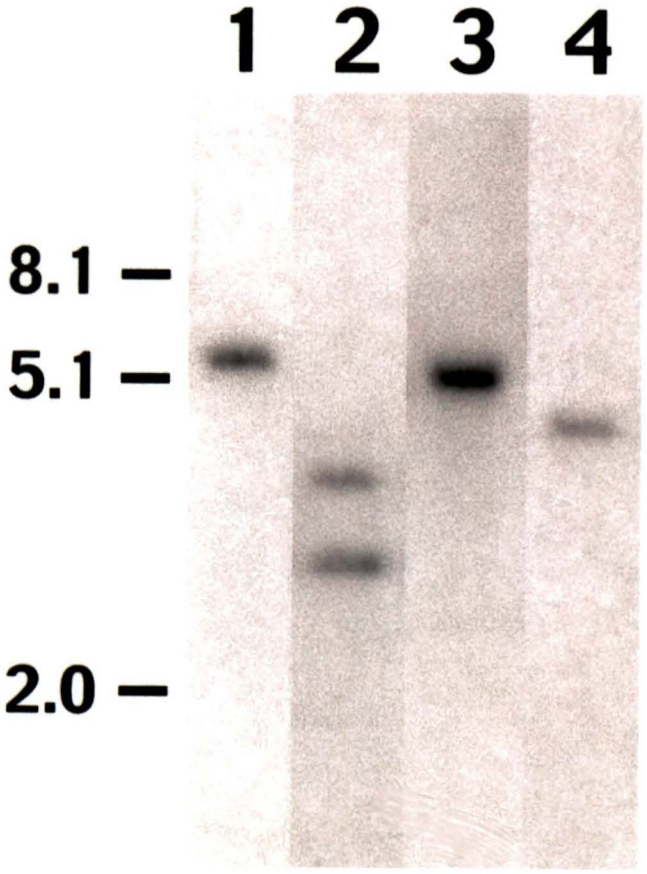
3.1.1 Discovery of new endoPG genes in C. purpureum

To detect new endopolygalacturonase genes in *C. purpureum*, the degenerate oligonucleotide primers Cpepgs-For/Cpepgs-Rev were designed based on two highly conserved regions of several ascomycete and two *C. purpureum* endoPGs. Amplification using this primer pair resulted in a major band of the expected size (~460 bp). Of twenty-six sequenced amplification products, ten aligned perfectly with *epgA*, two aligned closely with the Miyairi EndoPG I cDNA, five corresponded to a new, third endoPG, and two corresponded to another new, fourth endoPG. The remaining six clones did not show homology to any known endoPGs. The gene corresponding to the fragment aligning with EndoPG I cDNA was named *epgB*, the first new endoPG gene was named *epgC*, and the second new endoPG gene was named *epgD*. To further confirm that all the members of this gene family were uncovered, the identity of thirty additional amplification products were subjected to restriction analysis. Twenty-four of these matched the restriction pattern of one of the four endoPG fragments, and the remaining six corresponded to sequenced fragments which were not endoPGs.

3.1.2 Copy number of C. purpureum *endoPG genes*

Genomic Southern blots of *C. purpureum* 2128u DNA digested with *Bgl*II showed one hybridizing band when probed with PCR amplified fragments of *epg*s *A*, *C*, or *D*, and two hybridizing bands when probed with a PCR amplified fragment of *epgB* (Fig. 3). This suggested the presence of one copy each of *epgA*, *epgC* and *epgD*, and two copies of *epgB*. To confirm the presence of two copies of *epgB*, Southern blots of 2128u DNA were digested with *Eco*RI and probed with the *epgB* PCR fragment. These blots also showed one hybridizing band for *epg*s *A*, *C*, and *D*, and two hybridizing bands for *epgB*, at approximately 6.4 kb and 6.2 kb. The *Sac*I digest showed two *epgB*-hybridizing bands at 10 kb and 5.5 kb. Since the *epgB* probe DNA did not contain *Bgl*II, *Sac*I, or *Eco*RI restriction sites, it was concluded that each hybridizing band corresponded to a unique *epgB* gene locus.

FIGURE 3. Southern analysis of endoPG copy number in *Chondrostereum purpureum*. Total genomic DNA from isolate 2128u digested with *Bgl*II and probed with the ^{32}P-dCTP-labeled PCR fragment of *epgA* (lane 1), *epgB* (lane 2), *epgC* (lane 3), and *epgD* (lane 4). The *epgB* probe sequence was more than 99% identical between *epgB1* and *epgB2*, so this probe did not distinguish between the two *epgB* genes.


1
2
3
4
8.1
5.1
2.0

3.1.2 Sequencing the epgB1 and epgB2 genes

A recombinant λ clone hybridizing to the *epgB* probe was subcloned into plasmid vectors and sequenced. The DNA sequence of approximately 1300 bp of upstream promoter, 1722 bp *epgB1* coding region (Fig. 4), and 492 bp of terminator sequence were obtained. Since the *epgB1* sequence was similar (91.5% nucleotide identity) to the cDNA sequence of an endoPG (EndoPG I) identified in *C. purpureum* isolate ASP-4B (Miyairi et al., 1997), it was uncertain whether the differences between the *epgB1* and ASP-4B sequences were due to inter-strain differences or whether the two genomic fragments which hybridized to the *epgB* probe corresponded to the ASP-4B EndoPG I and *epgB1* gene. Recent deposition of the full length sequence of the ASP-4B gene encoding EndoPG I (named *cppg1*) into GenBank (Accession number AB043873) allowed comparison of the full length sequences of *epgB1* and *cppg1*, and several restriction sites which differed between the two genes were identified. A primer pair, B-For and B-Rev, specific to both *epgB* and *cppg1* but not the other three endoPG genes, was used to amplify a ~670 bp internal segment of 2128u genomic DNA, and numerous restriction enzymes were unable to segregate the two *epgB* gene copies in the PCR products. All of the restriction digest patterns corresponded to that expected for *epgB1* from the known sequence data. There were no sequences amplified that were more similar to the digest pattern espected for ASP-4B *cppg1*. This suggested the two *epgB* genes were very similar to each other, and indicated that the divergence between the ASP-4B *cppg1* and 2128u *epgB* genes was due to inter-strain differences.

Another tactic to uncover the second *epgB* gene was to return to the genomic library and select more clones. A second recombinant λ clone hybridizing to the *epgB* probe was DNA-purified and subcloned into plasmid vectors, and 105 bp of promoter sequence, 1737 bp

of *epgB2* coding region (Fig. 4), and 570 bp of terminator sequence were obtained for this gene. In the first 1527 bp of *epgB1* and *epgB2*, only four single synonymous changes were noted between the two genes, but from positions 1528 through to the 3' ends of the genes, there were marked differences between them. Within this ~200 bp region there were 22 nucleotide and 7 amino acid changes. In addition, two small insertions 3 and 2 amino acids in length were found in *epgB2*, at residues 359 and 368, respectively (Fig. 4).

In the 105 bp immediately upstream of the *epgB1* and *epgB2* start codons, four nucleotides differed between the sequences and one nucleotide was deleted in *epgB1*. All five changes were upstream of position -66 (Fig. 5a). The first 190 bp of terminator sequence was well conserved between *epgB1* and *epgB2*, with 15 single nucleotide differences between the two 3' regions. A 25 bp insertion in *epgB2* relative to *epgB1* 190 bp downstream of the 3' ends of the two genes was followed by a well conserved segment 92 bp long. Downstream of this conserved segment no similarities were evident between the two sequences (Fig. 5b).

3.1.3 Subcloning and sequencing the epgC and epgD genes

A total of 3.5 kb sequence information was obtained for *epgC*, which consisted of 1500 bp of promoter sequence, the 1602 bp coding region (Fig. 4), and 559 bp of terminator sequence. For *epg*D, 3.5 kb of sequence information was also obtained, consisting of 1500 bp of promoter sequence, the 1611 bp coding region (Fig. 4), and 421 bp of the 3' flanking region.

3.2 Sequence identity between the *C. purpureum* endoPGs

The alignment of the five endoPG DNA sequences isolated from *C. purpureum* 2128u and their deduced polypeptides are shown in Fig. 4. Among the polypeptide sequences, 200 of the 362 to 408 amino acids were conserved in all five isozymes. The polypeptides encoded by *epgB1* and *epgB2* differed by 12 amino acids, and shared 97.1 % identity. Pairwise sequence identity between the deduced *epgA*, *epgB1* or *epgB2*, *epgC*, and *epgD* polypeptides varied between 61.8 and 80.0 % (Table 1). This is similar to the range of amino acid identity between members of ascomycete endoPG multigene families (Parenicová et al., 1998; Wubben et al., 1999). The endoPG genes of *C. purpureum* shared from 25 to 41% amino acid identity with ascomycete endoPG sequences.

TABLE 1. Percent amino acid (aa) identity between the five *C. purpureum* endoPGs. Values were obtained from pairwise alignments based on the alignment shown in Fig. 4. The first value is the percent identity calculated when 3' extensions were included and treated as mismatched amino acids, and the value in brackets is the percent identity calculated when 3' extensions were omitted.

endoPG	EPGB1	EPGB2	EPGC	EPGD
EPGA	62.5 (69.6)	61.8 (69.6)	65.6 (65.8)	65.8 (67.4)
EPGB1		97.1	62.8 (69.8)	66.8 (73.3)
EPGB2			62.5 (69.6)	66.8 (73.3)
EPGC				80.0 (81.8)

FIG. 4. Sequence alignment of the *C. purpureum* endoPG genes *epgA*, *epgB1*, *epgB2*, *epgC*, and *epgD* and their deduced peptides. Introns are indicated by lower case letters. The conserved fungal polygalacturonase motif containing the active site histidine residue is boldfaced, as is the conserved R[I/V]K motif. Potential N-glycosylation sites are underlined. The differences between *epgB1* and *epgB2* are double underlined in both the gene and polypeptide sequences. Gene, upstream and downstream sequences for *epgA*, *epgB1*, *epgB2*, *epgC* and *epgD* are available in GenBank under accession numbers AF237653, AF237654, AF348967, AF237655, and AF237656, respectively.

```
EPGA     M  P  S  L  S  S  I  L  K  G  L  G  D  V  T  L  F  A  S  V  -  A  T  V  T  A  V    26
EPGB1    M  T  S  F  S  S  L  F  K  L  L  -  -  -  -  I  L  V  S  V  T  A  A  A  P  S  -    22
EPGB2    M  T  S  F  S  S  L  F  K  L  L  -  -  -  -  I  L  V  S  V  T  A  A  A  P  S  -    22
EPGC     M  S  S  I  S  S  L  L  K  L  L  A  -  -  -  L  S  S  V  V  -  A  A  M  P  S  R    23
EPGD     M  S  S  F  S  S  L  I  K  H  L  A  -  -  -  L  L  S  L  V  -  A  A  A  P  S  -    22

epgA     ATGCCTTCGTTATCCTCGATCCTTAAGGGCCTTGGCGACGTCACCCTTTTTGCCTCTGTG---GCTACTGTAACGGCCGT    77
epgB1    ATGACTTCGTTCTCTTCTCTGTTCAAACTCCTG----------ATTTTGGTGTCTGTCACTGCGGCAGCACCGTCT--    66
epgB2    ATGACTTCGTTCTCTTCTCTGTTCAAACTCCTG----------ATTTTGGTGTCTGTCACTGCGGCAGCACCGTCT--    66
epgC     ATGTCTTCCATCTCTTCTCTTCTCAAACTCCTGGCC---------CTCTCTTCGGTTGTT---GCGGCAATGCCTTCCAG    68
epgD     ATGTCTTCCTTCTCTTCCCTGATCAAACACCTTGCT---------CTTTTGTCTCTCGTT---GCCGCAGCCCCGTCC--    66

EPGA       P  -  K  R  A  S  C  T  V  A  S  V  S  D  A  A  N  I  S  G  C  T  S  V  T  I     51
EPGB1      -  -  K  R  A  T  C  T  V  K  S  V  D  D  A  K  D  I  A  G  C  S  A  V  T  L     46
EPGB2      -  -  K  R  A  T  C  T  V  K  S  V  D  D  A  K  D  I  A  G  C  S  A  V  T  L     46
EPGC       E  S  K  R  A  T  C  T  V  K  S  V  A  D  S  S  D  I  A  G  C  S  V  V  T  I     49
EPGD       -  -  K  R  A  T  C  T  V  K  S  V  A  D  S  N  D  I  A  G  C  A  A  V  T  I     46

epgA     GCCG---AAGCGAGCATCCTGCACCGTTGCCTCAGTTAGCGATGCTGCAAATATTTCGGGATGTACCTCTGTGACCATAA   154
epgB1    -------AAACGTGCTACTTGCACAGTCAAGTCTGTCGATGATGCCAAGGACATTGCAGGCTGCAGTGCGGTCACTCTGA   139
epgB2    -------AAACGTGCTACTTGCACAGTCAAGTCTGTCGATGATGCCAAGGACATTGCAGGCTGCAGTGCGGTCACTCTGA   139
epgC     AGAGTCTAAACGCGCTACTTGCACCGTCAAATCCGTTGCCGACAGCAGTGACATCGCTGGCTGCAGTGTTGTCACGATTG   148
epgD     -------AAACGTGCTACTTGTACCGTCAAGTCTGTCGCCGACAGCAATGACATCGCCGGGTGCGCTGCTGTCACCATTG   139

EPGA     K  S  F  T  V  P  S  G                                                             59
EPGB1    N  A  F  T  V  P  A  G                                                             54
EPGB2    N  A  F  T  V  P  A  G                                                             54
EPGC     D  S  F  T  V  P  A  E                                                             57
EPGD     D  S  F  T  V  P  A  G                                                             54

epgA     AATCGTTTACCGTACCTTCGGGCCgtgagtgctctcattttggttgtgcggttactgtggctcatcgaatccctag----   230
epgB1    ATGCGTTCACTGTACCTGCTGGAAgtaagcattgtctccgtcgctgcatgtaatggtcctattcaacgatttcaacag--   217
epgB2    ATGCGTTCACTGTACCTGCTGGAAgtaagcattgtctccgtcgctgcatgtaatggtcctattcaacgatttcaacag--   217
epgC     ACTCGTTCACTGTGCCTGCTGAAAgtaagcccatcttcgcagcgaagtatactctttttctaacaataccttttgaag---   225
epgD     ACTCGTTCACCGTACCTGCTGGAAgtaagtacttgcttccagtctactagtatacctctcctctaaacgtgtggtatcac   219

EPGA          Q  T  L  V  L  N                                                             65
EPGB1         K  T  L  V  L  N                                                             60
EPGB2         K  T  L  V  L  N                                                             60
EPGC          N  T  V  I  M  K                                                             63
EPGD          S  T  I  D  M  K                                                             60

epgA     -----AAACTCTTGTTCTAAACgtttgtacaccatgagaaaaccaatgctagtttctgattatctatccgacatgttagC   305
epgB1    -----AGACCCTTGTGTTGAACgtgcgttcttactctccacgaacacagatttcagtctattgaacatttccacag---C   289
epgB2    -----AGACCCTTGTGTTGAACgtgcgttcttactctccacgaacacagatttcagtctattgaacatttccacag---C   289
epgC     -----ACACTGTAATCATGAAGgtatgctaaggcttctcaccgaaatctaaatccttttcaccattatttcctcag--C   298
epgD     ccaagGTACCATCGACATGAAGgcaagtgtagttctccagagtatttgacacccctcgcataaccattcctttag----C   295

EPGA     P  S  D  G  T  T  V  A  M  V  G  D  V  T  F  A  K  T  T                          84
EPGB1    P  D  K  G  A  T  V  T  M  A  G  D  I  T  F  A  K  T  T                          79
EPGB2    P  D  K  G  A  T  V  T  M  A  G  D  I  T  F  A  K  T  T                          79
EPGC     P  D  S  G  A  T  V  T  M  T  G  D  I  T  F  D  K  T  T                          82
EPGD     P  D  S  G  A  T  V  T  M  A  G  D  L  T  F  A  K  T  T                          79

epgA     CCTCAGACGGTACCACCGTAGCTATGGTTGGCGATGTCACATTTGCAAAGACGACATgtaggtggtctcttgttaaaaat   385
epgB1    CCGACAAGGGCGCTACTGTCACCATGGCCGGCGACATTACATTCGCAAAGACCACCTgtgagtgttcctagaatagatct   369
epgB2    CCGACAAGGGCGCTACTGTCACCATGGCCGGCGACATTACATTCGCAAAGACCACCTgtgagtgttcctagaatagatct   369
epgC     CTGACAGTGGCGCGACTGTAACCATGACCGGCGACATCACATTCGACAAGACTACTTgtgagtgtgttcgccacattcct   378
epgD     CCGACAGTGGCGCGACTGTCACCATGGCCGGTGATCTTACATTTGCAAAGACCACTCgtgagtactttggatattctccg   375

EPGA                                       S  S  G  P  L  F  T  I  D  G  S  N  I  K  F  K   100
EPGB1                                      L  D  G  P  L  V  T  I  D  G  T  S  I  N  F  V    95
EPGB2                                      L  D  G  P  L  V  T  I  D  G  T  S  I  N  F  V    95
EPGC                                       L  D  G  P  L  F  T  I  Q  G  T  D  I  K  F  D    98
EPGD                                       L  D  G  P  L  F  T  I  E  G  T  N  I  K  F  V    95

epgA     gctatcccagttaacattctctattaag------CGTCTGGTCCTCTGTTTACCATTGATGGCTCGAACATCAAATTTAA   459
epgB1    tccgagcagttgctgaagtgtttctcccag----TGGACGGTCCCCTCGTCACTATTGATGGGACGAGCATTAACTTCGT   445
epgB2    tccgagcagttgctgaagtgtttctcccag----TGGACGGTCCCCTCGTCACTATTGACGGGACGAGCATTAACTTCGT   445
epgC     ctatgagcagctcatgactctgagcttctcccagTGGATGGCCCTCTCTTTACTATCCAGGGGACCGATATCAAATTTGA   458
epgD     actgcggagctaagaaaaatttctcacag-----TGGACGGTCCTCTCTTTACTATTGAGGGGACTAACATCAAGTTTGT   450
```

```
EPGA       G  A  G  H  K  F  D                                                                  107
EPGB1      G  A  D  H  K  F  D                                                                  102
EPGB2      G  A  D  H  K  F  D                                                                  102
EPGC       G  G  N  H  K  F  D                                                                  105
EPGD       G  A  N  H  K  F  D                                                                  102

epgA    AGGCGCAGGACATAAGTTCGACgtatgtttaagttcgcggaatcttcccttgtcttacctgacatttttttgttcgtag--   537
epgB1   TGGGGCCGACCACAAATTCGATgtacggtaactcactatggccacctctcctagctaatctatttgttgcag--------   517
epgB2   TGGGGCCGACCACAAATTCGATgtacggtaactcactatggccacctctcctagctaatctatttgttgcag--------   517
epgC    CGGTGGCAACCATAAATTTGATgtatgtgcctttagtatcgtcctattgactactctcagttgacgccaatcctcag---   535
epgD    CGGTGCGAACCACAAGTTCGATgtatgtactaccccgattgccttttttgaacacattgatttaacatcgaacctcgtcag   530

EPGA       G  N  G  A  K  Y  W  D  G  Q  G  T  N  G  G  V  T  K  P  H  P  F  L  K  F  K  G   134
EPGB1      G  N  G  A  L  Y  W  D  G  L  G  T  N  N  G  T  H  K  P  H  P  F  L  K  I  K  G   129
EPGB2      G  N  G  A  L  Y  W  D  G  L  G  T  N  N  G  T  H  K  P  H  P  F  L  K  I  K  G   129
EPGC       G  N  G  T  V  Y  W  D  G  L  G  G  N  G  G  N  H  K  P  H  P  F  L  K  F  K  G   132
EPGD       G  N  G  A  M  Y  W  D  G  L  G  T  N  N  G  T  H  K  P  H  P  F  L  K  F  K  G   129

epgA    GGAAATGGCGCAAAGTATTGGGATGGGCAGGGGACTAACGGTGGCGTTACCAAACCCCATCCTTTCCTCAAATTCAAAGG   617
epgB1   GGAAATGGCGCTTTGTATTGGGATGGACTGGGAACCAACAATGGCACTCACAAGCCGCACCCGTTCCTCAAGATCAAAGG   597
epgB2   GGAAATGGCGCTTTGTATTGGGATGGACTGGGAACCAACAATGGCACTCACAAGCCGCACCCGTTCCTCAAGATCAAAGG   597
epgC    GGAAATGGTACAGTGTACTGGGATGGCTTAGGAGGCAACGGTGGAACCCATAAGCCACATCCTTTCCTCAAATTCAAAGG   615
epgD    GGAAATGGTGCTATGTACTGGGACGGCTTGGGAACTAACAATGGAACTCACAAGCCGCATCCGTTCCTCAAGTTCAAGGG   610

EPGA       S  G  Q  Y  S  S  F  T  V  L  N  S  P  A  Q  A  I  S  I  G  N  S  D  G  -  L    159
EPGB1      S  G  T  Y  K  K  F  T  V  L  D  S  P  A  Q  A  I  S  V  G  P  T  D  A  H  L    155
EPGB2      S  G  T  Y  K  K  F  T  V  L  D  S  P  A  Q  A  I  S  V  G  P  T  D  A  H  L    155
EPGC       S  G  S  Y  S  N  F  E  V  L  N  S  P  A  H  A  I  S  V  G  P  T  T  G  N  I    158
EPGD       S  G  N  Y  A  N  F  E  V  L  N  S  P  A  Q  A  I  S  I  G  N  T  N  G  N  I    155

epgA    ATCAGGTCAATACTCGAGTTTCACTGTCCTAAACAGCCCTGCGCAGGCAATATCTATCGGAAATAGTGACGGC---CTAA   694
epgB1   ATCGGGCACCTACAAGAAATTCGAAGTTCTCAACAGCCCCGCACAAGCAATATCCGTTGGCCCTACCGACGCACACCTTA   677
epgB2   ATCGGGGACCTACAAGAAATTCGAAGTTCTCAACAGCCCCGCACAAGCAATATCCGTTGGTCCTACCGACGCACACCTTA   677
epgC    TTCTGGCAGCTATTCAAACTTTGAGGTGCTCAATAGCCCTGCCCATGCCATATCTGTTGGACCTACGACAGGGAACATTC   695
epgD    TTCGGGCAATTACGCCAACTTTGAGGTCCTCAACAGTCCCGCTCAAGCCATCTCCATTGGAAACACCAACGGAAATATCG   690

EPGA       T  F  D  T  V  T  V  D  N  S  A  G  D  S  G  S  L  G  H  N  T  D                 181
EPGB1      T  F  D  G  I  T  V  D  D  F  A  G  D  T  K  N  L  G  H  N  T  D                 177
EPGB2      T  F  D  G  I  T  V  D  D  F  A  G  D  T  K  N  L  G  H  N  T  D                 177
EPGC       L  F  D  T  V  T  I  D  N  S  A  G  D  V  D  S  L  G  H  N  T  D                 180
EPGD       V  F  D  T  V  T  V  D  D  S  A  G  D  T  N  S  Q  G  H  N  T  D                 177

epgA    CCTTTGACACTGTCACTGTGGACAATAGTGCCGGAGATAGTGGTAGTCTCGGGCACAACACTGATgtctgtacatttttt   774
epgB1   CCTTCGATGGGATCACTGTCGACGACTTCGCGGGAGACACCAAGAATCTAGGCCACAACACTGACgtaggtttttatctc   757
epgB2   CCTTCGATGGGATCACTGTCGACGACTTCGCGGGAGACACCAAGAATCTAGGCCACAACACTGATgtaggtttttatctc   757
epgC    TTTTTGATACTGTCACTATAGACAACAGTGCAGGAGATGTCGACAGTCTAGGTCACAATACTGACgtgcgcttcctactc   775
epgD    TCTTTGATACTGTCACTGTCGACGATAGTGCGGGGGACACCAACAGTCAAGGCCATAACACTGATgtacgatgccttgcc   770

EPGA                                                                G  F  D  V  S  A  D  N  V  T   191
EPGB1                                                               G  F  D  V  S  A  N  D  V  T   187
EPGB2                                                               G  F  D  V  S  A  N  D  V  T   187
EPGC                                                                G  F  D  V  S  A  D  N  V  T   190
EPGD                                                                G  F  D  V  S  A  D  N  V  T   187

epgA    ctcagtgcgattacatctattcatgatacctacctttcctag---------GGCTTCGATGTATCTGCGGACAATGTTAC   845
epgB1   gctcggtggcagcagcacttgtgctgattattgccctccttcag--------GGTTTCGACGTCTCCGCCAACGATGTGAC   830
epgB2   gctcggtggcagcagcacttgtgctgattattgccctccttcag--------GGTTTCGACGTCTCCGCCAACGATGTGAC   830
epgC    tcatggactacagtccccgtactaatactctgtccacgccgttcacgttagGGATTTGATGTATCTGCCGATAATGTGAC   855
epgD    ctttgcgcctcagctcttgtactaatcctatccaaattctccttcag----GGTTTTGATGTGTCCGCTGACAATGTGAC   846

EPGA       I  Q  N  S  V  V  K  N  Q  D  D  C  I  A  I  N  D  G  S  N  I  V                 213
EPGB1      I  Q  N  C  I  V  K  N  Q  D  D  C  I  A  I  N  D  G  N  N  I  R                 209
EPGB2      I  Q  N  C  I  V  K  N  Q  D  D  C  I  A  I  N  D  G  N  N  I  R                 209
EPGC       I  Q  N  C  I  I  K  N  Q  D  D  C  I  A  I  N  N  G  K  N  I  Q                 212
EPGD       I  K  N  C  I  V  K  N  Q  D  D  C  I  A  I  N  D  G  K  N  I  Q                 209

epgA    CATCCAGAATAGTGTTGTCAAAAACCAAGATGACTGCATTGCCATCAACGATGGCTCTAATATTGTgtatgtgtgcctcc   925
epgB1   CATCCAGAATTGCATTGTTAAGAACCAGGACGACTGCATCGCAATCAATGACGGCAAGAACATTCGgtatcttcagttct   910
epgB2   CATCCAGAATTGCATTGTTAAGAACCAGGACGACTGCATCGCAATCAATGACGGCAAGAACATTCGgtatcttcagttct   910
epgC    AATCCAGAACTGTATTATCAAAAATCAGGACGATTGTATTGCCATCAATAATGGAAAGAACATCCAgtatgtctattctt   935
epgD    TATTAAGAACTGCATCGTCAAGAACCAGGACGACTGCATTGCAATCAACGACGGAAAGAATATCCAgtatgtccatttac   926
```

```
EPGA                                                                F  Q  N  N  Q  C  S  G  G  H    223
EPGB1                                                               F  E  N  N  Q  C  S  G  G  H    219
EPGB2                                                               F  E  N  N  Q  C  S  G  G  H    219
EPGC                                                                F  L  N  N  Q  C  S  G  G  H    222
EPGD                                                                F  L  N  N  Q  C  S  G  G  H    219

epgA    ttcttgtccgcaactcatccacgactggcttcttattgcccttgcag-TTTCCAAAACAATCAGTGCTCTGGAGGGCACG 1004
epgB1   tctcctgttcatgtgctggaatggtggctgatatgtgtttgtccacagTTTCGAGAACAACCAGTGCTCTGGAGGACACG  990
epgB2   tctcctgttcatgtgctggaatggtggctgatatgtgtttgtccacagTTTCGAGAACAACCAGTGCTCTGGAGGACACG  990
epgC    cctaactctttttcgcgctggacgacaccgactgacttcgctgcag--GTTCTTGAACAATCAGTGCTCAGGGGGCCATG 1013
epgD    ccgctcatttaaaactgcaaggcgttgacctatcatcttccaccatagGTTCTTGAACAACCAATGCTCAGGAGGGCATG 1006

EPGA    G  I  S  V  G  S  I  A  S  G  K  H  V  S  G  V  A  I  K  G  N  T  V  T  N  S  M    250
EPGB1   G  I  S  I  G  S  I  A  T  G  K  H  V  S  N  V  K  I  K  G  N  T  V  T  R  S  M    246
EPGB2   G  I  S  I  G  S  I  A  T  G  K  H  V  S  N  V  K  I  K  G  N  T  V  T  R  S  M    246
EPGC    G  I  S  V  G  S  I  G  T  G  K  H  V  S  S  V  T  I  K  G  N  T  V  T  N  S  T    249
EPGD    G  I  S  V  G  S  I  A  T  G  K  H  V  S  S  V  T  I  K  G  N  T  V  T  N  S  M    246

epgA    GTATTTCTGTTGGTTCTATCGCCTCAGGAAAGCACGTATCGGGTGTAGCAATCAAAGGAAACACGGTCACAAACAGCATG 1084
epgB1   GAATTTCCATTGGCTCAATTGCCACTGGCAAGCACGTCTCCAACGTTGTCATCAAGGGTAACACGGTCACCCGCAGTATG 1070
epgB2   GAATTTCCATTGGCTCAATTGCCACTGGCAAGCACGTCTCCAACGTTGTCATCAAGGGTAACACGGTCACCCGCAGTATG 1070
epgC    GAATCTCTGTTGGATCGATTGGGACTGGAAAGCACGTCTCTTCCGTTACTATCAAAGGCAACACGGTCACCAATAGTACA 1093
epgD    GAATCTCTGTCGGGTCGATCGCGACAGGAAAACATGTCTCTTCCGTCACTATCAAAGGCAACACGGTCACCAATAGCATG 1086

EPGA     Y  G  M  R  V  K  V  K  A  A  A  T  S  A  S  V  S  A  V  T  Y  S  G  N  T  I  S    277
EPGB1    Y  G  I  R  I  K  A  Q  R  K  A  T  S  A  S  V  S  G  V  T  Y  D  G  N  T  I  S    273
EPGB2    Y  G  I  R  I  K  A  Q  R  K  A  T  S  A  S  V  S  G  V  T  Y  D  G  N  T  I  S    273
EPGC     N  G  L  R  V  K  V  Q  A  S  A  T  N  A  S  V  S  G  V  T  Y  D  N  N  T  L  S    276
EPGD     Y  G  L  R  V  K  V  Q  A  S  A  T  N  A  S  V  S  G  V  T  Y  D  G  N  T  L  S    273

epgA    TATGGCATGCGTGTCAAGGTCAAGGCTGCTGCAACCTCTGCTTCAGTCTCTGCTGTGACTTACAGTGGCAACACCATTTC 1164
epgB1   TACGGTATCCGCATCAAAGCCCAGCGCAAGGCTACCTCTGCCTCTGTGTCCGGCGTGACCTACGACGGAAACACGATCTC 1150
epgB2   TACGGTATCCGCATCAAAGCCCAGCGCAAGGCTACCTCTGCCTCTGTGTCCGGCGTGACCTACGACGGAAACACGATCTC 1150
epgC    AACGGCCTCCGTGTTAAAGTTCAAGCCTCTGCCACCAATGCTTCTGTTTCCGGTGTGACTTACGATAACAATACTCTTTC 1173
epgD    TACGGCCTCCGTGTGAAAGTTCAAGCCTCCGCTACCAATGCCTCTGTATCCGGCGTTACTTACGATGGAAACACCTTATC 1166

EPGA      A  I  A  K  Y  G  F  L  V  S  Q  S  Y  P  D  D                                  293
EPGB1     G  I  A  K  Y  G  L  L  I  S  Q  S  Y  P  D  D                                  289
EPGB2     G  I  A  K  Y  G  L  L  I  S  Q  S  Y  P  D  D                                  289
EPGC      G  I  T  K  Y  G  L  L  I  T  Q  S  Y  P  A  N                                  292
EPGD      G  I  T  K  F  G  L  L  V  S  Q  S  Y  P  A  D                                  289

epgA    TGCGATTGCCAAGTATGGCTTCCTCGTCTCGCAGTCTTACCCAGATGATGgtgcgtaacagagtcatattgaagagttat 1244
epgB1   TGGAATTGCTAAATACGGTCTCCTCATTTCGCAATCTTACCCTGATGATGgtgtgttgcgactcatccgtgttttccact 1230
epgB2   TGGAATTGCTAAATACGGTCTCCTCATTTCGCAATCTTACCCTGATGATGgtgtgttgcgactcatccgtgttttccact 1230
epgC    CGGAATCACCAAGTATGGTCTTTTGATTACACAGTCATATCCGGCAAATGgtgtgtgccacattgaaaggttttgtgtag 1253
epgD    TGGGATCACCAAATTCGGTCTTCTGGTCTCGCAATCTTACCCTGCCGATGgtacgcattgacttcgatgacttggttgtg 1246

EPGA                                  A  S  T  P  G  T  G  A  P  I  S  G  I  N  F  S  G  D    311
EPGB1                                 V  G  T  P  G  T  G  A  P  F  S  D  V  N  F  T  G  G    307
EPGB2                                 V  G  T  P  G  T  G  A  P  F  S  D  V  N  F  T  G  G    307
EPGC                                  E  G  E  A  G  T  G  G  P  I  S  D  I  N  F  T  G  G    310
EPGD                                  E  G  K  A  G  T  G  G  P  I  S  G  I  D  F  T  G  G    307

epgA    tctgatttggcattatatgggcag--CTTCAACTCCCGGAACTGGCGCCCCCATCTCGGGGATCAACTTCAGTGGTGATA 1322
epgB1   cgactgaccaccgagcttatcatcagTTGGAACTCCTGGGACTGGAGCCCCCTTTTCTGATGTCAACTTTACTGGAGGTG 1310
epgB2   cgactgaccaccgagcttatcatcagTTGGAACTCCTGGGACTGGAGCCCCCTTTTCTGATGTCAACTTTACTGGAGGTG 1310
epgC    gaaagactgacccaatcatcag----AGGGAGAAGCGGGGACTGGTGGTCCTATCTCTGACATCAACTTCACTGGAGGCG 1329
epgD    gttgctcaacttctcatttag-----AGGGAAAGGCTGGAACCGGCGGCCCCATCTCTGGTATCGACTTCACAGGAGGTG 1321

EPGA    T  T  N  I  K                                                                     316
EPGB1   A  T  T  I  K                                                                     312
EPGB2   A  T  T  I  K                                                                     312
EPGC    A  T  N  I  N                                                                     315
EPGD    A  T  N  I  K                                                                     312

epgA    CTACCAATATCAAgtacgtcttccttcggtagtggatcgaatctcctttctcaatcttatttttttattatttacttagA 1402
epgB1   CCACAACGATCAAgtaagccgctctgtaaatcactctggcatttcagctcctcaatagtattcctttctctgtatag--G 1388
epgB2   CCACAACGATCAAgtaagccgctctgtaaatcactctggcatttcagctcctcaatagtattcctttctctgtatag--G 1388
epgC    CCACAAACATCAAgtacaacatttttgtcacttctctgtgcaccatgatactcacaatgtactaatcttgtag-------C 1402
epgD    CCACAAACATCAAgtaagtgttcctcgcctgtagcgtgtcatcataggtagtaatatctgcctag--------------G 1387
```

```
EPGA     V  N  S  G  A  K  R  V  T  V  D  C  R  S  C  T  G  T  W  N  W  S  K  L  T  A  T   343
EPGB1    V  N  T  A  A  K  R  V  T  V  E  C  G  N  C  S  G  N  W  D  W  S  K  L  T  V  T   339
EPGB2    V  N  T  A  A  K  R  V  T  V  E  C  G  N  C  S  G  N  W  D  W  S  K  L  T  V  T   339
EPGC     V  T  S  K  A  K  R  V  S  V  N  C  A  A  C  T  G  N  W  D  W  S  Q  L  T  V  T   342
EPGD     V  N  T  G  S  K  R  V  S  V  N  C  A  A  C  S  G  T  W  D  W  S  Q  L  T  V  T   339

epgA     GTAAACAGTGGAGCGAAACGGGTCACTGTCGACTGTCGTAGTTGCACCGGCACGTGGAACTGGTCTAAGCTGACCGCTAC  1482
epgB1    GTCAATACTGCAGCCAAACGGGTCACTGTGGAATGCGGAAACTGCTCTGGAAATTGGGACTGGTCCAAACTCACTGTCAC  1468
epgB2    GTCAATACTGCAGCCAAACGGGTCACTGTGGAATGCGGAAACTGCTCTGGAAATTGGGACTGGTCCAAACTCACTGTCAC  1468
epgC     GTCACCAGTAAAGCGAAGCGTGTTTCCGTCAACTGTGCTGCCTGTACCGGTAACTGGGATTGGTCACAACTCACTGTTAC  1482
epgD     GTTAACACTGGATCAAAGCGTGTTTCGGTCAACTGTGCCGCCTGCAGTGGCACTTGGGACTGGTCTCAACTTACCGTCAC  1467

EPGA       G  G  S  A  G  T  I  S  S  D  K  A  K                                          356
EPGB1      G  G  K  A  G  T  I  K  S  D  K  A  K                                          352
EPGB2      G  G  K  A  G  T  I  K  S  D  K  A  K                                          352
EPGC       G  G  K  A  G  T  I  N  S  D  Q  A  Q                                          355
EPGD       G  G  K  A  G  T  I  N  S  D  Q  A  K                                          352

epgA     AGGGGGAAGTGCCGGCACTATCTCCAGTGACAAAGCTAAGgtactttgattcgaacgtactactgttagaaatgactgag  1562
epgB1    CGGTGGAAAGGCTGGTACAATTAAGAGTGACAAGGCCAAGgtcagatttagcccatcatctcagcatttcccaagctc  1548
epgB2    CGGTGGAAAGGCTGGTACAATTAAGAGTGACAAGGCCAAGgtcagatttagcccatcattttcagcatttcccttaagctc  1548
epgC     TGGCGGAAAGGCTGGGACTATCAATAGTGATCAGGCTCAGgtgaaacctgattccattcgcgtttttcttccatgcactc  1562
epgD     TGGTGGAAAGGCTGGCACTATTAACAGTGATCAAGCCAAGgtcagaataaactgcccgtctgtggtactgctttcatctc  1547

EPGA                         I  S  G  G  S  Y                                            362
EPGB1                        I  T  G  G  Q  Y  L  -  -  -  A  D  Q  P  A  G  -  -  N  D  I   368
EPGB2                        I  T  G  G  T  Y  L  E  D  D  M  D  Q  P  D  G  G  E  T  E  P   373
EPGC                         I  N  G  G  Q  Y  V                                          362
EPGD                         I  S  G  G  K  F  A  D  D  D  D  S  E  T  S                   367

epgA     cagccattctaacag---ATTTCTGGCGGCTCATACTAG                                          1598
epgB1    acttactttcctatgcagATCACTGGTGGCCAGTACTTGG---------CGGATCAGCCAGCGGGT------AACGACAT  1613
epgB2    acttactttcctatgtagATCACTGGTGGCACATACTTGGAAGATGATATGGATCAGCCTGATGGAGGGGAGACCGAACC  1628
epgC     accatcgattccatag--ATCAATGGTGGGCAATATGTGTAA                                        1602
epgD     accctgtctaccatag--ATTTCCGGTGGCAAGTTTGCAGATGACGACGACTCAGAGACATCATAA                1611

EPGB1      E  E  M  P  A  Q  D  P  N  D  P  E  D  P  D  T  A  M  Q  E  A  E  A  E  E  A    394
EPGB2      E  E  M  P  A  Q  D  P  N  D  P  E  D  P  D  T  A  M  Q  E  A  E  A  E  E  A    399

epgB1    AGAAGAGATGCCGGCTCAGGATCCTAATGACCCAGAGGACCCAGATACCGCTATGCAAGAGGCTGAGGCAGAAGAGGCTG  1693
epgB2    AGAAGAGATGCCTGCTCAGGATCCTAATGACCCAGAAGATCCAGATACCGCTATGCAAGAGGCTGAGGCAGAAGAGGCTG  1708

EPGB1      A  A  G  D  L  T  P  S  D                                                      403
EPGB2      A  A  G  D  P  T  P  S  D                                                      408

epgB1    CTGCAGGGGGACCTCACTCCTTCCGACTGA                                                    1722
epgB2    CTGCAGGGGATCCCACTCCTTCCGATTGA                                                     1737
```

FIGURE 5A. Alignment of 105 bp of promoter sequences for *epgB1* and *epgB2*. Differences are indicated by double underlining. The start codons of the genes are indicated in boldface.

FIGURE 5B. Alignment of terminator sequences of *epgB1* and *epgB2*. Differences are indicated by double underlining. The stop codons of the genes (TGA) are shown in boldface.

A)

```
                      -90              -75              -60
epgB1   GTTGACGGGGCTGA CTATAAAGAGC-CCC AGAATCAGCATTGTA GGAGCATCAG
epgB2   GTCGACGGGGCTGA CTATAAAGAGCACCC CGAACCAACATTGTA GGAGCATCAG

             -45              -30              -15          -1
epgB1   TTGTC TTCATTCAGCTGCTC TTCTACCCATACGAA TCATCGTACATTACA ATG
epgB2   TTGTC TTCATTCAGCTGCTC TTCTACCCATACGAA TCATCGTACATTACA ATG
```

B)

```
                      15               30               45               60
epgB1 TGA TTTCATTCTTTGGCT TGTTGGATATCTAAC GAGTCTGATTACTTT GTCACTCTAAATTCC
epgB2 TGA TTTCATTCTTTGCCT TGTTGGATGTCTAAC AAGTCGGATTACTTT GTCGCTCTAATTTCC

                      75               90              105              120
epgB1     TATCTTTTCATTCCG GGTTCTATGAAATAT ATATGTTAGCGAAAT CACCTCTTCTCTTCT
epgB2     TATTTTTTCACTCCG GGTTCTATGAAATAT ACATGTTAGCGAAAG CACCTCTTCCCTTTT

                     135              150              165              180
epgB1     GGATTACTAGTCTGA GAGTGTCAGATGTGC TTGTGGTGGTTGGCC TGACTCGCATGGTAT
epgB2     GGATTACTAGTCTGA GAGTGACAGATGTGC TTGTGGTGGCTGGCC TGACTCGCATGGTAT

                     195              210              225              240
epgB1     GTACAATAGG----- ---------------- -----TGGCTCCTTA ACTCTGTTATCAACA
epgB2     GTACATTAGGGGATT CCGCTACTCCTGCTT GATCTTGGCTCCCTA ACTCTGGTATCAACA

                     255              270              285              300
epgB1     AATACACGTAGTAGT GCATGCATATACTTC TGACAGTTGTGTTCA TCTTCATTCCTAGTT
epgB2     AATACACGTAGTAGT GCATGCATATACCTC TGGCAGTTGTGTTCA TCTTCATTCCTAGTT

                     315              330              345              360
epgB1     TCACTCAGTATCTTC TTGATCG-------- --------------T TTGTGCG-----AGA
epgB2     TCACTCAGTATCTTC TTGATCAGCCCAGTC AACGTCGAGGGTATC CTACCCGGGGGGAAT

                     375              390              405
epgB1     CTGATGAATGACAAC GACCTCTGGTGGTCA TTCACTCGTACTGTT TTGGCTT---------
epgB2     CACCTAACAGACGAT GAACAGATGCGTAAC GCCGCCTATGCCTAC TTGGCTTGTTGCGGAG
```

3.3 Coding region characteristics

The consensus fungal polygalacturonase active site, CXGGHGXSIGSVG, was found in all five *C. purpureum* endoPGs, with a slight modification to CSGGHGIS[V/I]GSI[A/G] (Fig. 4). The active site histidine corresponded to that identified in biochemical studies of *Aspergillus niger* polygalacturonases as essential for activity (Cooke et al., 1976; Rexová-Benková and Mracková, 1978). The RIK motif conserved in prokaryote, fungal and plant polygalacturonases was modified to RVK in three of the *C. purpureum* proteins (Fig. 4). The same modification was also found in a *Lycopersicon esculentum* (tomato plant) endoPG (Hong and Tucker, 1998), a *Botrytis cinerea* endoPG (Accession No. U68716), and an *Ophiostoma novo-ulmi* endoPG (Accession No. AF052061). This positively charged motif may be involved in ionic interactions with the carboxylate groups of the substrate (Bussink et al., 1991a).

Analysis of the codon usage of the endoPG sequences showed that 52 to 58 codons were used by each of the genes, and all four together used 60 of the possible 61 sense codons. Codons ending in a pyrimidine were preferred, with 37% ending in C and 31% ending in T.

Based on comparison of the derived *C. purpureum* amino acid sequences to those of other fungal endoPGs, the five *C. purpureum* endoPGs are likely synthesized as pre-propectinases. A two-step maturation has been suggested for the majority of endoPG proteins, with cleavage at the von Heijne (1986) peptidase cleavage site followed by monobasic or dibasic endopeptidase processing. In EPGA, the signal peptidase site was predicted to follow amino acid position 25, and in EPGB1, EPGB2, EPGC, and EPGD, this site was predicted to follow amino acid position 19. The pro-sequences are likely cleaved by a dibasic peptidase similar to the yeast KEX2 endoprotease (Calmels et al., 1991) after the Lys-Arg residues at positions 29, 24, 24, 27 and 24, in EPGs A, B1, B2, C, and D, respectively. N-terminal

sequencing of the mature protein of the EPGB homolog in isolate ASP-4B, EndoPG I, confirmed the cleavage point of the signal peptide (Miyairi et al., 1997) for this enzyme. The mature proteins were thus predicted to be 333, 379, 384, 335, and 343 amino acids long, respectively. Predicted isoelectric points varied considerably for the five mature polypeptides, being pH 7.97 for *epgA*, 4.71 for EPGB1, 4.49 for EPGB2, 5.19 for EPGC, and 4.87 for EPGD.

Relative to EPGD, the C-terminal ends of EPGB1 and EPGB2 included extensions of 39 and 41 amino acids, respectively. EPGD was slightly longer than EPGA and EPGC (Fig. 4). Aside from the C-terminal extensions, there were only three gaps in the alignment of the amino acid sequences: a deletion of one amino acid following position 158 in EPGA, and insertions at positions 359 and 366 in EPGB2.

Potential N-glycosylation sites were identified within the polypeptide sequences by searching with the motif (N-X{≠P}-[S/T]). Four potential sites were identified for EPGA, seven were identified for both EPGB1 and EPGB2, and three were identifed for both EPGC and EPGD (Fig. 4). These sites were not well conserved among the endoPG genes.

3.4 Intron position and splicing sequences

The introns of *epgB1*, *epgB2*, *epgC*, and *epgD* were identified based upon established intron positions for *epgA*, the cDNA sequence of EndoPG I (Miyairi et al., 1997), and alignments with endoPG protein sequences from other species. Intron positions were well conserved for all of the *C. purpureum* endoPGs, although the nucleotide sequences of the introns were only conserved in the two *epgB* genes (Fig. 4). The presence of nine introns in

the five *C. purpureum* endoPG genes was unusual, as the number ranged from zero to four for all previously characterized fungal PG genes. As the *C. purpureum* genes are the first endoPGs characterized from a basidiomycete, the difference in intron number may simply reflect a difference in gene organization between the ascomycetes and basidiomycetes. The introns ranged from 50 to 66 bp in length. The 5' consensus sequence, GTANGT, proposed for the introns of filamentous fungi (Ballance, 1986) was reduced to GY for the *C. purpureum* endoPG introns. The general consensus lariat sequence proposed for filamentous fungi [T/C]GCTAAC (Gurr et al., 1987) was not evident in the *C. purpureum* endoPG introns. The consensus 3' splice sequence, YAG, was consistent for all but two of the forty-five *C. purpureum* endoPG introns, where the Y was replaced by an A.

3.5 Promoter and terminator motifs

Several sequence motifs of possible functional significance were found in the 5' promoter and 3' terminator regions of each of the five endoPGs (Table 2). At least one putative TATA box was identified between 75 and 103 bp 5' of the start codon for each gene. One or two putative CAAT motifs were also identified between 63 and 300 bp 5' of the start codon of all five genes, and the sequence GGGCGG was found in the *epgA* promoter adjacent to the putative CAAT motif. CAAT motifs and GC boxes may play a role in determining the efficiency of the promoter, and can function in either orientation at distances that vary considerably from the startpoint (Lewin, 1997). A 32 bp pyrimidine segment interrupted by one adenine was found in the *epgA* promoter at positions -63 to -32. Such CT-rich regions are commonly found in the promoters of highly expressed genes (Ballance, 1986; Gurr et al., 1987). Additionally, a segment was identified in the promoter of *epgA* at positions -605 to

-584 with respect to the translation initiation codon which shared significant similarity to the upstream activating sequence UAS2 in *Saccharomyces cerevisiae* (Guarente et al., 1984). A similar motif was identified in several *Aspergillus niger* endoPG promoters (Bussink et al., 1992), and promoter deletions in the *A. niger pgaII* gene demonstrated the region containing this UAS2-like segment was necessary for high level expression of the gene (Bussink et al., 1992). The alignment of this segment in the *C. purpureum* epgA promoter with the yeast UAS2 and *A. niger pgaII*, *pgaI* and *pgaC* promoter segments is shown in Fig. 6. 14 of 23 bp were shared between the *epgA* and *S. cerevisiae* segments, although a single base pair insertion was necessary in the *epgA* segment to attain this level of identity.

The sequence flanking the start codon of the the *C. purpureum* endoPGs was T[T/A]ACN**ATG**NC. The consensus site proposed by Ballance (1986) for filamentous fungi was (TCA[C/A][A/C]**ATG**[G/T]C). The adenine at position -3 was identified as a key consensus site by Kozak (1984), who found a purine (79% adenine) at position -3 in 205 of 211 eukaryotic genes. For each endoPG there was an AATAAT or AATATA sequence 74 to 105 bps downstream from the TAG stop codon in the 3' nontranslated region. This is similar to the polyadenylation signal AATAAA found in the mRNA of higher eukaryotes (Ballance, 1986).

TABLE 2. Sequence motifs of possible functional significance in the promoters and terminators of the four *Chondrostereum purpureum* endoPG genes.

endoPG	Putative TATA box(es) and bp 5' of coding region	Putative CAAT box(es): bp 5' of coding region	GGGCGG motif(s): bp 5' of coding region	CT-rich region: bp 5' of coding region	Putative polyadenylation signal and bp 3' of coding region
*epg*A	TATAA 103 TAAAT 79	300[a]	294[a]	-63 to -32	AATAAT 75
epgB1	TATAA 88	63[a] 135	-	-	AATATA 85
epgB2	TATAA 89	63[a]	unknown	unknown	AATATA 85
epgC	TATAA 75	123[a]	-	-	AATATA 102
epgD	TATAA 79	160	-	-	AATATA 105

[a] CAAT or GGGCGG motif is in reverse orientation relative to the gene coding region, and position 5' of the coding region is from the C of CAAT or first G of GGGCGG to the translation start site of the gene.

FIGURE 6. Aligned putative upstream activating sequences from the *C. purpureum epgA* promoter, the *Saccharomyces cerevisiae CYC1* promoter, and the *Aspergillus niger pgaII, pgaI and pgaC* promoters. The positions of the 5' start and the 3' end of the sequences are given with respect to the translation-initiation codon and also indicate the orientation of the sequences (Bussink et al., 1992). The most well conserved portion of this sequence is indicated by asterisks above the aligned sequences.

```
                                 ********
Cp epgA        -605   GAGGATGAGTTG-TGGATCCATA  -584
Sc CYC1        -226   GAGCGTTGATTGGTGGATCAAGC  -204
An PGA II      -632   GACCGTTCATTGGTGGAACTAGC  -610
An PGA I       -206   CCTCCATTATTGGTGGAAAGACC  -228
An PGA C       -359   AGGTAATCATTGGTGGAGATAAC  -381
```

3.6 Evolution of *C. purpureum* endoPGs and PGs from other species

The single most parsimonious phylogenetic tree created using the amino acid sequences of 29 polygalacturonases clearly separated the bacterial and fungal PGs, and resolved the endo- and exo- specific enzymes into distinct groups. The five *C. purpureum* endoPGs formed their own monophyletic group; whereas the ascomycete endoPGs from multigene families were split between more than one monophyletic group. All of the ascomycete endoPGs were contained in one clade separate from the *C. purpureum* clade (Fig. 7). This suggested that (1) divergence of the basidiomycetes and ascomycetes occurred before the gene duplications leading to existing multigene families of endoPGs, and (2) the ancestral genes of certain ascomycete endoPGs were duplicated prior to ascomycete speciation events.

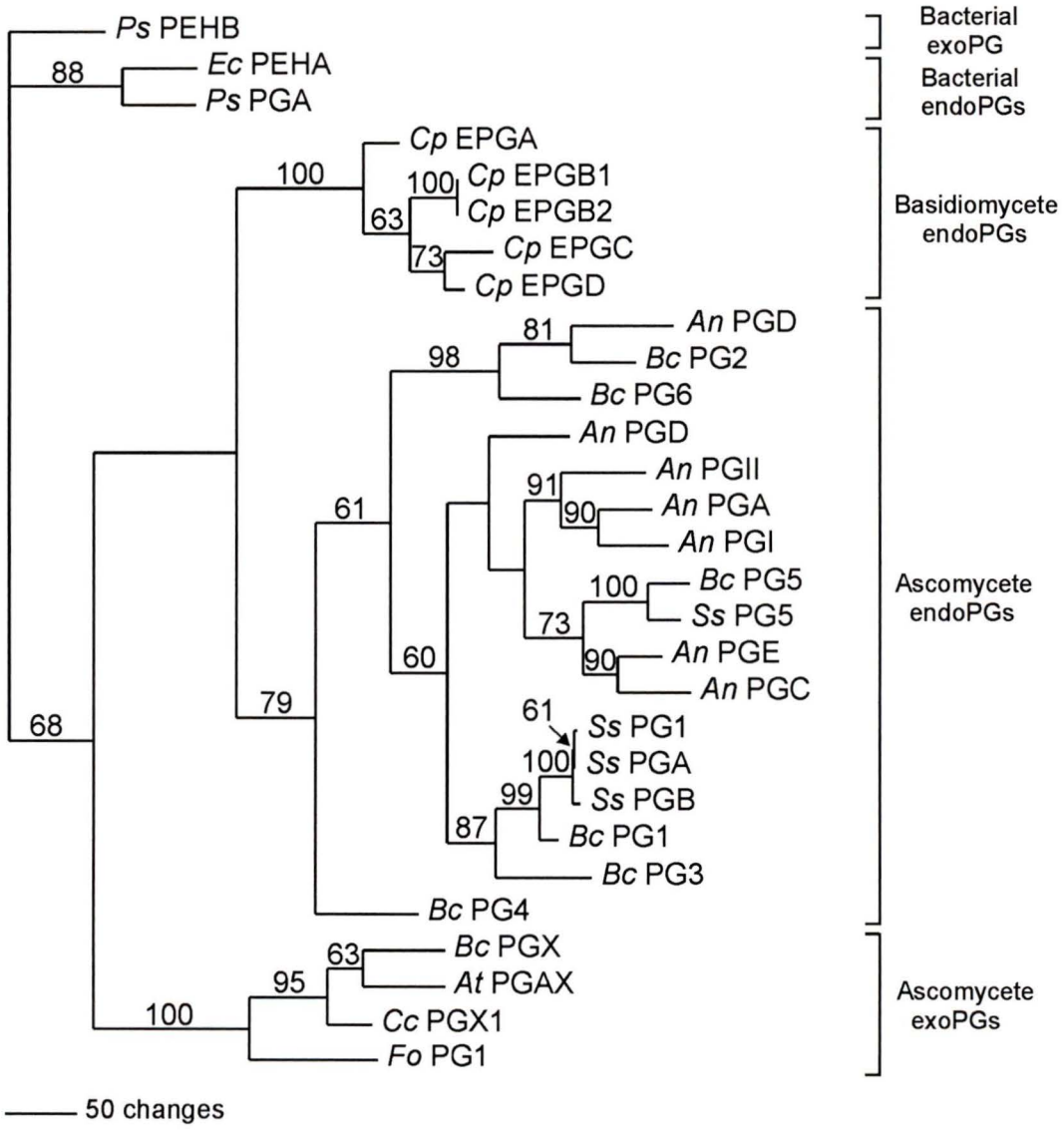
The bacterial PGs were the least similar to all other sequences and formed the root of the tree. With the exception of two nodes within the ascomycete endoPG clade, bootstrap support for the nodes ranged from 61 to 100%, with the majority of groups having at least 80% support (Fig. 7).

The high sequence similarity between *epgB1* and *epgB2* (97.6% nucleotide identity) suggests the two genes represent a recent duplication event. These two genes are much more similar to each other than to their homolog in isolate ASP-4B, *cppg1*, suggesting that *epgB* was duplicated in 2128u after the diverence of 2128u and ASP-4B. It is likely that *epgB1* more closely represents the ancestral gene, and the difference in length between *epgB1* and *epgB2* is due to insertions in the *epgB2* gene rather than deletions in the *epgB1* gene. This was deduced from a comparison of *epgB1*, *epgB2*, and the homologous gene in isolate ASP-4B, *cppg1*. The *epgB1* and *cppg1* genes share open reading frames identical in length and a

nucleotide identity of 91.5 %, with the differences between the genes fairly evenly dispersed throughout the gene sequences. In contrast, *epgB1* and *epgB2* are nearly identical in the first 1527 bp of the genes (99.7% nucleotide identity), but relatively more divergent in the remaining ~200 nucleotides of the genes (82.4% nucleotide identity). It is more plausible that after isolates 2128u and ASP-4B diverged, the existing *epgB1* gene was duplicated and insertions occurred in one of the duplicate copies, than that *epgB2* was duplicated and *epgB1* and *cppg1* independently lost the same segments of DNA.

It is possible that the 3' ends of the *epgB* genes are relatively widely divergent due to chance mutations, but it also may be that features near the 3' end of the *epgB2* gene make it more susceptible to mutation than the rest of the gene. Alternatively, purifying selection may have conserved the majority of the *epgB* gene sequence, but the 3' end may be relatively less important to the function of the enzyme and thus not subject to strong negative selection. However, the first eight of the nine exons are completely conserved between *epgB1* and *epgB2*, and it is unlikely these introns are subject to intense purifying selection. It is more plausible that the gene duplication occurred relatively recently and the 3'end of *epgB2* is relatively less stable than the rest of the gene.

FIGURE 7. Phylogram of the maximum parsimony tree showing the relatedness the five *C. purpureum* endoPGs to 17 ascomycete endoPGs, four ascomycete exoPGs, two bacterial endoPGs, and a bacterial exoPG. Abbreviated species names and publication sources of protein sequences are as follows: *An*, *Aspergillus niger*, PGI (Bussink et al., 1991b; Accession No. X58892), PGII (Bussink et al., 1990; X58893), PGA and PGB (Parenicová *et al.*, 2000a; Y18804 and Y18805), PGC (Bussink et al., 1992; X64356), PGD (Parenicová et al., 2000b; Y18806), PGE (Parenicová *et al.*, 1998; Y14386); *At*, *Aspergillus tubingensis*, PGAX (Kester et al., 1996; X99795); *Bc*, *Botrytis cinerea*, PG1 (ten Have *et al.*, 1998; U68715), PG2, PG3, PG4, PG5, and PG6 (Wubben et al., 1999; U68716, U68717, U68719, U68721, and U68722), PGX (AF145229); *Cc*, *Cochliobolus carbonum* PGX1 (Scott-Craig et al., 1998; L48982); *Cp*, *Chondrostereum purpureum*, EPGA, EPGB1, EPGB2, EPGC, EPGD (AF237653, AF237654, not yet available, AF237655, AF237656); *Ec*, *Erwinia carotovora*, PEHA (Saarilahti et al., 1990; X51701); *Fo*, *Fusarium oxysporum*, PG1 (AF136444); *Ps*, *Pseudomonas solanacearum*, PGA (Huang and Schell, 1990; M33692), PEHB (Huang and Allen, 1997; U60106); *Sc*, *Sclerotinium sclerotiorum*, PG1 (Reymond et al., 1994; L12023), PGA and PGB (Fraissinet-Tachet et al., 1995; L29040 and L29041), PG5 (Y13669). Bootstrap values indicate the percent of times the node appeared in 1000 random samples of the data. The two nodes within the ascomycete endoPG clade lacking bootstrap values were supported in less than 50% of the bootstrap trees.

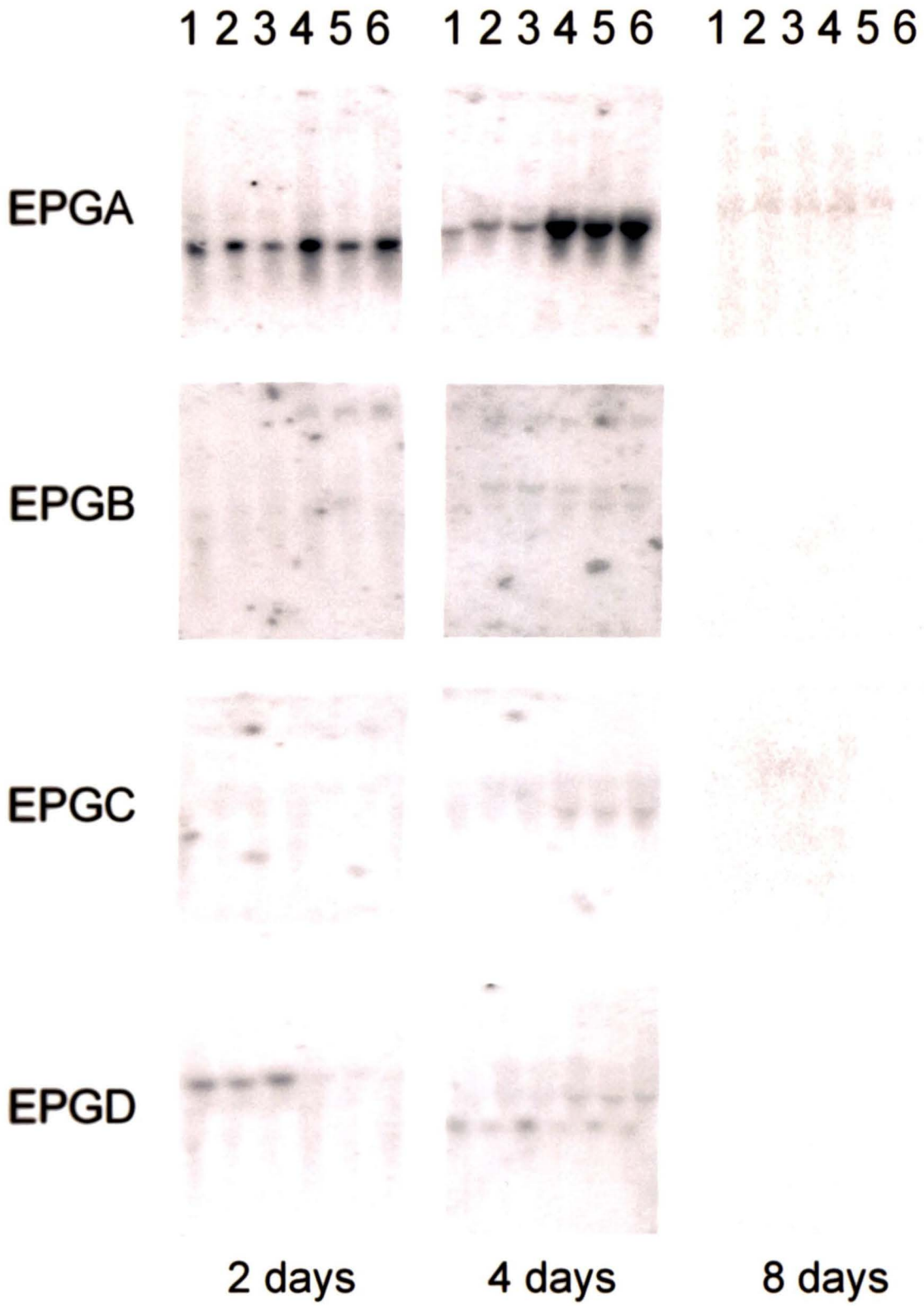

Ps PEHB
Ec PEHA
Ps PGA
Cp EPGA
Cp EPGB1
Cp EPGB2
Cp EPGC
Cp EPGD
An PGD
Bc PG2
Bc PG6
An PGD
An PGII
An PGA
An PGI
Bc PG5
Ss PG5
An PGE
An PGC
Ss PG1
Ss PGA
Ss PGB
Bc PG1
Bc PG3
Bc PG4
Bc PGX
At PGAX
Cc PGX1
Fo PG1
88
100
100
63
73
81
98
91
90
61
100
73
90
60
79
61
100
99
87
68
63
95
100
50 changes
Bacterial exoPG
Bacterial endoPGs
Basidiomycete endoPGs
Ascomycete endoPGs
Ascomycete exoPGs

3.7 Influence of carbon source on endoPG expression

The expression of the endoPG gene family was analyzed in media supplemented with a variety of carbon sources: glucose, sucrose, pectin, a combination of pectin and glucose or pectin and sucrose, and no carbon source. RNA blot analysis revealed that *epgA* is much more highly expressed than the other three endoPG genes, on all six media and at all three time intervals examined (Fig. 8). The transcript level of *epgA* was similar in all six media after 2 days, greater in the three pectin-containing cultures than in the minimal and simple sugar media after 4 days, and reduced in all media types after 8 days. At all three time periods tested, the *epgB* (*1* and *2*), *epgC*, and *epgD* genes were expressed at significantly lower levels than *epgA* or were not detected at all. The reasons for this are not clear, although these genes may be transiently expressed either earlier or later than the time points at which RNA was extracted from these cultures. In *Colletotrichum lindemuthianum,* expression of the *CLPG2* gene of was not detected *in planta* and only briefly detected in culture, whereas the *CLPG1* gene was detected *in planta* and over a longer time span in culture (Centis et al., 1997). Some endoPGs (Gao et al., 1996; Fraissinet-Tachet et al., 1995) have been identified which only appear to be induced in host plant tissues, and this may also be the case for some of the endoPG genes in *C. purpureum*. The differing isoelectric points of the four enzymes may be significant; *epgs B1, B2, C,* and *D* have acidic pIs, while the estimated pI for *epgA* is slightly basic (7.97). EPGA would thus be the only one of the four enzymes with a positive charge at physiological pH values. This charge and the higher number of basic residues in EPGA might affect the binding of the enzyme to the pectic substrate (Keon et al., 1987). Promoter analysis also revealed attributes unique to *epgA*; the promoter of this gene contained three identified elements associated with high level gene expression (Ballance, 1986; Gurr et al., 1987; Lewin, 1997;

Bussink et al., 1992) which were absent in the promoters of *epgB1*, *epgC*, and *epgD*. The promoter of *epgB2* was not fully compared to those of the other endoPGs because only 105 bp immediately 5' to *epgB2* was sequenced. The finding that *epgA* was strongly expressed and not catabolite-repressed by simple sugars suggests it may be one of the relatively uncommon constitutively produced endoPGs. It was suggested that constitutively expressed endoPGs may provide degradation products which induce expression of a battery of other pectin degrading enzymes (Shieh et al., 1997), including other members of a multigene family (Martel et al., 1998; Wubben et al., 1999). Gene disruption of the constitutively produced fungal endoPG in *Aspergillus flavus* (Shieh et al., 1997) and in *Botrytis cinerea* (ten Have et al., 1998) significantly reduced the ability of these pathogens to spread within host tissues. If constitutively produced endoPGs act as the 'leaders' of pectin degradation in multigene families of endoPGs, it would be expected that disruption of constitutively expressed genes might have the greatest effect on the virulence of the pathogen. Thus evidence from this study suggests *epgA* should be the first target of gene disruption studies to determine the role of endoPG in this pathogen; however, expression of the five endoPG genes needs to be studied in more detail, both in culture and in host plants.

FIGURE 8. Expression of the *C. purpureum* endoPG genes 2, 4 and 8 days post-inoculation on minimal (lane 1), glucose (lane 2), sucrose (lane 3), pectin (lane 4), pectin plus glucose (lane 5) and pectin plus sucrose (lane 6) media. The variation in level of hybridizing bands at different time intervals is due to variation in the length of time gels were run. The *epgB* probe would be expected to bind to both *epgB1* and *epgB2* transcripts if they were present.


1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
EPGA
EPGB
EPGC
EPGD
2 days
4 days
8 days

3.10 Summary and Conclusions

Endopolygalacturonase activity has been correlated with the virulence of numerous plant pathogens (reviewed in Alghisi and Favaron, 1997). Recently, molecular genetic techniques were used to confirm a role of endoPG in the virulence of a bacterial pathogen (Huang and Schell, 1990; Huang and Allen, 1997) and two fungal pathogens (Shieh et al., 1997; ten Have et al., 1998). Other endoPG gene disruptions (Scott-Craig et al., 1990; Gao et al., 1996) and introduction of an endoPG into isolates lacking the gene (Di Pietro and Roncero, 1998) did not have any apparent effect on the virulence of the transformed fungi.

Purified PG (both endo- and exo-), pectate lyase, pectin lyase, and pectinesterase from *Chondrostereum purpureum* were injected alone and in combination into apple trees. EndoPG was the only pectinase able to induce silver-leaf symptoms (Miyairi, 1988). Four endoPG enzymes were initially identified in this species based on differences in isoelectric points and size (Miyairi, 1988). Four types of endoPG genes have now been cloned and characterized from isolate 2128u, although it is unknown whether these correspond to the four enzymes described by Miyairi (1988). One of these genes appears to have recently undergone a gene duplication and is present in two nearly identical copies. *Chondrostereum purpureum* is the first basidiomycete fungus for which endoPG genes have been cloned and characterized. Phylogenetic analysis indicated that the five *C. purpureum* endoPGs are more closely related to each other than to any of the ascomycete endoPGs, whereas many ascomycete endoPGs were more closely grouped with endoPGs from other ascomycetes than with other members of their own multigene family. This suggests the basidiomycetes and ascomycetes diverged prior to the duplication of existing endoPG genes, whereas many ascomycete speciation events occurred after the duplication of ascomycete endoPGs. A study comparing the expression of the four *C.*

purpureum endoPGs indicated that EPGA was the major endoPG expressed *in vitro*, and was not repressed by simple sugars. Further expression studies are needed to establish whether *epgA* is constitutively expressed *in planta*. Disruption of a constitutively expressed endoPG gene has significantly reduced the virulence of two fungal pathogens, despite the presence of other endoPG genes (Shieh et al., 1997; ten Have et al., 1998). A similar disruption strategy may allow the role of endoPGs in *C. purpureum* to be elucidated.

FUTURE STUDIES

4.1 Development of a transformation system for *C. purpureum*

Much remains to be learned about the endopolygalacturonase family of *Chondrostereum purpureum*, including its importance in the pathogenicity of the fungus. Definitive proof of the role of endoPG can only be gained by comparison of the virulence levels of strains with disrupted or overexpressed endoPG genes with those of wild-type strains. However, a transformation system for *C. purpureum* remains to be developed.

Attempts to transform *Chondrostereum purpureum* to phleomycin resistance using a vector, pGft, containing a phleomycin resistance gene, *Sh ble* (*Streptoalloteichus hindustanus ble*), regulated by the promoter and terminator of the constitutively expressed glyceraldehyde-3-phosphate dehydrogenase GPD (*GPD*) gene were unsuccessful. Two similar vectors with the *S. commune GPD* promoter or terminator replaced with the like regulatory sequence from *Agaricus bisporus* also failed to transform *C. purpureum*. (All three vectors were kindly provided by Dr. F. Schuren, University of Groningen, The Netherlands.) Although these regulatory sequences were from basidiomycete fungi and all three vectors were used to transform *Schizophyllum commune* (Schuren et al., 1994), it is possible that these sequences were not recognized by the transcriptional machinery of *C. purpureum*. The *S. commune GPD* promoter was replaced with the *C. purpureum epgA* promoter, but this vector was also unsuccessful in transforming the *C. purpureum* genome to phleomycin resistance.

Numerous variations of one basic method were used in most of the *C. purpureum* transformation experiments. This method is the one most commonly used to transform fungi, and involves protoplasting the fungus and then introducing foreign DNA in the presence of

polyethylene glycol (PEG) and calcium ions (Fincham, 1989). The concentrations of PEG, calcium, and other chemicals reported to affect efficiency in other fungal transformation systems were variously adjusted, and transformation methods using linearized plasmids and restriction enzyme-mediated integration (REMI) were tested. Several attempts were made to electroporate the *Sh ble*-containing vectors into either protoplasts or mycelial fragments.

Having tried to transform *C. purpureum* using numerous variations of a few methods, I would suggest any future attempts to transform this species should use an entirely new approach. Many of my transformation experiments were rendered useless due to growth of the negative controls on phyleomycin concentrations well above previously established inhibitory concentrations. The same problem was also encountered by Tang (1996) in her attempts to transform *C. purpureum* to phleomycin resistance. It may be worthwhile to create an auxotrophic mutant of *C. purpureum* which would allow unequivocal determination of transformation back to the autotrophic state.

Recently, a new method has been developed to transform filamentous fungi. *Agrobacterium*-mediated transformation was used to transform *Aspergillus awamori*, *Aspergillus niger*, *Fusarium venenatum*, *Trichoderma reesei*, *Colletotrichum gloeosporioides*, *Neurospora crassa*, and the basidiomycete *Agaricus bisporus* (De Groot et al., 1998). This method has been used to transform a wide range of plant species, by exploiting the natural ability of *Agrobacterium* to insert a portion of its plasmid DNA into plant chromosomes. It is interesting that *A. tumefaciens* can also transfer DNA into fungi. Indeed, *Agrobacterium*-mediated transformation increased the transformation efficiency 600-fold as compared with conventional techniques for transformation of *A. awamori* protoplasts (De Groot et al., 1998). Gouka et al. (1999) showed that when Ti DNA shares homology with the *A. awamori* genome,

integration can occur by homologous recombination. Thus *Agrobacterium*-mediated transformation may be a possibility for gene-disruption of the *C. purpureum* endoPG genes.

REFERENCES

Albersheim, P. 1976. *The primary cell wall.* In *Plant Biochemistry* (J. Bonner and J.E. Varner, Eds.), pp. 225-273. Academic Press, New York.

Alghisi, P. and Favaron, F. 1995. Pectin-degrading enzymes and plant-parasite interactions. *European Journal of Plant Pathology* **101**:365-375.

Annis, S.L. and Goodwin, P.H. 1997. Recent advances in the molecular genetics of plant cell wall-degrading enzymes produced by plant pathogenic fungi. *European Journal of Plant Pathology* **103**:1-14.

Ballance, D.J. 1986. Sequences important for gene expression in filamentous fungi. *Yeast* **2**:224-236.

Bateman, D.R. and Basham, J.G. 1976. Degradation of cell walls and membranes by microbial enzymes. In *Encyclopedia of Plant Physiology Vol. 4* (R. Heitefuss and P.H. Williams, Eds.), pp. 316-355. Springer-Verlab, New York.

Beaulieu, C. and van Gijsegem, F. 1992. Pathogenic behavior of several mini-mu-induced mutants of *Erwinia chrysanthemi* on different plants. *Molecular Plant-Microbe Interactions* **5**:340-346.

Becker, E.M., Ball, L.A., and Hintz, W.E. 1999a. PCR-based genetic markers for detection and infection frequency analysis of the biocontrol fungus *Chondrostereum purpureum* on Sitka alder and trembling aspen. *Biological Control* **15**:71-80.

Becker, E.M., Ball, L.A., Dumas, M.T., Pitt, D.G., Wall, R.E., and Hintz, W.E. 1999b. *Chondrostereum purpureum* as a biological control agent in forest vegetation management. III. Infection survey of a national field trial. *Canadian Journal of Forest Research* **29**: 859-865.

Benen, J.A.E, Kester, H.C.M., and Visser, J. 1999. Kinetic characterization of *Aspergillus niger* N400 endopolygalacturonases I, II and C. *European Journal of Biochemistry* **259**: 577-585.

Benhamou, N., Lafitte, C., Barthe, J.P., and Esquerre, T.M.T. 1991. Cell surface interactions between bean leaf cells and *Colletotrichum lindemuthianum*: Cytochemical aspects of pectin breakdown and fungal endopolygalacturonase accumulation. *Plant Physiology* **97**:234-244.

Bergmann, C.W., Ito, Y., Singer, D., Albersheim, P., Darvill, A.G., Benhamou, N., Nuss, L., Salvi, G., Cervone, F., and De-Lorenzo, G. 1994. Polygalacturonase-inhibiting protein accumulates in *Phaseolus vulgaris* L. in response to wounding, elicitors and fungal infection. *Plant Journal* **5**:625-634.

Brooks, R.T. and Moore W.C. 1926. Silver-leaf disease. V. *Journal of Pomology and Horticultural Science* **5**:11-97.

Brown, R.L., Cleveland, T.E., Cotty, P.J., and Mellon, J.E. 1992. Spread of *Aspergillus flavus* in cotton bolls, decay of intercarpellary membranes, and production of fungal pectinases. *Phytopathology* **82**:462-467.

Bussink, H.J.D., Kester, H.C.M., and Visser, J. 1990. Molecular cloning, nucleotide sequence and expression of the gene encoding prepro-polygalacturonase II of *Aspergillus niger*. *FEBS Letters* **273**:127-130.

Bussink, H.J.D., Buxton, F.P., and Visser, J. 1991a. Expression and sequence comparison of the *Aspergillus niger* and *Aspergillus tubingensis* genes encoding polygalacturonase II. *Current Genetics* **19**:467-474.

Bussink, H.J.D., Brouwer, K.B., de Graaff, L.H., Kester, H.C.M., and Visser, J. 1991b. Identification and characterization of a second polygalacturonase gene of *Aspergillus niger*. *Current Genetics* **20**:301-307.

Bussink, H. J. D., Buxton, F. P., Fraaye, B. A., de Graaff, L. H., and Visser, J. 1992. The polygalacturonases in *Aspergillus niger* are encoded by a family of diverged genes. *European Journal of Biochemistry* **208**:83-90.

Byrde, R.J.W. 1979. Role of polysaccharides-degrading enzymes in microbial pathogenicity. In *Microbial Polysaccharides and Polysaccharases*. (R.C.W. Berkeley, G.W. Gooday, and D.C. Ellwood, Eds.), pp. 417-336. Academic Press, London.

Calmels, T.P.G., Martin, F., Durand, H., and Tiraby, G. 1991. Proteolytic events in the processing of secreted proteins in fungi. *Journal of Biotechnology* **17**:51-66.

Caprari, C., Richter, A., Bergmann, C., Cicero, S. L., Salvi, G., Cervone, F., and De Lorenzo, G. 1993. Cloning and characterization of a gene encoding the endopolygalacturonase of *Fusarium moniliforme*. *Mycological Research* **97**:497-505.

Carpita, N.C. and Gibeaut, D.M. 1993. Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *Plant Journal* **3**:1-30.

Centis, S., Dumas, B., Fournier, J., Marolda, M., and Esquerré-Tugayé, M-T. 1996. Isolation and sequence analysis of *Clpg1*, a gene coding for an endopolygalacturonase of the phytopathogenic fungus *Colletotrichum lindemuthianum*. *Gene* **170**:125-129.

Centis, S., Guillas, I., Séjalon, N., Esquerré-Tugayé, M-T., and Dumas, B. 1997. Endopolygalacturonase genes from *Colletotrichum lindemuthianum*: Cloning of *Clpg2* and comparison of its expression to that of *Clpg1* during saprophytic and parasitic growth of the fungus. *Molecular Plant-Microbe Interactions* **10**:769-775.

Cervone, F., De Lorenzo, G., Degrà, L., and Salvi, G. 1987. Elicitation of necrosis in *Vigna unguiculata* Walp. by homogeneous *Aspergillus niger* endo-polygalacturonase and by α-D-galacturonate oligomers. *Plant Physiology* **85**:626-630.

Cervone, F., Hahn, M.G., De Lorenzo, G.D., Darvill, A.G., and Albersheim, P. 1989. Host-pathogen interactions: XXXIII. A plant protein converts a fungal pathogenesis factor into an elicitor of plant defense responses. *Plant Physiology* **90**:542-548.

Cervone, F., De Lorenzo, G., Pressey, R., Darvill, A.G., and Albersheim, P. 1990. Can *Phaseolus* PGIP inhibit pectic enzymes from microbes and plants? *Phytochemistry* **29**:447-449.

Cervone, F., Hahn, M.G., De Lorenzo Caprari, C., Clark, A.J., Desiderio, A., Devoto, A., Leckie, F., Nuss, L., Salvi, G., and Toubart, P. 1993. The interaction between fungal endopolygalacturonase and plant cell wall PGIP (polygalacturonase-inhibiting protein). In *Mechanisms of Plant Defense Responses*. (B. Fritig and M. Legrand, Eds), pp. 64-67. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Charudattan, R. 1988. Inundative control of weeds with indigenous plant pathogens. In *Fungi in biological control systems* (M.N. Burge, Ed.), pp. 86-119. Manchester Univ. Press, Manchester and New York.

Chan, Y-H. and Sackston, W.E. 1970. Polygalacturonase production by virulent and avirulent isolates of *Sclerotium bataticola* in culture and in sunflowers. *Canadian Journal of Botany* **48**:1449-1463.

Cleveland, T.E., and Cotty, P.J. 1991. Invasiveness of *Aspergillus flavus* isolates in wounded cotton bolls is associated with production of a specific fungal polygalacturonase. *Phytopathology* **81**:155-158.

Collmer, A., and Keen, N.T. 1986. The role of pectic enzymes in plant pathogenesis. *Annual Review of Phytopathology* **24**:383-409.

Cooke, R.D., Ferber, C.E.M., and Kanagasabapathy, L. 1976. Purification and characterization of polygalacturonase from a commercial *Aspergillus niger* preparation. *Biochimica et Biophysica Acta* **452**:440-51.

Cooper, R.M. 1977. Regulation of synthesis of cell-degrading enzymes of plant pathogens. In *Cell wall biochemistry related to spccificity in host-plant pathogen interactions* (B. Solheim and J. Raa, Eds.), pp. 163-211. Universitetsforlaget, Oslo.

Cooper, R.M. 1983. The mechanisms and significance of enzymic degradation of host cell walls by parasites. *In* Biochemical Plant Pathology (J.A. Callow, Ed.), pp. 101-135. John Wiley and Sons, New York.

Cooper, R.M. 1984. The role of cell wall-degrading enzymes in infection and damage. In *Plant disease: infection, damage and loss* (R.K.S. Wood and G.J. Jellis, Eds.), pp. 13-27. Blackwell Scientific Publications, Oxford.

Cooper, R.M., Rankin, B., and Wood, R.K.S. 1978. Cell-wall degrading enzymes of vascular wilt fungi. II. Properties and modes of action of polysaccharides of *Verticillium alboatrum* and *Fusarium oxysporum f.sp. lycopersici*. *Physiological Plant Pathology* **13**:101-134.

Darvill, A., McNeil, M., Albersheim P., Delmer D. P. 1980. The primary cell walls of flowering plants. In *The Biochemistry of Plants, Vol. 1* (P.K. Stumpf and E.E. Conn, Eds.), pp. 91-162. Academic Press, New York.

Davis, K.R., Lyon, G.D., Darvill, A.G., and Albersheim, P. 1984. Host-pathogen interactions. XXV. Endopolygalacturonic acid lyase from *Erwinia carotovora* elicits phytoalexin accumulation by releasing plant cell wall fragments. *Plant Physiology* **74**:52-60.

Davis, K.R., Darvill, A.G., Albersheim, P., and Dell, A. 1986a. Host-pathogen interactions. XXIX. Oligogalacturonides released from sodium polypectate by endopolygalacturonic acid lyase are eleictiors of phytoalexins in soybean. *Plant Physiology* **80**:568-577.

Davis, K.R., Darvill, A.G., and Albersheim, P. 1986b. Host-pathogen interactions. XXXI. Several biotic and abiotic elicitors act synergistically in the induction of phytoalexin accumulation in soybean. *Plant Molecular Biology* **6**:23-32.

de Groot, M.J., Bundock, P., Hooykaas, P.J., and Beijersbergen, A.G. 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nature Biotechnology* **16**:839-42.

de Jong, M.D., Scheepens, P.C., Zadoks, J.C. 1990. Risk analysis for biological control: A Dutch case study in biocontrol of *Prunus serotina* by the fungus *Chondrostereum purpureum*. *Plant Disease* **74**: 189-194.

de Jong, M.D., 2000. The BioChon story: Deployment of *Chondrostereum purpureum* to suppress stump sprouting in hardwoods. *Mycologist* **14**:58-62.

Desiderio, A., Aracri, B., Leckie, F., Mattei, B., Salvi, G., Tigelaar, H., Van Roekel, J.S.C., Baulcombe, D.C., Melchers, L.S., De Lorenzo, G., and Cervone, F. 1997. Polygalacturonase-inhibiting proteins (PGIPs) with different specificities are expressed in *Phaseolus vulgaris*. *Molecular Plant-Microbe Interactions* **10**:852-860.

Di Pietro, A., and Roncero, M.I.G. 1998. Cloning, expression, and role in pathogenicity of *pg1* encoding the major extracellular endopolygalacturonase of the vascular wilt pathogen *Fusarium oxysporum*. *Molecular Plant-Microbe Interactions* **11**:91-98.

Dorworth, C.E. 1990. Use of indigenous microorganisms for forest weed biocontrol – The P.F.C. enhancement process. In *Alternatives to the chemical control of weeds. Proceedings of an international conference, Rotorua, New Zealand, July 25-27, 1989.* (C. Bassett, L.J.Whitehouse, and J.A. Zabkiewicz, Eds.), pp. 116-119. Ministry of Forestry, FRI Bulletin 155.

Ekramoddoullah, A.K.M., Shamoun, S.F. and Wall, R.E. 1993. Comparison of Canadian isolates of *Chondrostereum purpureum* with respect to temperature response, virulence, and protein profiles. *Canadian Journal of Plant Pathology* **15**: 7-13.

English, P.D., Jurale, J.B., and Albersheim, P. 1971. Host-pathogen interactions. II. Parameters affecting polysaccharide-degrading enzyme secretion by *Colletotrichum lindemuthianum* grown in culture. *Plant Physiology* **47**:1-6.

Favaron, F., Peretto, R., Bonfante, P., and Alghisi, P. 1993. Differential absorption and localization of two *Sclerotinia sclerotiorum* endo-polygalacturonases in soybean hypocotyls. *Physiological and Molecular Plant Pathology* **43**:353-364.

Fincham, J.R. 1989. Transformation in fungi. *Microbiological Reviews* **53**:148-70.

Fraissinet-Tachet, L., Reymond-Cotton, P., and Fèvre, M. 1995. Characterization of a multigene family encoding an endopolygalacturonase in *Sclerotinia sclerotiorum*. *Current Genetics* **29**:96-99.

Frittrang, A.K., Deising, H., and Mendgen, K. 1992. Characterization and partial purification of pectinesterase, a differentiation-specific enzyme of *Uromyces viciae fabae*. *Journal of General Microbiology* **138**: 2213-2218.

Gao, S. and Shain, L. 1995. Activity of polygalacturonase produced by *Cryphonectria parasitica* in chestnut bark and its inhibition by extracts from American and Chinese chestnut. *Physiological and Molecular Plant Pathology* **46**:199-213.

Gao, S., Choi, G.H., Shain, L., and Nuss, D.L. 1996. Cloning and targeted disruption of *enpg-1*, encoding the major *in vitro* extracellular endopolygalacturonase of the chestnut blight fungus, *Cryphonectria parasitica*. *Applied Environmental Microbiology* **62**:1984-1990.

Gosselin L., Jobidon R., and Bernier, L. 1995. Assessment of genetic variation with *Chondrostereum purpureum* from Quebec by random amplified polymorphic DNA analysis. *Mycological Research* **100**:151-158

Gosselin L., Jobidon R., and Bernier, L. 1999a. Genetic variability and structure of Canadian populations of *Chondrostereum purpureum,* a potential biophytocide. *Molecular Ecology* **8**: 113-122.

Gosselin L., Jobidon R., and Bernier, L. 1999b. Biological control of stump sprouting of broadleaf species in rights-of-way with *Chondrostereum purpureum*: Incidence of the disease on nontarget hosts. *Biological Control* **16**: 60-67.

Gouka, R.J., Gerk, C., Hooykaas, P.J., Bundock, P., Musters, W., Verrips, C.T., and de Groot, M.J. 1999. Transformation of *Aspergillus awamori* by *Agrobacterium tumefaciens*-mediated homologous recombination. *Nature Biotechnology* **17**:598-601.

Guarente, L., Lalonde, B., Gifford, P. and Alani, E. 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* **36**:503-11.

Gurr, S.J., Unkles, S.E., and Kinghorn, J.R. 1987. The structure and organization of nuclear genes of filamentous fungi. In *Gene Structure in Eukaryotic Microbes* (J. R. Kinghorn, Ed.), pp. 93-139. IRL Press, Oxford WA.

Hahn, M.G., Darvill, A.G., and Albersheim, P. 1981. Host-pathogen interactions VI. A single plant protein efficiently inhibits endopolygalacturonases secreted by *Colletotrichum lindemuthianum* and *Aspergillus niger*. *Plant Physiology* **51**:489-491.

Harper, G.J., Comeau, P.G., Hintz, W., Wall R.E., Prasad, R., and Becker, E.M. 1999. *Chondrostereum purpureum* as a biological control agent in forest vegetation management. II. Efficacy on Sitka alder and aspen in western Canada. *Canadian Journal of Forest Research* **29**:842-858.

Heiler, S., Mendgen, K., and Deising, H. 1993. Cellulolytic enzymes of the obligately biotrophic rust fungus *Uromyces viciae fabae* are regulated differentiation-specifically. *Mycological Research* **97**:77-85.

Hong, S.B., and Tucker, M.L. 1998. Genomic organization of six tomato polygalacturonases and 5' upstream sequence identity with *tap*1 and *win*2 genes. *Molecular and General Genetics* **258**:479-487.

Huang, Q., and Allen, C. 1997. An exo-poly-α-D-galacturonosidase, PehB, is required for wild-type virulence of *Ralstonia solanacearum*. *Journal of Bacteriology* **179**:7369-7378.

Huang, J.and Schell, M.A. 1990. DNA sequence analysis of *pglA* and mechanism of export of its polygalacturonase product from *Pseudomonas solanacearum*. *Journal of Bacteriology* **172**:3879-3887.

Jarvis, W.R. 1977. *Botryotinia* and *Botrytis* species: taxonomy, physiology, and pathogencity: a guide to the literature. Canada Department of Agriculture, Ottawa, Ontario, Canada.

Jin, D.F. and West, C.A. 1984. Characteristics of galacturonic acid oligomers as elicitors of casbene synthetase activity in castor bean seedlings. *Plant Physiology* **74**:989-992.

Jones, T.M., Anderson, A.J., and Albersheim, P. 1972. Host-pathogen interactions IV. Studies on the polysaccharide-degrading enzymes secreted by *Fusarium oxysporum* f. sp. *lycopersici*. *Physiological Plant Pathology* **2**:153-166.

Kamoen, O., and van der Cruyssen, G. 1996. Two secretions from *Botrytis cinerea* important for pathogenesis. *Mededelingen van de Faculteit Landbouwwetenschappen Rijksuniversiteit Gent* **61**:251-260.

Keon J.P.R., Byrde R.J.W., and Cooper R.M. 1987. Some aspects of fungal enzymes that degrade plant cell walls. In *Fungal Infection of Plants* (G. F. Pegg and P. G. Ayres, Eds.), pp. 133-157. University Press, Cambridge.

Kester, H.C., Kusters-van Someren, M.A., Muller,Y. and Visser, J. 1996. Primary structure and characterization of an exopolygalacturonase from *Aspergillus tubingensis*. *European Journal of Biochemistry* **240**:738-746.

Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research* **12**:857-872.

Leone, G., and van den Heuvel, J. 1987. Regulation by carbohydrates of the sequential *in vitro* production of pectic enzymes by *Botrytis cinerea*. *Canadian Journal of Botany* **65**: 2133-2141.

Lewin, B. 1997. *Genes VI*. Oxford University Press, Oxford, NY.

Marciano, P., Di Lenna, P., and Magro, P. 1982. Polygalacturonase isoenzymes produced by *Sclerotinia sclerotiorum in vivo* and *in vitro*. *Physiological Plant Pathology* **20**:201-212.

Markin, G.P. and Gardner, D.E. 1993. Status of biological control in vegetation management in forestry. *Canadian Journal of Forest Research* **23**:2023-2031.

Martel, M.B., Letoublon, R., and Fevre, M. 1996. Purification of endopolygalacturonases from *Sclerotinia sclerotiorum*: Multiplicity of the complex enzyme system. *Current Microbiology* **33**:243-248.

Martel, M.B., Letoublon, R., and Fevre, M. 1998. Purification and characterization of two endopolygalacturonases secreted during the early stages of the saprophytic growth of *Sclerotinia sclerotiorum*. *FEMS Microbiology Letters* **158**:133-8.

McNeil, M., Darvill, A.G., Fry, S.C., and Albersheim, P. 1984. Structure and function of the primary cell walls of plants. *Annual Review of Biochemistry* **53**:625-663.

Miyairi, K. 1988. Biochemical studies on the silver-leaf inducing substance of apple silver-leaf disease. *Bulletin of the Faculty of Agriculture of the Hirosaki University* **49**:145-148.

Miyairi, K., Fujita, K., Okuno T, and Sawai, K. 1977. A toxic protein causative of silver-leaf disease symptoms on apple trees. *Agricultural and Biological Chemistry* **41**:1897-1902.

Miyairi, K., Senda, M., Watanabe, M., Hasui,Y., and Okuno, T. 1997. Cloning and sequence analysis of cDNA encoding Endopolygalacturonase I from *Stereum purpureum*. *Bioscience Biotechnology and Biochemistry* **61**:655-659.

Mullen, J.M. and Bateman, D.F. 1976. Polysaccharide degrading enzymes produced by *Fusarium roseum* "Avanaceum" in culture and during pathogenesis. *Physiological Plant Pathology* **6**: 233-246.

Nothnagel, E.A., McNeil, M., Albersheim, P. and Dell, A. 1983. Host-pathogen interactions XXII. A galacturonic acid oligosaccharide from plant cell walls elicits phytoalexins. *Plant Physiology* **71**:916-926.

Parenicová, L., Benen, J.A.E., Kester, H.C.M., and Visser, J. 2000a. *pga*A and *pga*B encode two constitutively expressed endopolygalacturonases of *Aspergillus niger*. *Biochemistry Journal* **345**:637-644.

Parenicová, L., Kester, H.C., Benen, J.A. and Visser, J. 2000b. Characterization of a novel endopolygalacturonase from *Aspergillus niger* with unique kinetic properties. *FEBS Letters* **467**:333-336.

Parenicová, L., Benen, J.A.E., Kester, H.C.M., and Visser, J. 1998. *pga*E encodes a fourth member of the endopolygalacturonase gene family from *Aspergillus niger*. *European Journal of Biochemistry* **251**:72-80.

Peretto, R., Bettini, V., and Bonfante P. 1993. Evidence of two polygalacturonases produced by a mycorrhizal ericoid fungus during its saprophytic growth. *FEMS Microbiology Letters* **114**:85-92.

Ramsfield, T.D., Becker, E.M., Rathlef, S.M., Tang, Y., Vrain, T.C., Shamoun, S.F., and Hintz, W. E. 1996. Geographic variation of *Chondrostereum purpureum* detected by polymorphisms in the ribosomal DNA. *Canadian Journal of Botany* **74**:1919-1929.

Ramsfield, T.D., Shamoun, S.F., Punja, Z.K., and Hintz, W.E. 1999. Variation in the mitochondrial DNA of the potential biological control agent *Chondrostereum purpureum*. *Canadian Journal of Botany* **77**: 1490-1498.

Rexová-Benková, L. and Markovic, O. 1976. Pectic enzymes. In *Advances in Carbohydrate Chem and Biochem 33*. (R.S. Tipson and D. Horton, Eds.), pp. 323-385. Academic Press, New York.

Rexová-Benková L., and Mracková, M. 1978. Active groups of extracellular endo-D-galacturonase of *Aspergillus niger* derived from pH effect on kinetic data. *Biochimica et Biophysica Acta* **523**:162-169.

Reymond, P., Deléage, G., Rascle, C., and Fèvre, M. 1994. Cloning and sequence analysis of a polygalacturonase-encoding gene from the phytopathogenic fungus *Sclerotinia sclerotiorum. Gene* **146**:233-237.

Rodriquez-Palenzuela, P., Burr T.J., and Collmer, A. 1991. Polygalacturonase is a virulence factor in *Agrobacterium tumefaciens* biovar 3. *Journal of Bacteriology* **173**:6547-6552.

Saarilahti, H.T., Heino, P., Pakkanen, R., Kalkkinen, N., Palva, I., and Palva, E.T. 1990. Structural analysis of the *pehA* gene and characterization of its protein product, endopolygalacturonase, of *Erwinia carotovora* subspecies *carotovora*. *Molecular Microbiology* **4**:1037-1044.

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

Schafer, W. 1994. Molecular mechanism of fungal pathogenicity to plants. *Annual Review of Phytopathology* **32**:461-477.

Scheepens, P.C. 1980. Bestrijding van de Amerikaanse vogelkers met pathogen schimmels; een perspectief. *Cent. Agrobiol. Onderz.* Wageningen. Verslag Nr. 29.

Schell, M.A., Roberts, D.P., and Denny, T.P. 1988. Analysis of the *Pseudomonas solanacearum* polygalacturonase encoded by *pglA* and its involvement in phytopathogenicity. *Journal of Bacteriology* **170**:4501-8.

Schuren, F.H. and Wessels, J.G. 1994. Highly-efficient transformation of the homobasidiomycete *Schizophyllum commune* to phleomycin resistance. *Current Genetics* **2**:179-83.

Scott-Craig, J. S., Panaccione, D. G., Cervone, F. and Walton, J. D. 1990. Endopolygalacturonase is not required for pathogenicity of *Cochliobolus carbonum* on maize. *Plant Cell* **2**:1191-1200.

Scott-Craig, J. S., Cheng, Y-Q., Cervone, F., De Lorenzo, G., Pitkin, J. W., and Walton, J. D. 1998. Targeted mutants of *Cochliobolus carbonum* lacking the two major extracellular polygalacturonases. *Applied Environmental Microbiology* **64**:1497-1503.

Shamoun, S.F., Vrain, T.C. and Wall, R.E. 1991. Genetic variability among isolates of *Chondrostereum purpureum* (Abstr.). *Canadian Journal of Plant Pathology* **13**:285.

Shamoun, S.F. and Wall, R.E. 1996. Characterization of Canadian isolates of *Chondrostereum purpureum* by protein content, API ZYM and isozyme analyses. *European Journal of Forest Pathology* **26**:333-342.

Shieh, M-T., Brown, R.L., Whitehead, M.P., Cary, J.W., Cotty, P.J., Cleveland, T.E., and Dean, R.A. 1997. Molecular genetic evidence for the involvement of a specific polygalacturonase, P2c, in the invasion and spread of *Aspergillus flavus* in cotton bolls. *Applied Environmental Microbiology* **63**:3548-3552.

Spiers, A.G., Brewster, D.T., Slade, A., and Gardiner, S.E. 2000. Characterization of New Zealand isolates of *Chondrostereum purpureum* with regard to morphology, growth, pathogenicity and RAPD banding patterns. *Mycological Research* **104**: 395-402.

Suykerbuyk, M.E.G., Schaap, P.J., Stam, H., Musters, W. and Visser, J. 1995. Cloning, sequence and expression of the gene coding for rhamnogalacturonase of *Aspergillus aculeatus*; a novel pectinolytic enzyme. *Applied Microbiology and Biotechnology* **43**:861-870.

Swofford, DL. 2000. *PAUP *. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.* Sinauer Associates, Sunderland, Massachusetts.

Talmadge, K.W., Keegstra, K., Bauer, W.D., Albersheim, P. 1973. The structure of plant cell walls. I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiology* **51**:158-173.

Tang, Y. 1996. Cloning and molecular characterization of a polygalacturonase gene of the white-rot fungus *Chondrostereum purpureum*. M.Sc. thesis, University of Victoria, Canada.

Templeton, G.E., TeBeest, D.O., and Smith, R.J.Jr. 1979. Biological weed control with mycoherbicides. *Annual Review of Phytopathology* **17**:301-310.

Tenberge, K.B., Homann, V., Oeser, B., and Tudzynski, P. 1996. Structure and expression of two polygalacturonase genes of *Claviceps purpurea* oriented in tandem and cytological evidence of pectinolytic enzyme activity during infection of rye. *Phytopathology* **86**:1084-1097.

ten Have, A., Mulder, W., Visser, J., and van Kan, J.A.L. 1998. The endopolygalacturonase gene *Bcpg1* is required for full virulence of *Botrytis cinerea*. *Molecular Plant-Microbe Interactions* **11**:1009-1016.

Thompson, J.D., Higgins, D.G., and Gibson, T. J. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**:4673-4680.

van der Cruyssen, G., de Meester, E., and Kamoen, O. Expression of polygalacturonases of *Botrytis cinerea in vitro* and *in vivo*. *Mededelingen van de Faculteit Landbouwwetenschappen Rijksuniversiteit Gent* **59**:895-905.

von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14**:4683-4690.

Wall, R.E. 1986. Pathogenicity of *Chondrostereum purpureum* to yellow birch. *Plant Disease* **70**: 158-160.

Wall, R.E. 1990. The fungus *Chondrostereum purpureum* as a silvicide to control stump sprouting in hardwoods. *Northern Journal of Applied Forestry* **7**: 17-19.

Wall, R.E. 1991. Pathological effects of *Chondrostereum purpureum* in inoculated yellow birch and beech. *Canadian Journal of Plant Pathology* **13**: 81-87.

Wall, R.E., Prasad, R., and Shamoun, S.F. 1992. The development and potential role of mycoherbicides in forestry. *The Forestry Chronicle* **68**: 736-741.

Watson, A.K. 1989. Current advances in bioherbicide research, pp. 987-996. Proceedings of the Brighton Crop Protection Conference – Weeds.

Whitehead, M. P., Shieh, M.T., Cleveland, T.E., Cary, J.W., and Dean, R.A. 1995. Isolation and characterization of polygalacturonase genes (*pecA* and *pecB*) from *Aspergillus flavus*. *Applied Environmental Microbiology* **61**:3316-3322.

Wubben, J.P., Mulder, W., ten Have, A., van Kan, J.A.L., and Visser, J. 1999. Cloning and partial characterization of endopolygalacturonase genes from *Botrytis cinerea*. *Applied Environmental Microbiology* **65**:1596-1602.

VITA

Surname: Williams Given Names: Holly Louise

Place of Birth: Victoria, British Columbia, Canada

Educational Institutions Attended:

University of Victoria	1997-2001
Simon Fraser University	1991-1996

Degrees Awarded:

B.Sc. (Co-operative Education)	Simon Fraser University	1996

Honours and Awards:

Amelia Leith Memorial Fellowship	1998
M.C. Melburn Award	1997
NSERC Undergraduate Research Scholarship	1992
Abe Unrau Co-op Scholarship	1991

UNIVERSITY OF VICTORIA PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my thesis to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain by the University of Victoria shall not be allowed without my written permission.

Title of Thesis:

Characterization of the multigene family encoding endopolygalacturonase in the basidiomycete *Chondrostereum purpureum*

Author 

Holly Louise Williams

March 16, 2001