INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Isolation and Regulation of Genes Expressed During Douglas-fir Germination and Post-Germination

by

Timothy John Tranbarger

B.A, University of California, Santa Cruz, 1987
M.Sc., Washington State University, 1992

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry and Microbiology

We accept this dissertation as conforming to the required standard

---

Dr. Santosh Misra, Supervisor (Department of Biochemistry and Microbiology)

Dr. Robert Olafson, Departmental Member (Department of Biochemistry and Microbiology)

Dr. Terry Pearson, Departmental Member (Department of Biochemistry and Microbiology)

Dr. Chris Upton, Departmental Member (Department of Biochemistry and Microbiology)

Dr. William Hintz, Outside Member (Department of Biology)

Dr. Peter Pauls, External Examiner (Department of Crop Science, University of Guelph)

© Timothy John Tranbarger, 1998
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

Supervisor: Dr. Santosh Misra

ABSTRACT

To identify genes expressed during Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco) germination and early seedling development, a cDNA library was constructed with mRNA pooled from 4-6-day-old seedlings. The library was then screened differentially with cDNA probes synthesized using mRNA isolated from mature seeds and 6-day-old seedlings. Partial DNA sequence analysis and predicted amino acid sequences revealed cDNA clones that encoded polypeptides with similarity to several plant proteins including: a chaperonin 60ß (cpn60ß), a low molecular weight heat shock protein (LMW HSP), a luminal binding protein (BiP), a type II chlorophyll a/b-binding protein (CAB), and a cysteine protease (CysP). In northern blots, each cDNA clone detected transcripts that increased during seed germination. A clone detected RNA at similar levels in both mature seeds and in 6-day-old seedlings was isolated and found to share similarity to a NADPH-cytochrome P450 reductase (CPR) (EC 1.6.2.4). The cDNA clones encoding the CysP and the CPR were selected for further sequence and gene expression analysis.

The CysP cDNA consists of a 5' untranslated region (*UTR*) of 153 bp followed by an open reading frame (ORF) of 1362 bp encoding a putative mature CysP flanked by N- and C-terminal propeptides. A 364 bp 3' *UTR* contains multiple putative AU-rich elements that may be involved in the destabilization of transcripts. The CysP from Douglas-fir (pseudotzain) contains the same invariant amino acid residues that are involved in the catalytic reaction and make up the catalytic center of CysP from other plants and animals. Pseudotzain transcripts were most abundant in the megagametophyte (MG) after germination and were not detected in the MG or embryo during embryogenesis. Various osmotic stresses slightly enhanced pseudotzain transcript quantities during early seedling

development, whereas abscisic acid, gibberellic acid and other plant growth regulators and changes in environmental conditions had little or no effect. Pseudotzain transcripts were present in different amounts in the cotyledons, root and seed coat of 10-day-old seedlings, but were most abundant in the MG, suggesting a role for this protease in storage protein mobilization. Phylogenetic analysis of mature pseudotzain groups it with other angiosperm CysP having both N- and C-terminal propeptides, suggesting a conserved function and/or targeting of this subgroup of enzymes.

The CPR cDNA encodes a polypeptide of approximately 79.6 kDa. A cDNA probe detected a single transcript of 3 kb that was expressed differentially in cotyledons, radicle and MG. CPR transcript quantities were low during seed maturation, higher in mature seeds, and remained constant throughout germination and early seedling development before they declined in 14-day-old seedlings. An antiserum against a synthetic CPR-peptide was produced and western blot analysis detected a single 80 kDa polypeptide in the membrane fraction of microsomal extracts from seeds and seedlings. CPR accumulation during germination and early seedling development indicated regulation is at the transcriptional or post-transcriptional level. However, CPR activity (measured by NADPH-cytochrome-c reduction) present in the microsomes increases during stratification, germination and post-germination and decreases in 7-14-day-old seedlings. These results indicate CPR may be post-translationally activated during Douglas-fir stratification and germination.

This study describes the isolation of the first cDNAs that share identity with a CysP, cpn60ß, a LMW HSP, BiP and CPR (EC 1.6.2.4) from a gymnosperm. The developmental expression of these cDNAs suggests that their gene products play critical roles during the process of germination and post-germination and provides the necessary framework for future studies.

Examiners:

Dr. Santosh Misra, Supervisor (Department of Biochemistry and Microbiology)

Dr. Robert Olafson, Departmental Member (Department of Biochemistry and Microbiology)

Dr. Terry Pearson, Departmental Member (Department of Biochemistry and Microbiology)

Dr. Chris Upton, Departmental Member (Department of Biochemistry and Microbiology)

Dr. William Hintz, Outside Member (Department of Biology)

Dr. Peter Pauls, External Examiner (Department of Crop Science, University of Guelph)

TABLE OF CONTENTS

Abstract ii

Table of Contents v

List of Tables ix

List of Figures x

List of Abbreviations xv

Acknowledgements xvi

Dedication xvii

Chapter 1 Seed Germination and Post-Germination: Importance and Overview 1

- Introduction 1
- Seed Germination and Post-Germination Seedling Growth 1
- Gene Expression Patterns During Plant Development 3
- Seed Storage Protein Mobilization 8
- Seed Dormancy 11
- Roles of Abscisic Acid and Gibberellic Acid in Dormancy and Germination 14
- Differences Between Gymnosperm and Angiosperm Seeds 15
- Objectives 18
- Literature Cited 19

Chapter 2 The Molecular Characterization of a Set of cDNAs Differentially Expressed During Douglas-fir Germination and Post-Germination 28

- Abstract 28
- Introduction 29

Materials and Methods 31
Plant material and growth conditions 31
Construction and differential screening of a Douglas-fir cDNA library 31
Northern blotting and hybridization 32
DNA sequence analysis 33
Results 34
Isolation of germination associated cDNAs 34
Northern blot hybridization analysis during germination and early seedling development 34
Identification of germination clones through DNA sequence and similarity search analysis 40
Discussion 48
Conclusions 51
Literature Cited 52

Chapter 3 Structure and Expression of a Developmentally Regulated cDNA Encoding a Cysteine Proteinase (Pseudotzain) from Douglas-fir 58
Abstract 58
Introduction 59
Materials and Methods 61
Plant material and growth conditions 61
Northern blotting and hybridization 62
PM3-3 cysP cDNA clone isolation and DNA sequence analysis 62
Amino acid sequence alignments and phylogenetic analysis 63
Results 64
Cloning and sequence analysis of a cysteine proteinase cDNA from Douglas-fir 64
The cysteine proteinase amino acid sequence: alignment and phylogenetic analysis 69

Northern blot analysis of cysteine proteinase mRNA steady-state amounts 74
Discussion 82
Conclusions 85
Literature Cited 86

Chapter 4 A Douglas-fir NADPH-cytochrome P450 reductase expressed during germination and seedling development: expression pattern and evidence for post-translational regulation 91
Abstract 91
Introduction 93
Materials and Methods 96
Plant material and growth conditions 96
cDNA clone isolation and sequence analysis 96
RNA isolation and analysis 97
CPR-peptide antiserum production 98
Microsomal membrane isolation, protein extraction and analysis 98
NADPH-cytochrome c reductase assays 100
Immunofluorescence localization 100
Results 102
Isolation and characterization of the Douglas-fir CPR cDNAs 102
Developmental and tissue regulated *cpr1* expression 116
A CPR-peptide derived antiserum detects a single CPR in 2-dimensional SDS-PAGE 123
PMCPR localizes with the endoplasmic reticulum and NADPH-cytochrome c activity in sucrose gradients 123
Accumulation of PMCPR protein and activity in the microsomes during germination and early seedling development 129
Immunofluorescence localization of PMCPR 132

Discussion 136
Conclusions 142
Literature Cited 143

Chapter 5 Conclusions and Future Studies 151
Pseudotzain: Regulation, Processing and Role in Storage
Protein Mobilization 152
CPR and the Cytochrome P450 Monooxygenase System 156
Literature Cited 164

Appendix I. Total RNA Isolation Data (Samples, Concentrations and Yields) 168

Appendix II. Microsomal and Total Protein Isolation Data Isolation Data (Samples, Concentrations and Yields) 174

LIST OF TABLES

Table 2.1: Douglas-fir cDNA clones-summary 36

Table 2.2: Primers synthesized for sequencing of cDNA clones 41

Table 3.1: Primers synthesized for sequencing the *PM3-3/cysP* cDNA clone. 65

Table 4.1: Primers synthesized for sequencing the *CPR* cDNA clones 106

Table 4.2: Start codon consensus sequences of the *CPR* cDNA 110

LIST OF FIGURES

Figure 1.1: Gene sets expressed during stages of plant development 4

Figure 1.2: An overview and comparison of monocot, dicot and gymnosperm seed anatomy 6

Figure 2.1: Time course of Douglas-fir seed germination and early seedling development 35

Figure 2.2: Northern blot analysis of transcript levels detected by cDNAs during Douglas-fir germination and post-germination 37

Figure 2.3: Densitometric measurements of the northern blot analysis from Figure 2.2 38

Figure 2.4: Partial nucleotide and predicted amino acid sequences of clone *PM6-3* 42

Figure 2.5: Partial nucleotide and predicted amino acid sequences of clone *PM3* 43

Figure 2.6: Partial nucleotide and predicted amino acid sequences of clones *PM3-3* and *PM4-5* 46

Figure 2.7: Partial nucleotide and predicted amino acid sequences of clones *PM5-1* and *PM14* 47

Figure 3.1: The complete nucleotide sequence of *PM3-3* and its deduced amino acid sequence 67

Figure 3.2: Features of the *PM3-3/cysP* cDNA clone and predicted amino acid sequence 68

Figure 3.3: CysP are encoded as zymogens that can include N- and C-terminal extension peptides 70

Figure 3.4: Phylogenetic tree of representative mature CysP from protozoans, animals and plants 72

Figure 3.5: Alignment of plant CysP sequences that have N-terminal extension peptides 73

Figure 3.6: Northern blot analysis of *cysP* transcript levels during embryogenesis and early seedling development 76

Figure 3.7: Northern blot analysis of *cysP* transcript levels in 10-day-old seedling tissues 77

Figure 3.8: Northern blot analysis of *cysP* transcript levels in response to hormonal and environmental treatments 78

Figure 3.9: Northern blot analysis of *cysP* expression in response to wounding during germination 80

Figure 3.10: Northern blot analysis of *cysP* expression in response to desiccation treatments during germination 81

Figure 4.1: The cytochrome P450 monooxygenase system of eukaryotes 103

Figure 4.2: The nucleotide sequences of the cDNAs *PM5* and *PM14* and the deduced Douglas-fir CPR (PMCPR) amino acid sequence 105

Figure 4.3: Overview of the CPR cDNA clones 107

Figure 4.4: An agarose gel of purified *Eco*RI derived *PM5* fragments and restriction map of unique sites 109

Figure 4.5: An alignment of representative plant, animal and fungal CPR amino acid sequences 112

Figure 4.6: Phylogenetic tree of representative CPR predicted amino acid sequences from mammals, yeast and higher plants 114

Figure 4.7: An alignment of the divergent amino terminal sequences of plant CPRs 115

Figure 4.8: Northern blot analysis of *cpr* expression during embryogenesis and early seedling development 117

Figure 4.9: Expression of *cpr1* in 10 day old seedling tissues 119

Figure 4.10: The expression of *cpr1* is not modulated by light during germination and early seedling development, and is not detected in emerging needles in the spring 120

Figure 4.11: Northern blot analysis of *cpr1* expression in response to wounding 121

Figure 4.12: Northern blot analysis of *cpr1* expression in response to desiccation 122

Figure 4.13: An alignment of the amino acid sequence used to raise the CPR peptide-antiserum with the corresponding sequences from other plants 124

Figure 4.14: The CPR-peptide antiserum detects a single CPR in western blot analysis of microsomal proteins 125

Figure 4.15: PMCPR is a microsomal membrane associated protein 127

Figure 4.16: PMCPR is associated with the fraction containing the endoplasmic reticulum and NADPH-cytochrome c reductase activity 128

Figure 4.17: Western blot analysis of PMCPR protein levels during germination and early seedling development 130

Figure 4.18: A comparison of CPR enzyme activity and PMCPR protein levels during germination and early seedling development 131

Figure 4.19: A comparison of *cpr1* transcript and PMCPR protein amounts in 10-day-old seedling tissues 134

Figure 4.20: Immunolocalization of CPR in the developing radicles of 5-day-old seedlings 135

Figure 5.1: mRNA decay experiments 154

Figure 5.2: Primers synthesized to engineer the *Xba* and *Sst* restriction sites and the structure of the *cpr* gene construct used to transform tobacco plants 157

Figure 5.3: Western blot analysis of putative transgenic CPR tobacco plants 158

Figure 5.4: Northern blot analysis of putative CPR transgenic tobacco plants 159

Figure 5.5: Western blot analysis of the serine phosphorylation status of microsomal proteins from mature seed and 2-day-old seedlings 163

LIST OF ABBREVIATIONS

ABA, abscisic acid
ABI, ABA-insensitive
ARE, AU-rich element(s)
BiP, luminal binding protein
bp, base pair(s)
CAB, chlorophyll a/b-binding protein
cDNA, complementary DNA
cpn60ß, chaperonin 60ß
CPR, NADPH-cytochrome P450 reductase
cpr1, gene encoding PMCPR
cysP, gene encoding CysP
CysP, cysteine protease(s)
ER, endoplasmic reticulum
GA, gibberellic acid
HSP, heat shock protein
kb, kilobase(s) or 1000 bp
kDa, kilodaltons
LEA, late embryogenesis abundant
LMW HSPs, low molecular weight heat shock proteins
MG, megagametophyte
Mk, molecular mass markers
mRNP, messenger ribonucleoprotein
nt, nucleotide(s)
ORF, open reading frame
Pm, *Pseudotsuga menziesii*
RER, rough endoplasmic reticulum
rRNA, ribosomal RNA
uORF, upstream ORF(s)
UTR, untranslated region(s)

ACKNOWLEDGEMENTS

My childhood experiences in the redwood and Douglas-fir forests of the Santa Cruz Mountains and my surfing adventures along the Pacific West Coast instilled in me an appreciation of the mystery, power and beauty of nature. Through these passions, I developed and maintained the inspiration to pursue a scientific study of plant biology. I thank my boss, supervisor and friend Dr. Santosh Misra for her consistent encouragement and support, especially through the difficult periods I experienced in the last few years. I wish to express my gratitude to my committee members: Dr. William Hintz, Dr. Robert Olafson, Dr. Terry Pearson and Dr. Chris Upton, and to my external examiner Dr. Peter Pauls from the University of Guelph. Excellent support is greatly appreciated from Rozanne Poulson and the other Biochemistry and Microbiology office staff for guiding me through the UVic bureaucracy, Scott Scholz and Albert Labossiere for keeping everything running smoothly with their technical assistance and Bob Beecroft for assistance with the CPR antiserum production. Multitudes of thanks go to my lab colleagues past and present: Malinee Chatthai, Ben Forward, Dave Machander, Anna-mary Schmidt, Ivan Stefanov, Kris Wilde and Bill Yu. Final thanks go to Andrea for all her love, patience and support throughout my studies. We're "On the Road" again!

DEDICATION

I dedicate this work to my dad, John Edward Tranbarger, who always encouraged me in everything I pursued.

CHAPTER 1

Seed Germination and Post-Germination: Importance and Overview

Introduction

Seed germination and the early growth and development of the seedling are critical stages in the life cycle of plants. The genetic, physiological and biochemical attributes of a seed determine a seedling's potential to become established, compete and thrive in its habitat. Despite the importance of seed germination and early seedling growth to agriculture and forestry, the molecular mechanisms that underlie the events of germination have yet to be determined (Bewley, 1997). To date, most of the molecular biology of germination has focused on the economically important angiosperm crop species. The emphasis of this chapter is to examine some of the important processes that occur during seed germination and early seedling development (post-germination) and what is known about the molecular events that underlie these developmental changes. The information accumulated from work with angiosperms is compared to what is known about gymnosperms. The similarities and differences between angiosperms and gymnosperms are highlighted especially at the molecular level.

Seed Germination and Post-Germination Seedling Growth

Germination begins with the uptake of water (imbibition) by the mature dry seed and ends when the embryo radicle begins to elongate (Bewley and Black, 1994). When the radicle emerges from the seed, germination is complete and the post-germination stage of development is underway. Radicle elongation is a turgor-driven event that is a result of a yielding of the walls of cells that lie between the root cap and the base of the radicle. However, the osmotic potential of the radicle cells does not change prior to the initiation of

radicle growth (Welbaum and Bradford, 1990). During imbibition, the initial rehydration of the desiccated seed causes the cellular membranes to undergo a transition from a gel phase to a hydrated liquid-crystalline state. As the membranes become hydrated, a transient loss of solutes and low molecular weight metabolites from the cells occurs (Simon and Raja Harun, 1972). Electrolyte leakage stops as membrane integrity is regained and may be partially due to an increase in the membrane-stabilizing phospholipid *N*-acetylphosphatidylethanolamine (Sandoval et al., 1995). Seed rehydration also initiates metabolic activity including respiration, glycolysis, synthesis of proteins from stored mRNAs, repair of DNA and mitochondria, and the synthesis of new mRNAs. An initial phase of rapid polysomal formation occurs from ribosomes stored in the mature seed, however, new ribosomes are also produced and incorporated into polysomes early during imbibition (Dommes and Van der Walle, 1990). These early events of germination lead to cell elongation, cell division and the synthesis of nucleic acids and proteins essential for seedling growth.

During post-germination, the seedling is dependent on the nutrient reserves stored in the storage organs of the seed, in the form of oil, carbohydrates and proteins (Bewley and Black, 1994). The mobilization of storage compounds is one of the most important metabolic activities during post-germination. The composition of the storage reserves varies from species to species. In cereals such as rice, corn, barley, wheat and oats, the main storage product is in the form of carbohydrates (starch). In legumes, storage reserves can be high in proteins (vicilins and legumins), whereas in pine and castor bean, oil (triacylglycerols) represents the most abundant storage material. Hydrolytic enzymes (e.g. lipases, amylases, and proteinases) are required to convert the high molecular weight storage compounds into forms that are easily transported (e.g. sucrose and amino acids) to the developing organs of the seedling, namely the radicle and the emerging shoots.

Gene Expression Patterns During Plant Development

In angiosperms, distinct sets of genes have been characterized that are expressed during embryogenesis, seed maturation, germination and post-germination (Figure 1.1; Goldberg et al., 1989). Examples of the gene sets include: the constitutively expressed actin and tubulin mRNAs; embryo-specific mRNAs of unknown function; seed storage protein mRNAs expressed during embryogenesis; the late embryogenesis abundant (LEA) mRNAs thought to encode proteins with a role in seed desiccation and dormancy; mRNAs that overlap late embryogenesis and early germination and are stored in dormant mature seeds; the post-germination specific mRNAs for the enzymes such as isocitrate lyase and malate synthase of the glyoxylate cycle responsible for the mobilization of stored lipids, or the hydrolases involved in seed storage reserve mobilization (Weir et al., 1980; Harada et al., 1988; Comai et al., 1989; Dure et al., 1989; Goldberg et al., 1989; Hughes and Galau, 1989; Galua et al, 1991; Lane, 1991; Thomas, 1993). The identities of many of the developmental genes and the mechanisms of their regulation have yet to be determined.

During the rehydration of the mature dry seed, there are changes in the pattern of gene expression thought to underlie the shift in seed metabolism (Bewley and Marcus, 1990). Specific changes in mRNA populations and newly synthesized protein profiles have been observed during germination of monocots, dicots and gymnosperm seeds (Lalonde and Bewley, 1986; Sánchez-Martínez et al., 1986; Mullen et al., 1996). In *Brassica napus*, imbibition initiates transcriptional alterations that lead to a shift from an embryonic to a germination developmental program (Comai and Harada, 1990). A wheat protein referred to as "germin" accumulates in wheat embryo radicles just before elongation (Lane, 1991). Germin has sequence similarity to oxalate oxidases, enzymes with activity leading to the degradation of calcium oxalate and the production of calcium ions (Ca^{2+}) and hydrogen peroxide (H_2O_2). Both of these compounds function as secondary messengers at low

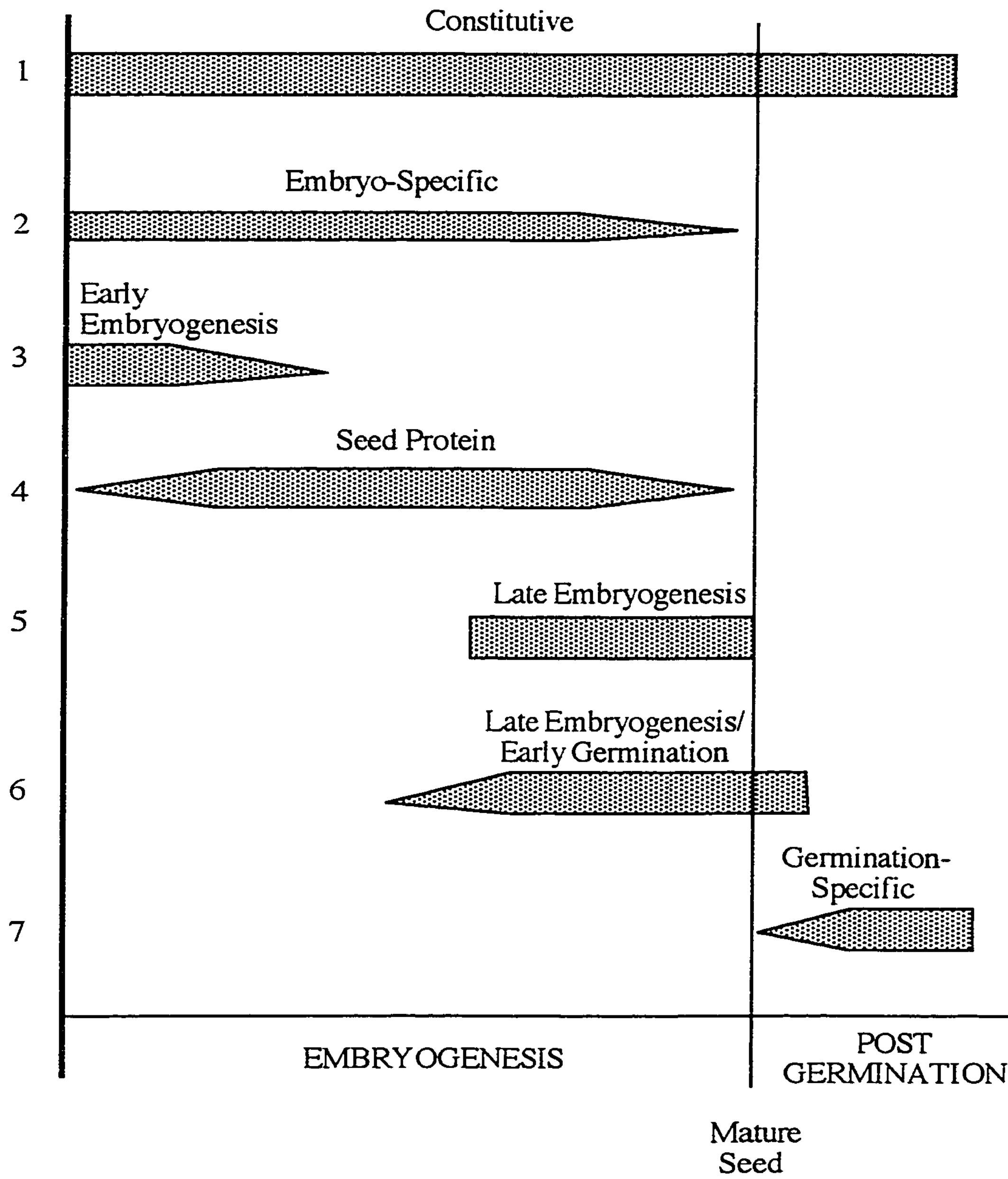


Figure 1.1. Gene sets expressed during stages of plant development. The figure was adapted from Goldberg et al., 1989.

concentrations and mediate cell-wall polymer cross-linking at higher concentrations (Cassab and Varner, 1988; Apostol et al., 1989; Lane et al., 1993; Luttrell, 1993; Showalter, 1993). Accumulation of germin transcripts, protein and activity in germinating wheat seeds coincide spatially and temporally in the outer tissues of the embryo, consistent with a role for germin in cell-wall restructuring (Caliskan and Cuming, 1998). However, the expression and accumulation of germin is not limited to germination, and its possible physiological roles (i.e. signal transduction or cell-wall polymer cross-linking via the production of Ca^{2+} and/or H_2O_2) appear to overlap germination and post-germination. Recently, a partial nucleotide sequence (GENBANK accession number AF049065) encoding a germin-like protein was isolated from *Pinus radiata* (Monterey pine) somatic embryos. It will be of interest to examine the expression of the gymnosperm germin-like gene in comparison to the germin genes of angiosperms. The components that are involved in signal transduction and the transcriptional activation of germination associated genes have not been identified, although work with developmental mutants is beginning to give some insight into the factors involved (see section on the Roles of Abscisic Acid and Gibberellic Acid below).

Distinct sets of genes expressed during germination and post-germination have recently been characterised in gymnosperms (Stabel et al., 1990; Gifford et al., 1991; Groome et al., 1991; Leal and Misra, 1993; Schneider and Gifford, 1994; Mullen et al., 1996). Several cDNAs for the glyoxylate cycle enzyme isocitrate lyase from *Pinus taeda* (loblolly pine) were recently characterized (Mullen and Gifford, 1997). The isocitrate lyase transcripts were detected in mature seeds, and accumulated to higher amounts during germination. There was a corresponding increase in the amount of isocitrate lyase protein and enzymatic activity. During the senescence of the megagametophyte (MG; the location of the main seed storage reserves of gymnosperms; Figure 1.2), steady-state amounts of

Seed Anatomy

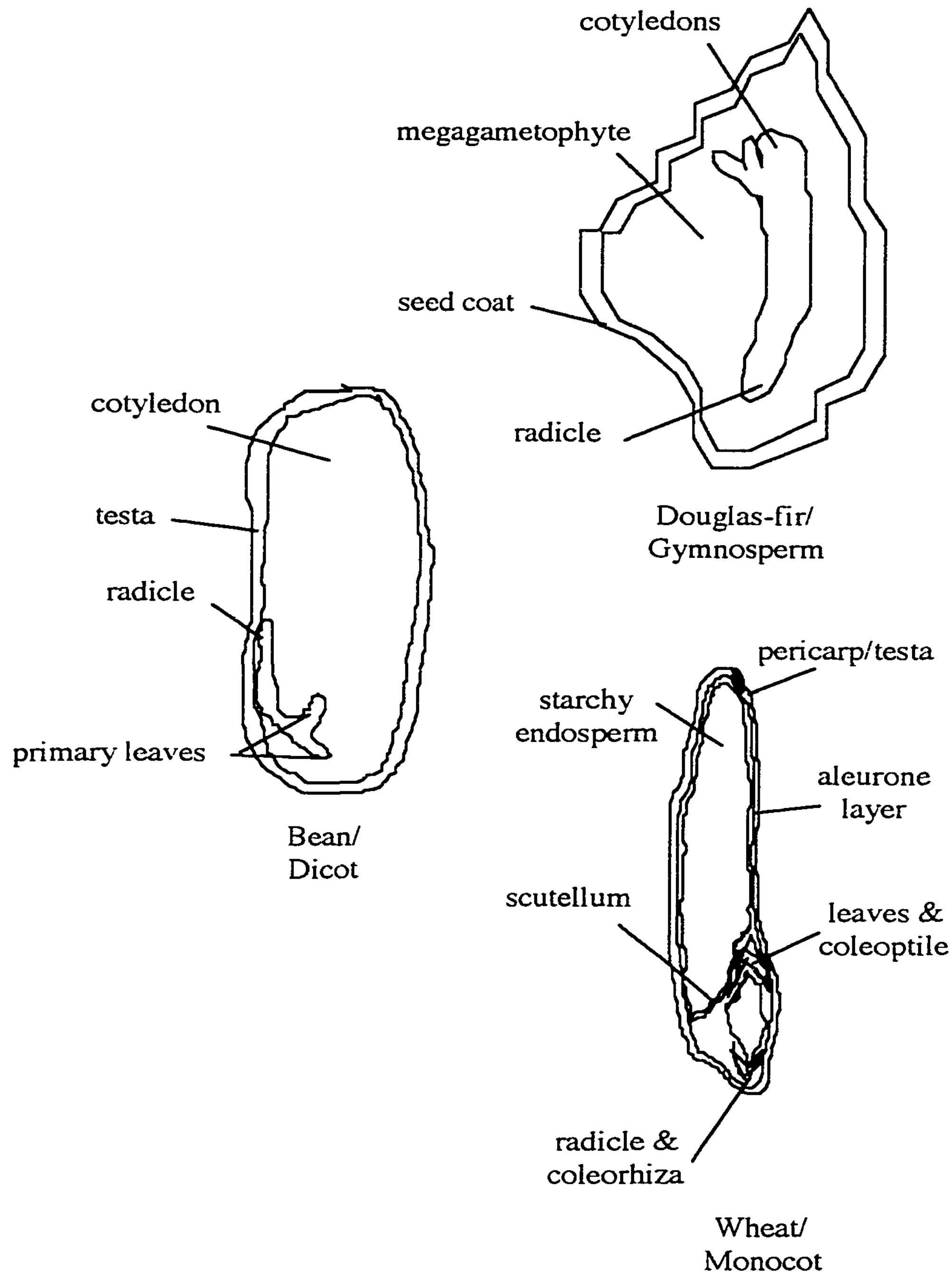


Figure 1.2. An overview and comparison of monocot, dicot and gymnosperm seed anatomy. The figure was adapted from Bewley and Black, 1994.

isocitrate lyase transcript and protein remained fairly constant while activity and the rate of *in vivo* isocitrate lyase protein synthesis declined. This pattern of expression is consistent for the role of isocitrate lyase in the mobilization of storage lipids during germination.

One of the most extensively studied classes of seed protein genes encode the LEA proteins that are thought to play a role in preventing damage from desiccation during the final stages of seed maturation (Dure et al., 1989). LEA transcripts increase during the late stage of embryogenesis when storage protein gene expression decreases, and then decline during germination. Some LEA genes are induced by desiccation as well as exogenous treatment with the phytohormone abscisic acid (ABA) during other developmental stages when they are thought to protect plant structures against water loss. The LEA proteins are hydrophilic, and have tandemly arranged repeated motifs believed to be important in their function as desiccation protectants (Dure et al., 1989). LEA proteins are grouped into classes based on their amino acid sequences and LEA cDNAs have been isolated from monocots, dicots and gymnosperms including rice, barley, wheat, cotton, rape (*B. napus*), radish, *Arabidopsis thaliana*, Douglas-fir (*Pseudotsuga menziesii*), and white spruce (*Picea glauca*) (Skriver and Mundy, 1990 and references therein; Jarvis et al., 1996; Dong and Dunstan, 1997).

In Douglas-fir, several LEA cDNAs were isolated from stratified seeds by differential screening of a cDNA library (Jarvis et al., 1996). Each LEA cDNA encoded a representative from three classes of LEA proteins identified from angiosperms (Dure et al., 1989). The LEA transcripts were present in the dry mature seed, declined in freshly imbibed seeds, then were induced transiently during the cold-moist treatment required to break seed dormancy (Jarvis et al., 1996; Jarvis et al., 1997; see section below on Seed Dormancy).

LEA cDNAs (PgEMB12, 14, 15) have also been isolated from white spruce somatic embryos by differential screening (Dong and Dunstan, 1997). The LEA transcripts were

induced by ABA, highly expressed in the cotyledonary embryo tissue during somatic embryogenesis and declined after germination. In contrast to the expression of angiosperm LEA genes, the white spruce LEA transcripts were detected early during embryogenesis in the immature embryos. This discrepancy may be a result of exogenous ABA application necessary for the stimulation of white spruce somatic embryogenesis. The transcripts were not detected in plantlets derived from the somatic embryos or in mature needles (Dong and Dunstan, 1996; Dong and Dunstan, 1997).

The expression of the LEA genes during embryogenesis and their induction by ABA appear to be common to angiosperms and gymnosperms. The ABA-responsive element (ABRE) involved in the transcriptional induction of ABA-inducible genes, including the LEA genes, have been characterized from angiosperms (Busk and Pages, 1998 and references therein). From more than 20 functional ABREs that have been characterized, the element is defined as a sequence of 8-10 base pairs with a core sequence of ACGT. The ABRE binds a number of basic leucine zipper transcriptional factors *in vitro*, but the DNA-binding protein responsible for ABA-induced gene transcription *in vivo* has not been identified. The regulatory sequences of gymnosperm LEA genes have yet to be characterized.

Seed Storage Protein Mobilization

The mobilization of seed storage proteins during germination and post-germination is a critical process that supplies amino acids for the synthesis of new proteins necessary for normal seedling development. Seed storage protein mobilization provides a model system for examining developmental patterns of gene expression that underlie changes in the metabolism during germination and post-germination. In cereal and dicots, storage protein mobilization involves the initial activity of metallo-proteinases, followed by a number of cysteine-proteinases (CysP) in a bulk hydrolysis phase, and finally by amino- and carboxy-

peptidases to yield amino acids (Shutov and Vaintraub, 1987). During germination, specific proteinases are synthesized *de novo* and targeted to storage protein vesicles where they initiate proteolysis of the protein reserves.

The most extensively studied plant proteinases have similarities to the mammalian lysosomal cathepsin class of CysP (Rogers et al., 1985; Shutov and Vaintraub, 1987; Koehler and Ho, 1988; Koehler and Ho, 1990a; Koehler and Ho, 1990b; Holwerda et al., 1990; Holwerda et al., 1992; Bethke et al., 1996). One of the best characterized of the plant CysP at the molecular level is aleurain of the barley aleurone cells. The aleurone layer of barley grains, a specialized secretory layer of cells surrounding the endosperm storage tissue, has provided insights into the regulation, targeting, secretion and activity of hydrolytic proteinases involved in protein mobilization. (Figure 1.2). During seed maturation, barley storage proteins (hordeins) are synthesized on rough endoplasmic reticulum (RER) and are targeted to and accumulate in specialized vacuolar derived membrane-bound vesicles referred to as protein bodies (Fincher, 1989). Aleurain gene expression is induced by gibberellic acid (GA), and inhibited by ABA, two plant growth regulators with opposing roles in germination (Rogers et al., 1985; see section on the Roles of ABA and GA below). Aleurain is encoded as a precursor polypeptide (proaleurain) that is targeted to the vacuolar compartment of the aleurone cells via two short amino acid sequences in the N-terminal propeptide, and is post-translationally processed to mature aleurain (Holwerda et al., 1990; Holwerda et al., 1992). A vacuolar targeting receptor that binds the N-terminal vacuolar targeting determinant of proaleurain was characterized from pea and is thought to direct aleurain from the Golgi to the vacuole via clathrin-coated vesicles (Kirsch et al., 1994). Two additional CysP, designated EP-A and EP-B, are induced by GA and secreted by the aleurone cells to the endosperm where they are involved in the degradation of the hordein, the native storage protein of barley (Koehler and Ho, 1990a; Koehler and Ho, 1990b; Davy et al., 1998). CysP are also implicated in the

degradation of the native storage proteins (zeins) from *Zea mays* (maize), and GA induced *de novo* synthesis of CysP occurs during rice germination (Abe et al., 1987; Arai et al., 1988; de Barros and Larkins, 1990; Watanabe et al., 1991; Watanabe et al., 1992).

Much less is known about the regulation, targeting, secretion, and enzymatic specificities of the hydrolases involved in storage protein mobilization of gymnosperms. The major seed storage proteins of gymnosperms share similar structural and solubility characteristics with the angiosperm seed proteins (Misra, 1994). Storage proteins are characterized by their solubilities in water (albumins), salt solutions (globulins), acid or alkali (glutelins), or aqueous alcohols (prolamins). The most abundant seed storage proteins of a number of gymnosperms are the crystalloid or "legumin-like" proteins which are similar to the 11S globulin (legumin) proteins of angiosperms. The activities of proteinases and peptidases implicated in the hydrolysis of *Pinus sylvestris* storage proteins have been characterized (Salmia and Mikola, 1975; Salmia and Mikola, 1976a, 1976b; Salmia et al., 1978; Salmia, 1980; Salmia and Mikola, 1980; Salmia, 1981a, 1981b; Misra, 1994). In resting (imbibed) *P. sylvestris* seeds, a pepstatin-sensitive (characteristic of an aspartic acid proteinase) proteinase accounts for the majority of the activity and is thought to be involved in the initial phase of storage protein hydrolysis that supplies amino acids for the *de novo* synthesis of other hydrolases (Salmia, 1981b). Based on inhibitor assays with *O*-phenanthroline (a chelator of divalent cations), a metallo-protease activity is present in germinating seeds, but little is known about this enzyme (Salmia et al., 1978). During germination and post-germination, the pepstatin-sensitive proteinase activity remains constant while an increase in two proteinase activities (proteinases I and II) with CysP characteristics increase in the MG. Proteinases I and II may be responsible for the bulk hydrolysis phase of the storage protein mobilization (Salmia, 1981b). Two groups of peptidases present in the storage tissue of resting *P. sylvestris* seeds are also implicated in protein reserve mobilization; one group includes alkaline peptidases that increase in activity

during mobilization (Salmia and Mikola, 1975); the second group includes aminopeptidases with activities that remain constant during mobilization (Salmia and Mikola, 1976b). A carboxypeptidase activity increases after the bulk of storage protein mobilization has occurred and is thought to play a role in senescence of the storage tissue (Salmia and Mikola, 1976a). Information on the molecular biology of these or other hydrolases involved in storage protein mobilization of gymnosperms is non-existent.

Seed Dormancy

Seed dormancy, a state of metabolic and developmental inactivity, serves as a mechanism that allows seeds to survive and germinate only at the right time and under specific conditions. Dormancy results from an internal block to germination that exists within the seed, and needs to be removed before germination can proceed. To overcome dormancy, a seed must encounter specific favourable environmental conditions, mainly light (quality or photoperiod), temperature (chilling or warming), and moisture (humidity) that vary between species. Once these conditions have been met, germination proceeds. Two types of dormancy have been defined. The first is referred to as "coat-imposed" dormancy where the embryo itself is not dormant, but the surrounding seed tissues impose a mechanical barrier to germination (Bewley and Black, 1994). Embryos dissected from these types of seeds will germinate readily. The second type is "embryo dormancy" when the embryo itself is dormant. Seeds can exhibit either one or both types of dormancy. The mechanisms that result in a dormant seed, especially the dormancy inherent in the embryo, are largely unknown (Bewley, 1997).

Changes in gene expression patterns and protein synthesis associated with the onset, maintenance and loss of seed dormancy have been examined in angiosperm and gymnosperm species (Morris et al., 1991; Hance and Bevington, 1992; Goldmark et al., 1992; Hong et al., 1992; Dyer, 1993; Li and Foley, 1994; Schneider and Gifford, 1994;

Johnson et al., 1995; Li and Foley, 1995; Mullen et al., 1996; Stacy et al., 1996; Jarvis et al., 1996; Jarvis et al., 1997). The steady-state amounts of some mRNAs and proteins increase while others decrease during seed imbibition. Different sets of genes are expressed in dormant versus non-dormant seeds. The identities of most of these genes are not known, and their association with dormancy is only correlative and not causative. For example, LEA proteins, thought to play a role in desiccation tolerance, accumulate normally during the late stages of seed maturation, are present in mature seeds and decrease during germination (Dure et al., 1989). However, there is only correlative evidence that LEA proteins have roles in the onset, maintenance or release of dormancy. A cDNA (pBS128) expressed preferentially in the embryos from hydrated dormant *Bromus secalinas* seeds was characterized (Goldmark, 1992). The pBS128 transcripts rapidly declined and disappeared in non-dormant seeds which subsequently germinated. Exogenous application of 50 μM ABA to non-dormant seeds inhibited germination and enhanced pBS128 transcript amounts. The transcript B15C from barley shares sequence similarity with pBS128 and increased in imbibed embryos from dormant grains (Stacy et al., 1996). The B15C transcript was down-regulated in the germinating embryo and by exogenous GA application. The protein (PER1) encoded by the B15C transcript is similar to a group antioxidants called peroxiredoxins. PER1 reduced oxidative damage *in vitro*, and may function as a scavenger of reactive oxygen species produced as by-products from desiccation and respiration occurring during late embryogenesis, imbibition and germination.

A set of dormancy and non-dormancy-associated cDNAs was isolated by differential display (Johnson et al., 1995). Five dormancy-associated gene transcripts increased in dormant embryos during the first 48 hours of imbibition, while another set of transcripts was more abundant in non-dormant embryos during imbibition. One of the non-dormancy-associated transcripts (AFN3) shared similarity to a glutathione peroxidase-like cDNA, but

none of the other transcripts shared similarity with sequences in the databases. The glutathione peroxidase-like protein may act as a scavenger for free radicals produced from the reinitiation of metabolism during germination.

A cDNA homologous to the transcriptional factor VIVIPAROUS 1 (VP1) was cloned from mature *Avena fatua* embryos (Jones et al., 1997). The expression of the *A. fatua* VP1 homolog (afVP1) was positively correlated to the dormant phenotype and the length of time required for after-ripening (warm, dry conditions that break dormancy). The regulation of some dormancy associated gene sets includes post-transcriptional mechanisms (Li and Foley, 1996). A set of mRNAs, including transcripts encoding a LEA protein, had longer half-lives in dormant than in non-dormant *Avena fatua* (wild oat) seeds. Only a slight difference in transcriptional activity was observed which suggests that these dormancy-associated genes are regulated at the level of mRNA stability.

Douglas-fir seeds are subject to dormancy that is overcome by a moist chilling treatment (stratification) for several weeks (Edwards, 1986; Taylor et al., 1993). For instance, loss of dormancy and high germination rates are only achieved after seeds are stratified at 4 °C for 6 weeks, whereas treatment at 20 °C is not effective in breaking dormancy (Taylor et al., 1993). During stratification of Douglas-fir seeds, the steady-state amounts of several transcripts increase. These include: LEA genes, histone H1 and for the ß-subunit of the 20S proteosome (Taylor and Davies, 1995). However, the expression of histone H1 and the 20S proteosome ß-subunit is also induced by pretreatment of seeds at 20 °C indicating that they may not be directly associated with the dormancy-breaking stratification requirement. The expression pattern of one of the LEA genes (DF65) correlated to the stratification requirement of Douglas-fir seeds, whereas expression of two LEA genes (DF6 and DF77) was independent of temperature (Jarvis et al., 1996; Jarvis et al., 1997). The expression DF65 was higher in the MG and occurred prior to that in the embryo. This

pattern of DF65 expression is consistent with the hypothesis that the surrounding MG tissue plays a role in imposing dormancy on the embryo (Jarvis et al., 1997). The function(s) of the LEA gene products in overcoming dormancy in Douglas-fir seeds requires further investigation.

Roles of Abscisic Acid and Gibberellic Acid in Dormancy and Germination

The plant growth regulators ABA and GA control seed dormancy and germination. ABA is thought to promote dormancy through the induction of gene sets expressed during maturation drying, whereas GA promotes and maintains germination (Skriver and Mundy, 1990; Bewley, 1997). The best evidence for this arises from the studies with the vivipary and germination mutants of *Zea mays* and *A. thaliana* (McCarty, 1995). Vivipary is the precocious sprouting of the seeds before a dormant state can be achieved. Most of these mutants have alterations in the synthesis of ABA or GA, or in the ability to perceive these growth regulators. The resultant mutants are either unable to germinate (dormant phenotype) or unable to achieve a dormant state (precocious germination phenotype). Molecular work with ABA-Insensitive (ABI) mutants has led to the identification of gene products involved in ABA signal transduction. The genes include those encoding transcriptional activators from *A. thaliana* (*abi3*; Giraudat et al., 1992) and maize (*vp1*; McCarty et al., 1991), a DNA-binding protein similar to the plant-specific APETALA2 type transcriptional regulators (*abi4*; Finkelstein et al., 1998) and two serine-threonine phosphatases from *A. thaliana* (*abi1* and *abi2*; Leung et al., 1994; Meyer et al., 1994). ABA normally inhibits wild-type *A. thaliana* seed germination. The *abi3*, *vp1*, *abi1*, and *abi2* mutations result in a decrease in the sensitivity to ABA, and mutant seeds germinate in the presence of exogenously applied ABA. However, the question remains why some wild-type seeds which are sensitive to ABA, fail to germinate under control conditions (i.e. in the absence of exogenous ABA application). The mutants lacking the transcriptional

activators encoded by *abi3* or *vp1* have reduced amounts of the Em protein (a LEA protein) and another LEA protein, both ABA inducible seed maturation-specific proteins (Paiva and Kriz, 1994). The significance of these mutational effects in relation to germination is not known. Another gene encoding the ß subunit of a farnesyl transferase was identified from mutants with an enhanced response to ABA (*era1*, Cutler et al., 1996). Farnesyl transferases consist of α and β subunits that dimerize to form an enzyme which catalyzes the attachment of farnesyl pyrophosphate to proteins containing a C-terminal targeting motif. Farnesylation of signal transduction proteins anchors them to lipids or proteins in the membrane. The *era1* mutation results in an increase in sensitivity to ABA and a loss of farnesyl transferase function which suggests that target protein(s) involved in ABA signal-transduction may require negative regulation through farnesylation. The results from the ABA mutant studies indicate that ABA perception involves a cascade of events that includes a number of factors.

Treatment with exogenous ABA inhibits germination of stratified Douglas-fir seeds (Jarvis et al., 1997). Endogenous ABA levels were similar in dormant and non-dormant seeds, however, non-dormant seeds had a reduced sensitivity to ABA. The LEA genes were not induced by exogenous ABA treatment of seeds. In contrast, exogenous methyl-jasmonate (Me-JA; a volatile plant growth regulator) promoted dormant Douglas-fir seed germination and induced LEA gene expression. The expression of a gene encoding a low molecular weight heat shock protein (LMW HSP) was also induced by exogenous Me-JA application during Douglas-fir seed stratification, but the significance of Me-JA induced gene expression in relation to seed dormancy and germination is unknown (Kaukinen et al., 1996).

Differences Between Gymnosperm and Angiosperm Seeds

Gymnosperm and angiosperm seeds have genetic, morphological and anatomical

differences (Figure 1.2). Gymnosperm seeds are comprised of a maternally derived diploid seed coat, the haploid MG storage tissue and the diploid embryo derived both maternally and paternally. The maternal:paternal gene contribution in gymnosperms, including Douglas-fir, is 4n:1n and results in a strong maternal control over seed germination (El-Kassaby et al., 1992). The main storage tissue of angiosperms can either be the cotyledons of the embryo derived from the fusion of the maternal and paternal nuclei, or the endosperm derived from the fusion of two maternal polar nuclei and the paternal pollen tube nucleus. In both cases, the storage tissues of angiosperms are maternally and paternally derived.

Gymnosperms and angiosperms differ in both the size and structure of their respective genomes (Kinlaw and Neale, 1997). The haploid genome sizes (expressed as picogram per haploid genome) for angiosperms range from 0.45 for rice and 1.0 for tomato, to 2.6 for maize and 2.7 for lettuce (Bernatzky and Tanksley, 1986; Landry et al., 1987; Causse et al., 1994; Shen et al., 1994). In the gymnosperm loblolly pine (*Pinus taeda*) haploid genome size is 22.0 (Devey et al., 1994). This may in part be due to higher gene copy numbers in gymnosperm. Several low copy and single copy genes of angiosperms, including those encoding chaperonin 60ß, thiolase, elongation factor 1α, acid phosphatase, actin-depolymerizing factor, and alcohol dehydrogenase occur as larger gene families in loblolly pine (Kinlaw and Neale, 1997). However, these studies are based on Southern blot hybridization which does not distinguish between functional and pseudo-genes. The existence of larger gene families in gymnosperms is not simply due to polyploidy common to angiosperms. All known species of pine (more than 100) are diploid, with 24 chromosomes in the diploid genome (Kinlaw and Neale, 1997). It is clear that the gymnosperm and angiosperm genomes have evolved differently but the significance of the genomic differences remains to be determined.

Gymnosperms differ from angiosperms in a number of fundamental developmental

processes. For example, light regulated photosynthesis associated gene sets of angiosperms are not regulated by light in gymnosperms (Alosi et al., 1990). The biosynthesis of lignin differs between gymnosperms and angiosperms (Whetten and Sederoff, 1995). The synthesis and accumulation of the lignin glucoside monomers *p*-hydroxycinnamyl alcohol glucoside, coniferin and syringin occurs almost exclusively in gymnosperms and may represent an ancestral form of lignin biosynthesis. The lignin monomers of gymnosperms are less methylated than those of angiosperms, a characteristic that makes gymnosperm wood pulp a more difficult and expensive source of the cellulose fibres used for paper making. The loblolly pine multifunctional *O*-methyltransferase differs in both structure and activity from the angiosperm enzymes involved in the methylation of lignin monomers (Laigeng et al., 1997). The pine enzyme methylates both caffeic acid and caffeoyl CoA, enzymatic functions that are mediated by two separate enzymes in angiosperms. The dual activity makes this enzyme a potential target for genetically altering the methylation level of the lignin monomers to make gymnosperm lignin more angiosperm-like.

A cDNA (PM2.1) encoding a novel metallothionein-related (MT-related) protein was isolated from developing Douglas-fir zygotic embryos (Chatthai et al., 1997). A comparison of the predicted amino acid sequences of the Douglas-fir and angiosperm MT-related proteins indicated divergence has occurred within conserved amino acid motifs during higher plant evolution. However, a comparison of the promoter of the corresponding PM2.1 gene with other angiosperm MT-related gene elements indicated the presence of conserved regulatory elements (Chatthai and Misra, unpublished data).

The temporal accumulation patterns of seed storage proteins differ between gymnosperms and angiosperms. In gymnosperms, seed storage proteins start to accumulate soon after fertilization and before the embryo has formed cotyledons (Misra, 1994; Chatthai and Misra, 1998). However, a comparison of the promoter sequence of the 2S albumin storage

protein genes of Douglas-fir revealed the presence of conserved regulatory elements (Chatthai and Misra, unpublished data). In angiosperms, storage protein accumulation begins during the mid-to late-maturation stage after the cotyledons have formed (Goldberg et al., 1989). In the gymnosperm *Ginkgo biloba*, the promoter regions of the legumin storage protein genes contain sequence motifs which are known to function as regulatory elements that direct seed-specific expression of angiosperm legumins (Häger et al., 1995). The legumin genes of the *G. biloba* contain four instead of the three conserved introns found in all known angiosperm legumin genes. This indicates that the evolution of legumin genes of higher plants may have involved the loss or gain of an intron.

Objectives

To understand the molecular mechanisms involved in the regulation of seed germination and early seedling growth of gymnosperms, it is necessary to isolate and identify genes expressed during these developmental stages. Therefore, the first objective of this dissertation was to isolate and identify genes (cDNAs) that are developmentally regulated during Douglas-fir seed germination and early post-germination. The cDNAs were used as probes to examine the expression of the corresponding genes, and their nucleic acid and deduced amino acid sequences were determined and compared to the sequence databases. Selected cDNA clones were used as probes in additional analyses that included: examinations of the temporal- and spatial-steady-state transcript amounts during embryogenesis, stratification, germination, post-germination, and the environmental and hormonal factors that control their expression. In order to examine the encoded polypeptides of selected clones, the production of antisera was pursued. Antisera were tested and used in examinations of the subcellular localization and accumulation of the encoded products.

Literature Cited

Abe, K., Kondo, H., and Arai, S. (1987) Purification and properties of a cysteine proteinase from germinating rice seeds. Agricultural and Biological Chemistry 51: 1509-1514

Apostol, I., Heinstein, P. F., and Low, P. S. (1989) Rapid stimulation of an oxidative burst during elicitation of cultured plant cells. Role in defense and signal transduction. Plant Physiology 90: 109-116

Alosi, M. C., Neale, D. B. and Kinlaw, C. S. (1990) Expression of cab genes in Douglas-fir is not strongly regulated by light. Plant Physiology 93: 829-832

Arai, S., Hosoyama, H., and Abe, K. (1988) Gibberellin-induced cysteine proteinase occurring in germinating rice seeds and its specificity for digesting oxidized insulin B-chain. Agricultural and Biological Chemistry 52: 2957-2959

Bernatzky, R., and Tanksley, S. D. (1986) Majority of random cDNA clones correspond to single loci in the tomato genome. Molecular and General Genetics 203: 8-14

Bethke, P. C., Hillmer, S., Jones, R.L. (1996) Isolation of intact protein storage vacuoles from barley aleurone. Plant Physiology 110: 521-529

Bewley, J. D. (1997) Seed germination and dormancy. Plant Cell 9: 1055-1066

Bewley, J. D., and Marcus, A. (1990) Gene expression in seed development and germination. Progress in Nucleic Acid Research and Molecular Biology 38: 165-193

Bewley, J. D., and Black, M. (1994) Seeds: Physiology of Development and Germination. Plenum Press, New York

Busk, P. K., and Pages, M. (1998) Regulation of abscisic acid-induced transcription. Plant Molecular Biology 37: 425-435

Caliskan, M., and Cuming, A. C. (1998) Spatial specificity of H_2O_2-generating oxalate oxidase gene expression during wheat embryo germination. The Plant Journal 15: 165-171

Cassab, G. I., and Varner, J. E. (1988) Annual Review of Plant Physiology and Plant Molecular Biology 39: 321-353

Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P. C., and Harrington, S. E. (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138: 1251-1274

Chatthai, M., Kaukinen, K. H., Tranbarger, T. J., Gupta, P. K., and Misra, S. (1997) The isolation of a novel metallothionein-related cDNA expressed in somatic and zygotic embryos of Douglas-fir: regulation by ABA, osmoticum, and metal ions. Plant Molecular Biology 34: 243-254.

Chatthai, M., and Misra, S. (1998) Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziesii* [Mirb.] Franco. Planta 206: 138-145

Comai, L., Dietrich, R. A., Maslyar, D. J., Baden, C. S. and Harada, J. J. (1989) Coordinate expression of transcriptionally regulated isocitrate lyase and malate synthase genes in *Brassica napus* L. Plant Cell 1: 293-300

Comai, L., and Harada, J. J. (1990) Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. Proceedings of National Academy of Sciences USA 87: 2671-2674

Cutler, S., Ghassemian, M., Bonetta, D., Cooney, S., and McCourt, P. (1996) A protein farnesyl transferase involved in abscisic acid signal transduction in *Arabidopsis*. Science 273: 1239-1241

Davy, A., Svendsen, I., Sørensen, S. O., Sørensen M. B., Rouster, J., Meldal, M., Simpson, D. J, and Cameron-Mills, V. (1998) Substrate specificity of barley cysteine endoproteases EP-A and EP-B. Plant Physiology 117: 255-261

de Barros, E. G., and Larkins, B.A. (1990) Purification and characterization of zein-degrading proteases from endosperm of germinating maize seeds. Plant Physiology 94: 297-303

Devey, M. E., Fiddler, T. A., Liu, B. H., Knapp, S. J., and Neale, D. B. (1994) An RFLP linkage map for loblolly pine based on a three-generation outbred pedigree. Theoretical and Applied Genetics 88: 273-278

Dong, J.-Z., and Dunstan, D. I. (1996) Expression of abundant mRNAs during somatic embryogenesis of white spruce [*Picea glauca* (Moench) Voss].

Dong, J.-Z., and Dunstan, D. I. (1997) Characterization of cDNAs representing five abscisic acid-responsive genes associated with somatic embryogenesis in *Picea glauca*, and their responses to abscisic acid stereostructure. Planta 203: 448-453

Dommes, J., and Van der Walle, C. (1990) Polysome formation and incorporation of new ribosomes into polysomes during germination of the embryonic axis of maize. Physiologia Plantarum 79: 289-296

Dure III, L., Crouch, M., Harada, J., Ho, T.-H. D., Mundy, J., Quatrano, R., Thomas, T., and Sung, Z. R. (1989) Common amino acid sequence domains among the LEA proteins of higher plants. Plant Molecular Biology 12: 475-486

Dyer, W. E. (1993) Dormancy-associated embryonic mRNAs and proteins in imbibing *Avena fatua* caryopses. Physiologia Plantarum 88: 201-211

Edwards, D. G. W. (1986) Special pre-chilling techniques for tree seeds. Journal of Seed Technology 10: 151-171

El-Kassaby, Y. A., Edwards, D. G. W., and Taylor, D. W. (1992) Genetic control of germination parameters in Douglas-fir and its importance for domestication. Silvae Genetica 41: 48-54

Fincher, G. B. (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. Annual Review of Plant Physiology and Plant Molecular Biology 40: 305-346

Finkelstein, R. R., Wang, M. L., Lynch, T. J., Rao, S., and Goodman, H. M. (1998) The *Arabidopsis* abscisic acid response locus ABI4 encodes an APETALA2 domain protein. Plant Cell 10: 1043-1054

Galua, G. A., Jakobsen, K. S., and Hughes, D. W. (1991) The controls of late dicot embryogenesis and early germination. Physiologia Plantarum 81: 280-288

Gifford, D. J., Dale, P. L., and Wenzel, K. A. (1991) Lodgepole pine seed germination. III. Patterns of protein and nucleic acid synthesis in the megagametophyte and embryo. Canadian Journal of Botany 69: 301-305

Giraudat, J., Hauge, B. M., Valon, C., Smalle, J., Parcy, F., and Goodman, H. M. (1992) Isolation of the Arabidopsis *ABI3* gene by positional cloning. Plant Cell 4: 1251-1261

Goldberg, R. B., Barker, S. J., and Perez-Grau, L. (1989) Regulation of gene expression during plant embryogenesis. Cell 56: 149-160

Goldmark, P. J., Curry, J., Morris, C. F., and Walker-Simmons, M. K. (1992) Cloning and expression of an embryo-specific mRNA up-regulated in hydrated dormant seeds. Plant Molecular Biology 19: 433-441

Groome, M. C., Axler, S. R., and Gifford, D. J. (1991) Hydrolysis of lipid and protein reserves in loblolly pine seeds in relation to protein electrophoretic patterns following imbibition. Physiologia Plantarum 83: 99-106

Häger, K. -P., Braun, H., Czihal, A., Müller, B., Bäumlein, H. (1995) Evolution of seed storage protein genes of *Ginkgo biloba*. Journal of Molecular Evolution 41: 457-466

Hance, B. A, and Bevington, J. M. (1992) Changes in protein synthesis during stratification and dormancy release in embryos of sugar maple (*Acer saccharum*). Physiologia Plantarum 86: 365-371

Harada, J. J., Baden, C. S., and Comai, L. (1988) Spatially regulated genes expressed during seed germination and postgerminative development are activated during embryogeny. Molecular and General Genetics 212: 466-473

Holwerda, B. C., Galvin, N. J., Baranski, T. J., and Rogers, J. C. (1990) *In vitro* processing of aleurain, a barley vacuolar thiol protease. Plant Cell 2: 1091-1106

Holwerda, B. C., Padgett, H. S., and Rogers, J. C. (1992) Proaleurain vacuolar targeting is mediated by short contiguous peptide interactions. Plant Cell 4: 307-318

Hong, B., Barg, R., and Ho, T.-H. D. (1992) Developmental organ-specific expression of an ABA- and stress-induced protein in barley. Plant Molecular Biology 18: 663-674

Hughes, D. W., and Galau, G. A. (1989) Temporally modular gene expression during cotyledon development. Genes and Development 3: 358-369

Jarvis, S. B., Taylor, M. A., MacLeod, M. R., and Davies, H. V. (1996) Cloning and characterisation of three cDNA clones of genes that are differentially expressed during dormancy-breakage in the seeds of Douglas fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 147: 559-566

Jarvis, S. B., Taylor, M. A., Bianco, J., Corbineau, F., and Davies, H. V. (1997) Dormancy-breakage in seeds of Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco).

Support for the hypothesis that LEA gene expression is essential for this process. Journal of Plant Physiology 151: 457-464

Johnson, R. R., Cranston, H. J., Chaverra, M. E., and Dyer, W. E. (1995) Characterization of cDNA clones for differentially expressed genes in embryos of dormant and nondormant *Avena fatua* L. caryopses. Plant Molecular Biology 28: 113-122

Jones, H. D., Peters, N. C. B., and Holdsworth, M. J. (1997) Genotype and environment interact to control dormancy and differential expression of the VIVIPAROUS 1 homologue in embryos of *Avena fatua*. Plant Journal 12: 911-920

Kaukinen, K. H., Tranbarger, T. J., and Misra, S. (1996) Post-germination-induced and hormonally dependent expression of low-molecular-weight heat shock protein genes of Douglas fir. Plant Molecular Biology 30: 1115-1128

Kinlaw, C. S., and Neale, D. B. (1997) Complex gene families in pine genomes. Trends in Plant Science 2: 356-359

Kirsch, T., Paris, N., Butler, J. M., Beevers, L., and Rogers, J. C. (1994) Purification and initial characterization of a potential plant vacuolar targeting receptor. Proceedings of National Academy of Sciences USA 91: 3403-3407

Koehler, S. M., and Ho, T.-H. D. (1988) Purification and characterization of gibberellic acid-induced cysteine endoproteases in barley aleurone layers. Plant Physiology 87: 95-103

Koehler, S. M., and Ho, T.-H. D. (1990a) Hormonal regulation, processing, and secretion of cysteine proteinases in barley aleurone layers. Plant Cell 2: 769-783

Koehler, S. M., and Ho, T.-H. D. (1990b) A major gibberellic acid-induced barley aleurone cysteine proteinase which digests hordein: Purification and characterization. Plant Physiology 94: 251-258

Laigeng, L., Popko, J. L., Zhang, X. H., Osakabe, K., Tsai, C. J., Joshi, C. P., and Chiang, V. L. (1997) A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. Proceedings of National Academy of Sciences USA 94: 5461-5466

Lalonde, L and Bewley, J. D. (1986) Desiccation during germination and its effects on the messenger RNA population. Journal of Experimental Botany 37: 754-764

Landry, B. S., et al. (1987) Comparison of restriction endonucleases and sources of probes for their efficiency in detecting restriction fragment length polymorphisms in lettuce (*Lactuca sativa* L.). Theoretical and Applied Genetics 74: 646-653

Lane, B. G. (1991) Cellular desiccation and hydration: developmentally regulated proteins, and the maturation and germination of seed embryos. FASEB Journal 5: 2893-2901

Lane, B. G., Dunwell, J. M., Ray, J. A., Schmitt, M. R., and Cuming, A. C. (1993) Germin, a protein marker of early plant development, is an oxalate oxidase. Journal of Biological Chemistry 268: 12239-12242

Leal, I., and Misra, S. (1993) Developmental gene expression in conifer embryogenesis and germination. III. Analysis of crystalloid protein mRNAs and desiccation protein mRNAs in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss). Plant Science 88: 25-37

Leung, J., Bouvier-Durand, M., Morris, P-C. Guerrier, D., Chefdor, F., and Giraudat, J. (1994) *Arabidopsis* ABA response gene *ABI1*: features of a calcium-modulated protein phosphatase. Science 264: 1448-1452

Li, B., and Foley, M. E. (1994) Differential polypeptide patterns in imbibed dormant and after-ripened *Avena fatua* embryos. Journal of Experimental Botany 45: 275-279

Li, B., and Foley, M. E. (1995) Cloning and characterization of differentially expressed genes in imbibed dormant and afterripened *Avena fatua* embryos. Plant Molecular Biology 29: 823-831

Li, B., and Foley, M. E. (1996) Transcriptional and posttranscriptional regulation of dormancy-associated gene expression by afterripening in wild oat. Plant Physiology 110: 1267-1273

Luttrell, B. M. (1993) The biological relevance of the binding of calcium ions by inositol phosphates. Journal of Biological Chemistry 268: 1521-1524

McCarty, D. R., Hattori, T., Carson, C. B., Vasil, V., Lazar, M., and Vasil I. K. (1991) The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. Cell 66: 895-905

McCarty, D. R. (1995) Genetic control and integration of maturation and germination pathways in seed development. Annual Review of Plant Physiology and Plant Molecular Biology 46: 71-93

Meyer, K., Leube, M., and Grill, E. (1994) A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. Science 264: 1452-1455

Misra, S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis, and seed germination: biochemical and molecular advances. Seed Science Research 4: 357-384

Morris, C. F., Anderberg, R. J., Goldmark, P. J., and Walker-Simmons, M. K. (1991) Molecular cloning and expression of abscisic-acid responsive genes in embryos of dormant wheat seeds. Plant Physiology 95: 814-821

Mullen, R. T., King, J. E., and Gifford, D. J. (1996) Changes in mRNA populations during loblolly pine (*Pinus taeda*) seed stratification, germination and postgerminative growth. Physiologia Plantarum 97: 545-553

Mullen, R. T., and Gifford, D. J. (1997) Regulation of two loblolly pine (*Pinus taeda L.*) isocitrate lyase genes in megagametophytes of mature stratified seeds and during postgerminative growth. Plant Molecular Biology 33: 593-604

Paiva, R., and Kriz, A. L. (1994) Effect of abscisic acid on embryo-specific gene expression during normal and precocious germination in normal and viviparous maize (*Zea mays*) embryos. Planta 192: 332-339

Rogers, J. C., Dean, D., and Heck, G. R. (1985) Aleurain: a barley thiol protease closely related to mammalian cathepsin H. Proceedings of National Academy of Sciences USA 82: 6512-6516

Salmia, M. A., and Mikola, J. J. (1975) Activities of two peptidases in resting and germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 33: 261-265

Salmia, M. A., and Mikola, J. J. (1976a) Localization and activity of a carboxypeptidase in germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 36: 388-392

Salmia, M. A., and Mikola, J. J. (1976b) Localization and activity of naphthylamidases in germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 38: 73-77

Salmia, M. A., and Mikola, J. J. (1980) Inhibitors of endogenous proteinases in the seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 48: 126-130

Salmia, M. A., Nyman, S. A., and Mikola, J. J. (1978) Characterization of the proteinases present in germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 42: 252-256

Salmia, M. A. (1980) Inhibitors of endogenous proteinases in Scots pine seeds: Fractionation and activity changes during germination. Physiologia Plantarum 48: 266-270

Salmia, M. A. (1981a) Fractionation of the proteinases present in the endosperm of germinating seed of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 51: 253-258

Salmia, M. A. (1981b) Proteinase activities in resting and germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 53: 38-47

Sánchez-Martínez, Puigdomènech, P., and Pagès, M. (1986) Regulation of gene expression in developing *Zea mays* embryos. Protein synthesis during embryogenesis and early germination of maize. Plant Physiology 82: 543-549

Sandoval, J. A., Huang, Z.-H., Garret, D. C., Gage, D. A., and Chapman, K. D. (1995) *N*-Acylphosphatidylethanolamine in dry and imbibing cotton seeds. Amounts, molecular species, and enzymatic synthesis. Plant Physiology 109: 269-275

Schneider, W. L., and Gifford, D. J. (1994) Loblolly pine seed dormancy. I. The relationship between protein synthesis and the loss of dormancy. Physiologia Plantarum 90: 246-252

Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, B., McCreery, T., Helentjaris, T., Baysdorfer, C., Almira, E., Ferl, R. J., Habben, J. E., and Larkins, B. (1994) Partial sequencing and mapping of clones from two maize cDNA libraries. Plant Molecular Biology 26: 1085-1101

Showalter, A. M. (1993) Structure and function of plant cell wall proteins. Plant Cell 5: 9-23

Shutov, A. D., and Vaintraub, I. A. (1987) Degradation of storage proteins in germinating seeds. Phytochemistry 26: 1557-1566

Simon, E. W., and Raja Harun, R. M. (1972) Leakage during seed imbibition. Journal of Experimental Botany 23: 1076-1085

Skriver, K., and Mundy, J. (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503-512

Stabel, P., Eriksson, T., and Engström, P. (1990) Changes in protein synthesis upon cytokinen-mediated adventitious bud induction and during seedling development in Norway spruce, *Picea abies*. Plant Physiology 92: 1174-1183

Stacy, R. A. P., Munthe, E., Steinum, T., Sharma, B., and Aalen, R. B. (1996) A peroxiredoxin antioxidant is encoded by a dormancy-related gene, *Per1*, expressed during late development in the aleurone and embryo of barley grains. Plant Molecular Biology 31: 1205-1216

Taylor, M. A., Davies, H. V., Smith, S. B., Abruzzese, A., and Gosling, P. G. (1993) Cold-induced changes in gene expression during dormancy breakage in seeds of Douglas-fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142: 120-123

Taylor, M. A., and Davies, H. V. (1995) Changes in gene expression during imbibition of seeds of Douglas fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142: 54-58

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

Watanabe, H., Abe, K., Emor, Y., Hosoyama, H., and Arai S. (1991) Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). Journal of Biological Chemistry 266: 16897-16902

Watanabe, H., Abe, K., and Arai, S. (1992) Gibberellin-responsive gene expression taking place with oryzains α and γ as cysteine proteinases of rice seeds. Bioscience, Biotechnology and Biochemistry 56: 1154-1155

Weir, E. M., Riezman, H., Grieneberger, J-M., Becher, W. M., and Leaver. C. J. (1980) Regulation of glyoxysomal enzymes during germination of cucumber. European Journal of Biochemistry 112: 469-477

Welbaum, G. E., and Bradford, K. J. (1990) Water relations of seed development and germination in muskmelon (*Cucumis melo* L.). V. Water relations of imbibition and germination. Plant Physiology 92: 1046-1062

Whetten, R., and Sederoff, R. (1995) Lignin Biosynthesis. Plant Cell 7: 1001-1013

CHAPTER 2

The Molecular Characterization of a Set of cDNAs Differentially Expressed During Douglas-fir Germination and Post-Germination

Abstract

To identify genes with increased expression during germination and early seedling development, a Douglas-fir cDNA library was differentially screened with cDNA probes synthesized using mRNA isolated from mature seeds and 6-day-old seedlings. Partial DNA sequence analysis and predicted amino acid sequences revealed cDNA clones encoding sequences with similarity to several plant proteins including: a chaperonin 60ß (cpn60ß), a low molecular weight heat shock protein (LMW HSP), a luminal binding protein (BiP), a type II chlorophyll a/b-binding protein (CAB), and a cysteine proteinase (CysP). In mature seeds, transcripts corresponding to each cDNA clone were detected by northern blot and their quantities increased after 2 days of exposure to the germination conditions. The clone with identity to the CysP *(PM3-3)* and clone *PM6-7* had peak transcript amounts during the germination phase, whereas the cpn60ß *(PM3)*, a LMW HSP *(PM4-5)*, BiP *(PM5-1)*, and CAB *(PM6-3)* showed peaks during the post-germination phase. All the transcripts were present at lower amounts in 10-week-old needles except for CAB. As a comparison, an additional clone *(PM14)* that was present at similar quantities in both mature seeds and in 6-day-old seedlings was isolated and found to be similar to a NADPH-cytochrome P450 reductase (CPR) (EC 1.6.2.4). CPR gene expression changed little during germination and early seedling development with lower transcript amounts in 10-week-old needles. The significance of the individual genes, their expression, and possible roles during germination is discussed.

Introduction

Desiccation, an essential part of normal seed development, may activate genes for germination and turn off the synthesis of embryogenesis related proteins (Misra and Bewley, 1985; Kermode, 1990). Several studies have revealed the presence of mRNA in mature seed tissues, referred to as stored mRNA, that may be immediately available on rehydration for translation of the proteins required for germination (Delseny et al., 1977; Harada et al., 1988; Whitmore, 1991; Masumori et al., 1992; Taylor et al., 1993). Imbibition of mature dormant angiosperm and gymnosperm seeds leads to the appearance of new and temporally distinct sets of mRNAs and gene products that are absent in mature seeds (Comai and Harada, 1990; reviewed by Lane, 1991; Taylor and Davies, 1995). Several of these gene products are involved in the transformation of storage compounds into energy and anabolic substrates for germination (Bewley and Marcus, 1990). Examples include enzymes of the glyoxylate cycle that have been identified from both angiosperm and gymnosperm sources (Weir et al., 1980; Comai et al., 1989; Mullen and Gifford, 1997) and enzymes involved in storage product mobilization and nitrogen assimilation (reviewed in Bewley and Marcus, 1990; Koehler and Ho, 1990; Washio and Ishikawa, 1992). In *Brassica napus*, the expression of malate synthase and isocitrate lyase genes is coordinately regulated at the transcriptional level (Comai et al., 1989). Defined mRNAs associated with axis growth e.g. proline and hydroxyproline rich cell-wall proteins, phytochrome etc. have also been identified (reviewed by Bewley and Marcus, 1990). However, the growth regulatory genes involved in germination have not been identified.

Seed germination is a complex process in conifers. Not only do seeds of many conifer species require long periods of chilling but also the germination rate and seed vigor are often low (Edwards, 1986). Compared with angiosperms, little is known about the regulation of gene activity during germination and early seedling development of conifer

seeds. In conifers, the haploid megagametophyte (MG) is the main tissue containing storage reserves mainly in the form of lipids (Ching, 1963; Owens et al., 1993) and proteins (Green et al., 1991; Misra and Green, 1991). The biochemical changes that occur in the MG and embryos of germinating conifer seeds have been investigated at a number of levels. Changes in the amounts of lipids, starch, proteins, and nucleic acids have been reported, and the characteristics of the enzymes responsible for starch and storage protein degradation, and lipid conversion have also been examined (reviewed by Misra, 1994 and references therein). To elucidate the regulatory mechanisms involved in the expression of genes during the germination of conifer seeds and early seedling development, the following study was initiated on the cloning and characterization of developmentally regulated genes. This section describes the cloning and analysis of a set of cDNAs isolated from a Douglas-fir germination-specific library. Northern blot analysis indicated that these mRNAs are differentially expressed during germination and early seedling development and are present in lower amounts in mature seeds and 10-week-old needles.

Materials and Methods

Plant material and growth conditions

Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco) seeds were imbibed overnight with sterile distilled water at 2-4 °C. The seeds were then blotted dry, placed into a plastic bag and stored at 2-4 °C for 3 weeks for stratification, a treatment that improves germination of several conifer species (Edwards, 1980) including Douglas-fir (Allen, 1941; El-Kassaby et al., 1992). Stratified seeds were germinated in trays in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hr. photoperiod. Cool White High Output fluorescent bulbs (F72T12/CW/1500) from Canadian General Electric Company Limited and 60 watt Country/Royale incandescent bulbs from Philips were used to provide a light quantity of about 300 mmol m^{-2} s^{-1}. Whole mature seeds, stratified seeds, germinating seeds at 2, 4, 6, 8 and 14 days after exposure to the germination conditions, and 10-week-old needles were collected, immediately frozen in liquid nitrogen and stored at -80 °C.

Construction and differential screening of a Douglas-fir cDNA library

Total RNA isolated according to Verwoerd et al. (1989), with modifications described previously (Leal and Misra, 1993a), was pooled from 4-6-day-old Douglas-fir seedlings for poly$(A)^+$ mRNA isolation. Five μg poly$(A)^+$ mRNA isolated with oligo (dT) cellulose spun columns (Pharmacia, Uppsala, Sweden) were used to synthesize the double stranded cDNA library (Pharmacia, Uppsala, Sweden). A total of 1.2 μg of double stranded cDNA was recovered, a portion of which was then ligated into the λ ZAP II arms (Stratagene, La Jolla, California, U.S.A.) and packaged *in vitro* into phage particles using the Gigapack system (Stratagene, La Jolla, California, U.S.A.).

To differentially screen the library, mRNAs from mature seeds and 6-day-old seedlings were used to synthesize single stranded cDNAs that were labeled with [α-^{32}P]dCTP according to Sambrook et al. (1989). A total of $1x10^4$ phage clones were transferred to nitrocellulose filters. Only the positive cDNA clones that hybridized exclusively with the single stranded [α-^{32}P]dCTP labeled cDNAs synthesized from mRNA isolated from 6-day-old seedlings were picked and excised from λ ZAP II as a recombinant pBluescript SK(-) plasmid (Stratagene, La Jolla, California, U.S.A.).

Northern blotting and hybridization

Total RNA was isolated from 1 g of whole seed or seedlings as described by Grimes et al. (1993). The RNA concentration was determined spectrophotometrically (A_{260}). The RNA was electrophoretically separated on agarose gels (20 μg total RNA per lane) containing formaldehyde (Sambrook et al., 1989), stained with ethidium bromide to assure equal loading, and transferred to a Zeta-Probe GT membrane (BIO-RAD, Mississauga, Ontario, Canada). Membranes were prehybridized for 5 minutes in prehybridization solution (7% SDS, 0.25 M Na_2HPO_4, pH 7.2) at 65 °C. The cDNA inserts were labeled by random priming with [α-^{32}P]dCTP and hybridized overnight in prehybridization solution at 65 °C. After hybridization, membranes were washed twice for 30 minutes in rinse solution #1 (40 mM Na_2HPO_4, 5% SDS, pH 7.2), twice for 30 minutes in rinse solution #2 (40 mM Na_2HPO_4, 1% SDS, pH 7.2), exposed to autoradiography film at -80 °C and developed.

DNA sequence analysis

The cDNA insert sequences were determined by the Taq Dye Primer Cycle Sequencing and the Taq DyeDeoxy Terminator Cycle Sequencing Kits using the Applied Biosystems Model 373A DNA Sequencing System (Applied Biosystems, Inc., Foster City, California, U.S.A.), and by the dideoxynucleotide chain termination method using Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio, U.S.A.). Primers used for sequencing were synthesized using the Applied Biosystems PCR-Mate DNA Synthesizer as described by the manufacturer (Applied Biosystems, Inc., Foster City, California, U.S.A.). DNA similarity searches were performed using the BLAST network service at NCBI (Altschul et al., 1990; Gish and States, 1993). The accession numbers for the sequence data deposited at the EMBL Nucleotide Database are as follows: *PM3*, cpn60ß, Z49766; *PM3-3*, CysP, Z49765; *PM4-5*, LMW HSP, Z49763; *PM5-1*, BiP, Z49764; *PM6-3*, typeII CAB, Z49749; *PM14*, CPR, Z49767).

Results

Isolation of germination associated cDNAs

To identify Douglas-fir genes that are developmentally regulated during germination and post-germination (Figure 2.1), a Douglas-fir cDNA library was constructed and differentially screened using cDNA probes synthesized with mRNA from mature seeds and 6-day-old seedlings. After a primary and secondary screening, 10 putative positive clones (*PM3*, *3-2*, *3-3*, *4*, *4-5*, *5-1*, *6-3*, *6-7*, *9*, *10*) were found to hybridize preferentially with the [α-^{32}P]dCTP labeled cDNA probes derived from 6-day-old seedlings and to a lesser degree with mature seed cDNA probes. Of these clones, nine had unique inserts based on their apparent molecular weight in agarose gels and by grouping together the clones that hybridized with each other during Southern blot analysis. Of these nine clones, six hybridized with transcripts that were present six days after germination (Table 2.1, Figures 2.2 and 2.3, clones *PM3*, *3-3*, *4-5*, *5-1*, *6-3*, and *6-7*). The clones *PM4*, *9* and *10* did not hybridize to transcripts during the germination or early seedling development in two separate northern blot analyses. They also had no sequence identity when compared to sequences in the data bases and no further work was done with these clones. Clone *PM14* was picked as a representative of a transcript present at similar amounts in both mature seeds and 6-day-old seedlings and was used for a comparison with the differentially regulated clones in gene expression studies.

Northern blot hybridization analysis during germination and early seedling development

To determine the expression pattern of the genes corresponding to the individual clones, total RNA was isolated from mature seeds, stratified seeds, 2-, 4-, 6-, 8-, and 14-day-old seedlings (Figure 2.1), and 10-week-old needles. Northern blot analysis demonstrated the

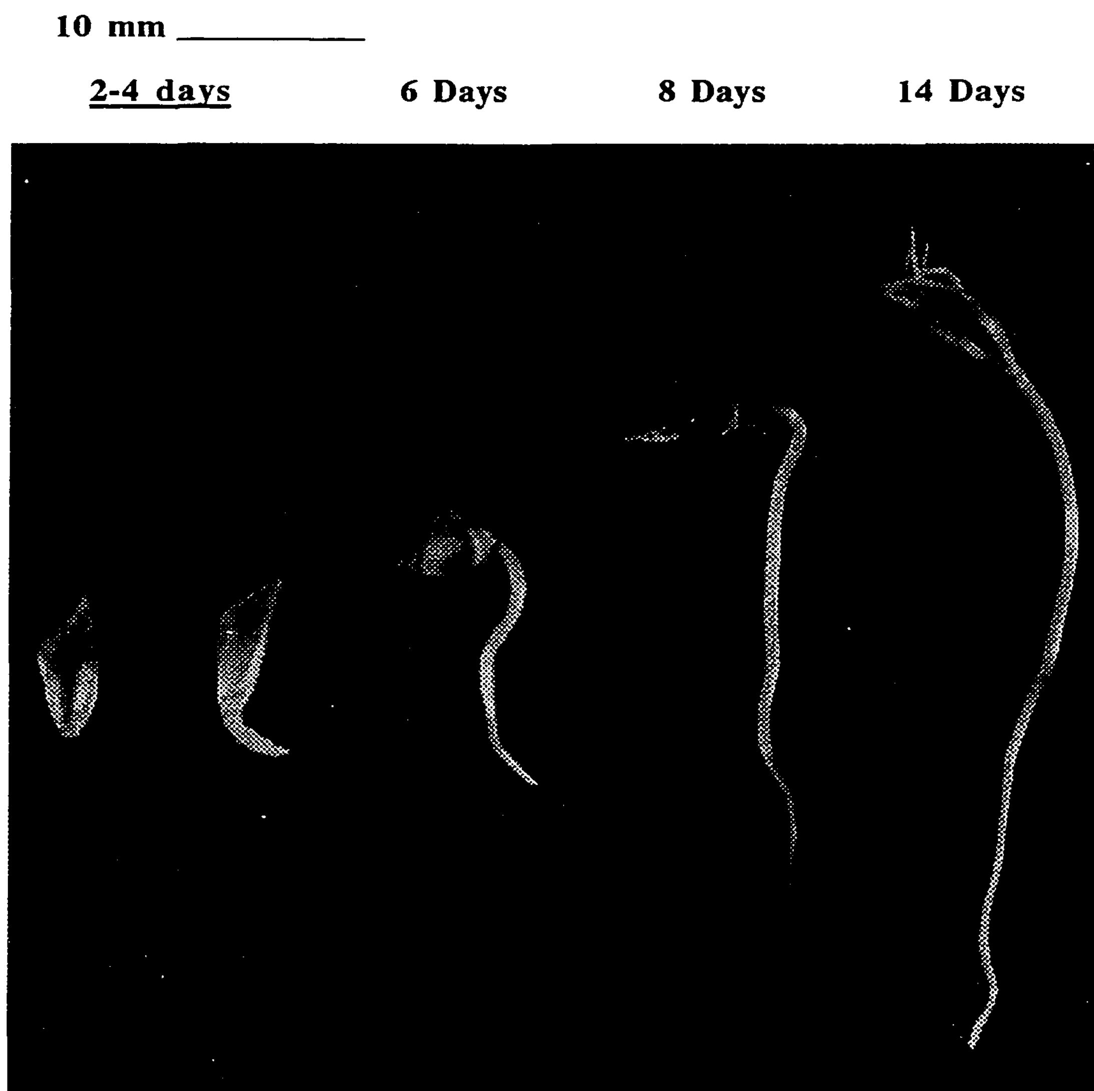


Figure 2.1 Time course of Douglas-fir seed germination and early seedling development. The germination phase of Douglas-fir occurs after an initial stratification treatment (3 weeks at 2-4 °C), lasts 2-4 days, and ends upon emergence of the radicle. During the following 5-14 days (the post-germination phase), the radicle increases in length, the hypocotyl emerges and grows, and the seed coat drops off.

Table 2.1. Douglas-fir cDNA clones-summary. The nucleotide (nt) and amino acid (aa) identities were determined from the alignments obtained from the BLAST network service at NCBI (Altschul et al., 1990; Gish and States, 1993). The insert sizes were estimated in agarose gels stained with ethidium bromide and the transcript sizes were estimated from northern blot analyses compared to molecular mass markers.

cDNA Clone	Insert Size (kb)	Transcript Size (kb)	Similarity	% nt Identity	% aa Identity	References
PM3	1.3	2.6	cpn60ß	79	94	Zabeleta et al. 1992
PM3-3	2.0	2.1	CysP	63	61	Koizumi et al. 1993
PM4-5	0.7	0.9	LMW HSP	54	72	Tseng et al. 1992
PM5-1	2.6	2.9	BiP	78	87	Denecke et al. 1991
PM6-3	0.9	1.2	typeII CAB	92	97	Kojima et al. 1992
PM6-7	1.6	1.4	?	?	?	?
PM14	1.1	3.0	CPR	76	78	Shet et al. 1993

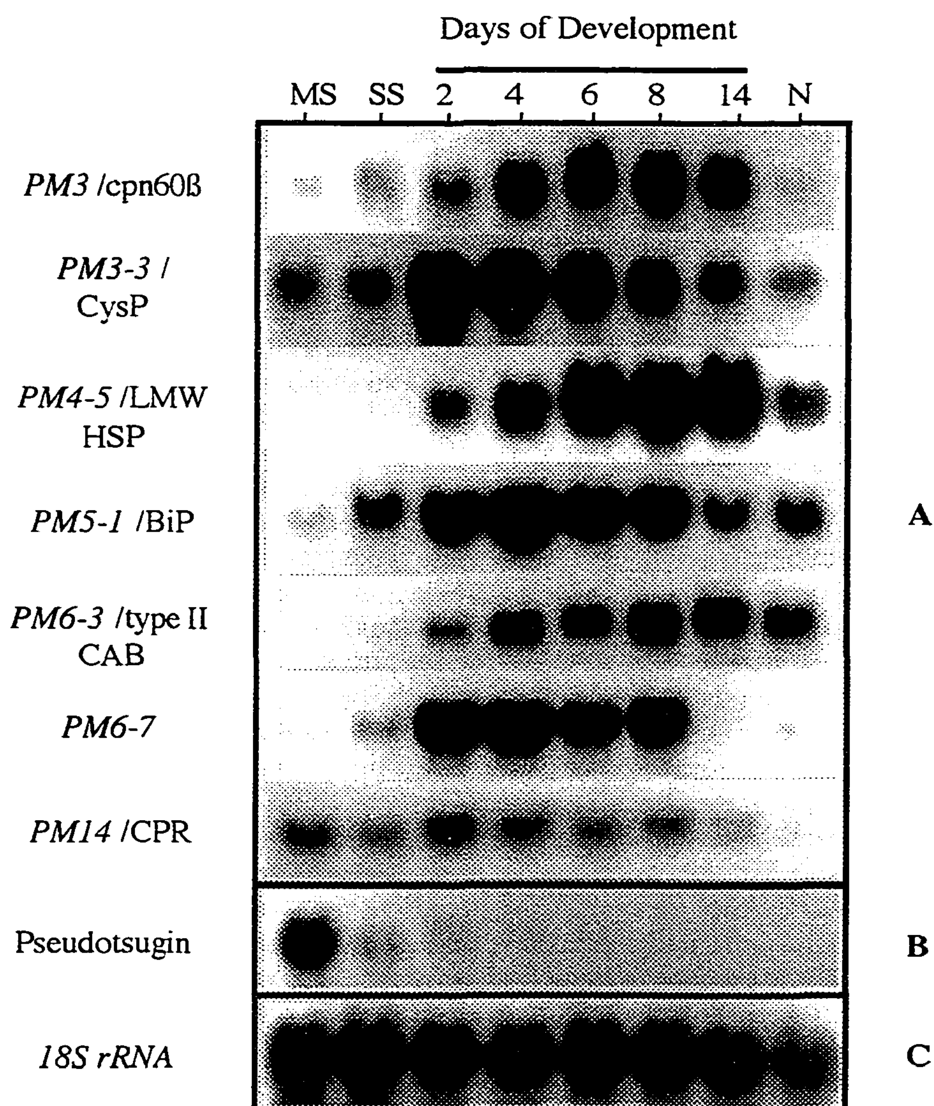


Figure 2.2 Northern blot analysis of transcript levels detected by cDNAs during Douglas-fir germination and post-germination. Twenty μg total RNA were loaded per lane extracted from whole mature seeds (MS), stratified seeds (SS), seedlings at 2, 4, 6, 8, and 14 days after the beginning of germination, and needles from 10-week-old seedlings (N). (A) Hybridization pattern of the six developmentally regulated cDNAs (*PM3*, *PM3-3*, *PM4-5*, *PM5-1*, *PM6-3*, *PM6-7*) isolated by differential screening and one constitutively expressed clone *PM14* isolated as described in the Materials and Methods section. Hybridization pattern detected with (B) a cDNA for pseudotsugin (a legumin-like clone) and (C) a Douglas-fir genomic clone for the 18S ribosomal RNA during the same time course.

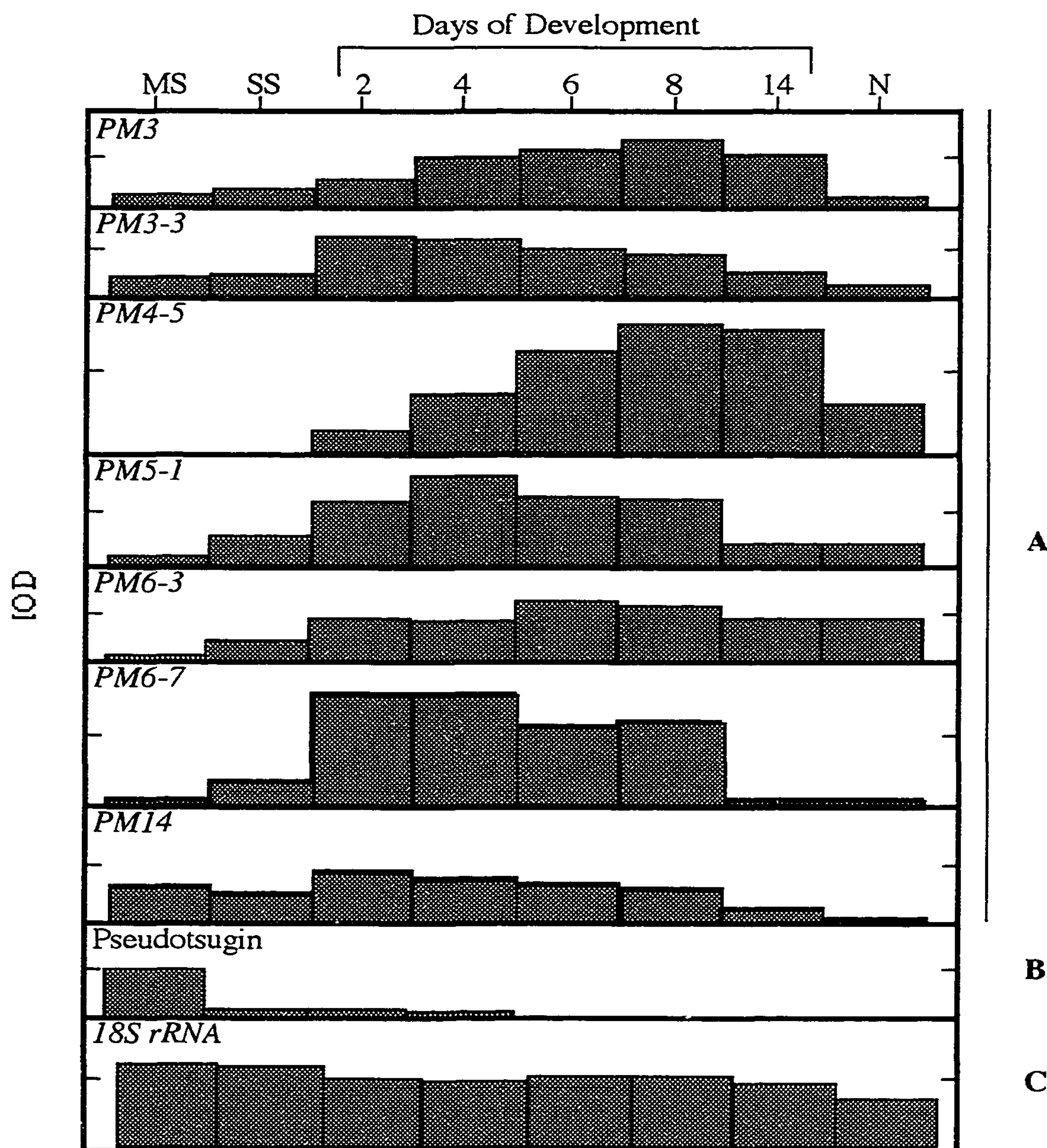


Figure 2.3. Densitometric measurements of the northern blot analysis from Figure 2.2. Mature seeds (MS); stratified seeds (SS); 10-week-old needles (N). Transcript signals from each hybridization were quantified using a Personal Densitometer (Molecular Dynamics). Variations in the total RNA loaded in each lane were corrected by dividing the lowest rRNA densitometric value by each rRNA value to obtain a ratio. The hybridization signal value for each clone was then multiplied by its respective rRNA ratio. The results are presented graphically using the arbitrary units of integrated optical density (IOD) (lower panel). Panels A, B and C are the same as in Figure 2.2.

cDNA probes *PM3*, *3-3*, *4-5*, *5-1*, *6-3* and *6-7*, hybridized with distinct transcripts that increased in abundance by 2 days and were relatively low in abundance in mature seeds (Figures 2.2 and 2.3).

The temporal pattern of increased expression of germination-associated clones was strikingly different to the pattern of decreasing transcript amounts detected by a cDNA that encodes pseudotsugin, a major storage protein in Douglas-fir (Figures 2.2 and 2.3; Leal and Misra, 1993b). The increase in transcript abundance of germination-associated clones was also in contrast to the relative steady-state amount of transcripts detected by clone *PM14* and with a probe made from a PCR-amplified 18S rDNA genomic clone of Douglas-fir.

In whole mature seeds, *PM3* and *3-3* had the highest transcript amounts, whereas *PM4-5*, *5-1*, *6-3* and *6-7* had the lowest transcript amounts relative to their respective peaks during the time course. The largest increases in transcript level were with clones *PM4-5*, *5-1* and *6-7*. Clones *PM5-1*, *6-3* and *6-7* transcripts increased slightly by the end of the stratification period. All the clones detected an increase in transcript quantities by 2 days after the beginning of germination, immediately before the emergence of the radicle (Figures 2.1, 2.2 and 2.3). All six clones isolated by differential screening therefore correspond to genes that exhibit increased expression during germination. The timing of the peak transcript amounts detected by the clones varied. *PM3-3* and *6-7* transcripts were highest at 2 to 4 days, *PM5-1* at 4 days, and *PM3*, *4-5*, and *6-3* at 6 to 8 days after the beginning of germination (Figures 2.2 and 2.3). The clones *PM3-3*, *5-1*, *6-7*, and *14* showed a decrease in the amount of transcript in 14-day-old seedlings and in 10-week-old needles. With the exception of *PM6-3*, transcript quantities of all the clones were reduced in 10-week-old needles relative to their peaks during germination.

Identification of germination clones through DNA sequence and similarity search analysis

The six germination associated clones and *PM14* were subjected to DNA sequence and similarity search analysis to determine their possible identities. DNA Primers were synthesized and used to confirm the nucleotide sequences of the cDNA clones in both directions (Table 2.2). Based on partial nucleotide sequence data (NCBI blastn) and predicted amino acid sequences (NCBI blastx), all clones except *PM6-7* were found to share identities with known plant genes (Table 2.1, Figures 2.4, 2.5, 2.6 and 2.7).

Clone *PM6-3* contained the complete coding region for a mature type II chlorophyll a/b binding protein (CAB) of photosystem II, along with a partial sequence that encodes the transit peptide needed for proper intracellular sorting to the chloroplast (Figure 2.4). The presumed CAB coding region and partial target sequence had the highest degree of similarity on both the nucleotide and the deduced amino acid sequence levels to the *cab-6* gene from *Pinus thunbergii* (Kojima et al., 1992) (Table 2.1 and Figure 2.4). The apparent mature CAB polypeptide encoded by the 687 bp region of the *PM6-3* sequence had 229 amino acid residues (approximate molecular weight of 25 kDa), which was the same number of residues for the type II CAB of *P. thunbergii* (Kojima et al., 1992). These results indicate that type II CAB coding sequences are conserved between the two genera *Pinus* and *Pseudotsuga*, and probably among gymnosperms. The predicted mature CAB sequence of *PM6-3* also showed high similarity to the CAB type II sequences from the dicots *Lycopersicon esculentum* (94.7%, Pichersky et al. 1987), *Gossypium hirsutum* (93.4%, Sagliocco et al. 1992), and *Pisum sativum* (93.0%, Falconet et al. 1991), and the monocot *Oryza sativa* (94.7%, Matsuoka 1990) indicating that CAB type II sequences are highly conserved among higher plants.

A 255 bp partial nucleotide sequence from clone *PM3* shared similarity with an internal sequence of the *Arabidopsis thaliana* 60-kDa chaperonin beta (cpn60ß) cDNA (Zabeleta et al., 1992) (Table 2.1, Figure 2.5). The predicted amino acid sequence from *PM3* showed

Table 2.2. Primers synthesized for sequencing of cDNA clones.

Name	Nucleotide Sequence (5'-3')	Clone Similarity
$3A_1$	TGGAGAGCGGAAAAGCCAATACT	cpn60ß
$3A_{1rev}$	AAGTATTGGCTTTTCCGCTCTCCA	cpn60ß
$3B_1$	CATCAATTATCAATAGCGGCTCCAA	cpn60ß
$3\text{-}3A_1$	GTGGCTGGCTCAGCATAAGAAAGC	CysP
$3\text{-}3B_1$	CTTTCCATGCATTGATTTAATTTCCTTC	CysP
$4\text{-}5A_1$	ACAGTCAGCACCCCGTTTTCCAT	LMW HSP
$4\text{-}5B_1$	GCTTCGGCCGTAGCTAATACTCAGAT	LMW HSP
$4\text{-}5A_{1rev}$	ATGGAAAACGGGGTGCTGACTGT	LMW HSP
$4\text{-}5B_{1rev}$	CTTTCCAGTCGATCTGAGTATTAGC	LMW HSP
$5\text{-}1A_1$	GAATCTCTTGTCCGTGGCCAATTTC	BiP
$5\text{-}1B_1$	TGATGCACAAAGACAGGCCACCA	BiP
$5\text{-}1B_{1rev}$	TGGTGGCCTGTCTTTGTGCATCA	BiP
$6\text{-}3A_1$	CAGCAGATGGGCGATGTTGGGA	typeII CAB
$6\text{-}3A_{1rev}$	TCCCAACATCGCCCATCTGCTG	typeII CAB
$6\text{-}3B_1$	CCTGGGTACAATGGGTCCAAC	typeII CAB
$6\text{-}3B_{1rev}$	GGTTGGACCCATTGTACCCAG	typeII CAB
$6\text{-}7A_1$	AGGTGAAAATAGAATTGGAGGAAGG	?
$6\text{-}7B_1$	TGGATCTTATGATGTTCCACCTG	?
$6\text{-}7A_2$	TTGTCCATTTCACAGAATAAATACAT	?
$6\text{-}7B_2$	GCTTAGAGTGTTGGCTTGACTG	?
$9A_1$	GTTCTATCAGCGCTACATTGATTCT	?
$10B_1$	CTGGTAGGTCCATCCAATTCTACAC	?
$14A_1$	GCATGATATTTAACTTTGCTGGTACT	CPR
$14B_1$	CTGGAGAGAAGCTTGGGCCAGC	CPR
$14B_{1rev}$	GCTGGCCCAAGCTTCTCTCCAG	CPR
$14B_{2rev}$	TGACATGTATTCTATTGGATGCATA	CPR

PM6-3 (CAB)

```
1    CGCAGGCTCGAATCACCATGCGAAGAACGGTCAGAAGCGCCCCTGAGAGCATTTGGTAT 59
       Q  A  R  I  T  M  R  R  T  V  R  S  A  P  E  S  I  W  Y
                   ^
60   GGACCTGATCGCCCCAAGTACCTGGGCCCCTTTTCGGAACAGACGCCGTCGTATCTCACC 119
      G  P  D  R  P  K  Y  L  G  P  F  S  E  Q  T  P  S  Y  L  T
                                             -
120  GGAGAATTTCCCGGGGATTACGGGTGGGACACCGCCGGACTCTCGGCCGACCCAGAGACC 179
      G  E  F  P  G  D  Y  G  W  D  T  A  G  L  S  A  D  P  E  T

180  TTCGCCAAAAATAGAGAGCTGGAGGTGATCCACAGCAGATGGGCGATGTTGGGAGCGCTC 239
      F  A  K  N  R  E  L  E  V  I  H  S  R  W  A  M  L  G  A  L
                                       -
240  GGCTGTGTTTTCCCGGAGCTGTTGGCCAAAAACGGGGTCAAATTTGGAGAAGCCGTGTGG 299
      G  C  V  F  P  E  L  L  A  K  N  G  V  K  F  G  E  A  V  W
            +
300  TTTAAGGCCGGGGCGCAGATATTCTCGGAGGGAGGCCTCGACTACCTGGGAAACCCCAAC 359
      F  K  A  G  A  Q  I  F  S  E  G  G  L  D  Y  L  G  N  P  N

360  CTGATCCACGCGCAGAGCATTCTAGCCATCTGGGCCTTCCAAGTTGTGCTCATGGGATTG 419
      L  I  H  A  Q  S  I  L  A  I  W  A  F  Q  V  V  L  M  G  L
                                          -
420  ATTGAAGGATACAGAGTGGGAGGAGGACCCCTTGGGGAGGGGTTGGACCCATTGTACCCA 479
      I  E  G  Y  R  V  G  G  G  P  L  G  E  G  L  D  P  L  Y  P

480  GGGGGTGCCTTCGACCCACTAGGGTTGGCCGACGACCCAGAGGCCTTCGCGGAGCTGAAG 539
      G  G  A  F  D  P  L  G  L  A  D  D  P  E  A  F  A  E  L  K

540  GTGAAGGAGCTAAAGAACGGGCGGCTGGCGATGTTCTCCATGTTCGGCTTCTTCGTGCAG 599
      V  K  E  L  K  N  G  R  L  A  M  F  S  M  F  G  F  F  V  Q
               +
600  GCAATCGTGACCGGCAAGGGCCCCATCGAAAACCTATTCGACCACTTGGCAGACCCCACT 659
      A  I  V  T  G  K  G  P  I  E  N  L  F  D  H  L  A  D  P  T
                                          +                    -
660  GCCAACAATGCCTGGGCCTATGCCACCAATTTCGTTCCCGGCAAGTGAagtgactaaaaa 719
      A  N  N  A  W  A  Y  A  T  N  F  V  P  G  K  *

720  taaagaggaaggctgtaatctgttatctgtgcattattaatttgaaaatggcc         772
```

Figure 2.4. Partial nucleotide and predicted amino acid sequences of clone *PM6-3*. The predicted amino acid sequence is compared to the CAB type II from *Pinus thunbergii* (Kojima et al., 1992). Arrow indicates the probable processing site of the precursor CAB polypeptide; overlined region indicates the partial putative transit sequence; stop codon (*); putative polyadenylation signal is underlined. Single letter codes in bold with a plus sign represent conserved amino acid replacements and bold with a minus sign represent unconserved replacements.

PM3 (cpn60ß)

```
1   AAAAGTGCAGAAAATAATCTGTATGTTGTAGAAGGAATGCAATTTGATCGTGGATATATT 60
     K  S  A  E  N  N  L  Y  V  V  E  G  M  Q  F  D  R  G  Y  I

61  TCTCCATACTTTGTGACAGACAGTGAGAAGATGTCTGTCGAGTATGAGAACTGCAAGCTG 120
     S  P  Y  F  V  T  D  S  E  K  M  S  V  E  Y  E  N  C  K  L
                                                +  +
121 TTATTGGTTGATAAGAAGATCACAAATGCAAGAGACATGATCACAGTATTGGAAGATGCC 180
     L  L  V  D  K  K  I  T  N  A  R  D  M  I  T  V  L  E  D  A
                                          +  +  -
181 ATCCGTGGTGGTTATCCCATTTTGATTATTGCGGAGGACATTGAACAAGAAGCCTTAGCT 240
     I  R  G  G  Y  P  I  L  I  I  A  E  D  I  E  Q  E  A  L  A

241 ACTCTTGTTGTGAAC                                              255
     T  L  V  V  N
```

Figure 2.5. Partial nucleotide and predicted amino acid sequences of clone *PM3*. A 255 bp partial nucleotide sequence of clone *PM3* and its predicted amino acid sequence shares identity with a cpn60ß from *Arabidopsis thaliana* (Zabeleta et al., 1992). Single letter codes in bold with a plus sign represent conservative amino acid residue replacements and bold with a minus sign represent unconserved replacements relative to the cpn60ß sequence from *A. thaliana* (Zabeleta et al., 1992).

a high percent identity to the sequence from the *A. thaliana* cpn60ß (Figure 2.5) and supports the earlier conclusion that cpn60ß genes are highly conserved among plants (Zabeleta et al., 1992). In addition, during the sequence similarity search analysis, *PM3* showed lower similarity to the cpn60α subunit (data not shown), which is consistent with the divergence found between cpn60α and cpn60ß in both *Brassica napus* and *A. thaliana* (Martel et al., 1990).

A partial 153 bp nucleotide sequence and predicted amino acid sequence from *PM3-3* showed the highest identity to a cysteine protease (CysP) that is induced by drought from *A. thaliana* (RD21; Koizumi et al., 1993) (Table 2.1, Figure 2.6). When the partial predicted amino acid sequence of *PM3-3* is compared to RD21, thirteen of the twenty amino acid replacements observed are conserved. The partial sequence also included several conserved catalytic residues found in CysP (Figure 2.6).

The partial predicted amino acid sequence of clone *PM4-5* was found to share identity to the carboxyl-terminal portion of a class I low molecular weight heat shock protein (LMW HSP) from *Oryza sativa* seedlings (Tseng et al., 1992) on both the nucleotide (138 bp) and amino acid levels (Table 2.1, Figure 2.6). The predicted amino acid sequence included the conserved Pro and Gly-Val-Leu sequence common to the LMW HSPs of plants and other eukaryotes (Lindquist and Craig, 1988).

The 255 bp nucleotide sequence of *PM5-1* aligned with an internal coding region of the BiP cDNA *(BLP4)* characterized from tobacco (Deneck et al., 1991) and is highly conserved on the amino acid level (Table 2.1, Figure 2.7).

Clone *PM14* shared the highest identity with a NADPH-cytochrome P450 reductase (CPR) (EC 1.6.2.4) from *Vigna radiata* (Shet et al., 1993) on both the nucleotide and the predicted amino acid level (Table 2.1, Figure 2.7). The predicted amino acid sequence from a 267 bp open reading frame shared identity with an internal sequence of the CPR containing a conserved region proposed as one of the NADPH binding sites. The

identification of clone *PM6-7* has not been determined.

A *PM3-3* (CysP)

```
1   GTGGGATATGGCTCAGAGTCTGGTATAGACTACTGGCTCGTGAAGAATTCTTGGGGAAAT 60
     V  G  Y  G  S  E  S  G  I  D  Y  W  L  V  K  N  S  W  G  N
                 +     +     -           +     +  ^  #  #     +
61  AGCTGGGGAGAGAAGGGTTTCATCAAGCTGCAGAGGAACCTTGAAGGGGCTTCGACGGGT 120
     S  W  G  E  K  G  F  I  K  L  Q  R  N  L  E  G  A  S  T  G
                 -     +  +  +  +  -        +  X  -  -     +
121 ATGTGCGGTATAGCAATGGAGGCATCGTACCCT                            153
     M  C  G  I  A  M  E  A  S  Y  P
```

B *PM4-5* (LMW HSP)

```
1   CTTCGGCGCTTCCGGCTGCCGGACAATGCGAAGGTTGAAGAGATCAAAGCGGCCATGGAA 60
     L  R  R  F  R  L  P  D  N  A  K  V  E  E  I  K  A  A  M  E
                          +     -     -     +           +
61  AACGGGGTGCTGACTGTGACAGTGCCGAAGCAGCCTGAACCGCAACCTCCTCAACCCAAA 120
     N  G  V  L  T  V  T  V  P  K  Q  P  E  P  Q  P  P  Q  P  K
                                   +  X        +  -     -  -
121 TCCATCGAGATCTCTGGT                                           138
     S  I  E  I  S  G
           +     +
```

Figure 2.6. Partial nucleotide and predicted amino acid sequences of clones *PM3-3* and *PM4-5*. (A) The predicted amino acid sequence from a 153 bp nucleotide sequence of *PM3-3* shares identity to a CysP from *Arabidopsis thaliana* (RD21; Koizumi et al., 1993). Overlined amino acids represent conserved catalytic site; (^), catalytic center; (#), catalytic amino acids. (B) A partial nucleotide sequence of *PM4-5* and predicted amino acid sequence which shares identity with a class I LMW HSP from *Oryza sativa* (Tseng et al., 1992). Conserved eukaryotic LMW HSP residues are overlined. Single letter codes in bold with a plus sign represent conserved amino acid replacements; bold with a minus sign represent unconserved replacements; (X), inserted amino acid.

A *PM5-1* (BiP)

```
1   TTTGATGTGAAACGGTTGATTGGAAGAAAGTATGAGGACAAGGAGGTGCAAAAAGACATC 60
     F  D  V  K  R  L  I  G  R  K  Y  E  D  K  E  V  Q  K  D  I
                                  +  +                 +     +
61  AAACTTTTGCCCTACAAAATTGTAAACAAAGATGGGAAGCCTTACATTCAGGTGAAGATC 120
     K  L  L  P  Y  K  I  V  N  K  D  G  K  P  Y  I  Q  V  K  I
           +
121 AGGGATGGTGAAATCAAAGTTTTTAGTCCCGAGGAAATTAGTGCAATGATTTTGTTGAAA 180
     R  D  G  E  I  K  V  F  S  P  E  E  I  S  A  M  I  L  L  K
     +           -     +                                   -
181 ATGAAGGAAACAGCTGAGTCCTACCTTGGAAGGAAAATCAAGGATGCAGTTGTTACAGTT 240
     M  K  E  T  A  E  S  Y  L  G  R  K  I  K  D  A  V  V  T  V
                       +           +
241 CCAGCATATTTCAAT                                              255
     P  A  Y  F  N
```

B *PM14* (CPR)

```
1   CCCAGGTATGCATCCAATAGAATACATGTCACATGTGCTCTGGTTTATGGGCCCAGTCCA 60
     P  R  Y  A  S  N  R  I  H  V  T  C  A  L  V  Y  G  P  S  P
           +        -     +                                +
61  ACTGGCAGAATTCACAAAGGTGTTTGTTCCAATTGGATGAAGAATTCAGTGCCTTCCGAG 120
     T  G  R  I  H  K  G  V  C  S  N  W  M  K  N  S  V  P  S  E
                                 -              +  +
121 AAAAGCCATGATTGTAGCTGGGCACCAGTCTTTGTTAGACAGTCAAATTTCAAATTGCCA 180
     K  S  H  D  C  S  W  A  P  V  F  V  R  Q  S  N  F  K  L  P
           -           -        +     +     -
181 TCGGATCCTTCAGTTCCTATTGTTATGGTGGGCCCTGGAACTGGTTTAGCACCTTTTAGA 240
     S  D  P  S  V  P  I  V  M  V  G  P  G  T  G  L  A  P  F  R
     -     -     +        +
241 GGCTTTTTGCAGGAAAGAGCTGCAATT                                  267
     G  F  L  Q  E  R  A  A  I
                       -     +
```

Figure 2.7. Partial nucleotide and predicted amino acid sequences of clones *PM5-1* and *PM14*. (A) The 5' end of clone *PM5-1* and its corresponding predicted amino acid sequence shares identity to the BiP (*BLP4*)from tobacco (Deneck et al., 1991). (B) Partial sequence from *PM14* shares identity with a NADPH-cytochrome P450 reductase from *Vigna radiata*(Shet et al., 1993). Conserved putative NADPH binding site overlined. Single letter codes in bold with a plus sign represent conserved amino acid replacements; bold with a minus sign represent unconserved replacements; (X), inserted amino acid.

Discussion

cDNAs for molecular chaperones (e.g. cpn60ß and BiP), a LMW HSP, a CysP, and a type II CAB were isolated from a Douglas-fir germination cDNA library. The expression of this set of cDNAs is developmentally regulated during germination. A cDNA for a CPR, an essential component of the P450 monooxygenase system, was also isolated that appears to be constitutively expressed in mature conifer seeds and seedlings.

A major metabolic event during seed germination is the production of enzymes responsible for the mobilization and utilization of storage reserves. The proteinases and peptidases that are responsible for the storage protein breakdown have been characterized at a biochemical level in several *Pinus sp.* (Salmia, 1981a and b; Gifford and Tolley, 1989; Gifford et al., 1989; and reviewed by Misra, 1994). However, there are no reports detailing the molecular characterization and regulation of these enzymes. A cDNA clone (*PM3-3*) that shares identity with a CysP from *A. thaliana* (Koizumi et al., 1993) was isolated. The transcripts of this proteinase *(PM3-3)* were detected in mature seeds and their amounts increased during germination, suggesting a role in the mobilization of storage proteins during germination and early seedling development.

Following germination, the expression of nuclear genes encoding components of the photosynthetic apparatus may increase to facilitate an autotrophic state. Douglas-fir cDNA *PM6-3* encoded a key component of photosynthesis, CAB, that binds chlorophyll a and b in the thylakoid membrane. Although CAB expression in gymnosperms differs from that in angiosperms in that it is light independent (Alosi et al., 1990; Yamamoto et al., 1991; Kojima et al., 1992), the increase in CAB transcript levels during early seedling development suggests that CAB expression in gymnosperms, as in angiosperms, is under developmental control (Simpson et al., 1986; Karlin-Neumann et al., 1988). Peak CAB transcript levels in Douglas-fir occurred during post-germination phase, which is similar to the CAB expression in Norway spruce (Sundås et al., 1992).

During germination and early seedling growth, there is a high degree of germination-specific protein synthesis, sorting and targeting to the various organelles of the cell. To fold into their native state, some newly synthesized proteins require assistance from other proteins referred to as molecular chaperones (Hemmingsen et al., 1988). In the present study, cDNAs that encode several molecular chaperones, e.g. cpn60ß subunit *(PM3)*, LMW HSP *(PM4-5)*, and BiP *(PM5-1)*, a member of the HSP 70 gene family have been isolated. In Douglas-fir the expression of their corresponding mRNAs increased during germination or post-germination. Clone *PM3* shares identity with the ß subunit of cpn60, the ribulose bisphosphate carboxylase (Rubisco) large subunit binding protein involved in holoenzyme formation (Barraclough and Ellis, 1980). The HSP 70 family members have also been characterized by their ability to bind polypeptides and their key role in the folding and stabilization of some proteins. BiP is a HSP 70 member that functions in the ER, is specifically involved in the secretory pathway and is not induced by heat shock. In angiosperms, BiP mRNA levels are high in tissues undergoing large numbers of cell divisions (Denecke et al., 1991). For example, in tobacco, BiP transcript levels were high in 4-day-old seedlings that were characterized by meristematic cells actively secreting high levels of cell-wall proteins (Denecke et al., 1991). A similar expression pattern was seen in the present study where BiP mRNAs in Douglas-fir were also highest in 4-day-old seedlings. BiP mRNA levels also increased during the stratification of Douglas-fir seeds. One explanation for high BiP expression is that it may be involved in the folding and secretion of proteins synthesized during stratification (Taylor et al., 1993) and at the onset of germination.

Based on partial nucleotide sequence analysis, the clone *PM4-5* is related to the most predominant HSPs found in a range of angiosperms referred to as LMW HSPs (Mansfield and Key, 1987). Recent reports have implicated mammalian (Jacob et al., 1993) as well as plant LMW HSPs (Vierling, 1995, Collada et al., 1997) as molecular chaperones *in vitro*.

The induction of LMW HSPs is not limited to heat shock (Vierling, 1991). In angiosperms, their expression is differentially regulated in developing somatic embryos (Györgyey et al., 1991) and in maturing seeds (Almoguera and Jordano, 1992). In Douglas-fir, the LMW HSP transcripts *(PM4-5)* were first detected in 2-day-old seedlings and their high level expression continued throughout the post-germination phase. Transcripts corresponding to the *PM4-5* clone were not detected in developing zygotic embryos of Douglas-fir (Kaukinen et al., 1996).

The clone *PM14* described here shares similarity with CPR, an integral membrane flavoprotein that is an essential component of microsomal cytochrome P450 monooxygenase system of eukaryotes (Lu et al., 1969). In castor bean, CPR activity in the ER and glyoxysomal membranes reached a maximum 2 days after germination and corresponded with increased activity of glyoxylate cycle enzymes (Alani et al., 1990). In Douglas-fir, CPR transcripts were detected in mature seeds and remained at a similar level throughout germination and post-germination.

In summary, this section describes the isolation of the first cDNAs that share identity with a cpn60ß, a CysP, a LMW HSP and a BiP from a conifer. In addition, a cDNA for CAB has also been isolated that has identity to a CAB previously identified from conifers and other plant species. The developmental expression of this set of cDNAs suggests their gene products may play critical roles during the process of germination and early seedling development. While the expression of all the cDNA clones were developmentally regulated during germination, distinct differences in the patterns of their expression suggest the presence of multiple regulatory mechanisms. The availability of homologous probes for these genes allows future experiments that will address the hormonal, environmental and developmental regulation of these genes in relation to their roles during germination and post-germination.

Conclusions

(1) Six developmentally regulated and one constitutively expressed cDNA clones were isolated.

(2) The cDNA clones were identified through BLAST sequence database search and share similarities with a cpn60ß, CysP, LMW HSP, BiP, Type II CAB, and CPR.

(3) Transcriptional and/or post-transcriptional changes occur during stratification, germination, and post-germination that increase the steady-state amounts of transcripts.

(3) The distinct expression patterns of each cDNA clone suggest multiple mechanisms of gene regulation function during stratification, germination and post-germination.

Literature Cited

Alani, A. A., Luster, D. G., and Donaldson, R. P. (1990) Development of endoplasmic and glyoxysomal membrane redox activities during castor bean germination. Plant Physiology 94: 1842-1848

Almoguera, C., and Jordano, J. (1992) Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and Lea mRNAs. Plant Molecular Biology 19: 781-792

Allen, G. S. (1941) A standard germination test for Douglas-fir seed. Forestry Chronicle 17: 75-78

Alosi, M. C., Neale, D. B., and Kinlaw, C. S. (1990) Expression of *cab* genes in Douglas-fir is not strongly regulated by light. Plant Physiology 93: 829-832

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403-410

Barraclough, R., and Ellis, J. R. (1980) Protein synthesis in chloroplasts. IX. Assembly of newly-synthesized large subunits into ribulose bisphosphate carboxylase in isolated intact pea chloroplasts. Biochimica et Biophysica Acta 608: 19-31

Bewley, J. D., and Marcus, A. (1990) Gene expression in seed development and germination. Progress in Nucleic Acids Research and Molecular Biology 38: 165-193

Ching, T. M. (1963) Fat utilization in germinating Douglas-fir seed. Plant Physiology 38: 722-728

Collada, C., Gomez, L., Casado, R., and Aragoncillo, C. (1997) Purification and *in vitro* chaperone activity of a class I small heat-shock protein abundant in recalcitrant chestnut seeds. Plant Physiology 115: 71-77

Comai, L., and Harada, J. J. (1990) Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. Proceedings of National Academy of Sciences USA 87: 2671-2674

Comai, L., Dietrich, R. A., Maslyar, D. J., Baden, C. S., and Harada, J. J. (1989) Coordinate expression of transcriptionally regulated isocitrate lyase and malate synthase genes in *Brassica napus* L. Plant Cell 1: 293-300

Delseny, M., Aspart, L., and Guitton, Y. (1977) Disappearance of stored polyadenylic acid mRNA during early germination of radish (*Raphanus sativus L.*) embryo axes. Planta 135: 125-128

Denecke, J., Goldman, M. H. S., Demolder, J., Seurinck, J., and Botterman, J. (1991) The tobacco luminal binding protein is encoded by a multigene family. Plant Cell 3: 1025-1035

Edwards, D. G. W. (1980) Maturity and quality of tree seeds - a state-of-the-art- review. Seed Science and Technology 8: 625-657

Edwards, D. G. W. (1986). Special pre-chilling techniques for tree seeds. Journal of Seed Technology 10: 151-171

El-Kassaby, Y. A., Edwards, D. G. W., and Taylor, D. W. (1992) Genetic control of germination parameters in Douglas-fir and its importance for domestication. Silvae Genetica 41: 48-54

Falconet, D. R., White, M. J., Fristensky, B. W., Dobres, M. S., and Thompson, W. F. (1991). Nucleotide sequence of Cab-215, a type II gene encoding a photosystem II chlorophyll a/b-binding protein in *Pisum*. Plant Molecular Biology 17: 135-139.

Gifford, D. J., and Tolley, M. C. (1989) The seed proteins of white spruce and their mobilization following germination. Physiologia Plantarum 77: 254-261

Gifford, D. J., Wenzel, K. A., and Lammer, D. L. (1989) Lodgepole pine seed germination. I. Changes in peptidase activity in the megagametophyte and embryonic axis. Canadian Journal of Botany 67: 2539-2543

Gish, W., and States, D. J. (1993) Identification of protein coding regions by database similarity search. Nature Genetics 3: 266-72

Green, M. J., Mcleod, J. K., and Misra, S. (1991) Characterization of Douglas-fir protein body composition by SDS-PAGE and electron microscopy. Plant Physiology and Biochemistry 29: 49-55

Grimes, H. D., Tranbarger, T. J., and Franceschi, V. R. (1993) Expression and accumulation patterns of nitrogen-responsive lipoxygenase in soybeans. Plant Physiology 103: 457-466

Györgyey, J., Gartner, A., Németh, K., Magyar, Z., Hirt, H., Heberle-Bors, E., and Dudits, D. (1991) Alfalfa heat shock genes are differentially expressed during somatic embryogenesis. Plant Molecular Biology 16: 999-1007

Harada, J. J., Baden, C. S., and Comai, L. (1988) Spatially regulated genes expressed during seed germination and postgerminative development are activated during embryogeny. Molecular and General Genetics 212: 466-473

Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., and Dennis, D. T. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature 333: 330-334

Jakob, U., Gaestel, M., Engel, K., and Buchner, J. (1993) Small heat shock proteins are molecular chaperones. Journal of Biological Chemistry 268: 1517-1520

Karlin-Neumann, G. A., Sun, L., and Tobin, E. M. (1988) Expression of light-harvesting chlorophyll a/b-protein genes is phytochrome-regulated in etiolated *Arabidopsis thaliana* seedlings. Plant Physiology 88: 1323-1331

Kaukinen, K. H., Tranbarger, T. J., and Misra, S. (1996) Post-germination-induced and hormonally dependent expression of low-molecular-weight heat shock protein genes of Douglas fir. Plant Molecular Biology 30: 1115-1128

Koehler, S. M., and Ho, T.-H. D. (1990) Hormonal regulation, processing, and secretion of cysteine proteinases in barley aleurone layers. Plant Cell 2: 769-783

Koizumi, M., Yamaguchi-Shinozaki, K., Tsuji, H., and Shinozaki, K. (1993) Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. Gene 129: 175-182

Kojima, K., Yamamoto, N., and Sasaki, S. (1992) Structure of the pine (*Pinus thunbergii*) chlorophyll a/b-binding protein gene expressed in the absence of light. Plant Molecular Biology 19: 405-410

Lane, B. G. (1991) Cellular desiccation and hydration: developmentally regulated proteins, and the maturation and germination of seed embryos. FASEB Journal 5: 2893-2901

Lindquist, S., and Craig, E. A. (1988) The heat shock proteins. Annual Review of Genetics 22: 631-677

Leal, I., and Misra, S. (1993) Developmental gene expression in conifer embryogenesis and germination. III. Analysis of crystalloid protein mRNAs and desiccation protein mRNAs in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss). Plant Science 88: 25-37

Leal, I., and Misra, S. (1993) Molecular cloning and characterization of a legumin-like storage protein cDNA of Douglas-fir seeds. Plant Molecular Biology 21: 709-715

Lee, G. J., Pokala, N., and Vierling, E. (1995) Structure and *in vitro* molecular chaperone activity of cytosolic small heat shock proteins from pea. Journal of Biological Chemistry 270: 10432-10438

Lu, A. Y. H., Junk, K. W., and Coon, M. J. (1969) Resolution of the cytochrome P-450-containing ω-hydroxylation system of liver microsomes into three components. Journal of Biological Chemistry 244: 3714-3721

Mansfield, M. A., and Key, J. L. (1987) Synthesis of the low molecular weight heat shock proteins in plants. Plant Physiology 84: 1007-1017

Masumori, M. Yamamoto, N., and Sasaki, S. (1992) Stored mRNA in pine seeds: Prolonged preservation in dry seeds and disappearance during germination. Plant and Cell Physiology 33: 485-488

Matsuoka, M. (1990) Classification and characterization of a cDNA that encodes the light-harvesting chlorophyll a/b binding protein of photosystem II from rice. Plant and Cell Physiology 31: 519-526

Martel, R., Cloney, L. P., Pelcher, L. E., and Hemmingsen, S. M. (1990) Unique composition of plastid chaperonin-60: α and ß polypeptide-encoding genes are highly divergent. Gene 94: 181-187

Misra, S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis, and seed germination: biochemical and molecular advances. Seed Science Research 4: 357-384

Misra, S., and Bewley, D. (1985) Reprogramming of protein synthesis from a developmental to a germinative mode induced by desiccation of the axes of *Phaseolus vulgaris*. Plant Physiology 78: 876-882

Misra, S., and Green, M. J. (1991) Developmental gene expression in conifer embryogenesis and germination. II. Crystalloid protein synthesis in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss.). Plant Science 68: 163-173

Mullen, R. T., and Gifford, D. J. (1997) Regulation of two loblolly pine (*Pinus taeda L.*) isocitrate lyase genes in megagametophytes of mature stratified seeds and during postgerminative growth. Plant Molecular Biology 33: 593-604

Owens, J. N., Morris, S. J., and Misra, S. (1993) The ultrastructural, histochemical and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii* (Mirb.) Franco. Canadian Journal of Forest Research 23: 816-827

Pichersky, E., Hoffman, N. E., Malic, V. S., and Bernatzky, R. (1987) The tomato Cab-4 and Cab-5 genes encode a second type of CAB polypeptides localized in photosystem II. Plant Molecular Biology 9: 109-120

Sagliocco, F., Kapazoglou, A., and Dure, L. III. (1992) Sequence of cab-151, a gene encoding a photosystem II type II chlorophyll a/b-binding protein in cotton. Plant Molecular Biology 18: 841-842

Salmia, M. A. (1981) Fractionation of the proteinases present in the endosperm of germinating seed of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 51: 253-258

Salmia, M. A. (1981) Proteinase activities in resting and germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 53: 38-47

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) In Molecular Cloning: A Laboratory Manual. 2nd Ed. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y. ISBN 0-87969-309-6

Shet, M. S., Sathasivan, K., Arlotto, M. A., and Mehdy, M. C. (1993) Purification, characterization, and cDNA cloning of an NADPH-cytochrome P450 reductase from mung bean. Proceedings of National Academy of Sciences USA 90: 2890-2894

Simpson, J., Montagu, M. V., and Herrera-Estrella, L. (1986) Photosynthesis-associated gene families: differences in response to tissue-specific and environmental factors. Science 233: 34-38

Sundås, A., Tandre, K., Holmstedt, E., and Engström, P. (1992) Differential gene expression during germination and after the induction of adventitious bud formation in Norway spruce embryos. Plant Molecular Biology 18: 713-724

Taylor, M. A., Davies, H. V., Smith, S. B., Abruzzese, A., and Gosling, P. G. (1993) Cold-induced changes in gene expression during dormancy breakage in seeds of Douglas-fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142: 120-123

Taylor, M. A., and Davies, H. V. (1995) Changes in gene expression during imbibition of seeds of Douglas fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142: 54-58

Tseng, T. S., Yeh, K.W., Yeh, C. H., Chang, F. C., Chen,Y. M., and Lin, C. Y. (1992) Two rice (*Oryza sativa*) full-length cDNA clones encoding low molecular weight heat shock proteins. Plant Molecular Biology 18: 963-965

Yamamoto, N., Mukai, Y., Matsuoka, M., Kana-Murakami, Y., Tanaka, Y., Ohashi, Y., Ozeki, Y., and Odani, K. (1991) Light-independent expression of cab and rbcS genes in dark-grown pine seedlings. Plant Physiology 95: 379-383

Verwoerd, T. C., Dekker, M. M., and Hoekema, A. (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Research. 17: 2362

Vierling, E. (1991) The roles of heat shock proteins in plants. Annual Review of Plant Physiology and Molecular Biology 42: 579-620

Washio, K., and Ishikawa, K. (1992) Structure and expression during the germination of rice seeds of the gene for a carboxypeptidase. Plant Molecular Biology 19: 631-640

Weir, E. M., Riezman, H., Grieneberger, J-M., Becher, W. M., and Leaver. C. J. (1980) Regulation of glyoxysomal enzymes during germination of cucumber. European Journal of Biochemistry 112: 469-477

Whitmore, F. W. (1991) Stored messenger RNA and stratification in eastern white pine (*Pinus strobus* L.). Seed Science and Technology 19: 341-346

Zabeleta, E., Oropeza, A., Jiménez, B., Salerno, G., Crespi, M., and Herrera-Estrella, L. (1992) Isolation and characterization of genes encoding chaperonin 60ß from *Arabidopsis thaliana*. Gene 111: 175-181

CHAPTER 3

Structure and Expression of a Developmentally Regulated cDNA Encoding a Cysteine Proteinase (Pseudotzain) from Douglas-fir

Abstract

The complete sequence and expression of a cDNA clone *(PM3-3)* encoding a cysteine proteinase (CysP) from *Pseudotsuga menziesii* [Mirb] is reported. The sequence consists of a 5' untranslated region *(UTR)* of 153 bp followed by an open reading frame (ORF) of 1362 bp encoding a putative mature CysP flanked by N- and C-terminal propeptides. A 364 bp 3' *UTR* contains multiple AU-rich elements *(ARE)* that may be involved in the destabilization of transcripts. Pseudotzain contains the same invariant amino acid residues that are involved in the catalytic reaction and make up the catalytic center of CysP from plants and animals. Northern blot analysis showed that *cysP* transcripts were most abundant in the megagametophyte (MG) after germination and not detected in the MG or embryo during embryogenesis. Osmotic stress treatments slightly enhanced *cysP* steady-state transcript amounts during early seedling development, whereas abscisic acid (ABA), gibberellin (GA) and other plant growth regulators and environmental conditions had little or no effect. The *cysP* transcripts were present in different amounts in the cotyledons, root and seed coat of 10-day-old seedlings, but were most abundant in the MG, suggesting a role for this proteinase in storage protein mobilization. Phylogenetic analysis of mature CysP groups pseudotzain with other angiosperm CysP having both N- and C-terminal propeptides, suggesting a conserved function and/or targeting of this subgroup of enzymes.

Introduction

The main storage tissue in conifer seeds is the haploid, maternally derived MG. In contrast to angiosperm storage tissues that originate and develop after fertilization, the MG of Douglas-fir and other conifers is fully developed before fertilization (Allen and Owens, 1972). Storage protein accumulation in the MG and the embryo starts during early and mid embryogenesis, respectively, and terminates during the desiccation phase of seed maturation (Owens et al., 1993; Misra, 1994). During germination, proteins are mobilized from the MG to provide free amino acids to the embryo (Lammer and Gifford, 1989). In angiosperms, CysP (EC 3.4.22) are thought to catalyze the initial stages of storage protein mobilization during seed germination (reviewed by Shutov and Vaintraub, 1987). CysP activity is also associated with protein mobilization during conifer seed germination (Salmia, 1981 a and b), but little is known about proteinase expression during conifer embryogenesis or germination. Both mono- and dicotyledonous plants express multiple *cysP* genes that are differentially regulated during germination; however, the mechanisms controlling their distinct expression are not fully understood (Dietrich et al., 1989; Kalinski et al., 1990; Koehler and Ho, 1990; Watanabe et al., 1991; Kalinski et al., 1992; Becker et al., 1994). In germinating rice seeds (Watanabe et al., 1991) and in barley aleurone, *cysP* gene expression is induced by gibberellin (GA) but inhibited by abscisic acid (ABA, Rogers et al., 1985; Koehler and Ho, 1990), whereas chick pea *cysP* is induced by ethylene (Cervantes et al., 1994). *CysP* gene expression is not restricted to induction by hormones. In *Arabidopsis* seedlings and in pea shoots, *cysP* is induced by osmotic stress independently of ABA (Guerroro et al., 1990; Koizumi et al., 1993; Williams et al., 1994), in response to cold in tomato fruit and *Arabidopsis* plants (Schaffer and Fischer, 1988; Williams et al., 1994), and during senescence in pea ovary tissues (Granell et al., 1992).

This section describes the complete sequence of the cDNA clone *PM3-3* encoding a putative CysP (pseudotzain), its evolutionary relationship to angiosperm and animal CysP, and analysis of the factors regulating its expression during embryogenesis and germination.

Materials and Methods

Plant material and growth conditions

Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco) seeds were stratified and germinated as described previously (Chapter 2, Material and Methods; Tranbarger and Misra, 1995). MGs and embryos were dissected from developing cones collected (from Pacific Forest Products Limited, Saanichton, B. C.) May 31 (M), June 19 (J), July 17 (Jy), and August 7 (A). Mature seeds (seed lot # 952 from the Ministry of Forest Tree Seed Center, Surrey, B.C.) were taken from 4 °C storage and dissected or germinated. Stratified seeds were germinated for 4 (4D) and 10 days (10D) on Whatman 3MM paper in trays with distilled water in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hour photoperiod. Seed coats were removed at all time points. For hormonal and environmental effects on *cysP* expression, 14-day-old seedlings were treated for 48 hours on Whatman 3MM paper in 150 x 15 mm Petri dishes (50-60 seedlings per dish) immersed in the following solutions: PEG, 24% polyethylene glycol; NaCl, 250 mM sodium chloride; dry, seedlings were desiccated by leaving them on Whatman 3MM paper in the absence of solution; ABA, 10^{-4} M abscisic acid; GA, 10^{-4} M gibberellic acid (GA_3); IAA, 10^{-4} M indoleacetic acid; Kinetin, 10^{-4} M kinetin; SA, 10 mM salicylic acid. Control plants were incubated in water. Wounding was done by pinching stratified seeds with forceps and allowing seeds to incubate in water for 48 hours under germination conditions described in the Materials and Methods section of Chapter 2. HS, heat shock treatment was at 40 °C for 3 hours. HS control was incubated for 3 hours at 24 °C. CS, cold shock treatment was at 4 °C for 16 hours. CS control was incubated at 24 °C for 14 hours.

Seedlings subjected to HS and CS treatments and controls were performed in water in the dark.

Northern blotting and hybridization

Total RNA was isolated according to Verwoerd et al. (1989), with modifications described previously (Leal and Misra, 1993). The RNA concentration was determined spectrophotometrically (A_{260}), electrophoretically separated on 1% agarose gels with 2.2 M formaldehyde (Sambrook et al., 1989), and transferred to a Zeta-Probe GT membrane (BIO-RAD, Mississauga, Ontario, Canada). Both probes were labeled by random priming with [α-^{32}P]dCTP (BIO-RAD, Mississauga, Ontario, Canada). See the Materials and Methods section in Chapter 2 for the northern blot analysis procedure.

PM3-3 cysP cDNA clone isolation and DNA sequence analysis

The *PM3-3* cDNA clone was isolated by differentially screening a germination cDNA library from Douglas-fir as previously described (Chapter 2, Material and Methods; Tranbarger and Misra, 1995). The cDNA clone *PM3-3* was contained in the *Eco*RI site of a pBluescript SK(-) plasmid (Stratagene, La Jolla, California, USA). The nucleotide sequence was determined by the Taq Dye Primer Cycle Sequencing and the Taq DyeDeoxy Terminator Cycle Sequencing Kits using the Applied Biosystems Model 373A DNA Sequencing System (Applied Biosystems, Inc., Foster City, California, U.S.A.), and by the dideoxynucleotide chain termination method using Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio, U.S.A.). Primers used for sequencing were synthesized using the Applied Biosystems PCR-Mate DNA Synthesizer as described by the manufacturer (Applied Biosystems, Inc., Foster City, California, U.S.A.). The nucleotide

sequence reported here appeared in the GenBank database under accession number U41902.

Amino acid sequence alignments and phylogenetic analysis

The phylogenetic tree was constructed using the PHYLIP package (Felsenstein, 1989) based on mature CysP amino acid sequences aligned using CLUSTAL V (Higgins and Sharp, 1989).

Results

Cloning and sequence analysis of a cysteine proteinase cDNA from Douglas-fir

The nucleotide sequence of the CysP cDNA clone was analyzed and confirmed in both directions using synthetic DNA primers designed from *PM3-3* (Table 3.1). Clone *PM3-3* contains an insert of 1882 bp with an ORF of 1362 bp encoding a predicted polypeptide comprised of 454 amino acids (Figure 3.1 and 3.2). A search of the databases using BLAST (Altshul et al., 1990; Gish and States, 1993) confirmed that *PM3-3* encodes a polypeptide with similarity to plant and animal CysP (Figures. 3.1, 3.2 and 3.5). The insert is short of being a full length clone because the size of the transcript is estimated to be approximately 2.1 kb (Tranbarger and Misra, 1995). The nucleotide sequence includes a 153 bp 5' *UTR* based on the use of the first start codon in the ORF (Figures 3.1 and 3.2). Three possible start codons are present within the first 120 nucleotide of the ORF at positions 154, 196, and 274. Although the third start codon (nucleotide 274) is most similar to both the monocot (GCAMCCatgG) and dicot (AAAAAMatgG) consensus sequences (M=A/C; R=A/G; Cavener and Ray, 1991), this start position would eliminate the hydrophobic N-terminal signal sequence and significantly alter the alignment with other putative CysP (Figure 3.5). The second possible start codon (nucleotide 196) would result in a truncated N-terminal signal sequence and would also affect optimal amino acid alignment. Therefore, the start codon at nucleotide 154 is the most probable for translation initiation. Clone *PM3-3* has a short uORF from nucleotide 31 to 57 (Figure 3.1). The 364 bp 3' *UTR* contains several putative *ARE* that may have a role in determining mRNA stability. These include the sequence TTATTTTTT which is one of the imperfect nonamer motifs recently found to mediate mRNA destabilization (Zubiaga et al., 1995), and two TATTTAT sequence motifs that are involved in the destabilization of transcripts (Lagnado

Table 3.1. Primers synthesized for sequencing the *PM3-3/cysP* cDNA clone. Primers were synthesized as described in the Materials and Methods section.

Name	Nucleotide Sequence (5'-3')	Position and Direction Within Nucleotide Sequence
3-3A$_1$	GTGGCTGGCTCAGCATAAGAAAGC	(291->314)
3-3A$_{1rev}$	GTTGTTGTGCTGGTGAATATACAG	(393<-370)
3-3A$_2$	GATTTGCCAGAGTCCATTGACT	(544->565)
3-3A$_{2rev}$	GTGACGATTTGATTGATGCCT	(671<-651)
3-3A$_3$	TGACTACGAAGATGTGCCTGAG	(876->897)
3-3B$_3$	CAAGCTGAGTTCCGCAGTTAC	(1024<-1004)
3-3B$_2$	TCATACATGCAGCAGCAGGTG	(1316<-1296)
3-3B$_{2rev}$	ACGGGTATGTGCGGTATAGC	(1156->1175)
3-3B$_{1rev}$	CCATTTGGAACGAAGATGCTG	(1456->1476)
3-3B$_1$	CTTTCCATGCATTGATTTAATTTCCTTC	(1630<-1603)
3-3A$_4$	GCATTCCTGCACATATGGCT	(1654->1673)

Figure 3.1. The complete nucleotide sequence of *PM3-3* and its deduced amino acid sequence. The amino acid sequence of the coding region is shown in single letter code below the nucleotide sequence. The ORF is capitalized and the *UTR* are in lower case. The double underlined sequence in the 5' *UTR* indicates the uORF. The three possible start codons and their surrounding consensus regions are underlined and italicized. The potential *N*-glycosylation sites are underlined and in bold. The putative *ARE* elements in the 3' *UTR* are underlined. The potential hydrophobic leader sequence is bracketed. The probable N- and C-terminal processing sites of the mature CysP are indicated by (^). (O), amino acid that constitutes the catalytic center; (@), amino acid involved in the catalytic reaction; (=), cysteine residues which form disulfide bridges (all symbols are under the corresponding amino acid).

Figure 3.1.

```
ccctttcgtgggtcattgccttcttggcccatgcatgtgggtttgaatcttttatgagat   60
tggtttggttttcagcgaggatttcgcccaggccagtggagaagaattccctctgcgagc  120

ctgattcaataaaacaaataagattttatcagtATGGGAATCCTTTTGCTGTTTGCTGTG  180
                                [M  G  I  L  L  L  F  A  V      9

CTTGCATTGTCTGCGATGGCAGGCAGTGCTTCAAGAGCAGATTTCTCCATCATCAGCTAC  240
 L  A  L  S  A  M  A  G  S  A  S] R  A  D  F  S  I  I  S  Y    29

GATAGCCAAGATCTGATAGGAGATGACGCGATCATGGAGCTCTATGAACTGTGGCTGGCT  300
 D  S  Q  D  L  I  G  D  D  A  I  M  E  L  Y  E  L  W  L  A    49

CAGCATAAGAAAGCCTACAATGGTCTTGACGAGAAGCAGAAGAAGTTCTCTGTATTTAAA  360
 Q  H  K  K  A  Y  N  G  L  D  E  K  Q  K  K  F  S  V  F  K    69

GACAATTTTCTGTATATTCACCAGCACAACAACCAGGGGAATCCATCCTACAAATTGGGG  420
 D  N  F  L  Y  I  H  Q  H  N  N  Q  G  N  P  S  Y  K  L  G    89

CTGAACCAGTTTGCAGATCTGAGCCATGAGGAGTTCAAGGCTGCATATCTGGGTACCAAG  480
 L  N  Q  F  A  D  L  S  H  E  E  F  K  A  A  Y  L  G  T  K   109

CTGGATGCCAAGAAACGCTTGTCCAGGTCTCCTAGCCCTCGATACCAGTATTCGGTGGGC  540
 L  D  A  K  K  R  L  S  R  S  P  S  P  R  Y  Q  Y  S  V  G   129

GAGGATTTGCCAGAGTCCATTGACTGGAGGGAAAAGGGAGCCGTGACTGCCGTTAAGAAC  600
 E  D  L  P  E  S  I  D  W  R  E  K  G  A  V  T  A  V  K  N   149
       ^
CAGGGCTCCTGCGGAAGTTGTTGGGCATTCTCGACTGTGGCAGCTGTTGAAGGCATCAAT  660
 Q  G  S  C  G  S  C  W  A  F  S  T  V  A  A  V  E  G  I  N   169
 o        =        @
CAAATCGTCACCGGCAATTTGACTTCGCTGTCCGAGCAGGAATTGGTGGACTGTGATACT  720
 Q  I  V  T  G  N  L  T  S  L  S  E  Q  E  L  V  D  C  D  T   189
                                                    =
TCTTACAACCAAGGATGCAATGGCGGTCTCATGGATTATGCTTTCCAGTTTATCATAAGC  780
 S  Y  N  Q  G  C  N  G  G  L  M  D  Y  A  F  Q  F  I  I  S   209
                =
AACGGTGGGCTTGACAGCGAGGATGATTACCCTTACAAGGCCAACAATGGCAGCTGTGAC  840
 N  G  G  L  D  S  E  D  D  Y  P  Y  K  A  N  N  G  S  C  D   229
                                                       =
GCTTACAGAAAAAATGCCCATGTGGTGACAATCGATGACTACGAAGATGTGCCTGAGAAC  900
 A  Y  R  K  N  A  H  V  V  T  I  D  D  Y  E  D  V  P  E  N   249

GATGAGAAGTCGCTGAAGAAGGCCGCGGCGAATCAGCCAATTAGCGTGGCCATCGAAGCC  960
 D  E  K  S  L  K  K  A  A  A  N  Q  P  I  S  V  A  I  E  A   269

AGCGGAAGGGCGTTCCAGTTTTACGAATCTGGCGTGTTTACTAGTAACTGCGGAACTCAG 1020
 S  G  R  A  F  Q  F  Y  E  S  G  V  F  T  S  N  C  G  T  Q   289
                                                 =
CTTGACCACGGTGTGACTCTGGTGGGATATGGCTCAGAGTCTGGTATAGACTACTGGCTC 1080
 L  D  H  G  V  T  L  V  G  Y  G  S  E  S  G  I  D  Y  W  L   309
    o  @
GTGAAGAATTCTTGGGGAAATAGCTGGGGAGAGAAGGGTTTCATCAAGCTGCAGAGGAAC 1140
 V  K  N  S  W  G  N  S  W  G  E  K  G  F  I  K  L  Q  R  N   329
       @  o  @
CTTGAAGGGGCTTCGACGGGTATGTGCGGTATAGCAATGGAGGCATCGTACCCTGTAAAG 1200
 L  E  G  A  S  T  G  M  C  G  I  A  M  E  A  S  Y  P  V  K   349
                         =
AAGGGTGCCAACCCTCCAAACCCTGGGCCTTCACCCCCATCCCCTGTGAAGCCCCCTACC 1260
 K  G  A  N  P  P  N  P  G  P  S  P  P  S  P  V  K  P  P  T   369
       ^
GTGTGTGATAACTACTATTCCTGCCCGGAGAGCAACACCTGCTGCTGCATGTATGACTTT 1320
 V  C  D  N  Y  Y  S  C  P  E  S  N  T  C  C  C  M  Y  D  F   389

GGGGGTTACTGCTATGCATGGGGCTGCTGTCCTCTCAACTCTGCCACCTGCTGCGACGAT 1380
 G  G  Y  C  Y  A  W  G  C  C  P  L  N  S  A  T  C  C  D  D   409

CACTACAGCTGCTGCCCTAGCGATCATCCTGTTTGCGATCTCGATGCACAGACATGCCTC 1440
 H  Y  S  C  C  P  S  D  H  P  V  C  D  L  D  A  Q  T  C  L   429

AAGAGTCGCAAGGATCCATTTGGAACGAAGATGCTGAAGCGTACCCCTGCCAAACCTTAC 1500
 K  S  R  K  D  P  F  G  T  K  M  L  K  R  T  P  A  K  P  Y   449

TGGTCCCTATCTGGTtaggaaacgataactgaaaggacccagtatagctgatttctcttt 1560
 W  S  L  S  G  *                                             454
gcaatccttattttttaactcaacaacagactggtccgaatggaaggaaattaaatcaat 1620
gcatggaaagttcttaaacttatatatcgattagcattcctgcacatatggctcctcctc 1680
caaatcggcccagatgcttcccctcagtgttctcgtcaattctatttctatttctattta 1740
tgtttctctcgtcaattctatttctatttatgtgtaatgtacagatcttgctatgtatga 1800
ccattattaaagctgaagcctgtctctgaacagaacagtgtatgattctccgatctatta 1860
tcaatttaaaatccaggtttca                                       1882
```

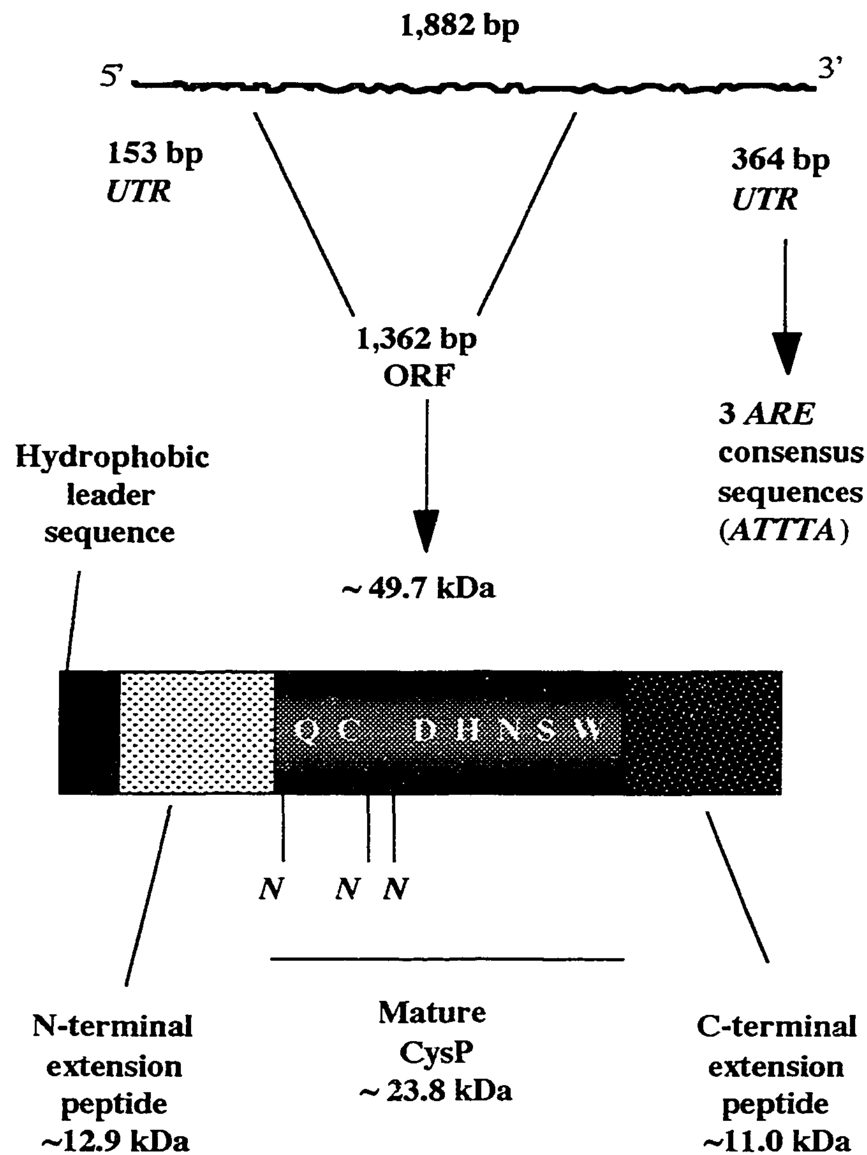


Figure 3.2. Features of the *PM3-3/cysP* cDNA clone and predicted amino acid sequence. The conserved catalytic amino acids within the mature CysP are indicated in single letter code. The relative positions of the putative *N*-glycosylation sites on the mature CysP are indicated by (*N*). The predicted molecular masses are in kilodaltons (kDa) and the nucleotide sequence lengths are in base pairs (bp).

et al., 1994). In addition, the 3' *UTR* contains three ATTTA motifs that have been correlated with sequence destabilization in mammals and plants (Shaw and Kamen, 1986; Ohme-Takagi et al., 1993; Abler and Green, 1996). The deduced amino acid sequence of *PM3-3* contains three potential *N*-glycosylation sites (Asn^{83}, Asn^{175} and Asn^{225}), the latter two are in comparable positions to those in aleurain (Rogers et al., 1985).

The cysteine proteinase amino acid sequence: alignment and phylogenetic analysis

Based on optimal amino acid alignment with other angiosperm CysP (e.g. papain and actinidin, Kamphuis et al., 1985; aleurain, Rogers et al., 1985; oryzain A, B and C, Watanabe et al., 1991) the name pseudotzain (derived from *Pseudotsuga menziesii*) for the CysP from Douglas-fir is proposed. CysP are commonly targeted to the secretory system and encoded as zymogens of various configurations (Figure 3.3). Pseudotzain contains N- and C-terminal propeptides in addition to features common to other CysP described below (Figures 3.1, 3.2, 3.3, 3.4, and 3.5). The N- and C-terminal ends of the mature form of pseudotzain are predicted to be Leu^{132} and Ala^{352}, respectively (Figure 3.1). In addition, the sequence shows in proper positions, the conserved amino acid residues constituting the catalytic triad (Cys^{156}, His^{292}, Asn^{312}), three conserved pairs of cysteine residues involved in disulfide bridge formation (Cys^{156}-Cys^{195}, Cys^{187}-Cys^{228}, Cys^{187}-Cys^{338}), and other amino acid residues (Gln^{150}, Ser^{313}, Trp^{314}) that are involved in the catalytic activity of CysP (Kamphuis et al., 1985; Baker and Drenth, 1987). Mature pseudotzain shares from 37 to 71% sequence identity with protozoan, animal and plant CysP but shares the highest identity and forms a phylogenetic subgroup with the angiosperm CysP that have both N- and C-terminal propeptides (Figures 3.3, 3.4 and 3.5). The hydrophobic vacuolar targeting motif (Asn-Xaa-Ile-Xaa-Xaa) found in several N-terminal propeptides including

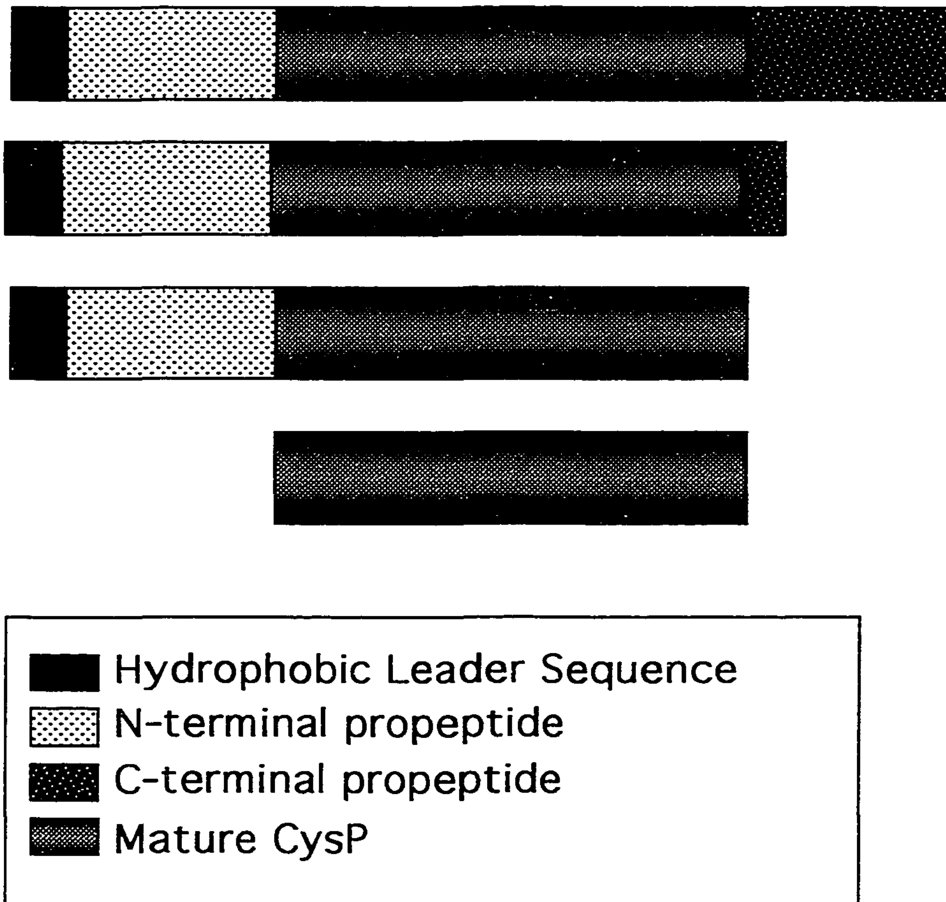
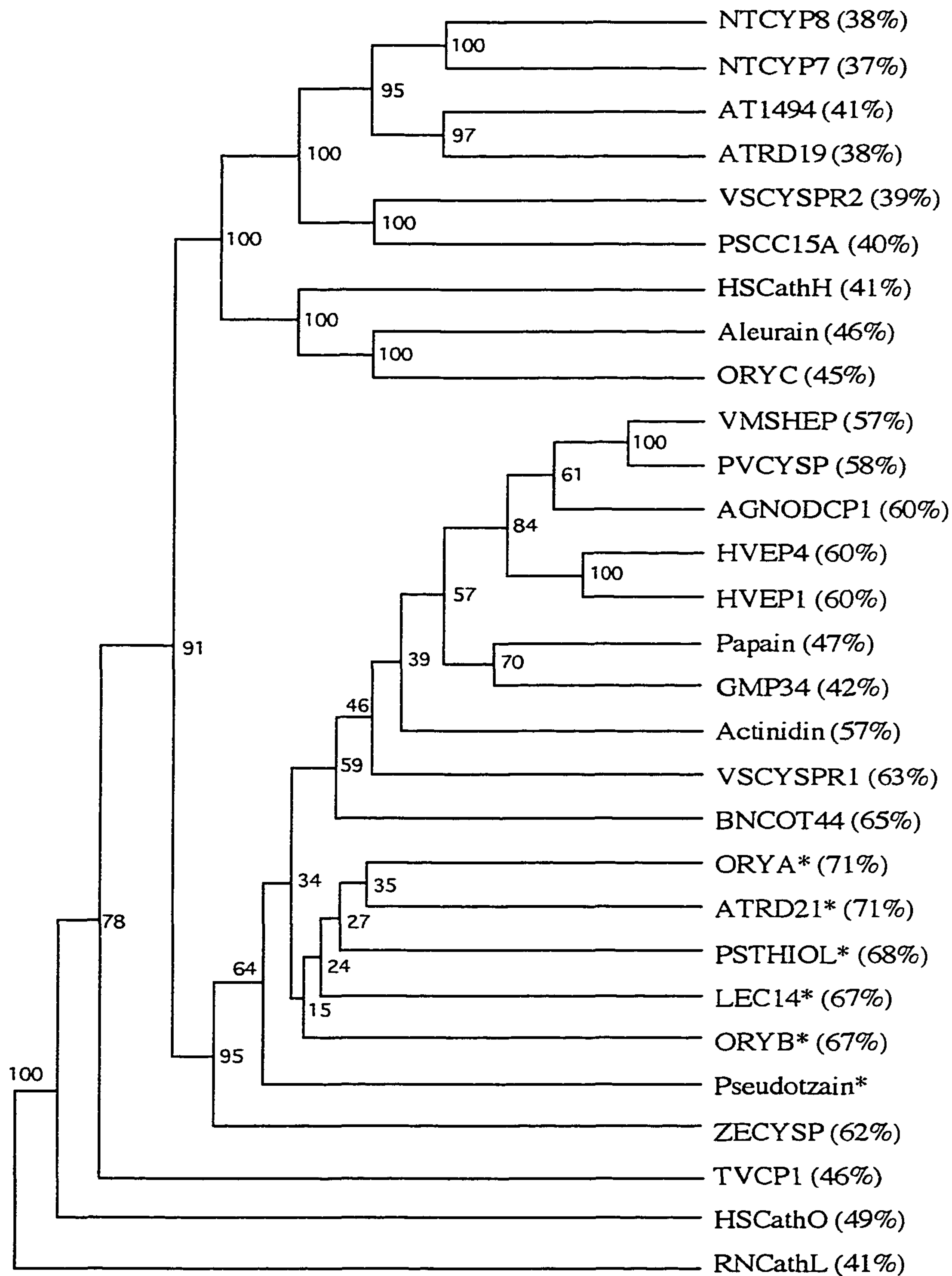


Figure 3.3. CysP are encoded as zymogens that can include N- and C-terminal extension peptides. Many also have a hydrophobic leader sequences targeting it to the endoplasmic reticulum. The results of the phylogenetic analysis presented in Figure 3.4 were based on the amino acid sequences of the mature portion of the CysP.

Figure 3.4. Phylogenetic tree of representative mature CysP from protozoan, animals and plants. Numbers at branches are bootstrap values out of 100. Numbers in parenthesis indicate the percent identity to the Douglas-fir CysP (pseudotzain). Asterisks indicate the preproprotein form of the CysP has both N- and C-terminal propeptides (actinidin also has both, although the C-terminal propeptide is only 34 amino acids in length compared to over 100 amino acids for the CysP indicated by an asterisk). The phylogenetic tree was constructed using the PHYLIP package (Felsenstein, 1989) based on mature CysP amino acid sequences aligned using CLUSTAL V (Higgins and Sharp, 1989). The following sequences were used to construct the phylogenetic tree: AT1494 (*Arabidopsis thaliana*, GenBank accession number X74359), NTCYP7 and NTCYP8 (tobacco, *Nicotiana tabacum*, Z13959 AND Z13964), BNCOT44 (*Brassica napus*, SWISS-PROT P25251), VSCYSPR1 and VSCYSPR2 (*Vicia sativa* L., GenBank X75749 and Z30338), LEC14 (tomato, *Lycopersicon esculentum*, M21444), ATRD19 and ATRD21 (*Arabidopsis thaliana*, D13042 and D13043), ORYA, ORYB, and ORYC (rice, *Oryza sativa*, D90406, D90407 and D90408), PSTHIOL, (pea, *Pisum sativum*, X66061), Pseudotzain, (this report, *Pseudotsuga menziesii.*, U41902), HVEP1 and HVEP4 (barley, *Hordem vulgare*, SWISS-PROT P25249 and P25250), VMSHEP (*Vigna mungo*, GenBank X15732), PVCYSP (*Phaseolus vulgaris*, X56753), AGNODCP1 (*Alnus glutinosa*, U13940), Papain, (*Carica papaya*, M15203), GMP34 (*Glycine max*, J05560), Actinidin (*Actinidia deliciosa*, X16466), ZECYSP (*Zinnia elegans*, U19267), PSCC15A (*Pisum sativum*, X54358), HSCathH (*Homo sapiens*, cathepsin H, X16832), Aleurain (barley, *Hordem vulgare*, X05167), TVCP1 (*Trichomonas vaginalis*, X77218), RNCathL (rat, *Rattus norvegicus*, L14776), HSCathO (human, U13665).

Figure 3.4.



```
PMCysP      MGILLLFAVLALSAMAGSASRA-DFSIISYDSQDLI------GDDAIMEL
RD21        MGFLKPTMAILFLAM-VAVSSAVDMSIISYDEKHGVSTTGGRSEAEVMSI
PSThiol     M--LLSKLTILFITLTFTLSLALDMCIISYDKTHPDKSTP-RTNDQVLTM
ORYA        MRISMALAAAALLLL-LSLAAA-DMSIVSYGE---------RSEEEARRL
            *          .  .  . .  * *  *.**.            .    .

PMCysP      YELWLAQHKKA--YNGLDEKQKKFSVFKDNFLYIHQHNNQGN---PSYKL
RD21        YEAWLVKHGKAQSQNSLVEKDRRFEIFKDNLRFVDEHNEK-N---LSYRL
PSThiol     YEEWLVKHGK--NYNALGEKEKRFEIFKDNLGFIDEHNSK-N---LSFRL
ORYA        YAEWKAEHGK--SYNAVGEEERRYAAFRDNLRYIDEHNAAADAGVHSFRL
            *  *   * *    *.. * ....  *.**. ....**    .    *..*

PMCysP      GLNQFADLSHEEFKAAYLGTKLDAKKRLSRSPS--PRYQYSVGEDLPESI
RD21        GLTRFADLTNDEYRSKYLGAKMEK-KGERRT-S--LRYEARVGDELPESI
PSThiol     GLNRFADLTNEEYRTRFLGTRINPNRRNRKVNSQTNRYATRVGDKLPESV
ORYA        GLNRFADLTNEEYRDTYLGLRNKP-RRERKV-S--DRYLAADNEALPESV
            ** .****...*..  .** .    .   .   *    **     . ****.

PMCysP      DWREKGAVTAVKNQGSCGSCWAFSTVAAVEGINQIVTGNLTSLSEQELVD
RD21        DWRKKGAVAEVKDQGGCGSCWAFSTIGAVEGINQIVTGDLITLSEQELVD
PSThiol     DWRKEGAVVGVKDQGSCGSCWAFSAIAAVEGVNKLATGDLISLSEQELVD
ORYA        DWRTKGAVAEIKDQGGCGSCWAFSAIAAVEDINQIVTGDLISLSEQELVD
            ***  ***   .*.**.********...***..*..  **.* .********

PMCysP      CDTSYNQGCNGGLMDYAFQFIISNGGLDSEDDYPYKANNGSCDAYRKNAH
RD21        CDTSYNEGCNGGLMDYAFEFIIKNGGIDTDKDYPYKGVDGTCDQIRKNAK
PSThiol     CDTSYNEGCNGGLMDYAFEFIINMVALTPEEDYPYRAIDGRCDQNRKNAK
ORYA        CDTSYNEGCNGGLMDYAFDFIINNGGIDTEDDYPYKGKDERCDVNRKNAK
            ******.***********.***     ..   . ****.. .   **  ****

PMCysP      VVTIDDYEDVPENDEKSLKKAAANQPISVAIEASGRAFQFYESGVFTSNC
RD21        VVTIDSYEDVPTYSEESLKKAVAHQPISIAIEAGGRAFQLYDSGIFDGSC
PSThiol     VVSIDQYEDVPAYDEGALKKAVANQVIAVAVEGGGREFQLYDSGVFTGRC
ORYA        VVTIDSYEDVTPNSETSLQKAVRNQPVSVAIEAGGRAFQLYSSGIFTGKC
            **.** ****     * .*.**  .* ...*.*..** **.* **.* . *

PMCysP      GTQLDHGVTLVGYGSESGIDYWLVKNSWGNSWGEKGFIKLQRNLEGASTG
RD21        GTQLDHGVVAVGYGTENGKDYWIVRNSWGKSWGESGYLRMARNIASS-SG
PSThiol     GTALDHGVAAVGYGTENGKDYWIVRNSWGGSWGEAGYIRLERNLATSKSG
ORYA        GTALDHGVAAVGYGTENGKDYWIVRNSWGKSWGESGYVRMERNIKAS-SG
            ** *****  ****.*.* ***.*.**** **** *.... **.   . .*

PMCysP      MCGIAMEASYPVKKGANPPNPGPSPPSPVKPPTVCDNYYSCPESNTCCCM
RD21        KCGIAIEPSYPIKNGENPPNPGPSPPSPIKPPTQCDSYYTCPESNTCCCL
PSThiol     KCGIAIEPSYPIKNGLNPPKPAPSPPSPVKPPSVCDSY-SCAEGSTCCCI
ORYA        KCGIAVEPSYPLKKGENPPNPGPTPPSPTPPPTVCDNYYTCPDSTTCCCI
             ****.*.***.*.* ***.*.*.****  **. **.* .*... ****.

PMCysP      YDFGGYCYAWGCCPLNSATCCDDHYSCCPSDHPVCDLDAQTCLKSRKDPF
RD21        FEYGKYCFAWGCCPLEAATCCDDNYSCCPHEYPVCDLDQGTCLLSKNSPF
PSThiol     FDYGGSCFEWGCCPLESATCCDDHYSCCPHEYPVCDTYAGLCRKNKNNPL
ORYA        YEYGKYCYAWGCCPLEGATCCDDHYSCCPHEYPICNVQQGTCLMAKDSPL
            ...*  *. ******..*******.***** . *.*.      *    .  *.

PMCysP      GTKMLKRTPAKPYWSL--------SG
RD21        SVKALKR---KPATPF-WSQGRKNIA
PSThiol     GVKSFKRTPAKPHFAIE-GKNKMGSV
ORYA        AVKALKRTLAKPNLSFLFGNGKKSSA
            . * .**   **  ..
```

Figure 3.5. Alignment of plant CysP sequences that have N-terminal extension peptides. The species and accession numbers for the amino acid sequences aligned with CLUSTAL V (Higgins and Sharp, 1989) are PMCysP, (pseudotzain, this report, *Pseudotsuga menziesii.*, U41902), RD21 (*Arabidopsis thaliana*, D13043), PSTHIOL, (pea, *Pisum sativum*, X66061), and ORYA (rice, *Oryza sativa*, D90406). An identical amino acid is indicated by (*) and a conserved amino acid is indicated by (.).

aleurain (Chrispeels and Raikhel, 1992) was not present in pseudotzain or in any of the plant CysP having both N- and C-terminal propeptides (Figures 3.4 and 3.5). The mature portion of pseudotzain also lacks the ER retention signal (His/Lys-Asp-Glu-Lue, HDEL/KDEL) seen at the C-terminal end of some CysP (Figure 3.4, VMSHEP, Akasofu et al., 1989; PVCYSP, Tanaka et al., 1991).

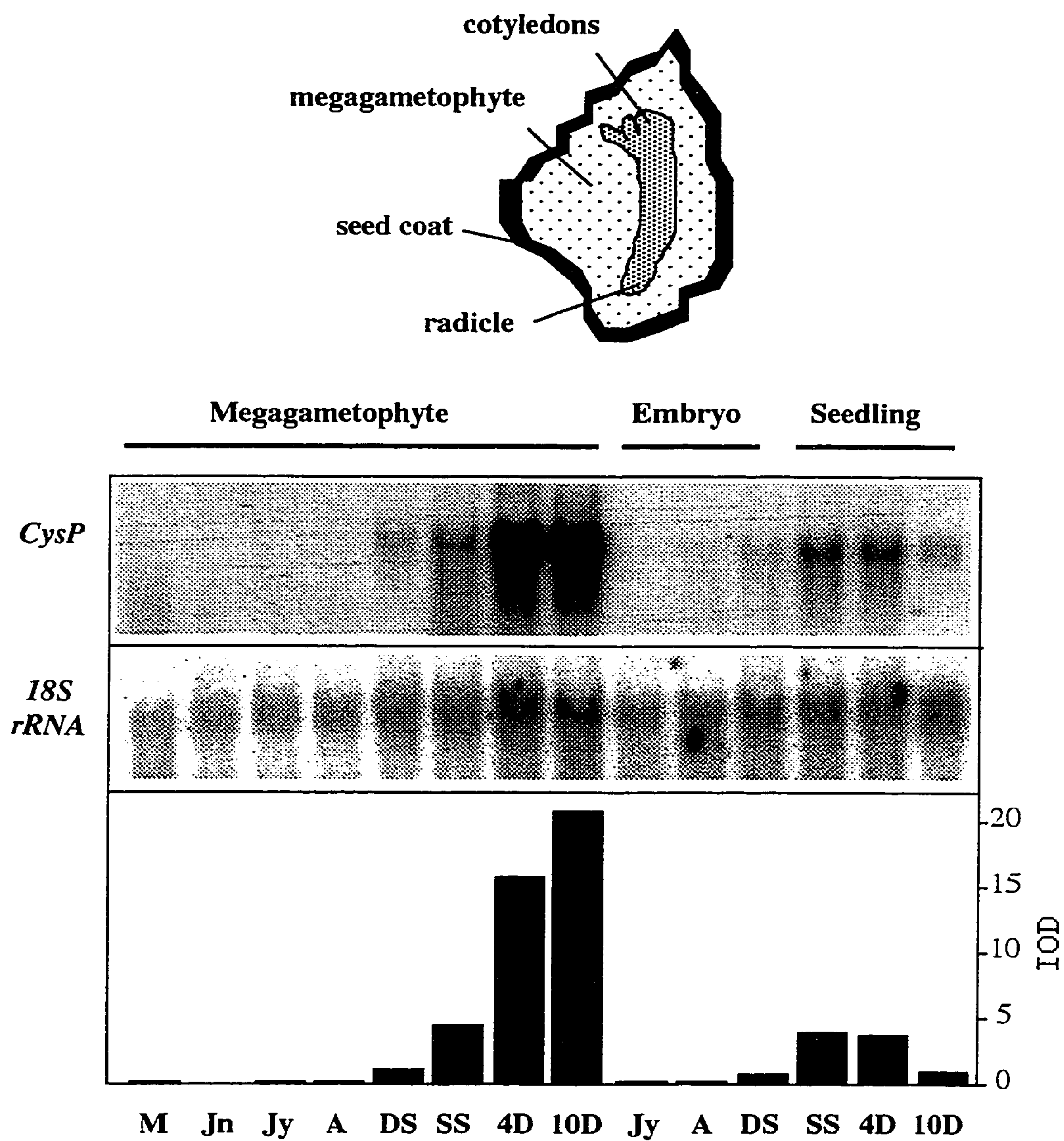
Northern blot analysis of cysteine proteinase mRNA steady-state amounts

To elucidate the pattern of *cysP* expression, steady-state transcript amounts from developing seedling tissues, and in response to various plant growth regulators and environmental conditions were analyzed by northern blot. The expression of *cysP* was compared with 18S rRNA gene expression. The *cysP* transcripts were not detected in prefertilized MG or early stages of embryogenesis (Figure 3.6, upper and lower panels, lanes marked by M, Jn, Jy, and A) but were detected in low quantities in MG and embryos of mature seeds (Figure 3.6, upper and lower panels, lanes marked by DS). The *cysP* transcripts increased slightly during stratification in both the MG and the embryo but the largest increase in transcript amounts occurred in the MG after germination (Figure 3.6, upper and lower panels, DS, SS, 4D, 10D). In seedlings, *cysP* transcripts increased until day 4 and decreased by day 10 (Figure 3.6, 4D and 10D).

Analysis of seed tissues showed that in 10-day-old seedlings highest *cysP* transcript quantities were found in the MG whereas cotyledons, roots and seed coats had lower amounts (Figure 3.7). The increase in *cysP* transcript amounts in 10-day-old seedlings did not correlate with the ability for a seed to germinate because amounts found in ungerminated seeds (seeds without emerging radicles kept under germination conditions for ten days) were similar to those in 10-day-old seedlings with different root lengths (Figure 3.7). Fourteen-day-old seedlings were subjected to various hormonal and environmental

Figure 3.6. Northern blot analysis of *cysP* transcript levels during embryogenesis and early seedling development. Total RNA was isolated according to Verwoerd et al. (1989) from MG, embryos and seedlings (12 μg total RNA per lane). The membrane was hybridized with the full length *cysP/PM3-3* insert (upper panel), stripped and rehybridized with a genomic clone for a portion of the 18S rRNA (middle panel). Transcript signals from each hybridization were quantified using a Personal Densitometer (Molecular Dynamics). The variations in the total RNA loaded in each lane were corrected by dividing the lowest rRNA densitometric value by each rRNA value to obtain a ratio. Each *cysP/PM3-3* hybridization signal value was then multiplied by their respective rRNA ratio. The results are presented graphically using the arbitrary units of integrated optical density (IOD) (lower panel). Northern blot analysis were repeated twice. May 31 (M), June 19 (J), July 17 (Jy), August 7 (A). stratified seeds (SS), mature seeds (DS), 4-day-old seedlings (4D), 10-day-old seedlings (10D).

Figure 3.6.



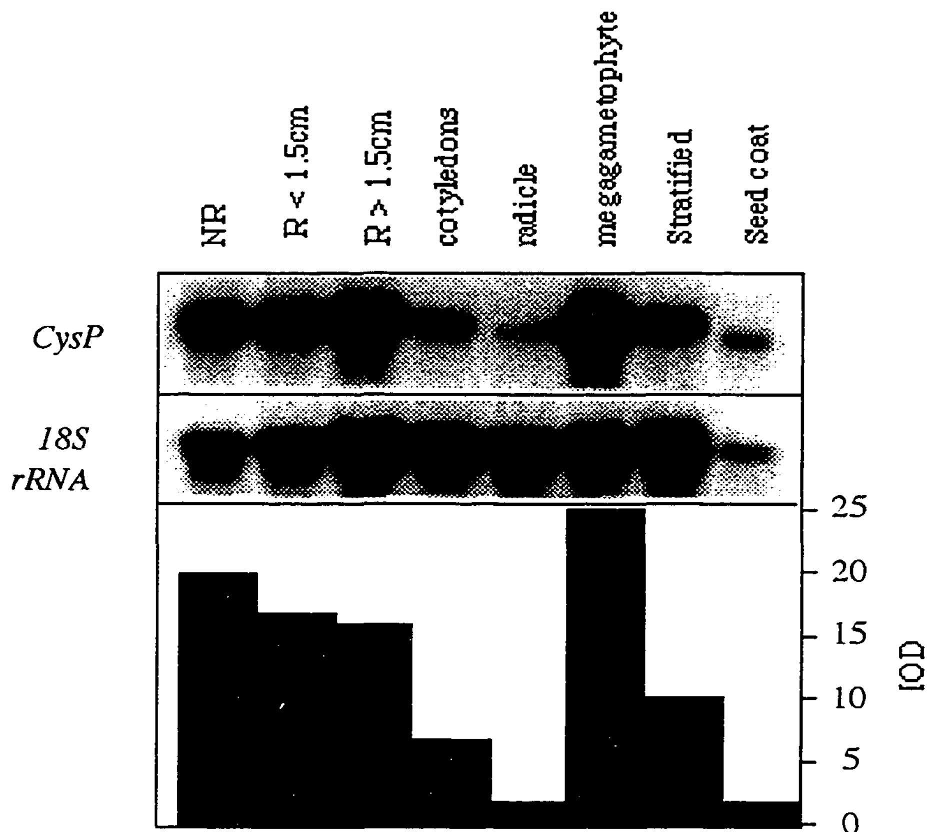


Figure 3.7. Northern blot analysis of *cysP* transcript levels in 10-day-old seedling tissues. NR, no radicle; R, radicle visible with lengths less than (<) or greater than (>) 1.5 cm; Stratified, stratified seeds. Panels are as described in the legend for Figure 3.6. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds or from separated tissues as described previously (Chapter 2, Materials and Methods; Tranbarger and Misra, 1995). Twenty μg of total RNA was loaded per lane.

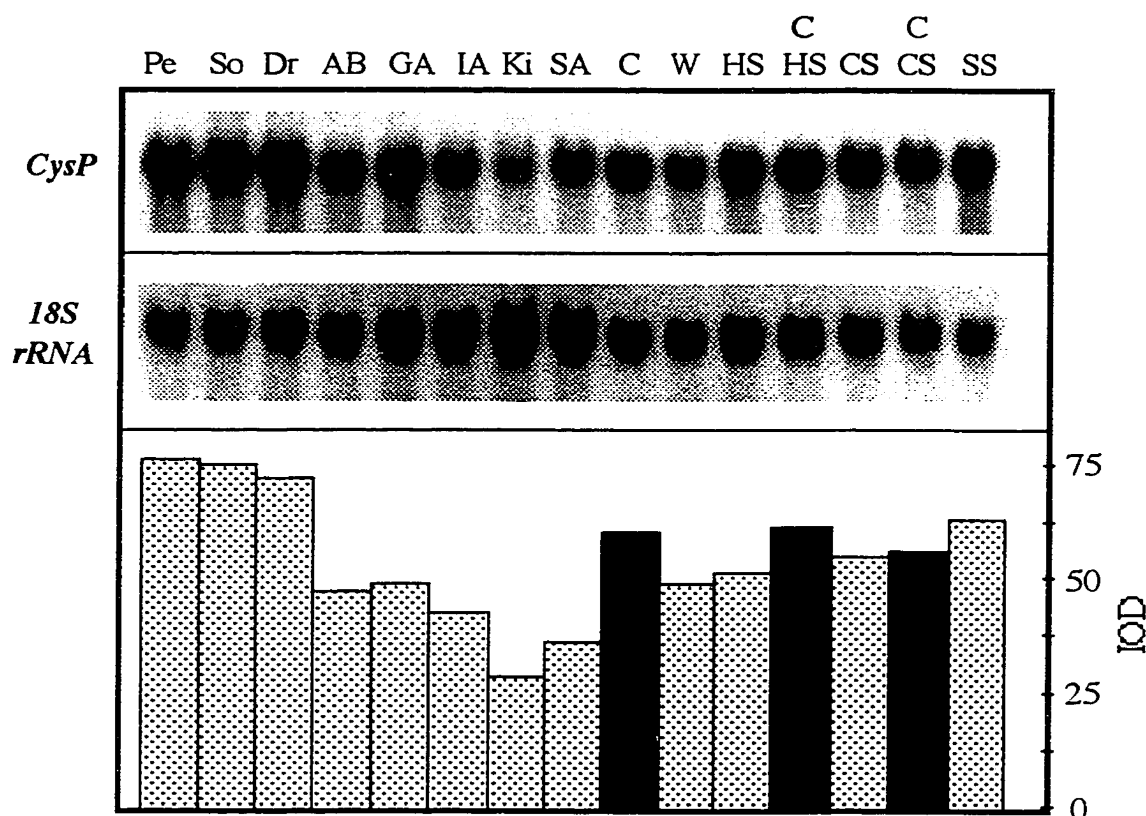


Figure 3.8. Northern blot analysis of *cysP* transcript levels in response to hormonal and environmental treatments. Panels and germination conditions are as described in the legend for Figure 3.6. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds, whole seedlings or from separated tissues as described previously (Chapter 2, Materials and Methods; Tranbarger and Misra, 1995). Seedlings were treated as described under "Materials and Methods." (Pe) PEG, 24% polyethylene glycol; (So) NaCl, 250 mM sodium chloride; (Dr) dry; (AB) 10^{-4} M abscisic acid; (GA), 10^{-4} M gibberellic acid (GA_3); (IA), 10^{-4} M indoleacetic acid; (Ki) Kinetin, 10^{-4} M kinetin; (Sa) SA, 10 mM salicylic acid; (C) Control plants in water; (W) wounding; (HS), heat shock; (CHS) HS control; (CS), cold shock; (CCS) CS control; (St) Stratified, stratified seeds. Twenty μg of total RNA was loaded per lane.

treatments and northern blot analysis showed that the treatments that affected the water status of the seedlings had a minor but consistent effect on increasing the *cysP* steady-state transcript amounts (Figure 3.8, upper and lower panels, lanes marked by Pe, So, Dr, and C). In contrast, seedlings incubated in GA, ABA or other plant growth regulators showed a slight decrease in *cysP* transcript quantities (Figure 3.8, upper and lower panels, lanes marked by AB, GA, IA, Ki, SA, and C). Wounding, heat and cold treatment had no effect on *cysP* transcript amounts (Figure 3.8, upper and lower panels, lanes marked by HS, CHS, CS and CCS). The amount of *cysP* transcript found in seedlings after 16 days of growth decreased to a similar amount as that found in stratified seeds (Figure 3.8, upper and lower panels, lanes marked by SS and C).

Additional experiments were done to examine the effect of wounding and desiccation on the expression of *cysP* during germination and early seedling growth (Figures 3.9 and 3.10). The seedlings showed a marked decrease in the amount of *cysP* transcripts as early as 1-hour after wounding (Figure 3.9). A decrease in the amount of 18S ribosomal RNA was also seen indicating that a general inhibition of transcription or degradation of RNA was caused by the wounding treatment. Smaller 18S rRNA fragments appear during the time course suggesting a specific pattern of degradation in response to wounding. Wounding does not increase *cysP* steady-state transcript amounts during germination or early seedling development. Desiccation and incubation of germinating seeds and seedlings with sodium chloride had only a minor effect on the amount of *cysP* transcripts (Figure 3.10). *CysP* transcripts increased normally during this developmental period (1-4 days; Figure 3.6). Therefore, an increase of *cysP* transcript amounts in response to desiccation may not be detectable above the developmental associated accumulation.

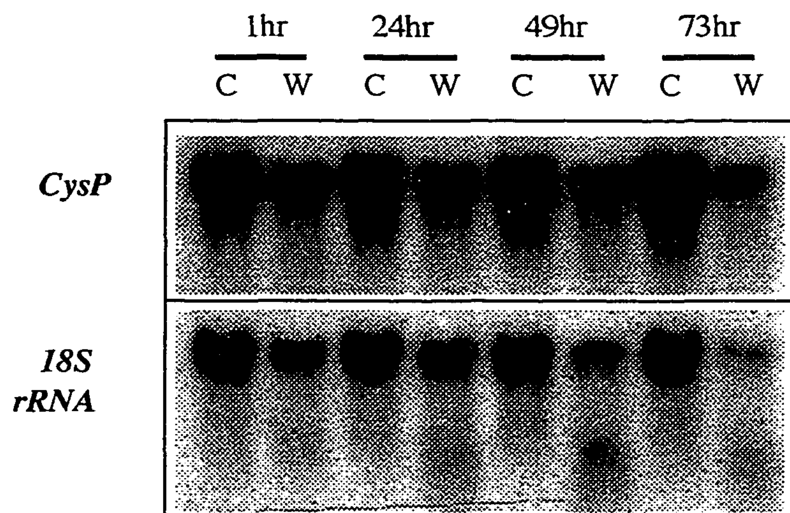


Figure 3.9. Northern blot analysis of *cysP* expression in response to wounding during germination. W, wounded seeds (pinched with forceps) and C, unwounded control seeds were incubated in water in petri plates in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hr. photoperiod for 1, 24, 49 and 73 hours. Panels are as described in the legend for Figure 3.6. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds as described previously (Chapter 2, Materials and Methods; Tranbarger and Misra, 1995). Twenty μg of total RNA was loaded per lane.

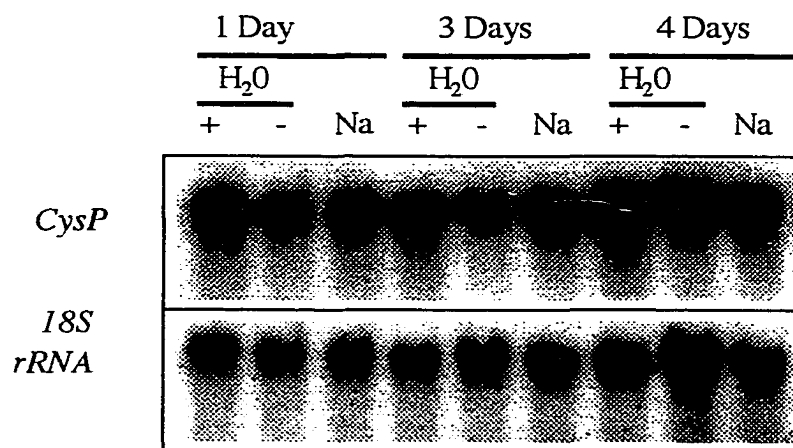


Figure 3.10. Northern blot analysis of *cysP* expression in response to desiccation treatments during germination. Seeds were germinated in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hr. photoperiod for 1, 3, and 4 days in petri plates with (+) and without (-) H_2O, and in an aqueous solution of 250 mM sodium chloride (Na). Panels are as described in the legend for Figure 3.6. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds and seedlings as described previously (Chapter 2, Materials and Methods; Tranbarger and Misra, 1995). Twenty μg of total RNA was loaded per lane.

Discussion

The mobilization of storage reserves from the storage tissues of the seed to the developing seedling is one of the most important metabolic activities that takes place during germination and post-germination. Before the seedling can establish photosynthetic self-sufficiency and produce the materials required for growth, it is dependent on the mobilization of stored seed reserves. This section describes the isolation of a cDNA clone that encodes a CysP (pseudotzain) from Douglas-fir that is expressed preferentially in the storage tissues of the seed during the time of reserve mobilization. The amino acid sequence of pseudotzain is similar to CysP that have both N- and C-terminal extension peptides. This type of CysP appears to be unique to plants. The physiological significance of having both N- and C-terminal propeptides and their roles in intracellular targeting are not known. However, their common occurrence suggests a conserved mechanism of targeting and/or functional role for this subgroup of CysP that evolved before the divergence of angiosperms and gymnosperms and after the divergence of plants and animals.

The presence of a hydrophobic leader sequence in the N-terminal region implies targeting of pseudotzain to the ER, whereas putative *N*-glycosylation sites similar to aleurain indicates the possibility of further post-translational processing (Rogers et al., 1985). Vacuolar targeting information of plant proteins resides in either the C- or N-terminal propeptide regions, or within the mature enzyme itself and not in glycans as with animals (Chrispeels and Raikhel, 1992). Other targeting information may also reside in the mature enzyme and/or the N- and C-terminal extension peptides of pseudotzain. Finally, further post-translational processing presumably results in the mature form of pseudotzain predicted by the conserved amino acid residues Leu^{132} and Ala^{352}.

Based on the present study, the expression of Douglas-fir *cysP* is affected by osmotic

stresses but does not appear to be regulated by ABA. Osmotic stress induces *cysP* gene expression in pea and *Arabidopsis* and the effect is independent of ABA (Guerroro et al., 1990; Koizumi et al., 1993; Williams et al., 1994). ABA-responsive genes are thought to have a role in the plant's response to osmotic stress although it now appears that an ABA-independent gene response to osmotic stress also occurs (Skriver and Mundy, 1990; Guerroro et al., 1990; Bostock and Quatrano, 1992; Koizumi et al., 1993; Williams et al., 1994). GA induces *cysP* expression in several angiosperm species (Watanabe et al., 1991; Rogers et al., 1985; Koehler and Ho, 1990) but did not increase steady-state *cysP* transcript amounts in the present study. In *A. thaliana*, the expression of *cysP* genes (*RD19* and *21*) was not affected by GA (Koizumi et al., 1993). In pea, the expression of the *cysP* gene *tpp* was negatively affected by GA (Granell et al., 1992). Thus, the mechanisms of *CysP* gene regulation appear to vary from species to species and within the developmental contexts in which the expression occurs.

Several features in the nucleic acid sequence of *cysP* suggest the possibility of post-transcriptional regulation. The CysP cDNA clone *PM3-3* contains an uORF in the 5' UTR, a feature found in up to 10% of plant gene leader sequences (Fütterer and Hohn, 1996). A number of recent studies point to possible functions of uORFs. For example, translation of the mRNA of the maize transcriptional activator OPAQUE 2 is inhibited by uORFs present in the leader sequence (Lohmer et al., 1993). The removal of two AUG codons of an uORF in a plasma membrane ATPase transcript increased translation in wheat germ extracts and protoplasts (Michelet, et al., 1994). In yeast, the presence of an uORF accelerated mRNA degradation (Oliveira and McCarthy, 1995). The significance of the uORF in the 5' UTR of *PM3-3* remains to be determined.

Recent studies of mRNA turnover have defined specific *ARE* in the 3' *UTR* that affect mRNA stability and longevity (Lagnado et al., 1994; Zubiaga et al., 1995). Similar *ARE* motifs have also been shown to function as instability determinants of plant transcripts

(Ohme-Takagi et al., 1993; Abler and Green, 1996). The *PM3-3* cDNA clone has multiple *ARE* elements in the 3' end that may direct post-transcriptional regulation at the level of mRNA turnover. An interaction of the *ARE* with yet to be identified destabilization factors may in part regulate expression of *cysP* during development and/or in response to osmotic stress post-transcriptionally.

The temporal and preferential expression of *cysP* transcript in the MG supports the hypothesis that pseudotzain is involved in the mobilization of storage proteins during germination and early seedling growth. In addition, pseudotzain presumably undergoes a series of post-translational processing steps to remove the N- and C-terminal extension peptides, although nothing is known about the processing of CysP with both types of extension peptides (Kamphuis et al., 1985; Koehler and Ho, 1990; Kalinski et al., 1992). Future experiments will address the mechanisms that regulate *cysP* expression, the processing steps that lead to the mature form of pseudotzain, and the role in storage protein mobilization during germination and post-germination.

Conclusions

(1) The cDNA clone *PM3-3* encodes a CysP containing both N- and C-terminal propeptides and is most similar to other CysP isolated from angiosperms containing N- and C-terminal propeptides.

(2) The mature CysP that have both N- and C-terminal propeptides form a distinct phylogenetic subgroup.

(3) There are putative mRNA destabilization motifs in both the 5' (uORF) and 3' (*ARE*) *UTR* of clone *PM3-3*.

(4) During embryogenesis, *cysP* transcript amounts are lower when storage proteins are accumulating during seed maturation and higher in the MG and embryo at the end of seed maturation.

(5) During early seedling development:

(i) The *cysP* transcripts increase in abundance mainly in the MG, the main storage organ of conifer seeds.

(ii) In 10-day-old seedlings, *cysP* transcripts are detected in the roots, shoots, and seed coat but the highest amounts are seen in the MG.

(iii) Osmotic stress treatments and desiccation slightly enhance the steady-state amount of *cysP* transcripts.

(iv) Exogenous treatment with ABA, GA, and other plant growth regulators have no effect or slightly decrease the steady-state amount of *cysP* transcripts.

Literature Cited

Abler, M. L., and Green, P. J. (1996) Control of mRNA stability in higher plants. Plant Molecular Biology 32: 63-78

Akasofu, H., Yamauchi, D., Mitsuhashi, W., and Minamikawa, T. (1989) Nucleotide sequence of cDNA for sulfhydryl-endopeptidase (SH-EP) from cotyledons of germinating *Vigna mungo* seeds. Nucleic Acids Research. 17: 6733

Allen, S. A., and Owens, J. N. (1972) The life history of Douglas fir. Information Canada, Ottawa

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403-410

Baker, E. N., and Drenth, J. (1987) The thiol proteases: structure and mechanism. *In* Biological Macromolecules and Assemblies. *Eds* Jurnak, F. A., and McPherson, A. John Wiley & Sons, New York pp. 314-368

Becker, C., Fischer, J., Nong, V. H., and Münitz K. (1994) PCR cloning and expression analysis of cDNAs encoding cysteine proteinases from germinating seeds of *Vicia sativa L.* Plant Molecular Biology 26: 1207-1212

Bostock, R. M., and Quatrano, R. S. (1992) Regulation of *Em* gene expression in rice. Plant Physiology 98: 1356-1363

Cavener, D. R., and Ray, S. C. (1991) Eukaryotic start and stop translation sites. Nucleic Acids Research. 19: 3185-3192

Cervantes, E., Rodríguez, A., and Nicolás, G. (1994) Ethylene regulates the expression of a cysteine proteinase gene during germination of chickpea (*Cicer arietinum L.*). Plant Molecular Biology 25: 207-215

Chrispeels, M. J., and Raikhel, N. V. (1992) Short peptide domains target proteins to plant vacuoles. Cell 68: 613-616

Dietrich, R. A., Maslyar, D. J., Heupel, R. C., and Harada, J. J. (1989) Spatial patterns of gene expression in *Brassica napus* seedlings: Identification of a cortex-specific gene and localization of mRNAs encoding isocitrate lyase and a polypeptide homologous to proteinases. Plant Cell 1: 73-80

Felsenstein, J. (1989) PHYLIP-- Phylogeny inference package (version 3.2). Cladistics 5: 164-166

Fütterer, J., and Hohn, T. (1996) Translation in plants - rules and exceptions. Plant Molecular Biology 32: 159-189

Gish, W., and States, D. J. (1993) Identification of protein coding regions by database similarity search. Nature Genetics 3: 266-72

Granell, A., Harris, N., and Pisabarro, A. G. (1992) Temporal and spatial expression of a thiolprotease gene during pea ovary senescence, and its regulation by gibberellin. Plant Journal 2: 907-915

Guerroro, F. D., Jones, J. T., and Mullet, J. E. (1990) Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted. Sequence and expression of three inducible genes. Plant Molecular Biology 15: 11-26

Higgins, D. G., and Sharp P. M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS 5: 151-153

Kalinski, A., Weisemann, J. M., Mathews, B. F., and Herman, E. M. (1990) Molecular cloning of a protein associated with soybean seed oil bodies that is similar to thiol proteases of the papain family. Journal of Biological Chemistry 265: 13843-13848

Kalinski, A., Melroy, D. L., Dwivedi, R. S., and Herman, E. M. (1992) A soybean vacuolar protein (P34) related to thiol proteases is synthesized as a glycoprotein precursor during seed maturation. Journal of Biological Chemistry 267: 12068-12076

Kamphuis, I. G., Drenth, J., and Baker, E. N. (1985) Thiol proteases - comparative studies based on the high-resolution structures of papain and actinidin, and on amino acid sequence information for cathepsins B and H, and stem bromelain. Journal of Molecular Biology 182: 317-329

Koehler, S. M., and Ho, T.-H. D. (1990) Hormonal regulation, processing, and secretion of cysteine proteinases in barley aleurone layers. Plant Cell 2: 769-783

Koizumi, M., Yamaguchi-Shinozaki, K., Tsuji, H., and Shinozaki, K. (1993) Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. Gene 129: 175-182

Lagnado, C. A., Brown, C. Y., and Goodall, G. J. (1994) AUUUA is not sufficient to promote Poly(A) shortening and degradation of an mRNA: the functional sequence within AU-rich elements may be UUAUUUA(U/A)(U/A). Molecular and Cellular Biology 14: 7984-7995

Lammer, D. L., and Gifford, D. J. (1989) Lodgepole pine seed germination. II. The seed proteins and their mobilization in the megagametophyte and embryonic axis. Canadian Journal of Botany 67: 2544-2551

Leal, I., and Misra, S. (1993) Developmental gene expression in conifer embryogenesis and germination. III. Analysis of crystalloid protein mRNAs and desiccation protein mRNAs in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss). Plant Science 88: 25-37

Lohmer, S., Maddaloni, M., Motto, M., Salamini, F., and Thompson, R. D. (1993) Translation of the mRNA of the maize transcriptional activator *opaque-2* is inhibited by upstream open reading frames present in the leader sequence. Plant Cell 5: 65-73

Misra, S. (1994) Conifer zygotic embryogenesis, somatic embryogenesis, and seed germination: Biochemical and molecular advances. Seed Science Research 4: 357-384

Michelet, B., Lukaszewicz, M., Dupriez, V., and Boutry, M. (1994) A plant plasma membrane proton-ATPase gene is regulated by development and environment and shows signs of a translational regulation. Plant Cell 6: 1375-1389

Ohme-Takagi, M., Taylor, C. B., Newman, T. C., and Green, P. J. (1993) The effects of sequences with high AU content on mRNA stability in tobacco. Proceedings of National Academy of Sciences USA 90: 11811-11815

Oliveira, C. C., and McCarthy, J. E. G. (1995) The relationship between eukaryotic translation and mRNA stability - a short upstream open reading frame strongly inhibits translational initiation and greatly accelerates mRNA degradation in the yeast *Saccharomyces cerevisiae*. Journal of Biological Chemistry 270: 8936-8943

Owens, J. N., Morris, S. J., and Misra S. (1993) The ultrastructural, histochemical, and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. Canadian Journal of Forest Research 23: 816-827

Rogers, J. C., Dean, D., and Heck, G. R. (1985) Aluerain: a barley thiol protease closely related to mammalian cathepsin H. Proceedings of National Academy of Sciences USA 82: 6512-6516

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY. ISBN 0-87969-309-6

Salmia, M. A. (1981a) Fractionation of the proteinases present in the endosperm of germinating seed of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 51: 253-258

Salmia, M. A. (1981b) Proteinase activities in resting and germinating seeds of Scots pine, *Pinus sylvestris*. Physiologia Plantarum 53: 39-47

Schaffer, M. A., and Fischer, R. L. (1988) Analysis of mRNAs that accumulate in response to low temperature identifies a thiol protease gene in tomato. Plant Physiology 87: 431-436

Shaw, G., and Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46: 659-667

Shutov, A. D., and Vaintraub, I. A. (1987) Degradation of storage proteins in germinating seeds. Phytochemistry 26: 1557-1566

Skriver, K., and Mundy, J. (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2: 503-512

Tanaka, T., Yamauchi, D., and Minamikawa, T. (1991) Nucleotide sequence of cDNA for an endopeptidase (EP-C1) from pods of maturing *Phaseolus vulgaris* fruits. Plant Molecular Biology 16: 1083-1084

Tranbarger, T. J., and Misra S. (1995) The molecular characterization of a set of cDNAs differentially expressed during Douglas-fir germination and early seedling development. Physiologia Plantarum 95: 456-464

Verwoerd, T. C., Dekker, B. M. M., and Hoekema, A. (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Research. 17: 2362

Watanabe, H., Abe, K., Emor, Y., Hosoyama, H., and Arai S. (1991) Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). Journal of Biological Chemistry 266: 16897-16902

Williams, J., Bulman, M., Huttly, A., Phillips, A., and Neill, S. (1994) Characterization of a cDNA from *Arabidopsis thaliana* encoding a potential thiol protease whose expression is induced independently by wilting and abscisic acid. Plant Molecular Biology 25: 259-270

Zubiaga, A. M., Belasco, J. G., and Greenberg, M. E. (1995) The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. Molecular and Cellular Biology 15: 2219-2230

CHAPTER 4

A Douglas-fir NADPH-cytochrome P450 reductase expressed during germination and seedling development: expression pattern and evidence for post-translational regulation

Abstract

NADPH-cytochrome P450 reductase (CPR; NADPH-ferrihemoprotein reductase; EC 1.6.2.4) is a key enzyme that transfers electrons from NADPH to the cytochrome P450 family of enzymes. To characterize the expression and regulation of this enzyme in Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco), a full length cDNA was isolated from a seedling λ ZAP cDNA library. Northern blot analysis indicated *cpr* expression is regulated both developmentally before seed maturation and differentially in the cotyledons, radicle and MG of seedling tissues. An antiserum raised against a synthetic CPR-peptide detected a single 80 kDa polypeptide in microsomes isolated from seeds and seedlings in western blot analysis. Microsomal membrane fractionation in sucrose gradients indicated CPR is associated with membranes of the endoplasmic reticulum and NADPH:cytochrome c reductase activity. The increase in steady-state mRNA quantities do not correlate with similar amounts of CPR protein present in the developmental stages examined, indicating either undetected *cpr* genes are expressed and translated, or transcriptional and/or post-transcriptional regulation of *cpr* gene activity. Furthermore, NADPH:cytochrome c reductase activity in microsomes increased during stratification, germination and early

seedling development, indicating that CPR enzyme may be stored in mature seeds in an inactive form and activated by events during stratification and germination.

Introduction

The cytochrome P450 monooxygenase (P450) gene superfamily of eukaryotes encodes numerous microsomal associated heme-containing proteins that catalyze a range of oxidation reactions often important in the detoxification of xenobiotic compounds and drug metabolism in mammals (reviewed by Nebert and Gonzalez, 1987; Porter and Coon, 1991). In plants, P450s are implicated in reactions necessary for normal plant development, growth and defense including the biosynthesis of gibberellins, jasmonates, salicylic acid, lignin, fatty acids, alkaloids, phytoalexins, glucosinolates, phenylpropanoids, and terpenoids, in addition to reactions leading to herbicide detoxification (for reviews see Donaldson and Luster, 1991, Bolwell et al. 1994, Schuler, 1996).

In mammals and yeast, all P450 enzyme activities are dependent on the transfer of electrons originating from NADPH to the P450s by another microsomal integral membrane protein, NADPH-cytochrome P450 reductase (CPR; NADPH-ferrihemoprotein reductase; EC 1.6.2.4), encoded by a single gene (Simmons et al., 1985; Sutter and Loper, 1989). CPR is a unique type of flavoprotein characterized by the presence of the two noncovalently bound cofactors FAD and FMN. Electrons from NADPH are transferred first to FAD and then to FMN before sequential single electron transfer to the heme group of the P450 enzymes (Vermilion et al., 1981). A functionally active cytochrome P450:CPR complex occurs at a 1:1 ratio (Miwa and Lu, 1984; Miwa et al., 1979), whereas the amount of P450 enzyme is in excess of up to 25 times the CPR content (Peterson et al., 1976). Cytochrome P450 isozymes appear to compete to bind CPR (Cawley et al., 1995) and are therefore thought to limit cytochrome P450 metabolic activities.

In plants at least two distinct genes encode CPRs in *Arabidopsis thaliana* (Urban et al.,

1997; Mizutani and Ohta, 1998), parsley (Koopmann and Hahlbrock, 1997), and poppy (Rosco et al., 1997). CPR isoforms have been immunologically characterized in microsomes of *Helianthus tuberosus* (Jerusalem artichoke), and the isolation of at least two distinct partial cDNA clones were reported (Benveniste et al., 1989, 1991; Lesot et al. 1995). In contrast, a partial length CPR cDNA clone from *Catharanthus roseus* detected only one gene copy in Southern blot analysis (Meijer et al. 1993), and only a single CPR cDNA has been isolated so far from *Vigna radiata* (Shet et al. 1993). The significance of plant CPR diversity on cytochrome P450 metabolism is unknown.

Animal, insect and plant P450 genes can be expressed according to a developmental program, regulated by specific substrates or xenobiotics, or in response to wounding and pathogen infection (Porter and Coon 1991; Bolwell et al. 1994; Teutsch et al. 1993; Frank et al., 1996; Mizutani et al. 1997), whereas the developmental or tissue-specific regulation of mammalian CPR expression appears to be rare (Shephard et al. 1992). Some plant CPR expression is increased in response to wounding, infection and fungal elicitors (Meijer et al. 1993; Lesot et al. 1995; Koopmann and Hahlbrock 1997; Mizutani and Ohta, 1998). Little is known, however, about the regulation of CPR gene expression in plants with respect to CPR activity and the activities of the different P450 enzymes.

Recent studies implicate plant P450 enzymes in an early biosynthetic step of gibberellins (Winkler and Helentjaris, 1995), the biosynthesis of the brassinosteriod hormone brassinolide essential for cell elongation (Szekeres et al. 1996), key enzymes leading to the synthesis of lignin (Meyer et al., 1996), and fatty acid hydroxylation reactions unique to plants (Cabello-Hurtado et al., 1998), all critical processes during plant growth and development. In conifers, cytochrome P450 enzymes are involved in the biosynthesis of diterpenoid resin, an important component in the defense against insects and pathogens

(Funk and Croteau, 1994; Funk et al., 1994). Very little is known about the cytochrome P450 monooxygenase system of conifers, its regulation and roles during germination and the early stages of seedling growth and development. Several studies have shown changes in gene expression patterns of Douglas-fir seeds during germination and early seedling development (Chapters 2 and 3; Taylor et al., 1993; Taylor and Davies, 1995; Tranbarger and Misra, 1995; Tranbarger and Misra, 1996; Kaukinen, et al., 1996; Jarvis et al., 1996; Jarvis et al., 1997). Although these studies clearly demonstrate that transcriptional changes occur, they focus mainly on the expression of genes without examining their corresponding gene product and its activity.

The present section describes, to the best of my knowledge, the isolation of the first full length CPR cDNA from a gymnosperm, the analysis of its expression, and the accumulation and localization of its gene product and activity. The cDNA and a CPR-peptide antiserum were used to determine the amounts of CPR mRNA and enzyme, respectively, during Douglas-fir zygotic embryogenesis, germination and early seedling development. The CPR enzyme was localized in the membrane fraction of the microsome and was associated with the ER and NADPH:cytochrome c activity. CPR expression was examined in relation to the amounts of CPR enzyme and activity (quantified by NADPH:cytochrome c reduction) and results suggest that regulation of CPR activity may occur post-translationally during stratification and the early stages of germination.

Materials and Methods

Plant material and growth conditions

Douglas-fir (*Pseudotsuga menziesii* [Mirb] Franco) seeds were imbibed overnight with sterile distilled water at 2-4 °C. The seeds were then blotted dry, placed in a plastic bag and stratified at 2-4 °C for 3-4 weeks, a treatment which improves germination of several conifer species (Edwards, 1980, 1986) including Douglas-fir (El-Kassaby et al., 1992). Stratified seeds were germinated in trays in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hour photoperiod. Cool White High Output fluorescent bulbs (F72T12/CW/1500) from Canadian General Electric Company Limited and 60 watt Country/Royale incandescent bulbs from Philips were used to provide a fluence rate of about 300 μmol m^{-2} s^{-1}. Megagametophytes (MG) and embryos were dissected from developing cones collected from a seed orchard of Pacific Forest Products Limited, Saanichton, B. C. All plant material collected was frozen in liquid nitrogen and stored at -80 °C.

cDNA clone isolation and sequence analysis

A 1.1 kb partial CPR cDNA clone *(PM14)* previously isolated (Tranbarger and Misra, 1995) was labelled by random priming with [α-^{32}P]dCTP and used to rescreen the cDNA library for full length clones. A total of 1 x 10^4 phage clones were transferred to nitrocellulose filters and putative clones were picked and excised from λ ZAP II as recombinant pBluescript SK(-) plasmid (Stratagene).

The cDNA insert sequences were determined by the Taq Dye Primer Cycle Sequencing

and the Taq DyeDeoxy Terminator Cycle Sequencing Kits using the Applied Biosystems Model 373A DNA Sequencing System (Applied Biosystems, Inc., Foster City, California, U.S.A.), and by the dideoxynucleotide chain termination method using Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio, U.S.A.). Primers used for sequencing were synthesized using the Applied Biosystems PCR-Mate DNA Synthesizer as described by the manufacturer (Applied Biosystems, Inc., Foster City, California, U.S.A.). DNA similarity searches were performed using the BLAST network service at NCBI (Altschul et al., 1990, Gish and States, 1993). CPR amino acid sequences were aligned using CLUSTAL W (Thompson et al., 1994). The "strong" groups of conserved amino acids are STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW and the "weak" groups are CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, HFY.

RNA isolation and analysis

Total RNA was isolated according to Verwoerd et al. (1989) with minor modifications (Leal and Misra, 1993), The RNA concentration was determined spectrophotometrically (A_{260}). The RNA was electrophoretically separated on agarose gels containing formaldehyde (Sambrook et al., 1989). Gels were stained with ethidium bromide to assure equal loading (total RNA per lane), and the RNA was transferred to a Zeta-Probe GT membrane (BIO-RAD, Mississauga, Ontario, Canada). The cDNA inserts were labelled by random priming with [α-^{32}P]dCTP and hybridized (BIO-RAD, Mississauga, Ontario, Canada). *CPR* transcript signals from each hybridization were compared to the hybridization signal of the 18S ribosomal RNA. See the Materials and Methods section in Chapter 2 for the northern blot analysis procedure.

CPR-peptide antiserum production

The synthetic CPR peptide was synthesized at the University of Victoria Protein Micro-Chemistry Centre using an Applied Biosystems Model 430A automatic peptide synthesizer (Foster City, CA) with FastMoc chemistry software. The peptide was synthesized with an additional cysteine at the C-terminus in order to couple the peptide to maleimide-activated keyhole limpet hemocyanin (KLH) carrier protein using the Imject® kit according to the manufactures instructions (Pierce). The CPR-peptide KLH conjugate was mixed with Freunds adjuvant, injected into rabbits, and antiserum titer was determined by ELISA with the free CPR-peptide and by western blot analysis with protein extracts from Douglas-fir.

Microsomal membrane isolation, protein extraction and analysis

Microsomes were prepared according to Shet et al. 1993 with minor modifications. Tissue samples stored at -80 °C were ground in liquid nitrogen to a fine powder with a mortar and pestle. Two volumes of Buffer A (100 mM Tris-HCl, pH 7.4; 250 mM sucrose; 2.8 mM ß-mercaptoethanol; 1 mM EDTA; 1 mM Phenylmethylsulfonyl fluoride; 2% Polyclar AT) per gram of tissue was added to the ground sample with a pinch of acid washed sand followed by further grinding. The slurry was filtered through miracloth, and the filtrate was centrifuged at 25,000 x g for 20 minutes at 4 °C in a JA20 rotor. The supernatant (SN1) was collected and centrifuged at 100,000 x g for 60 minutes at 4 °C. Supernatant (SN2) was removed, the pellets were washed gently in Buffer A, air dried on ice and resuspended in ice cold Buffer B (50 mM sodium phosphate, pH 7.5; 20% glycerol, 10 mM ß-mercaptoethanol). The samples were stored in Buffer B at -80 °C until

analyzed.

The microsomes and the supernatants were assayed for protein concentration with Bio-Rad Protein Assay reagent according to the manufacturer's directions (Bio-Rad Laboratories). Protein samples were resolved by 2-dimensional SDS-PAGE according to O'Farrell and O'Farrell (1977) and 1-dimensional SDS-PAGE according to Sambrook et al. (1989). Extracts were mixed 1:1 (v/v) with 2X Laemmli sample buffer (Laemmli, 1970), heated at 90 °C for 5 min, and centrifuged at 13,000 x g for 5 seconds, and the samples were then loaded onto duplicate SDS-PAGE gels on an equal protein basis. Gels with resolved proteins were either stained with Coommassie Blue or electroblotted to nitrocellulose at 25 V for 16 hours in transfer buffer containing no SDS (39 mM glycine, 48 mM Tris base, 20% methanol, pH 8.3). Blots were blocked for 2-4 hours at 4 °C with TTBS (0.5% Tween-20 in TBS: Tris-buffered saline, 20 mM Tris, 500 mM NaCl, pH 7.5). After blocking, the blots were incubated with CPR-peptide antiserum (1:2,000 dilution) in TTBS for 1-3 hours. Blots were rinsed twice for five minutes in TTBS before they were incubated in secondary antibody (goat anti-rabbit IgG alkaline phosphatase; Sigma) in TTBS at a concentration of 1:20,000 for 1 hour. Blots were rinsed once in TTBS for five minutes, and once in TBS for five minutes and were then developed in carbonate buffer (pH 9.8) with the addition of 66 μL NBT (Nitro blue tetrazolium, 0.5g in 10 mL 70% dimethylformamide) and 33 μL BCIP (5-bromo-4-chloro-3-indolylphosphate, 0.5g in 10 mL 100% dimethylformamide).

Sucrose gradient fractionation of the microsomes was carried out as described by Coughlan et al. (1996) with some modifications. The resuspended microsomes were loaded onto a preformed sucrose gradient (2 mL step of 20%, mass/vol., sucrose, over a 10 mL linear gradient of 30-60% sucrose (mass/vol.). The tubes were loaded onto a SW41

rotor and centrifuged for 2 hours at 25,000 rpm (112,700 x g) in a Beckman L8-80 ultracentrifuge. One mL fractions were collected from the bottom of the tube at 4 °C and analyzed for protein content as described above. Microsomes were separated into soluble and membrane fractions according to Fujiki et al. (1982) by incubating in ice-cold 100 mM sodium carbonate buffer (pH 11.5), followed by centrifugation at 229,600 x g in a Ti70.1 rotor. The membrane pellet was resuspended in Buffer B and the proteins in the supernatant were precipitated with TCA.

NADPH-cytochrome c reductase assays

NADPH-cytochrome c reductase activity in the microsomes and in the sucrose gradient fractions was determined spectrophotometrically as described by Madyastha and Coscia (1979) with some modifications. The increase in absorbance at 550 nm was monitored at 25 °C, in 250 mM Tris-Cl pH 7.5, 50 μM cytochrome c from bovine heart, 100 μM NADPH, 5 μM FMN, 5 μM FAD and 100 μM KCN (to inhibit cytochrome c oxidase activity from possible mitochondrial contamination) in a total reaction volume of 1 mL. The reaction was initiated by the addition of the microsomal extract or sucrose gradient fraction and was monitored for 3 minutes. An extinction coefficient of 28 $mM^{-1}cm^{-1}$ for bovine heart cytochrome c was used for the calculations.

Immunofluorescence localization

The immunolocalization procedure was done as described previously (Forward et al., submitted 1998). Root tips of 5 day old Douglas-fir seedlings were fixed for 1-2 hours at room temperature with slow agitation in 1% formaldehyde (Polysciences 16% stock,

Warrington, Pennsylvania, USA), 0.1% glutaraldehyde (Polysciences 50% stock) in 20 mM Pipes (1,4- piperazinediethanesulfonic acid), then washed 3 times over 1 hour in 50 mM Pipes containing 5 mM EGTA and 2 mM $MgSO_4$. Samples were dehydrated to 100% ethanol stepwise at 20 minute intervals, 10, 25, and 50% at 0 °C and 75, 95, 100% (3 changes) at -20 °C. Samples were embedded according to previously published methods (Baskins et al. 1992). To illustrate quality of preservation, some sections were stained with toluidine blue (1% w/v in 1% borax solution).

Sections attached to slides were immersed in pure acetone for 15 min to remove resin and then rehydrated in decreasing ratios of ethanol: 0.85% NaCl (100%, 80%, 50%, 30%, 1 minute each), transferred to PBS (20 minutes), PBS plus 1% BSA (10 minutes), and then into primary CPR-peptide antibody diluted 1:3000 with PBS or preimmune serum overnight at 4 °C followed by 1 hour at 37 °C. Sections were washed 3 times with PBS (20 minutes each) and then treated with secondary antibody, anti-rat IgM Cy3-conjugate (Jackson, West Grove, Pennsylvania, USA) diluted 1:400 for 1 hour at 37 °C. Sections were rinsed thoroughly with PBS, treated with DAPI (1 μg/ml) to stain nuclei and chromosomes and mounted in FluorSave (Calbiochem).

Results

Isolation and characterization of the Douglas-fir CPR cDNAs

CPR is an essential component of the P450 monooxygenase system (Figure 4.1) and is thought to be encoded by a single gene in mammals and yeast (Simmons et al., 1985; Sutter and Loper, 1989), whereas some plants have multiple CPR genes (Rosco et al., 1997; Koopmann and Hahlbrock, 1997; Mizutani and Ohta, 1998). In order to isolate a full length CPR cDNA clone, a partial length cDNA clone (*PM14*) that encodes a portion of a CPR (Tranbarger and Misra, 1995) was used to to rescreen a Douglas-fir cDNA library constructed from 4-6-day-old seedling mRNA. After a primary and secondary screening of six plates containing 10^{-4} pfu per plate, three putative positive clones were found to hybridize strongly with the [α-^{32}P]dCTP labelled *PM14* cDNA. Based on the restriction pattern of the putative CPR plasmids with *Eco*RI followed by Southern blot analysis of the plasmids with *PM14*, two of the clones were within the range of a full length CPR clone. Partial DNA sequence analysis of the three putative CPR cDNAs revealed 100% sequence identity within a portion of the clones that overlapped the *PM14* cDNA. These results suggest that only one *CPR* (designated *cpr1*) is expressed in 4-6-day-old Douglas-fir seedlings.

The complete DNA sequences of one of the largest clones *(PM5)* and *PM14* were determined and are identical for the 1130 bases where they overlap (Figure 4.2). The nucleotide sequences were confirmed in both directions using primers designed from the putative CPR clones (Table 4.1) and a summary of the features of the CPR clones is shown in Figure 4.3. The *PM14/PM5* hybrid sequence is a total length of 2779 bp, close to the 3 kb estimate for the CPR transcript previously reported (Tranbarger and Misra, 1995). Clone *PM5* is 2631 bp in length, overlaps the *PM14* clone from nucleotide 1650 to

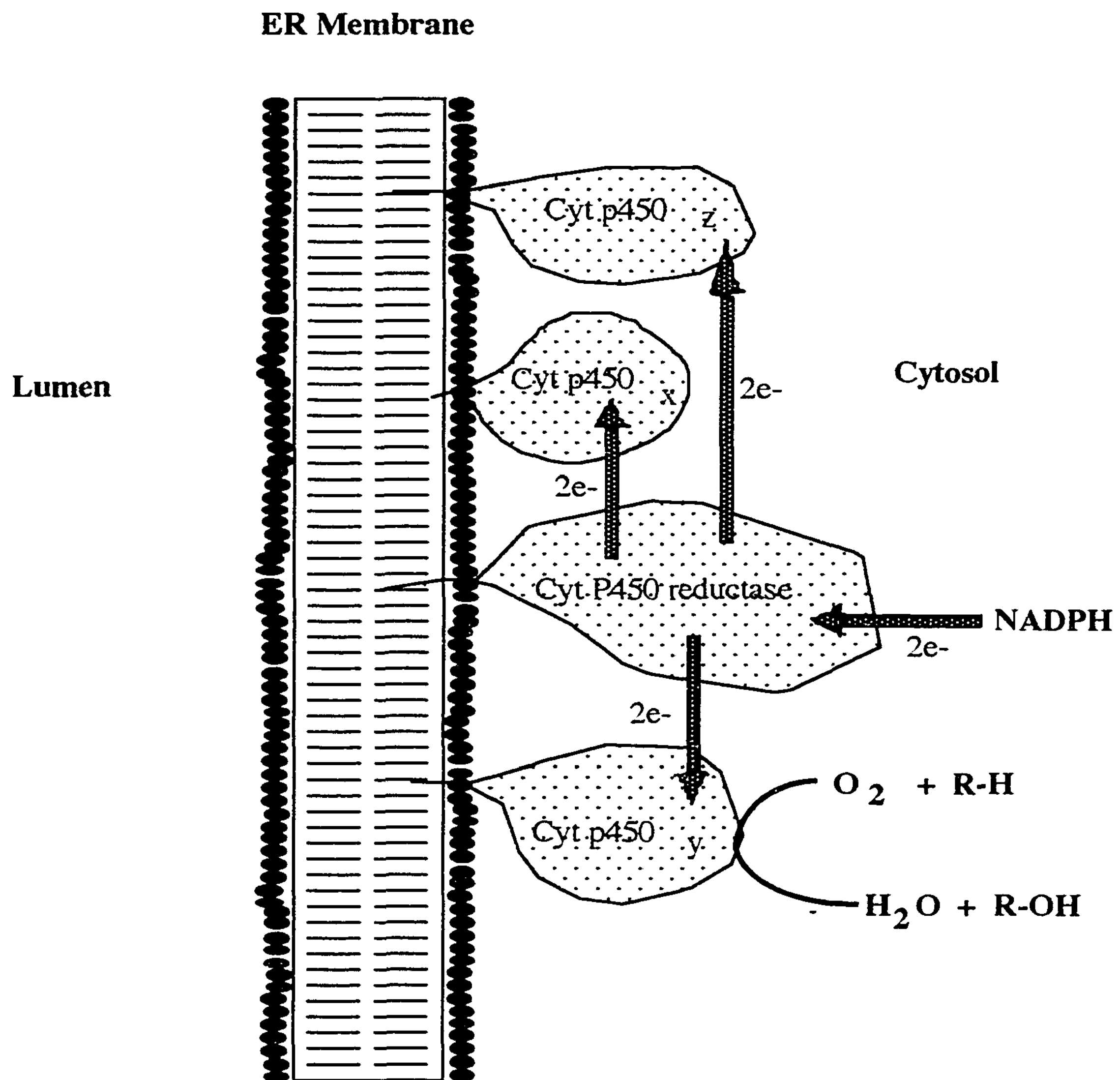


Figure 4.1. The cytochrome P450 monooxygenase system of eukaryotes. The figure was adapted from Donaldson and Luster, 1991.

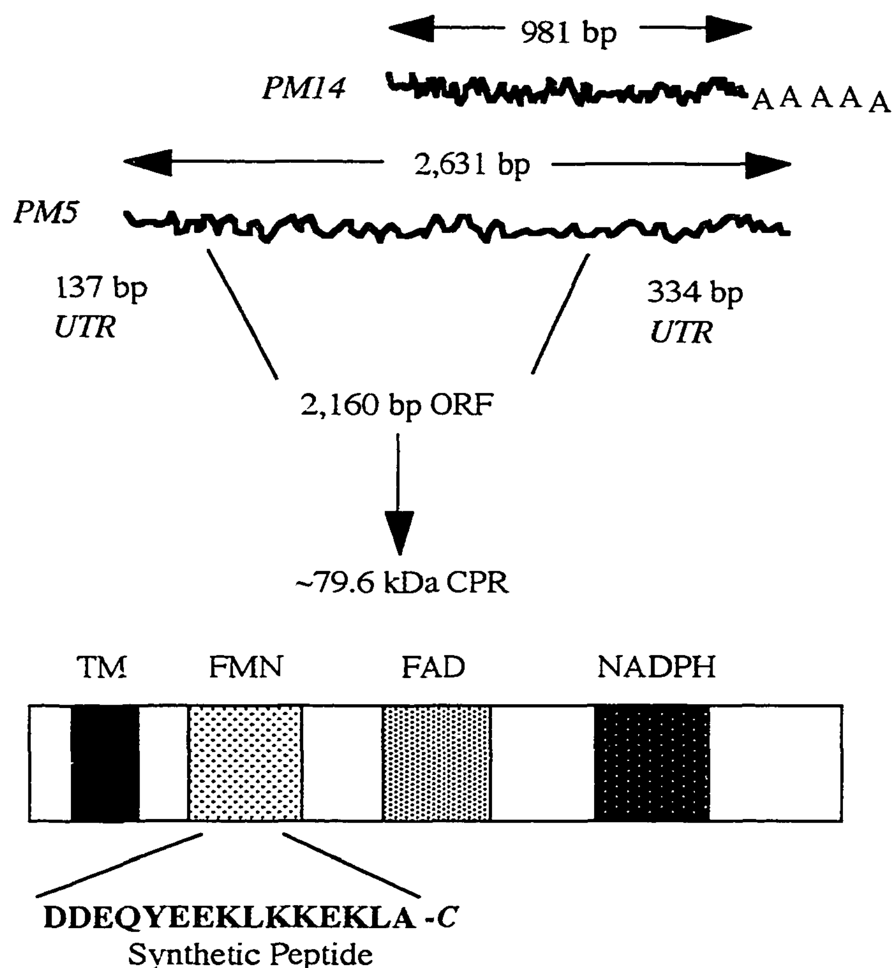
Figure 4.2. The nucleotide sequences of the cDNAs *PM5* and *PM14* and the deduced Douglas-fir CPR (PMCPR) amino acid sequence. The amino acid sequence is shown in single letter code below the nucleotide sequence. The three methionine residues representing alternative translation start sites are underlined. The potential *N*-glycosylation site is underlined and italisized. The serine phosphorylation consensus sequence is bracketed and italisized. The amino acid sequence used for the synthetic peptide is bracketed The putative polyadenylation signal sequence in the 3' *UTR* is underlined. The trans-membrane sequence is in parenthesis. The stop codon is represented by (*). Numbering of the nucleotide sequence is in relation to the beginning of the *PM5* cDNA sequence.

Figure 4.2.

```
tttgaggttgacgaagttccggtgacaataaaactcccccgtcttagcaggcaggtcaca   60
ggtcacaggttcgaatcgaaaacccgagagagatcgtgtttctcgtctccttaccctggt  120
tttcttgattcgtccgaatgcagtctaattcgaacaccgtggagagtggcacccagggct  180
                   M  Q  S  N  S  N  T  V  E  S  G  T  Q  G  L   15
tgggcaagagcttgctggagatatctcggcttgaccatgtttttcccttttcgtgaatg  240
 G  K  S  L  L  E  I  S  R  L  D  H  V  F  S  L  F  V  N  G   35
gcaagggggatctagggggaatgatggagtcggcggtcgttttgagggagaattcggagc  300
 K  G  D  L  G  G  M  M  E  S  A  V  V  L  R  E  N  S  E (L   55
tgttgatggtcttgactacttctctggccgtcttgattgcttgcgttttgctctttgttt  360
 L  M  V  L  T  T  S  L  A  V  L  I  A  C  V  L  L  F  V  W)  75
ggcggagaggaggatctgctccctcgaagctaccggagaagccgacgcctctggggagag  420
[R  R  G  G  S] A  P  S  K  L  P  E  K  P  T  P  L  G  R  V   95
tagaggaagaggaagaggaggaggatgattcggggaaaaagaaagtaacggttttctttg  480
 E  E  E  E  E  E  E  D  D  S  G  K  K  K  V  T  V  F  F  G  115
ggactcagaccgggacggctgaaggatttgccaaggctctggttgaggaagccaaagcta  540
 T  Q  T  G  T  A  E  G  F  A  K  A  L  V  E  E  A  K  A  R  135
ggtatgacaaggctgtctttaaagtagtggatctggatgattatgcagcagatgatgaac  600
 Y  D  K  A  V  F  K  V  V  D  L  D  D  Y  A  A  D  D  E [Q  155
agtatgaagaaaagttaaagaaggaaaaattagcatttttatgctcgcaacgtatggag  660
 Y  E  E  K  L  K  K  E  K  L  A] F  F  M  L  A  T  Y  G  D  175
atggagagccaactgacaatgcagcaaggttttataagtggttttctgagacggaaaaag  720
 G  E  P  T  D  N  A  A  R  F  Y  K  W  F  S  E  T  E  K  E  195
agagagggccttggctttccaaccttccgtacggggtatttggtttaggcaatagacagt  780
 R  G  P  W  L  S  N  L  P  Y  G  V  F  G  L  G  N  R  Q  Y  215
atgagcatttcaataaggttgcttgcacagtagatgaagcattaaatgagcaaggtggga  840
 E  H  F  N  K  V  A  C  T  V  D  E  A  L  N  E  Q  G  G  K  235
aacgtcttgttccagtgggccttggtgatgatgatcagtgcattgaagatgactttactg  900
 R  L  V  P  V  G  L  G  D  D  D  Q  C  I  E  D  D  F  T  A  255
cttggcgagagcaggtttggcctgagctagatcagttactgcgcgacgaagatgatcagc  960
 W  R  E  Q  V  W  P  E  L  D  Q  L  L  R  D  E  D  D  Q  P  275
cttctagtgctacaccttatacagcggctatacctcagtaccgggtagaaatccatgatc 1020
 S  S  A  T  P  Y  T  A  A  I  P  Q  Y  R  V  E  I  H  D  P  295
ctccagttccgctttacgaagaaacttatgcaacgcagaatggtcaagcctttattgata 1080
 P  V  P  L  Y  E  E  T  Y  A  T  Q  N  G  Q  A  F  I  D  I  315
tccatcacccttgccgagcaaatgtagcggtgcaaaggaacttcatacgccattgtctg 1140
 H  H  P  C  R  A  N  V  A  V  Q  R  E  L  H  T  P  L  S  D  335
accgttcttgcatccatttagaatttgatatctcaggcactggcctaacgtacgagactg 1200
 R  S  C  I  H  L  E  F  D  I  S  G  T  G  L  T  Y  E  T  G  355
gagatcatgttggtgtctatgcagaaaacttaagcgataccgtggaagaagctgcaaaat 1260
 D  H  V  G  V  Y  A  E  N  L  S  D  T  V  E  E  A  A  K  L  375
tgttaggctacccactagacacaataatttcagtacatagtgacaaagaagatggtacgc 1320
 L  G  Y  P  L  D  T  I  I  S  V  H  S  D  K  E  D  G  T  P  395
ctcttggtggaagctctttgctaccaccattccctggtccatgtaccctgcagactgctc 1380
 L  G  G  S  S  L  L  P  P  F  P  G  P  C  T  L  Q  T  A  L  415
ttgctcgttatgcagatttgttgaatcctccgcgaaaggctgccttggttgcattggcga 1440
 A  R  Y  A  D  L  L  N  P  P  R  K  A  A  L  V  A  L  A  S  435
gtcatgcatctgatccagctgaagcagagaggttgaagttcctttcatcaccagccggaa 1500
 H  A  S  D  P  A  E  A  E  R  L  K  F  L  S  S  P  A  G  K  455
aggatgagtactcccaatggattactgtgagtcaaagaagtcttttagaagtaatggcag 1560
 D  E  Y  S  Q  W  I  T  V  S  Q  R  S  L  L  E  V  M  A  E  475
aatttccatcagcaaaaccacctcttggtgtcttctttgcagctgttgccccgcgtctgc 1620
 F  P  S  A  K  P  P  L  G  V  F  F  A  A  V  A  P  R  L  Q  495
aacctcgatattattctatttcttcctctcccaggtatgcatccaatagaatacatgtca 1680
 P  R  Y  Y  S  I  S  S  S  P  R  Y  A  S  N  R  I  H  V  T  515
catgtgctctggtttatgggcccagtccaactggcagaattcacaaaggtgtttgttcca 1740
 C  A  L  V  Y  G  P  S  P  T  G  R  I  H  K  G  V  C  S  N  535
attggatgaagaattcagtgccttccgagaaaagccagtattgtagctgggcaccagtct 1800
 W  M  K  N  S  V  P  S  E  K  S  Q  Y  C  S  W  A  P  V  F  555
ttgttagacagtcaaatttcaaattgccatcggatccttcagttcctattgttatggtgg 1860
 V  R  Q  S  N  F  K  L  P  S  D  P  S  V  P  I  V  M  V  G  575
gccctggaactggtttagcacctttagaggcttttgcaggaaagagctgcaattcaaa 1920
 P  G  T  G  L  A  P  F  R  G  F  L  Q  E  R  A  A  I  Q  K  595
aatctggagagaagcttgggccagctgttcttttctttggatgcaggaatcgccaaatgg 1980
 S  G  E  K  L  G  P  A  V  L  F  F  G  C  R  N  R  Q  M  D  615
attacatttatgaagatgaactgaagagctatgtggaaaatggtgtattgactgagcttg 2040
 Y  I  Y  E  D  E  L  K  S  Y  V  E  N  G  V  L  T  E  L  V  635
ttcttgctttctctcgcgaaggagcaaccaaagagtatgtgcagcacaagattacagaga 2100
 L  A  F  S  R  E  G  A  T  K  E  Y  V  Q  H  K  I  T  E  K  655
agggatcctatatctggaatctgattgctcaaggtggctatctctatgtatgtggtgatg 2160
 G  S  Y  I  W  N  L  I  A  Q  G  G  Y  L  Y  V  C  G  D  A  675
ccaagggcatggctagggatgtgcacaggacactacatagtatagtccaagagcaggaat 2220
 K  G  M  A  R  D  V  H  R  T  L  H  S  I  V  Q  E  Q  E  S  695
ccgtggatagcaccagcgcagaggctacagtgaagaaattacagacagaaggaagatatt 2280
 V  D  S  T  S  A  E  A  T  V  K  K  L  Q  T  E  G  R  Y  L  715
tacgagatgtatggtaattgagtgaatttgttctaggtcagttcaatctcatctcgtgca 2340
 R  D  V  W  *                                                719
gattggtaagaagatcgaggaaaagagaaagcctagccaagtatttatttgggaatgaaa 2400
tctcttggcaaactaatgcttaagttcatttacccacggaaatgttgggattctttacac 2460
cttcaaatactgcattccaggatttgctatcactgtggccattcgaattttctagcccca 2520
gtataacattagtactggtctagaattgtattcttaattcatttgtgtctaagagtttag 2580
taccagcaaagttaaatatcatgcctttgggttagaagacaactctccacgtcatagaat 2640
ttttaatatatcttggtatgatctttgcataccatacttgacctacttctcattatctat 2700
atgggaaagggttcctttccagtttgatgttgtgtttctcttccttctataatgtcatat 2760
aagaatcgatactggattgaaaaaaaaaaaaaaaaaaaaaaaa                  2779
```

Table 4.1. Primers synthesized for sequencing the *CPR* cDNA clones. Primers were synthesized as described in the Materials and Methods section.

Name	Nucleotide Sequence (5'-3')	Position and Direction Within *CPR/PM5* Nucleotide Sequence
$5A_5$	TCGGAGCTGTTGATGGTCTT	(294->313)
$5A_{5rev}$	AGAGGCGTCGGCTTCTCC	(412<-395)
$5A_{7rev}$	AGACAGCCTTGTCATACCTAGC	(558<-537)
$5A_6$	ATGCTCGCAACGTATGGAGA	(642->661)
$5A_{6rev}$	AATGCTCATACTGTCTATTGCCTA	(789<-766)
$5A_3$	AGATCAGTTACTGCGCGACG	(929->948)
$3\text{-}1A_{1rev}$	CAATAAAGGCTTGACCATTCTGC	(1077<-1055)
$3\text{-}1A_2$	GACTGGAGATCATGTTGGTGTCT	(1196->1218)
$5A_{2rev}$	AATGGTGGTAGCAAAGAGCTTC	(1351<-1330)
$5A_7$	CCCAATGGATTACTGTGAGTCA	(1513->1534)
$14B_{2rev}$	TGACATGTATTCTATTGGATGCATA	(1680<-1656)
$14B_1$	CTGGAGAGAAGCTTGGGCCAGC	(1924->1945)
$14B_{1rev}$	GCTGGCCCAAGCTTCTCTCCAG	(1945<-1924)
$5A_{8rev}$	TAATCTTGTGCTGCACATACTC	(2094<-2073)
$5A_4$	GGTCAGTTCAATCTCATCTCGTG	(2316->2338)
$5A_{4rev}$	CACGAGATGAGATTGAACTGACC	(2338<-2316)
$14A_1$	GCATGATATTTAACTTTGCTGGTACT	(2604<-2579)

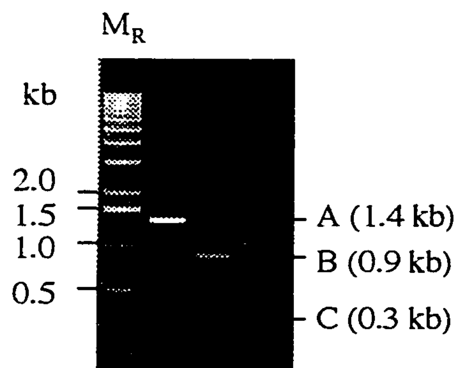


TM =Transmembrane region
FMN = Flavin Mononucleotide binding region
FAD = Flavin Adenine Dinucleotide binding region
NADPH = Nicotinamide Adenine Dinucleotide Phosphate binding region

Figure 4.3. Overview of the *CPR* cDNA clones. The amino acids of the synthetic peptide are indicated in single letter code. The predicted molecular mass of CPR is in kilodaltons (kDa) and the nucleotide sequence lengths are in base pairs (bp).

nucleotide 2631 and does not include a poly(A)$^+$ 3' end (Figure 4.2 and 4.3). Clone *PM14* has an additional 148 bp in the 3' untranslated region which contains a putative polyadenylation sequence ATAAGA 15 nucleotides upstream of the site of poly(A) addition. The overlapped *PM5/PM14* sequence contains an ORF of 2160 nucleotides with 5' and 3' untranslated flanking regions of 137 and 334 nucleotides respectively. Clone *PM5* has three *Eco*RI restriction sites that result in four fragments, three of which can be resolved on an agarose gel and visualized by staining with ethidium bromide (Figure 4.4). The longest ORF of *PM5* predicts a protein of 720 amino acids with a molecular mass of 79.6 kDa, and a pI of 5.09. There are three additional in-frame start codons that would result in proteins of 75.5, 75.3 and 73.8 kDa with pIs of 5.07, 5.07, and 5.12 respectively (Figure 4.2). It is assumed that the first AUG is the site of translation initiation because none of the other AUG start site flanking sequences had a higher probability for being used, based on a comparison with the consensus sequences compiled for higher plants (Table 4.2; Joshi et al., 1997). The predicted amino acid sequence of *PM5* contains a putative serine phosphorylation consensus region at Arg^{76}, and a glycosylation site at Asn^{364}.

The Douglas-fir CPR (designated PMCPR) has a hydrophobic N-terminal membrane-binding domain and the hyrophilic C-terminal catalytic domain common to all known CPRs. The catalytic domain shares conserved amino acids with predicted CPR sequences from plants, mammals and yeast that correspond to the NADPH, FAD and FMN cofactor binding sites (Figure 4.5) inferred from the three-dimensional structure (Wang et al., 1997). The PMCPR shares the highest identity with other plant CPRs (*V. radiata*, VR, 70%), (*A. thaliana*, AT2, 67%), (*C. roseus*, CR, 67%), and (*A. thaliana*, AT1, 66%),



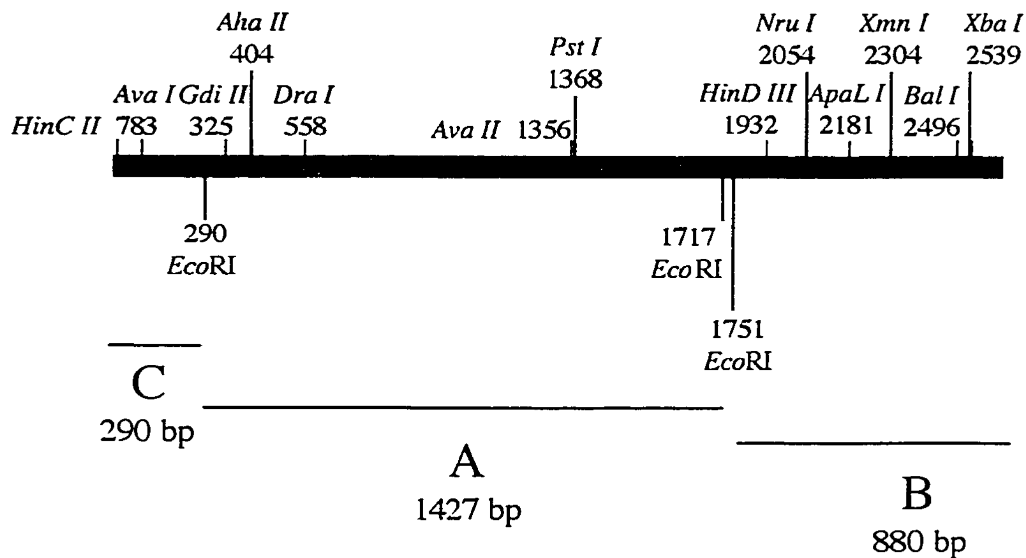


Figure 4.4. An agarose gel of purified *Eco*RI derived *PM5* fragments and restriction map of unique sites.

Table 4.2. Start codon consensus sequences of the *CPR* cDNA. The nucleotides of the *CPR* cDNA consensus sequences surrounding each putative start codon were compared with the consensus sequences compiled for non-plant species (Non-Plant), dicot species (Di), monocot species (Mono), lower plants (LP), and higher plants (HP), and identical nucleotides were scored (Joshi et al., 1997).

ATG	Consensus Sequence	Non-Plant	Di	Mono	LP	HP
#1	ATTCGTCCGA**ATG**CAGTCTAATT	7	6	5	6	8
#2	TCTAGGGGGA**ATG**ATGGAGTCGG	7	3	7	8	4
#3	AGGGGGAATGA**ATG**GAGTCGGCGG	8	6	12	10	7
#4	GGAGCTGTTGA**ATG**GTCTTGACTA	2	4	8	3	6

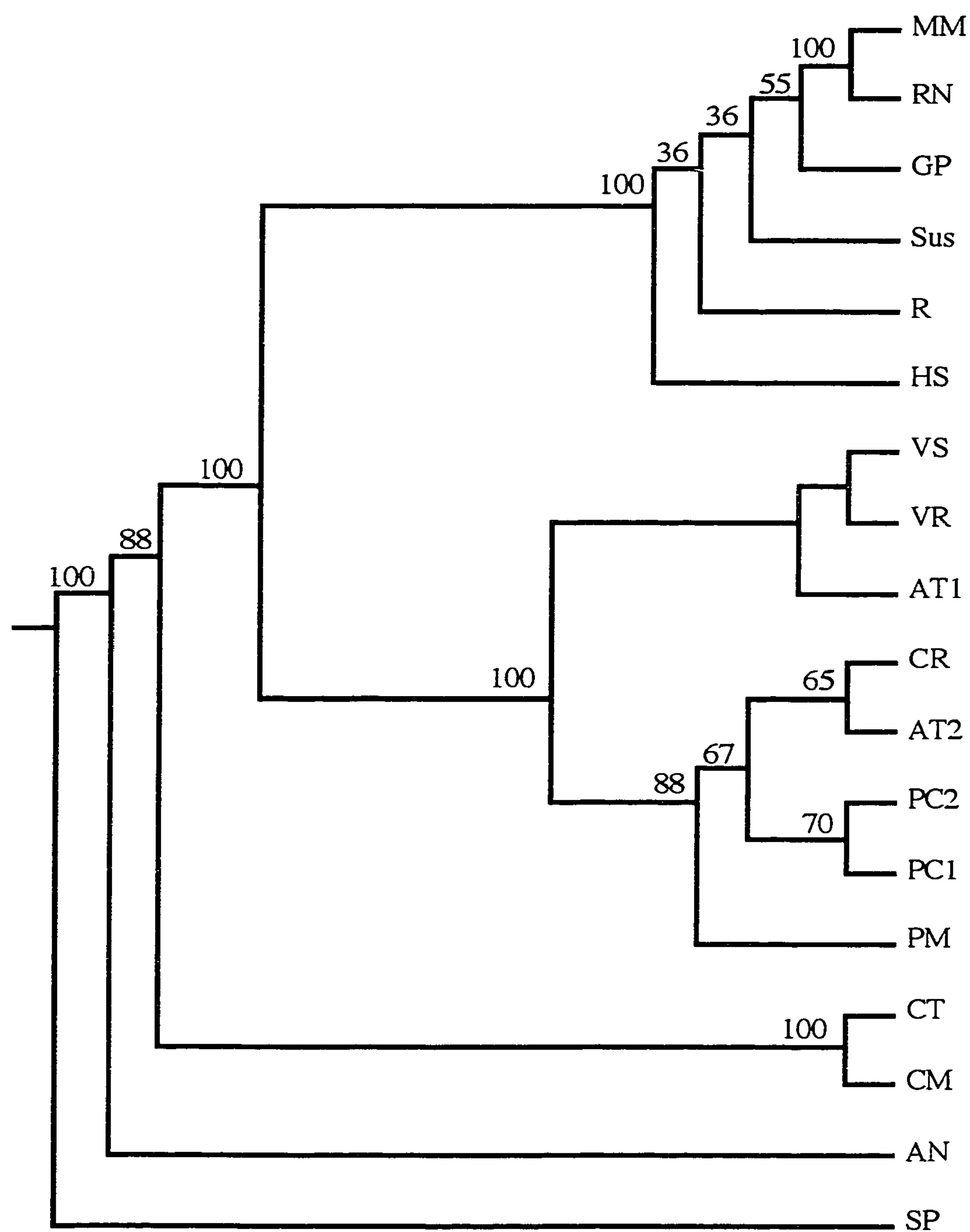
Figure 4.5. An alignment of representative plant, animal and fungal CPR amino acid sequences. An identical amino acid is indicated by (*), and a conserved amino acid is indicated by (:) for a strong group or (.) for a weak group as defined under Materials and Methods section. The following CPR amino acid sequences were aligned with CLUSTAL W (Thompson et al., 1994): PM (PMCPR, *Pseudotsuga menziesii*, this report); CR (*Catharanthus roseus*, X69791); VR, (*Vigna radiata*, A47298); AT (AT1, *Arabidopsis thaliana*, X66016); RN (rat, *Rattus norvegicus*, M10068); MM (mouse, *Mus musculus*, D17571); HS (human, A60557); CM (*Candida maltosa*, G436053).

Figure 4.5.

```
PM  MQSNSNTVESGTQGLGKSLLEISRLDHVFSLFVNGKGDLGGHMESAVVLREN--SELLMVLTTSLAVLIACVLLFVWRRG
CR  -------MDSSSEKLS-PFEIMSAILKGAKLDGSNSSDSGVAVSPAVMAMLLENKELVMILTTSVAVLIGCVVVLIWRRS
VR  ---------------------MAS---NSDLVRAVESFLGVSLGDSVS------DSLLLIATTSAAVVVG-LLVFLWKKS
AT  ---------------------MTSALYASDLFKQLKSIMGT---DSLSD-----DVVLVIATTSLALVAG-FVVLLWKKT
RN  ------------------------------MGDSHEDTSATMPEAVA------EEVSLFSTTDMVLFSLIVGVLTYWFI
MM  ------------------------------MGDSHEDTSATVPEAVA------EEVSLFSTTDIVLFSLIVGVLTYWFI
HS  ------------------------------MGDSHVDTSSTVSEAVA-------EEVSLFSMTDMILFSLIVGLLTYWFL
CM  ---------------------------------------------------------MALFKLDLYVIIVLAVAVAAYFAK
                                                          : : :.          . :  :

PM  -GSAPSKLPE--KPTPLGRVEEEEEE--EDDSGKKKVTVFFGTQTGTAEGFAKALVEEAKARYDKAVFKVVDLDDYAADD
CR  SGSGKKVVEPPKLIVPKSVVEPEEID-----EGKKKFTIFFGTQTGTAEGFAKALAEEAKARYEKAVIKVIDIDDYAADD
VR  -SDRSKEVKP--VVVPRDLMMEEEEE-VDVAAGKTKVTIFFGTQTGTAEGFAKALAEEIKARYEKAAVKVVDLDDYAADD
AT  TADRSGELKP--LMIPKSLMAKDEDDDLDLGSGKTRVSIFFGTQTGTAEGFAKALSEEIKARYEKAAVKVIDLDDYAADD
RN  FRKKKEEIPEFSKIQTTAPPVKESSFVEKMKKTGRNIIVFYGSQTGTAEEFANRLSKDAH-RYGMRGMS-ADPEEYDLAD
MM  FKKKKEEIPEFSKIQTTAPPVKESSFVEKMKKTGRNIIVFYGSQTGTAEEFANRLSKDAH-RYGMRGMS-ADPEEYDLAD
HS  FRKKKEEVPEFTKIQTLTSSVRESSFVEKMKKTGRNIIVFYGSQTGTAEEFANRLSKDAH-RYGMRGMS-ADPEEYDLAD
CM  -NQFLDEPQDTGFLSNDTAGGNSRDILETLKKNNKNTLLLFGSQTGTAEDYANKLSREIHSRFGLKTMV-ADFADYDWDN
     .           .           .  :::*:****** :*: * .: : *:     .   *  :*   :

PM  EQYEEKLKKEKLAFFMLATYGDGEPTDNAARFYKWFSETEKERGPWLSNLTYGVFGLGNRQYEHFNKVACTVDEALNEQG
CR  EEYEEKFRKETLAFFILATYGDGEPTDNAARFYKWFVE-GNDRGDWLKNLQYGVFGLGNRQYEHFNKIAKVVDEKVAEQG
VR  DLYEEKLKKESLVFFMLATYGDGEPIDNAARFYKWFTEGKDERGIWLQKLTYGVFGLGNRQYEHFNKIGKVVDEELAEQG
AT  DQYEEKLKKETLAFFCVATYGDGEPTDNAARFSKWFTE-ENERDIKLQQLAYGVFALGNRQYEHFNKIGIVLDEELCKKG
RN  LSSLPEIDK-SLVVFCMATYGEGDPTDNAQDFYDWLQE----TDVDLTGVKFAVFGLGNKTYEHFNAMGKYVDQRLEQLG
MM  LSSLPEIDK-SLVVFCMATYGEGDPTDNAQDFYDWLQE----TDVDLTGVKFAVFGLGNKTYEHFNAMGKYVDQRLEQLG
HS  LSSLPEIDN-ALVVFCMATYGEGDPTDNAQDFYDWLQE----TDVDLSGVKFAVFGLGNKTYEHFNAMGKYVDKRLEQLG
CM  --FGDIPND-ILVFFIVATYGEGEPTDNADEFHTWLTD----EADTLSTLRYTVFGLGNSTYEFYNAIGRKFDRVFEEKG
          .  *..* :*****:*:* ***  *  *: :          *  : : **.***  **.:* :.  .*.. :*

PM  GKRLVPVGLGDD-DQCIEDDFTAWREQVWPELDQLLRDEDDQPSSATPYTAAIPQYR-VEIHD-PPVPLYEETYATQ---
CR  GKRIVPLVLGDD-DQCIEDDFAAWRENVWPELDNLLRDEDDT-TVSTTYTAAIPEYR-VVFPD-KSDSLISEANGHANG-
VR  AKRLVAVGLGDD-DQSIEDDFSAWKESLWSELDQLLRDEDDANTVSTPYTAAILEYR-VVIHD-PTAASTYDNHSTVA--
AT  AKRLIEVGLGDD-DQSIEDDFNAWKESLWSELDKLLKDEDDK-SVATPYTAVIPEYR-VVTHD-PRFTTQKSMESNVA--
RN  AQRIFELGLGDD-DGNLEEDFITWREQFWPAVCEFFGVEATG----E--ESSIRQYELVVHEDMDVAKVYTGEMGRLKS-
MM  AQRIFELGLGDD-DGNLEEDFITWREQFWPAVCEFFGVEATG----E--ESSIRQYELVVHEDMDTAKVYTGEMGRLKS-
HS  AQRIFELGLGDD-DGNLEEDFITWREQFWPAVCEHFGVEATG----E--ESSIRQYELVVHTDIDAAKVYMGEMGRLKS-
CM  GERFADYGEGDDGTGTLDEDFLTWKDNVFDTLKNDLNFEERE-LKYEPNVKLTERDDLTVD-D-SEVSLGEPNKKYIQSE
    .:*:         ***     :::** :*::..:  : : :  *                     .   :   *

PM  ---NGQAFIDIHHPCRANVAVQRELHTPLSDRSCIHLEFDISGTGLTYETGDHVGVYAENLSDTVEEAAKLLGYPLDTII
CR  -YANGNTVYDAQHPCRSNVAVRKELHTPASDRSCTHLDFDIAGTGLSYGTGDHVGVYCDNLSETVEEAERLLNLPPETYF
VR  ---NGNTEFDIHHPCRVNVAVQKELHKPESDRSCIHLEFDISGTSITYDTGDHVGVYAENCNETVEETGKLLGQNLDLFF
AT  ---NGNTTIDIHHPCRVDVAVQKELHTHESDRSCIHLEFDISRTGITYETGDHVGVYAENHVEIVEEAGKLLGHSLDLVF
RN  -YENQKPPFDAKNPFLAAVTANRKLN-QGTERHLMHLELDISDSKIRYESGDHVAVYPANDSALVNQIGEILGADLDVIM
MM  -YENQKPPFDAKNPFLAAVTTNRKLN-QGTERHLMHLELDISDSKIRYESGDHVAVYPANDSTLVNQIGEILGADLDVIM
HS  -YENQKPPFDAKNPFLAAVTTNRKLN-QGTERHLMHLELDISDSKIRYESGDHVAVYPANDSALVNQLGKILGADLDVVM
CM  EIDLTKGPFDHTHPYLAKISKTRELF-ASKERNCVHVEFDVSESNLKYTTGDHLAVWPSNSDENIAKFIKCFGLQ-DKIN

PM  SVHSDKEDGTPLGGSSLLPPFPGPCTLQTALARYADLLNPPRKAALVALASHASDPAEAERLKFLSSPAG--KDEYSQWI
CR  SLHADKEDGTPLAGSSLPPPFP-PCTLRTALTRYADLLNTPKKSALLALAAYASDPNEADRLKYLASPAG--KDEYAQSL
VR  SLHTDKDDGTSLGG-SLLPPFPGPCSLRTALARYADLLNPPRKAALLALATHASEPSD-ERLKFLSSPQG--KDEYSKWV
AT  SIHADKEDGSPLES-AVPPPFPGPCTLGTGLARYADLLNPPRKSALVALAAYATEPSEAEKLKHLTSPDG--KDEYSQWI
RN  SLN-NLDEESNKKH-----PFPCPTTYRTALTYYLDITNPPRTNVLYELAQYASEPSEQEHLHKMASSSGEGKELYLSWV
MM  SLN-NLDEESNKKH-----PFPCPTTYRTALTYYLDITNPPRTNVLYELAQYASEPSEQEHLHKMASSSGEGKELYLSWV
HS  SLN-NLDEESNKKH-----PFPCPTSYRTALTYYLDITNPPRTNVLYELAQYASEPSEQELLRKMASSSGEGKELYLSWV
CM  TVFELKALDSTYQI-----PFPNPITYGAVVRHHLEISGPVSRQFFLAIAGFA--PDE--ETKKTFTRIGNDKQEFANKI

PM  TVSQRSLLEVMAEFPSAKPPLGV---FFAAVAPRLQPRYYSISSSPRYASNRIHVTCALVYGPSPTGRIHKGVCSNWMKN
CR  VANQRSLLEVMAEFPSAKPPLGV---FFAAIAPRLQPRFYSISSSPRMAPSRIHVTCALVYEKTPGGRIHKGVCSTWMKN
VR  VGSQRSLVEVMAEFPSAKPPLGV---FFAAIAPRLQPRYYSISSSPRFAPQRVHVTCALVYGPTPTGRIHKGVCSTWMKN
AT  VASQRSLLEVMAAFPSAKPPLGV---FFAAIAPRLQPRYYSISSCQDWAPSRVHVTSALVYGPTPTGRIHKGVCSTWMKN
RN  VEARRHILAILQDYPSLRPPID----HLCELLPRLQARYYSIASSSKVHPNSVHICAVAVEYEAKSGRVNKGVATSWLRA
MM  VEARRHILAILQDYPSLRPPID----HLCELLPRLQARYYSIASSSKVHPNSVHICAVAVEYEAKSGRVNKGVATSWLRT
HS  VEARRHILAILQDCPSLRPPID----HLCELLPRLQARYYSIASSSKVHPNSVHICAVVVEYETKAGRINKGVATNWLRA
CM  TRKKLNVADALLFASNGRPWSDVPFEFIIEMVPHLQPRYYSISSSSLSEKQTINITAVVEVEEEADGRAVTGVVTNLLKN

PM  SVP----SEK--------SHDCS-----WAPVFVRQSNFKLPSDPSVPIVMVGPGTGLAPFRGFLQERAAIQKSGEKLGP
CR  AIP----LEE--------SRDCS-----WAPIFVRQSNFKLPADPKVPVIMIGPGTGLAPFRGFLQERLALKEEGAELGT
VR  AIP----SEK--------SQDCS-----SAPIFIRPSNFKLPVDHSIPIIMVGPGTGLAPFRGFLQERYALKEDGVQLGP
AT  AVP----AEK--------SHECS-----GAPIFIRASNFKLPSNPSTPIVMVGPGTGLAPFRGFLQERMALKEDGEELGS
RN  KEP----AGE--------NGGRA-----LVPMFVRKSQFRLPFKSTTPVIMVGPGTGIAPFMGFIQERAWLREQGKEVGE
MM  KEP----AGE--------NGRRA-----LVPMFVRKSQFRLPFKPTTPVIMVGPGTGVAPFMGFIQERAWLREQGKEVGE
HS  KEP----AGE--------NGGRA-----LVPMFVRKSQFRLPFKATTPVIMVGPGTGVAPFIGFIQERAWLRQQGKEVGE
CM  IEIEQNKTGEKPVVHYDLSGPRNKFNKFKLPVHVRRSNFKLPKNTTTPVILIGPGTGVAPLRGFVRERVQQVKNGVNVGK

PM  AVLFFGCRNRQMDYIYEDELKSYVEN-GVLTELVLAFSREG-ATKEYVQHKITEKGSYIWNLI-AQGGYLYVCGDAKGMA
CR  AVFFFGCRNRKMDYIYEDELNHFLEI-GALSELLVAFSREG-PTKQYVQHKMAEKASDIWRMI-SDGAYVYVCGDAKGMA
VR  ALLFFGCRNRQMDFIYEDELKSFVEQ-GSLSELIVAFSREG-AEKEYVQHKMMDKAAHLWSLI-SQGGYLYVCGDAKGMA
AT  SLLFFGCRNRQMDFIYEDELNNFVDQ-GVISELIMAFSREG-AQKEYVQHKMMEKAAQVWDLI-KEEGYLYVCGDAKGMA
RN  TLLYYGCRRSDEDYLYREELARFHKD-GALTQLNVAFSREQ-AHKVYVQHLLKRDREHLWKLIHEGGAHIYVCGDARNMA
MM  TLLYYGCRRSDEDYLYREELARFHKD-GALTQLNVAFSREQ-AHKVYVQHLLKRDKEHLWKLIHEGGAHIYVCGDARNMA
HS  TLLYYGCRRSDEDYLYREELAQFHRD-GALTQLNVAFSREQ-SHKVYVQHLLKQDREHLWKLI-EGGAHIYVCGDARNMA
CM  TVLPYGCRNEHDDFLYKKEWSEYASVLGENFEMFTAFSRQDPSKKVYVQDKIAENSKLVNDLL-NEGAIIYVCGDASRMA

PM  RDVHRTLHSIVQEQESVDSTSAEATVKKLQTEGRYLRDVW-
CR  RDVHRTLHTIAQEQGSMDSTQAEGFVKNLQMTGRYLRDVW-
VR  RDVHRTLHSIVQEQENVDSTKAEAIVKKLQMDGRYLRDVW-
AT  RDVHRTLHTIVQEQEGVSSSEAEAIVKKLQTEGRYLRDVW-
RN  KDVQNTFYDIVAEFGPMEHTQAVDYVKKLMTKGRYSLDVWS
MM  KDVQNTFYDIVAEFGPMEHTQAVDYVKKLMTKGRYSLDVWS
HS  RDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS
CM  RDVQSTIAKIVAKHREIQEDKAVDLVKSWKVQNRYQEDVW-
```

Figure 4.6. Phylogenetic tree of representative CPR predicted amino acid sequences from mammals, yeast and higher plants. The following sequences were aligned with CLUSTAL W and used to construct the phylogenetic tree using the PHYLIP package (Felsenstein, 1989): PM (PMCPR, *Pseudotsuga menziesii*, this report); SP (*Schizosaccharomyces pombe*, X64702); AN (*Aspergillus niger*, Z26938); CT (*Candida tropicalis*, M35199); CM (*Candida maltosa*, G436053); R (Rabbit, *Oryctolagus cuniculus*, A25505); HS (human, A60557); GP (guinea pig *Cavia porcellus*, D10498); MM (mouse, *Mus musculus*, D17571); RN (rat, *Rattus norvegicus*, M10068); PC1 (parsely, *Petroselinum crispum*, AF0204634); PC2 (parsely, *Petroselinum crispum*, AF024635); CR (*Catharanthus roseus*, X69791); AT1 (*Arabidopsis thaliana*, X66016); AT2 (*Arabidopsis thaliana*, X66017); VS (*Vicia sativa*, Z26252); VR (*Vigna radiata*, A47298). Numbers at branches are bootstrap values out of 100.

Figure 4.6.



```
                                                            oooooooooooooooooooo
PM   MQSNSNTVESGTQGLGKSLLEISRLDHVFSLFVNGKGDLGGMMESAVVLRENSELLMVLTTSLAVLIACVLLFVWRRGGSAPS---KLPEKPTPLGRVEEE
CR   MDSSSEKLSP-FE-LMSAILKGAKLDGSNSSDSGV--AVSPAV--MAMLLENKELVMILTTSVAVLIGCVVVLIWRRSSGS-GK--KVVEPPKLIVPKSVV
AT2  MSSSSSSSTSMID-LMAAIIKGEPVIVSDPANASAYESVAAELSS--MLIENRQFAMIVTTSIAVLIGCIVMLVWRRSGSG---NSKRVEPLKPLVIKP--
AT1  MTS--------------ALYASDLFKQLKSIM--GTDSLSDDV----------VLVIATTSLALVAG-FVVLLWKKTTADRSGELKPLMIPKSLMAKDED
VS   MTS--------------S--NSDLVRTIESAL--GI-SLGDSVSDS--------VVIIATTSAAVIIG-LLVFLWRKSP-DRSRELRPVIVPKFTV-KHED
VR   MAS-----------------NSDLVRAVESFL--GV-SLGDSVSDS--------LLLIATTSAAVVVG-LLVFLWKKSS-DRSKEVKPVVVPRDLM-MEEE
PC2  MGGESLATS-------LPATL-----------------------------LENRDLLMLLTTSIAVLIGCAVVLVWRRSSL---RSVKSVEPPKLIVPKVEI
PC1  MQSESMEVSP-VD-LLASILKIDSVES------------------MTLLLENRDVLMLLTTSFAVLIGLGLVMMWRRSTTM-TKSAKKLEPAKIVIPKFEM
     *  .                                                 .. *** *.. .   .. .*..                .
```

Figure 4.7. An alignment of the divergent amino terminal sequences of plant CPRs. The methionine residues representing alternative translation start sites are underlined. The trans-membrane sequence is indicated by (o). An identical amino acid is indicated by (*) and a conserved amino acid is indicated by (.). The following CPR amino acid sequences were aligned with CLUSTAL V: PM (PMCPR, *Pseudotsuga menziesii*, this report); CR (*Catharanthus roseus*, X69791); AT1 (*Arabidopsis thaliana*, X66016); AT2 (*Arabidopsis thaliana*, X66017); VS (*Vicia sativa*, Z26252); VR (*Vigna radiata*, A47298); PC1 (parsely, *Petroselinum crispum*, AF0204634); PC2 (parsely, *Petroselinum crispum*, AF024635).

lower identity with mammalian CPRs (MM, *Mus musculus*, 37%), (RN, rat, *Rattus norvegicus*, 37%), and (HS, human, 37%), and the lowest identity with yeast CPR (CM, *Candida maltosa*, 31%). The N-terminal domain has the highest amount of divergence among the CPR sequences, whereas the C-terminal domain is the most conserved. Phylogenetic analysis of selected CPR predicted amino acid sequences from mammals, yeast and plants segregates PMCPR within the plant CPR subgroup (Figure 4.6). Within the plant CPR subgroup, there appears to be two distinct subtypes of CPRs as suggested previously (Urban et al., 1997), although the two parsley CPRs do not segregate to different groups and PMCPR does not clearly fall into either group. The predicted trans-membrane sequence contains the highest amount of similarity within the amino terminus of the plant CPRs (Figure 4.7). Seven out of eight of the CPR sequences have additional methionine residues that could be alternative sites of translation initiation. Many of the methionine residues occur upstream of the trans-membrane segment.

Developmental and tissue regulated *cpr1* expression

To determine the expression of *cpr1* during Douglas-fir seed development, germination and early seedling growth, developing seeds were collected monthly and mature Douglas-fir seeds were germinated for two weeks as described (Chapter 3, Materials and Methods). Equal amounts of total RNA were loaded per lane and visualized by staining with ethidium bromide prior to northern blot analysis (data not shown). The expression of *cpr1* was compared with the expression of the gene for the 18S rRNA. The MG and embryos or emerging seedlings were dissected from developing or germinating seeds, and total RNA was isolated and analyzed by northern hybridization using a region of the *PM5* cDNA containing the conserved FMN binding domain as a probe (Figure 4.8). Expression of

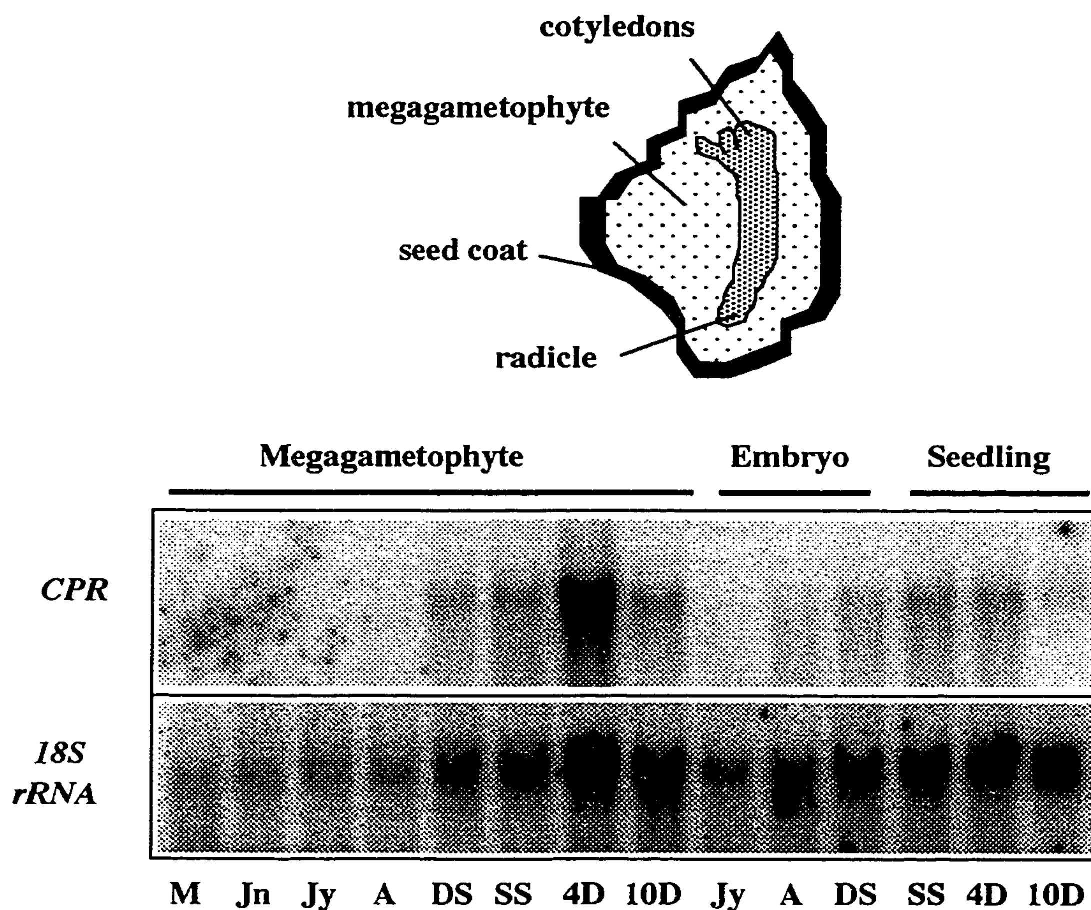


Figure 4.8. Northern blot analysis of *cpr* expression during embryogenesis and early seedling development. Total RNA (12 μg per lane) isolated from Douglas-fir tissues was analyzed as described in the Materials and Methods section (upper panel). Membranes were stripped and rehybridized with a genomic clone for a portion of the 18S rRNA (lower panel). May 31 (M), June 19 (J), July 17 (Jy), August 7 (A). stratified seeds (SS), mature seeds (DS), 4-day-old seedlings (4D), 10-day-old seedlings (10).

cpr1 was highest in the MG of 4-day-old seedlings. In contrast, only small amounts of *cpr1* transcripts in the embryo and none in the MG were detected during seed development, whereas both the MG and embryos from dry seeds contained detectable amounts of *cpr1* transcript. *Cpr1* transcripts were detectable in the embryos of developing seeds collected in August (during the later stages of seed development) and not in the MG, suggesting expression of CPR is initiated in the embryo. In 10-day-old seedling tissues, *cpr1* transcripts were highest in the MG, present in reduced amounts in the cotyledons and radicle, and not detected in the seed coat (Figure 4.9). *Cpr1* expression in 10-day-old seedlings did not correlate with the ability for a seed to germinate because amounts found in ungerminated seeds (seeds without emerging radicles kept under germination conditions for ten days) were similar to those in 10-day-old seedlings with different root lengths (Figure 4.9). The amount of *cpr1* transcript was similar in light- and dark-grown seedlings and transcripts were not detected in needles emerging in the spring (Figure 4.10). These results indicate that *cpr1* transcripts accumulate during the final stages of seed maturation, and expression is under both tissue and developmental regulation during germination and early seedling development, presumably at the transcriptional and/or post-transcriptional level.

The expression of *CPR* genes and the accumulation of CPR protein has been shown to be induced by wounding in *Helianthus tuberosus* (Lesot et al., 1995). Seeds were wounded immediately after stratification, germinated and *cpr1* transcript amounts were examined by northern blot analysis (Figure 4.11). *Cpr1* transcripts decreased by 1-hour after wounding and continued to be lower than controls for 73 hours. The 18S ribosomal RNA amount also decreased, with degradation intermediates accumulating by 24 hours after wounding, indicating a general decrease in transcription as a result of wounding. The expression of *cpr1* was not affected by salt or desiccation treatments during germination or early seedling

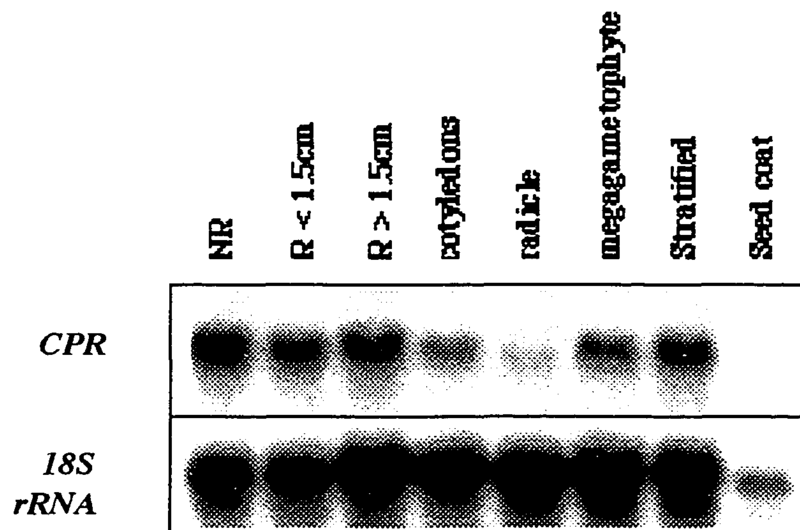


Figure 4.9. Expression of *cpr1* in 10-day-old seedling tissues. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds or from separated tissues as described in the Materials and Methods section. NR, no radicle; R, radicle visible with lengths less than (<) or greater than (>) 1.5 cm; Stratified, stratified seeds. Upper and lower panels are as described in the legend for Figure 4.8. Twenty μg of total RNA was loaded per lane.

Figure 4.10. The expression of *cpr1* is not modulated by light during germination and early seedling development, and is not detected in emerging needles in the spring. Twenty μg of total RNA was loaded per lane isolated from light- (L) and dark-grown (D) 2-day-old (2D), 6-day-old seedlings (6D) and emerging needles from mature trees (EN) as described in the Materials and Methods section. Upper and lower panels are described in the legend for Figure 4.8. Twenty μg of total RNA was loaded per lane.

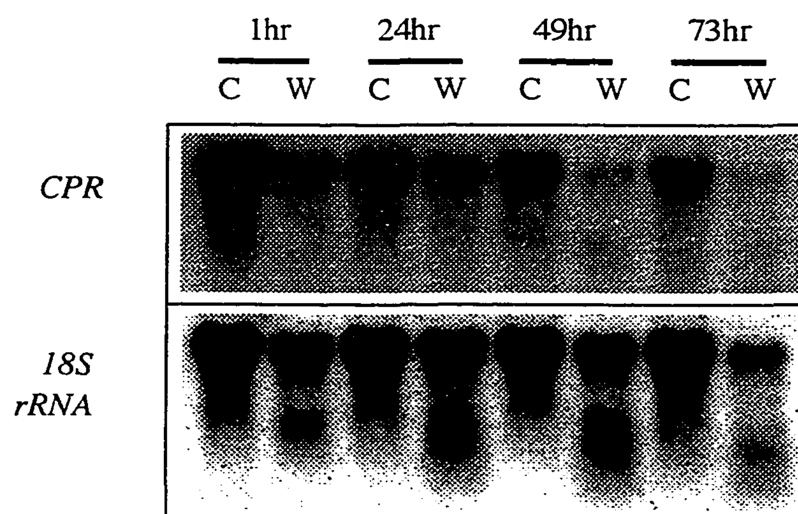


Figure 4.11. Northern blot analysis of *cpr1* expression in response to wounding. W, wounded seeds (pinched with forceps) and C, unwounded control seeds were incubated in water in petri plates in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hr. photoperiod for 1, 24, 49 and 73 hours. Panels are as described in the legend for Figure 4.8. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds as described in the Materials and Methods section. Twenty μg of total RNA was loaded per lane.

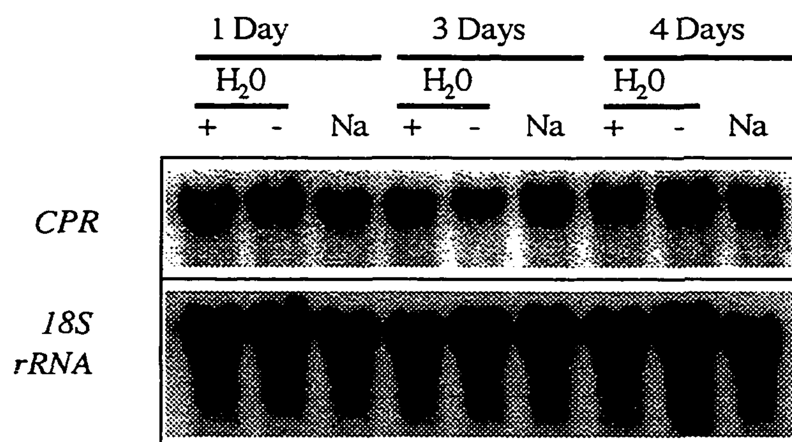


Figure 4.12. Northern blot analysis of *cpr1* expression in response to desiccation. Seeds were germinated in a controlled environmental incubator providing a day/night temperature of 30/20 °C and an 8-hr. photoperiod for 1, 3, and 4 days in petri plates with (+) and without (-) H_2O, and in an aqueous solution of 250 mM sodium chloride (Na). Panels are as described in the legend for Figure 4.8. Total RNA was isolated from 1 g (20-30 seeds or seedlings) of whole seeds and seedlings as described in the Materials and Methods section. Twenty μg of total RNA was loaded per lane.

development (Figure 4.12).

A CPR-peptide derived antiserum detects a single CPR in two-dimensional SDS-PAGE

To examine the CPR protein predicted by the *PM5* cDNA, a peptide based on a region of the predicted CPR amino acid sequence was synthesized, conjugated to the carrier protein KLH, and an antiserum was raised (Figure 4.4). The synthetic CPR-peptide contains conserved amino acids that are common to other plant CPRs and should be able to detect CPR isoforms (Figure 4.13). Microsomes from stratified seeds were isolated and microsomal proteins were separated by 2-dimensional SDS-PAGE and blotted to nitrocellulose for western blot analysis (Figure 4.14). The antiserum against the synthetic CPR-peptide detected a single polypeptide with a molecular mass between 70-80 kDa and a pI between 5-6. These values are comparable to those predicted by the *PM5* cDNA. Similarly, no isoforms of CPR were detected in mature and germinating seeds (data not shown). The amino acid sequence of the synthetic peptide used to raise the antiserum to the Douglas-fir CPR has identities that range from 45% to 85% to the same region in other plant CPRs (Figure 4.13). Given this relatively high amount of identity across plant species, the CPR-peptide antiserum should detect CPR isoforms. These results indicate that the CPR-peptide antiserum detects a single CPR in the microsomes of Douglas-fir seeds in the developmental stages examined.

PMCPR localizes with the endoplasmic reticulum and NADPH-cytochrome c activity in sucrose gradients

The presence of a hydrophobic trans-membrane sequence in the N-terminal region of PMCPR, similar to those found in other CPR polypeptides, indicates a membrane

```
PMCPR  DDEQYEEKLKKEKL
CR     DDEEYEEKFRKETL
AT2    DDDEYEEKLKKEDV
AT1    DDDQYEEKLKKETL
VS     DDDQYEEKLKKETL
VR     DDDLYEEKLKKESL
PC2    EDEEYEAKFKKESF
PC1    DDDEYETKLKKESM
       .*.  ** *..**
```

Figure 4.13. An alignment of the amino acid sequence used to raise the CPR peptide-antiserum with the corresponding sequences from other plants. The amino acids of the synthetic peptide correspond to amino acid numbers 155-166 of the predicted PMCPR amino acid sequence shown in Figure 4.2. The following plant CPR amino acid sequences were aligned with CLUSTAL V: PM (PMCPR, *Pseudotsuga menziesii*, this report); CR (*Catharanthus roseus*, X69791); AT1 (*Arabidopsis thaliana*, X66016); AT2 (*Arabidopsis thaliana*, X66017); VS (*Vicia sativa*, Z26252); VR (*Vigna radiata*, A47298); PC2 (parsely, *Petroselinum crispum*, AF024635); PC1 (parsely, *Petroselinum crispum*, AF0204634).

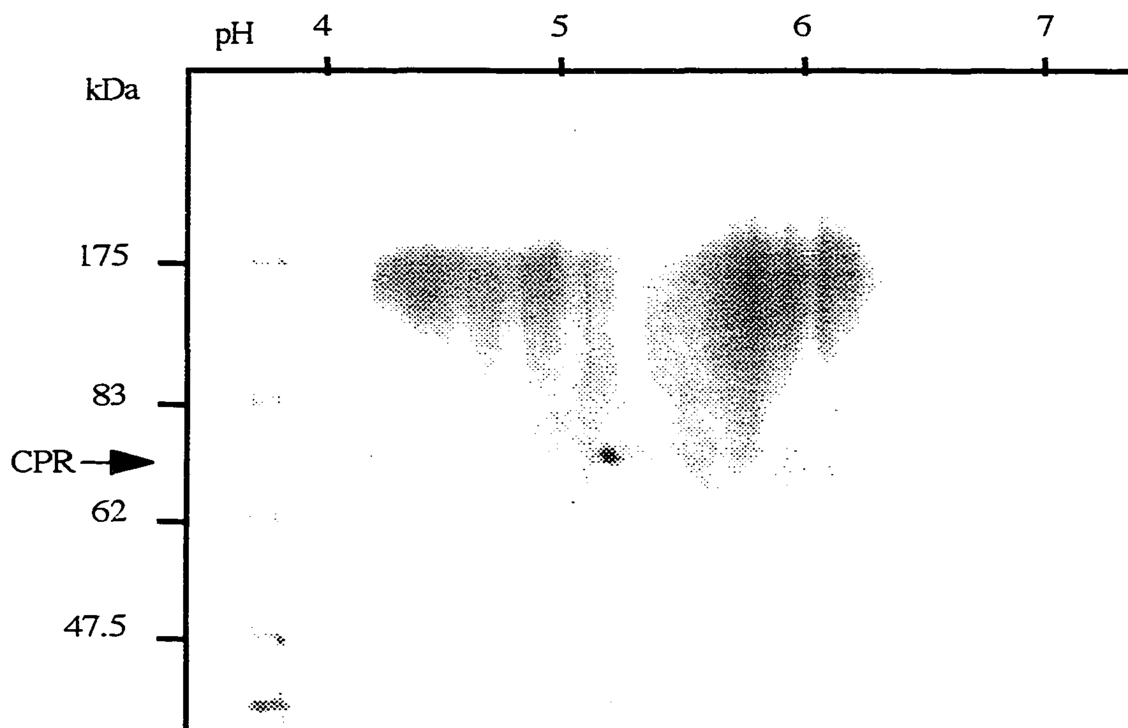


Figure 4.14. The CPR-peptide antiserum detects a single CPR in western blot analysis of microsomal proteins. Thirty μg of microsomal protein extracted from stratified seeds was resolved by 2-dimensional SDS-PAGE, electroblotted to nitrocellulose and analyzed using the CPR-peptide antiserum as described in the Materials and Methods section. Molecular mass markers are on the left in kilodaltons, and pH units are on top.

association (Figure 4.7). To determine the subcellular localization of PMCPR, microsomes were treated with a sodium carbonate buffer to separate the lumenal and integral membrane protein fractions (Fujiki et al., 1982; Figure 4.15). The fractions were then separated by SDS-PAGE and blotted to nitrocellulose for western blot analysis with the CPR-peptide antiserum. The results show PMCPR is enriched in the microsomal fraction. Twice as much protein from the soluble fraction (SN1, after the initial 100,000 x g centrifugation step) was loaded per lane in order for PMCPR to be detectable. When the microsomes were washed three successive times with sodium carbonate, PMCPR persisted in the membrane fraction. These results indicate that PMCPR is a membrane associated protein of the microsome. To determine the composition of the microsomes isolated from Douglas-fir seeds, the components of the microsomes were separated by sucrose gradient centrifugation (Coughlan et al., 1996; Robinson et al., 1994; Figure 4.16). There were two diffuse bands of protein visible in the sucrose gradient after centrifugation. A less apparent band corresponded to the 34% sucrose fraction, and a major band at the 20/30% sucrose interface that contained membranes derived from the endoplasmic reticulum (ER; Coughlan et al., 1996; Robinson et al., 1994). The CPR-peptide antiserum detected PMCPR at the 20/30% sucrose interface, indicating an association with the ER as found with mammalian CPRs (Figure 4.16, upper panel). The sucrose gradient fractions were assayed for NADPH-cytochrome c activity, a standard assay commonly used for the analysis of CPR activity and a marker for the ER (Madyasha et al., 1976; Robinson et al., 1994). The NADPH-cytochrome c reductase activity was also found exclusively in the 20/30% interface fraction, along with the PMCPR enzyme (Figure 4.16, lower panel). These results indicate the PMCPR polypeptide predicted by the *PM5* cDNA is an integral membrane protein that cofractionates in sucrose gradient fractions containing membranes

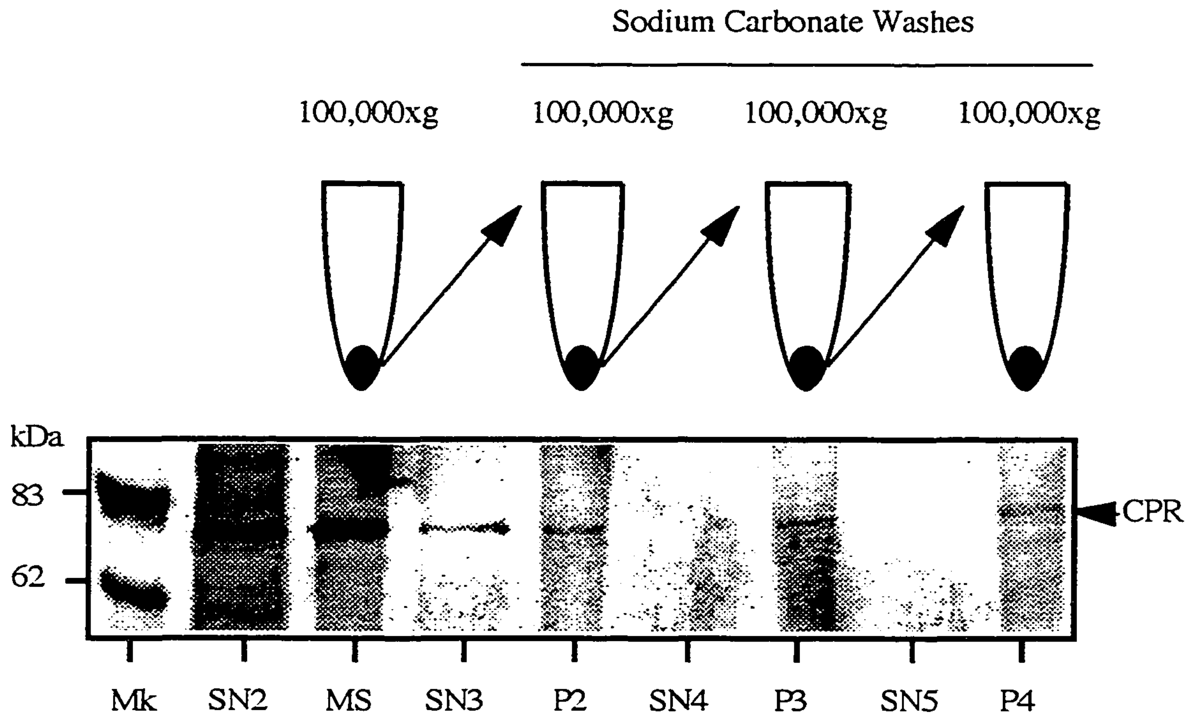


Figure 4.15. PMCPR is a microsomal membrane associated protein. Western blot analysis of soluble and membrane proteins from stratified seed microsomes were analyzed by SDS-PAGE. SN2 and MS are the supernatant and microsomal pellet, respectively, after initial 100,000 x g centrifugation. SN3, 4, and 5 and P2, 3, and 4 are supernatants and pellets, respectively, after microsomal pellet (P2) was diluted with 8 mL sodium carbonate buffer (pH 11.5) and centrifuged at 229,600 x g. Sixty and 30 μg of protein from SN2 and MS respectively was loaded per lane. Five μL of resuspended TCA precipitated proteins from SN3, 4 and 5 (resuspended in total of 10 μL Buffer B), and 5 μL of resuspended pellets, P2, 3 and 4 (resuspended in total of 10 μL Buffer B) were loaded per lane. Mk, Molecular mass markers in kilodaltons.

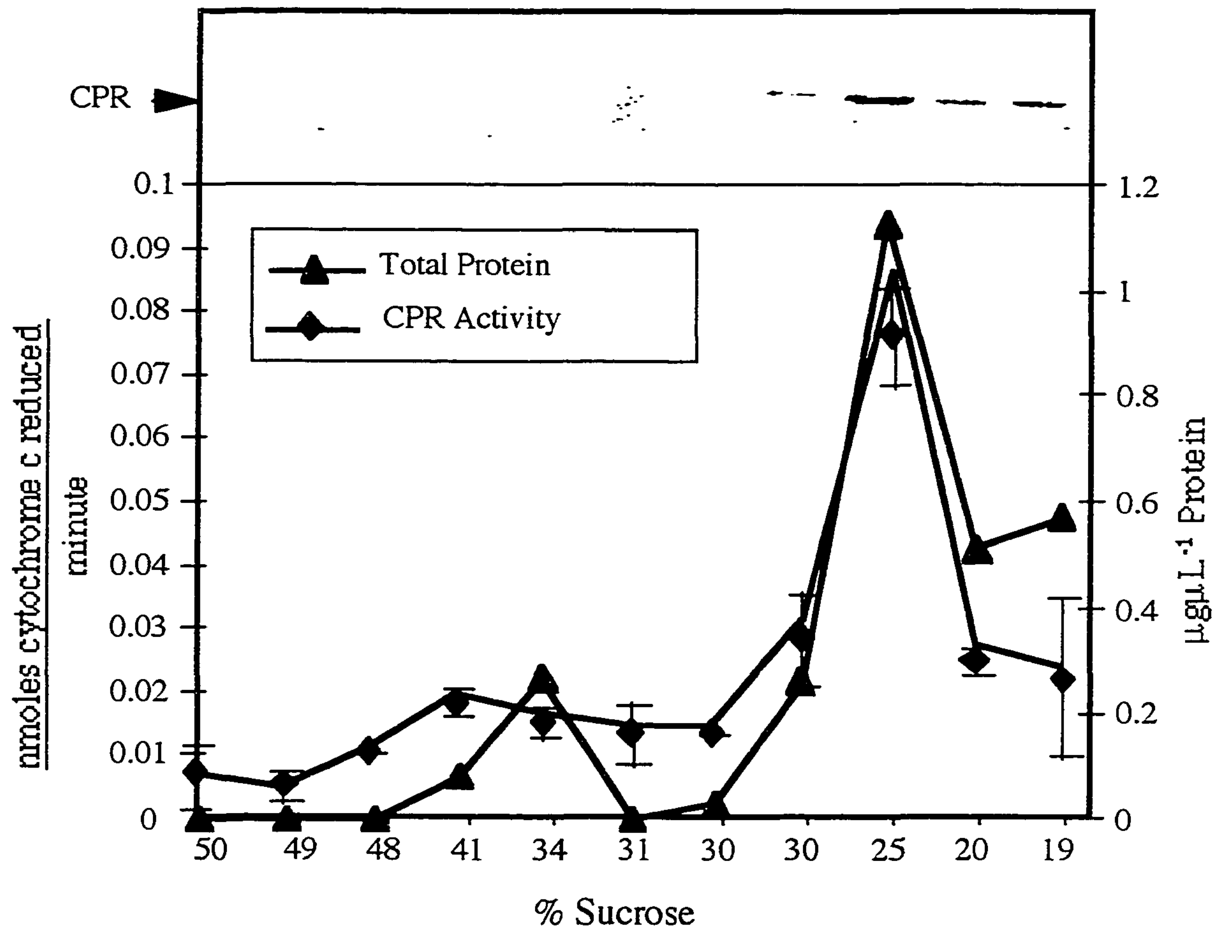


Figure 4.16. PMCPR is associated with the fraction containing the endoplasmic reticulum and NADPH-cytochrome c reductase activity. Sucrose gradient fractions were assayed for CPR content with the CPR-peptide antiserum by western blot analysis and for CPR activity spectrophotometrically by NADPH-cytochrome c reductase assays.

derived from the ER and NADPH-cytochrome c reductase activity.

Accumulation of PMCPR protein and activity in the microsomes during germination and early seedling development

To assess the regulation of *cpr1* gene expression, microsomal and soluble proteins from mature, stratified, and germinating seeds and seedlings were isolated, separated by SDS-PAGE, and blotted to nitrocellulose for western blot analysis (Figure 4.17). To ensure that an equal amount of protein was loaded in each lane, a duplicate gel was stained with Coomassie Blue and evaluated visually (data not shown). PMCPR enzyme was present in similar amounts in the microsomes isolated from mature, stratified and germinating seeds, and then declined in 14-day-old seedlings. The decline in PMCPR enzyme by 14 days appears to be the result of degradation because lower molecular mass proteins are appearing at the same time the PMCPR band intensity is decreasing. No PMCPR was detected in the supernatants throughout the time course when an equal amount of protein (30μg per lane) was loaded. The CPR-peptide antiserum detected PMCPR in the supernatant when 60 μg protein was loaded per lane (Figure 4.15, SN2). The pattern of PMCPR enzyme accumulation follows the constitutive expression of *cpr1* during the same development time course, which suggests PMCPR enzyme amounts are controlled at a level prior to translation. It is not known if the degradation of PMCPR seen during early seedling development is occurring *in vivo* and is a mechanism for regulating enzyme quantity.

A peak of NADPH:cytochrome c reductase activity was seen at 4 days that was approximately five times that seen in mature seeds (Figure 4.18). The 4 day peak activity of 9.7 nmoles cytochrome c reduced mg^{-1} protein min^{-1} was lower but within the same magnitude reported in the literature for microsomes from pig kidney (40 nmoles

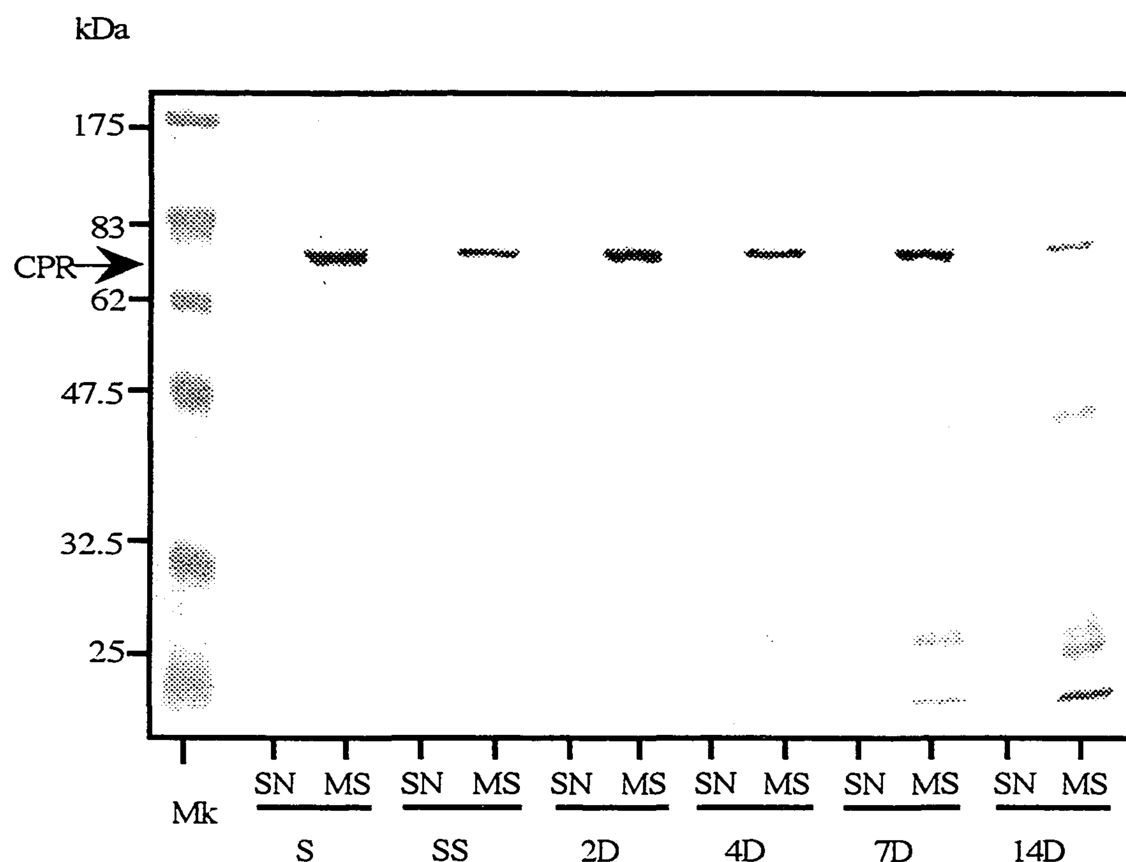


Figure 4.17. Western blot analysis of PMCPR protein levels during germination and early seedling development. Thirty μg of soluble and microsomal protein was loaded per lane. Two hundred μg of soluble and microsomal protein was assayed for activity. SN, supernatant; MS, microsomal fraction; S, mature seeds; SS, stratified seeds; 2D, 4D, 7D, and 14D, 2-, 4-, 7-, and 14-day-old seedlings respectively. Mk, molecular mass markers are on the left in kilodaltons.

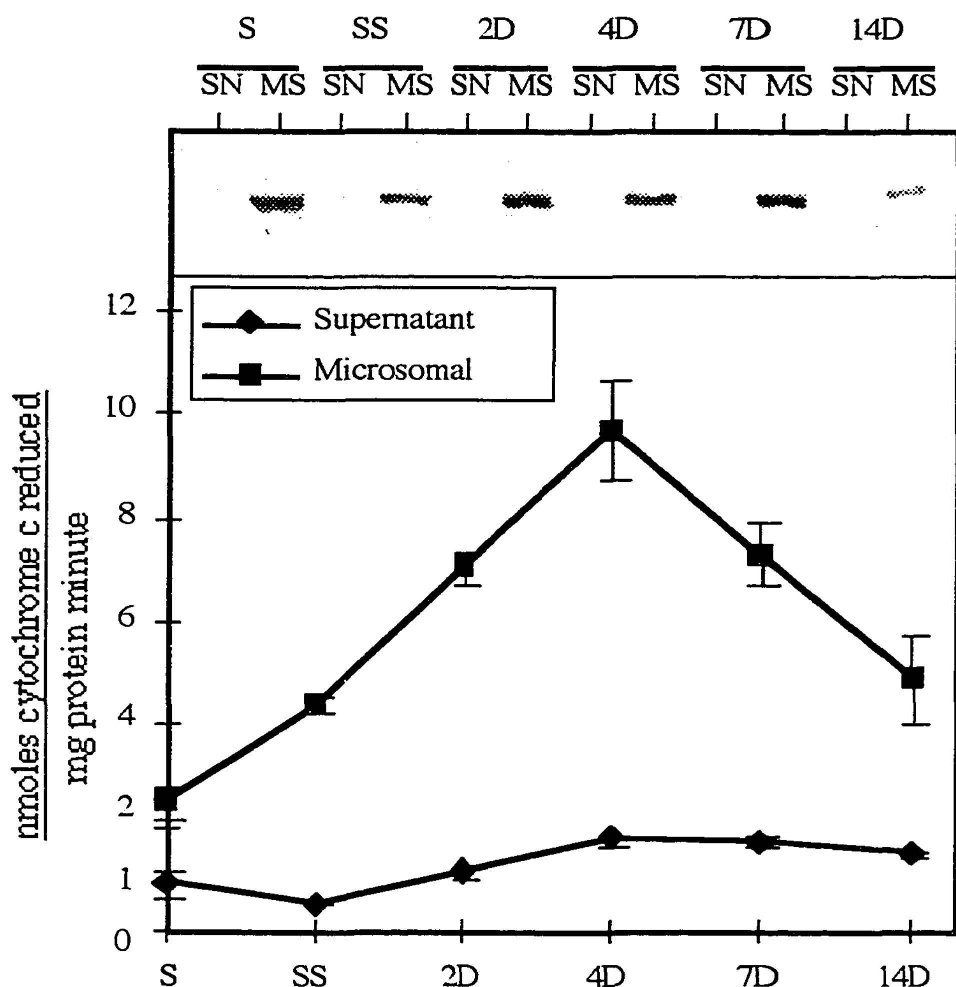


Figure 4.18. A comparison of CPR enzyme activity and PMCPR protein levels during germination and early seedling development. PMCPR enzyme is present in similar amounts (panel A), while CPR activity increases (panel B) in seeds and during germination and early seedling growth. Soluble and microsomal proteins extracted from mature, stratified and germinating seeds and developing seedlings were assayed for CPR content with the CPR-peptide antiserum by western blot analysis and for CPR activity spectrophotometrically by NADPH-cytochrome c reductase assays. Activity and PMCPR accumulation trends were repeated three times. SN, supernatant; MS, microsomal fraction; S, mature seeds; SS, stratified seeds; 2D, 4D, 7D, and 14D, 2-, 4-, 7-, and 14-day-old seedlings respectively.

cytochrome c reduced mg^{-1} protein min^{-1}, Yasukochi et al., 1980), sweet potato roots (34 nmoles cytochrome c reduced mg^{-1} protein min^{-1}, Fujita and Asahi, 1985), wounded Jerusalem-artichoke (*Helianthus tuberosus*) tubers (140 nmoles cytochrome c reduced mg^{-1} protein min^{-1}, Benveniste et al., 1986), tobacco cell suspensions (30 nmoles cytochrome c reduced mg^{-1} protein min^{-1}, Imaishi et al., 1995) spearmint glandular trichomes (222 nmoles cytochrome c reduced mg^{-1} protein min^{-1}, Ponnamperuma and Croteau, 1996) and chickpea cell suspensions (93 nmoles cytochrome c reduced mg^{-1} protein min^{-1}, Loerwald et al., 1996). Although the amount of CPR enzyme found in the mature seeds was similar to that found at 4 days, the activity was lower (Figure 4.18). Either another undetected CPR is responsible for this increase in activity, or CPR protein accumulates as an inactive or less active enzyme and is activated post-translationally during stratification and germination. The decline in the amount of NADPH-cytochrome c activity by 7-14 days follows the apparent degradation of PMCPR enzyme as detected by western blot analysis (Figures 4.17 and 4.18). Additional evidence for post-translation regulation comes from a comparison of *cpr1* transcript and PMCPR protein in 10-day-old seedling tissues (Figure 4.19). The lowest amount of *cpr1* transcript was found in the radicles, whereas the highest amount of PMCPR protein at this stage was detected in the radicles.

Immunofluorescence localization of PMCPR

In order to address the subcellular localization of PMCPR, a preliminary study was done on sections from 5-day-old Douglas-fir seedling radicles prepared for immunofluorescence microscopy. The CPR peptide-antiserum detected antigen that was within the cytosolic

portion of the cells (Figure 4.20 A). The staining pattern was not evenly spread throughout the cell, but at discrete points, often just outside of the nuclear envelope (Figure 4.20 A and B). No staining was seen with the preimmune serum.

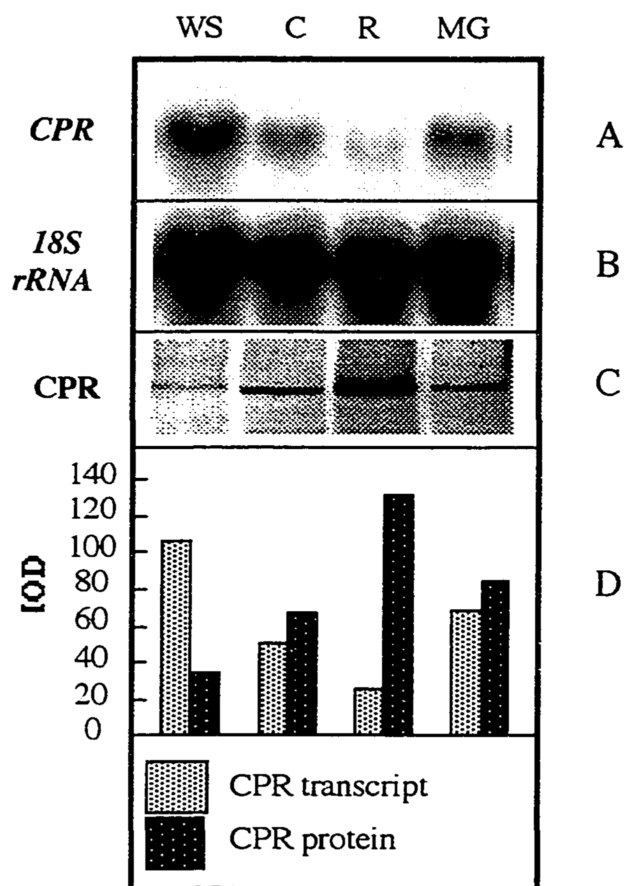


Figure 4.19. A comparison of *cpr1* transcript and PMCPR protein amounts in 10-day-old seedling tissues. Total RNA (20 μg per lane) isolated from Douglas-fir tissues was analyzed by hybridization with the full length *cpr/PM5* insert (panel A), stripped and rehybridized with a genomic clone for a portion of the 18S rRNA (panel B) as described in the Materials and Methods section. Total soluble proteins (30 μg per lane) isolated from Douglas-fir tissues were resolved by SDS-PAGE, blotted to nitrocellulose, and probed with the CPR peptide-antiserum as described in the Materials and Methods section (panel C). The CPR transcript and protein amounts shown in panels A and C were quantified by densitometry and the results are presented graphically (panel D). WS, whole mature seed; C, cotelydon; R, radicle; MG, megagametophyte.

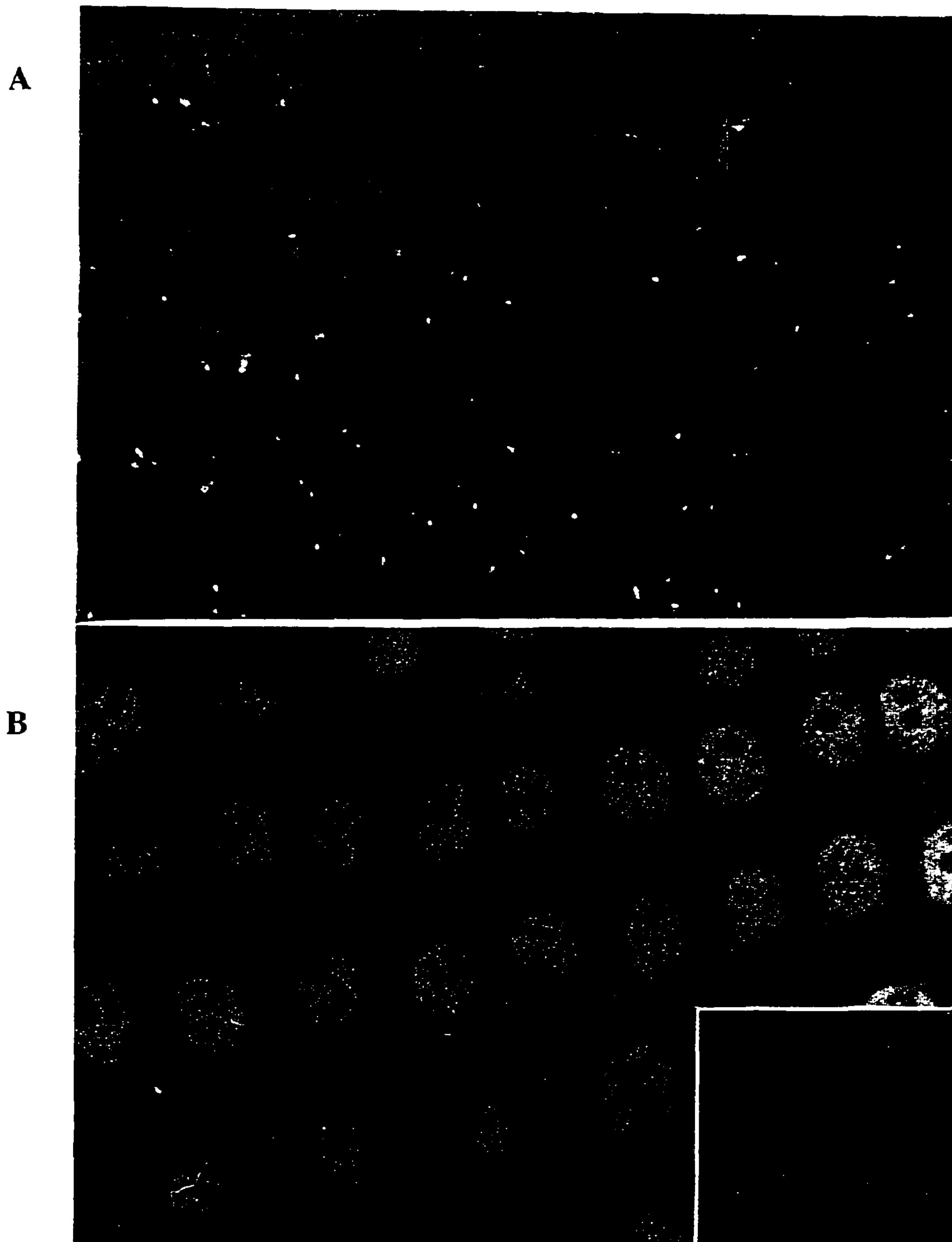


Figure 4.20. Immunolocalization of CPR in the developing radicles of 5-day-old seedlings. (A) Radicle sections were stained with the CPR peptide-antiserum, (B) with DAPI to visualize nuclei, and with preimmune (lower right hand corner) as a control.

Discussion

As an essential component of the P450 monooxygenase system of the ER, CPR is a multifaceted enzyme that associates with NADPH, FAD and FMN in order to transfer electrons to a multitude of individual P450 enzymes. In mammals and yeast this is carried out by a single CPR (Simmon et al., 1985; Sutter and Loper, 1989). Recent evidence clearly shows that some plants have CPR isoforms encoded by distinct genes which exhibit differential expression in response to wounding, UV, infection and elicitors (Koopman and Hahlbrock, 1997; Mizutani and Ohta, 1998, Lesot et al., 1995). To the best of my knowledge, this section describes the first cDNA encoding a CPR isolated from a gymnosperm. A partial cDNA clone previously isolated that contained the 3' *UTR* and a portion of the CPR coding sequence detected only a single type of CPR cDNA, in a cDNA library made from 4-6 day old Douglas-fir seedlings. A single CPR gene appears to be expressed during germination and early seedling development because only one CPR transcript size was detected, whereas two sizes were found in *Helianthus tuberosus* (Lesot et al., 1995). In addition, CPR expression patterns were similar irrespective of whether a probe corresponding to a conserved region in the ORF or the less conserved 3'UTR was used (data not shown). Furthermore, a CPR-peptide antiserum derived from a conserved region detected only a single polypeptide in soluble and microsomal fractions in all the developmental stages tested. Two-dimensional SDS-PAGE analysis of microsomal proteins from seeds and seedlings confirmed the presence of only one CPR. This is in contrast to biochemical and immunological evidence of multiple CPRs in some higher plants (Benveniste et al., 1989; Benveniste et al., 1991, Ponnamperuma and Croteau, 1996). One cannot rule out that there may be more than a single Douglas-fir CPR expressed in other tissues or developmental stages and it is also possible that the CPR-

peptide antiserum only detects a single CPR isoform. When the CPR-antiserum was tested against tobacco and *A. thaliana* protein extracts, only a single CPR band was detected (data not shown). However, based on the presence of a single transcript size in northern analysis, and the isolation of only one cDNA type from the cDNA library, *cpr1* appears to be the only *cpr* gene expressed at the stages of development examined. Therefore, the expression of the *PM5* gene results in the accumulation of PMCPR in mature and stratified seeds and during germination and early seedling development.

CPR represents a unique enzyme that is membrane associated where it interacts with different P450s, contains both FAD and FMN as cofactors and is thought to have evolved from the fusion of individual FAD and FMN flavoprotein genes (Porter et. al. 1990). The present results indicate that the gymnosperm CPR shares highly conserved regions for binding the flavin cofactors as do angiosperm, yeast and mammalian CPRs. In addition, the data suggest that higher plant CPRs have diverged from their mammalian and yeast counterparts especially within the N-terminal region containing the hydrophobic trans-membrane sequence, and have continued to diverge since the separation of angiosperms and gymnosperms. The questions remain what significance multiple CPR genes in some plants have, at what point during plant evolution did they appear, and whether differences in the N-terminal regions reflects significant functional divergence between plant, yeast and mammalian CPRs.

An interest in *cpr* expression stems from efforts to elucidate the regulatory mechanisms involved in the expression of genes during conifer germination and early seedling development. The P450 monooxygenase system is responsible for a number of essential reactions including fatty acid hydroxylation, secondary metabolite biosynthesis, herbicide detoxification, and gibberellin synthesis (reviewed by Donaldson and Luster, 1991;

Bolwell et al., 1994) during germination and early seedling development. Previously *cpr* expression was found to be constitutive during germination and early seedling development when total RNA extracted from whole seeds and seedlings was examined by northern blot analysis (Chapter 2; Tranbarger and Misra, 1995). In the present study it is shown that *cpr1* expression is developmentally regulated prior to the mature seed stage because transcripts are lower or undetectable during seed maturation and are present in mature seeds in both the MG and the seedling. In addition, *cpr1* expression is dependent on tissue type because transcripts are higher in the MG at 4 days than in the seedling, and vary in the cotyledons, radicle and MG from 10-day-old seedlings. Developmental and tissue level regulation of *cpr* expression have also been observed in *C. roseus*, parsley, and *A. thaliana* (Meijer et al., 1993; Koopmann and Hahlbrock, 1997; Mizutani and Ohta, 1998). The results described here support these previous observations indicating *cpr* expression in plants differs from mammals, where expression is under negligible developmental or tissue-specific control and the *cpr* promoter resembles that of house-keeping genes (Shephard et al., 1992; O'Leary et al., 1994).

Most studies to date have examined the transcriptional regulation of plant CPR genes, whereas only a few studies have examined both the expression of CPR and the accumulation of CPR enzyme and corresponding activity. An increase in the *HTR1* and *HTR2* CPR transcripts of *Helianthus tuberosus* in response to wounding results an accumulation of two CPR proteins (Lesot et al., 1995). The increase in *CPR1* and *CPR2* gene expression in parsley cells in response to elicitor treatment results in the accumulation of CPR protein and an increase in total CPR activity (Koopman and Hahlbrook, 1997). In the present study, PMCPR amounts remain steady during stratification, germination and early seedling development. This pattern of PMCPR accumulation differs from that

predicted by from northern blot analysis of total RNA from dissected MGs and embryos because *cpr1* transcripts increase in the MG by 4 days (Figure 4.8). A study that examined the regulation of a CPR gene induced in response to thyroid hormone found a large increase in CPR transcript only resulted in a small increase of CPR protein and corresponding activity (Ram and Waxman, 1992). It was suggested that post-translational regulation by differential CPR stability and/or translation efficiency may explain the discrepancy between CPR transcript induction and enzyme amounts. Likewise, the present results indicate a discrepancy between the increase in *cpr1* transcript and the amount of PMCPR during seedling development by 4 days which could also be explained by post-translation regulation. Another possible explaination is that other *cpr* genes and their products are not detected by the *PM5* cDNA and the CPR-antiserum probes used in these studies. The discrepancy between PMCPR transcript and PMCPR protein quantities could also be due to the comparison of northern analysis of dissected tissue versus microsomal proteins from whole seeds and seedlings. Unfortunately, the quantity of dissected material that would be required for microsomal preparations and enzyme assays is prohibitive.

Plant CPRs are believed to be localized to the ER based on the presence of a trans-membrane domain in their N-terminal region and the ER localization of mammalian CPRs (Okada et al., 1982). CPR activity, measured by NADPH:cytochrome c reduction (Madyastha and Coscia, 1979), is commonly used as a marker enzyme for the ER (Benveniste et al., 1978; Robinson et al., 1994). CPRs from different plant sources have been purified and their NADPH:cytochrome c activities examined (Fujita and Asahi, 1985; Menting et al., 1994; Loerwald et al, 1996), and reconstitution *in vitro* with various P450 activities has been demonstrated (Madyastha and Coscia, 1979; Benveniste et al., 1986; Kochs and Grisebach, 1989; Imaishi et al., 1995; Ponnamperuma and Croteau, 1996).

CPR cDNAs from different plant sources have also been isolated and expressed in various systems and their activities reconstituted *in vitro* with P450 activities (Shet et al., 1993; Mizatani and Ohta, 1998), but a study of the subcellular localization of a plant CPR identified through the isolation of its cDNA has not been carried out. The present study uses a CPR-peptide antiserum based on the cDNA sequence of *PM5* to examine the subcellular localization of its corresponding gene product. The data indicate that the product encoded by *PM5* is a microsomal membrane protein, fractionates at the same sucrose density as the ER and NADPH:cytochrome c reductase activity. These results support the conclusions that *PM5* encodes a CPR and that plant CPR enzyme and activity can be used as a marker for the microsomal membranes containing ER (Benveniste et al., 1978). The CPR-peptide antiserum was used to examine the cellular localization of PMCPR in radicle cells by immunofluorescence. The staining pattern indicated PMCPR is not spread throughout the cytosol as other ER associated proteins (e.g. BiP, Forward et al., submitted 1998). PMCPR is localized in distinct areas, possibly adjacent to the outer nuclear envelop which is continuous with the ER. This pattern of staining is consistent with the conclusion that PMCPR associates with the ER, but localization may be directed to certain ER membranes at sites close to the outer nuclear membrane.

Very little is known about the regulation of gene expression during gymnosperm seed germination and early seedling development at the molecular level. Gymnosperm seeds differ both anatomically and genetically from angiosperms and the development of molecular probes to assess the differences in the molecular events that may distinguish these two higher plant groups is essential. It is clear from recent studies that changes in gene expression patterns occur during stratification, germination and early seedling development, but the metabolic consequences of these changes are still not known. During

germination and early seedling development, the P450 monooxygenase system is responsible for a number of essential reactions including fatty acid hydroxylation, secondary metabolite biosynthesis, herbicide detoxification, and gibberellin synthesis. During germination there is an increase in seed metabolic activity and the production of NADPH (Bewley and Black, 1994). It remains to be determined whether the increase in CPR activity found in the present study is coordinated with the increase of other metabolic pathways as suggested earlier (Alani et al., 1990) and with the P450 activities occurring during germination.

The development of molecular probes to detect both CPR transcripts and protein in a gymnosperm is important to further elucidate the mechanisms of CPR regulation in relation to the P450 monooxygenase system and their similarities or differences from angiosperms. Studies of CPR's binding and interaction with NADPH, FAD and FMN, its insertion and retention in the membrane where it interacts with the multitude of P450s, and the role of CPR diversity in plants will also lead to a greater understanding of the P450 monooxygenase system and its evolution in higher plants.

Conclusions

(1) The cDNA clone *PM5* corresponds to a gene (*cpr1*) that encodes a CPR.

(2) *Cpr1* is under developmental and tissue specific control possibly at the transcriptional and/or post-transcriptional levels.

(3) *Cpr1* steady-state transcript amounts are highest during protein mobilization in the MG.

(4) A CPR peptide-antiserum detects a single polypeptide (PMCPR) in 1 and 2-D SDS-PAGE.

(5) PMCPR is associated with the microsomal membrane.

(6) CPR localizes with the ER and cytochrome c activity in sucrose gradient fractions.

(7) Post-translational modifications may regulate CPR activity during germination.

Literature Cited

Alani, A. A., Luster, D. G., and Donaldson, R. P. (1990) Development of endoplasmic and glyoxysomal membrane redox activities during castor bean germination. Plant Physiology 94: 1842-1848

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403-410

Baskins, T.I., Busby, C.H., Fowke, L.C., Summut, M., and Gubler, F. (1992) Improvements in immunostaining samples embedded in methacrylate: localization of microtubules and other antigens throughout developing organs in plants of diverse taxa. Planta 187: 405-413

Benveniste, I., Salaün, J. P., and Durst, F. (1978) Phytochrome-mediated regulation of a monooxygenase hydroxylating cinnamic acid in etiolated pea seedlings. Phytochemistry 17: 359-363

Benveniste, I., Gabriac, B., and Durst, F. (1986) Purification and characterization of the NADPH-cytochrome P-450 (cytochrome c) reductase from higher-plant microsomal fraction. Biochemistry Journal 235: 365-373

Benveniste, I., Lesot, A., Hasenfratz, M. P., and Durst, F. (1989) Immunochemical characterization of NADPH-cytochrome P-450 reductase from Jerusalem artichoke and other higher plants. Biochemistry Journal 259: 847-853

Benveniste, I., Lesot, A., and Hasenfratz, M. P., and Durst, F. (1991) Multiple forms of NADPH-cytochrome P450 reductase in higher plants. Biochemical and Biophysical Research Communications 177: 105-112

Bewley, J. D., and Black, M. (1994) Seeds: Physiology of Development and Germination. Plenum Press, New York

Bolwell, G. P., Bozak, K., and Zimmerlin, A. (1994) Plant cytochrome P450. Phytochemistry 37: 1491-1506

Cabello-Hurtado, F., Batard, Y., Salaün, J. P., Durst, F., Pinot, F., and Werck-Reichhart, D. (1998) Cloning, expression in yeast, and functional characterization of CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation of fatty acids. Journal of Biological Chemistry 273: 7260-7267

Cawley, G. F., Batie, C. J., and Backes, W. L. (1995) Substrate-dependent competition of different P450 isozymes for limiting NADPH-cytochrome P450 reductase. Biochemistry 34: 1244-1247

Coughlan, S. J., Hastings, C., and Winfrey, R. J. (1996) Molecular characterization of plant endoplasmic reticulum: Identification of protein disulfide-isomerase as the major reticuloplasmin. European Journal of Biochemistry 235: 215-224

Donaldson, R. P., and Luster, D. G. (1991) Multiple forms of plant cytochromes P-450. Plant Physiology 96: 669-674

Edwards, D. G. W. (1980) Maturity and quality of tree seeds - a state-of-the-art- review. Seed Science and Technology 8: 625-657

Edwards, D. G. W. (1986) Special pre-chilling techniques for tree seeds. Journal of Seed Technology 10: 151-171

El-Kassaby, Y. A., Edwards, D. G. W., and Taylor, D. W. (1992) Genetic control of germination parameters in Douglas-fir and its importance for domestication. Silvae Genetica 41: 48-54

Forward, B. S., Fowke, L. C., and Misra, S. The Douglas-fir luminal binding protein (BiP) is seasonally regulated and is temporally expressed during seed development and germination. (Submitted to Plant Journal, 1998)

Frank, M. R., Deyneka, J. M., and Schuler, M. A. (1996) Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea. Plant Physiology 110: 1035-1046

Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: Application to the endoplasmic reticulum. Journal of Cell Biology 93: 97-102

Fujita, M., and Asahi, T. (1985) Purification and properties of sweet potato NADPH-cytochrome c (P-450) reductase. Plant and Cell Physiology 26: 397-405

Funk, C., and Croteau, R. (1994) Diterpenoid resin acid biosynthesis in conifers: characterization of two cytochrome P450-dependent monooxygenases and an aldehyde dehydrogenase involved in abietic acid biosynthesis. Archives of Biochemistry and Biophysics 308: 258-266

Funk, C., Lewinsohn, E., Vogel, B. S., Steele, C. L., and Croteau, R. (1994) Regulation of oleoresinosis in Grand fir (*Abies grandis).* Plant Physiology 106: 999-1005

Gish, W., and States, D. J. (1993) Identification of protein coding regions by database similarity search. Nature Genetics 3: 266-72

Higgins, D. G., and Sharp P. M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS 5: 151-153

Imaishi, H., Yamada, T., and Ohkawa, H. (1995) Purification and immunochemical characteristics of NADPH-cytochrome P-450 oxidoreductase from tobacco cultured cells. Biochimica et Biophysica Acta 1246: 53-60

Jarvis, S. B., Taylor, M. A., MacLeod, M. R., and Davies, H. V. (1996) Cloning and characterisation of three cDNA clones of genes that are differentially expressed during dormancy-breakage in the seeds of Douglas fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 147: 559-566

Jarvis, S. B., Taylor, M. A., Bianco, J., Corbineau, F., and Davies, H. V. (1997) Dormancy-breakage in seeds of Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco). Support for the hypothesis that LEA gene expression is essential for this process. Journal of Plant Physiology 151: 457-464

Joshi, C. P., Zhou, H., Huang, X., and Chiang, V. L. (1997) Context sequences of translation initiation codon in plants. Plant Molecular Biology 35: 993-1001

Kaukinen, K. H., Tranbarger, T. J., and Misra, S. (1996) Post-germination-induced and hormonally dependent expression of low-molecular-weight heat shock protein genes of Douglas fir. Plant Molecular Biology 30: 1115-1128

Kochs, G., and Grisebach, H. (1989) Phytoalexin synthesis in soybean: Purification and reconstitution of cytochrome P450 3,9-dihydroxypterocarpan 6a-hydroxylase and separation from cytochrome P450 cinnamate 4-hydroxylase. Archives of Biochemistry and Biophysics 273: 543-553

Koopmann, E., and Hahlbrock, K. (1997) Differentially regulated NADPH:cytochrome P450 oxidoreductases in parsley. Proceedings of National Academy of Sciences U. S. A. 94: 14954-14959

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

Leal, I., and Misra, S. (1993) Developmental gene expression in conifer embryogenesis and germination. III. Analysis of crystalloid protein mRNAs and desiccation protein mRNAs in the developing embryo and megagametophyte of white spruce (*Picea glauca* [Moench] Voss). Plant Science 88: 25-37

Lesot, A., Hasenfratz, M. P., Batard, Y., Durst, F., and Benveniste, I. (1995) Two messenger RNAs are encoding for NADPH-cytochrome P450 reductases in *Helianthus tuberosus* tuber tissues. Plant Physiology and Biochemistry 33: 751-757

Loerwald, B., Clemens, S., and Barz, W. (1996) Purification and characterization of the NADPH-cytochrome P-450 (cytochrome c) reductase from Chickpea. Journal of Plant Physiology 149: 633-635

Madyastha, K. M., and Coscia, C. J. (1979) Detergent-solubilized NADPH-cytochrome c (P-450) reductase from the higher plant, *Catharanthus roseus*. Journal of Biological Chemistry 254: 2419-2427

Meijer, A. H., Cardoso, M. I. L., Voskuilen, J. T., de Waal, A., Verpoorte, R., and Hoge, J. H. C. (1993) Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH-cytochrome P-450 reductase, an enzyme essential for reactions catalysed by cytochrome P-450 mono-oxygenases in plants. Plant Journal 4: 47-60

Menting, J. G. T., Cornish, E., and Scopes, R. K. (1994) Purification and partial characterization of NADPH-cytochrome c reductase from *Petunia hybrida* flowers. Plant Physiology 106: 643-650

Meyer, K., Cusumano, J. C., Somerville, C., and Chapple, C. C. S. (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. Proceedings of National Academy of Sciences USA 93: 6869-6874

Miwa, G. T., and Lu, A. Y. H. (1984) The association of cytochrome-P-450 and NADPH-cytochrome P-450 reductase in phospholipid membranes. Archives of Biochemistry and Biophysics 234: 161-166

Miwa, G. T., West, S. B., Huang, M. T., and Lu, A. Y. H. (1979) Studies on the association of cytochrome P-450 and NADPH-cytochrome c reductase during catalysis in a reconstituted hydroxylating system. Journal Biological Chemistry 254: 5695-5700

Mizutani, M., and Ohta, D. (1998) Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Plant Physiology 116: 357-367

Mizutani, M., Ohta, D., and Sato, R. (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis thaliana* and its expression manner in planta. Plant Physiology 113: 755-763

Nebert, D. W., and Gonzalez, F. J. (1987) P450 genes: Structure, evolution, and regulation. Annual Review of Biochemistry 56: 945-93

O'Farrell P. H., and O'Farrell, P. Z. (1977) Two-dimensional polyacrylamide gel electrophoretic fractionation. Methods in Cell Biology 16: 407-420

Okada, Y., Frey, A. B., Guenthner, T. M., Oesch, F., Sabatini, D. D., and Kreibich, G. (1982) Studies on the biosynthesis of microsomal membrane proteins. European Journal of Biochemistry 122: 393-402

O'Leary, K. A., Beck, T. W., and Kasper, C. B. (1994) NADPH cytochrome P-450 oxidoreductase gene: Identification and characterization of the promoter region. Archives of Biochemistry and Biophysics 310: 452-459

Peterson, J. A., Ebel, R. E., O'Keeffe, D. H., Matsubara, T., and Estabrook, R. W. (1976) Temperature dependence of cytochrome P-450 reduction. Journal of Biological Chemistry 251: 4010-4016

Ponnamperuma, K., and Croteau, R. (1996) Purification and characterization of an NADPH-cytochrome P450 (cytochrome c) reductase from spearmint (*Mentha spicata*) glandular trichomes. Archives of Biochemistry and Biophysics 329: 9-16

Porter, T. D., Beck, T. W., and Kasper, C. B. (1990) NADPH-cytochrome P-450 oxidoreductase gene organization correlates with structural domains of the protein. Biochemistry 29: 9814-9818

Porter, T. D., and Coon, M. J. (1991) Cytochrome P-450: Multiplicity of isoforms, substrates, and catalytic regulatory mechanisms. Journal of Biological Chemistry 266: 13469

Ram, P. A., and Waxman, D. J. (1992) Thyroid hormone stimulation of NADPH P450 reductase expression in liver and extrahepatic tissues. Journal of Biological Chemistry 267: 3294-3301

Robinson, D. G., Hinz, G., and Oberbeck, K. (1994) Isolation of endo-and plasma membranes. *In* Plant Cell Biology: A Practical Approach. *Eds* Harris, N., and Oparka, K.J. IRL Press, New York pp. 245-272

Rosco, A., Pauli, H. H., Priesner, W., and Kutchan, T. M. (1997) Cloning and heterologous expression of NADPH-cytochrome P450 reductases from the *Papaveraceae*. Archives of Biochemistry and Biophysics 348: 369-377

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY. ISBN 0-87969-309-6

Schuler, M. (1996) The role of cytochrome P450 monooxygenases in plant-insect interactions. Plant Physiology 112: 1411-1419

Shephard, E. A., Palmer, C. N. A., Segall, H. J., and Phillips, I. R. (1992) Quantification of cytochrome P450 reductase gene expression in human tissues. Archives of Biochemistry and Biophysics 294: 168-172

Shet, M. S., Sathasivan, K., Arlotto, M. A., Mehdy, M. C., and Estabrook, R. W. (1993) Purification, characterization, and cDNA cloning of an NADPH-cytochrome P450 reductase from mung bean. Proceedings of National Academy of Sciences U. S. A. 90: 2890-2894

Simmons, D. L., Lalley, P. A., and Kasper, C. B. (1985) Chromosomal assignments of genes coding for components of the mixed-function oxidase system in mice. Journal of Biological Chemistry 260: 515-521

Sutter, T. R., and Loper, J. C. (1989) Disruption of the *Saccharomyces cerevisiae* gene for NADPH-cytochrome P450 reductase causes increased sensitivity to kenoconazole. Biochemical and Biophysical Research Communications 160: 1257-1266

Szekeres, M., Németh, K., Koncz-Kálmán, Z., Mathur, J., Kauschmann, A., Altmann, T., Rédei, G. P., Nagy, F., Schell, J., and Koncz, C. (1996) Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in *Arabidopsis*. Cell 85: 171-182

Taylor, M. A., Davies, H. V., Smith, S. B., Abruzzese, A., and Gosling, P. G. (1993). Cold-induced changes in gene expression during dormancy breakage in seeds of Douglas-fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142:120-123

Taylor, M. A., and Davies, H. V. (1995) Changes in gene expression during imbibition of seeds of Douglas fir (*Pseudotsuga menziesii*). Journal of Plant Physiology 142: 54-58

Teutsch, H. G., Hasenfratz, M. P., Lesot, A., Stoltz, C., Garnier, J. M., Jeltsch, J. M. Durst, F., and Werck-Reichhart, D. (1993) Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. Proceedings of National Academy of Sciences U. S. A. 90: 4102-4106

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

Tranbarger, T. J., and Misra S. (1995) The molecular characterization of a set of cDNAs differentially expressed during Douglas-fir germination and early seedling development. Physiologia Plantarum 95: 456-464

Tranbarger, T. J., and Misra, S. (1996) Structure and expression of a developmentally regulated cDNA encoding a cysteine protease (pseudotzain) from Douglas fir. Gene 172: 221-226

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., and Pompon, D. (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. Journal of Biological Chemistry 272: 19176-19186

Vermilion, J. L., and Coon, M. J. (1981) Purified liver microsomal NADPH-cytochrome P-450 reductase: Spectral characterization of oxidation-reduction states. Journal of Biological Chemistry 253: 2694-2704

Verwoerd, T. C., Dekker, B. M. M., and Hoekema, A. (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Research. 17: 2362

Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J-J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes. Proceedings of National Academy of Sciences U. S. A. 94: 8411-8416.

Winkler, R. G., and Helentjaris, T. (1995) The Maize *dwarf3* gene encodes a cytochrome P450-mediated step in gibberellin biosynthesis. Plant Cell 7: 1307-1317

Yasukochi, Y., Okita, R. T., and Masters, B. S. S. (1980) Comparison of the properties of detergent-solubilized NADPH-cytochrome P-450 reductases from pig liver kidney immunochemical, kinetic, and reconstitutive properties. Archives of Biochemistry and Biophysics 202: 491-498

CHAPTER 5

Conclusions and Future Studies

To examine the regulation of gene expression during Douglas-fir seed germination and post-germination, a set of developmentally regulated cDNAs was isolated by differentially screening a seedling cDNA library (Chapter 2). The cDNAs were identified through sequence database searches and were found to encode proteins with similarities to a chaperonin 60ß (cpn60ß), cysteine proteinase (CysP), low molecular weight heat shock protein (LMW HSP), luminal binding protein (BiP), and chlorophyll a/b-binding protein (Type II CAB). A constitutively expressed cDNA with similarity to a NADPH-cytochrome P450 reductase (CPR) was also isolated. The individual expression patterns of these genes suggested transcriptional and/or post-transcriptional changes occur during stratification, germination, and post-germination of Douglas-fir seeds (Tranbarger and Misra, 1995, 1996; Kaukinen et al., 1996; Forward et al., submitted 1998). Many questions concerning the regulation and functions of these genes during germination and post-germination remain unanswered and require further investigation.

One of the aims of future studies is to identify the *cis*-acting promoter elements that interact with trans-acting factors involved in the transcriptional initiation of the developmentally regulated genes. The cDNAs described in this study provide probes for the isolation of the corresponding genes and their regulatory elements. The comparison of the gymnosperm genes with those of angiosperms will help identify the molecular divergence that has occurred during higher plant evolution. For example, the legumin

storage protein genes of angiosperms have one less intron than gymnosperm legumin-like genes, which may be due to a loss of an intron during evolutionary divergence (Häger et al., 1995).

Several of the developmentally regulated Douglas-fir transcripts are present in the dormant mature seed (Chapter 2, Figures 2.2 and 2.3). The localization of mRNAs during development is an important mechanism for the regulation and targeting of proteins during eukaryotic development (Micklem, 1995). Only a few studies have addressed the mechanisms involved in the storage of mRNAs in mature seeds and their specific cellular destinations during development. Some transcripts stored in dormant seeds associated with proteins as messenger ribonucleoprotein (mRNP) complexes, while others were localized in the nucleus and were polyadenylated during germination (Hammett and Katterman, 1975; Ajtkhozhin et al., 1976; Harris and Dure, 1978; Peumans et al., 1979). *In situ* hybridization studies on sections of mature, stratified and germinating seeds and on seedling tissues during development will provide information on the changes in the subcellular localization of these transcripts.

Pseudotzain: Regulation, Processing and Role in Storage Protein Mobilization

The presence of multiple AU-rich elements (*ARE*) in the 3' *UTR* of the *cysP* transcript (Chapter 3, Figure 3.1 and 3.2) suggest pseudotzain gene expression is post-transcriptionally regulated at the level of mRNA turnover. Both *in vitro* and transgenic approaches have shown that destabilization elements in 3' *UTR* have a role in post-transcriptional regulation of certain mammalian, yeast and plant genes (Sullivan and Green, 1993; Chen and Shyu, 1995; Ross, 1995; Caponigro and Parker, 1996; Abler and Green,

1996; Gil and Green, 1996). A determination of the *in vitro cysP* mRNA decay rates at different developmental time points, and when incubated with various crude cellular extracts (e.g. the cytosol, polysomes, or mRNP) would provide basic information about the stability of *cysP* transcripts during development (Figure 5.1A). The results from this experiment would determine whether *cysP* transcripts are differentially stable during development and indicate the cellular fraction that may contain factors involved in destabilization. A transgenic approach could then be used to identify the elements in the *cysP* transcript that mediate destabilization. This would involve construction of chimeric genes containing a common promoter (e.g. the 35S promoter from the cauliflower mosaic virus) fused to test genes with different 3' ends derived from the *cysP* cDNA *UTR* sequence (Figure 5.1B). The decay rates of the test transcripts can then be determined from mRNA isolated from transformed plant cells treated with the transcription inhibitor actinomycin D (Figure 5.1A). If the *cysP* 3' *UTR* confers differential stability to the test transcripts, the sequence elements that direct the destabilization could be used in mobility-shift assays with various cellular extracts to identify putative trans-acting factors that are involved in mRNA destabilization.

A unique feature of the predicted amino acid sequence of pseudotzain is the presence of both N- and C-terminal extension peptides. The large C-terminal extension peptide appears to be unique to plant CysP, but no functions (e.g. targeting, enzymatic etc.) have been attributed to its presence. To determine the roles of the extension peptides in the subcellular targeting of pseudotzain, gene constructs containing the coding region of a reporter gene (e.g. the gene for ß-glucuronidase, GUS) flanked by either the N- or C-terminal extension peptides or both can be expressed in transgenic tobacco under the control of a constitutive promoter (e.g. the 35S promoter). The localization of GUS activity in relation to the

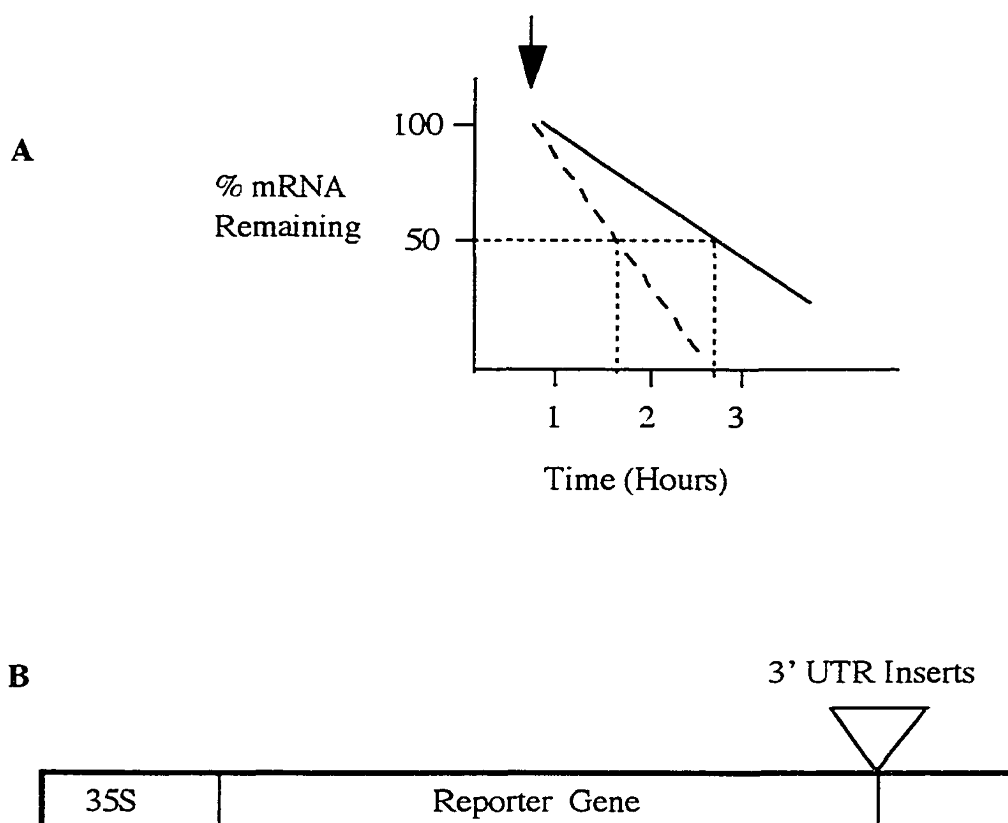


Figure 5.1 mRNA decay experiments. (A) The graph depicts an idealized mRNA decay time course of two transcripts (dashed and solid line). Arrow indicates beginning of *in vitro* incubation of mRNAs with cellular extracts or point at which transformed cells are treated with actinomycin D. mRNA half-life values can be extrapolated from the graph (dotted lines). (B) A representative construct used to determine the effects of different *cysP* 3' *UTR* on reporter transcript stability. The reporter gene indicates any transcript that is detectable by northern hybridization analysis. 35S, promoter from the cauliflower mosaic virus.

presence or absence of the extension peptides would determine whether the N- and C-terminal extension peptides are involved in the subcellular targeting of pseudotzain.

The N-terminal extension peptides of several CysP are post-translationally removed during development, but there are no reports of the processing of CysP that have both N- and C-terminal extension peptides (Koehler and Ho, 1990; Holwerda et al., 1990; Kalinski et al., 1992). A pseudotzain antiserum is required to determine the subcellular location of the native pseudotzain in Douglas-fir tissues and to elucidate the processing steps of the pro-pseudotzain to the mature form. An attempt was made to raise an anti-pseudotzain antiserum based on a synthetic peptide from the cDNA *PM3-3*, but the antiserum did not detect any antigens in various Douglas-fir protein extracts. In addition, an antiserum for an angiosperm CysP was tested on Douglas-fir protein extracts and no specific band was detected in SDS-PAGE/western blot analysis. An alternative approach is to overexpress the mature portion of pseudotzain in an expression vector and raise an antiserum against the purified protein.

Several questions remain unanswered concerning the role of pseudotzain during Douglas-fir germination (Chapter 3). Northern blot analysis demonstrated that pseudotzain expression increases during germination and is highest in the megagametophyte (MG) during post-germination. This result is consistent for the proposed role of pseudotzain in the mobilization of storage proteins. In order to assign a conclusive role for proteinases in the degradation of storage proteins, several criteria have been established (Shutov and Vaintraub, 1987). These include the ability to hydrolyse native storage proteins, localization of the proteinase and the substrate in the same cell compartment, and enzyme characteristics that match the corresponding cell compartment environment. Immunolocalization and western blot analysis of protein body extracts incubated with

pseudotzain will confirm whether it associates with and degrades native storage proteins of Douglas-fir.

CPR and the Cytochrome P450 Monooxygenase System

It has been proposed that P450 enzymes compete for a limited amount of CPR which may limit P450 enzymatic activities (Cawley et al., 1995). To test this hypothesis, experiments were initiated to determine whether elevated CPR expression in transgenic tobacco plants leads to higher CPR accumulation, activity and ultimately P450 enzyme activities. CPR constructs were made with the coding region of the PMCPR cDNA behind a double 35S promoter in the pBI121 plasmid (Figure 5.2A and B). Eleven putative transgenic plants were selected and screened for elevated amounts of CPR by western blot with the CPR peptide-antiserum (Figure 5.3). The Douglas-fir antiserum cross-reacted with a polypeptide of expected size in tobacco protein extracts from control nontransgenic plants and from eleven transformants. A qualitative assessment of CPR amounts indicated none of the transformants had higher CPR accumulation than the control non-transgenic plants. Lanes that appear to have higher CPR amounts also had higher total protein loaded per lane based on Coomassie Blue stained gels (data not shown). One of the transgenic plants was selected for northern blot analysis with the CPR cDNA (*PM5*) as a probe to determine the amount of *cpr* mRNA. The results indicated that the *cpr* gene was not expressed in various tissues of the transgenic tobacco plant examined (Figure 5.4). However, results from Southern blot analysis with the CPR cDNA *PM5* demonstrated the *cpr* coding sequence was stably integrated into the tobacco genome (data not shown). Based on these data, it appears that the *cpr* gene construct was integrated but is not expressed in the transgenic tobacco. The post-transcriptional silencing of genes expressed

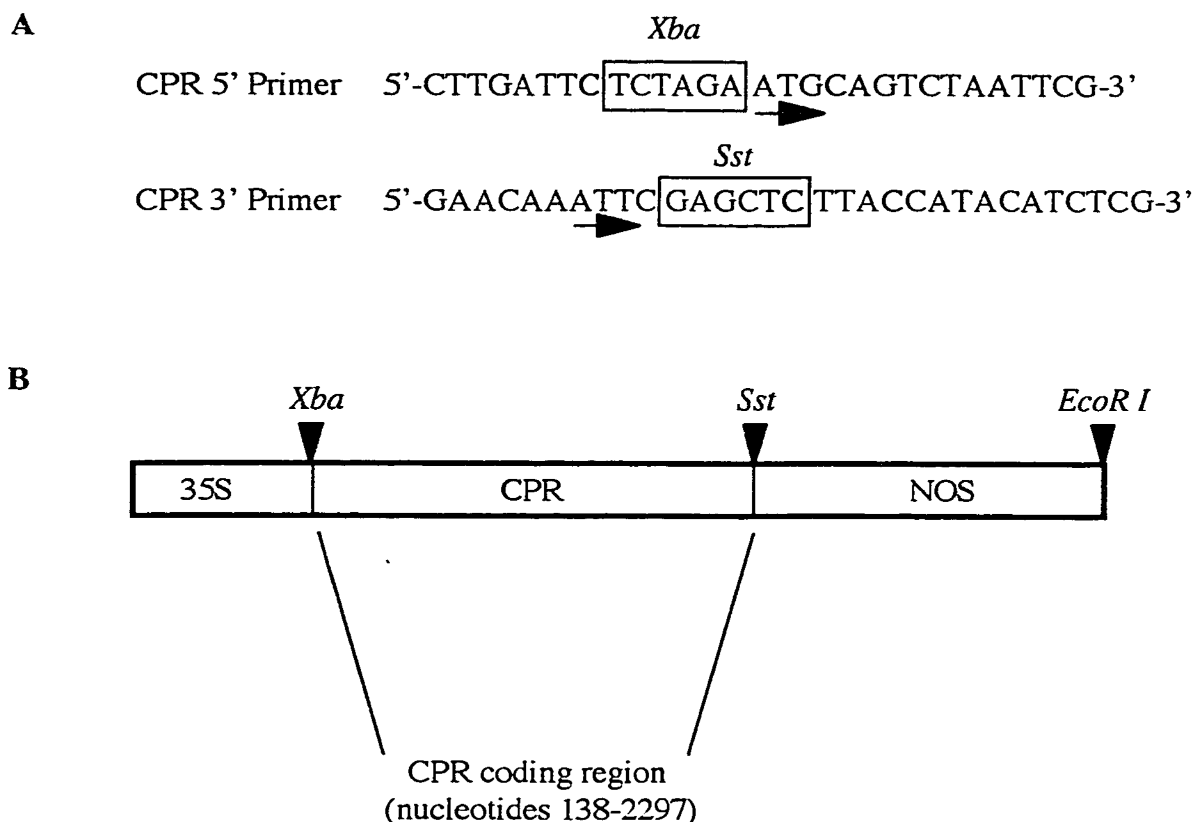


Figure 5.2. Primers synthesized to engineer the *Xba* and *Sst* restriction sites and the structure of the *cpr* gene construct used to transform tobacco plants. (A) The start (CPR 5' primer) and stop codons (CPR 3' primer) of the *cpr* coding sequence are underlined with arrows. (B) The *cpr* gene construct. 35S, 35S promoter from the cauliflower mosaic virus; CPR, the CPR coding sequence; NOS, the nopaline synthase 3' end; *Xba*, restriction site; *Sst*, restriction site; *Eco*RI, restriction site. Tobacco leaf discs were incubated with *Agrobacterium tumefaciens* containing the constructs, and transformed cells were selected for their ability to form callus tissue on medium containing kanamycin. Transformed callus was used to regenerate the transgenic plants.

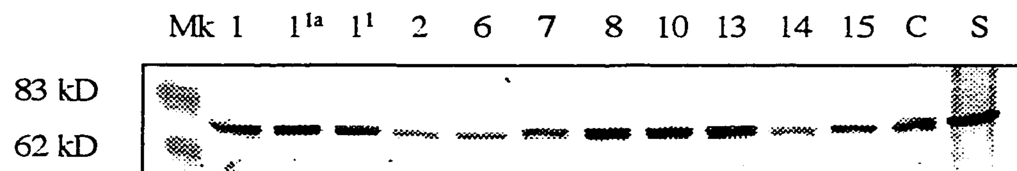


Figure 5.3. Western blot analysis of putative transgenic CPR tobacco plants. Soluble protein extracts (approximately 15 μg per lane) isolated from leaves of putative transgenic and control tobacco plants were resolved by SDS-PAGE, blotted to nitrocellulose and probed with the CPR peptide-antiserum. Mk, molecular mass markers; 1, 1^{1a}, 1^{1}, 2, 6, 7, 8, 10, 13, 14 and 15 are labels for putative transgenic tobacco plants; C, control, nontransgenic tobacco plant; S, mature Douglas-fir seed.

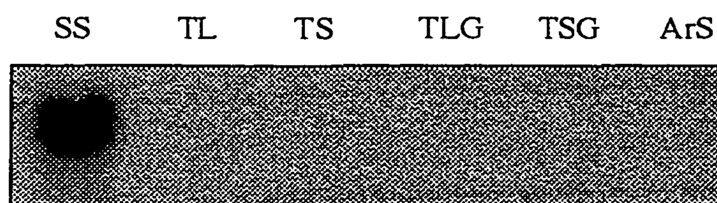


Figure 5.4. Northern blot analysis of putative CPR transgenic tobacco plants. The *PM5* cDNA was used to probe total RNA (15 μg loaded per lane) isolated from the seeds and tissues of a putative transgenic tobacco plant (1^1), and control seeds and plant tissues. SS, stratified Douglas-fir seed; TL, putative transgenic tobacco plant 1^1 leaves; TS, putative transgenic tobacco plant 1^1 seeds; TLG, control GUS transgenic plant leaves; TSG, control GUS transgenic plant seeds; ArS, *Arabidopsis thaliana* mature seeds.

in transgenic plants is an important but poorly understood phenomenon that may explain why the *cpr* gene is not expressed in transgenic tobacco (Baulcombe, 1996). Northern blot analysis of other putative transformants is in progress.

A unique aspect of the plant cytochrome P450 monooxygenase system is the multiplicity of *cpr* genes in some plant species (Benveniste et al., 1989, 1991; Lesot et al. 1995; Koopmann and Hahlbrock, 1997; Rosco et al., 1997; Urban et al., 1997; Mizutani and Ohta, 1998). In the present study, only a single CPR protein was identified, but the question remains whether there are multiple functional *cpr* genes in Douglas-fir. A Southern blot analysis of Douglas-fir genomic DNA using the CPR cDNA as a probe gave inconclusive results (data not shown). In addition, because of the high number of pseudogenes in gymnosperm genomes, Southern blot data cannot be used exclusively to determine functional gene copy number (Kinlaw and Neale, 1997). An alternative to Southern blot analysis includes using different portions of the CPR cDNA, including the conserved cofactor regions and the less conserved 5' or 3' *UTRs*, to isolate CPR gene sequences from a genomic library. Gene specific probes used in northern blot analysis or RNA protection assays of RNA isolated from different developmental stages and tissues would determine whether the genes are expressed.

Regulation of PMCPR appears to occur post-translationally because CPR activity during germination does not correlate with the amount of PMCPR protein. The amino acid sequence of PMCPR contains a putative serine phosphorylation (RRGGS, Chapter 4, Fig. 4.2.) and a *N*-glycosylation site. In an attempt to elucidate mechanisms of post-translational regulation of CPR activity during germination, studies were initiated to examine the glycosylation and the phosphorylation status of PMCPR. To determine whether PMCPR is glycosylated, microsomal proteins extracted from 2-day-old seedlings

were treated with the enzyme Endo-H to remove N-linked glycosyl-moieties. Preliminary results from this experiment indicated no shift in the molecular mass of PMCPR in SDS-PAGE/western blot analysis after Endo-H treatment (data not shown). Based on this result, PMCPR does not appear to be glycosylated. The activities of some P450 enzymes of mammals are modulated post-translationally through phosphorylation (Porter and Coon, 1991). The presence of a putative serine phosphorylation site in the PMCPR prompted an examination of the phosphorylation status of Douglas-fir microsomal proteins during development. Microsomal protein extracts from mature seeds and 2-day-old seedlings were resolved by SDS-PAGE, blotted to nitrocellulose and probed with the antiphosphoserine antiserum (Figure 5.5). The antiphosphoserine antiserum did not detect any protein band that corresponded with the molecular mass of PMCPR. However, there were several lower molecular mass bands that increased in the 2-day-old seedling samples indicating that phosphorylation of some proteins at serine residues may occur during germination. These proteins are of interest because *A. thaliana* mutants having constitutive serine-threonine phosphatase activities germinate in the presence of exogenously applied ABA (Leung et al., 1994; Meyer et al., 1994). The phosphorylation of some proteins at serine residues may be an important signal transduction process that is related to seed germination, and a further examination of these phosphorylated proteins is warranted. To elucidate the factors that regulate CPR activity during development, future work is needed on the purification and characterization of PMCPR activity *in vitro*.

Based on the data presented in this thesis, transcriptional, post-transcriptional and post-translational mechanisms appear to control genes expressed during germination and post-germination. The availability of molecular probes forms the basis for future experiments that will clarify the specific regulatory mechanisms that function during seed dormancy,

germination and post-germination of Douglas-fir. The results of these experiments will allow a better understanding of the molecular and metabolic differences that have developed during the evolutionary divergence of angiosperms and gymnosperms.

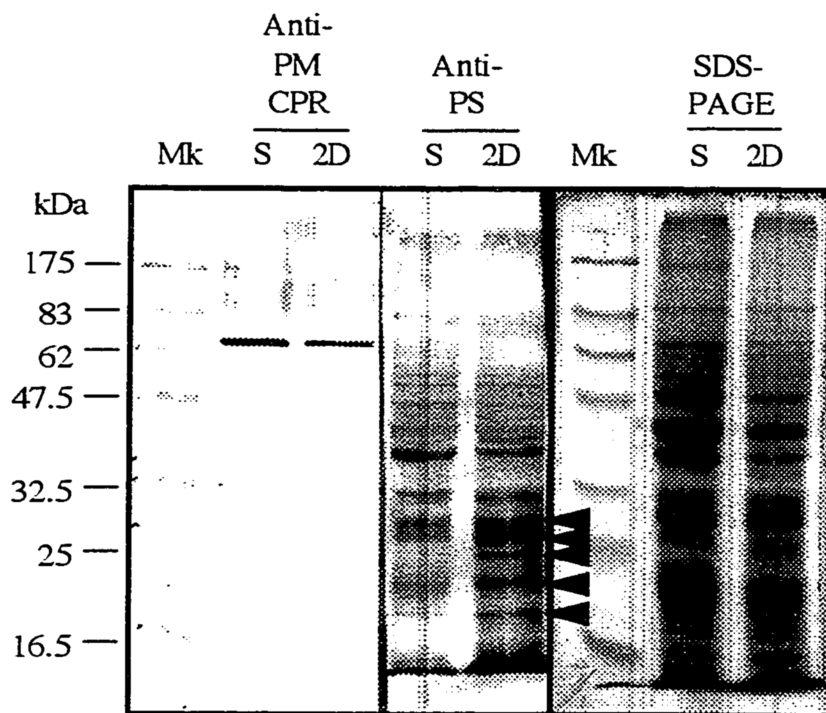


Figure 5.5. Western blot analysis of the serine phosphorylation status of microsomal proteins from mature seeds and 2-day-old seedlings. Microsomal proteins (30 μg loaded per lane) were resolved by SDS-PAGE, and blotted to nitrocellulose. Anti-PMCPR, western blot probed with the CPR peptide-antiserum; Anti-PS, western blot probed with the anti-phosphoserine (1/500 dilution); SDS-PAGE, resolved proteins stained with Coomassie Blue; Mk, molecular mass markers. The anti-phosphoserine detected bands were visualized by incubation with anti-mouse (IgM +IgG) conjugated to alkaline phosphatase (1,2000 dilution). Arrows indicate proteins that appear to be phosphorylated during germination. See Chapter 4 Materials and Methods section for SDS-PAGE and western blot procedures.

Literature Cited

Abler, M. L., and Green, P. J. (1996) Control of mRNA stability in higher plants. Plant Molecular Biology 32: 63-78

Ajtkhozhin, M. A., Doschanov, Kh. J., and Akhanov, A. U. (1976) Informosomes as a stored form of mRNA in wheat embryos. FEBS Letters 66: 124-126

Baulcombe, D. C. (1996) RNA as a target and an initiator of post-transcriptional gene silencing in trangenic [sic] plants. Plant Molecular Biology 32: 79-88

Benveniste, I., Lesot, A., Hasenfratz, M. P., and Durst, F. (1989) Immunochemical characterization of NADPH-cytochrome P-450 reductase from Jerusalem artichoke and other higher plants. Biochemistry Journal 259: 847-853

Benveniste, I., Lesot, A., and Hasenfratz, M. P., and Durst, F. (1991) Multiple forms of NADPH-cytochrome P450 reductase in higher plants. Biochemical and Biophysical Research Communications 177: 105-112

Caponigro, G., and Parker R. (1996) Mechanisms and control of mRNA turnover in *Saccharomyces cervisiae*. Microbiological Reviews 60: 233-249

Cawley, G. F., Batie, C. J., and Backes, W. L. (1995) Substrate-dependent competition of different P450 isozymes for limiting NADPH-cytochrome P450 reductase. Biochemistry 34: 1244-1247

Chen, C.-Y. A., and Shyu, A.-B. (1995) AU-rich elements: characterization and importance in mRNA degradation. Trends in Biochemical Science 20: 465-470

Forward, B. S., Fowke, L. C., and Misra, S. The Douglas-fir luminal binding protein (BiP) is seasonally regulated and is temporally expressed during seed development and germination. (Submitted, 1998)

Gil, P., and Green, P. J. (1996) Multiple regions of the Arabidopsis SAUR-AC1 gene control transcript abundance: the 3' untranslated region functions as an mRNA instability determinant. The EMBO Journal 15: 1678-1686

Häger, K. -P., Braun, H., Czihal, A., Müller, B., Bäumlein, H. (1995) Evolution of seed storage protein genes of *Ginkgo biloba*. Journal of Molecular Evolution 41: 457-466

Hammett, J. R., and Katterman, F. R. (1975) Storage and metabolism of poly(adenylic acid)-mRNA in germinating cotton seeds. Biochemistry 14: 4375-4379

Harris, B., and Dure, L. (1978) Developmental regulation in cotton seed germination: Polyadenylation of stored messenger RNA. Biochemistry 17: 3250-3256

Holwerda, B. C., Galvin, N. J., Baranski, T. J., and Rogers, J. C. (1990) *In vitro* processing of aleurain, a barley vacuolar thiol protease. Plant Cell 2: 1091-1106

Kalinski, A., Melroy, D. L., Dwivedi, R. S., and Herman, E. M. (1992) A soybean vacuolar protein (P34) related to thiol proteases is synthesized as a glycoprotein precursor during seed maturation. Journal of Biological Chemistry 267: 12068-12076

Kaukinen, K. H., Tranbarger, T. J., and Misra, S. (1996) Post-germination-induced and hormonally dependent expression of low-molecular-weight heat shock protein genes of Douglas fir. Plant Molecular Biology 30: 1115-1128

Kinlaw, C. S., and Neale, D. B. (1997) Complex gene families in pine genomes. Trends in Plant Science 2: 356-359

Koehler, S. M., and Ho, T.-H. D. (1990) Hormonal regulation, processing, and secretion of cysteine proteinases in barley aleurone layers. Plant Cell 2: 769-783

Koopmann, E., and Hahlbrock, K. (1997) Differentially regulated NADPH:cytochrome P450 oxidoreductases in parsley. Proceedings of National Academy of Sciences U. S. A. 94: 14954-14959

Lesot, A., Hasenfratz, M. P., Batard, Y., Durst, F., and Benveniste, I. (1995) Two messenger RNAs are encoding for NADPH-cytochrome P450 reductases in *Helianthus tuberosus* tuber tissues. Plant Physiology and Biochemistry 33: 751-757

Leung, J., Bouvier-Durand, M., Morris, P-C., Guerrier, D., Chefdor, F., and Giraudat, J. (1994) *Arabidopsis* ABA response gene *ABI1*: features of a calcium-modulated protein phosphatase. Science 264: 1448-1452

Meyer, K., Leube, M., and Grill, E. (1994) A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. Science 264: 1452-1455

Micklem, D. R. (1995) mRNA localisation during development. Developmental Biology 172: 377-395

Mizutani, M., and Ohta, D. (1998) Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Plant Physiology 116: 357-367

Peumans, W. J., Caers, L. I., and Carlier, A. R. (1979) Some of the aspects of the synthesis of long-lived messenger ribonucleoproteins in developing rye embryos. Planta 144: 485-490

Porter, T. D., and Coon, M. J. (1991) Cytochrome P-450: Multiplicity of isoforms, substrates, and catalytic regulatory mechanisms. Journal of Biological Chemistry 266: 13469

Rosco, A., Pauli, H. H., Priesner, W., and Kutchan, T. M. (1997) Cloning and heterologous expression of NADPH-cytochrome P450 reductases from the Papaveraceae. Archives of Biochemistry and Biophysics 348: 369-377

Ross, J. (1995) mRNA stability in mammalian cells. Microbiological Reviews 59: 423-450

Shutov, A. D., and Vaintraub, I. A. (1987) Degradation of storage proteins in germinating seeds. Phytochemistry 26: 1557-1566

Sullivan, M. L., and Green, P. J. (1993) Post-transcriptional regulation of nuclear-encoded genes in higher plants: the roles of mRNA stability and translation. Plant Molecular Biology 23: 1091-1104

Tranbarger, T. J., and Misra S. (1995) The molecular characterization of a set of cDNAs differentially expressed during Douglas-fir germination and early seedling development. Physiologia Plantarum 95: 456-464

Tranbarger, T. J., and Misra, S. (1996) Structure and expression of a developmentally regulated cDNA encoding a cysteine protease (pseudotzain) from Douglas fir. Gene 172: 221-226

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., and Pompon, D. (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. Journal of Biological Chemistry 272: 19176-19186

Appendix I

Total RNA Isolation Data (Samples, Concentrations and Yields)

#	Sample Description	Tissue Fresh Weight (grams)	Total RNA Concentration [$\mu g \mu L^{-1}$]	Yield (μg; pellets dissolved in 100 μL)	Yield Ratio (μg total RNA per gram Fresh weight tissue)
1	Mature Seeds	1.0	3.7	370	370
2	Stratified Seeds (3 weeks, 4 °C)	1.0	3.3	330	330
3	2 Day Old Seedlings	1.0	2.7	270	270
4	4 Day Old Seedlings	1.0	2.8	280	280
5	6 Day Old Seedlings	1.0	2.2	220	220
6	8 Day Old Seedlings	1.0	2.3	230	230
7	14 Day Old Seedlings	1.0	4.4	440	440
8	Needles (14 Days)	1.0	11.0	1,100	1,100
9	Radicles (14 Days)	0.5	1.0	100	200

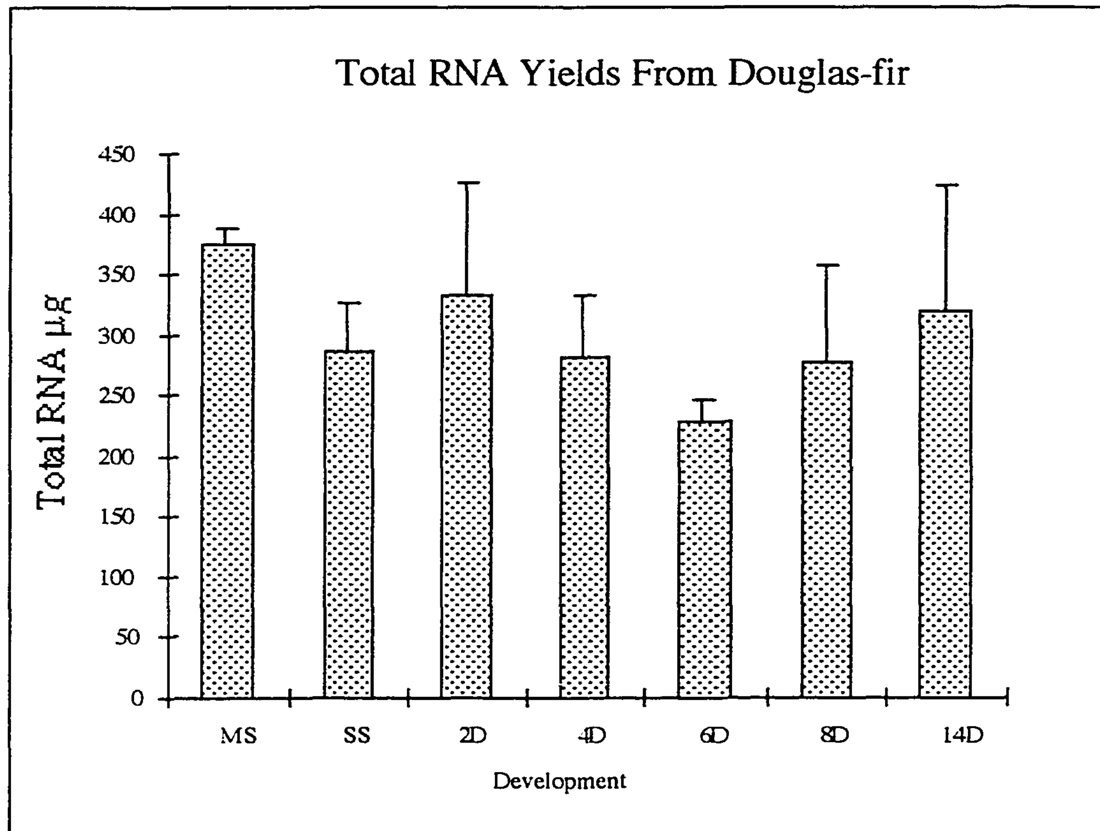
Run I. Total RNA isolated for Douglas-fir germination time course via modified Perez-Grau procedure (Grimes et al., 1993). For northern blots see Chapter 2.

#	Sample Description	Tissue Fresh Weight (grams)	Total RNA Concentration [$\mu g \mu L^{-1}$]	Yield (μg; pellets dissolved in 100 μL, *80 μL, or @90 μL)	Yield Ratio (μg total RNA per gram Fresh weight tissue)
1	Mature Seeds	1.0	3.9	390	390
2	Stratified Seeds (3 weeks, 4 °C)	1.0	2.5	250	250
3	2 Day Old Seedlings	1.0	2.9	290	290
3'	2 Day Old Seedlings	1.0	2.9	*232	232
4	4 Day Old Seedlings	1.0	2.5	250	250
5	6 Day Old Seedlings	1.0	3.7	370	370
6	8 Day Old Seedlings	1.0	2.5	250	250
7	14 Day Old Seedlings	1.0	2.3	230	230
8	10 week Old Needles	1.0	1.25	100	100
9	10 Week Old Roots	1.0	0.86	69	69
10	10 Week Old Stems	1.0	1.8	@162	162

Run II. Total RNA isolated for Douglas-fir germination time course via modified Perez-Grau procedure (Grimes et al., 1993). For northern blots see Chapter 2.

#	Sample Description	Tissue Fresh Weight (grams)	Total RNA Concentration [$\mu g \mu L^{-1}$]	Yield (μg; pellets dissolved in 50 μL)	Yield Ratio (μg total RNA per gram Fresh weight tissue)
1	Mature Seeds	1.0	7.3	365	365
2	Stratified Seeds (3 weeks, 4 ^{o}C)	1.0	5.5	275	275
3	2 Day Old Seedlings	1.0	8.8	440	440
4	4 Day Old Seedlings	1.0	6.7	335	335
5	6 Day Old Seedlings	1.0	4.3	215	215
6	8 Day Old Seedlings	1.0	4.7	235	235
7	14 Day Old Seedlings	1.0	5.4	270	270
8	Needles (14 Days)	1.0	8.4	420	420

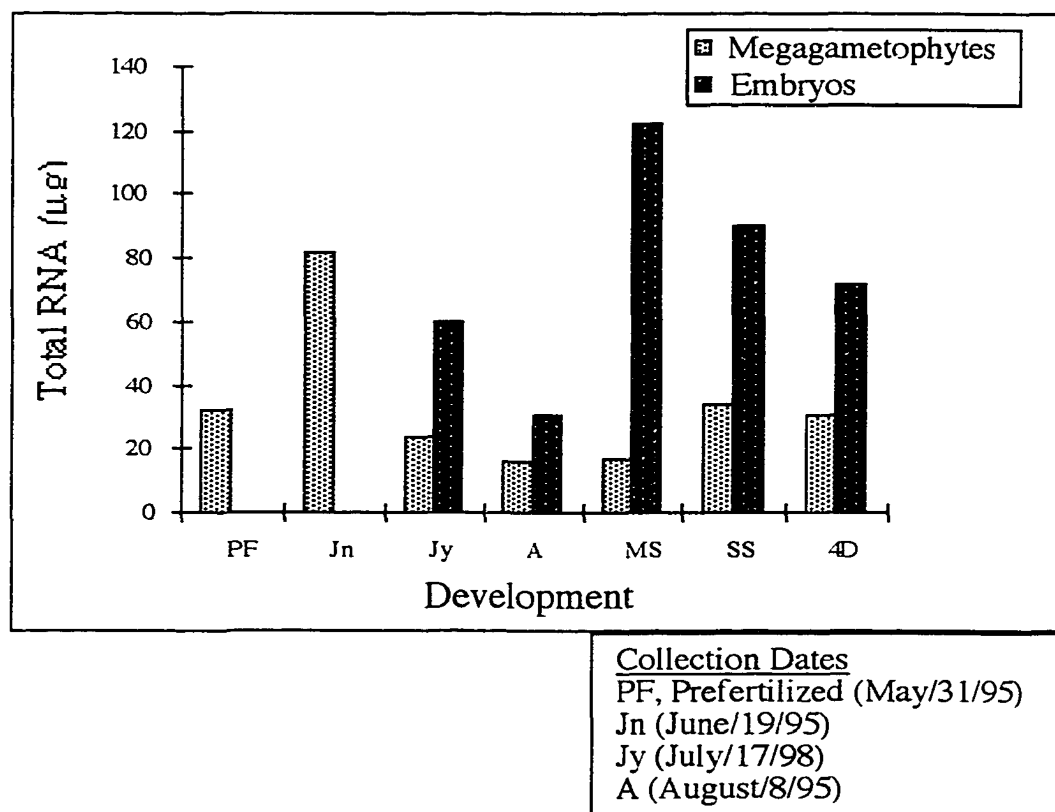
Run III. Total RNA isolated for Douglas-fir germination time course via modified Perez-Grau procedure (Grimes et al., 1993). For northern blots see Chapter 2.



Summary of Yields. Total RNA isolated for Douglas-fir germination time course via modified Perez-Grau procedure (Grimes et al., 1993). For northern blots see Chapter 2.

#	Sample Description	Tissue Fresh Weight (grams)	Total RNA Concentration [$\mu g \mu L^{-1}$]	Yield (μg; pellets dissolved in 20 μL)	Yield Ratio (μg total RNA per gram Fresh weight tissue)
1	Prefertilized (May/31/95)	0.1	1.6	32	320
2	MG (June/19/95)	0.1	4.1	82	820
3	MG (July/17/98)	0.1	1.2	24	240
4	MG (August/8/95)	0.1	0.81	16	160
5	MG Mature Seeds	0.1	0.85	17	170
6	MG Stratified Seeds	0.1	1.7	34	340
7	MG 4 Day Old Seedlings	0.1	1.5	30	300
8	MG 10 Day Old Seedlings	–	–	–	–
9	Embryo (July/17/98)	0.7	3.0	60	86
10	Embryo (August/8/95)	0.7	1.5	30	43
11	Embryo Mature Seeds	0.7	6.1	122	174
12	Embryo Stratified Seeds	0.7	4.5	90	128
13	Embryo 4 Day Old Seedlings	0.7	3.6	72	104
14	Embryo 10 Day Old Seedlings	–	–	–	–

Total RNA isolated for Douglas-fir embryogenesis and germination time course via modified Verwoerd miniprep procedure (Leal and Misra, 1993). For northern blots see Chapters 3 and 4.



Summary graph of Total RNA isolated for Douglas-fir embryogenesis and germination time course via modified Verwoerd miniprep procedure (Leal and Misra, 1993). For northern blots see Chapters 3 (Figure 3.) and Chapter 4 (Figure 4.)

Appendix II

Microsomal and Total Protein Isolation Data Isolation Data (Samples, Concentrations and Yields)

#	Sample Description	Tissue Fresh Weight (grams)	Microsomal Protein (MS) Concentration [$\mu g \mu L^{-1}$]	Soluble Protein (SN2) Concentration [$\mu g \mu L^{-1}$]	Microsomal Protein Yield [μg MS protein per g Fresh weight; pellet resuspended in 100 μL or *50 μL)
1	Mature Seeds	5	14.5	10.8	290
2	Stratified Seeds	5	10.3	7.7	206
3	2 Day old Seedlings	5	12.8	7.8	256
5	7 Day old Seedlings	5	6.8	8.5	136
6	14 Day old Seedlings	5	12.8	1.2	256
7	14 Day old Seedlings	5	13.0	0.7	260

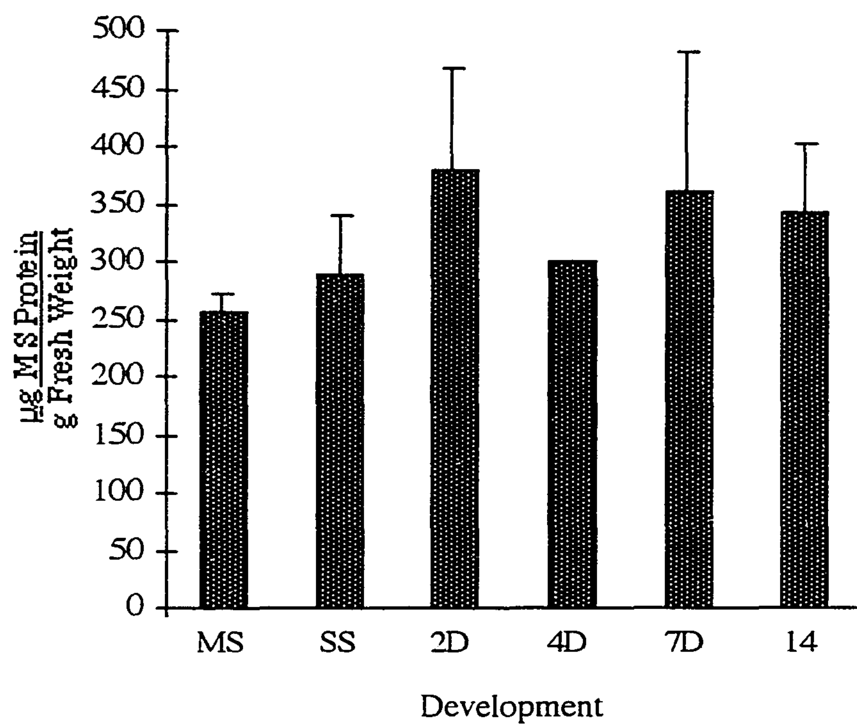
Run I. Microsomes were isolated according to Shet et al., 1993 as described in the "Materials and Methods" section of Chapter 4, For western blots see Chapter 4.

#	Sample Description	Tissue Fresh Weight (grams)	Microsomal Protein (MS) Concentration [$\mu g \mu L^{-1}$]	Soluble Protein (SN2) Concentration [$\mu g \mu L^{-1}$]	Microsomal Protein Yield [μg MS protein per g Fresh weight; pellet resuspended in 50 μL)
1	Mature Seeds	5	24.5	12	245
2	Stratified Seeds	5	38.0	14.4	380
3	Stratified Seeds	5	38.5	13.4	385
4	2 Day old Seedlings		55.0	15	550
5	7 Day old Seedlings	5	55.0	16.6	550
6	14 Day old Seedlings	5	46	7.2	460

Run II. Microsomes were isolated according to Shet et al., 1993 as described in the "Materials and Methods" section of Chapter 4, For western blots see Chapter 4.

#	Sample Description	Tissue Fresh Weight (grams)	Microsomal Protein (MS) Concentration [$\mu g \mu L^{-1}$]	Soluble Protein (SN2) Concentration [$\mu g \mu L^{-1}$]	Microsomal Protein Yield [μg MS protein per g Fresh weight; pellet resuspended in 50 μL)
1	Mature Seeds	5	14.5	24.0	230
2	Stratified Seeds	5	17.4	31.2	278
3	2 Day old Seedlings	5	20.6	26.0	330
4	4 Day old Seedlings	5	18.8	18.6	301
5	7 Day old Seedlings	5	24.6	18.8	394
6	14 Day old Seedlings	5	19.4	15.8	310

Run III. Microsomes were isolated according to Shet et al., 1993 as described in the "Materials and Methods" section of Chapter 4, For western blots see Chapter 4.



Summary graph of microsomal protein yields from Douglas-fir germination time course.