uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Shuyou Han
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biology)
GRADE / DEGREE

Department of Biology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Development of Soybean Seed Coat as a Bioreactor for the Production of Industrial Enzymes

TITRE DE LA THÈSE / TITLE OF THESIS

Douglas Johnson
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

John Arnason

Peter Pauls

Christine Charest

Tim Xing

Brian Miki

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

DEVELOPMENT OF SOYBEAN SEED COAT AS A BIOREACTOR FOR THE PRODUCTION OF INDUSTRIAL ENZYMES

Shuyou Han

Thesis submitted to the

School of Graduate Studies and Research

University of Ottawa

in partial fulfillment of the requirement for the

PhD degree in the

Ottawa-Carleton Institute Biology

April 2009

Your file Votre référence
ISBN: 978-0-494-51798-7
Our file Notre référence
ISBN: 978-0-494-51798-7

ABSTRACT

We have investigated the use of the soybean seed coat as a bioreactor for the production of valuable proteins. In this system heterologous proteins are targeted to the hourglass cells. While seed coats currently have little commercial value, the introduction of high valued bio-products may generate new markets for Canadian industries and contribute towards the growth of rural economies.

The seed coat system is based upon a promoter responsible for the production of soybean peroxidase (SBP) in the hourglass cell layer of the seed coat. SBP protein is thought to be synthesized as a pre-protein containing N- and C-terminal propeptides responsible for the intracellular localization of the mature SBP to the vacuole. We used this promoter and these signals to direct the synthesis and targeting of horseradish peroxidase (HRPC) from *Armoracia rusticana* and two laccases (LCC1 and LCC4) from the fungus *Trametes versicolor*. The cDNA sequences were codon-optimized for soybean, and fused with the SBP N- and C-terminal propeptides. A HIS tag was added to the C-terminal end of the mature protein to facilitate purification. Expression of the chimeric genes was driven by the Ep 1.5 kb fragment cloned from soybean peroxidase promoter.

Fourteen lines of HRPC transgenics, 16 lines of LCC1 transgenics and 40 lines of LCC4 transgenics were generated by particle bombardment. The T_1 and T_2 seeds of the transgenics were screened for peroxidase and laccase activity with the highest expressing lines further analysed by PCR and Western blots. The peroxidase activities measured in HRPC transgenics ranged from 0.02 to 3.09 U/mg protein while the laccase activities measured in laccase transgenics varied from 0 to 0.099 U/mg protein. In general these values are low compared to activities measured in the native species or in other

transgenics. We have found that expression of HRPC and laccase as a trait is inheritable from T_0 to T_2 generation. The anti-LCC4, anti-HRPC antibodies were used to detect the expression of LCC4 and HRPC by Western blotting. A 35.4 kDa protein reactive with anti-HRPC and anti-HIS tag antibodies was found in the seed coat of an HRPC transgenic while a 77.6 kDa protein reactive with anti-LCC4 and anti-HIS tag antibodies was found in the seed coat of an LCC4 transgenic.

These results support the use of the soybean seed coat as a bioreactor for the production of foreign proteins. The factors which affect foreign protein expression are fully discussed and future directions have been proposed to optimize the system.

RÉSUMÉ

Nous avons étudié le tégument de la graine de soya utilisé en tant que bioréacteur pour la production de protéines à valeur ajoutée. Dans ce système, les protéines hétérologues ont été ciblées dans les ostéoscléréides ('hourglass cells'), cellules à constriction médiane du tégument. Alors que les téguments des graines ont peu de valeur marchande, l'introduction de bio-produits à valeur ajoutée a le potentiel de générer de nouveaux marchés pour les industries canadiennes et contribuer ainsi à la croissance des économies rurales.

Le système du tégument de la graine de soya est basé sur un promoteur responsable de la production de la peroxydase de soya (SBP) dans la couche des ostéoscléréides. On pense que la protéine SBP est d'abord synthétisée comme une pré-protéine avec des propeptides à terminales -N et -C qui sont responsables de la localisation intracellulaire de la SBP mature dans la vacuole. Nous avons utilisé ce promoteur et ses signaux pour optimiser et diriger la synthèse de la peroxydase du raifort, *Armoracia rusticana*, (HRPC), et de deux laccases (LCC1 et LCC4) du champignon *Trametes versicolor*. Les codons des séquences d'ADN ont été optimisés pour le soya, et unis aux propeptides SBP à terminales –N et –C. Un marqueur HIS a été ajouté à la terminale -C de la protéine mature pour faciliter sa purification. L'expression des gènes chimériques a été dirigée par le fragment PE de 1,5 kb qui a été clôné à partir du promoteur de la peroxydase de soya.

Quatorze lignées transgéniques de HRPC, 16 lignées transgéniques de LCC1 et 40 lignées transgéniques de LCC4 ont été générées par bombardement de particules. Les

graines des générations transgéniques T_1 et T_2 ont été sélectionnées à partir des activités les plus élevées de la peroxydase et des laccases, puis ces lignées ont été analysées par PCR et Western Blot. Les activités peroxydasiques mesurées dans les HRPC transgéniques vont de 0,02 à 3,09 U/mg de protéines, celles des laccases de 0 à 0,099 U/mg de protéines. En général, ces valeurs sont faibles par rapport aux activités mesurées dans des espèces indigènes ou encore dans d'autres lignées transgéniques. Nous avons observé que l'expression des laccases et de l'HRPC est un trait héréditaire parmi les générations T_0 à T_2. Les anticorps anti-LCC4 et anti-HRPC ont été utilisés pour détecter l'expression de LCC4 et d'HRPC par Western Blot. Nous avons trouvé une protéine de 35,4 kDa réagissant à l'anti-HRPC et à l'anti-marqueur HIS dans le tégument des graines d'une HRPC transgénique ainsi qu'une protéine de 77,6 kDa réagissant à l'anti-LCC4 et à l'anti-marqueur HIS dans les téguments des graines d'une LCC4 transgénique.

Ces résultats valident notre hypothèse que le tégument de la graine de soya peut être utilisé comme bioréacteur pour la production de protéines étrangères. Nous discutons ici des facteurs qui affectent l'expression des protéines étrangères et proposons des axes de recherche pour optimiser encore plus ce système dans l'avenir.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my supervisor, Dr. Douglas Johnson, for his patient guidance and encouragement during the development of this thesis. Dr. Johnson provided me with insight, critical discussion and opportunities to attend exhilarating conferences. I would also like to thank my advisor, Dr. Brian Miki, for access to the research facilities at Agriculture and Agri-Food Canada and who mentored me in the philosophy of science and provided encouragement.

I would also like to thank my committee members, Dr. Linda Bonen and Dr. Tim Xing for their helpful comments and suggestions throughout the course of this thesis. I am grateful to have the support of several scientists from AAFC, in particular, Dr. Daina Simmonds, Dr. Johann Schernthaner and Dr. Shea Miller for their technical support and helpful comments.

I appreciated the kind help from Dr. Dan Brown (AAFC, London), Dr. Federica Brandizzi (University of Saskatchewan), who allowed me to use their confocal microscopy facilites. I am very grateful to Mr. Ann-fook Yang and Ms. Denise Chabot (AAFC, Ottawa) for their time and effort in training me on the microscopic techniques. I am indebted to Dr. Christiane Charest for the translation of the Abstract.

Thanks also go to my lab mates both in school and at the farm, Jaimie Schnell, Loreta Gudynaite-Savitch, Hélène Labbé, Teresa Martin, and Ming Hu who gave me great help on many aspects of my work. The friends I made at University of Ottawa are the source of happiness and wisdom. I will miss the laughter and good humor.

Nothing is more important than the encouragements from my parents and my wife. I could not finish my thesis without emotional and financial supports from my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviations	Full Name
35S	cauliflower mosaic virus 35S promoter
aa	amino acid
Amp	ampicillin
bp	base pair
BSA	bovine serum albumin
cDNA	complementary DNA
CTPP	C-terminal propeptide
cv.	cultivar
D	day
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dNTP	dinucleotide triphosphate
dTTP	deoxythymidine triphosphate
dH_2O	distilled water
ddH_2O	distilled de-ionized water
DNA	deoxyribonucleic acid
dNTP	deoxyribonuleotide triphosphate, any of the four bases
dpa	days post anthesis
DTT	dithiothreitol
DV	dense vesicle

E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EpEp	soybean cultivar homozygous for the *Ep* gene with high soybean peroxidase activity, such as cv. Harosoy 63
epep	soybean cultivar homozygous for the ep gene with low soybean peroxidase activity, such as cv. Jack
ER	endoplasmic reticulum
g	gram
gDNA	genomic DNA
GFP	green fluorescent protein
GMO	genetically modified crops
HGCs	hourglass cells
hr	hour
HRPC	horseradish peroxidase C from *Armoracia rusticana*
IPTG	isopropyl β-D-thiogalactosidase
ISAAA	international service for the acquisition of agri-biotech applications
Kan	kanamycin
Kan^{50}	50mg/L kanamycin
Kan^{R}	kanamycin resistant
Kan^{S}	kanamycin sensitive
Kb	kilobase
KCI	potassium chloride

L	liter
LB	Luria broth
LCC1	laccase I from *Trametes versicolor*
LCC4	laccase IV from *Trametes versicolor*
LV	lytic vacuole
M	molar
mg	milligram
$MgCl_2$	magnesium chloride
min	minute
mL	milliliter
mM	millimolar
mon	month
mRNA	messenger RNA
MWt	molecular weight
NaCI	sodium chloride
ng	nanogram
nt	nucleotide
NaOAc	sodium acetate
NaOH	sodium hydroxide
NTPP	N-terminal propeptide
OD	optical density
ORF	open reading frame
PBS	phosphate buffered saline solution

PCR	polymerase chain reaction
pI	isoelectric point
PSV	protein storage vacuole
rDNA	ribosomal DNA
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
rpm	revolution per minute
RT	room temperature
RT-PCR	reverse-transcriptase polymerase chain reaction
SBP	soybean peroxidase (from *Glycine max* L. Merrill)
SDS	sodium dodecyl sulfate
sec	second
T	metric tonnes
TAE	1X tris-acetate buffer (pH 8.0)
TE	10 mM Tris, 1mM EDTA (pH 8.0)
T-DNA	transferred DNA
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
U	unit
µ g	microgram
µ l	micro liter
µ M	micro molar

LIST OF APPENDICES

CHAPTER ONE: INTRODUCTION

1.1 Soybean as an Important Crop Plant

Soybean (Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine max L. Merrill) has been used as food for more than 1,000 years in Asian countries and today is consumed world wide. Unique among crop plants, it has both high fat (20%) and protein (40%) content with the remainder consisting of 35% carbohydrate and 5% ash (Hajduch et al., 2005). Soybean accounts for 56% of the world oilseed production (Kerley and Allee, 2003), an industry-leading 29% of all vegetable oil production, and 67% of all vegetable protein meal production (Gardner and Payne, 2003). The oil is used mainly for human consumption and industrial uses while the protein is used mainly as animal feed and human foodstuffs (Rao et al., 1998). Soy protein is considered to be complete with all the essential amino acids. Moreover, its amino acid composition is comparable to that of protein from milk or meat (Takamatsu et al., 2004). The high usage of soybean meal in poultry and swine diets can be attributed to its relatively high concentration of protein and the excellent profile of highly-digestible amino acids (Bruce et al., 2006). Recently, various soybean specialty products such as soymilk and tofu have gained popularity. Tofu, a cottage-cheese like soybean curd, has a high nutritional value and is rich in proteins, vitamins, and minerals, particularly calcium (Rao et al., 1998). In Canada, the Harovinton cultivar was designated as a tofu-type soybean (Buzzell et al., 1991). Soybeans are also used in industrial products including oils, soap, cosmetics, resins, plastics, inks, crayons, solvents, clothing, and biodiesel and have even been used as fermenting stock to make a brand of vodka (http://www.matr.net/article-5744.html). Several health-related claims have been made for soybean products including

cholesterol-lowering effects or as anti-obesity foods (Velasquez and Bhathena, 2007). In addition they may relieve some lifestyle-related diseases such as cancer, osteoporosis and menopausal disorder (Takamatsu et al., 2004).

One of the world's premier agriculture crops, global soybean production was 227 Mt (million tonnes) (http://www.soybean.on.ca/Marketing/World_Supply_dist.pdf) [1] in 2007 while the estimated production in Canada in the 2007 crop year was 2.87 Mt (http://www.canadiansoybeans.com/production.php)[2]. In 2007 Ontario farmers planted more than 0.89 million ha of soybean, an increase of 3.9% over 2006. (http://www.statcan.ca/Daily/English/070626/d070626a.htm)[3] and just 34,400 ha short of the record set in 2004 and 2005. Soybean is an important cash crop in Canada, ranking fourth in economic importance behind wheat, canola, and corn. Canadian soybean production contributes $769 million (average price of seed was $270 per tonne) to the Canadian economy. Soybean in Canada is mainly used in food and industry (54% of production), feed (10%) and for stock and other uses (36%) (http://www.canadiansoybeans.com/production.php).

Transgenic soybean is the largest single transgenic crop in the world accounting for 64% in terms of global adoption rate, followed by cotton (43%), corn (24%) and canola (20%). In 2007, the global area of transgenic soybean reached ~60 million ha (http://www.isaaa.org/resources/publications/briefs/37/pptslides/Brief37slides.pdf). GM soybeans were first planted in Canada in 1997 and by 2005 accounted for 60% of the area devoted to soybean (Brookes and Barfoot, 2006). GM soybean basically is herbicide-tolerant

[1] 227 million tonnes
[2] 2824 kt = 2.87 Mt ; 1 tonne (metric) = 0.9842 ton (long UK)
[3] 2.2 million acres = 0.89 million ha; 1 hectare = 2.4711 acre

soybean, mainly Roundup Ready soybean (Monsanto, USA). In Ontario, ~40-50% of soybeans are genetically engineered.

Currently, soybean seed coats are removed during the processing of the seed and they have little value to soybean seed processors. It is known that soybean seed coats, or hulls, represent 8 to 10% of the weight of soybean grain (Sessa and Wolf, 2001). According to Statistics Canada, the national soybean production in 2007 was 2,848,000 tonnes, suggesting that 256,320 tonnes of seed coat were produced. Soybean seed coats have little intrinsic value and may be discarded as waste, although their value can be as much as 73.8-83.6 USD per tonne if added back to soybean meal or to other food as a source of fiber (http://www.mnsoy.com/market_meal.htm).

In this thesis, we were attempting to target the soybean seed coat for the production of useful foreign proteins. These products could be introduced into a market that already accepts transgenic soy. Success in this project could have a tremendous economic impact.

1.2 Transgenic Soybean

The commercial varieties and the lines currently in field trails have predominantly been modified for herbicide tolerance using bacterial genes, for instance, Roundup Ready soybean from Monsanto expressed throughout the plant from the CaMV 35S promoter/enhancer. This is fairly straight forward compared with the complex manipulations required for tissue-specific expression of value-added traits. In North America, tissue-specific modifications of transgenic traits have been limited to the seed embryo. The advantages of using seeds for the production of recombinant proteins with plant-based expression systems have been demonstrated for several species including maize (Hood et al., 2003), rice (Yasuda et al., 2006), and tomato (Zheng et al., 2007) . Soybean has the potential

for large-scale recombinant protein production. Recently, the utility of soybean as an efficient production platform for vaccines that can be used for oral delivery was demonstrated. The B subunit of the heat labile toxin of enterotoxigenic *Escherichia coli* (LTB) was used as a model immunogen for production in soybean seed. LTB expression was under the control of soybean seed-specific glycinin promoter and was directed to the endoplasmic reticulum (ER) of seed storage parenchyma cells for sequestration. Pentameric LTB accumulated to 2.4% of the total seed protein at maturity and was stable in desiccated seed (Moravec et al., 2007). It was also proven that soybean could produce the major K99 fimbrial subunit FanC for use as an edible subunit vaccine (Piller et al., 2005). A synthetic version of FanC, codon-optimized for soybean, accumulated to nearly 0.5% of the total soluble protein when expressed in the cytosol of mesophyll cells in the leaf tissue.

The evaluation of processes for the purification of recombinant protein has also been performed in soybean (Robic et al., 2006). This study showed that even β-glucuronidase (GUS) with an acidic pI can be successfully separated from native soybean proteins which also have acidic pIs.

1.3 Soybean Seed Development

1.3.1 Soybean Seed Coat Structure

Relative to the much larger body of knowledge on the development of the soybean embryo, the seed coat is poorly characterized. Studies have been conducted on the biochemical composition of the seed coat (Buzzell et al., 1987) but information on the specific cell types and their development is limited (Miller et al., 1999). Only a few seed coat-specific genes and regulatory elements have been described with soybean peroxidase being a prominent exception (Gijzen, 1997).

The novel features of the mature soybean seed coat, its structure, development and composition were recognized in early studies (Corner, 1951; Carlson and Lersten, 1987; Moïse et al., 2005). As shown in Figure 1.1, the seed coat consists of several prominent cell layers. Outermost is the epidermal layer, which consists of a single layer of palisade cells, or macrosclereids that are elongated and perpendicular to the surface of the seed. Inside the palisade layer is a single-celled hypodermal layer of thick-walled hourglass cells (HGCs) or osterosclereids, except for the hilum region. More specifically, the hourglass cell layer is composed of large vacuolated cells. It is densely cytoplasmic (Jin, 2008) and appears to be fluid and viscous in the immature seed coat (Gijzen et al., 1993). The presence of numerous starch grains in the hourglass cells during embryogenesis may indicate that the seed coat is responsible for synthesis of solutes that are beneficial to the developing embryo (Algan and Buyukkartal, 2000; Wang and Grusak, 2005). The innermost portion of the seed coat is a multi-cellular layer of partially flattened parenchyma cells. Each of the above layers is maternally derived from the outer (palisade cell layer, hourglass cell layer,) and inner (thick walled parenchyma layer, thin cell walled parenchyma cell layer, and endothelium layer) integuments. Immediately inside the inner parenchyma is the aleurone layer, which is derived by double fertilization and thought to be part of the endosperm. It is tightly compressed against the seed coat by the expansion of the cotyledons of the embryo (Miller et al., 1999).

1.3.2 Soybean Seed Coat Development

HGCs in the soybean seed coat arise from the outer-cell layer of the outer integument and differentiate from the hypodermis near the top of the seed at 12 dpa (Jin, 2008). In

mature soybean seed coats, HGCs are separated by wide intercellular air spaces and are larger than palisade and parenchyma cells (Souza and Marcos-Filho, 2001). The main function of hourglass cells is providing mechanical strength to the seed coat (Moïse et al., 2005; Jin, 2008).

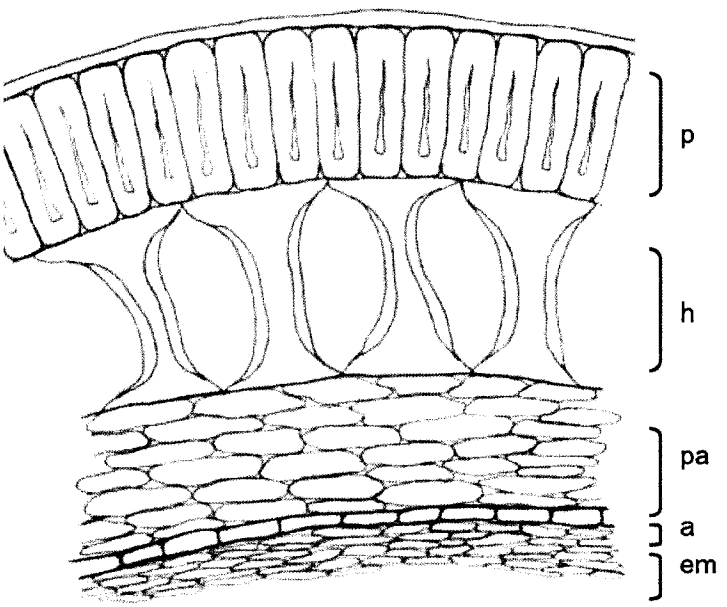
Morphological studies using light microscopy and histochemistry were done to describe changes that occur in the seed coat after fertilization when most of the dramatic changes occur (Miller et al., 1999). Jin (2008) described the development of hourglass cells using both scanning and transmission electron microscopy.

At anthesis, the seed coat does not show the characteristic features of the mature tissue. A cuboidal epidermis surrounding several layers of undifferentiated thin-walled parenchyma forms the outer integument which is 2-4 cell layers thick; the thickness in the micropyle and hilum regions is exceptional, where a greater number of cell layers are present. Following fertilization, both periclinal and anticlinal divisions occur so that the outer integument becomes 12-15 cell layers thick. The inner integument consists of a cuboidal endothelium which provides the inner boundary of the seed coat.

At 6 days post anthesis (dpa), the structure of the seed coat changes significantly (Jin, 2008). The epidermal cells of the outer integument are elongating perpendicular to the seed coat surface to differentiate into the thick walled palisade layer. The thick-walled parenchyma has expanded and becomes four to six cells thick. The lateral vascular bundles are extending to form the vascular region of the thin-walled parenchyma of the outer integument.

Figure 1.1 Schematic Drawing of Mature Soybean Seed Coat

p, palisade layer; h, hourglass cells; pa, partially crushed parenchyma; a, aleurone; em, crushed endosperm. The palisade layer and the hourglass cells comprise the outer integument while the inner integument consists of parenchymatous cells. This drawing is from Moïse et al. (2005).



At 9 dpa, the seed has started to enlarge. The palisade, counter palisade that is connection part between seed coat and funiculous (Miller et al., 1999) ,and tracheid bar at the base of the funiculus start to show their characteristic features. The outer layer of thin-walled parenchyma and the inner layer of thick-walled parenchyma have differentiated from the outer integument; intercellular air spaces develop within the thin-walled parenchyma cell layers (Miller et al., 1999). Parenchyma may facilitate active transport or secretion (Thorne, 1981) with the presence of plasmodesmata interconnecting parenchyma cells indicating a possible symplastic route for nutrients (Yaklich et al., 1998; Jin, 2008). Epidermis and hypodermis appear in the cells opposite the hilum, at the bottom of the seed. The cellular development of hilum is different from the rest of the seed coat. The epidermal cells of outer integument and the funiculus result in a double layer of palisade cells. According to Jin (2008), at 9 days post anthesis most hypodermal cells show the rectangular cell shape and are tightly packed without intercellular air spaces.

At 12 dpa, the palisade cells have distinct thin- and thick-walled cells and clear vascular region develops. Although HGCs begin to differentiate from the hypodermis near the top, or hilum portion, of the seed, they do not show characteristics of mature hourglass cells. The central vacuole in hourglass cell occupies ~40% of the cell and several mitochondria, ER, proplastids, Golgi and proplastids with starch grains can be seen (Jin, 2008). There are still no obvious air spaces between HGCs.

At 15 dpa, the seed coat shows three typical layers. The cell wall of palisade and HGCs becomes thicker. Obvious air spaces are visible between HGCs when they differentiate. The thin- and thick-walled parenchyma and vascular region are clearly

distinguishable (Miller et al., 1999). Large air spaces are also starting to appear between the thin-walled parenchyma cells. At this stage, HGC walls are thicker (Jin, 2008).

At 18 dpa, changes to the structure of the soybean seed coat are less dramatic. HGCs are separated at their mid points by large air spaces but at their top and bottom the cell walls still adjoin closely. Inclusions in the central vacuole can be seen. HGCs have significant peroxidase activity at 19 dpa, while palisade layers contained trace amounts (Jin, 2008). At 21 dpa, the cell walls and cell shapes of the hourglass cell do not change much. The central vacuole still occupies most of the HGC, the content of vacuolar inclusions is much less than at 18 dpa (Jin, 2008). Small electron opaque vesicles are found in the cytoplasm of both hourglass cells and parenchyma cells. At 30 dpa, the parenchyma is beginning to be crushed (Miller et al., 1999). The ultrastructure of HGC is unchanged. HGCs contain many ER, Golgi, proplastids and vacuolar inclusions. At 45 dpa, the seed coat is mature. The most prominent features are the HGCs and palisade layer and only a few layers of slightly flattened thin-walled parenchyma remains (Jin, 2008).

The above observations provide the fundamental information about the structure and development of the hourglass cells that is needed for the analysis of foreign proteins synthesized in HGCs such as HRPC and laccases as described in this thesis.

1.4 The Hourglass Cell as a Place of Protein Storage

Hourglass cells have been described in the seed coats of many legumes including lupin (Clements et al., 2004) , pea (Van Dongen et al., 2003), milkvetch (Miklas et al., 1987), common bean (Yeung, 1990), white popinac (Serrato-Valenti et al., 1995), *Senna corymbosa* (Rodriguez-Pontes, 2007) and many others. The HGC structure is similar among the legumes and likely is common to all of them.

HGCs in soybean are interesting in that they appear to store large amount of peroxidase as a single isozyme of seed coat peroxidase (SBP) (Gijzen et al., 1993; Welinder and Larsen, 2004). Soybean cultivars can be classified into high-peroxidase and low-peroxidase phenotypes (Buttery and Buzzell, 1968), the former due to the presence of a dominant gene, designated *Ep.* Soybean cultivars homozygous for the recessive gene, *epep*, e.g., cv. "Jack", express peroxidase at only 1% peroxidase activity of that of the high-peroxidase phenotype (Gijzen, 1997). Peroxidase deficient *epep* lines produce normal-looking HGCs and do not under-perform *EPEP* soybean varieties such as Harosoy 63 in field trials (Gijzen et al., 1993). The *ep* mutation contains a deletion of 87 bp in which the putative transcription start site remains intact (Gijzen, unpublished data) while 9 bp of the 5'UTR and 78 bp encoding the first 26 amino acids are deleted (Gijzen, 1997).

The *epep* cv. "Jack" was chosen as the host for our experiments. Its low peroxidase level allows for easier detection of transgenic peroxidases. In Jack, the seed coat is 9.92 ± 0.38% of whole seed weight (N = 11) and the protein content is 10.68 ± 0.098 % (N = 4) of the seed coat total weight.

1.5 The Expression of Foreign Proteins in Heterologous Systems

1.5.1 Foreign Protein Expression Systems

Foreign proteins can be expressed in a number of different systems including bacteria, fungi, plants, and animals (Table 1.1). Generally speaking, the ideal expression system must: 1) produce large amounts of the desired, biologically-functional products; 2) be cost-effective; 3) allow for convenient storage and distribution of desired product; 4) come

with little or no risk environmentally; 5) and be brought to production within a reasonable time (Norris, 2005).

Prokaryotic and eukaryotic systems are the two general categories. Prokaryotic systems are generally easier to handle and are satisfactory for most purposes. However, there are serious limitations in using prokaryotic cells for the production of eukaryotic proteins. For example, many of the eukaryotic proteins undergo a variety of post-translational modifications like proper folding, glycosylation, phosphorylation, formation of disulphide bridges, etc. that are not found in typical prokaryotes.

All heterologous protein expression systems have advantages but also limitations, therefore, the choice of system necessarily requires tradeoffs (Frommer and Ninnemann, 1995; Verma et al., 1998; Macauley-Patrick et al., 2005; Sorensen and Mortensen, 2005; Vitale and Pedrazzini, 2005; Boehm, 2007). For instance, *E. coli* was the first organism to be employed for recombinant protein production because of its long tradition as a model scientific organism, the ease of genetic manipulations and the availability of well-established fermentation procedures (Gellissen et al., 2005). However, limitations in protein secretion and the lack of glycosylation impose restrictions on general use, especially for expressing plant and mammalian proteins. Furthermore, recombinant products are often retained as inclusion bodies. Although inclusion bodies may represent a good starting material for purification and downstream procedures, such as protein refolding, they often contain the recombinant proteins as insoluble, biologically inactive aggregates (Villaverde and Carrio, 2003).

Correct disulfide bond formation and glycosylation are important for heterologous protein production. Prokaryotes have, in general, the limited capacity of forming disulfide

bonds. If one or more disulfide bonds are necessary for target protein's activity, a eukaryotic system would be the better choice. If the target protein requires N- terminal glycosylation for proper function, prokaryotic systems are also disqualified. In spite of these disadvantages, *Escherichia coli* (Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae) can be used for production of small protein in which post translational modification is not important (Dietmar et al., 1988). Fungal and yeast systems can also be used to produce enzymes and other proteins. Using the fungi *Aspergillus niger* and *Aspergillus oryzae*, protein production can be increased to a very high level. But the process of fermentation and need for equipment can be expensive (Macauley-Patrick et al., 2005). Still 43% of the biologics on the market are produced in bacteria and yeast including insulin, human growth hormone and many others (Boehm, 2007). Up to 2006, 43 biopharmaceuticals produced in *E coli* and 23 biopharmaceuticals produced in yeast, were approved in the United States and Europe (http://www.nature.com/nbt/journal/v24/n7/full/nbt0706-769.html).

Foreign proteins could be produced in animals or animal cell culture, but the resulting product is often prohibitively expensive, requiring bioreactors that cost many millions of dollars when production is scaled up to commercial levels. Despite that, the majority of the biologics (57%) is produced in different mammalian cell cultures where high quality proteins can be produced (Boehm, 2007). Overall, there is no universal expression system for heterologous proteins. All expression systems have some advantages as well as some disadvantages that should be considered (Table 1.1). Choosing the best one requires evaluating the options – from yield to glycosylation, proper folding, economics of scale-up (Frommer and Ninnemann, 1995; Rai and Padh, 2001), the availability of laboratory facilities, lab expertise, toxicity of expressed protein and so on (Verma et al., 1998).

Table 1. 1Comparison of Several Commonly Used Heterologous Protein Expression Systems

System	Cost	Speed	Glycosylation	Folding	Contamination	Product quality	Scale-up	Storage cost
Bacteria (*E coli*)	Low (Boehm, 2007)	fast	none	worst	endotoxins (Boehm, 2007)	low	high	moderate
Yeast (*S. cerevisiae*)	medium	medium	incorrect (Ma et al., 2003)	medium	low risk	medium	high (Macauley -Patrick et al., 2005)	moderate
Plant	very low	long	minor difference	medium	low risk	high	very low (Fischer et al., 2001)	inexpensive
Cell Culture (mammalian)	high	long	correct	best	Virus, prions and oncogenic DNA	very high	very high	expensive

Reference: http://www.proteinsciences.com/technology/technology_why.htm

1.5.2 Plant Systems

Recent advances in plant molecular biology have led to use of plants as safe and inexpensive alternatives for the industrial scale production of recombinant proteins (Ko and Koprowski, 2005; Vitale and Pedrazzini, 2005). Transgenic plants, as natural bioreactors for the production of industrial and chemical products, obtain increasing attention from the industry (Yoshida and Shinmyo, 2000). As with animals, plants are complex, multicellular organisms and capable of protein synthesis that is more similar to that of animals than those of bacteria or yeast, which are not capable of producing complex proteins. Glycosylation and other post-translational modifications in plants are similar to those in mammals (Faye et al., 2005) and the proteins that are produced in a plant may accumulate to high levels throughout the plant or in specific tissues (Hood and Jilka, 1999; Bailey et al., 2004). The use of plants avoids the risk of contamination with animal pathogens, such as viruses, since to date no plant viruses have been found to be pathogenic to humans. Purification of the desired product is often easier than from bacteria, which can be labour- and cost-intensive (Liu et al., 2005). From an economic point of view, the production cost of biomolecules using plants is much less than that of fermentation processes since transgenic plants can be grown on an agricultural scale using solar energy, CO_2, and inorganic chemicals to drive plant growth. Therefore, the cost of producing crude recombinant protein in plants could be three orders of magnitude lower than that of the mammalian cell system, and 10-fold less than microbial fermentation (Hood and Howard, 2002). Another advantage of using plant expression systems is that plant material can be used directly in food, feed or as an industrial feedstock without purification.

Currently, plant protein expression systems include corn, soybean, canola, alfalfa, tobacco, tomatoes, potatoes, and safflower (Hood, 2004). The foreign protein could be expressed in any organ of the plant, such as tobacco leaf (Schillberg et al., 1999), potato tuber (Beaujean et al., 2001), and rice seed (Yasuda et al., 2006), maize seed (Hood et al., 2003), soybean seed (Robic et al., 2006). However, the low cost of production, rapid scale-up, simple distribution and ease of handling make the seed a favourable organ for producing foreign proteins. For example transgenic soybean seeds expressing the novokinin Arg-Pro-Leu-Lys-Pro-Trp which can reduce systolic blood pressure were generated and the amount of novakinin accounted for 0.5% of total soluble protein (Yamada et al., 2008). Furthermore, high methionine grain legumes were obtained using foreign genes encoding Met-rich proteins (Muntz et al., 1998). We propose to produce enzymes with potential commercial value such as horseradish peroxidase and laccases in soybean seed coats.

1.6 Protein Modification and Targeting

The economics of expressed products are dictated by gene expression patterns and amounts. Expression of genes depends on many factors including transcription, translation, targeting and importantly the ability of the plant to accumulate the protein. In the case of foreign proteins, for example an industrial enzyme such as peroxidase, it is desirable to sequester the protein as much as possible into a specific target tissue (Hood et al., 2003). This could be accomplished with the use of tissue specific promoters that are active only in limited tissue types. A striking example was the use of the endosperm-specific *GluB*1 promoter to express a major house dust mite antigen (Der p 1) in seed endosperm of rice for use as an edible vaccine (Yang et al., 2008). An embryo-preferred promoter was used for the expression of laccase (LCC1) in corn seed (Hood et al., 2003). Here, we are using a

promoter derived from the SBP gene which drives expression in the hourglass cell layer of soybean seed coat.

In addition to promoter specificity, targeting proteins to subcellular compartments such as ER, protein storage vacuole or extracellular space is useful as protein accumulation is favoured by some subcellular locations (Hood et al., 2003). The cellular machinery responsible for targeting proteins is under intense study and has been elegantly reviewed (Vitale and Hinz, 2005; Muntz, 2007). There is no universal protein targeting working model (Jolliffe et al., 2005) but Brandizzi et al (2004) have provided us with a good working model (Figure 1.2). The NTPP (N-terminal propeptide), normally 20-50 amino acids in length with an internal stretch of at least six hydrophobic amino acids, is known to act as signal targeting proteins to the ER. This process involves translocation of the protein through the membrane into the cisternal space of ER accompanied by cleavage of the signal peptide and then *N*-linked glycosylation. Correctly folded proteins then enter the Golgi via dense vesicles. If the NTPP is the only signal present, the protein will be secreted (Vitale and Denecke, 1999) while the presence of the C-terminal extension HDEL or KDEL leads to retention in the lumen of the ER.

The Golgi apparatus plays a major role in post-translational modification of secretory proteins such as oligomers being transformed to glycans or proteins previously glycoslyated in the ER being modified to more complex glycans using several glycosidases and glycosyl-transferases (Kis et al., 2004) . From the Golgi the protein can be transported either to vacuole or to the extracellular space.

Two types of vacuolar sorting signals have been recently described. One sequence can be located anywhere within a protein and shares a conserved amino acid signature NPIR

(Jolliffe et al., 2005; Vitale and Hinz, 2005). Proteins that contain this specific signal are targeted to lytic vacuoles in clathrin coated vesicles (Vitale and Raikhel, 1999). The second type is a C-terminal vacuolar sorting signal which has no conserved sequence (Jolliffe et al., 2005; Vitale and Hinz, 2005). The C-terminal vacuolar sorting signal is believed to be responsible for directing protein to protein storage vacuole (PSV) through dense vesicles. The CTPP (C-terminal propeptide) is cleaved during protein maturation.

Storage vacuoles are found in storage tissues, for instance, seeds, tubers, and their functions are to store proteins for long time periods suggesting that they may be suitable reservoirs for foreign proteins. Studies of the expression of antibody fragments in tobacco and laccase I in maize demonstrated that the secretory pathway-targeted proteins are more stable and accumulate to higher levels than cytosol-localized proteins (Schillberg et al., 1999; Hood et al., 2003).

The development of strategies for manipulating seed coats requires an understanding of the potential signaling peptides that target the corresponding proteins. It is suspected that soybean seed coat peroxidase contains a transit peptide signal sequence (N-terminus) and a vacuolar signal sequence (C-terminus) that would target it to the vacuole via the ER. Two lines of analysis support this model. The primary protein sequence (Gijzen et al., 1993, 1997) suggests their presence and the sequence of mature SBP lacks these sequences as expected if SBP is transported via the ER to the vacuole (Gray et al., 1996; Welinder and Larsen, 2004). Furthermore, the presence of mannose, xylose and fucose suggests that SBP had passed through the Golgi on its way to the vacuole (Gray et al., 1996).

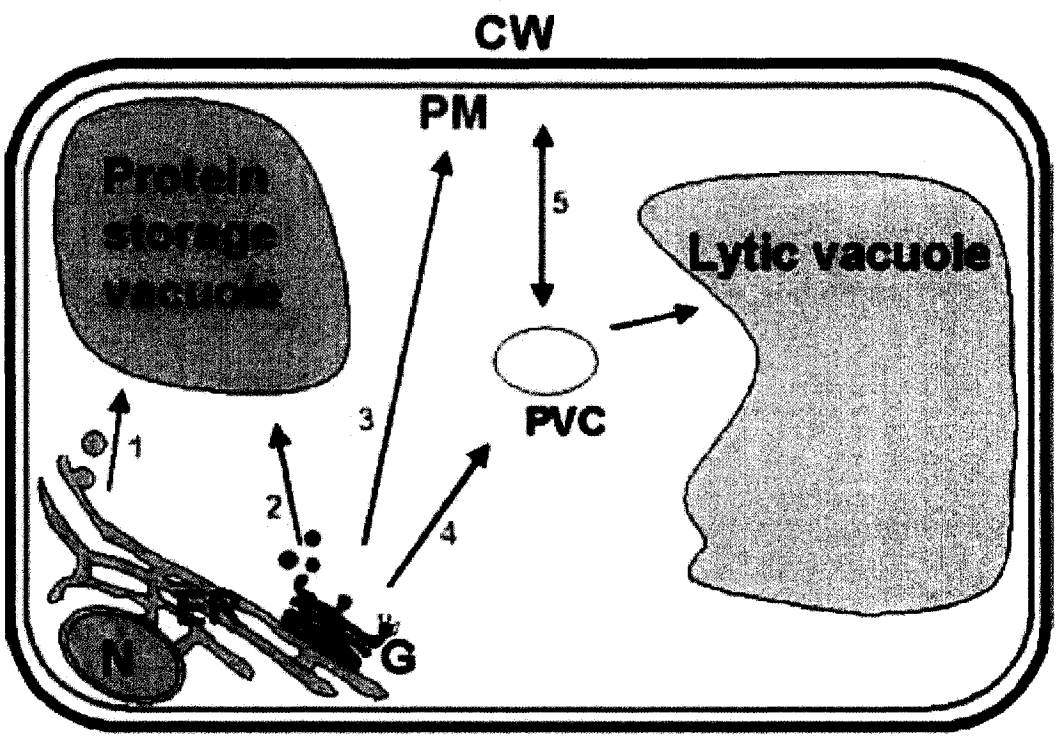
The structure of SBP protein is very unique to soybean. It was found that *Arabidopsis thaliana* (Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids

II; Brassicales; Brassicaceae) peroxidase gene family has no peroxidase orthologous to SBP (Welinder et al., 2002; Duroux and Welinder, 2003). There is no research on the functions of SBP N-, C-terminal peptides up to now. Experimentation is needed to confirm the functions of NTPP and CTPP before SBP signalling peptides can be used to target foreign protein to specific compartments in the hourglass cells. These results can be directly applied to the manipulation of foreign proteins in *epep* mutant soybean line.

The functionality of NTPPs and CTPPs can be tested using model transgenic systems. For example, to investigate the roles of pro-peptides in intact plants, Kis et al. (2004) used tobacco to express different versions of synthetic horseradish peroxidase (isozyme C) f rom *Armoracia rusticana* (Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae). The gene was engineered to include additional sequences coding for either the natural N-terminal or C-terminal extension or both from HRPC. The constructs were placed under the control of CaMV-35S promoter. They found that the N-terminal extension was essential for the production of high levels of recombinant protein mainly in the cytosol, while the C-terminal extension had little effect. In this project, we will verify the functions of Ep signal peptides in order to express our foreign proteins in soybean seed coat. There are two potential pro-peptides in SBP an NTPP of 26 amino acids and a CTPP of 20 amino acids (Welinder and Larsen, 2004). These signalling peptides are removed when SBP matures.

Figure 1. 2 Schematic Diagram of the Endomembrane System in Plant Cells

This figure illustrates the endomembrane organelles of the plant secretory pathway. CW, cell wall; ER, endoplasmic reticulum; G, Golgi apparatus; PM, plasma membrane; N, nucleus; PVC, prevacuolar compartment (Brandizzi et al., 2004).



1.7 Optimizing Expression of Foreign Proteins

The expression of functional proteins in heterologous hosts is a cornerstone of modern biotechnology. For several reasons however, proteins are often difficult to express outside their original context, e.g. expression of mammalian genes in *E. coli* (Feng et al., 2000) which may contain codons that are rarely used in the desired host (Twyman et al., 2003; Gustafsson et al., 2004). Improvements in the speed and cost of gene synthesis have facilitated the design of gene sequences to maximize the likelihood of high protein expression. These strategies include alteration of mRNA structural elements and use of different codon biases. The translation rate can be maximized by codon usage in some transgenes (Koziel et al., 1996) since the efficiency (rate and accuracy) of translation is expected to be proportional to the concentration of the cognate tRNA (Shpaer, 1989; Sorensen and Mortensen, 2005; Hu et al., 2006). For instance, level of the codon-optimised version of the *Plasmodium yoelii* merozoite surface protein 4/5 (PyMSP4/5) expressed in tobacco for controlling malaria accumulated to a level six fold higher than the native *P. yoelii* sequence (Wang et al., 2008).

Since differences in CG content between soybean and fungi (45% in soybean versus 63% in *Trametes versicolor*) (Fungi; Dikarya; Basidiomycota; Agaricomycotina; Agaricomycetes; Agaricomycetes incertae sedis; Polyporales; Coriolaceae) is very large, codon optimization may be critical for protein expression in the soybean seed coat.

1.8 Methods for Measuring Targeting

To visualize protein distribution within individual cells and tissues they are labeled with either a reporter or an antigenic tag, or detected by specific antibodies. Enzyme reporters, such as β-glucuronidase (GUS) (Jefferson, 1989) cannot provide subcellular

protein localization at high resolution or in real time as they are detected indirectly via chromogenic reaction products (Taylor, 1997). The use of antigenic tags, such as the T_7 or HA epitopes and specific antibodies, requires labor-intensive and time-consuming techniques such as immunocytochemistry and electron microscopy. The green fluorescent protein (GFP) originally cloned from the jelly fish *Aequorea victoria* overcomes many of these limitations.

1.9 Green Fluorescent Protein as a Reporter in Plant Cells

GFP can be used and detected in plants, animals, fungi, and prokaryotes (Chalfie, 1995; Brandizzi et al., 2003; Brandizzi et al., 2004). Its advantages reside in its simple and reproducible detection under standardized conditions in a large variety of organisms and cell types. Due to its autofluorescence and high quantum yields, GFP and its spectral variants (Griesbeck et al., 2001) provide sensitive and convenient tools to track biological molecules in real time in yeast, animal and plant systems (Hanson and Kohler, 2001). Furthermore, GFP fluorescence can be both qualitative and quantitative (Lissemore et al., 2000; Mortensen and Larsson, 2001; Gu et al., 2004). Thus GFP is widely used in the study of protein targeting (Zhou et al., 2004).

Usually, GFP is expressed as a chimeric protein either with selected peptide targeting sequences or as a fusion with complete proteins. Expression of GFP constructs can be observed in living material using conventional or preferably, confocal laser scanning epifluorescence microscopes (Brandizzi et al., 2004). GFP has a number of intrinsic properties that make it attractive as a marker for cell biological studies. For instance, it is highly fluorescent and requires no co-factors or substrates to fluoresce. It is relatively stable,

shows no detectable level of cytotoxicity and in many instances does not interfere with the functioning of native proteins when attached to them as chimeric constructs. Moreover, it has been genetically modified to alter its spectral and other properties such as folding rate, to make it more efficient (Cubitt et al., 1995). Overall, these advantages have facilitated protein targeting research (Neuhaus and Boevink, 2001); however potential problems with GFP such as photo-bleaching (Tamura et al., 2003) have been observed.

In our project, we are using pCambia 1302 vector, which contains mGFP5 (absorption 395/475 nm, emission 510 nm). The advantages of mGFP5 are that it does not require splicing, generates equal absorption peaks, and it is more thermo-stable (Neuhaus et al., 2001). The mGFP5 has been widely used in *Arabidopsis* (Yang et al., 2005; Gendra et al., 2004; Taira et al., 2004; Abdenoor et al., 2003) but analysis of large numbers of transformants of rice and *Arabidopsis* at CAMBIA showed that the fluorescence produced by the *mGFP5* protein could be quite faint (pCambia webpage).

Arabidopsis thaliana has been chosen as our model for testing the ability of the putative SBP protein targeting signals to direct GFP to specific subcellular locations. We use *Arabidopsis* leaf mesophyll cells for protein targeting studies because 1) *Arabidopsis* is a widely accepted model plant. More importantly, both soybean and *Arabidopsis* are dicots and being closer in terms of phylogeny, they may share the same mechanism of protein targeting; 2) the sizes of mesophyll cells are relatively large compared with palisade cells, and amount of chloroplasts is relatively fewer (Mr. Ann-Fook Yang, AAFC, personal communication).

1.10 Expressing Heterologous Protein Genes in the Soybean Seed Coat

The *epep* cultivar Jack was chosen as the host since it has little peroxidase background in the soybean seed coat, which makes the detection of HRPC and LCC proteins expressed in soybean seed coat easier. The design of our soybean constructs contained several important features: 1) to restrict gene expression to the HGC the *Ep* promoter was employed. A 1.5kb DNA fragment upstream from the *Ep* gene can restrict GUS gene expression to the soybean HGC (Dr. D Simmonds, unpublished data). 2) The NTPP and CTPP of the *Ep* gene were used to target expressed proteins to the subcellular location occupied by SBP (this thesis; Moïse, unpublished). 3) To aid folding of the expressed proteins, the four amino acid adaptor GGAA was inserted between the NTPP and CTPP and the mature protein (Argos, 1990). 4) A HIS tag was inserted between GGAA adaptor and CTPP to assist protein isolation. 5) The coding sequence of the mature protein along with the signal sequences from the *Ep* gene was codon-optimized for soybean at GeneScript, USA (Batard et al., 2000). The DNA cassette was cloned into the modified pCambia 1300 vector, between *Nco* I and *Bst* E II sites.

1.11 Use of Horseradish Peroxidase and Fungal Laccases to Verify the Seed Coat System

1.11.1 Horseradish Peroxidase Structure and Function

Horseradish Peroxidase (HRPC) belongs to class III of the plant peroxidase super-family that includes SBP. It can use a variety of organic compounds such as phenolics, as electron donors, while peroxide is reduced to water. HPRC contains a heme cofactor in its

active site. HRPC is one of the most valuable and widely used industrial enzymes (Ghasempur et al., 2007). It is used in medical diagnostics because of its superior temperature stability, longer shelf life and high sensitivity (Yoshida et al., 2003) and is widely used in bio-sensing and immunoassays. Due to its ability to oxidize a wide range of substrates there are several areas where it could replace current chemical oxidation process as a substitute for chlorine bleach in the process of pulp and paper bleaching, soil remediation and on-site waste destruction.

Generally speaking, the function and practical applications of HRP are the same as SBP since HRPC and SBP share strikingly similar three-dimensional structures with more than 60% sequence homology (Kamal and Behere, 2003). Although HRPC and SBP have many structural stabilizing factors in common (Gajhede et al., 1997; Henriksen et al., 2001), SBP has a higher melting temperature (86°C vs 74°C for HRPC) (Chattopadhyay and Mazumdar, 2000) and a 20 fold higher catalytic efficiency than HRPC at their pH optima using ABTS as a substrate at 25°C (Kamal and Behere, 2003). Structural differences may account for the difference of their activity and stability (Kamal and Behere, 2003).

1.11.2 HPRC Signalling Peptides

The *HRPC* (M37156) cDNA sequence is 1062 bp (Fujiyama et al., 1988; Fujiyama et al., 1990) and encodes a 354 aa protein (AAA33377) (Welinder, 1979; Gray et al., 1998).. By comparing the HRPC deduced protein sequence with that of the purified protein, two putative signal peptides were identified: an N-terminal propeptide MHFSSSSTLFTCITLIPLVCLILHASLSDA(30aa) and a C-terminal propeptide LLHDMVEVVDFVSSM(15aa), both of which are predicted by bioinformatics (Emanuelsson and von Heijne, 2001). The N-and C-terminal signalling peptide are cleaved

during maturation and both are necessary for efficient sorting of the protein into vacuoles in tobacco BY2 cells (Matsui et al., 2003). HRPC1a derived from a pre-protein lacking CTPP is shunted into the secretory pathway (Matsui et al., 2003; Matsui et al., 2006). The HPRC promoter has been characterized (Kawaoka et al., 1992). The subcellular localization of HRPC in horseradish root has not been reported. Based upon the presence of potential signal peptides (Emanuelsson and von Heijne, 2001) and studies of transgenic HRPC expressed in tobacco cells, it likely accumulates in protein storage vacuoles (Matsui et al., 2003).

1.11.3 HRPC Expression in Heterologous Systems

In horseradish plants 86% of the HRPC activity is present in the tap and lateral roots and 14% activity is found in leaf petiole. The HPRC activity from roots of different cultivars measured using guaiacol as the substrate ranged from 2.07 to 3.71 U/mg (Kushad et al., 1999) which is equivalent to 6.46-11.58 ABTS U/mg (this thesis, Results). HRPC has been expressed in a number of heterologous plant systems, e.g., 0.025-0.3 pyrogallol U/mg in tobacco BY2 cells (Matsui et al., 2006) which is equivalent to 0.042-0.498 ABTS U/mg (Table 1.2). While HPRC has been expressed in other plant systems, a specific activity was not given so direct comparisons cannot be made. When HRPC was expressed in transgenic aspen an activity of 1,408 guaiacol units (4,393 ABTS units) was obtained in leaves, a value 10 times higher than that of WT poplar leaves (Kawaoka et al., 2003).

In this project, we are expressing codon-optimized HRPC in the soybean seed coat.

1.11.4 Fungal Laccase: Structure and Function

Laccases (EC 1.10.3.2) are copper-containing oxidase glycoproteins that are found in many plants, fungi, and microorganisms (Thurston, 1994). They can catalyze oxidation of

various phenolic compounds, aromatic amines and even certain inorganic compounds by reducing molecular oxygen into water. Their substrate versatility makes laccases highly interesting for various applications, including textile dye bleaching, pulp bleaching (http://www.iogen.ca) and bioremediation, where enzymatic catalysis could serve as a more environmentally and benign alternative to the currently used chemical processes (Kiiskinen, 2004).

Fungal laccases show much broader substrate specificity than other members of the family. The essential sequence features of fungal laccases have been analyzed based on multiple sequence alignments of more than 100 sequences (Kumar et al., 2003). There are a set of four un-gapped sequence regions, L1-L4, as the overall signature sequences that can be used to identify the laccases, distinguishing them within the broader class of multi-copper oxidases. Recently, the crystal structure of an active laccase containing a full complement of copper, and the complete polypeptide chain together with seven carbohydrate moieties, was reported (Piontek et al., 2002) . The geometry of the trinuclear copper cluster in the *Trametes versicolor* laccase is similar to that found in other organisms, suggesting a common reaction mechanism for the copper oxidation and the O_2 reduction. In contrast to most blue copper proteins, the type-1 copper in the *T. versicolor* laccase (LCC1) has no axial ligand and is only 3-fold coordinated (Piontek et al., 2002). The LCC1 structure is a monomer, organized in three sequentially arranged domains (Bertrand et al., 2002).

Table 1. 2 Expression of HRPC in Endogenous and in Selected Heterologous Transgenic Systems

Host	Plant Tissue	Substrate used for Assay	HRPC Activity (U/mg)	Equivalent HRPC Activity (U/mg) with ABTS	Reference
Armoracia rusticana	Root	guaiacol	1.1	3.43	(Regalado et al., 1996)
Armoracia rusticana	Root	guaiacol	47.4	147.89	(Miranda et al., 1998)
Armoracia rusticana	Root	guaiacol	2.07-3.71	6.46-11.58	(Kushad et al., 1999)
Tobacco	leaves	pyrogallol	0.43-23	0.71-36.8	(Matsui et al., 2006)
Tobacco	BY2 cells	pyrogallol	0.025-0.3	0.042-0.50	(Matsui et al., 2006)
Tobacco	leaves	guaiacol	1-30 ΔA_{420} units/min /g fresh weight	Cannot be calculated. No protein concentration was given.	(Kis et al., 2004)
Tobacco	leaves	guaiacol	0.8 ΔA_{420} units/min/ 100ul of extract	Cannot be calculated. No protein concentration was given.	(Pellegrineschi et al., 1995)
Tobacco	leaves	guaiacol	10 times higher than WT	Cannot be calculated. No protein concentration was given.	(Heggie et al., 2005)
Aspen	leaves	guaiacol	WT: 140.8; Best transgenic line: 10 times higher	4,393	(Kawaoka et al., 2003)
Insect	Larvae	guaiacol	9	28.08	(Loustau et al., 2008)
Saccharomyces cerevisiae and *Pichia pastoris*		guaiacol, ABTS	5.4-fold higher activity for ABTS and 2.3-fold higher activity for guaiacol.	Cannot be calculated. No protein concentration was given.	(Morawski et al., 2000)

For horseradish peroxidase activity, 1 $U_{guaiacol}$ = 3.12 U_{ABTS}; 1 $U_{Pyrogallol}$ = 1.66 U_{ABTS}. These conversion values were performed in our lab (see Chapter 3).

1.11.5 Laccase Signalling Peptides

Trametes versicolor is one of the best studied white-rot fungi and its laccases have been intensively studied. The cDNAs encoding the *T. versicolor* laccase *LCC*1 (U44430) and a putative laccase *LCC*4(U44431) have been characterized (Ong et al., 1997). *LCC1* from *T. versicolor* encodes an isoenzyme of 519 amino acids preceded by 20aa NTPP (MGLQRFSFFVTLALVARSLA; Lanthier and Yaguchi, unpublished). A putative NTPP of 27aa (MGKFHSFVNVVALSLSLSGRVFGAIGP; Theresa White, Iogen, personal communication) has been suggested for *LCC* IV. These NTPPs might play an important role during laccase secretion. The NTPPs of laccases were replaced with the SBP NTPP and CTPP to direct the expressed laccases to the vacuole of the HGC of soybean seed coat.

1.11.6 Laccases Expressed in Endogenous and Heterologous Systems

Fungal laccase protein sequences are highly conserved (Necochea et al., 2005) suggesting that their activities should be similar (Theresa White, Iogen, personal communication). Up to now, there are several reports of laccase activity from other *Trametes*. Crude extracts from *Trametes trogii* cultures showed laccase activity at 2.4 ABTS U/mg (Garzillo et al., 1998). A laccase purified from *Trametes versicolor* 951022 was found to be much higher at 437.4 U/mg. With a MWt of 90 kDa it is a new laccase different from LCC1 and LCC4 (~60 kDa) (Han et al., 2005). Recently, it was found that the activity of LCC1 in *Pichia pastoris* can be up to 2.6 U/mg (Koschorreck et al., 2008). As mentioned above for HRPC, it is difficult to convert literature values for laccases into

specific activity as these values are expressed as units per volume without any protein concentrations being given.

Laccase, as a ligninolytic enzyme, has been difficult to over-express in active form in heterologous hosts. Detection of active recombinant laccase has been reported in the yeasts *Saccharomyces cerevisiae* (Kojima et al., 1990) and *Pichia pastoris* (Jönsson et al., 1997) and in the filamentous fungi *Trichoderma reesei* (Kiiskinen et al., 2004) and *Aspergillus oryzae* (Yaver et al., 1996). Yeasts and filamentous fungi are generally good hosts for protein secretion but they are more prone to errors in glycosylation, and the quality of expressed proteins is not as good as that expressed in plants. They are not ideal laccase expression systems.

Laccase has been expressed in a number of plant systems, such as maize seed (Hood et al., 2003; Bailey et al., 2004), rice seed (de Wilde et al., 2008), tobacco leaf (Sonoki et al., 2005; Sato and Whetten, 2006) and *Arabidopsis* leaf (Wang et al., 2004). Unfortunately, laccase gene expression in maize seed although regulated by an embryo-preferred promoter was not restricted to corn seed and hindered growth. Cotton plant laccase was also expressed in *Arabidopsis* at 0.005-0.008 U/ mg protein, which was fairly low. *Pycnoporus cinnabarinus* laccase was expressed in rice endosperm at 0.14 U/mg and *Melanocarpus albomyces* laccase activity could be as high as 0.58 U/mg proteins in rice seed endosperm. The heterologous expression of laccase was summarised in Table 1.3.

In this project, we are expressing soy-optimized fungal laccase, LCC1 and LCC4 in the soybean seed coat

Table 1. 3 Expression of Laccase in Endogenous and in Selected Heterologous Transgenic Systems

Hosts	Plant Tissue	LCC Activity (ABTS U/mg).*	References
Trametes versicolor	---	1.3 ΔOD/min for LCC1 & 1.3 ΔOD/min for LCC2 following purification; wavelength at 420 nm.	(Bourbonnais et al., 1995)
Trametes trogii	---	2.4 U/mg for total laccase activities following culture infiltrate; wavelength at 420 nm.	(Garzillo et al., 1998)
Trametes sp. strain AH28-2	---	16.1 U/mg for laccase A, and B following culture infiltrate; wavelength at 420 nm.	(Xiao et al., 2003)
Trametes versicolor 951022	---	437.4 U/mg for one laccase of size at 90 kDa following cell free extract; wavelength at 420 nm.	(Han et al., 2005)
Pichia pastoris	---	2.6 U/mg for the expressed LCC1 (*Trametes versicolor*); wavelength at 420 nm.	(Koschorreck et al., 2008)
Rice	endosperm	*Melanocarpus* laccase: 0.14; *Pycnoporus* laccase : 0.58; wavelength at 420 nm.	(de Wilde et al., 2008)
Maize	embryo	Concentration of the active laccase was determined by comparison to purified recombinant *Trametes* laccase I from *Aspergillus* assuming equal specific activity of these laccases; Fifty ppm dry weight of aqueously extractable laccase was obtained; No laccase activity was specified in this publication.	(Bailey et al., 2004)
Maize	embryo	25-30 ng laccase (*Trametes versicolor* laccase I) per mg dry seed weight; Laccase activity was determined by comparison with purified recombinant *Trametes* laccase I from *Aspergillis.*	(Hood et al., 2003)
Tobacco	root	Laccase (*Coriolus versicolor*) removed 20 μmol bisphenol A or pentachlorophenol per gram dry weight. It did not use ABTS as substrate. It can not be converted	(Sonoki et al., 2005)
Arabidopsis	leaves	Expressed cotton laccase (*LAC*1): 0.005-0.008	(Wang et al., 2004)
Poplar	leaves	Expressed cotton laccase (*GaLAC*1) in poplar, transgenic plants exhibited a 2.1-13.2-fold increased laccase activity than control plant.	(Wang et al., 2008)

* A *Trametes sp.* produces more that one laccase. Thus the total activity of all laccases is measured. When a specific gene is cloned, the activity measure is the activity of the specific transgene

1.12 Thesis Objectives and Hypotheses

The soybean seed coat has been proposed as a vehicle for the production of foreign proteins to enhance soybean seed value (Moïse et al., 2005). The availability of regulatory elements derived from the *Ep* gene and the ability to transform soybean has made this goal feasible for testing. The objective of this thesis was to develop the soybean seed coat as a bioreactor to produce commercially useful proteins. To achieve this goal, I have

1) Tested the ability of Ep signalling peptides to function as targeting peptides using mesophyll cells of *A. thaliana* as a model system;
2) Expressed the proteins laccase I, laccase IV and HRPC in soybean seed coat using vectors designed to optimize expression;
3) Evaluated the ability of soybean seed coats to accumulate foreign proteins;
4) Tested the enzymatic activities of expressed proteins.

CHAPTER TWO: MATERIALS AND METHODS

2.1 Source of Chemicals and Materials

The chemicals and materials used in this project are listed below in alphabetical order, showing catalogue number, and manufacturer.

2.1.1 Sources of Chemicals

Acetone (BDH, #BDH1101-4LP); Agarose (BioShop, # AGA001); Ammonium persulfate (BIO-RAD, #: 161-0700); Anhydrous ethyl alcohol (ethanol)(Commercial Alcohols Inc.); 2,2'-Azino-bis (3-Ethylbenzthiazoline-6-Sulfonic Acid) (ABTS) (Sigma, #: A9941); Acrylamide 30% Solution (BioShop, #ACR009); Bacto Tryptone (Triptona Bacto, #DF0123173); Bacto Yeast Extract (Triptona Bacto, #DF0127179); BisTris (MP Biomedicals, #101038); Bromphenol blue (Sigma, #B-5525); 4-Chloro-1-maphthol (Sigma-Aldrich, #C8890-5G); Coomassie brilliant blue R-250 (PIERCE, #20278); Cupric sulphate ($CuSO_4$) (Fisher Biotech, #C49010); 3, 3-Diminobenzidine (Fluka, #32741); 2, 7-Diaminofluorine (Aldrich, #D17106-1G,); 1, 8-Diaminonaphthalene (Fluka, #33170); Dimethyl Sulfoxide (DMSO) (Fisher Biotech, #BP231-1); Ethidium bromide (Fisher Biotech, #BP102-5); Glycerol (enzyme grade) (Fisher Biotech, #BP229-1); Glycine (Fisher Chemicals, #G46-500); Guaiacol (Sigma, #G5502-250G); Hydrogen peroxide (H_2O_2) (Sigma, # H-1009); Javex (Bleach) (Colgate-Palmolive Inc.); 2-Mercaptoethanol (β-Me) (Sigma-Aldrich, #M6250-100ML); Methanol (Fisher Scientific, #A452-4); Murashige and Skoog Basal (MS salt) (Sigma, #M5524); Non-fat dry milk, local supermarket; Ponceau S (Sigma, #P3504); Potassium dihydrogen orthophosphate (KH_2PO_4) (BDH Chemicals, #ACS657); Propidium iodide (Calbiochem, #537059);

pyrogallol (Sigma, # P0381); Sodium azide, (Fluka, through Sigma, #S2002); Sodium chloride (EMD, #SX0420-3); Sodium Dodecyl Sulfate (SDS) (Bio-Rad, #161-0302) ; Sodium phosphate diabasic ($Na_2HPO_4 \cdot 2H_2O$) (Fisher Scientific, #S-374); Sucrose (EMD, #SX1075-3); Syringaldazine (Sigma, #S7896-1G); Temed (BIO-RAD, #TEM001); Trichloroacetic Acid (TCA) (Sigma, # T 9159); Tris (BioShop, #TRIS 001).

2.1.2 Sources of Materials

Agrobacterium tumefaciens strain (Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group) GV3101 competent cells (Dr. Brian Miki's lab, AAFC, prepared according to Hellens et al., 2000); Ampicillin (Sigma, #A9518-25G); Tetra-HIS HRP Conjugate (QIAGEN, #34470); Anti-HRPC antibody (Abcam, #: ab2115); Anti-LCC1 antibody (Southern Alberta Cancer Research Institute [SACRI] Antibody Services, University of Calgary); Anti-LCC4 antibody (Southern Alberta Cancer Research Institute [SACRI] Antibody Services, University of Calgary); BugBuster HT (Novagen, #70922-4); Coomassie (Bradford) Protein Assay Kit (PIERCE, #1856209); DNeasy Plant Maxi kit (QIAGEN, #68163); DNeasy plant mini kit (50) (QIAGEN, #69104); Subcloning efficiency DH5α chemically competent *E. coli* cells (Invitrogen, #18265-017) ; *E coli* Top 10 cells (Invitrogen, # C4040-10,); Flamingo fluorescent gel stain (QIAGEN, #161-0490); Goat anti-mouse IgG AP conjugate (Sigma, # A-2429); Goat anti-mouse IgG HRP conjugate (BIO-RAD, # 170-6516); Goat anti-rabbit IgG AP conjugate (BIORAD, #: 170-6518); Goat anti-rabbit IgG HRP conjugate (Jackson, #111-036-003); (GeneRuler^{TM} 100 bp DNA ladder plus (Fermentas, #: SM0321); GeneRuler^{TM} 1kb DNA ladder plus (Fermentas, #SM0311); HRPC (Sigma, # P-8250) HRP Color Development Buffer (BIO-RAD, #L9702833 Rev B,); HRP Color

Reagent A (BIO-RAD, # L9701118 Rev C,); HRP Color Reagent B (BIO-RAD, #L9701120 Rev D,); Hygromycin B (Roche, #10843555001); Kanamycin A (Sigma, #K1372); Laccase from *Trametes versicolor* (Biochemika, #38429); LCC1, LCC4, and HRPC synthetic genes (GenScript, USA); NBT/BCIP Reagent Kit (Molecular Probes, #N6547); OneStep RT-PCR kit (QIAGEN, #: 210212); pCAMBIA1300 and other pCAMBIA plasmids (www.cambia.org); pGEMT-Easy plasmid (Promega); pPRX-1 Plasmid (Dr. Mark Gijzen, London, ON, AAFC); pTrcHis2 A (www.invitrogen.com, gift from Dr. Steve Gleddie); pLacZi (Clontech); Precision plus protein all blue standards (BIO-RAD, #161-0373); Precision protein all blue standards (BIO-RAD #:161-0318); Protease inhibitor cocktail (Sigma, # P9599-1ML); PVDF membrane (Fisher, #: BP172-5); QIAexpress® Ni-NTA Fast Start kit (QIAGEN, # 30600); QIAquick Gel Extraction kit (QIAGEN, #28706); QIAquick PCR Purification Kit (50) (QIAGEN, #28104); QIAEX II Gel Extraction Kit (QIAGEN, #20051); REDExtract-N-Amp™ Plant PCR kits(Sigma, #XNAP-1KT); RNeasy Plant Mini Kit (QIAGEN, #74905); RNase-free DNase set (50) (QIAGEN, #79254); Restriction enzymes (Fermentas); SBP (Sigma, # P1432); Taq DNA Polymerase (Fermentas, #EP0402); Soybean cv. Harosoy 63, seeds (Dr. E. Cober, AAFC); Soybean cv. Jack seeds (Dr. D. Brown, AAFC); T4 DNA Ligase (Promega, #M1804); tCUP promoter (Dr. B. Miki, AAFC); Timentin (3.1 g/vial, Glaxo Smith Kline); Wizard® Plus Minipreps DNA Purification System (Promega, #A7100).

2.2 Plant Materials and Propagation

2.2.1 *Arabidopsis* Transformation and Growth

The *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* (ecotype Columbia) was performed by the floral dip method (Clough and Bent, 1998). Seed was

normally collected after 3 weeks. All *Arabidopsis* plants were grown in a Conviron E15 growth cabinet under the following conditions: 16 h, 22°C, 100 μE m^{-2} s^{-1} day / 8h, 22°C night; relative humidity of 50%.

The seed sterilization procedures were as follows: 1 ml of 70% ethanol was added into an Eppendorf tube containing 100-200 μl of seeds. After shaking for 2 min on a shaker, the tube was spun for 30 sec to collect the seeds. The ethanol was removed and 1 ml of 50% bleach was added. Seeds were re-suspended by vortexing. After shaking for 10 min, seeds were collected by centrifugation and the bleach was removed. Then 1 ml of autoclaved, distilled water was added, the tube was shaken for 2 min and the seeds were collected. Washing with water was repeated 2 more times. Finally the seeds were re-suspended in 8-12 ml of 0.1% agarose. Transformed plants were selected by spreading on half-strength MS medium (Murashige and Skoog, 1962) containing 1% sucrose and 0.8% agar, pH 5.7-5.8, supplemented with 30 μg/ml of hygromycin and 300 μg/ml of timentin. After 2-3 weeks of growth on selective medium transgenic plants were transferred to soil and grown in a Conviron E15 growth cabinet.

2.2.2 Soybean Transformation and Growth

Embryogenic suspension cultures of soybean [*Glycine max* (L.) Merrill cv. "Jack"] were initiated and transformed by particle bombardment (Finer and McMullen, 1991) as modified by Dr. Sijun Zhou in the laboratory of Dr. Dan Brown (S. Zhou, D. Brown, manuscript in preparation). One to two weeks after bombardment, embryogenic tissue was placed in a liquid proliferation medium containing hygromycin. Four to six weeks after bombardment, lobes of yellow-green, hygromycin-resistant tissue, which began as outgrowths on brown clumps of hygromycin-sensitive tissue, were isolated and cultured

to give rise to clones of transgenic embryogenic material. Transgenic, fertile plants can be routinely produced from the proliferating transgenic embryogenic clones. The presence of the *hpt*II gene was verified by PCR (Tara Rintoul, personal communication).

Seeds from soybean cultivars including Harosoy 63, Jack and all transgenics, were germinated in moist vermiculite in a growth cabinet (Conviron E15) for 7 days until soybean plant had 3-4 trifoliate (ternate) leaves and then was transferred to soil in 12-inch pots. All soybean plants grown from seedlings were grown under the following conditions: 14 h, 25°C, 300 μE m^{-2} s^{-1} day / 10h, 20°C night; relative humidity of 80%. After one week the light period was shortened to 13 h, and two weeks later, the light period was reduced to 12 h until the plants set seeds (Dr. D. Simmonds, personnel communication), at which time the relative humidity was reduced to 70%. At approximately 5 weeks after flowering, watering was terminated to allow for seed drying (Miller et al., 1999).

2.3 DNA Isolation and Manipulation

2.3.1 Isolation of Plasmid DNA

The Promega Wizard Prep Kit was used to isolate plasmids from *E. coli* liquid cultures according to the manufacturer's recommendations. DNA was eluted from the resin with 50 μl of sterile water or TE buffer supplied by the Promega. DNA concentration was quantified using Spectronic BioMate 3 and 1.5% agarose gel electrophoresis was used to confirm the presence of plasmid DNA. Samples were stored at –20°C for further use.

2.3.2 Isolation of Genomic DNA

Genomic DNA was isolated from plant tissues (usually leaves) using the QIAGEN DNeasy Plant Mini Kit according to the manufacturer's recommendations. Typically 100 mg (wet weight) of tissue were processed (DNeasy Plant Mini, and Maxi Handbook, www.QIAGEN.com). DNA was eluted from the resin by DNA elution buffer (buffer AE) provided in the kit. To get high yield of DNA, a second elution was performed and the two eluates were combined. DNA concentration was quantified at 260 nm using Spectronic BioMate 3. Samples were stored at –20°C for further use.

2.3.3 Restriction Digestion and Gel Electrophoresis

Typically, restriction digestions were performed for one hour at 37°C in a final volume of 50 µl containing 0.5-1.0 µg of DNA, the specific reaction buffer as indicated by the manufacturer, BSA with final concentration at 1 µg/ml, and 1 µl of restriction enzyme (5-10 units). The DoubleDigest™ WEB page was referenced (http://www.fermentas.com/doubledigest/index.html) if two restriction enzymes did not share the same restriction enzyme buffer. After digestion 1% agarose gel electrophoresis was used to obtain the inserts (Sambrook et al., 1989); the band(s) of interest were cut out of the gel and stored in -20°C for future purification.

2.3.4. Purification of DNA Fragments from Agarose Gels

The gel slice containing the DNA band of interest was excised with a clean, sharp scalpel. DNA was isolated using the QIAquick or QIAEX II Gel Extraction Kit, depending on the fragment size, according to the manufacturer's recommendations. The DNA was eluted using 50 µl of the supplied buffer or water. The purified DNA

fragments were stored in -20°C for cloning or other uses after quantifying DNA concentration at A260 nm by using a Spectronic BioMate.

2.3.5 PCR Product Purification

PRC products were purified with the QIAquick PCR purification kit according to the manufacturer's recommendations. The DNA was eluted using 50 µl of the supplied buffer or water. The purified DNA fragments were stored in -20°C for cloning or other uses after quantifying DNA concentration.

2.3.6 Ligation

DNA fragments generated by restriction enzyme digestion or PCR products digested with restriction enzymes were ligated into pCAMBIA1302 or modified pCAMBIA1301 vectors. A1:3 ratio of vector: insert (concentration) was generally used to maximize ligation. Typical ligation conditions were 1x Rapid Ligation Buffer (60 mM Tris-HCl, pH 7.8, 20 mM MgCl2, 20 mM DTT, 1mM ATP 10% PEG, Promega), 50 ng of vector DNA, 150ng of insert, 3 Weiss units of T4 DNA ligase in a final volume of 10 µl. The ligation reaction was incubated overnight at 4°C and stored at –20°C until needed for transformation.

2.4 PCR Reaction Conditions

2.4.1 Standard PCR

For standard PCR the reaction volume was 20 µl and contained 1X Taq buffer (100mM Tris.HCl pH 8.3, 500mM KCl, 15mM $MgCl_2$, 10X PCR buffer), 0.2 mM each dATP, dCTP, dGTP, dTTP, 0.5 µM each forward and reverse primers, 1 U of Taq DNA polymerase, and 0.5 µg of template DNA. The annealing temperature was optimized by

gradient PCR. Extension time was estimated based on the synthesis speed for Taq polymerase of 1000 bp per min (rscott.myweb.uga.edu/protocols/CEP_13.pdf). The number of cycles was 30-35.

2.4.2 Colony PCR

To screen *E. coli* colonies for inserts, colonies were grown 15 hours (overnight) at 37°C. A single, labelled colony picked with either a sterile loop or pipette tip and added to the PCR reaction mix as the template DNA. The PCR recipe and conditions were the same as the standard PCR.

2.4.3 Transgene Screening

Genomic DNA was isolated from leaves using the REDExtract-N-Amp™ Plant PCR kits (Sigma) according to manufacturer's recommendations. Transgenic plants were PCR screened for different transgene elements such as Ep promoter, transgene itself, NOS terminator, tCUP promoter and hpt II gene. The primers for are found in Table 1.

2.5 Bacterial Transformation

2.5.1 *E. coli* Transformation

DH5α *E. coli* competent cells were thawed on ice for 10 min. Plasmid DNA (about 0.5 μg) was mixed gently with 50 μl cells and incubated on ice for 30 min. Following a heat shock at 42°C for 45 sec, the mixture was placed on ice for 2 min. Then 1 mL of SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) was added to the transformed cells and the mixture was incubated at 37 °C for one hour with shaking. For selection, aliquots were plated on freshly prepared LB plates (7.5 g agar per 500 ml Luria-Bertani broth) containing

50mg/L kanamycin. Plates were incubated overnight at 37°C, and colonies were selected for preparing liquid cultures to isolate the recombinant plasmids or for colony PCR.

2.5.2 *Agrobacterium tumefaciens* Transformation

Agrobacterium tumefaciens competent cells (GV3101) were thawed on ice for 10 min. Plasmid DNA (1 μg) was mixed gently with 100 μl cells. The mixture was frozen in liquid nitrogen for 5 min, and then thawed for 5 min at 37°C. After 1 ml of LB was added the mixture was transferred to a 15ml Falcon tube and incubated for 2 h on a roller drum at 30°C. The mixture was poured into 1.5 ml microfuge tube, and the tube was spun for 5 min to pellet cells. The supernatant was removed and the pellet was resuspended in 100 μl LB. All of the suspension was plated on LB plates containing 50mg/L kanamycin and incubated at 30°C. Transformed colonies were visible on the fourth day of incubation.

2.6 Plasmid Construction

2.6.1 GFP Constructs for Protein Targeting

Four constructs were made to test the ability of putative signal peptides from the Ep gene to target GFP to their predicted subcellular locations in *A. thaliana*. The base vector was pCAMBIA1302 containing the *mgfp5* version of the *Aequoria victoria* green fluorescent protein (Siemering et al., 1996). pCAMBIA1302 was cut with *Sph* I, which resulted in a DNA fragment containing LacZ MCS, 35S promoter, mGFP, HIS tag, and NOS poly A. The resulting DNA fragment was subcloned into the *Sph* I site of pGEMT-Easy to create intermediate plasmid pGEMT-GFP. To generate the CTPP fusion a 60 bp fragment encoding the C-terminal region of SBP was generated by PCR of the plasmid pPRX-3 (Ep peroxidase cDNA cloned into pCMV-BK; gift of Dr. M. Gijzen, London,

AAFC) using the primers PRX-gfp3-For and PRX-gfp3-Rev (Table 2.1). To create the GFP3'prx construct the PCR product was purified by QIAGEN PCR Purification Kit, digested with *Nhe*1 + *Bst*E II for CTPP and ligated into pGEMT-GFP digested with the same enzyme pair. To create 5'prxGFP construct, the N-terminal targeting signal pro-peptide (NTPP) DNA fragment was amplified by PCR using the pPRX-3 as the template and the primer pair PRX-gfp5-For + PRX-gfp5-Rev. The resulted PCR products were digested with *Nco* I + *Spe* I, and were subcloned in the pCAMBIA 1302 vector cut with the same restriction enzymes. To create the 5'prxGFP3'prx construct, we used *Xba* I and *Spe* I to cut 5'prxGFP and cloned the released fragment into GFP3'prx linearized with the same two enzymes.

In total, four GFP constructs were obtained, namely, A) pCAMBIA1302 that was a control vector, without any Ep NTPP or CTPP; B) GFP3'prx with the CTPP fused with GFP; C) 5'prxGFP with the NTPP fused with GFP; and D) 5'prxGFP3'prx with both NTPP and CTPP fused with GFP (Figure 3.2).

2.6.2 Ep (S) Series Constructs for Soybean Transformation

The pCAMBIA series of plasmids used in this project are described at http://www.cambia.org/daisy/cambia/materials/vectors.html. pCAM1301-tCUP plasmid was made by replacing the CaMV 35S promoter that drives expression of hpt II gene by the tCUP promoter (Malik et al., 2002). The peroxidase promoter (Prx) was synthesized based on the Prx sequence (Gijzen, 1997) at NRC-Plant Biotechnology Institute (PBI), Saskartoon and cloned into *Hind*III/*Nco*I sites of plasmid pCAM1301-tCUP to give the base plasmid S8. In S8 derivatives the order of the expression cassette was hpt II←tCUP - NOS -synthetic gene ←prx NOS. In S18 derivatives, ~2.7kb fragment derived from

pLacZi (Clontech) was inserted between the tCUP and Prx promoters to give the order hpt II←tCUP - LacZ - prx→ synthetic gene – NOS. These constructs were built by Dr. Loreta Gudynaite-Savitch.

To improve the potential expression efficiency of HRPC, LCC1 and LCC4 in soybean, the codon usage of the transgenes was optimized to meet the codon bias of soybean genome (Kotula and Curtis, 1991; Chang et al., 2006, 2006; Hu et al., 2006). The three transgenes were synthesized by GenScript Corporation (USA) (http://www.genscript.com). The synthetic genes in pUC57 were digested with *Nco*I + *Bst*E II and the fragment cloned into the above vectors. Fragments cloned into S8 result in the plasmid series S8-HRPC, S8-LCC1, S8-LCC4; fragments cloned into S18 result in the plasmid series S18-HRPC, S18-LCC1, S18-LCC4 (Figure. 3.3).

2.6.3 Laccase Constructs for Antibody Production

The base vector used for laccase cloning was pTrcHis2 A (Invitrogen, http://www.invitrogen.com) a gift from Dr. Steve Gleddie. Primers were designed to amplify DNA encoding the mature protein parts of LCC1 (Ong et al., 1997) and LCC4 (Dr. Theresa White, Iogen, unpublished data). LCC1 DNA was amplified with the primer pair LCC1-*Nco* I-OP-LCC1 + LCC1-*Sal* I-OP- LCC1 with the synthetic LCC1 as the template, and the product purified by QIAGEN PCR purification kit. After digestion with *Nco*I + *Sal*I, it was cloned into pTrcHis2 A digested with *Nco*I + *Sal*I (http://tools.invitrogen.com/content/sfs/vectors/ptrchis2a_mcs.pdf). Mature LCC4 was cloned in the same manner but with the template synthetic LCC4 and the primer pair LCC4-*Nco* I-For and LCC4-*Sal* I-Rev (Figure 3.1).

2.6.4 Laccase Constructs for *Arabidopsis* Transformation

The base vector for LCC expression in *Arabidopsis* was pCAMBIA1302. Plasmid DNA from S18 LCC1 or S18 LCC4 was cut with *Nco* I -*Bst* E II. The inserts were separated on an agarose gel, purified using QIAGEN gel purification kit and subcloned into pCAMBIA1302 resulting in two 35S::synthetic LCC gene-NOS cassettes designated pCAMBIA1302-LCC1 and pCAMBIA1302-LCC4 (Figure 4).

2.7 Techniques for Transgene Analysis

2.7.1 Confocal Microscopy

Confocal microscopy involved the use of a Zeiss Meta 510 microscope (Agriculture and Agri-Food Canada, Ottawa) fitted with a 63X 1.2 water-immersion objective (C-Apochromat, Zeiss) and Argon ion laser. Fresh leaf squares 0.2cm X 0.2 cm were cut using a shaver blade and cell membrane was stained using 8.2 mM FM4-64 (T13320, Molecular Probes) for 10 min. Samples were mounted on glass slide and 0.17mm cover slip in 0.01% toluidine blue as a background auto fluorescence counterstain. The slides were excited with the line 488 nm (6.2A, 28.8 % power) and emission was detected at 505 to 530 nm for GFP and at 575 to 750 nm for cell membrane (AAFC, Ottawa). This work was done in collaboration with Dr. Jaimie Schnell.

The confocal images taken at University of Saskatchewan used the same settings except FM4-64 was not added in order to stain plasma membranes.

2.7.2 Assay of Peroxidase and Laccase Enzyme Activity

The specific activity of peroxidase or laccase was calculated by determining the rate of oxidation of 2,2'-Azino-bis (3-Ethylbenzthiazoline-6-Sulfonic Acid) (ABTS). As

ABTS is oxidized, it changes colour from clear or pale green to an intense blue-green (Pütter and Becker, 1983). One unit of activity produces 1.0 μmol of oxidized ABTS per minute at the appropriate pH at 25°C. The complete protocols are given in Appendix 2 and Appendix 3. The Honours student, Ms. Sadia Karim, helped to screen HPRC transgenic soybean seeds for peroxidase enzyme activity.

2.8 Protein Extraction and Quantification

2.8.1 Isolation of Protein from *E. coli* for Antibody Production

The *E. coli* proteins (both soluble proteins and inclusion bodies) were extracted using BugBuster Protein Extraction Kit (Novagen) based on manufacturer's recommendation. To 100 ml of LB, 1 ml of overnight cultured LCC transformed *E coli* was added. The absorbance value reached around 0.6 after culturing at 37°C with shaking at 250 rpm. Then production of LCC proteins was induced with 1 mM IPTG. The 2 ml of pTrcHis2 A LCC1 or pTrcHis2 A LCC4 liquid culture was taken out before IPTG induction, and 1, 2, 3, 4 h after IPTG induction. The soluble and insoluble proteins were extracted from pTrcHis2 A LCC1 or pTrcHis2 A LCC4 *E coli* culture, and SDS gel analysis was performed with 15 μg proteins in each lane. HIS tagged proteins were purified by using QIAexpress Ni-NTA Fast Start kit according to manufacturer's recommendations. Then, the expressed protein bands were purified by SDS-gel electrophoresis followed by excision from the gel. The gel fragment was sent to SACRI for antibody production. The titres of LCC antibodies raised in rabbit were verified in Calgary by Sampson Law.

2.8.2 Isolation of Protein from Plant Tissue

Protein extracts were obtained by grinding 100 mg of plant tissue (soybean seed coat, embryo or *Arabidopsis* leaves) in liquid nitrogen with the aid of a homogenizer (IKA Werke). For the peroxidase activity assay, 150 µl of 0.067 M phosphate buffer (0.144 g $Na_2HPO_4 \cdot 2H_2O$ + 0.798 g KH_2PO_4 in 100 mL deionized water, pH to 6.0) was used with 1 µl protease inhibitor cocktail (Sigma) added to prevent protein degradation. The protein extraction buffer for laccase activity assay was 0.1 M sodium acetate pH 5.0 (1.36 g sodium acetate in 100 ml deionized water). The extract was centrifuged at 15,000 rpm for 15 min at 4°C. Supernatants were stored at -20°C until required for enzyme assay or gel electrophoresis.

2.8.3 Protein Concentration Quantification

Protein concentration was measured on Spectronic BioMate 3 (Thermo Electron Corporation) using Coomassie (Bradford) Protein Assay Kit (Pierce). For the standard curve, a stock Bovine Albumin Standard (BSA) was diluted in 0.067 M phosphate buffer to give a range from 0 to 2,000µg/ml. Absorbance at 595 nm was measured using a Spectronic BioMate 3. The BSA standard curve was used to determine protein concentrations of samples.

2.8.4 Protein Electrophoresis and Western Transfer

Protein extracts were concentrated by the addition of 100 % trichloroacetic acid (500 g TCA into 350 ml ddH_2O) to achieve a final concentration of 20% TCA. After incubating at 4°C for 10 min the pellet was collected by centrifugation at 14, 000 rpm for 5 min in a cold room (4°C). The supernatant was removed, the pellet was washed using

200 ul of cold acetone for 5 times, collected by centrifugation as before and dried at 95°C for 5 min. Phosphate buffer was added to achieve the final concentration of 2 mg/ml.

Proteins were separated on a 12% SDS-PAGE gel in a BioRad mini gel apparatus (Laemmli, 1970). The proteins were denatured by 5 min-incubation at 100°C heat-block. Electrophoresis was carried out for 10 min at 60 V and then 100 V for 2 h. Following electrophoresis, the gel was removed from the apparatus and equilibrated in a Western transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3) for 30 min at room temperature. For transfer, a cassette consisting of two fibre pads, two filter papers (all soaked in Western transfer buffer for 10 to 20 minutes), the SDS gel and a PVDF membrane (wetted in 100% methanol, then soaked in Western transfer buffer) was assembled in a BioRad Gel transfer apparatus. Transfer to PVDF membrane was achieved using 100 V (350 mA) for 1 h. The apparatus was then disassembled, the membrane was rinsed in 1X PBS (10X PBS is 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4, 2.4 g KH_2PO_4, water to 1 L, pH 7.4) for 5 to 10 minutes and then overnight at 4°C in Western blocking solution (5% nonfat dry milk in PBS) to block non-specific binding on a shaker. The next day, the membrane was washed 3X for 5 min each in PBST (PBS with 0.2% Tween-20).

Plant protein samples were analysed using either commercially available antibodies such as QIAexpress Anti-HIS tag antibody (QIAGEN) (conjugated with HRPC), and rabbit polyclonal anti-HRPC antibody (ab2115, Abcam) or those prepared specifically for this project (laccase antibodies produced in rabbit). The membrane was incubated for 1 h at room temperature in the primary antibody diluted to the required concentration in antibody buffer (5% nonfat dry milk in PBST, 0.02% sodium azide) on a

shaker. Following 3 washes for 5 min each in PBST, the membrane was incubated for 30 min at room temperature with the secondary antibody (conjugated with AP or HRPC) diluted with 5% nonfat dry milk in PBST to the required concentration. The membrane was washed 3X for 5 min in PBST, then washed in PBS for 5 min to remove the Tween-20 from the surface. To detect alkaline phosphatase the membrane was overlaid with NBT/BCIP solution for several minutes until sufficiently developed. To detect HRPC, HRP Color Reagent A, and B were used. To stop colour development, the membrane was submerged in water and rinsed 2X for 5 min in water. Membrane was photographed and dried on filter paper for permanent storage.

2.8.5 DNA Sequencing and Bioinfomatics

The sequences of all clones were verified by DNA sequencing at PBI (DNA Technologies Unit, Plant Biotechnology Institute, NRC, Saskatoon, SK, Canada). Both strands were sequenced. For DNA analyses, the software was DNAMAN 5.2.10 (Lynnon Corporation), DNA Star (DNASTAR), NTI 10 (Invitrogen) and free software in NCBI website (online).

Table 2. 1 Primers Used in This Project

Primer Name	Sequence	Usage
Ep 1.5 F32	5'-TAGATAAAAAAATGGGATATAATTTTTCTCAG 3'	Transgene Screening
Ep-1.5 F	5'-CGCAAGCTTTAGATAAAAAAATGGGATATAATTTTTC-3'	Transgene Screening
Ep-1.2 F	5'-CGCAAGCTTTTGTAATCTCACCTTTTTCATTTAA-3'	Transgene Screening
Ep-0.9 F	5'- CGCAAGCTTATTTGTACCATCATTATTTCCC-3'	Transgene Screening
Ep-0.6 F	5'- CGCAAGCTTCAATGTTCCAAAACCTAATGC-3'	Transgene Screening
Ep-0.3F	5'- CGCAAGCTTCCAACCACATTTAAGAGATTATAG-3'	Transgene Screening
Ep-Ori-Rev	5'-CCCATGGTTTGAGTTAATATG-3'	Transgene Screening
HptII (99) For	5' GATTTGTGTACGCCCGACAGT 3'	Transgene Screening
HptII (614) Rev	5' CCGCAAGGAATCGGTCAAT 3'	Transgene Screening
HRPC-RT-F	5' TGACATTGGCTGGAGGTCCT 3'	Transgene Screening
HRPC-RT-R	5' AGCCACAGAAGCAAGTCCGA 3'	Transgene Screening
HRPC (361)Rev	5' TGACTCCACAGCAGCCTTCA 3'	Transgene Screening
HRPC(1047his)Rev	5' TGATGATGATGATGATGAGC 3'	Transgene Screening
HRPC (243) Rev	5' TCGCAACCATTCACGAAG 3'	Transgene Screening
HRPC-Not1-For	5' TTGCGGCCGCATGGGTTCTATGAGGCTTTTGGTGGTT 3'	Cloning, screening
HRPC-Xba1-Rev	5' AA TCTAGA TT ACTTAGATTG AGCCACCAAC TTTTG 3'	Cloning, screening
LCC1-RT-F	5'ACACCGGGCGGAGTTGATAA 3'	Transgene Screening
LCC1-RT-R	5'ATGGTGATGGTGATGAGCGG 3'	Transgene Screening
LCC1(HIS)RT-Rev	5' TGATGGTGATGGTGAGCGGCT 3'	Transgene Screening
LCC1-*Nco* I-OP	5' TTCCATGGCTATCGGACCTGTGGCAA 3'	LCC1 Cloning
LCC1-*Sal* I - OP	5' TTGTCGACTTGGTTCGCTTCAGAAAG 3'	LCC1 Cloning
LCC1 (198) Rev	5' AGTGGCGATGGAACAACAC 3'	Transgene Screening
LCC1-Not 1 For	5' TTGCGGCCGCATG GGTTCCATGCGCCTGCTGGTGGTG3'	Cloning, screening
LCC1-Xba 1 Rev	5' AATCTAGATTACTTTGATTGGGCAACCAGTTTTTG3'	Cloning, screening
LCC4-RT-F	5'GTTCCACCAACTGTGCCTGTT 3'	Transgene Screening
LCC4-RT-R	5' AAGCCAAACCGAAATGGTGA 3'	Transgene Screening
LCC4(1622his)Rev	5'GGTGATGATGATGGTGTGCAG 3'	Transgene Screening
LCC4-*Nco* I-For	5'AACCATGGTTACTGATCTTACAATTTCC 3'	LCC4 Cloning
LCC4-*Sal* I-Rev	5' AAGTCGACCAAATCGGAGGAATCAAGAGC 3'	LCC4 Cloning
LCC4 (214) Rev	5' GAACTCGTCACCTTTGTTTCC 3'	Transgene Screening
LCC4-Not1 For	5' TTGCGGCCGCATGGGTAGCATGAGACTTCTC3'	Cloning, screening
LCC4-Xba1 Rev	5' AATCTAGATCACTTAGACTGAGCCACCAACTTCTGTTT 3'	Cloning, screening
PRX-gfp5-For	5'-ACTGGCCATGGGTTCCATGCGTCTATTAG-3'	GFP cloning
PRX-gfp5-Rev	5'-CCACTAGTAAGCTGAGCATAAGAGAC-3'	GFP cloning
PRX-gfp3-For	5'-CAAGCTAGCAATGGAGACTCGTTTGGA3'	GFP cloning
PRX-gfp3-Rev	5'-CAAGGTGACCTTATTTAGATTGAGCAACAAG3'	GFP cloning
NOS (13) For	5' AATCCTGTTGCCGGTCTT 3'	Transgene Screening
NOS (192) Rev	5' TTATCCTAGTTTGCGCGCT 3'	Transgene Screening
SoyTub2For	5' GTGACTTGAACCATCTGATCTAGC 3'	Assess DNA quality
SoyTub2Rev	5' GTTGAAGCCATCCTCAAGCCAG 3'	Assess DNA quality
tCUP(20) For	5' GGGATCTTCTGCAAGCATCTC 3'	Transgene Screening
tCUP(432) Rev	5' TGGCGGTTGGTGAGAAAT 3'	transgene Screening

CHAPTER THREE: RESULTS

In this chapter, prior to a description of the results, we first describe the DNA constructs used for transgenic work (3.1) and then the validation of the tools and reagents that were generated for the analyses of these transgenics (3.2).

In all cases, the sequences of the transgenes within the vectors were verified by sequencing at Plant Biotechnology Institute (PBI), Saskatoon.

3.1 Vector Constructions

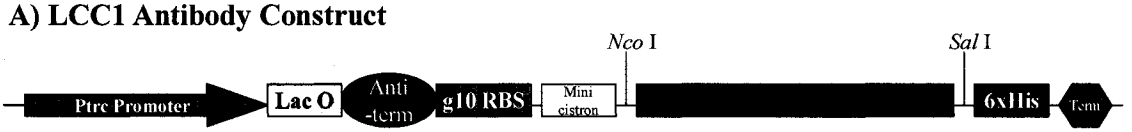
3.1.1 Constructs for Antibody Production

We attempted to express high levels of laccases I and IV in *E. coli* so that the purified proteins could be used as antigens for the production of specific anitsera that could be used for protein analyses. Based on the DNA sequences of the synthetic LCC1 and LCC4 (Appendix 1), we created primer pairs (LCC1-*Nco* I-OP/LCC1-*Sal* I–OP and LCC4-*Nco* I-For/ LCC4-*Sal* I-Rev) to amplify ~1.5 kb DNA fragments which correspond to mature LCC1 protein (499aa) and LCC4 mature protein (500aa) respectively. The PCR products were purified, digested with *Nco* I and *Sal* I restriction enzymes and subcloned into pTrcHis2 A digested with the same enzymes. The two major reasons we used the pTrcHis2 A (Invitrogen) vector were that it is IPTG inducible and the HIS tag at carboxyl terminus of the gene expression cassette allows easier detection and purification (Kehlenbach et al., 1999; Li et al., 2002). The resultant constructs were named as LCC1 antibody construct and LCC4 antibody construct respectively and transformed into *E coli* Top 10 cells.

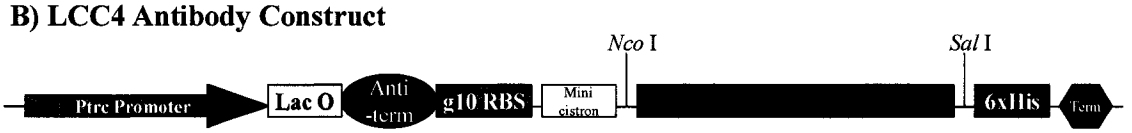
Figure 3. 1 Schematic Maps of LCC1 (A.) and LCC4 (B.) Constructs Used to Produce Proteins for Antibody Production.

In these constructs, the components of constructs and their uses are designated as follows: Ptrc: the trc promoter is a strong, hybrid promoter combining elements of the ~35 region, ~10 region, lac operator, rrnB antitermination signal, gene 10 region, ribosome binding site, and minicistron ORF; LCC1, or LCC4 are mature part of synthetic transgenes, followed with polyhistidine tag (6aa). *Ptrc* promoter is composed of -35 (trpB) and -10 (lacUV5) hybrid promoters for high-level expression of fusion protein (Santos et al., 2001); *Lac* O: lac operator permits binding of the Lac repressor to repress transcription; Anti-term: *rrn*B anti-termination signal reduces the level of premature transcription termination (Li et al., 1984); g10 RBS: bacteriophage gene 10 translational enhancer that optimizes translation initiation of minicistron (Olins et al., 1988); minicistron: minicistron and reinitiation ribosome binding site that contains a second ribosome site for efficient reinitiation of translation into the laccase gene; 6xHis: 6 His residues used for purification of recombinant fusion proteins on metal-chelating resins. The restriction enzyme sites used for cloning are indicated in the maps. For more information see http://tools.invitrogen.com/content/sfs/vectors/ptrchis2_map.pdf.

A) LCC1 Antibody Construct



B) LCC4 Antibody Construct



3.1.2 Constructs for GFP Targeting

The aim of building GFP constructs was to test the functionality of the putative signaling peptides of SBP as suggested by Welinder and Larsen (2004) by investigating the ability of NTPP and CTPP to target GFP within *Arabidopsis* leaf mesophyll cells. In SBP (genomic accession #: AF014502; cDNA accession #: L78163; protein accession #: AAL40127), the length of putative NTPP is 26 aa and the CTPP is 20 aa.

pCAMBIA1302 (accession #: AF234298) was used as the control vector, and it was then further modified for the testing of targeting signals (Figure 3.2). In this vector GFP with a HIS tag was expressed from the CaMV 35S promoter/enhancer.

The cloning of the CTPP was complicated by the presence of two *Nhe* I sites in the vector used for *Agrobacterium* transformation. Thus a cloning strategy was developed in which an 850 bp *Sph* I fragment containing LacZ alpha, MCS, 35S promoter, mGFP, HIS tag, and NOS poly A signal was subcloned into the *Sph* I site of pGEMT-Easy to create the intermediate plasmid pGEM-T-GFP. To generate the CTPP fusion a 60 bp fragment encoding the C-terminal region of SBP was generated by PCR of the plasmid pPRX-3 (Ep peroxidase cDNA cloned into pCMV-BK; gift of Dr. M. Gijzen, London, AAFC) using the primers PRX-gfp3-For and PRX-gfp3-Rev (Table 2.1). The PCR products were purified, digested with the appropriate restriction enzyme pairs and cloned into pGEM-T-GFP digested with the same enzymes. Once the desired fusions had been made in pGEM-T-GFP they were cut with *Sph* I and cloned back into pCAMBIA1302, thus making the GFP3'prx.

To generate the NTPP, a 78 bp DNA fragment encoding the N-terminal region of SBP was generated by PCR of the plasmid pPRX-3 using the primers PRX-gfp5-For and

PRX-gfp5-Rev (Table 2.1). We directly subcloned the *Nco* I – *Spe* I DNA fragment into pCAMBIA 1302 vector cut with the same restriction enzymes to generate 5'prxGFP.

To generate GFP with both signal peptides, we used *Xba* I and *Spe* I to cut 5'prxGFP and cloned the released fragment into GFP3'prx linearized with the same two enzymes, therefore making 5'prxGFP3'prx construct.

3.1.3 Constructs for *Arabidopsis* Transformation

The LCC1, and LCC4 synthetic transgenes were cleaved from pUC57 (Genescript) using *Nco* I and *Bst*E II and then directly ligate into pCAMBIA 1302, resulting in two constructs for *Arabidopsis* transformations, named by pCAMBIA 1302 LCC1, and pCAMBIA 1302 LCC4 (Figure 3.3).

3.1.4 Constructs for Soybean Transformation

The three synthesized transgenes (LCC1, LCC4, and HRPC; Appendix I), flanked by an *Nco* I site at the 5' end and an *Bst*E II site at the 3' end, were cloned in the *Sma* I site of pUC57 (Genscript). The general pattern of components for the transgenes was as follows: NTPP-GGAA-soy codon optimized mature transgene-GGAA-HIS tag-CTPP. The GGAA adaptors were used to facilitate the protein folding properly while the HIS tag was present to aid purification. Digestion with *Nco* I and *Bst*E II released a fragment that could be directly subcloned into the vectors S8 or S18 (Figure 3.4) and used for soybean transformation. S8 and S18 are versions of pCAMBIA1300 modified by Dr. L. Gundynaite-Savitch in which the tCUP promoter replaces the 35S promoter to drive hpt II (hygromycin) expression used for the selection of transgenic plants. We selected the region approximately 1.5 kb upstream from the *Ep* gene as the promoter to drive

expression of foreign genes. However, since the *Ep* promoter has not yet been fully characterized, it is possible that elements necessary for the high expression levels and seed coat specificity of the *Ep* gene were not included in the *Ep* 1.5 promoter.

The major reason for the use of these vectors was to minimize or even eliminate promoter-promoter interactions between *tCUP* and *Ep* promoters that could potentially negate the tissue-specificity of the *Ep* promoter. In the S8 vector this was achieved by cloning two promoters in the head to tail configuration. In the S18 vectors this was achieved by cloning 2.7 kb of yeast LacZ DNA between the two promoters. The resulting constructs were named S8 LCC1, S8 LCC4, and S8 HRPC when cloned into the vector S8 and S18 LCC1, S18 LCC4, and S18 HRPC when cloned into S18. All 6 soybean constructs were shipped to Dr. Dan Brown, AAFC, London, Ontario for transformation. At the time of writing only some S18 transgenic plants were available for analysis.

Figure 3. 2 Schematic Maps of mGFP Fusions with Signal Peptides of SBP

A) GFP control vector (pCAMBIA 1302) without any signals from SBP;

B) GFP3'prx: fusion protein of mGFP with the CTPP of SBP;

C) 5'prxGFP: fusion protein of mGFP with the NTPP of SBP;

D) 5'prxGFP3'prx: fusion protein of mGFP with the SBP NTPP at 5' end and the SBP CTPP at 3'end.

In these constructs, the components of constructs and their uses are designated as follows: 35 S promoter: promoter from CaMV, used to drive the expression of GFP and GFP fusions; NTPP: N-terminal transit peptide of SBP; CTPP: C-terminal transit peptide of SBP; NOS: nopaline synthase 3'UTR and polyA signal.

The restriction enzyme sites used for cloning are indicated in the maps.

A) GFP

B) GFP3'prx

C) 5'prxGFP

D) 5'prxGFP3'prx

Figure 3. 3 Schematic Maps of Laccase Constructs for *Arabidopsis* Transformation

A) pCAMBIA 1302 LCC1;

B) pCAMBIA 1302 LCC4;

In these constructs, the components of constructs and their uses are designated as follows: 35 S promoter: promoter from CaMV, used to drive the expression of laccase synthetic genes; NOS: nopaline synthase 3'UTR and polyA signal.

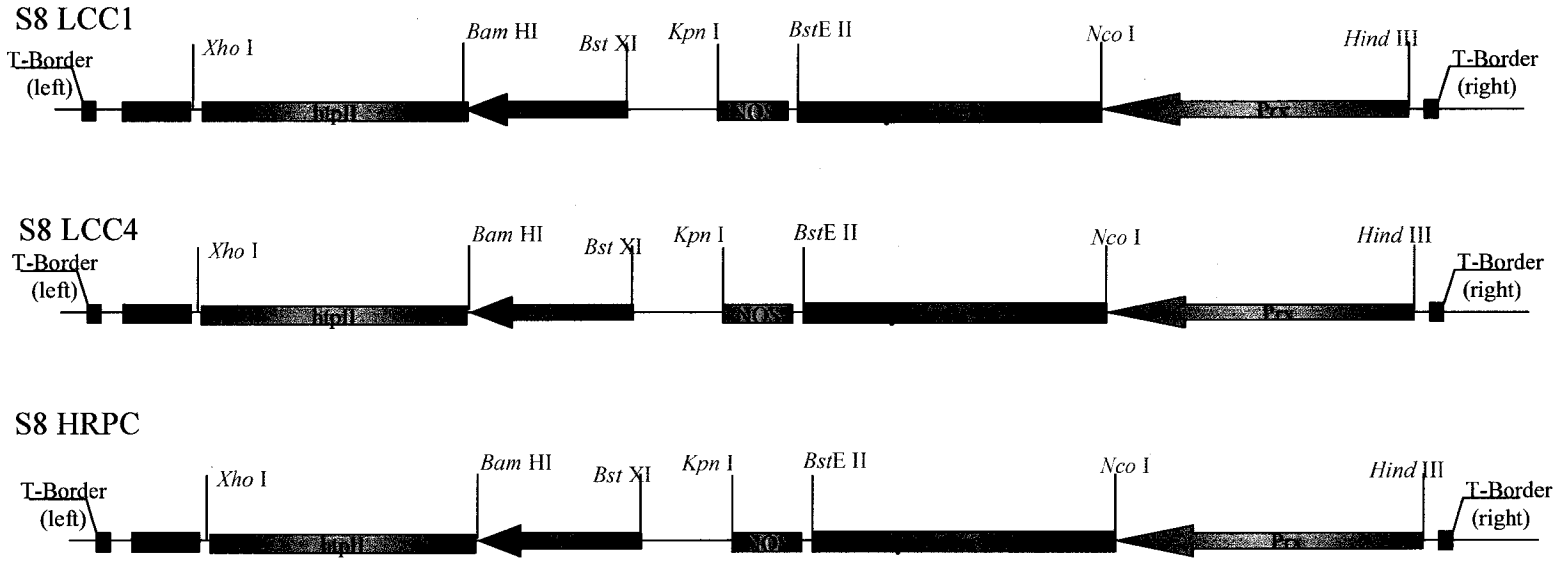
A) pCAMBIA 1302 LCC1

B) pCAMBIA 1302 LCC4

Figure 3. 4 Schematic Maps of Soybean Vectors for Transformation

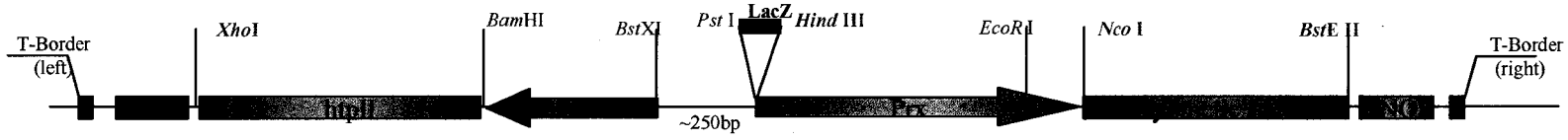
Soybean codon optimized genes were cloned into S8 and S18 vectors for transformation. In these constructs, the components and their uses are designated as follows. *35 S* promoter is *35S* promoter from CaMV; *hpt* II is hygromycin resistance gene; *tCUP* is a constitutive promoter derived from tobacco, and it is used for driving hygomycin gene expression; *NOS* is nopaline synthase terminator; *HRP* or *LCC*1 or *LCC*4: synthetic genes for horseradish peroxidase or laccase I or laccase IV respectively optimized for soybean codon usage, with the addition of HIS tags and N- and C-terminal transit peptides from SBP; Prx: 1.5 kb DNA fragment with the SBP promoter; *Lac*Z is 2.7 kb of yeast LacZ DNA, a reporter gene that codes for β-galactosidase. The restriction enzyme sites used for cloning are indicated in the maps.

S8 Soybean Transformation Constructs

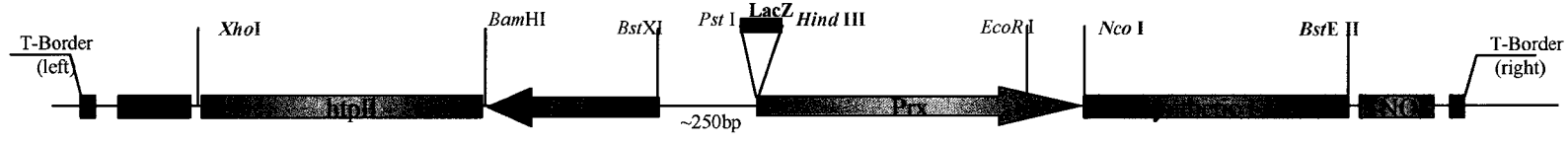


S18 Soybean Transformation Constructs

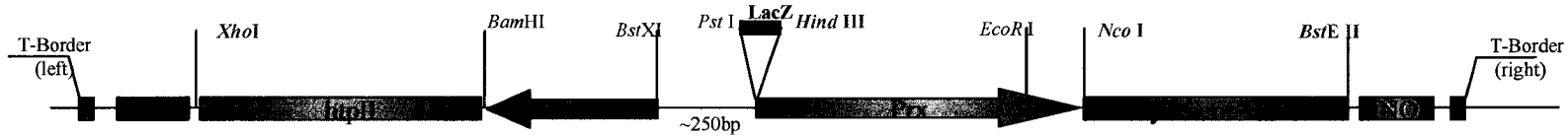
S18 LCC1



S18 LCC4



S18 HRPC



3.2 Molecular Tools for Characterizing Transgenic Plants

3.2.1 PCR Tools

PCR was used to screen transgenic plants for the presence of key DNA elements including *Ep* promoter, the different transgenes, *NOS*, *tCUP*, and *hpt* II regions. As a control for genomic DNA (gDNA) quality we used *GmTub2* gene. Standard PCR conditions are described in Materials and Methods, section 2.4, and the primers used for transgenic soybean screening are shown in Table 2.1. The optimized PCR working conditions are tabulated in Table 3.1.

Two DNA extraction kits were tested for the ability of the gDNA to be amplified by PCR-the Sigma REDExtract-N-Amp^{TM} Plant PCR kit and the QIAGEN DNeasy Plant Mini Kit. Their effectiveness was compared by isolating gDNA from a leaf from a control (cv. Jack) and transgenic (S8 LCC1 4-2-1) soybean using Sigma and QIAGEN kits and then testing for PCR amplification of regions containing *GmTub2* (tubulin positive control for gDNA) *NOS, tCUP* and *hpt* II.

As expected only the transgenic plant had the *NOS*, *tCUP*, and *hpt* II PCR products (Figure 3.5 lanes 3, 4) indicating that the primers are specific under these conditions. The results showed that DNA prepared with either the Sigma or QIAGEN kits was equally good for PCR screening. The Sigma kit was thus used throughout.

Testing for the presence of the *Ep* promoter is complicated by the presence of the *ep* allele in cv. Jack that has the same putative promoter and indeed most of the *Ep* gene (Gijzen, 1997; Gijzen, unpublished data). In order to detect these parts of the transgenes, we adopted a strategy of designing a forward primer complementary to the 5' end of the *Ep* 1.5 promoter (Ep-1.5 F32) and we designed a reverse primer within the transgene

~200 bp downstream of *Nco* I site for each transgene and within the DNA sequence encoding the mature protein. The PCR (Figure 3.6) results showed the primer pair of Ep-1.5 F32/transgene reverse primer did not detect the *Ep* promoter in WT (Jack) plant gDNA (Figure 3.6 lane 2, 4, 6), but did detect bands of Ep-LCC1, Ep-LCC4, and Ep-HRPC (Figure 3.6 lane 3, 5, 7) with the expected sizes of 1748, 1764, 1793 bp respectively. These results indicate the presence of *Ep* promoter connected to the transgene can be detected in the transgenic plant.

Table 3. 1 Optimization of PCR Conditions for the Different Transgene Components

Target Sequence	Primers	Product Size (bp)	Optimized Td (°C)	Extension Time (MM:SS)	Cycles
Tub 2	SoyTub2For / SoyTub2Rev	430	60	00:30	30
tCUP	tCUP (20) For/ (432) Rev	413	60	00:30	30
Hpt II	Hpt II (99) For/(614) Rev	516	60	00:30	30
NOS	NOS (13) For/(192) Rev	180	60	00:30	30
Ep0.3kb	Ep-0.3 F/ Ep-Ori-Rev	332	55	00:30	30
Ep0.6kb	Ep-0.6 F/ Ep-Ori-Rev	639	55	01:00	35
Ep0.9kb	Ep-0.9 F/ Ep-Ori-Rev	944	55	01:00	35
Ep1.2kb	Ep-1.2 F/ Ep-Ori-Rev	1243	55	01:00	35
Ep1.5kb	Ep-1.5 F/ Ep-Ori-Rev	1548	55	01:30	35
Ep1.5kb– LCC1	Ep-1.5 F32 / LCC1 (198) Rev	1748	60	02:00	35
Ep1.5kb– LCC4	Ep-1.5 F32 / LCC4 (214) Rev	1764	60	02:00	35
Ep1.5kb– HRPC	Ep-15 F32 / HRPC (243) Rev	1793	60	02:00	35
LCC1	LCC1-Nco I-OP / LCC1-Sal I - OP	1512	60	01:30	35
LCC4	LCC4-Nco I-For / LCC4-Sal I-Rev	1515	60	01:30	35
HRPC	HRPC-Not1-For / HRPC-Xba1-Rev	1116	60	01:30	35

For each primer pair the Td (annealing temperature) was estimated by gradient PCR. The table represents the combined efforts Dr. L. Gudynaite-Savitch, Shuyou Han and Jaimie Schnell.

Figure 3. 5 Comparing the Effectiveness of Sigma and QIAGEN gDNA as Templates for Transgenic Soybean Screening by PCR

gDNA was prepared from WT (Jack), and S8 LCC1 4-2-1 plants using either the Sigma or QIAGEN kits and was screened by PCR for GmTub2, NOS, tCUP, hptII, and LCC1 sequences. Standard PCR conditions were used with an extension time summarized in Table 3.1. The expected sizes are indicated on the right.

Lane 1: gDNA isolated from cv. Jack (WT) using the Sigma kit.

Lane 2: gDNA isolated from cv. Jack (WT) using the Qiagen kit.

Lane 3: gDNA isolated from S8 LCC1 4-2-1 using the Sigma kit.

Lane 4: gDNA isolated from S8 LCC1 4-2-1 using the Qiagen kit.

Lane 5: S8 LCC1 plasmid DNA (positive control)

The primers used for amplification (Table 1) were NOS (13) For + NOS (192) Rev for NOS amplification; SoyTub2For + SoyTub2Rev for GmTub2 amplification; tCUP (20) For + tCUP (432) Rev for tCUP amplification; HptII (99) For + HptII (614) Rev for hptII amplification.

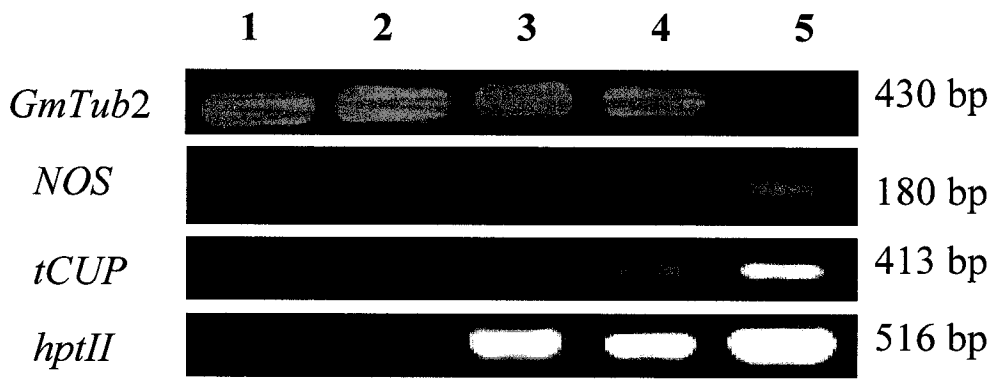


Figure 3. 6 Specific PCR Identification of the *Ep* 1.5 Promoter Fused with Transgenes

Genomic DNA was isolated using the Sigma kit. Aliquots were PCR amplified under using standard conditions with a Td of 60°C and an extension time of 2 min. The expected sizes are indicated.

Lane 1: GeneRuler^{TM} 100 bp (# SM0321) (Fermentas). Sizes are in base pairs.

Lane 2: gDNA from cv. Jack (WT), amplified with primers Ep-1.5F32 + LCC1 (198) Rev

Lane 3: S8 LCC1 plasmid DNA, amplified with primers Ep-1.5F32 + LCC1 (198) Rev

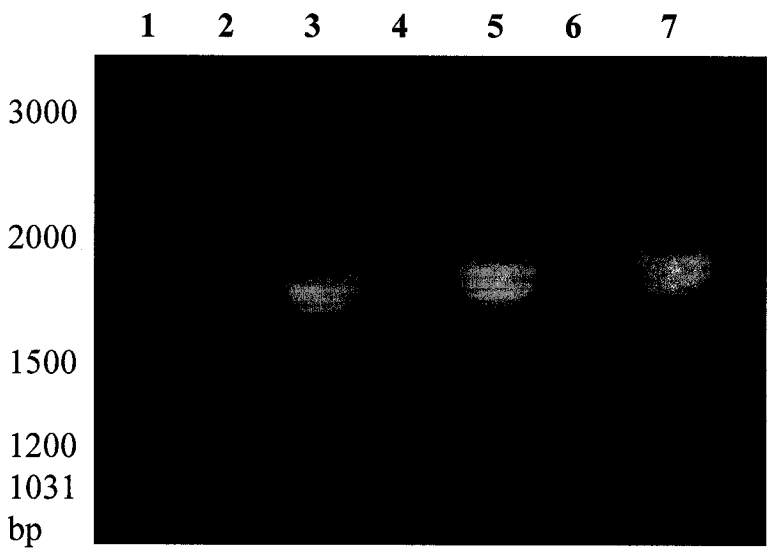
Lane 4: gDNA from cv. Jack (WT), amplified with primers Ep-1.5 F32 + LCC4 (214) Rev

Lane 5: S8 LCC4 plasmid DNA, amplified with primers Ep-1.5 F32 + LCC4 (214) Rev

Lane 6: gDNA from cv. Jack (WT), amplified with primers Ep-1.5 F32 + HRPC (243) Rev.

Lane 7: S8 HRPC plasmid DNA amplified with primers Ep-1.5 F32 + HRPC (243) Rev.

The sequences of primers were listed in Table 1.



3.2.2. Production and Purification of LCC1 and LCC4 Proteins in *E. coli* for Antiserum Production

To produce specific antisera against laccases (anti-HRPC is commercially available), we expressed laccases that had been cloned into the vector pTrcHis2 A as described in Section 3.1.1. The vector pTrcHis2 A contains an IPTG-inducible promoter that drives expression of the cloned gene. Each construct was grown and induced with IPTG as described in Materials and Methods, section 2.8.1. Analysis by SDS-PAGE revealed no major differences when 1-4 mM IPTG was used (data not shown). When a culture containing pTrcHis2 A LCC1 was induced a band of 54.9 kDa (determined by SDS gel) was detected (Figure 3.7A) in the soluble fraction based upon the fractionation procedure as defined by the BugBuster kit. This size is close to the calculated size of 54.6 kDa for this fusion protein. Purification was attempted based upon the presence of the HIS-tag and several proteins were detected by SDS-PAGE gel (Figure 3.7B). Surprisingly, several bands were observed with sizes of 82.1, 75.7, 23.3 kDa on SDS gel, and no strong band was observed at the expected size of 54.6 kDa. This experiment was repeated 3 times with similar results. While all the proteins in lane 6 are expected to have a HIS-tag their origins are unknown. They are unlikely to be from the bacterial host; using the tools available at NCBI (BLASTP; accessed April 7, 2009) we could not find any matches to *E coli* proteins. The band at 23.3 kDa may result from degradation and the higher molecular weight proteins of 82.1 kDa and 75.7 kDa may represent soluble proteins that are unexpectedly resistant to SDS and are not totally denatured on SDS gel (Dr. P. Telmer, AAFC, oral communications). As will be demonstrated later (Figure 3.9C), these bands are recognized by a cross-reacting antiserum prepared against LCC4

and therefore likely represent LCC1. The upper two bands were isolated from a preparative SDS gel and sent for antibody production.

LCC4 was expressed in *E coli* employing the same approach as for LCC1. A pTrcHis2 A LCC4 culture was induced with IPTG and fractionated into a soluble fraction and the inclusion body. In this case LCC4 was expressed in inclusion bodies based upon the fractionation procedure as defined by the BugBuster kit. In Figure 3.8A, LCC4 with an observed size of 64.3 kDa (Figure 3.8A) determined by SDS gel was similar to the predicted size of 54.4 kDa. LCC4 protein was purified with a QIAGEN HIS tag column. Several bands were identified with calculated sizes of 62.0, 43.7, 28.0, and 22.5 kDa (Figure 3.8 B). It is interesting that these sizes are similar to those sizes (55.4, 48.9, 31.2, 19.8 kDa) calculated from the protein sequence of the fusion protein assuming that internal initiation at methionine is occurring (Santagata et al., 2000). The differences between protein size determined by aa sequence and protein size calculated based on the protein ladder standard curve might be due to migration distance measurement and equation curve fit (Plikaytis et al., 1986).

The bands with sizes of 62.0 kDa were cut from several gels and sent for antibody production.

Figure 3.7 A Time Course of LCC1 Protein Induction in *E. coli*

An exponential culture of pTrcHis2 A LCC1 *E. coli* growing at 37°C with shaking was induced with the addition of 4mM IPTG. 15 μg of the soluble fraction were loaded onto an SDS-PAGE gel. The arrow indicates a protein of the size expected for the LCC1-His fusion. Lane 1: Protein Ladder (BIO-RAD #:161-0373) (5 μl); Lane 2: Before IPTG induction; Lane 3: IPTG induction for 1h; Lane 4: IPTG induction for 2h;Lane 5: IPTG induction for 3h; Lane 6: IPTG induction for 4h.

Figure 3.7 B Purification of HIS Tagged LCC1 Protein

LCC1 protein was purified with QIAGEN QIAexpress® Ni-NTA Fast Start kit according to the manufacturer's protocol, and fractions were run on an SDS gel. Protein sizes were determined using a standard curve based upon the protein ladder. Lane 1: Protein Ladder (BIO-RAD #:161-0373) (5 μl); Lane 2: 15μg of *E coli* cell lysate;
Lane 3: 15μg of flow through;Lane 4: 15μg of first wash; Lane 5: 0.5μg of second wash; Lane 6: 1μg of first elution;

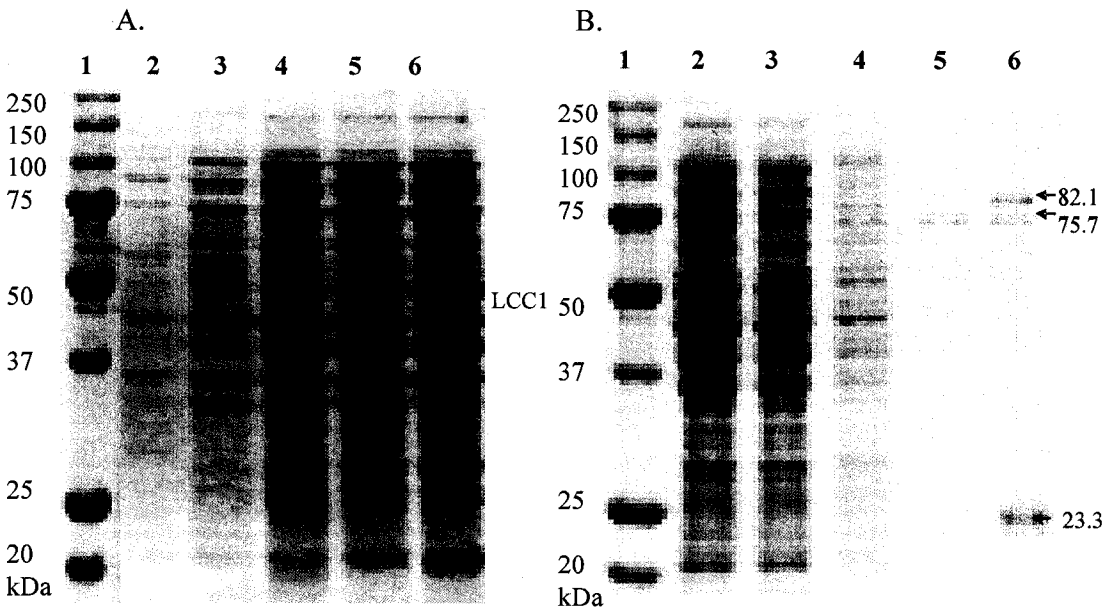
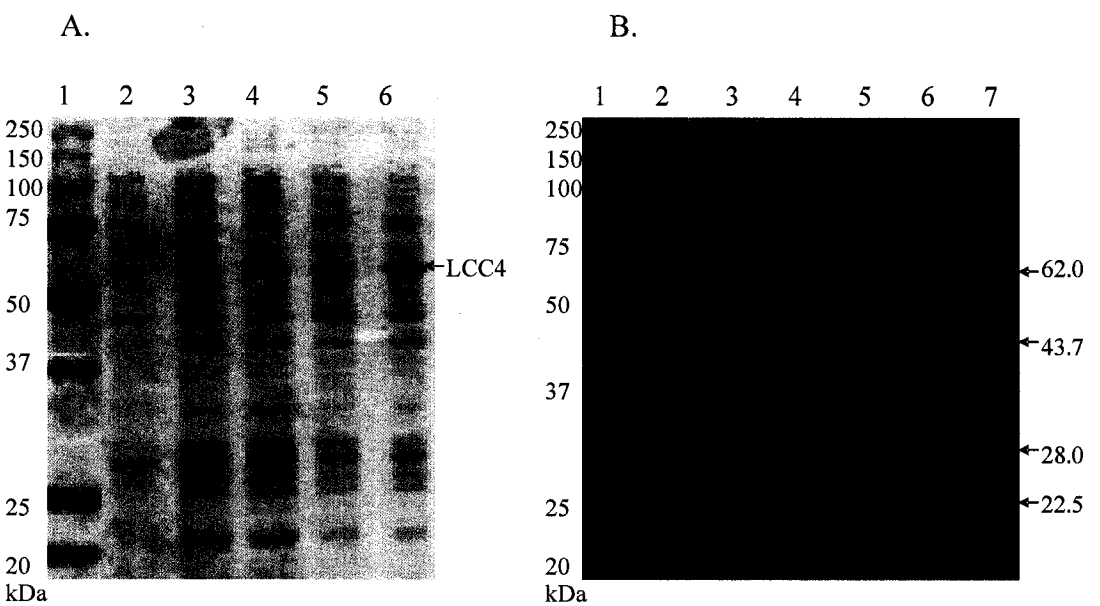


Figure 3.8 A Time Course of LCC4 Protein Induction in *E. coli*

An exponential culture of pTrcHis2 A LCC4 *E. coli* growing at 37°C with shaking was induced with the addition of 1 mM IPTG. 15 μg of the insoluble fraction were loaded onto an SDS-PAGE gel. The arrow indicates a protein of the size expected for the LCC4-His fusion. Lane 1: Protein Ladder (BIO-RAD #:161-0373) (5 μl); Lane 2: Before IPTG induction; Lane 3: IPTG induction for 1h; Lane 4: IPTG induction for 2h; Lane 5: IPTG induction for 3h; Lane 6: IPTG induction for 4h.

Figure 3.8 B Purification of HIS Tagged LCC4 Protein Expressed in *E coli* Inclusion Bodies

LCC4 protein was purified with QIAGEN QIAexpress® Ni-NTA Fast Start kit according to the manufacturer's protocol, and fractions were run on an SDS gel. Protein sizes were determined using a standard curve based upon the protein ladder. Lane 1: Protein Ladder (BIO-RAD #:161-0373) (5 μl); Lane 2, *E coli* inclusion bodies (15 μg); Lane 3, Flow through (15 μg); Lane 4, Wash one (10 μg); Lane 5, Wash two (0.5 μg); Lane 6, Elution one (1 μg); Lane 7, Elution two (1 μg);



3.2.3 Tests of Antibody Functionality and Specificity

3.2.3.1 Functionality and Specificity of Laccase Antibodies

The sensitivity of the LCC1, and LCC4 antibodies were tested using commercial laccase from from *Trametes versicolor* (Biochemika, #38429), and at the same time, soybean leaf, seed coat, and embryo proteins from cv. Jack were used to test the specificity of LCC1, and LCC4 antibodies.

The anti-LCC1 antibody was used to probe the Biochemika laccase protein dilution series on a Western blot using AP conjugate colorimetric detection system (Figure 3.9 A); The anti-LCC1 antibody detects double bands with protein sizes around at 65 kDa, consistent with previous results (Hood et al., 2003) which showed two LCC1 protein bands were expressed in *Trametes versicolor* at 64 and 66 kDa. Anti LCC1 antibody detected no protein signals in extracts from soybean (cv. Jack) leaf, seed coat, and embryo (Figure 3.9 B). However, the anti-LCC1 antibody could detect laccase on the same blot. In terms of sensitivity, the anti-LCC1 antibody was not a very good antibody since the intensities of protein bands on the Western blots did not change as a function of the amount of Biochemika laccase loaded. Also signal detection required a long exposure time, e.g., up to 14h (Figure. 3.9B).

Since the anti-LCC1 antibody was incapable of efficiently detecting Biochemika laccase we tried using the Anti-LCC4 antibody to detect LCC1 proteins expressed in pTrcHis2 A LCC1 *E coli.* The results indicated anti LCC4 could detect LCC1 in a dilution series (Figure 3.9 C) as well as a band of 35 kDa within 15 min. The sizes of the bands in the doublet (Figure 3.9D) are similar to the two upper bands detected with anti-LCC1 antibody (Figure.3.7B). The origin of this later band is not known but it may be

non-specific. In conclusion, the anti-LCC1 antibody could act as an alternative to detect LCC1 proteins (Figure 3.9 C). The potential reasons for this might be that 1) LCC1 and LCC4 protein sequences share 70% identity, some part of N-terminal regions of these two proteins share up to almost 100% similarity; 2) the conserved motifs, such as copper binding regions, are immunogenic.

The LCC1 proteins expressed in soybean have a HIS tag; therefore, the functionality of anti-HIS tag antibody was tested on the HIS tagged LCC1 proteins expressed in pTrcHis2 A LCC1 *E. coli.* The results showed that anti-HIS tag antibody could efficiently detect HIS tagged LCC1 proteins at 50 ng (Figure 3.9 D), and this antibody may provide an additional way to detect this protein. Surprisingly, the truncated LCC1 protein bands were not picked up, and only full-length protein band showed up, the reason might be full-length LCC1 proteins were relatively highly expressed, and protein expression levels of other two truncated LCC1 protein bands were not detectable by anti-HIS tag antibody.

Figure 3.9 A Test of Anti LCC1 Antibody Sensitivity

The dilution series of Biochemika laccase were made at range from 500 ng to 1 ng, and Western blot was done using AP colorimetric detection. Anti-LCC1 antibody was diluted at 500X, and secondary antibody Goat anti-rabbit IgG with AP was diluted at 3000X, the overlay time was 5 h.

Figure 3.9 B A Test of Anti LCC1 Antibody Specificity

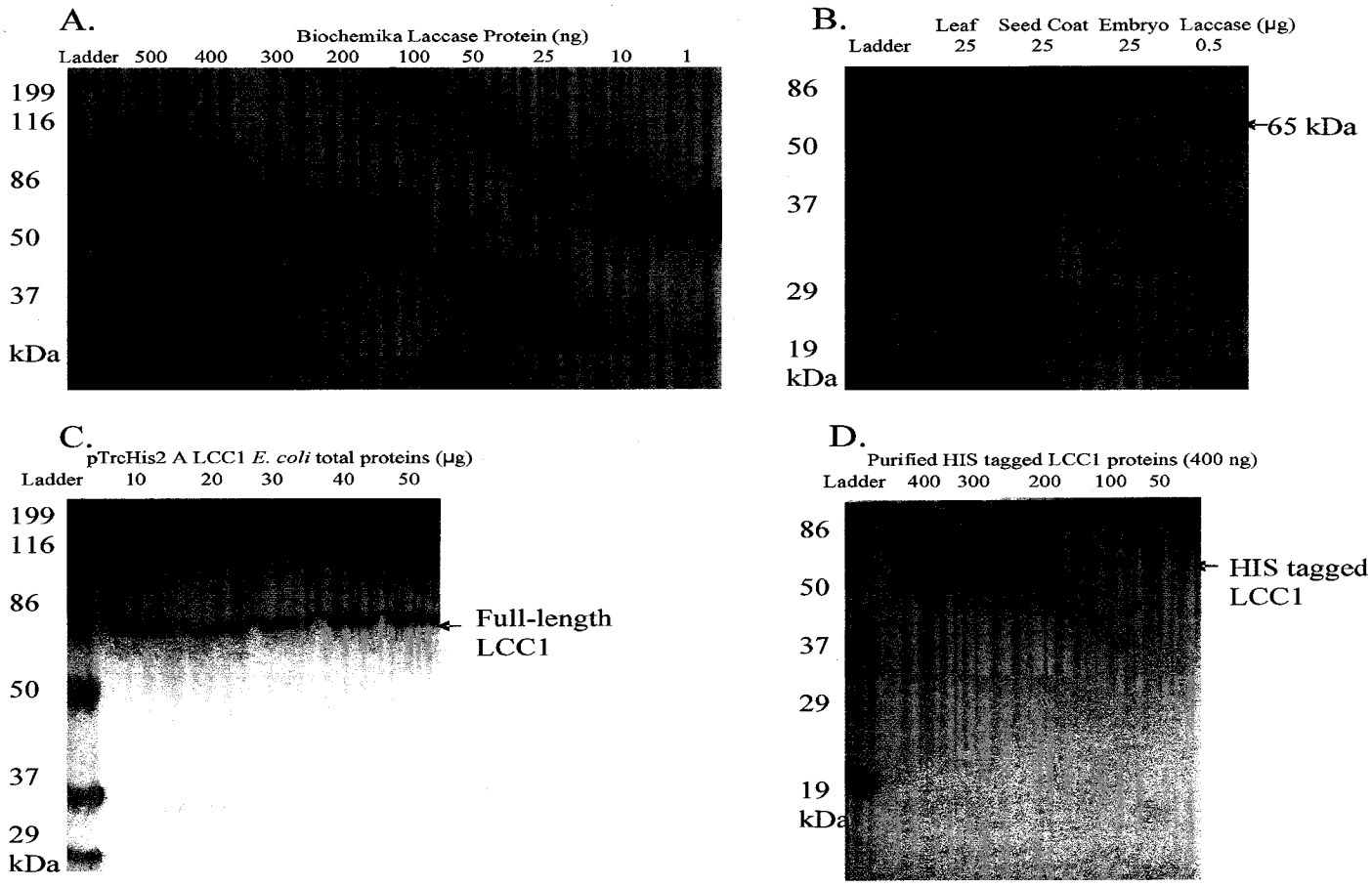
Soybean (cv. Jack) leaf, seed coat, and embryo proteins were used to test the specificity of anti-LCC1 antibody. The dilution factor of anti-LCC1 antibody was 250X, and that of goat anti-rabbit IgG AP conjugate was 3,000X. The overlay time was 14 h. Anti-LCC1 antibody can detect Biochemika laccase protein in lane 5.

Figure 3.9 C the Anti-LCC4 Antibody Could Detect LCC1

The series of total LCC1 proteins were made at range from 10 to 50 μg and Western blot was done using AP colorimetric detection. The dilution factor of anti-LCC4 antibody was 1000X, and that of goat anti-mouse IgG AP conjugate (Sigma, A-1293) was 3,000X. The overlay time was 15 min.

Figure 3.9 D Anti-HIS Tag Antibody Could Detect HIS Tagged LCC1 Proteins

Anti-HIS antibody (QIAGEN) can detected HIS tagged LCC1 proteins expressed in *E coli.* The series of purified HIS tagged LCC1 were made at range from 400 ng to 50 ng, and Western blot was done using AP colorimetric detection system. The dilution factor of the anti-HIS antibody was 1000X, and that of goat anti-mouse IgG AP conjugate (Sigma, A-1293) was 3,000X. The overlay time was 15 min.


A.
Biochemika Laccase Protein (ng)
Ladder 500 400 300 200 100 50 25 10 1
199
116
86
50
37
kDa
B.
Leaf Seed Coat Embryo Laccase (µg)
Ladder 25 25 25 0.5
86
50
37
29
19
kDa
65 kDa
C.
pTrcHis2 A LCC1 E. coli total proteins (µg)
Ladder 10 20 30 40 50
199
116
86
50
37
29
kDa
Full-length
LCC1
D.
Purified HIS tagged LCC1 proteins (400 ng)
Ladder 400 300 200 100 50
86
50
37
29
19
kDa
HIS tagged
LCC1

The functionality and specificity of the anti-LCC1 antibody were tested using the same approach as that of the anti-LCC1 antibody. The results showed that anti-LCC4 antibody was fairly robust and could detect nanogram levels of protein (Figure 3.10 A). The detected protein band intensities varied against the amount of Biochemika laccase loaded both in colorimetric (Figure 3.10A) and chemiluminescent detection system (Figure 3.10B). When the relationship between the band intensity in Figure 3.10B and laccase concentration was plotted in a graph (Figure 3.10 C) the best fit was to a logarithmic curve. The anti-LCC4 antibody specificity data (Figure 3.10 D) showed that the anti-LCC1 antibody detected one low molecular weight band in soybean leaf protein of around 35 kDa, two protein bands in seed coat with sizes of 75 kDa and 50 kDa, and one protein band in embryo with a size of about 75 kDa. This 50 kDa band might be a putative soybean laccase (AY113187.2, AAM54731.1, 589 aa). No data for expression of this gene in seed coats is available, only a single EST for root and hypocotyl (http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Gma&CID=8441). The presence of these bands may make the detection of the expressed laccase difficult since the 50 kDa is similar to the size as the Biochemika laccase (lane 5) and the predicted size (by aa sequence) of our LCC1 construct at 59.71 KDa with both SBP signalling peptides present and 53.59 kDa without them; for the LCC4 construct the values are 59.47 kDa with SBP signalling peptides and 53.32 kDa without them. The sizes measured on SDS gels may be affected by the extent of glycosylation (Gillikin and Graham, 1991).

Figure 3.10 A Functionality Test of Anti-LCC4 Antibody (Colorimetric)

The dilution series of Biochemika laccase proteins were made at range from 500 ng to 1 ng, and Western blot was done using AP conjugate colorimetric detection. Anti-LCC4 antibody was diluted at 1,000X, and secondary antibody Goat anti-rabbit IgG with AP was diluted at 3,000X, the overlay time was 14 h.

Figure 3.10 B Functionality Test of Anti-LCC4 Antibody (Chemiluminescent)

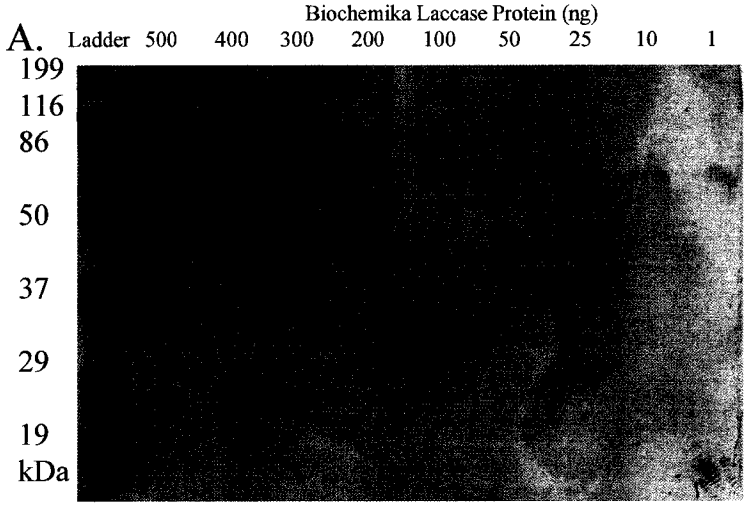
The dilution series of Biochemika laccase proteins were made at range from 500 ng to 1 ng, and Western blot was done using HRPC conjugate chemiluminescent detection. The anti-LCC4 antibody was used at 1000X, and secondary antibody goat anti-rabbit IgG with HRP (Dr. Simmonds') was diluted at 50,000X, the exposure time was 1min.

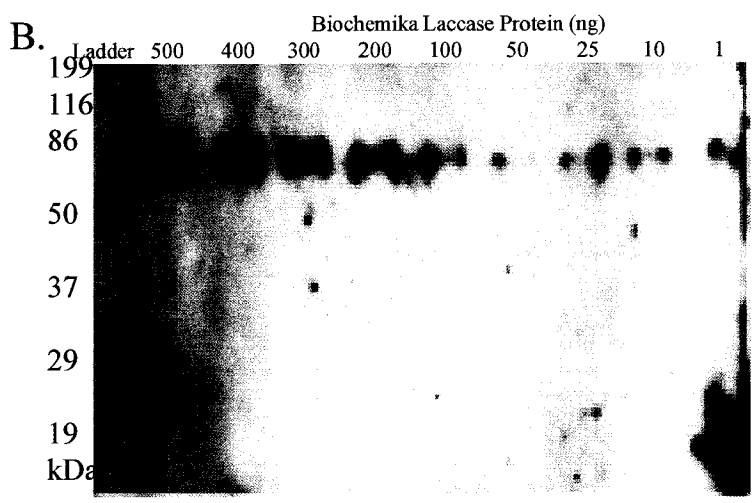
Figure 3.10 C Standard Curve for Laccase Proteins

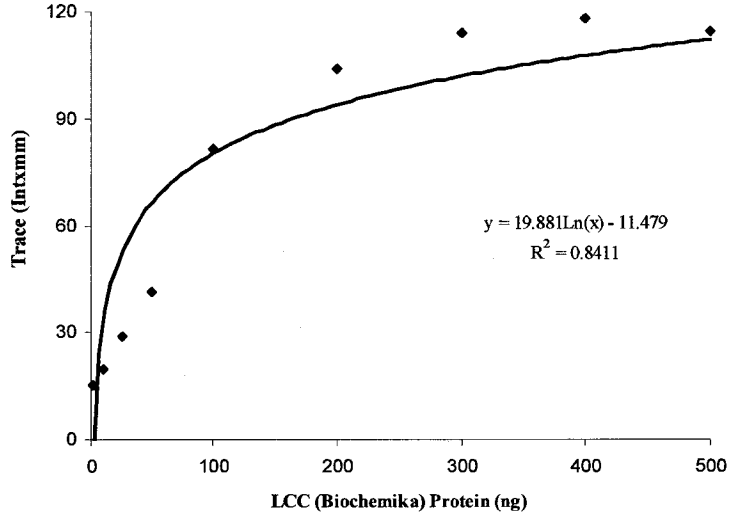
Based on the scanned Western picture (Figure 3.10 B), the band intensities were quantified using the software Quantity One (Bio-Rad), the relationship of protein band intensity and amount of protein were graphed, fitting to a logarithmic curve.

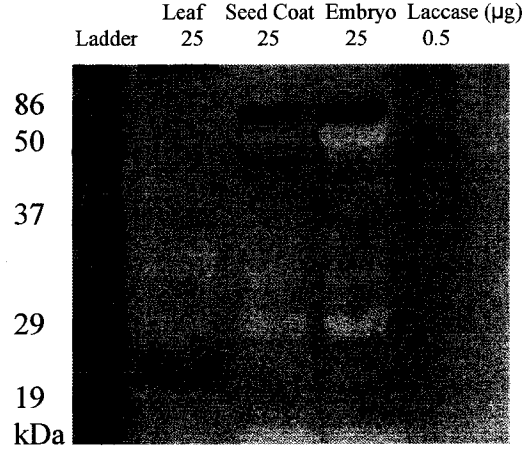
Figure 3.10 D The Specificity Test of Anti-LCC4 Antibody

Soybean (cv. Jack) leaf, seed coat, and embryo proteins were used to test the specificity of anti-LCC4 antibody. The dilution factor of anti-LCC4 antibody was 1000 X, and that of goat anti-rabbit IgG AP conjugate was 3,000X. The overlay time was 14 h. Anti-LCC4 antibody can detect Biochemika laccase protein in lane 5.


A.
Biochemika Laccase Protein (ng)
Ladder 500 400 300 200 100 50 25 10 1
199
116
86
50
37
29
19
kDa


B.
Biochemika Laccase Protein (ng)
Ladder 500 400 300 200 100 50 25 10 1
199
116
86
50
37
29
19
kDa


C.
120
90
60
30
0
Trace (Intxmm)
0 100 200 300 400 500
LCC (Biochemika) Protein (ng)
y = 19.881Ln(x) - 11.479
R2 = 0.8411


D.
Ladder
Leaf 25
Seed Coat 25
Embryo 25
Laccase (µg) 0.5
86
50
37
29
19
kDa

3.2.3.2 Functionalities and Specificities of Anti-HRPC and Anti-HIS tag Antibodies

Anti-HRPC antibody from Abcam, and anti-HIS tag antibody from QIAGEN were also tested. The HRPC (Sigma) protein was used to test functionality of the anti-HRPC antibody. At the same time we used cv.Jack soybean leaf, seed coat, and embryo proteins to test the specificities of anti-HRPC, and anti-HIS tag antibodies.

The anti-HPRC antibody (Abcam) functioned very well to detect 10ng of HRPC commercial proteins (Sigma) (Figure 3.11 A, B), the HPRC protein bands on the Western blot changed dramatically against the different amount of HRPC protein loaded. The relationship between band intensity and amount of HPRC proteins was plotted in a graph (Figure 3.11 C) based on the Figure 3.11 B, showing a significant linear tread ($R^2 = 0.97$).

Soybean leaf, seed coat, and embryo proteins from cv. Jack were used to test the specificity of anti-HRPC antibody and many bands were detected in protein extracts from the tested plant tissues (Figure 3.11 D). There are hundreds of peroxidases in plant genomes mainly to help protect against diseases (Bakalovic et al., 2006; Passardi et al., 2007) and the anti-HPRC antibody could be detecting them, e.g., peroxidases have the well conserved Fe^{++} binding motifs that might be immunogenic (Zubieta et al., 2007). Therefore, anti-HRPC antibody is not suitable to detect the expressed HPRC protein signals from soybean seed coat total proteins since the high background will make detection of the expressed enzymes difficult. However, it would still be useful to confirm the presence of the expressed HPRC after purification based upon the presence of the HIS tag.

The cv. Jack soybean leaf, seed coat, and embryo proteins were used to test the specificity of anti-HIS tag antibody (Figure. 3.11E). The results showed anti-HIS tag antibody detected no proteins from leaf and seed coat extracts but did detect 30.0 kDa protein in soybean embryo. The anti-HIS tag antibody is potentially useful for our project although it detects a 30.0 kDa band in soybean seed embryo because 1) the seed coat is our focus of research, and no protein was detected in the non-transgenic controls in either leaf or seed coats; 2) the size of the detected protein, 30kDa, is different from the expected size of HIS tagged proteins which should be at least 35.2 kDa for the HRPC and 53.3 kDa for LCC4 laccases if NTPP and CTPP of SBP are removed (Appendix 1).

Figure 3.11 A Functionality Test of Anti-HRPC Antibody (I)

The dilution series of HRPC were made at range from 500 ng to 1 ng, and Western blot was done using AP conjugate colorimetric detection. Anti-HRPC antibody was diluted at 3,000X, and secondary antibody Goat anti-rabbit IgG with AP was diluted at 3,000X, the overlay time was 3.5 min.

Figure 3.11 B Functionality Test of Anti-HRPC Antibody (II)

The dilution series of HRPC were made at range from 500 ng to 1 ng, and Western blot was done using HRPC conjugate chemiluminescent detection. The anti-HRPC antibody was used at 30,000X, and secondary antibody mouse anti-goat IgG with HRP was diluted at 50,000X, the exposure time was 20 sec.

Figure 3.11 C Graph of HPRC (Sigma) Protein Band Intensities (Int x mm) as the Function of the Amount of Laccase Proteins Loaded

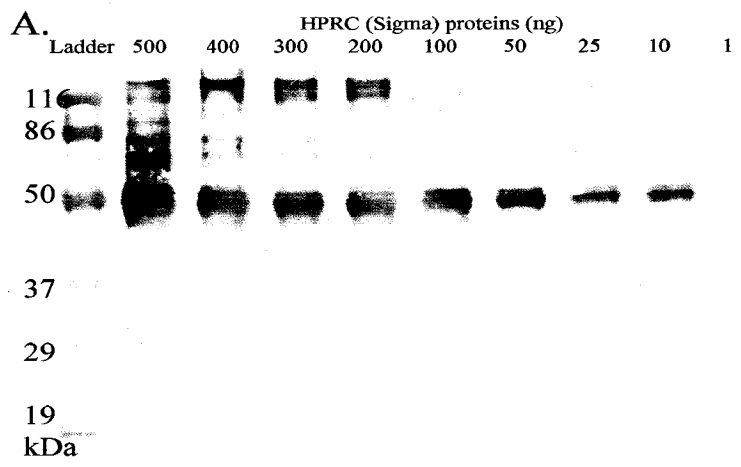
Based on the scanned Western picture (Figure 3.11 B), the band intensities were quantified using the software Quantity One (Bio-Rad), the relationship of protein band intensity and amount of protein were graphed, showing a strong linear trend.

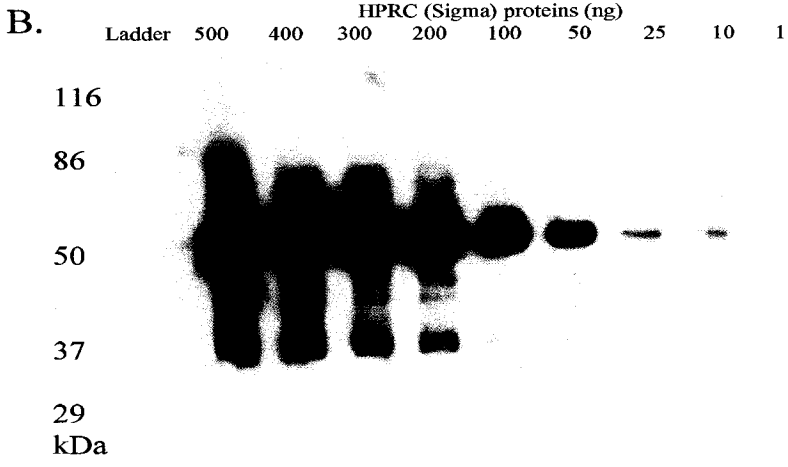
Figure 3.11 D Specificity Test of Anti-HRPC Antibody

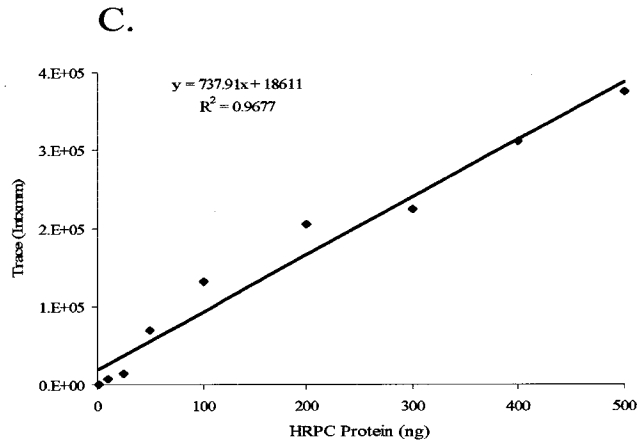
Soybean (cv. Jack) leaf, seed coat, and embryo proteins were used to test the specificity of anti-HRPC antibody. The dilution factor of anti-HRPC antibody was 6,000X, and that of goat anti-rabbit IgG HRPC conjugate was 20,000X. The exposure time was 1 min.

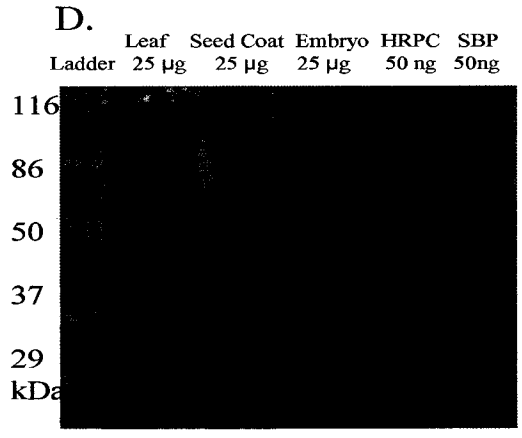
Figure 3.11 E Specificity Test of Anti HIS Tag Antibody

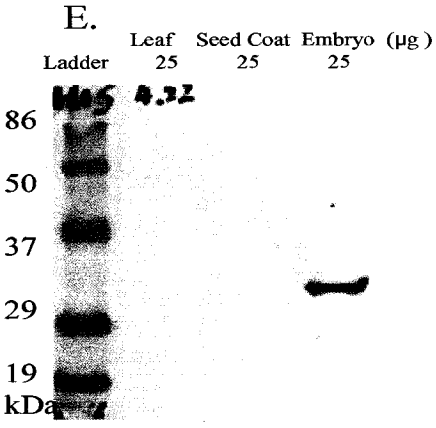
Soybean (cv. Jack) leaf, seed coat, and embryo proteins were used to test the specificity of anti-HIS tag antibody. The dilution factor of HIS antibody was 1000X, and that of goat anti-mouse IgG AP conjugates (Sigma, A-1293) was 3,000X. The overlay time was 15 min.


A.
HPRC (Sigma) proteins (ng)
Ladder 500 400 300 200 100 50 25 10 1
116
86
50
37
29
19
kDa


B.
HPRC (Sigma) proteins (ng)
Ladder 500 400 300 200 100 50 25 10 1
116
86
50
37
29
kDa


C.
y = 737.91x + 18611
R² = 0.9677
Trace (Intensity)
4.E+05
3.E+05
2.E+05
1.E+05
0.E+00
0
100
200
300
400
500
HRPC Protein (ng)


D.
Ladder
Leaf 25 µg
Seed Coat 25 µg
Embryo 25 µg
HRPC 50 ng
SBP 50ng
116
86
50
37
29
kDa


E.
Ladder
Leaf 25
Seed Coat 25
Embryo (µg) 25
86
50
37
29
19
kDa

3.3 Functionality of SBP Signalling Peptides

Several lines of evidence summarized in the introduction suggest that SBP contains both a NTPP and a CTPP. This hypothesis was tested by the analysis of transformed *Arabidopsis thaliana* (Columbia) plants expressing one of the four GFP constructs described in 3.1.2 In total, 20 independent lines were obtained for pCAMBIA 1302; 27 lines for 5'prxGFP, 32 lines for GFP3'prx and 26 lines for 5'rpxGFP3'prx. A large number of transformants were generated to increase our chances for recovery of high expressing lines. Transgene copy number and position effects generate a wide range of transgene expression levels. Preliminary screening by confocal microscopy identified 3 independent GFP-positive lines for each of the pCAMBIA1302 (L3, L4, L9), 5'prxGFP (L8, L10, L15), and GFP3'prx (L6, L18, L20) constructs but only one 5'prxGFP3'prx line (L26). The results were consistent for all lines with the same construct.

Transgenic *Arabidopsis* plants were transported to Dr. F. Brandizzi's lab for confocal microscopy. We found that unmodified GFP (vector control) was localized in the cytoplasm of mesophyll cells (Figure. 3.12 I A). The GFP with the SBP NTPP (5'prxGFP) was targeted to the "exterior" of the cell (Figure 3.12 I B), although it was difficult to distinguish between cell membrane, cell wall or apoplast locations. The pattern of GFP fluorescence with the SBP CTPP (GFP3'prx) appears to be cytoplasmic (Figure 3.12 I C), similar to the vector control lacking signal peptides. Finally, the location of GFP with both the SBP NTPP and CTPP (5'prxGFP3'prx) was similar to that of (5'prxGFP), the "exterior" of the cell (Figure 3.12 I D). In conclusion, while NTPP can target GFP to "exterior" of cell, the exact target location(s) was not conclusive and might be cell wall, plasma membrane, or apoplast. The CTPP retained GFP signals in the

cytoplasm. For the two constructs with the NTPP, the exact location(s) were difficult to determine (Figure 3.12). Therefore, FM-64, a plasma membrane dye (T13320, Molecular Probes) was used to stain leaves for 10 min before examination. Under these conditions the green GFP signal can be distinguished from the plasma membrane which stains red (Ashtamker et al., 2007). For GFP, 5'prxGFP, and GFP3'prx transgenic plants, we screened 9 plants from 3 independent lines for each construct with consistent results. We found that GFP signals were restricted within plasma membranes for GFP and GFP3'prx constructs but GFP was targeted to outside of the plasma membranes, i.e., the apoplast for the 5'prxGFP and 5'prxGFP3'prx transgenic plants. Figure 3.13 A shows the GFP in cytosol with the GFP surrounded by the plasma membrane as revealed by the red line. The GFP signals were fainter compared with the GFP targeted to apoplast in Figure 3.13 B and D. A likely reason might be dilution or the GFP was unstable when it was exposed to proteases residing in cytosol. Figure 3.13 B illustrates GFP targeted to the apoplast by NTPP of SBP. GFP is clearly surrounded by two lines of plasma membranes in this micrograph. In Figure 3.13 B the GFP fluorescence does not appear to be as strong in the other photos. A potential reason is photo bleaching by strong light of confocal microscope (Tamura et al., 2003; Ward and Brandizzi, 2004). We found that GFP declined gradually under strong light. Figure 3.13 C indicates that GFP was localized in a narrow cytosol zone created by the large central vacuole which pushed other cellular organelles into the smaller space between the tonoplast (vacuole membrane) and plasma membrane. Figure 3.13 D shows that GFP signals were sandwiched by plasma membranes, indicating GFP is targeted to apoplast for 5'prxGFP3'prx transgenic plants. In conclusion, NTPP by itself or with CTPP can target GFP to apoplast. CTPP, as well as GFP itself (without any protein signalling peptide), retains GFP proteins in the cytosol.

Figure 3.12 Functionality of SBP Signalling Peptides

The scale bar (10 μm) was indicated on the images. A. GFP (Transgenic E-L3); B. 5'prxGFP (Transgenic 5'prx-L8); C. GFP3'prx (Transgenic 3'prx-L20); and D. 5'prxGFP3'prx (Transgenic 5'3'prx-L26).

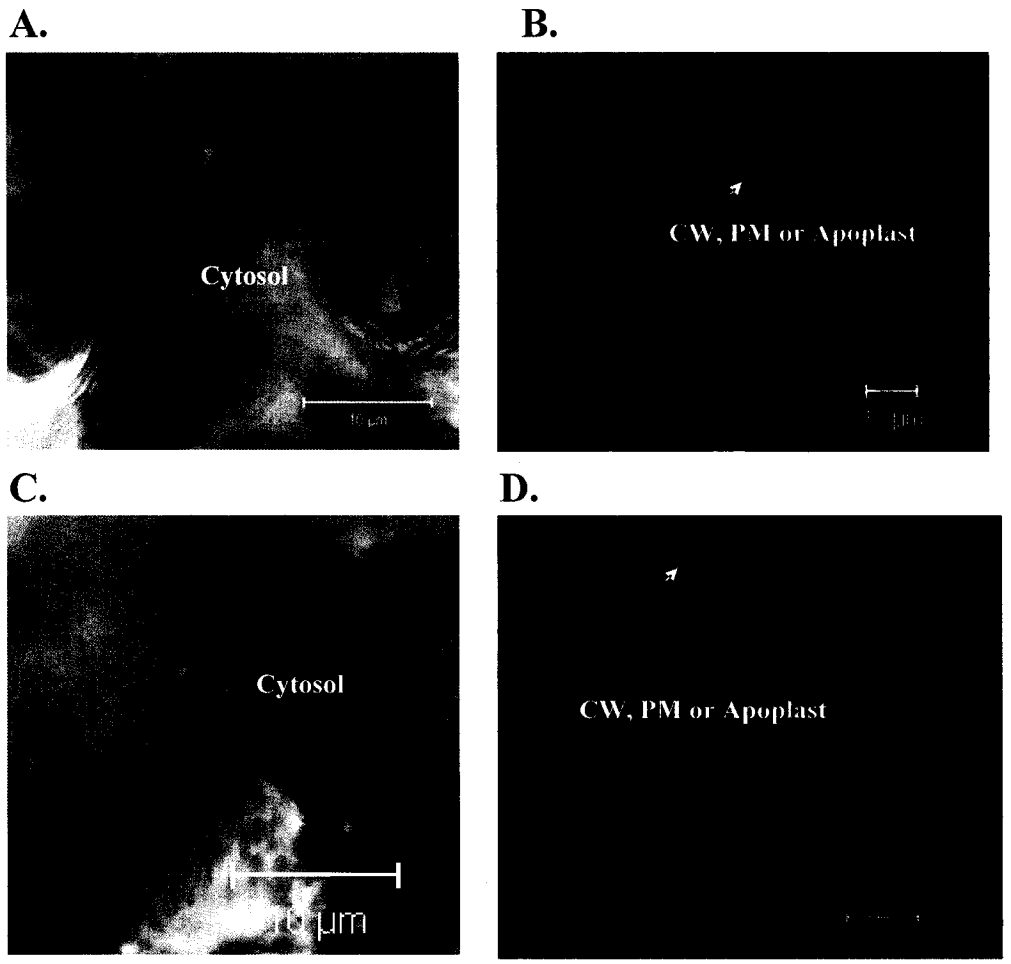
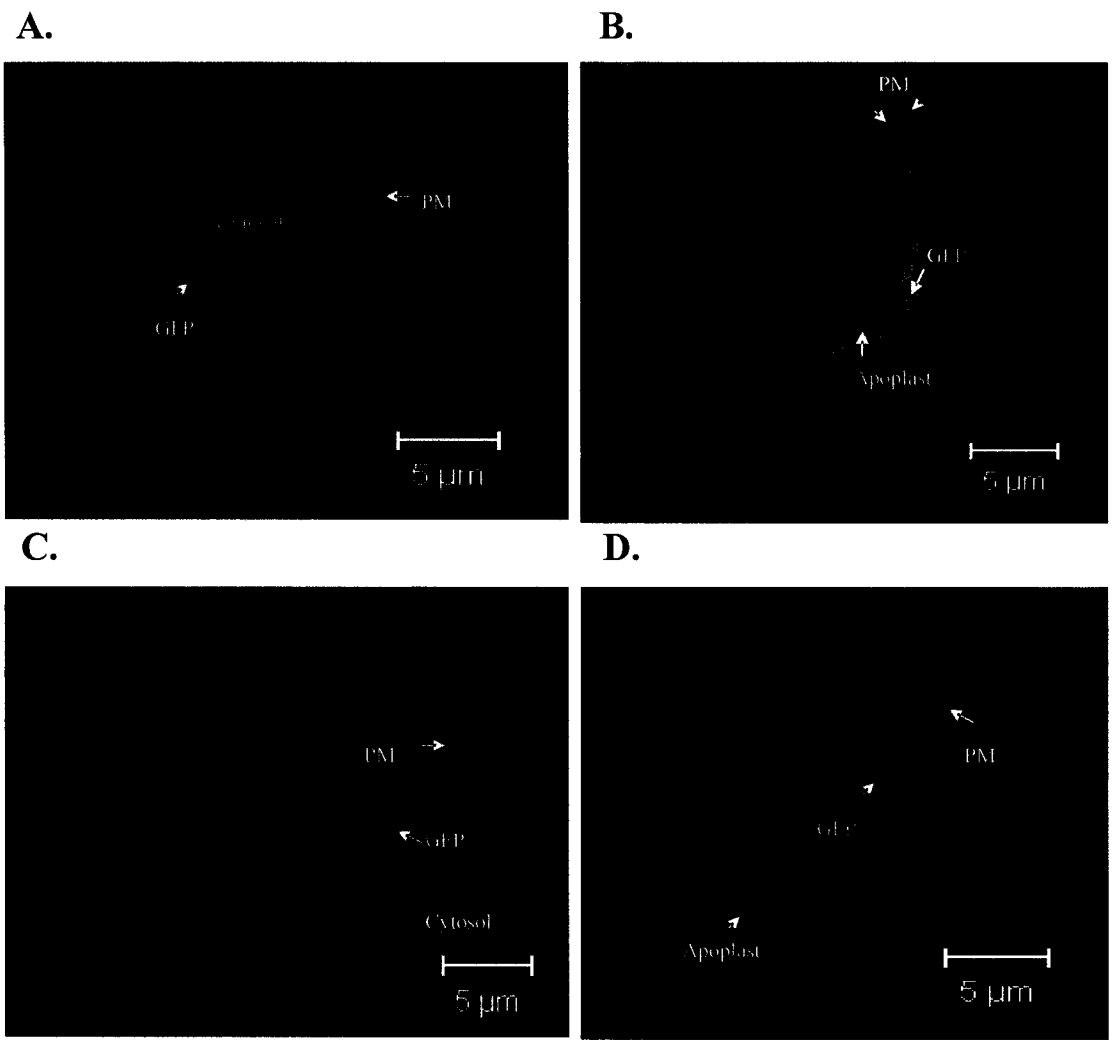


Figure 3.13 Functionality of SBP Signalling Peptides (with PM Stained)

The scale bar (5 μm) was indicated on the images. This work was performed in collaboration with Dr. Jaimie Schnell and Mr. Ann-Fook Yang at ECORC, AAFC, Ottawa. A. GFP (Transgenic E-L3); B. 5'prxGFP (Transgenic 5'prx-L8); C. GFP3'prx (Transgenic 3'prx-L20); and D. 5'prxGFP3'prx (Transgenic 5'3'prx-L26).



3.4 Screening Transgenic Soybean for HRPC and Laccase Activity

3.4.1 Collection of Plant and Seed Materials

The analysis of seed coat enzyme activity and the simultaneous recovery of embryos to ensure the next generation of plants are difficult challenges. As shown in Table 3.2, when the seed is in the T_1 generation, the seed coat and leaf are in the T_0 generation. Some T_0 plants did not give rise to seed and were lost from analysis while others yielded only 1 or 2 seeds. These plants were taken to the next generation before analysis to generate more seeds; however, the transgene was potentially segregating among the progeny.

Variation in environmental factors such as growth conditions, presence of bacteria and insects, and differences in growth cabinets may affect protein production (Mahan et al., 2004). In addition seeds may imbibe water and germinate at different rates. Since we germinated seeds on wet filter paper in Petri dishes, preliminary experiments were performed with cv. Harosoy 63 (*EpEp*) to determine whether the activity of SBP in the seed coat decreased with longer soaking times. While there was substantial variability in SBP activity between times and replicates (data not shown), overall seed coat SBP activity did not decrease with longer times of incubation in water. As a result we developed the following protocol: seeds were germinated on wet filter paper in Petri dishes at room temperature in the dark. Between 3 and 4 days the seed coats were removed manually into liquid nitrogen for eventual storage at -80°C and the sprouts were planted in vermiculite at RT for 1-2 days. Following peroxidase enzyme assays, the plants with the highest specific activity were transferred in the growth cabinets to the next generation.

3.4.2 Peroxidase Activity Units as Determined with Different Substrates

The peroxidase activity can be determined with different substrates, with the most commonly-used being ABTS, guaiacol, and pyrogallol (Morawski et al., 2000; Kawaoka et al., 2003; Matsui et al., 2006). In order to compare the different peroxidase activity units determined with different substrates, we assayed HRPC and SBP from Sigma with each of the 3 substrates (Table 3.3) under our standard conditions. In summary, one ABTS unit equalled 3.12 guaiacol units and 1.66 pyrogallol units if HRPC (Sigma) was used for assay. For SBP peroxidase activity, one ABTS unit was equivalent to 2.75 guaiacol units, and 3.49 pyrogallol units if SBP (Sigma) was used for the assay. These equivalencies were used for comparisons with published HRPC enzyme activity data measured with different substrates (also see Table 1.2).

Table 3. 2 Delineation of Generations of Transgenic Soybean Plants and Their Seeds

	Plants	Leaf	Seed	Embryo	Seed Coat[1]
T_0 Plant	Derived from callus	T_0 leaf	---	---	---
T_1 Seed			Produced from T_0 plant	T_1	T_0
T_1 Plant	Grown from T_1 seed	T_1 leaf			
T_2 Seed			Produced from T_1 plant	T_2	T_1
T_2 Plant	Grown from T_2 seed	T_2 leaf			
T_3 Seed			Produced from T_2 plant	T_3	T_2

[1] Seed coat is derived from maternal tissue and is the same generation as its parental plant.

Table 3. 3 Comparison of Peroxidase Activity Units with Three Peroxidase Substrates

Peroxidase	ABTS	Guaiacol	Pyrogallol
HRPC (Sigma)	506.9 ±11.4	162.4 ±13.0	304.8 ± 2.5
SBP (Sigma)	300.5 ±18.3	109.2 ± 7.8	86.2 ± 4.5

Notes: The average and standard error are presented (N =3). The values are in units per milligram protein where 1 unit is defined as 1 μmol of ABTS substrate oxidized per 1 min.

3.4.3 HRPC Transgenic Peroxidase Activity

We screened all T_1 seeds of 14 HRPC transgenic lines shipped from London, Ontario (AAFC) for a total of 80 seeds. Typically, we grew 3 to 7 T_1 plants per line; however, some T_1 HRPC seeds were advanced to the T_2 generation if the number of T_1 seed was limited. T_1 plant leaf tissue DNA was screened for the presence of HRPC transgenes by PCR. We did not advance all T_1 seed to an additional generation due to limited growth cabinet space. Our strategy was to assay seed coat peroxidase and advance only seedlings that had higher peroxidase enzyme activities in their seed coats to the next generation. The peroxidase activities of the seed coats are indicated in parenthesis after the plant names in Table 3.4 (the raw data are attached in Appendix IV). In total 192 T_1 and T_2 seeds were screened by the peroxidase enzyme assay. The peroxidase activity in cv. Jack soybean seed coat was 0.05± 0.01 U/mg protein (N=10). Among the lines we screened, we found the peroxidase activities in HRPC transgenic seed coat up to 60 times higher (3.09 U/mg) than that of control cv. Jack (0.05 ± 0.01 U/mg, N=10). The highest activity measured was 3.09 U/mg for S8 HRPC # 3-2-1. This level was low, but was comparable to levels found in horseradish roots by some authors (range from 3.43 to 147.89; see Table 1.2). It was also comparable to heterologous HRPC expression in leaf tissue from stable tobacco transformants ranging from 0.43-23 U /mg with pyrogallol as the substrate and 0.71-36.8 U /mg with ABTS as the substrate (Matsui et al., 2006).

Horseradish peroxidase activity differed from seed to seed in the same plant, for instance, we analysed 11 seeds of S8 HRPC # 3-2-1 plant (T_2), and the range of seed coat peroxidase activity ranged from 0.11 to 3.09 U/mg with average at 0.91 ± 0.32 U/mg. T_2 seed coat peroxidase activity (0.91U/mg) on average was higher than that of T_1

generation (0.47U/mg). Thus, screening of individual seed is necessary for understanding the range of protein accumulation in T_1 seeds.

The seed coat activity varied more in T_1 and T_2 generation, and less in the T_3 generation (Table 3.4). Overall, seed coat peroxidase activity was stably inherited, for instance, S8 HRPC #3-11 in Table 3.4. The average peroxidase activity of T_1 seed was 0.33 ± 0.13 U/mg (N = 4), that of T_2 seeds was 1.01 ± 0.25 U/mg (N=7), and that of T_3 seeds was 0.82 ± 0.11 U/mg (N = 6). The standard deviation of peroxidase activity among seed coats of T_3 S8 HRPC # 3-11 seeds was lower than that in T_1 and T_2 generations.

Advancing plants to the next generation should lead to segregation of the transgene according to Mendel's Laws. For plant HRPC 1-1 (T_2), 7 T_2 seed coats (T_1) out of 38 T_2 seeds (18.4%) had peroxidase activities that were indistinguishable from that in cv. Jack. This may indicate that the transgene is being inherited as a single locus at the expected ratio of 25% although gene silencing may contribute to this ratio.

Table 3. 4 Screening Transgenic Soybean Seed Coats for Peroxidase Activity

HRPC Transgenic Soybean	# Seed Analyzed	Average (U/mg) (± SE) [6]	Range (U/mg)	Generation [2]
cv. Jack	10	0.05 ± 0.01	0.03 - 0.09	---
S8 HRPC # 1 [1]	4	0.19 ± 0.04	0.09 ~ 0.26	T_1
S8 HRPC # 1-1 (0.26)	38	0.43 ± 0.09	0.03 ~ 2.66	T_2
S8 HRPC # 1-2 (0.23)	5	0.77 ± 0.2	0.18 ~ 1.29	T_2
S8 HRPC # 3-2	2	0.47 ± 0.22	0.25 ~ 0.68	T_1
S8 HRPC # 3-2-1 (0.68)	11	0.91 ± 0.32	0.11 ~ 3.09	T_2
S8 HRPC # 3-2-2 (0.25)	3	1.26 ± 0.14	1.11~1.54	T_2
S8 HRPC # 3-4	4	0.48 ± 0.06	0.36 ~ 0.66	T_1
S8 HRPC # 3-4-1 (0.66)	5	0.22 ± 0.02	0.14 ~ 0.26	T_2
S8 HRPC # 3-4-2 (0.45)	3	0.47 ± 0.01	0.18 ~ 0.64	T_2
S8 HRPC # 3-4-3 (0.43)	2	0.52 ± 0.1	0.42 ~ 0.61	T_2
S8 HRPC # 3-5	2	--- [4]	---	T_1
S8 HRPC # 3-5-1 (---)	3	0.32 ± 0.04	0.25 ~ 0.39	T_2
S8 HRPC # 3-5-2 (---)	5	0.26 ± 0.05	0.10 ~ 0.36	T_2
S8 HRPC # 3-6	15	0.31 ± 0.05	0.02 ~ 0.67	T_1
S8 HRPC # 3-6-1 (0.67)	4	0.33 ± 0.05	0.26 ~ 0.48	T_2
S8 HRPC # 3-6-2 (0.56)	3	0.55 ± 0.09	0.43 ~ 0.73	T_2
S8 HRPC # 3-6-3 (0.55)	3	1.03 ± 0.42	0.50 ~ 1.86	T_2
S8 HRPC # 3-8	1	0.55	0.55	T_1
S8 HRPC # 3-8-1 (0.55)	5	0.53 ± 0.13	0.28 ~ 1.02	T_2
S8 HRPC # 3-8-1-1 (1.02)	3	0.11 ± 0.007	0.22 ~ 0.28	T_3
S8 HRPC # 3-8-2 (---)	5	0.69 ± 0.19	0.36 ~ 1.33	T_2
S8 HRPC # 3-8-2-4 [5] (0.95)	7	0.44 ± 0.06	0.21 ~ 0.65	T_3
S8 HRPC # 3-10	2	---	---	T_1
S8 HRPC # 3-10-2 (---)	4	0.33 ± 0.10	0.22 ~ 0.62	T_2
S8 HRPC # 3-11	4	0.33 ± 0.13	0.04 ~ 0.68	T_1
S8 HRPC # 3-11-1 (0.68)	7	1.01 ± 0.25	0.13 ~ 2.20	T_2
S8 HRPC # 3-11-1-4 (1.41)	6	0.82 ± 0.11	0.56 ~ 1.34	T_3
S8 HRPC # 3-11-2 (0.33)	2	0.65 ± 0.37	0.28 ~ 1.01	T_2
S8 HRPC # 5 [3]	13	0.23 ± 0.05	0.02 ~ 0.46	T_1
S8 HRPC # 5-1 (0.46)	3	0.54 ± 0.08	0.39 ~ 0.68	T_2
S8 HRPC # 12 [3]	7	---	---	T_1
S8 HRPC # 13 [3]	15	0.22 ± 0.03	0.03 ~ 0.40	T_1
S8 HRPC # 14 [3]	6	0.20 ± 0.04	0.10 ~ 0.35	T_1
S8 HRPC # 15 [3]	1	---	---	T_1
S8 HRPC # 15-1 (---)	5	0.06 ± 0.01	0.03 ~ 0.08	T_2
S8 HRPC # 16 [3]	2	0.29 ± 0.17	0.12 ~ 0.45	T_1

Notes:

The numbering system includes the name of the vector, the transgene, the bombardment event name. For example, S8 HRPC #3-8 designates the vector S8 (Figure 3.4), the HPRC transgene (Appendix I) and the event # 3-8 (in some cases a single digit is given). The designations S8 HRPC #3-8-1 (0.55) and S8 HRPC #3-8-1-1 (1.02) refer to subsequent generations that derived from event S8 HRPC #3-8. The number in the bracket refers to seed coat peroxidase enzyme activity measured in the seed that plant is derived from.

[1] The names of T_1 HRPC transgenic soybean seeds are in the bold, representing the independent transformation events;

[2] In this table, the generation in the last column refers the generation of embryo. The relationship between generations of seed embryo and seed coat is that the generation of seed coat is that of embryo minus one (see Table 3.2).

[3] We advanced seven T_2 plants for each of S8 HRPC # 5, #12, #13, and three T_2 plants for each of S8 HRPC # 14, #16. The T_2 seeds of those plants were sent to Guelph University for further research;

[4] T_1 seed coat enzyme data were not available;

[5] The plants highlighted with yellow were chosen for further molecular characterization based on the higher level of peroxidase activity and completeness of transgene expression cassette they harboured.

[6] average peroxidase activity for all seed coats assayed in the defined generation without subtracting the value for cv. Jack.

3.4.4 Screening Transgenes in HRPC Plants by PCR

T_1 plants were selected for higher peroxidase activity in T_0 seed coats. Leaf samples were collected and gDNA was isolated using Sigma kit for PCR as described in Section 3.2. We first tested the gDNA quality by testing for the presence of the soybean tubulin gene employing GmTub2 primers. The primer pairs for several regions of the transgene and *hpt* II (Table 3.1) were used for the higher seed coat peroxidase expressors.

None of the HRPC transgenic plants that showed higher peroxidase activity tested positively for all components of the transgene vector suggesting that transformation had lead to vector rearrangement in these plants. For example both S8 HRPC # 3-2-1-2 (2.74 U/mg) and S8 HRPC # 3-6-3-2 (1.86 U/mg) lacked detectable HRPC sequences. Nor did they appear to have fragments corresponding to the Ep promoter, hptII gene and tCUP promoter. In addition, HRPC # 3-6-3-2 lacked the NOS sequence (Discussed in section 4.3).

Other higher peroxidase expressors such as S8 HPRC 3-8-1-1 (1.02 U/mg), S8 HRPC # 3-8-2-4 (0.95 U/mg), S8 HRPC # 3-11-1-4 (1.41 U/mg) and S8 HPRC # 16-1 (0.45 U/mg) had all of the components tested except for the Ep1.5 kb promoter which was truncated, with between 0.6 and 1.2 kb remaining (Table 3.5). Seed coat peroxidase was observed in two lines that lacked a full length HRPC gene. This phenomenon is discussed in Section 4.3. None had the Ep 1.5 kb promoter. We therefore used PCR to test for the presence of deleted forms of the promoter (Table 3.1). These results showed that S8 HRPC # 3-8-2-1 and # 3-8-2-4 contained at least 1.2 kb (but not 1.5 kb) of the promoter. S8 HRPC # 3-11-1-4 contained at least 0.9 kb (but not 1.2 kb). S8 HRPC # 16-

1 (0.45 U/mg) contained 0.6 kb Ep promoter. Both S8 HRPC 3-2-1-2 and S8 HRPC # 3-6-3-2 had less that 0.3kb of promoter sequence, which is our limit of detection.

We sought to find the HRPC transgenic plants with high peroxidase seed coat activity, longer Ep promoters (at least 0.9 kb) and full-length HRPC synthetic coding DNA. Out of 15 independent HRPC transformants, we found a T_2 plant, S8 HRPC # 3-8-2-4 (0.95 U/mg) that harboured the full-length HRPC synthetic gene driven by 1.2 kb Ep promoter and another T_2 plant, S8 HRPC # 3-11-1-4 (1.41 U/mg) that had a 0.9 kb Ep promoter followed by the full-length HRPC synthetic gene (Table 3.5). It has previously been shown that the 0.6 kb Ep promoter could restrict the GUS expression to the hourglass cell layer of soybean seed coats (Dr. L. Gudynaite-Savitch, unpublished data). The activity and specificity of truncated promoters needs to be tested in this project.

Table 3. 5 Screening Selected HRPC Transgenic Plants by PCR

Genes	***GmTub2***	***NOS***	***tCUP***	***hpt*II**	**Length of *Ep* promoter**	**Full-length *HRPC* gene**
WT (Jack)	+	-	-	-	-	-
S8 HRPC Plasmid	-	+	+	+	+	+
S8 HRPC # 3-2-1-2	+	+	-	-	-	-
S8 HRPC # 3-6-3-2	+	-	-	-	-	-
S8 HRPC # 3-8-1-1	+	+	+	+	(Ep1.2 kb)	+
S8 HRPC # 3-8-2-4	+	+	+	+	(Ep1.2 kb)	+
S8 HRPC # 3-11-1-4	+	+	+	+	(Ep0.9 kb)	+
S8 HRPC # 16-1	+	+	+	+	(Ep0.6 kb)	+

Notes: + represents the presence of the expected PCR product, - represents the absence of a product. The primers are given in Table 3.1

Our PCR screening results showed that the Ep 1.5 kb promoter existed in the T_0 plants S8 HRPC # 3-2, and S8 HRPC # 3-6 but lost in T_1 plants that were grown from T_1 seeds shipped from London, ON (Table 3.6). The fact that all T_1 plants of S8 HRPC # 3-2, S8 HRPC # 3-4, and S8 HRPC # 3-6 were Ep-promoter-PCR-product negative, and all the T_1 plants of S8 HRPC # 3-5, and S8 HRPC # 14 (Table 3.6) were HRPC-PCR-product negative indicated that the reason for Ep promoter loss was not only due to the transgene segregation but might be generated by rearrangement, transgene excision or deletion. Transgene loss was reported in *Phytophthora parasitica* (Scheid et al., 1991) and aspen (Melander et al., 2006). It is believed that multiple-copy integration and subsequent gene loss were the major reasons for the complex segregation behaviours of the transgene in transgenic plants (Zhao et al., 2007). Furthermore, it was found that HRPC transgene was inherited to from the T_0 to the T_1 generation of S8 HRPC 3-2, but it was lost in the T_2 generation. In S8 HRPC 3-4, one of T_1 plants lost the HRPC transgene while other 2 plants inherited it from T_0 plant. This scenario also happened to the lines of HRPC # 3-6. The transgene loss in transgenic soybean by particle bombardment was also found in Dr. D Simmonds' lab (AAFC, Ottawa). They concluded that if transgenes were found in T_3 plants, the transgene would likely be inherited stably (Ogasawara et al., 2005) because the transgene rearrangements were most frequent in T_1 and T_2 generations (Dr. Rugang Li, oral communication).

Table 3. 6 Transgene Segregation Through Generations

Plants	Plant Generation	*Ep* 1.5 promoter plus HRPC [3]	Full-length HRPC gene [4]
Jack	---	-	-
S8 HRPC plasmid	---	+ [1]	+
S 8 HRPC # 3-2*	T_0	+	+
S 8 HRPC # 3-2-1	T_1	- [2]	+
S 8 HRPC # 3-2-2	T_1	-	-
S 8 HRPC # 3-2-1-2	T_2	-	-
S 8 HRPC # 3-4	T_0	+	+
S 8 HRPC # 3-4-1	T_1	-	-
S 8 HRPC # 3-4-2	T_1	-	-
S 8 HRPC # 3-4-3	T_1	-	+
S 8 HRPC # 3-5	T_0	-	+
S 8 HRPC # 3-5-1	T_1	-	-
S 8 HRPC # 3-5-2	T_1	-	-
S 8 HRPC # 3-6*	T_0	+	+
S 8 HRPC # 3-6-1	T_1	-	+
S 8 HRPC # 3-6-2	T_1	-	+
S 8 HRPC # 3-6-3	T_1	-	+
S 8 HRPC # 3-6-1-4	T_2	-	+
S 8 HRPC # 3-6-3 -1	T_2	-	+
S 8 HRPC # 3-6-3 -2	T_2	-	-
S 8 HRPC # 14	T_0	-	+
S 8 HRPC # 14-1	T_1	-	-
S 8 HRPC # 14 -2	T_1	-	-
S 8 HRPC # 14 -3	T_1	-	-

Notes:

[1] "+" represents the presence of the expected PCR product;

[2] "-" represents the absence of a product;

[3] the length of Ep promoter was determined by the primer set Ep-15 F32 / HRPC (243) Rev shown in Table 3.1;

[4] the full-length HRPC gene was determined using the primer set HRPC-Not1-For / HRPC-Xba1-Rev indicated in Table 3.1.

3.4.5 Western Blot Analysis of HRCP Transgenic Plant S8 HRPC 16-1

We attempted to use an anti-HRPC antibody to detect the expression of HRPC protein in HRPC transgenic soybean. Extracts from seed coat and embryo of transgenic seeds were compared to extracts of the same tissues from cv. Jack (*epep*, negative controls). Commercial HRPC and SBP (Sigma) proteins were used as positive controls. Given the relatively low levels of expression, we pooled seed coat protein samples from related plants with higher peroxidase activities, measured protein concentration and then ran SDS gels. Based on our controls and the levels of activity that were measured, we were close to the limit of detection. HRPC proteins could not be detected with either HRPC or His-tag antibodies used as probes (data not shown).

We were able to take advantage of the HIS tag to purify HRPC from seed coats and embryos of 80 pooled samples from S8 HRPC 16-1 (T_2), a plant with a truncated Ep promoter (0.6 kb) but HRPC activity. The QIAGEN HIS tag column enriches for proteins with 3 consecutive HIS (oral communication, Dr. H. Schernthaner, AAFC, Ottawa) and recombinant HRPC has a C-terminal with 6XHis residues (Appendix 1). In Figure 3.14 A, the Western blot was probed with anti-HRPC antibody and in Figure. 3.14 B a replicate blot was probed with anti-His-tag antibody.

Figure 3.14 A, lane 2 represents the purified seed coat proteins from S8 # HRPC 16-1 probed with anti-HRPC antibody. To compare HIS tag purified band patterns (lane 2) with that of total cv. Jack seed coat proteins, we ran cv. Jack seed coat protein as the negative control in lane 3. The anti-HRPC antibody detected 6 bands with estimated molecular weights of ca 76.0, 64.6, 39.4, 37.4, 35.4 and, 30.0 kDa (lane 2 of Figure 3.14 A) whereas many bands were detected in lane 3 indicating that purification by QIAGEN

HIS tag column had worked. Probing of the replicate transfer in Figure 3.14 B with His-tag antibody revealed one faint band as indicated by the arrow with an estimated molecular weight of 35.4 kDa. When this blot was repeated a size of 34.7 kDa was determined (data not shown). This band may correspond to the band at 35.4 kDa band in lane 2 (Figure 3.14 A) detected with the anti-HRPC antibody (indicated by the arrow). Unfortunately we did not have enough material to do multiple repeats. This band was not seen in the untransformed control. Therefore both antibodies detect a protein of about the same size that contains epitopes for HRPC and His, suggesting that it may be HRPC. This putative HPR C has a much lower molecular weight than HPR C from Sigma run in lane 5. The size of the full length HRPC from *A. rusticana* calculated from the DNA sequence is 38.7 kDa for a protein of 353 aa (accession #: AAA33377). Several lines of evidence suggest that HPR C has NTPP and CTPP groups that are removed during targeting to the vacuole in transgenic systems (Kis et al., 2004; Matsui et al., 2006). When Fijiyama (Fujiyama et al., 1988) removed these putative signals (NTPP, 30 aa and CTPP, 15 aa) from their deduced protein sequence, a size of 33.9 kDa (308 aa) was determined. And the mature protein isolated from *A. rusticana* has a size of 33.9 kDa (308 aa) as determined from the protein sequence (Welinder, 1979). Based on our S8 HRPC soybean construct (Appendix 1), and assuming that the NTPP and CTPP were removed, the mature HRPC, including GGAA adaptors and 6XHIS tag will have a size of 35.2 kDa (322 aa) if glycosylation is not taken into consideration. Thus the size of the expected protein is 35.2 kDa. However, the size of HPR C (Sigma) estimated by SDS gel electrophoresis is ~50 kDa which is similar to previous estimates of 43.8 kDa (Regalado et al., 1996), 50 kDa (Loustau et al., 2008) and 44 kDa (Welinder, 1979) using the same

technique. The difference between the calculated and measured sizes may in part reflect the extensive glycosylation of native HRPC, estimated at 18% of the mass (Shannon et al., 1966). Sizes of glycosylated proteins are often over-estimated by SDS gels (Bohnsack et al., 2000). If so, then the putative HPRC identified in Figure 14 may not be extensively glycosylated. We also ran the HIS tag purified embryo proteins from S8 # HRPC 16-1 in lane 6 while total protein of cv. Jack embryo was used as a control. The two patterns are very similar. The 35.4 kDa band found in lane 2 was not seen in lane 6, and 7 of Figure 3.14 A, suggesting the 35.4 kDa band was only expressed in the HRPC transgenic soybean seed coat. In lanes 6 and lane 7 of Figure 3.14 B, a band with an estimated molecular weight 33.0 kDa (32.7 kDa when repeated) was detected with anti-HIS tag antibody. A protein of this size was also detected during experiments to determine antibody specificity (Figure 3.11 E). This band is not likely to be the putative HPR C expressed in the embryo for two reasons. The size of the protein in lane 7 is smaller that the size in lane 2 as calculated for each of the two gels. In addition, if this was HRPC then we would expect a band of the same size to be revealed with the anti-HRPC antibody. As indicated above, no band of the corresponding size is observed (Figure 3.14A lane 6). The band in the soybean embryo protein sample (lane 7) detected with anti-HIS tag antibody may represent the detection of an endogenous soybean protein since 45 hits using HHHHH as the probe were found in soybean genome (BLASTP: accessed at http://blast.ncbi.nlm.nih.gov/Blast.cgi, April 16, 2009). While this evidence suggests that HRPC expression can be detected by Western blot analysis, more replicates with different transgenic events are needed to confirm the result. Since the level of HPRC

expression is relatively low, we will need to accumulate a lot of seed for seed coat extraction-in the above experiments 80 seeds were processed.

3.5 Screening Laccase Transgenics

3.5.1 Laccase Transgenic Lines

T_0 S8 LCC1 and LCC4 transgenic soybean plants were grown in London Ontario, T_1 seeds were harvested and shipped to Ottawa. In total, we obtained 5 lines of S8 LCC1 transgenic soybean, and 36 lines of S8 LCC4 transgenic soybean. We also had T_0 plants shipped from London, including 11 lines of S18 LCC1 transgenic soybeans, and 4 lines of S18 LCC4 transgenic soybeans (Appendix 5).

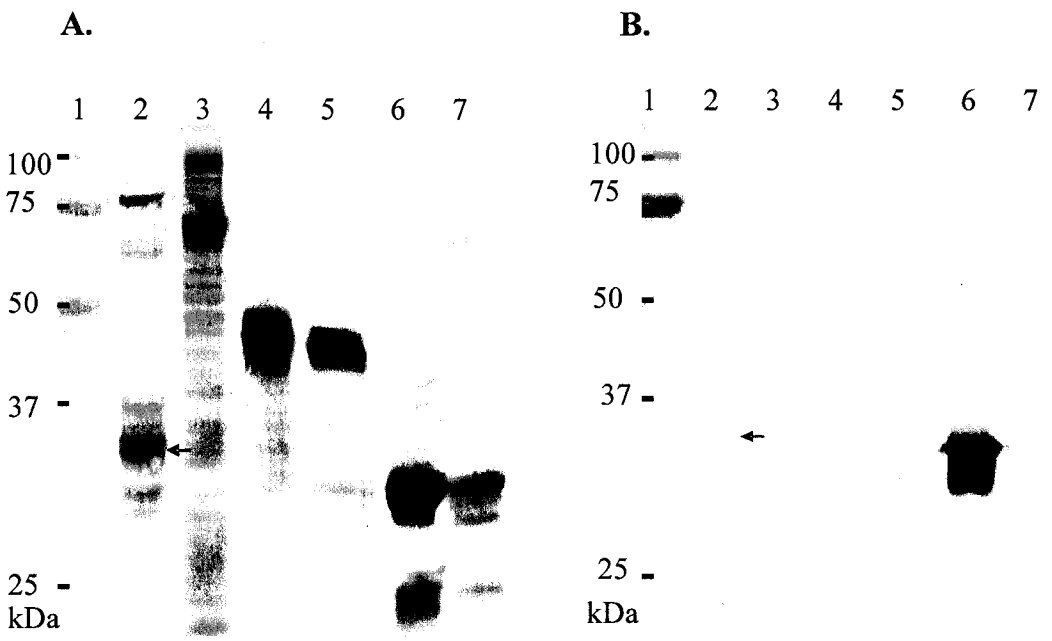
To get more transgenic soybean seeds to work with we advanced some of the T_1 seeds of S8 LCC1 and S8 LCC4 transgenic soybean seeds to the T_2 generation. PCR screening of T_1 plants was performed to minimize unnecessary seed coat laccase enzyme assays, and to focus on promising transgenics with intact genes since a known side effect of bombardment transformation is transgene rearrangement (Finer and McMullen, 1991; Patnaik and Khurana, 2001; Reddy et al., 2003; Zhao et al., 2007) as seen for the HRPC transgenics. We used LCC1, LCC4 and hpt II primer sets. Ten out of 36 S8 LCC4 lines, 2 out of 5 S8 LCC1 lines and 1 out of 11 S18 LCC1 lines were transgene negative and discarded. These results demonstrated the need to monitor transgene stability usually required to obtain stable homozygous lines when T_0 plants were taken up to the T_3 or T_4 generation (Verweire et al., 2007).

Figure 3.14 A Western Blots of Purified HRPC Proteins Detected with Anti-HRPC Antibody

A Western blot was probed with anti-HRPC antibody at 3000X, detected with goat anti-rabbit IgG with AP antibody at 1000X (Molecular Probes). The proteins signal was recorded 15 min after overlay. L1, Protein ladder (BioRad # 161-0373); L2, Purified HRPC proteins from S8 HRPC 16-1 seed coat, (100 ng); L3, Jack Seed Coat proteins, (25 μg); L4, HRPC (Sigma) protein, (50 ng); L5, SBP (Sigma) protein, (50 ng); L6, Purified HRPC proteins from S8 HRPC 16-1 embryo, (100 ng); L7, Jack embryo proteins, (25 μg)

Figure 3.14 B Western Blots of Purified HRPC Proteins Detected with Anti-HIS Tag Antibody

A Western blot was probed with anti-His tag antibody at 1000X, detected with goat anti-mouse IgG with AP antibody at 1000X (Molecular Probes). The proteins signal was recorded 30 min after overlay. The arrow indicates the protein discussed in the text. L1, Protein ladder (BioRad # 161-0373); L2, Purified HRPC proteins from S8 HRPC 16-1 seed coat, (100 ng); L3, Jack Seed Coat proteins, (25 μg); L4, HRPC (Sigma) proteins, (50 ng); L5, SBP (Sigma) proteins, (50 ng); L6, Purified HRPC proteins from S8 HRPC 16-1 embryo, (100 ng); L7, Jack embryo proteins, (25 μg)



Seed coats were assayed for laccase activity using ABTS as the substrate. For T_1 LCC transgenic soy, we chose to assay 15 seeds for each independent line. For the lines with fewer than 15 seeds, we screened them all. Promising plants with higher laccase activities in the seed coat were advanced to the T_2 generation. Leaf material was harvested for gDNA analysis and seeds for laccase enzyme assays. The detailed information on the number of seeds assayed, the mean value of laccase activity, the range of laccase within a line or in specific T_2 or T_3 plants were tabulated in Table 3.7 and Table 3.8. In total 404 laccase T_1 and T_2 transgenic soybean seeds were assayed.

The laccase activity in cv. Jack soybean seed coat was 0.01± 0.003 U/mg (N=10) when assayed as described in Appendix 3 using ABTS as the substrate. The raw data of laccase activity in individual seed coats can be found in Appendix 5. Laccase activity in the transgenic soybean varied from line to line, and seed to seed. The laccase activity in S8 LCC1, S18 LCC1, S8 LCC4 and S18 LCC4 transgenic seed coats was up to 9 fold higher (0.099 U/mg) than that of cv. Jack.

Based upon data from laccase enzyme activity and transgene screening, we chose S8 LCC1# 4-2-2-1, and S8 # LCC1 5-1-2 plants as the candidates for S8 LCC1 transgenic soybean, and S8 LCC4 # 16-1-1, and S8 # LCC4 30-1-4-2 plants as the promising candidates for S8 LCC4 further research since all them contain 1.2 kb or 1.5 kb promoters, and full-length of LCC1 or LCC4 synthetic genes (Table 3.9). These were indicated in yellow.

Table 3.7 The Screening of S8 LCC1, and S8 LCC4 Transgenic Soybean Seeds

HRPC Transgenic Soybean	# Seed Analyzed	Average (U/mg) (± SE) [4]	Range (U/mg)	Generation [2]
Jack	10	0.01 ± 0.003	0 ~ 0.021	---
S8 LCC1 # 4-1 [1]	8	0.028 ± 0.003	0.01 ~ 0.035	T_1
S8 LCC1 # 4-1-1 (0.035)	3	0.011 ± 0.001	0.009 ~ 0.012	T_2
				T_1
S8 LCC1 # 4-2	7	0.031 ± 0.002	0.023 ~ 0.035	
S8 LCC1 # 4-2-2 (0.035)	5	0.032 ± 0.004	0.021 ~ 0.04	T_2
S8 LCC1 # 4-2-2-1 (0.04) [3]	10	0.030 ± 0.003	0.014 ~0.049	T_3
				T_1
S8 LCC1 # 5	15	0.029 ± 0.003	0.004 ~ 0.035	
S8 LCC1 # 5-1 (0.045)	4	0.014 ± 0.005	0.009 ~ 0.027	T_2
S8 LCC1 # 5-1-2 (0.027)	13	0.010 ± 0.002	0.006 ~ 0.024	T_3
				T_1
S8 LCC1 # 10-2	11	0.001 ± 0.001	0 ~ 0.005	
S8 LCC1 # 12-2	8	0.001 ± 0.001	0 ~ 0.005	T_1
S8 LCC4 # 1	10	0.02 ± 0.003	0 ~0.04	T_1
S8 LCC4 # 2	8	0.008 ± 0.003	0~0.023	T_1
S8 LCC4 # 3	1	0.03	0.03	T_1
S8 LCC4 # 5	10	0.033 ± 0.003	0.009 ~0.043	T_1
S8 LCC4 # 9-2	2	0.026 ± 0.005	0.02 ~0.031	T_1
S8 LCC4 # 11	2	0.021 ± 0.019	0.002 ~0.04	T_1
S8 LCC4 # 13	3	0.002 ± 0.002	0 ~0.006	T_1
S8 LCC4 # 14	3	0.002 ± 0.002	0 ~0.005	T_1
S8 LCC4 # 15-2	2	0.035 ± 0.005	0.03 ~0.04	T_1
S8 LCC4 # 15-4	1	0.04	0.04	T_1
S8 LCC4 # 15-6	1	0.005	0.005	T_1
S8 LCC4 # 16	2	0.035 ± 0.005	0.03 ~0.04	T_1
S8 LCC4 # 16-1 (0.04)	13	0.031 ± 0.007	0 ~0.099	T_2
S8 LCC4 # 16-1-1 (0.099)	11	0.025± 0.004	0.009 ~ 0.05	T_3
S8 LCC4 # 17-1	12	0.017 ± 0.004	0~ 0.032	T_1
S8 LCC4 # 17-2	5	0.015 ± 0.006	0.001 ~0.036	T_1
S8 LCC4 # 18	13	0.011 ± 0.002	0 ~ 0.023	T_1
S8 LCC4 # 19-1	10	0.008 ± 0.002	0 ~ 0.019	T_1
S8 LCC4 # 19-2	10	0.005 ± 0.002	0 ~ 0.016	T_1
S8 LCC4 # 21	9	0.019 ± 0.004	0 ~ 0.032	T_1
S8 LCC4 # 24	8	0.023 ± 0.008	0.002 ~ 0.066	T_1
S8 LCC4 # 26	7	0.023 ± 0.005	0 ~ 0.04	T_1
S8 LCC4 # 27	15	0.009 ± 0.003	0 ~ 0.04	T_1
S8 LCC4 # 30-1	4	0.023 ± 0.006	0.01 ~ 0.04	T_1
S8 LCC4 # 30-1-4 (0.04)	4	0.028 ± 0.006	0.01 ~ 0.04	T_2
S8 LCC4 # 30-1-4-2 (0.04)	10	0.037± 0.004	0.023 ~ 0.063	T_3
S8 LCC4 # 31	10	0.016 ± 0.004	0 ~ 0.03	T_1
S8 LCC4 # 32-1	6	0.027 ± 0.008	0 ~ 0.05	T_1
S8 LCC4 # 32-2	15	0.018 ± 0.005	0 ~ 0.052	T_1
S8 LCC4 # 33	1	0.01	0.01	T_1

The numbering system includes the name of the vector, the transgene, the bombardment line name. For example, S8 LCC4 # 16 designates the vector S8 (Figure 3.4), the LCC4 transgene (Appendix 1) and the line (in some cases a single digit is given). The designations S8 LCC4 #16-1 (0.04 U/mg) and S8 LCC4 #16-1-1 (0.099 U/mg) refer to subsequent generations that derived from line S8 LCC4 #16, and the 0.04 U/mg means the seed coat laccase activity of the seed in which plant S8 LCC4 # 16-1 was derived from.

[1] The names of T_1 laccase transgenic soybean, representing the independent transformation events;

[2] In this table, the generation in the last column refers to plant generation, or the generation of seed (embryo);

[3] The plants highlighted in yellow were chosen for further molecular characterization based on the higher level of laccase activity and completeness of transgene expression cassette they harboured.

[4] average laccase activity for all seed coats assayed in the defined generation without subtracting the value for cv. Jack.

Table 3. 8 The Screening of S18 LCC1, and S18 LCC4 Transgenic Soybean Seeds

HRPC Transgenic Soybean	# Seed Analyzed	Average (U/mg) (± SE)	Range (U/mg)	Generation [2]
Jack	10	0.01 ± 0.003	0 ~ 0.021	---
S18 LCC1# 3A [1]	15	0.023 ± 0.004	0 ~ 0.045	T_1
S18 LCC1# 2B	5	0.004 ± 0.001	0 ~ 0.007	T_1
S18 LCC1# 3B	11	0.001 ± 0	0 ~ 0.005	T_1
S18 LCC1# 2C	5	0.035 ± 0.004	0.018 ~ 0.041	T_1
S18 LCC1# 3C	15	0	0 ~ 0.003	T_1
S18 LCC1# 2E	3	0	0	T_1
S18 LCC1# 2F	15	0.001 ± 0	0 ~ 0.006	T_1
S18 LCC1# 2G	15	0.025 ± 0.003	0 ~ 0.05	T_1
S18 LCC1# 2J	14	0.005 ± 0.001	0 ~ 0.01	T_1
S18 LCC1# 2H	15	0.001 ± 0	0 ~ 0.005	T_1
S18 LCC1# 2I	12	0.03 ± 0.001	0.022 ~ 0.034	T_1
S18 LCC4# 1A	15	0.006 ± 0.004	0 ~ 0.05	T_1
S18 LCC4# 1B	6	0.005 ± 0	0.004 ~ 0.006	T_1
S18 LCC4# 1C	11	0.01 ± 0.003	0 ~ 0.023	T_1
S18 LCC4#1D	15	0.03 ± 0.002	0 ~ 0.04	T_1

[1] The names of T_1 laccase transgenic soybean, representing the independent transformation events;

[2] In this table, the generation in the last column refers to plant generation, or the generation of seed (embryo);

Table 3. 9 Screening Selected S8 LCC1 and S8 LCC4 Transgenic Plants by PCR

Genes	*GmTub2*	*NOS*	*tCUP*	*hpt*II	Length of *Ep* promoter [3]	Full-length *LCC1*/*LCC4* gene[4]
WT (Jack)	+ [1]	- [2]	-	-	-	-
S8 LCC1 4-2-2-1	+	+	+	+	1.2 kb	+
S8 LCC1 5-1-2	+	-	-	+	1.2 kb	+
S8 LCC4 16-1-1	+	+	+	+	1.5 kb	+
S8 LCC4 30-1-4-2	+	+	+	+	1.5 kb	+

Notes:

[1]"+" represents the presence of the expected PCR product;

[2] "-" represents the absence of a product;

[3]The length of Ep promoter was determined with Ep-1.5 F32 / LCC1 (198) Rev for S8 LCC1 transgenic plants, and Ep-1.5 F32 / LCC4 (214) Rev for S8 LCC4 transgenic plants (Table 2.1); PCR working conditions were tabulated in Table 3.1.

[4] Full-length synthetic genes were determined with LCC1-Nco I-OP / LCC1-Sal I – OP for LCC1 synthetic gene, and LCC4-Nco I-For / LCC4-Sal I-Rev for LCC4 synthetic gene (Table 2.1); PCR working conditions were tabulated in Table 3.1.

3.5.2 Laccase Activity Heritability from Generation to Generation

To determine whether laccase activity was stably inherited, we monitored the changes of laccase activities from T_1 to T_3 generation for selected lines. The variation in laccase activity in laccase transgenic soybean seed coats was much less than in HRRC transgenic plants. The data in the Table 3.10 showed that laccase activity as a trait of laccase transgenic seed coat could be stably inherited. In T_1 generation, the range of seed coat laccase activity in the different lines varied from 0.023 to 0.031 U/mg proteins. It was found that the seed coat laccase activity fluctuated a great deal from generation to generation, for instance, laccase activities of S8 LCC1 # 5-1-2 (0.032 U/mg, T_3) was more than 2-fold higher than that in T_2 generation (0.014 U/mg). We found the S8 LCC4 # 16-1 plant with 0.099 U/mg laccase, which was the highest level of expression found to date. Among of the seed coats tested, some had little or no laccase activity indicating that gene loss or silencing may have occurred in these specific plants (discussed in Section 4.2).

Table 3. 10 The Laccase Activity Heritability of Generations

Plants	T_1	T_2	T_3
S8 LCC1 #4-2-2-1	0.031 ± 0.002 (N =7)	0.032 ± 0.004 (N =5)	0.027± 0.004 (N =10)
S8 LCC1 #5-1-2	0.029 ± 0.003 (N=10)	0.014 ± 0.005 (N =4)	0.032 ± 0.005 (N=13)
S8 LCC 4 #16-1-1	0.035 ± 0.005 (N=2)	0.031 ± 0.007 (N =13)	0.025± 0.004 (N=11)
S8 LCC4 #30-1-4-2	0.023 ± 0.006 (N=4)	0.028 ± 0.006 (N=4)	0.037± 0.004 (N=10)

Laccase activity in soybean seed coat (cv. Jack): 0.01 ± 0.003 U/mg (N =10)

3.5.3 Western Blot Analysis of LCC4 Transgenic Plant S8 LCC4 # 16-1

Since the level of expressed laccase proteins in soybean seed coat was too low to be seen in Western blots (data not shown), we purified the expressed LCC4 proteins. We purified LCC4 from seed coats of 80 pooled samples from S8 LCC4 16-1 (T_2), a plant with full length of Ep promoter and full length synthetic LCC4 gene, and higher laccase activity in its seed coat. The presence of the HIS tag linked to the expressed laccase allows the use of QIAGEN HIS tag columns to purify HIS tagged LCC4 proteins.

In Figure 3.15 A, the Western blot was probed with anti-LCC4 antibody and in Figure. 3.15 B a replicate blot was probed with anti-His-tag antibody. Enough protein was recovered to run duplicate blots. In Figure 3.15 C a Western blot developed by chemiluminescence was shown. Figure 3.15 A, lane 2 represented the purified seed coat proteins from S8 LCC4 # 16-1 probed with anti-LCC4 antibody. To compare HIS tag purified band patterns (lane 2) with that of total cv. Jack seed coat proteins, we ran cv. Jack seed coat protein as a control (lane 3). Note that in the protein ladder lane (lane 1, Figure 3.15 A) there was a band of 64.7 kDa detected by anti-LCC4 antibody that was not present in the protein ladder (lane 0). Nor was this band present in the blots probed with anti-HRPC antibody but was present in some blots previously probed with the anti LCC4 antibody (Figure 3.10 A, lane 1) suggesting that a protein contaminant not visible in the stained protein marker reacts with anti-LCC4 antibody. The HIS tagged LCC4 protein expressed in *E coli* was detected with anti-LCC4 antibody (Figure 3.15 A, lane 5) as a strong band with an estimated size of 59.1 kDa (59.7 kDa in replicate blot). Its size as determined on the blot with anti-HIS antibody was 58.2 kDa (59.1 kDa in a replicate blot, lane 5 of Figure. 3.15 B and C). These values were slightly larger than the calculated

size of 54.4 kDa (509 aa). This result demonstrated that these two antibodies could detect the HIS tagged LCC4 protein produced in *E coli*.

The anti-LCC4 antibody detected 2 bands with estimated molecular weights of 77.6 kDa, and 64.7 kDa (lane 2 of Figure 3.15 A, labelled with solid black dots) in the seed coat from the transgenic plant. The sizes of these two bands on replicate gel run with the same protein extract were 75.5 and 66.3 kDa respectively (data not shown). Probing of the replicate transfer in Figure 3.15 B with His-tag antibody revealed one faint band, as indicated by the arrow, with an estimated molecular weight of 77.6 kDa (75.5 kDa in Figure 3.15 C). This value was very close to the size of the upper band (77.6 kDa) detected with the anti-LCC4 antibody as indicated by the arrow in band in lane 2 of Figure 3.15 A). A band with a similar size was seen in the untransformed control probed with anti-LCC4 antibody (lane 3, Figure 3.15 A). The origin of this band was not known but a band of this size was not seen in the same control probed with anti-HIS tag antibody (lane 3, Figure 3.15 B).Therefore both antibodies detect a protein of about the same size that contains epitopes for LCC4 and HIS, suggesting that it may be the expressed LCC4.

The size of the full length LCC4 protein from *T. vesicolor* calculated from the DNA sequence is 56.1 kDa (527 aa) (Ong et al., 1997). With the removal of the NTPP of 27 aa (Dr. Theresa White, Iogen, Ottawa, personal communications) the expected final protein size would be 53.2 kDa (500 aa). Laccase has no predicted CTPP and is expected to be secreted (Jönsson et al., 1997; Ong et al., 1997; Guo et al., 2005). Commercial laccase (Biochemika) has a size of 64 ~ 66 kDa (Hood et al., 2003) or 64.5 kDa determined here (lane 4, Figure 3.15 A) although the band is diffuse. Similar size

differences have been previously observed and may be due in part to the contribution of glycosylation (Sardana et al., 2007).

Based upon our S8 LCC4 soybean construct (Appendix 1), and assuming that the NTPP and CTPP of SBP were removed, the mature LCC4 (501 aa), including GGAA adaptors (8 aa) and 6XHIS tag (6 aa) has a predicted size of 54.6 kDa (515 aa) before glycosylation is taken into consideration. The size of putative expressed LCC4 protein as measured by SDS gel was 77.6 kDa suggesting that glycosylation also is playing a role in the seed coat system as has been observed in other transgenic systems, e.g. LCC4 from *T. versicolor* expressed in *Pichia pastoris* had a size of around 85 kDa as determined by SDS gel and contained ~ 40% carbohydrate (Brown et al., 2002); when LCC1 of *Trametes versicolor* was expressed in maize seeds, the measured size was 63 to 65 kDa compared to the predicted size of 53.0 kDa (Hood et al., 2003). There are 13 potential N-linked glycosylation sites in the LCC4 mature protein sequence (Ong et al., 1997) and the molecular mass of LCC4 expressed in soybean might depend on the level of glycosylation.

In summary, we found that LCC4 protein was expressed in soybean seed coat with a size of 77.6 kDa, suggesting that the LCC4 gene fused to the Ep 1.5 kb promoter was successfully expressed in the seed coat.

Figure 3.15 A LCC4 Proteins Detected with Anti-LCC4 Antibody

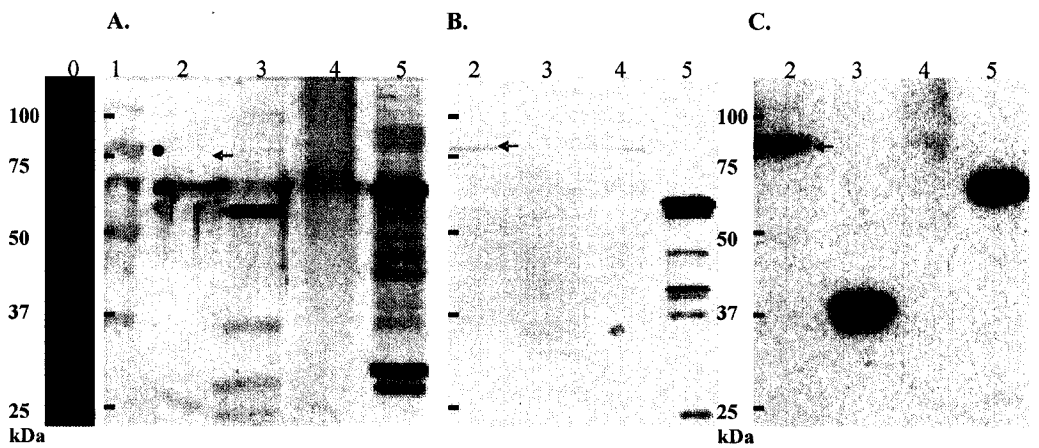
Western blot was probed with anti-LCC4 antibody at 1000X, and then it was detected with goat anti-rabbit IgG conjugated with AP at 3000X. NBT/BCIP Reagent Kit (Molecular Probes) was used to detect the protein signals. The overlay time was overnight. Lane 0 is a BioRad protein ladder lane from another SDS gel that is included for comparison. It was not probed with anti LCC4 antibody. Lane 1, Protein Ladder (5 µl), (BioRad # 161-0373); Lane 2, Purified LCC4 from S8 LCC4 16-1-1 seed coat (70 ng); Lane 3, Jack seed coat protein (25 µg); Lane 4, Laccase (Biochemika) (500 ng); Lane 5, pTrc 2A LCC4 *E coli* extract (6 µg);

Figure 3.15 B LCC4 Proteins Detected with Anti-HIS Tag Antibody (colorimetric)

Western blot was probed with anti-HIS tag antibody at 1000X, and then it was detected with goat anti-mouse IgG conjugated with AP at 3000X. NBT/BCIP Reagent Kit (Molecular Probes) was used to detect the protein signals. The lane setting was the same as Figure 3.15 A.

Figure 3.15 C LCC4 Proteins Detected with Anti-HIS Tag Antibody (chemiluminescent)

Western blot was probed with anti-HIS tag antibody at 1000X, and then it was detected with goat anti-mouse IgG conjugated with HRP (BIO-RAD, # 170-6516) at 30,000X for chemiluminescent detection with ECL plus Western blotting system (Amersham, # RPN2132).Lane 1, Protein Ladder (5 µl); (BioRad # 161-0373); Lane 2, Purified LCC4 from S8 LCC4 16-1-1 seed coat (100 ng); Lane 3, LCC1 Transgenic maize seed total proteins, (50 µg); Lane 4, Laccase (Biochemika), (1 µg); Lane 5, pTrc 2A LCC4 *E coli* extract (6 µg);



CHAPTER FOUR: DISCUSSION

4.1 Overall Conclusions:

In this project, we tested the functionality of the SBP signalling peptides (NTPP and CTPP) by GFP targeting in *Arabidopsis* mesophyll cells. The results showed that NTPP-GFP was targeted to the apoplast while GFP- CTPP was found in the cytosol.

Two different vectors S8 and S18 (Figure 3.4), were used to introduce HPRC, LCC1 and LCC4 transgenes into soybean by particle bombardment. Seed coats from T_1 seeds were screened for peroxidase or laccase enzyme activities from 14 S8 HRPC transgenic plants, 5 S8 LCC1 transgenic plants, 36 S8 LCC4 transgenic plants, 11 S18 LCC1 transgenic plants and 4 S18 LCC4 transgenic plants. Activities up to 3.09 U/mg for HRPC transgenics and up to 0.099 U/mg for laccase transgenics were measured, suggesting that the transgenes were successfully expressed in soybean seed coat. No differences between vectors expressing the same transgene (S8 vs S18) were observed.

Transgenic plants expressing high level enzyme activity were propagated to the T_2 and T_3 generations and characterized both by PCR to monitor transgene stability and inheritance, and enzyme analysis. The loss of the full-length gene was found at 40% (Table 3.8) in T_1 generation among five independent lines tested. Monitoring the presence of transgene by PCR is necessary until the homozygous plants are obtained.

Western blot analysis was performed with protein purified from seed coats of one HPRC and one LCC4 transgenic using commercial anitsera (anti-HPRC or anti-HIS tag antibodies) or an antiserum we produced (anti-LCC4 antibodies). A putative HRPC with a size of 35.4 kDa and a putative LCC4 with a size of 77.6 kDa were identified.

In summary, these results suggest that the soybean seed coat can be used to express useful foreign proteins.

4.2 Protein Targeting

Transgenic plants are attractive vehicles for producing recombinant proteins. Plant cells compartmentalize and store metabolites and proteins in vacuoles, and they often need to be targeted to the correct compartments to accumulate in a stable fashion (Jiang and Sun, 2002). Although initial unpublished experiments by Drs. M. Gijzen and D. Simmonds (AAFC) had shown that the Ep 1.5kb promoter directs HGC-specific expression of the β–Glucuronidase (GUS) reporter gene, experiments to verify the ability of the signal peptides to target the endogenous protein into vacuoles had not been done prior to this thesis.

In our project, we characterized the functionality of the SBP signalling peptides and found that the NTPP targeted GFP to the apoplast; whereas, the CTPP retained GFP in the cytoplasm (Figure 3.13). Unfortunately, we did not find GFP localized in protein storage vacuoles when GFP was fused with both NTPP and CTPP signals (Figure 3.13). The acidic environment of the central vacuole of *Arabidopsis* leaf could lead to GFP degradation (oral communication, Dr. H. Zheng, McGill University) (Gallie and Kado, 1989). We tried two approaches reported to overcome this problem. We either incorporated FM-64, a protease inhibitor, into our protocol to prevent GFP degradation or we incubated *Arabidopsis* overnight (Tamura et al., 2003). Neither approach produced a sharp image that could confirm GFP localized in the vacuole. A possible explanation might be that the pathway targeting GFP to the vacuole was saturated and excess of GFP spilled over to the apoplast by a default secretion pathway. Another reason for not seeing

GFP signals in vacuoles could be that other unidentified vacuolar targeting sequences in SBP DNA sequences might be needed to complement NTPP and CTPP. A precedent for this possibility was the finding that multiple vacuolar sorting determinants exist in soybean 11S globulin (Tamura et al., 2003).

The GFP targeting in *Arabidopsis* mesophyll cells was good as a model but it might not reflect the situation in soybean hourglass cells. It has been shown that GFP intensity could vary in the vacuoles of different plants (Maruyama et al., 2006). Protein targeting was believed to be plant tissue independent. This concept was challenged by the findings that *Aspergillus niger* phytase was efficiently secreted in rice leaf cells, however, other studies revealed that within endosperm cells it was retained in protein storage vacuoles (Di Sansebastiano et al., 2007). Therefore, we need to more fully identify and characterize the functionality of the SBP signalling peptides in the hourglass cells of soybean seed coat before making final conclusions on the value of targeting strategies for enzyme production in hourglass cells.

We assumed that the HRPC and laccase were targeted to protein storage vacuoles in hourglass cells but protein storage vacuoles have not been identified so far (Z. Jin, 2008, unpublished data). Based on the findings in *Arabidopsis*, expressed HRPC and laccase could be secreted to the apoplast and that location may also be appropriate for foreign protein accumulation (Drakakaki et al., 2006). This is consistent with other studies in which LCC1 targeting to the maize seed cell wall was achieved with high expression levels (Hood et al., 2003; Hood et al., 2007; Streatfield, 2007). However targeting enzymes such as laccases to the apoplast may lead to problems for the plant (Hood et al., 2003).

Recently, it was found that the ER retention signals could increase foreign protein yield because there was less protease activity in ER to degrade the expressed proteins. Combining an apoplast targeting sequence with an endoplasmic reticulum retention sequence could result in the foreign protein being sequestered in the ER (Hood et al., 2003). However, there are still concerns with ER retained proteins. For instance, the space in the ER might be a limiting factor for highly expressed foreign proteins since dilation and proliferation of the ER network were observed when overexpressing the hepatitis B surface antigen in the ER in soybean cell culture (Streatfield et al., 2003). Other potential drawbacks for ER-retained proteins is that the expressed proteins may not have the proper post-translational modifications since the expressed proteins may be further glycosylated in Golgi bodies and the addition of the C-terminal signal needed for retention. These may affect activities of certain proteins.

4.3 Transgene Expression and Stability

Predictable and high level transgene expression from generation to generation is necessary. We found that the HRPC transgenes detected in T_0 plant genomes were frequently lost in the T_1 and T_2 generations. For example, the Ep promoter-HRPC coding region fusion was present in the T_0 plant genomes of S8 HRPC # 3-2, S8 HRPC # 3-4 and S8 HRPC # 3-6, which were 3 independent lines, but absent in T_1 generation of plants such as S8 HRPC # 3-2-1, S8 HRPC # 3-2-2, S8 HRPC # 3-4-1, S8 HRPC # 3-4-2, S8 HRPC # 3-4-3, S8 HRPC # 3-6-1, S8 HRPC # 3-6-2, and S8 HRPC # 3-6-3. The percentage of Ep promoter loss was 60% (3/5). In addition deletions of the HRPC coding region were found in 40% (2/5) among 5 independent HRPC lines tested (Table 3.8). The

presence of the hygromycin gene in the T_0 generation was confirmed by PCR (D. Brown, AAFC, London); however, by the T_1 generation, hygromycin was already lost in some plants.

Transgene instability was also monitored for the laccase plants. We used PCR to screen T_1 plants for the presence of the transgene and the hygromycin resistance gene. Ten out of 36 S8 LCC4 lines, 2 out of 5 S8 LCC1 lines and 1 out of 11 S18 LCC1 lines were transgene negative. The inheritance of laccase transgene from T_0 to T_1 generations was not monitored as no T_0 laccase transgenic gDNA was available although during the regeneration of transformed plants the presence of the hygromycin gene was confirmed.

Particle bombardment transformation can lead to transgene rearrangements (Finer and McMullen, 1991; Patnaik and Khurana, 2001; Reddy et al., 2003; Latham et al., 2006; Zhao et al., 2007) Since the same transgene components are found in the S8 and S18 vectors used for soybean transformation (except the synthetic transgene), we would expect that similar results for S18 transgenic plants. Transgene loss was also found in PHB transgenics (J. Schnell, unpublished) and SBP transgenics (J. Schnell and D. Johnson, unpublished) in our research group. This was not unexpected as the literature showed that the recombination events in bovine β-casein trangenic soybean by particle bombardment resulted in the trangene loss (Choffnes et al., 2001).

In our project, particle bombardment was used to transform soybean (cv. Jack) because of the inefficiency of *Agrobacterium* mediated transformation of soybean (Finer and McMullen, 1991; Mello-Farias and Chaves, 2008). Furthermore, the transformation efficiency achieved by our colleagues using particle bombardment was fairly high (Dr. S Zhou, unpublished data). Unlike *Agrobacterium*-mediated transformation, where the

integration of the introduced DNA into the plant chromosomal DNA is defined by the T-DNA borders, particle bombardment results in the introduction of random pieces of the vector DNA. The potential for false positives in the T_0 generation can not be ruled out because the template DNA might be located somewhere other than the nucleus within cell and might not be integrated into the plant chromosomal DNA. Extra-chromosomal maintenance of introduced DNA has been observed (Praitis, 2006). Similar phenomena are also found in *Agrobacterium*-mediated transgenic plants (Frary and Earle, 1996; Tzfira et al., 1997; Stewart JR, 2008). The integration of HRPC transgenes into chromosomal DNA has to be confirmed by Southern blot hybridization analysis. Our strategy was to first identify plants with high peroxidase activity in the seed coat. Due to the large numbers of plants screened it was more efficient to do the Southern blotting analysis or gene mapping for only promising HRPC transgenic plants after the initial screen. Molecular characterization of T_3 and subsequent generation by Southern blot is one of our future directions, but has been restricted due to the long generation times for soybean.

Transgene rearrangements are common in plants generated by particle bombardment. The introduced DNA can be integrated at single sites or multiple sites on the chromosomal DNA, as partial copies or multiple copies in varying orientations. The inversely-repeated transgenes could be excised by recombination (Assaad and Signer, 1992). In addition, the introduced DNA may be mixed or interspersed with plant genomic DNA (Pawlowski and Somers, 1998). Transgene elimination has been demonstrated in transgenic dry bean and soybean, in which DNA inserted into the rDNA region, was eliminated by the disruption of rDNA unit (Romano et al., 2005). Complete loss of

transgenes was also reported in transgenic guar (Joersbo et al., 1999). Transgene rearrangement has been documented in commercial soybean, e.g., when Roundup Ready soybean (event 40-3-2) was analysed, a 254 bp EPSPS(enoylpyruvate shikimate synthase) gene fragment was found in addition to the EPSPS transgene described in the original application (Windels et al., 2001).

Our results appear to parallel findings that have been widely reported and indicate that further stages of selection are needed to generate lines with stable and heritable expression of our genes. In the future, we should monitor transgene stability up to at least the T_3 or T_4 generations until we obtain stable, homozygous transgenic lines (Verweire et al., 2007) as confirmed by the presence of the transgene by PCR and stable enzyme activity. In our case, it might take one more generation to get the homozygous transgenic soybean plant since seed coat is maternal tissue and transgene segregation is one generation later than that in embryo. Then Southern analysis can be done, not only to characterize the transgene copy number but also prove that the transgene is integrated into the plant genome. If DNA analysis confirms the presence of the transgene but the enzyme activity is too low to be detected, it suggests that transgene silencing might have occurred. Monitoring the presence of transgene from generation to generation and enzyme activity are keys to guarantee the transgene stability and inheritability.

The variation in peroxidase activity between lines might be associated with transgene position effects and/or silencing in addition to instability. The HRPC transgene, if inserted into in more active regions of soybean genomic DNA, would be likely expressed at higher levels, while integration in less active areas could lead to lower expression levels as has been observed in tobacco (Peach and Velten, 1991; Mlynarova et

al., 1994) and many other species, such as maize (Singh et al., 2008) and rice (Mlynarova et al., 1994) Homology-dependent silencing is possible since the introduced Ep 1.5 kb promoter is the same as the endogenous one. It was found that homologous promoters could cause gene silencing by methylation (Matzke et al., 1993; Jakowitsch et al., 1999; De Jaeger et al., 2002). Multiple copies of the transgene is fairly common in transgenic plant transformed with particle bombardment and this could also lead to expression silencing (Travella et al., 2005).

Seed to seed variation can result from transgene segregation and eventually cause transgene elimination (Romano et al., 2005). It was found that the expression of *gal*4 (transcriptional factor) fused with GFP showed the variation of GFP expression from seed to seed that were from the same parental rice plants in T_1 generation (J.M. Hibberd WEB site, accessed at http://www.plantsci.cam.ac.uk/Hibberd/pdf/alex1.pdf). The GUS expression variation from seed to seed that were from the same parental plant in T_1 transgenic wheat (*pina::uidA* lines) generation was also observed (Wiley et al., 2008). In our project, we have found there were variations of enzyme activity from seed to seed in T_1 generation; this is might be due to transgene segregation or transgene elimination triggered by a process of genome defence.

Screening every single seed by enzyme assay in T_1 and T_2 generations is necessary to obtain the promising transgenic plants. The T_3 generation was important because the seed coat enzyme activity varied more widely in T_1 and T_2 generation, and variation of seed coat enzyme activity in T_3 generation was found to narrower. This phenomenon was true both for HRPC (Table 3.4, S8 # HRPC 3-8) and laccase

transgenics (Table 3.10), indicating transgene segregation occurred in T_1 and T_2 generations, and homozygous plants were more likely to be obtained in the T_3 generation.

Based upon the literature (Verweire et al., 2007) transgene homozygozity may be reached in the T_3 generation (T_2 generation seed coat). Due to time constraints we could not advance all our lines to the T_3 generation but preliminary experiments in a few lines suggest that seed coat peroxidase activity was stably inherited. For instance, S8 HRPC # 3-11 in Table 3.4, yielded average peroxidase activity in T_1 seed at 0.33 ± 0.13 U/mg (N = 4), in T_2 seeds at 1.01 ± 0.25 U/mg (N=7), and in T_3 seeds was 0.82 ± 0.11 U/mg (N = 6) although clearly the number of plants (seed coats) analysed was small. We found that the peroxidase activity in T_3 seeds was significantly higher than that in T_1 seeds ($P < 0.05$). We also monitored the changes of laccase activities from T_1 to T_3 generation (Table 3.10). The mean values of laccase activity in T_2 and T_3 generations were not significantly different from that in T_1 generation. In both case the removal of plants that have segregated out the transgene will have affected these numbers. Our results showed that high laccase activity could be stably inherited.

It appears that in some of our plants we have peroxidase activity in the absence of the transgene as determined by PCR (Table 3.5). Possible explanations are i) During transformation the vector was fragmented and truncated HRPC DNA was inserted downstream of an unknown promoter. In this scenario, transgene expression might not be seed coat specific. This model predicts that truncation has led to loss of primer sites for HPRC PCR, therefore, we can not expect to get PCR products of HRPC transgene since the 5' end of HRPC transgene was gone. And any expressed protein may not react with anti HIS tag antibody but should react with the anti-HPRC antibody. HPRC sequences

should be detectable with primers that target the region of the HPRC sequence needed for activity. ii) Variation induced by tissue culture may have also activated other peroxidase genes (Kaeppler et al., 2000; Morcillo et al., 2006). The expression of peroxidase genes could be induced by stress conditions in which plant tissues were immersed in liquid culture with selection pressures of hygromycin and this unregulated expression persisted throughout generations. This model predicts that no HPRC is produced and thus is not detectable. iii) The truncated Ep promoter (< 300 bp) might have activated other peroxidases by insertion upstream of another peroxidase gene. There are hundreds of peroxidases in soybean genome (Gijzen et al., 1993; Santos et al., 2008). This model does not make strong predictions about the distribution of peroxidase expression since we do not know the specificity of the truncated promoter. It is unlikely that anti HRP antibody and anti HIS antibody will react with this peroxidase since we do not know which peroxidases are activated. iv) tCUP promoter fragments could also activate other peroxidase gene expression since tCUP promoter is one of components in our soybean transformation constructs. The predictions are similar to those in iii.

4.4 The Expressed HRPC Proteins in Soybean Seed Coat

Horseradish Peroxidase (HRPC) belongs to the class III plant peroxidase super-family. It is one of the most widely-used industrial enzymes (Ghasempur et al., 2007) in applications ranging from paper bleaching to medical diagnostics (Yoshida et al., 2003).

Enzymatic oxidation is a critical technology for industrial biotechnology. It offers significant improvements over traditional chemical processes when produced in plants by minimizing the use of toxic chemicals, lowering capital costs, producing cleaner products, and increasing cost efficiencies (Clough et al., 2006). We are the first research group to

express the active HRPC in soybean seed coats. To achieve this, the NTPP and CTPP of HRPC were replaced with sequences from SBP and codon usage was optimized for soybean. Furthermore, the expression of the HRPC synthetic gene was driven by the Ep 1.5 kb promoter. The HRPC gene was therefore strategically designed to be expressed in the soybean seed coat hourglass cells.

Following particle-bombardment transformation, 14 independent HRPC transgenic soybean lines were obtained. We screened T_1 HRPC transgenic seeds for peroxidase enzyme activity and the HRPC transgenic seeds with high peroxidase enzyme activities were advanced into the T_2 generation. Among of 192 T_1 and T_2 seeds screened, we found that the peroxidase activities in HRPC transgenic seed coat could be up to 60 times higher than that of cv. Jack. Following screening by PCR and HRPC enzyme activities, we chose several candidates for further analysis. At the time of writing, only S8 HPRC # 16-1 with an activity of 0.45 U/mg, had provided us with enough material for Western analysis.

In our transgenics, HRPC activity averaged more than 10 times that found in cv. Jack with the highest level being 3.09 U/mg found in S8 HRPC # 3-2-1. This value is comparable to activities found in root extracts from *Armoracia rusticana* (Table 1.2) with values of 3.43, 147.89, and 6.46 to 11.58 having been reported. HPRC has been expressed in other heterologous plant systems (see Table 1.2), such as tobacco BY2 cells (Matsui et al., 2006) at a level of 0.04-0.50 ABTS U/mg which is lower than that expressed in the seed coat. Unfortunately literature values for regenerated tobacco transgenics cannot be expressed as units per mg protein. HRPC has also been expressed in transgenic aspen with peroxidase activities up to 4,393 ABTS units/mg have been

reported in leaf, a value 10 times higher than WT plants (Kawaoka et al., 2003) and much higher that achieved to date in the seed coat. Thus overall we can conclude that the level of expression is better than has been reported for tobacco in culture and comparable to root extracts from *Armoracia rusticana*. This has been achieved during our first attempt to express a foreign enzyme in the soybean seed coat.

To provide more evidence that HRPC was being expressed in seed coats we turned to Western blotting. HRPC was purified from S8 HRPC # 16-1 (T_2) based on the presence of the HIS-tag added to the carboxy terminal end. A presumptive HPRC protein of 35.4 kDa (Figure 3.14) was detected using both anti HRPC, and anti-HIS tag antibodies. This protein was not found in embryo tissue from the transgenic nor in cv. Jack and has a size similar to the 35.2 kDa calculated for the HPRC transgene with 322 aa. The Western blot data provides additional support for the production of HRPC in the soybean seed coat.

HRPC has been expressed in heterogolous plant systems. Mature HRPC expressed in tobacco plants was reported to have sizes of 45 kDa (Pellegrineschi et al., 1995; Kis et al., 2004) and 38 kDa (Matsui, 2006). It is also known that signalling peptides did not affect the sizes of the expressed HRPC, which indicated that NTPP and CTPP were properly removed when HRPC proteins matured (Kis et al., 2004). The size of commercial HRPC is estimated to be between 43.8 kDa (Regalado et al., 1996) and 50 kDa (Loustau et al., 2008). The levels of heterogeneity of glycans at each N-glycosylation site of horseradish peroxidase might contribute to these differences (Gray et al., 1998). Since the size of expressed HRPC (35.4 kDa) was fairly close to the calculated size at 35.2 kDa, glycosylation might be different in the seed coat system. For

example, the expressed proteins might be targeted to the apoplast after being synthesized in ER (Figure 1.2) without further modification in Golgi bodies.

It was reported that SBP has a 20 fold higher catalytic efficiency than HRPC at their respective pH optima using ABTS as a substrate at 25°C (Kamal and Behere, 2003). It may follow that a 20 fold higher HRPC expression would be needed if the HRPC and SBP were to achieve the same level of peroxidase activity. HRPC expression in transgenic soybean seed coat must be robust to achieve higher peroxidase activity. An activity of 0.82 U/mg in S8 HRPC # 3-11-1-4, 16 times higher than in cv. Jack, may be very significant. For this system to be commercially viable, higher levels of expression may be necessary. The expression level of HRPC was roughly estimated as follows: eighty seed coats yielded 15 ml of extract at 500 μg/ml total proteins, or a total of 7,500 μg. HIS tag column chromatography yielded 1 ml at effluent as 12 μg /ml purified proteins, for a total of 12 μg. HRPC protein (%) expressed in TSP in seed coat protein was at (12/7500) x 100% = 0.16 %. Although this level is far below the economically accepted level of 1% (Daniell et al., 2001), it is still within an acceptable range and may be enhanced by crossing our elite lines and selecting for expression level. We could reach 1.6% if the expressed protein yield could be enhanced by 10 times. Given the variation that we have witnessed we believe that a 10-fold increase is possible.

It was reported that overexpression of the potato *Stprx2* gene conferred increased salt-tolerance (Collinge and Boller, 2001), and that of the sweet potato *swpa4* gene enhanced oxidative-stress tolerance (Kim et al., 2008). Interestingly, a growth stimulation was observed in transgenic tobacco and hybrid aspen upon overexpression of horseradish peroxidase (Yoshida et al., 2003). So far, we have not found any negative pleiotropic

effects of the expressed HRPC on soybean growth. Hopefully positive pleiotropic effects of the expressed HRPC on soybean growth, such as the enhanced resistance to plant diseases, and the growth stimulation will be a possibility in the future research. In conclusion, the biologically-active HRPC proteins were successfully produced in the soybean seed coats, producing the first evidence in support of the soybean seed coat as a bioreactor for the production of novel proteins.

4.5 The Expressed Laccase Proteins in Soybean Seed Coat

Laccases catalyze oxidation of various phenolic compounds, aromatic amines and even certain inorganic compounds by using molecular oxygen as the electron acceptor (Thurston, 1994). Substrate versatility makes laccases highly interesting for various applications, including textile dye bleaching, pulp bleaching and bioremediation (Claus et al., 2002; Claus, 2003, 2004), where enzymatic catalysis could serve as a more environmentally benign alternative to the currently used chemical processes (Kiiskinen, 2004). In this project, we have expressed soy-optimized fungal laccase, LCC1 and LCC4 in the soybean seed coat.

To produce laccase antibodies in rabbits, we first expressed LCC1 and LCC4 proteins in Top 10 *E coli* cells using the Ptrc promoter and induction with IPTG. The HIS tag purified laccase proteins were used to inject rabbits for laccase antibody production. The results showed that anti LCC4 antibody was fairly robust and could be used at a 1000X dilution while anti LCC1 antibody was of poorer quality. Fortunately, the anti LCC4 antibody could cross-react with LCC1. Testing of the anti LCC4 antibody demonstrated that it could detect laccase protein expressed in *E coli*, *Trametes versicolor*, and LCC1 transgenic maize seeds.

We also attempted to express laccases in *Arabidopsis* using by CaMV 35S promoter, but high laccase expressors were not found. Some of the transgenic *Arabidopsis* showed growth retardation, yellow leaves, and seed sterility. The strength of CaMV 35S promoter might have generated toxic levels of laccase. It is known that leaky expression of laccase in maize caused maize seed colour to change from yellow to brown, which was an indication of lignification (Hood et al., 2003).

In this project, we expressed soy-optimized fungal laccases (LCC1, and LCC4) in the soybean seed coat. The cDNA portion corresponding to mature laccase was fused with SBP signalling peptides and driven by Ep 1.5 kb promoter to target the expressed laccase proteins to hourglass cells of the soybean seed coat. In total, we obtained 5 lines of S8 LCC1 transgenic soybean and 36 lines of S8 LCC4 transgenic soybean. We also had T_0 plants shipped from London, including 11 lines of S18 LCC1 transgenic soybeans and 4 lines of S18 LCC4 transgenic soybeans. To get more transgenic soybean seeds to work with we advanced some of the T_1 seeds of S8 LCC1 and S8 LCC4 transgenic soybean seeds to the T_2 generation.

Seed coats were assayed for laccase activity using ABTS as the substrate. Plants with higher laccase activities in the seed coat were advanced to the T_2 generation. In total, 404 laccase T_1 and T_2 transgenic soybean seeds were assayed. Generally the laccase activity in S8/S18 LCC1, and S8/S18 LCC4 transgenic soybean seeds (Table 3.7) could be up to 9 fold higher than that of Jack. Our S8 LCC4 16 line yielded 0.099 U/mg laccase activity in the T_2 generation, which was comparable to LCC1 transgenic maize seeds in our hands (Hood et al., 2003).

The Western blot data provides additional support for the production of LCC4 in the soybean seed coat. We found that LCC4 protein was expressed in soybean seed coat, with a size of 77.6 kDa (Figure 3.15) as estimated from SDS gel electrophoresis. We repeated this experiment twice and got the same result. The predicted size based upon the amino acid sequence (515aa) was 54.6 kDa before glycosylation is taken into consideration. As suggested before the difference between the predicted and measured sizes suggest that glycosylation might be playing a role in the seed coat system as has been observed in other transgenic systems. For instance, LCC4 from *T. versicolor* expressed in *Pichia pastoris* had a size of around 85 kDa as determined by SDS gel and contained ~ 40% carbohydrate (Brown et al., 2002); when LCC1 of *Trametes versicolor* was expressed in maize seeds, the measured size was 63 to 65 kDa compared to the predicted size of 53.0 kDa (Hood et al., 2003). There are 13 potential N-linked glycosylation sites in the LCC4 mature protein sequence (Ong et al., 1997) and the molecular mass of LCC4 expressed in soybean might depend on the level of glycosylation. This explanation does not preclude the possibility that mis-targeting may have lead to a protein containing the targeting peptides, although their retention would only add a small amount to the size (retention of the CTPP would only add 20 aa equivalent to 2.1 kDa).

Laccase, as a ligninolytic enzyme, has been difficult to over-express in active form in heterologous hosts. Detection of active recombinant laccase has been reported in the yeasts *Saccharomyces cerevisiae* (Kojima et al., 1990), *Pichia pastoris* (Jönsson et al., 1997), in the filamentous fungi *Trichoderma reesei* (Kiiskinen et al., 2004), *Aspergillus oryzae* (Yaver et al., 1996).

Laccase also has been expressed in plants (Hood et al., 2003; Sonoki et al., 2005). In maize, the embryo-preferred promoter was employed to target laccase to maize seed. Unfortunately, laccase gene expression was not restricted to seed, and hindered plant growth. Cotton laccase was expressed in *Arabidopsis* at 0.005-0.008 U ABTS/ mg protein, which was much lower than ours. *Pycnoporus cinnabarinus* laccase was expressed in rice endosperm at 0.14 U ABTS /mg and *Melanocarpus albomyces* laccase activity could be as high as 0.58 U ABTS /mg proteins in rice seed endosperm (de Wilde et al., 2008). Hood et al. (2003) expressed LCC1 (*Trametes versicolor*) in maize seed. We measured their laccase activity in transgenic maize at 0.09 U ABTS/mg with the gift seeds we obtained from them. The laccase level of our best plant seeds (S8 LCC4 # 16-1) was the same as that of LCC1 transgenic maize seeds (Hood et al., 2003). In conclusion, the laccase expression level (0.027 to 0.099 U/mg) in our laccase transgenic soybean seed coat was within the range of published data and this level has been achieved during our first attempt.

It is unknown why the expression of laccases is low in most heterologous systems but there are many hypothetical reasons. Some laccases have very narrow expression patterns, for instance, LCC4 is not expressed in normal culture condition (Brown et al., 2002). Laccases usually contain a CpG-dinucleotide motif and it has been suggested that decreasing the motif content could increase the yield of expressed laccase (Hirai et al., 2008). It is known that plants tend to sequester the copper for essential functions during limited copper supply (Abdel-Ghany and Pilon, 2008), which might limit access for copper-dependent laccase synthesis. High laccase activity may interfere with the growth and development of plants.

In plants, laccases are not only involved in lignin synthesis (Ranocha et al., 2002), but also in plant defence such as the wounding response (Alcalde, 2007). For instance, cucumber multicopper oxidases were found to be abundant in cucumber fruit tissue against plant diseases (Ohkawa et al., 1989). Overexpression of polyphenol oxidase in transgenic tomato plants resulted in enhanced bacterial disease resistance (Li and Steffens, 2002). In an attempt to minimize negative pleiotropic effects, we employed the seed coat specific promoter *Ep* to restrict the expression of laccase to the seed coat. It is possible that it may also contribute to stress resistance; however, this was not examined. An apparent contradiction is that plant laccase promotes lignin synthesis, whereas fungal laccase depolymerises lignin. It is unknown how laccases control opposite functions in plant and fungi.

Although recombinant laccase was expressed at relatively low levels, it is still a good source for industrial enzymes production. Compared with HRPC, laccases require oxygen as an oxidant, which is much less expensive than the hydrogen peroxide required by peroxidase enzymes (Widsten and Kandelbauer, 2008). Moreover, laccase does not seem to be affected by wastewater constituents whereas HRPC was significantly affected by wastewater constituents (Auriol et al., 2008). From the perspective of cost effectiveness, laccase may present important advantages over HRP for industrial applications. However this cost advantage may disappear when the cost of ABTS, added along with laccase as being a reaction mediator (Bourbonnais et al., 1997) to increase oxidation efficiency via its redox potential (Johannes and Majcherczyk, 2000) is considered.

In summary, we successfully expressed biologically active laccases in soybean seed coat in this project, which proved that soybean seed coat could be good bioreactor to produce the novel proteins.

4.6 Future Directions of Research

4.6.1 Improvement in Transformation Technology is Essential

The development of methodologies for the delivery of genes into intact plant tissues by particle bombardment has revolutionized the field of plant transformation. The concept of accelerating DNA-coated particles into cells and tissues has evolved from novelty to an established tool in plant molecular biology (Klein et al., 1988), for instance, soybean transformation (Klein et al., 1992). Low copy number gene transfer and stable expression were both accomplished in a commercial wheat cultivar via particle bombardment (Yao et al., 2006; Sailaja et al., 2008). Recently, it was also found that glass beads can replace gold particle to achieve higher transformation efficiency (Feng et al., 2008).

In our project, full length vector DNA was used to deliver the HRPC and laccase transgenes by particle bombardment. One drawback is the integration of the vector backbone sequences into the host genome along with the transgenes. Backbone sequences often have negative effects on transgene expression, silencing being one of them and it could also lead to transgene rearrangements. Studies showed that extensive stretches of prokaryotic vector sequences are not well tolerated by higher eukaryotic genomes (Matzke and Matzke, 1998). When integrated into plant DNA, excess vector sequences are often densely methylated and this could spread into the neighbouring transgenes

(Jakowitsch et al., 1999). It is possible that unusual sequence composition, such as high GC content, could make long stretches of vector sequences conspicuous to plant methyltransferases (Astua-Monge et al., 2002). Particle bombardment using only transgene expression cassettes may therefore improve the stability of transgene expression.

Agrobacterium-mediated transformation is an alternative to direct gene delivery methods. One of advantages of the *Agrobacterium* system is the insertion of a discrete segment of DNA into the recipient genome; the T-DNA is delivered by right and left border sequences. *Agrobacterium* -mediated soybean transformation is less efficient than particle bombardment but it is improving (Wang and Xu, 2008; Yukawa et al., 2008). It would be helpful to improve *Agrobacterium*-mediated transformation in the recalcitrant soybean genotypes. It may eventually generate plants with more predictable insertion patterns and therefore more stable expression patterns. At this time it is inefficient and would not have yielded the large numbers of transgenic lines needed in this study.

Current transformation techniques are a significant limitation in the development of complex biotechnologies for soybean because of the inefficiencies and difficulties in recovering useful material for study. Improvements are essential for future research.

4.6.2 Traditional Breeding Techniques May Help to Enhance Accumulation of Foreign Proteins

Soybean seeds, as a bioreactor, offer the opportunity for producing bulk quantities of proteins. The risk of contamination by pollen is reduced since soybean is largely self-pollinating (Robic et al., 2006). Self-pollination to produce homozygous plants and

further improvement of agronomic characteristics through breeding may increase the levels of enzymes in future generations (Hood et al., 2007). Stable inheritance of the transgene, consistent expression patterns and competitive agronomic properties of transgenic crops are important parameters for transgenic crops. It was found that higher production levels of recombinant β-glucanase would be identified in homozygous transgenic T_3 barley plants and these remained constant over a 3-year period. These homozygous plants had improved 1000-grain weight values with respect to those of the original transformants (Horvath et al., 2001). It was also reported that expression levels of cellulase could be increased 100-fold from the T_1 generation through breeding and selection. It was expected to reach cost-effective production levels in a few generations (Hood et al., 2007). Breeding with elite transgenic lines could produce good transgenic materials with improved seed yields, field performance and most importantly, increased accumulation of the target protein (Streatfield et al., 2001; Hood et al., 2003). Tissue culture and transformation processes may produce undesirable field characteristics in transgenic plant (Gilbert et al., 2005), therefore, field evaluation is necessary for selecting genetically stable and agronomically acceptable material for commercial use.

For our higher HRPC or laccase lines we should advance them to T_3 or T_4 generations to get homozygous plants with stable expression levels. We are confident that the HRPC and laccase expression levels could be significantly enhanced to reach acceptable levels (1% TSP) (Daniell, 2006) by means of conventional breeding practices. For the high expressors, we are going to perform Southern blotting analysis to determine the transgene copy number and to do transgene mapping if possible.

4.6.3 Other Considerations to Boost Transgene Expression Levels

The expression of heterologous proteins in transgenic plants is an established technology and continues to undergo improvement. The classes of proteins that have been successfully expressed in plant systems at very high levels include industrially-useful enzymes, viral proteins, pharmaceutical proteins, and antibodies (Hood et al., 2007). The factors affecting the level of accumulation for each of these protein classes differ. The main factors include gene regulatory elements, codon usage, transgene copy number, position effects, protein targeting, and protein degradation (Streatfield, 2007). Among them, codon usage, promoter choice and protein targeting are critical to achieve high levels of target proteins in the tissue of interest.

4.6.3.1 Ep Promoter and Transgenes

Promoters are regions of the DNA generally upstream of a gene's coding region. It contains specific sequences recognized by proteins involved in the initiation of transcription (Potenza et al., 2004). Promoters regulate the gene expression levels and spatial patterns; therefore choosing a suitable promoter to drive the genes of interest is a key to obtaining optimal expression.

In seeds, only a few seed coat-specific promoters have been studied. For instance, barley seed coat specific promoter (Wu et al., 2000) and pea seed coat specific promoter (Buchner et al., 2002) were well characterized. The Ep promoter is a soybean seed coat-specific promoter. This thesis describes the first attempt to use the Ep promoter for plant biotechnology. Our intention was to target gene expression to the hourglass cells anticipating minimal disruption of plant growth. As an endogenous promoter, we hoped

for greater developmental and spatial control and stronger activity as it is native to the soybean HGC. We were aware of the potential risk that endogenous promoters could lead to gene silencing by co-suppression

The present data suggest that the tissue specificity of the Ep 1.5 kb promoter is good but its strength must be improved to achieve high levels of expression. Gene expression levels are governed by combinations of elements that control the rate of transcription initiation and transcript turnover. It is governed by a number of regulatory sequences additional to those generally considered to encompass a promoter. For instance, the 3' UTR can contribute to the control of gene expression (Sunilkumar et al., 2002). It was also found that the size of the promoter fragment could have an effect on gene expression (Taylor, 1997) since the removal of sequences with potential for regulation that are further 5' to the *cis* acting enhancer elements can eliminate interactions between them. Thus optimization of the Ep promoter must be accomplished by appropriate vector and transgene design. It should be noted that the Ep promoter could be affected by abiotic factors; for example, cold treatments (personal communications, Dr. M. Gijzen). The Ep promoter has *cis*-acting elements consistent with defence and stress responsiveness suggesting Ep it could be induced by environmental factors, such as abiotic and biotic stress. For these initial experiments we used the Ep 1.5kb promoter as gene expression is largely restricted in hourglass cell layer; however we realize that optimization is an important step.

Several of these design features were incorporated in our vectors. In constructs S8 and S18 we retained the 5' UTR region of the Ep gene (ATATTAACTCAAACC, 15 bp), because it was found that 5' UTRs enhanced gene expression in rice (Potenza et al., 2004)

and increased efficiency of mRNA translation in maize (Lu et al., 2008). As gene expression silencers may be found in 3'UTR regions (Mardanova et al., 2007), we did not include any 3' UTR region of SBP gene. Although some plant intron sequences enhance the expression of native genes (Kruys et al., 1987; Ortega et al., 2006) , it is a common practice to remove the intron splice sites, message destabilizing sequences and transcript termination sequences when the transgenes are synthesized (Giani et al., 2008) to eliminate any unexpected factors. In our soybean constructs, cDNA sequences of HRPC and laccase transgenes were synthesized using codons optimized for soybean.

There are several algorithms developed to optimize the codon usage of synthetic genes. The codons of HRPC, LCC1, and LCC4 were optimized for soybean. Codon optimization can dramatically increase foreign protein expression levels. It is reported that the codon-optimised version of *Plasmodium yoelii* merozoite surface protein 4/5 (PyMSP4/5) accumulated to levels six fold higher than the native *P. yoelii* sequence (Cai et al., 2008).

Matrix attachment regions (MARs) can be used to flank the transgenes in the vector to shield them from RNA silencing (Wang et al., 2008). Furthermore, many plant viruses contain translation-enhancing sequences that allow them to compete successfully with host messenger RNAs for the translation machinery (Rakotondrafara and Miller, 2008). Translational enhancers can be used in vectors to boost protein expression levels (Mlynarova et al., 2003; Butaye et al., 2004). Both strategies should be considered in the future studies.

4.6.3.2 Foreign Protein Degradation

Plants are an attractive host system for expressing foreign proteins although proteins can be produced using a range of heterologous expression vehicles. Being eukaryotes, plants correctly fold and assemble complex proteins and have the capacity for post-translational modifications such as glycosylation although the glycosylation in plants may be different from that in animals (Ma et al., 2003). The use of plants also eliminates the risks related to the use of potentially infectious animals, animal cells or animal-derived culture materials in the production process (Twyman et al., 2003). Low protein yield has been a significant problem limiting the commercial use of recombinant plant systems. Maximum product levels of 0.01 ~ 0.1% TSP or less were commonly reported (Daniell et al., 2001). Therefore, improving foreign protein accumulation is crucial for enhancing the commercial success of plant-based production systems (Doran, 2006). At any given time, the amount of foreign protein found in plant tissues is a balance between protein synthesis and protein degradation or loss.

Foreign protein degradation has been found to occur in many plant host species, for instance, soybean (Smith et al., 2002). In our Western blots, we did not see low-molecular-weight fragments, but this does not rule out the possibility of degradation since the digested fragments could be as small as 6–9 amino acids (Russell et al., 2005), which was below our levels of resolution. A number of factors could influence the susceptibility of foreign proteins to protease attack in plant cells. Differences between plant and fungal glycans that affect the folding and quaternary structure of laccase proteins (Ishimizu et al., 2004) rendering plant-derived proteins more vulnerable to protease activity. Mis-folding and lack of proper disulphide cross linking in plant-derived proteins can also increase the

likelihood of protease attack (Callis, 1995). Plant vacuoles, usually considered to be safe location for the expressed proteins, contain a variety of proteases that are active under mildly acidic conditions. They were identified as a possible site of recombinant protein degradation in potato (Dolja et al., 1998). Proteolytic degradation of foreign proteins can be minimized by targeting synthesis to the ER rather than the cytosol (unpublished results, Jaimie Schnell). In the future, we could integrate the ER retention signal (KDEL or HDEL) within our transgene because the plant cell ER contains few proteases and provides a relatively protective environment with molecular chaperones and stabilizing agents that interact with foreign proteins to enhance proper folding and assembly (Outchkourov et al., 2003). ER retention has resulted in foreign protein yields that are 10~100 times greater than those obtained when the protein is allowed to enter the secretory pathway (Nuttall et al., 2002).

The co-expression of protease inhibitors might be a potential strategy for lowering foreign protein degradation in planta. For instance, transgenic lettuce plants were developed with reduced protease activity due to the expression of heterologous protease inhibitor (Hellwig et al., 2004). The protease inhibitors did not significantly hinder normal plant growth and development. The use of HIS tag fused in N-terminus of a transgene was also found to enhance the stability of expressed protein (Gruis et al., 2002).

In summary, we have used a number of traditional strategies to enhance the production of enzymes in seed coats; however, several options exist for improvement in the future.

4.6.3.3 Potential Uses of Expressed HRPC and Laccases

Peroxidases are a family of isozymes found in all plants. They are heme-containing monomeric glycoproteins that utilize H_2O_2 to oxidize a wide variety of molecules (Reed et al., 2005). Horseradish peroxidase has been the most studied and widely used peroxidase. Horseradish peroxidases have attractive biocatalytic properties and are used in biosensing and immunoassays (Kamal and Behere, 2008). Hydroxylated aromatic compounds (HACs) are considered to be primary pollutants in a wide variety of industrial wastewaters. Horseradish peroxidase is suitable for the removal of these toxic substances. The results showed that HRP acted faster than SBP (Bódalo et al., 2006). Apart from phenols, HRPC can also oxide carbon monoxide, formate, and acetate (Carlsson et al., 2005). A working model has been proposed to improve phenol removal by applying minimum enzyme concentrations, it has achieved the highest removal in comparison with previous studies (Ghasempur et al., 2007).

The laccase of the fungus *Trametes versicolor* was able to polymerize various halogen-, alkyl-, and alkoxy-substituted anilines, showing substrate specificity similar to that of the horseradish peroxidase. The product profiles obtained by high-pressure liquid chromatography were similar for the two enzymes. In addition to horseradish peroxidase, a *T. versicolor* laccase can also polymerize aniline derivatives (Hoff et al., 1985). The HRPC and laccase produced in soybean seed coat may provide the bulk quantities needed for industrial applications, such as paper making industry, and waste water treatment. This could generate new industries and create new uses for soybean seed coats.

4.6.4 Summary of Future Research Directions

The results presented in this thesis suggest several lines of study both for HRPC and laccase transgenics, including: 1. Improvements to the expression cassette in order to achieve higher levels of enzyme are necessary. In our experiments the Ep 1.5 kb promoter was used for practical reasons. In the future, a clone containing a longer region of upstream DNA 5' driving expression of marker genes such as GUS or enhanced GFP could be the starting point for deletion analysis to discover which part(s) of the fragment are responsible for high level and tissue-specific expression. These findings could be used to improve the transformation vectors. Other regulatory elements such as transcriptional and/or translational enhancers could be added to the transgenes. For example the inclusion of introns into transgenes (Rose and Last, 1997; Fiume et al., 2004; Samadder et al., 2008) increases expression as does the translational enhancer from signal from soybean chlorotic mottle virus (Hasegawa et al., 1989).

2. Excess vector sequences are often densely methylated and this could spread into the neighbouring transgenes, and may contribute to transgene silencing. This hypothesis can be tested by Southern blotting analysis in which gDNA fragments are generated by the restriction enzymes that cut only vector backbone DNA sequence but not transgene expression cassette (for instance, *Xho* I). At the same time, the gDNA from the untransformed cv. Jack cut with the same restriction enzymes is used as the negative control. If the same fragment hybridized to the original pCAMBIA 1300 vector and the transgene, it indicates that the transgene is flanked with vector DNA sequence. We can also use PCR to test if the transgene is flanked with vector backbone DNA sequence. The solution should be to design primer pairs that cover the 5' and 3' boundaries, for example,

one primer flanks the insert, and other flanks the vector DNA sequence. If the PCR result is not positive, we might need to clone whole insert to determine if the transgene is flanked with the excess of vector DNA sequence. Another strategy is to use the methylation sensitive restriction enzymes, such as *Hpa*II and *Msp*I, to digest the gDNA of HRPC transgenic soybean. Methlyation of the transgene would be indicated if the transgene can be digested with *Msp*I but not by *Hpa*II (Quemada et al., 1987; Tang and Leisner, 1998). In this case, the plasmid DNA of S8 HRPC cut with the same enzymes can be used as positive control.

3. Determine experimentally the number of generations needed to recover stable lines and to determine whether the enzyme activity in the seed coat remains stable. This can be done by selfing plants up to generations T_3 or T_4 until homozygozity has been reached (as judged by monitoring the inheritance of the transgene as was done in section 3.4.4) and testing the enzyme activity in many seeds we have done in (section 3.4.3). In addition we can test the hypothesis that crossing lines will increase activity by crossing the elite lines using the traditional breeding strategy.

4. If plants from T_3 or T_4 generation expressing higher levels of enzyme can be isolated, we can investigate whether tissue-specific expression of the transgene has been achieved. This can be done by RT-PCR for transgene mRNA in the main organs of transgenic soybean, such as seed pod, seed coat, seed embryo, flower, stem, leaf, hypocotyl, and root. It may be possible to detect activity directly in seed coats by staining and by Western blotting using anti-HRPC, anti-laccase, and anti-HIS tag antibodies. In this thesis we require purified protein from 80 seeds for Western analysis but improvements in expression and in detection are still possible.

REFERENCES CITED

Abdel-Ghany SE, Pilon M (2008) Micro RNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in *Arabidopsis*. J Biol Chem **283:** 15932-15945

Alcalde M (2007) Laccases: biological functions, molecular structure and industrial application. *In* Industrial Enzymes. Springer, pp 461-476

Algan G, Buyukkartal H (2000) Ultrastructure of seed coat development in the natural tetraploid *Trifolium pratense* L. Journal of Agronomy and Crop Science **184:** 205-213

Argos P (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J Mol Biol **211:** 943-958

Ashtamker C, Kiss V, Sagi M, Davydov O, Fluhr R (2007) Diverse subcellular locations of cryptogein-induced reactive oxygen species production in tobacco bright yellow-2 cells. Plant Physiol **143:** 1817-1826

Assaad FF, Signer ER (1992) Somatic and germinal recombination of a direct repeat in *Arabidopsis*. Genetics **132:** 553-566

Astua-Monge G, Lyznik A, Jones V, Mackenzie SA, Vallejos CE (2002) Evidence for a prokaryotic insertion-sequence contamination in eukaryotic sequences registered in different databases. Theor Appl Genet **104:** 48-53

Auriol M, Filali-Meknassi Y, Adams CD, Tyagi RD, Noguerol TN, Pina B (2008) Removal of estrogenic activity of natural and synthetic hormones from a municipal wastewater: efficiency of horseradish peroxidase and laccase from Trametes versicolor. Chemosphere **70:** 445-452

Bódalo A, Gomez JL, Gomez E, Bastida J, Maximo MF (2006) Comparison of commercial peroxidases for removing phenol from water solutions. Chemosphere **63:** 626-632

Bailey MR, Woodard SL, Callaway E, Beifuss K, Magallanes-Lundback M, Lane JR, Horn ME, Mallubhotla H, Delaney DD, Ward M, Van Gastel F, Howard JA, Hood EE (2004) Improved recovery of active recombinant laccase from maize seed. Applied Microbiology and Biotechnology **63:** 390-397

Bakalovic N, Passardi F, Ioannidis V, Cosio C, Penel C, Falquet L, Dunand C (2006) PeroxiBase: a class III plant peroxidase database. Phytochemistry **67:** 534-539

Batard Y, Hehn A, Nedelkina S, Schalk M, Pallett K, Schaller H, Werck-Reichhart D (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. Arch Biochem Biophys **379:** 161-169

Beaujean A, Issakidis-Bourguet E, Catterou M, Dubois F, Sangwan RS, Sangwan-Norreel BS (2001) Integration and expression of Sorghum C(4) phosphoenolpyruvate carboxylase and chloroplastic NADP(+)-malate dehydrogenase separately or together in C(3) potato plants(1). Plant Sci **160:** 1199-1210

Bertrand T, Jolivalt C, Briozzo P, Caminade E, Joly N, Madzak C, Mougin C (2002) Crystal structure of a four-copper laccase complexed with an arylamine: insights

into substrate recognition and correlation with kinetics. Biochemistry **41:** 7325-7333

Boehm R (2007) Bioproduction of therapeutic proteins in the 21st century and the role of plants and plant cells as production platforms. Ann N Y Acad Sci **1102:** 121-134

Bohnsack JF, Takahashi S, Hammitt L, Miller DV, Aly AA, Adderson EE (2000) Genetic polymorphisms of group B streptococcus *scpB* alter functional activity of a cell-associated peptidase that inactivates C5a. Infect. Immun. **68:** 5018-5025

Bourbonnais R, Paice MG, Freiermuth B, Bodie E, Borneman S (1997) Reactivities of various mediators and laccases with kraft pulp and lignin model compounds. Appl. Environ. Microbiol. **63:** 4627-4632

Bourbonnais R, Paice MG, Reid ID, Lanthier P, Yaguchi M (1995) Lignin oxidation by laccase isozymes from *Trametes versicolor* and role of the mediator 2,2'-azinobis(3-ethylbenzthiazoline-6-sulfonate) in kraft lignin depolymerization. Appl Environ Microbiol **61:** 1876-1880

Brandizzi F, Hanton S, DaSilva LL, Boevink P, Evans D, Oparka K, Denecke J, Hawes C (2003) ER quality control can lead to retrograde transport from the ER lumen to the cytosol and the nucleoplasm in plants. Plant J **34:** 269-281

Brandizzi F, Irons SL, Johansen J, Kotzer A, Neumann U (2004) GFP is the way to glow: bioimaging of the plant endomembrane system. J Microsc **214:** 138-158

Brookes G, Barfoot P (2006) GM Crops: the first ten years - global socio-economic and environmental impacts. ISAAA (International Service for the acquisition of agri-biotech applications) Briefs

Brown MA, Zhao Z, Grant Mauk A (2002) Expression and characterization of a recombinant multi-copper oxidase: laccase IV from *Trametes versicolor.* Inorganica Chimica Acta **331:** 232-238

Bruce KJ, Karr-Lilienthal LK, Zinn KE, Pope LL, Mahan DC, Fastinger ND, Watts M, Utterback PL, Parsons CM, Castaneda EO, Ellis M, Fahey GC, Jr. (2006) Evaluation of the inclusion of soybean oil and soybean processing by-products to soybean meal on nutrient composition and digestibility in swine and poultry. J Anim Sci **84:** 1403-1414

Buchner P, Rochat C, Wuilleme S, Boutin JP (2002) Characterization of a tissue-specific and developmentally regulated beta-1,3-glucanase gene in pea (*Pisum sativum*). Plant Mol Biol **49:** 171-186

Butaye KMJ, Goderis IJWM, Wouters PFJ, Pues JMTG, Delaur SL, Broekaert WF, Depicker A, Cammue BPA, De Bolle MFC (2004) Stable high-level transgene expression in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. The Plant Journal **39:** 440-449

Buttery B, Buzzell R (1968) Peroxidase activity in the seeds of soybean varieties. Crop Science **8:** 722-725

Buzzell R, Anderson T, Hill A, Welacky T (1991) Harovinton soybean. Can J Plant Sci **71:** 525-526

Buzzell RI, Buttery BR, MacTavish DC (1987) Biochemical genetics of black pigmentation of soybean seed. J Hered **78:** 53-54

Cai Y, Sun J, Wang J, Ding Y, Tian N, Liao X, Xu W (2008) Optimizing the codon usage of synthetic gene with QPSO algorithm. J Theor Biol **254:** 123-127

Callis J (1995) Regulation of protein degradation. Plant Cell **7:** 845-857

Carlson J, Lersten N (1987) Reproductive morphology. *In* Je Wilcox, ed, Soybeans: improvement, production and uses. ASA-CSSA-SSSA Publishers, Madison, Wisconsin, pp 97–102

Carlsson GH, Berglund GI, Hajdu J, Nicholls P, Svistunenko D (2005) Complexes of horseradish peroxidase with formate, acetate, and carbon monoxide. Biochemistry **44:** 635-642

Chalfie M (1995) Green fluorescent protein. Photochem Photobiol **62:** 651-656

Chang SW, Lee GC, Shaw JF (2006) Codon optimization of *Candida rugosa* lip1 gene for improving expression in *Pichia pastoris* and biochemical characterization of the purified recombinant LIP1 lipase. J Agric Food Chem **54:** 815-822

Chang SW, Lee GC, Shaw JF (2006) Efficient production of active recombinant *Candida rugosa* LIP3 lipase in *Pichia pastoris* and biochemical characterization of the purified enzyme. J Agric Food Chem **54:** 5831-5838

Chattopadhyay K, Mazumdar S (2000) Structural and conformational stability of horseradish peroxidase: effect of temperature and pH. Biochemistry **39:** 263-270

Choffnes D, Philip R, Vodkin L (2001) A transgenic locus in soybean exhibits a high level of recombination. In Vitro Cellular and Development Biology - Plant **37:** 756-762

Claus H (2003) Laccases and their occurrence in prokaryotes. Arch Microbiol **179:** 145-150

Claus H (2004) Laccases: structure, reactions, distribution. Micron **35:** 93-96

Claus H, Faber G, Konig H (2002) Redox-mediated decolorization of synthetic dyes by fungal laccases. Appl Microbiol Biotechnol **59:** 672-678

Clements JC, Zvyagin AV, Silva KKMBD, Wanner T, Sampson DD, Cowling WA (2004) Optical coherence tomography as a novel tool for non-destructive measurement of the hull thickness of lupin seeds. *In*, Vol 123, pp 266-270

Clough RC, Pappu K, Thompson K, Beifuss K, Lane J, Delaney DE, Harkey R, Drees C, Howard JA, Hood EE (2006) Manganese peroxidase from the white-rot fungus *Phanerochaete chrysosporium* is enzymatically active and accumulates to high levels in transgenic maize seed. Plant Biotechnol J **4:** 53-62

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735-743

Collinge M, Boller T (2001) Differential induction of two potato genes, *Stprx*2 and *StNAC*, in response to infection by *Phytophthora* infestans and to wounding. Plant Mol Biol **46:** 521-529

Corner E (1951) The leguminous seed. Phytomorphology **1:** 117-150.

Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) Understanding, improving and using green fluorescent proteins. Trends Biochem Sci **20:** 448-455

Daniell H (2006) Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. Biotechnology Journal **1:** 1071-1079

Daniell H, Streatfield SJ, Wycoff K (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends in Plant Science **6:** 219-226

De Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G (2002) Boosting heterologous protein production in transgenic

dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. Nature Biotechnology **20:** 1265-1268

de Wilde C, Uzan E, Zhou Z, Kruus K, Andberg M, Buchert J, Record E, Asther M, Lomascolo A (2008) Transgenic rice as a novel production system for Melanocarpus and *Pycnoporus laccases*. Transgenic Research **17:** 515-527

Di Sansebastiano G, Renna L, Gigante M, De Caroli M, Piro G, Dalessandro G (2007) Green fluorescent protein reveals variability in vacuoles of three plant species. Biologia Plantarum **51:** 49-55

Dietmar B, Claus B, Heinz H (1988) Pharmaceutical Proteins. Angewandte Chemie International Edition in English **27:** 207-225

Dolja VV, Peremyslov VV, Keller KE, Martin RR, Hong J (1998) Isolation and stability of histidine-tagged proteins produced in plants via potyvirus gene vectors. Virology **252:** 269-274

Doran PM (2006) Foreign protein degradation and instability in plants and plant tissue cultures. Trends Biotechnol **24:** 426-432

Drakakaki G, Marcel S, Arcalis E, Altmann F, Gonzalez-Melendi P, Fischer R, Christou P, Stoger E (2006) The intracellular fate of a recombinant protein is tissue dependent. Plant Physiol **141:** 578-586

Duroux L, Welinder KG (2003) The peroxidase gene family in plants: a phylogenetic overview. J Mol Evol **57:** 397-407

Emanuelsson O, von Heijne G (2001) Prediction of organellar targeting signals. Biochim Biophys Acta **1541:** 114-119

Faye L, Boulaflous A, Benchabane M, Gomord V, Michaud D (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. Vaccine **23:** 1770-1778

Feng L, Chan WW, Roderick SL, Cohen DE (2000) High-level expression and mutagenesis of recombinant human phosphatidylcholine transfer protein using a synthetic gene: evidence for a C-terminal membrane binding domain. Biochemistry **39:** 15399-15409

Feng S, Xue L, Liu H, Lu P (2008) Improvement of efficiency of genetic transformation for *Dunaliella salina* by glass beads method. Mol Biol Rep

Finer J, McMullen M (1991) Transformation of soybean via particle bombardment of embryogenic suspension culture tissue. *In Vitro* Cellular & Developmental Biology - Plant **27:** 175-182

Fischer R, Schillberg S, Emans N (2001) Molecular farming: a field of growing promise. Outlook Agr. **30:** 31-36

Fiume E, Christou P, Giani S, Breviario D (2004) Introns are key regulatory elements of rice tubulin expression. Planta **218:** 693-703

Frary A, Earle ED (1996) An examination of factors affecting the efficiency of *Agrobacterium*-mediated transformation of tomato. Plant Cell Reports **16:** 235-240

Frommer WB, Ninnemann O (1995) Heterologous expression of genes in bacterial, fungal, animal, and plant cells. Annual Review of Plant Physiology and Plant Molecular Biology **46:** 419-444

Fujiyama K, Takemura H, Shibayama S, Kobayashi K, Choi JK, Shinmyo A, Takano M, Yamada Y, Okada H (1988) Structure of the horseradish peroxidase isozyme C genes. Eur J Biochem **173:** 681-687

Fujiyama K, Takemura H, Shinmyo A, Okada H, Takano M (1990) Genomic DNA structure of two new horseradish-peroxidase-encoding genes. Gene **89:** 163-169

Gajhede M, Schuller DJ, Henriksen A, Smith AT, Poulos TL (1997) Crystal structure of horseradish peroxidase C at 2.15 A resolution. Nat Struct Biol **4:** 1032-1038

Gallie DR, Kado CI (1989) A translational enhancer derived from tobacco mosaic virus is functionally equivalent to a Shine-Dalgarno sequence. Proceedings of the National Academy of Sciences of the United States of America **86:** 129-132

Gardner JC, Payne TL (2003) A soybean biotechnology outlook. AgBioForum **6:** 1-3

Garzillo AM, Colao MC, Caruso C, Caporale C, Celletti D, Buonocore V (1998) Laccase from the white-rot fungus *Trametes trogii*. Appl Microbiol Biotechnol **49:** 545-551

Gellissen G, Kunze G, Gaillardin C, Cregg JM, Berardi E, Veenhuis M, van der Klei I (2005) New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica* - a comparison. FEMS Yeast Res **5:** 1079-1096

Ghasempur S, Torabi SF, Ranaei-Siadat SO, Jalali-Heravi M, Ghaemi N, Khajeh K (2007) Optimization of peroxidase-catalyzed oxidative coupling process for phenol removal from wastewater using response surface methodology. Environ Sci Technol **41:** 7073-7079

Giani S, Altana A, Campanoni P, Morello L, Breviario D (2008) In trangenic rice, alpha- and beta-tubulin regulatory sequences control GUS amount and distribution through intron mediated enhancement and intron dependent spatial expression. Transgenic Res (published online)

Gijzen M (1997) A deletion mutation at the *ep* locus causes low seed coat peroxidase activity in soybean. Plant J **12:** 991-998

Gijzen M, Van Huystee R, Buzzell RI (1993) Soybean seed coat peroxidase (A comparison of high-activity and low-activity genotypes). Plant Physiol **103:** 1061-1066

Gilbert RA, Gallo-Meagher M, Comstock JC, Miller JD, Jain M, Abouzid A (2005) Agronomic evaluation of sugarcane lines transformed for resistance to sugarcane mosaic virus strain E. *In*, Vol 45, pp 2060-2067

Gillikin JW, Graham JS (1991) Purification and developmental analysis of the major anionic peroxidase from the seed coat of *Glycine max*. Plant Physiol. **96:** 214-220

Gray JS, Yang BY, Montgomery R (1998) Heterogeneity of glycans at each N-glycosylation site of horseradish peroxidase. Carbohydr Res **311:** 61-69

Griesbeck O, Baird GS, Campbell RE, Zacharias DA, Tsien RY (2001) Reducing the environmental sensitivity of yellow fluorescent protein: mechanism and applications. J Biol Chem **276:** 29188-29194

Gruis D, Selinger DA, Curran JM, Jung R (2002) Redundant proteolytic mechanisms process seed storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases. Plant Cell **14:** 2863-2882

Gu Y, Di WL, Kelsell DP, Zicha D (2004) Quantitative fluorescence resonance energy transfer (FRET) measurement with acceptor photobleaching and spectral unmixing. J Microsc **215:** 162-173

Guo M, Lu F, Du L, Pu J, Bai D (2005) Optimization of the expression of a laccase gene from *Trametes versicolor* in *Pichia methanolica*. Appl Microbiol Biotechnol: 1-5

Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol **22:** 346-353

Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. Plant Physiol **137:** 1397-1419

Han MJ, Choi HT, Song HG (2005) Purification and characterization of laccase from the white rot fungus *Trametes versicolor*. J Microbiol **43:** 555-560

Hanson MR, Kohler RH (2001) GFP imaging: methodology and application to investigate cellular compartmentation in plants. J Exp Bot **52:** 529-539

Hasegawa A, Verver J, Shimada A, Saito M, Goldbach R, Van Kammen A, Miki K, Kameya-Iwaki M, Hibi T (1989) The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. Nucleic Acids Res **17:** 9993-10013

Heggie L, Jansen MA, Burbridge EM, Kavanagh TA, Thorneley RN, Dix PJ (2005) Transgenic tobacco (*Nicotiana tabacum* L. cv. Samsun-NN) plants over-expressing a synthetic HRP-C gene are altered in growth, development and susceptibility to abiotic stress. Plant Physiol Biochem **43:** 1067-1073

Hellwig S, Drossard J, Twyman RM, Fischer R (2004) Plant cell cultures for the production of recombinant proteins. Nat Biotech **22:** 1415-1422

Henriksen A, Mirza O, Indiani C, Teilum K, Smulevich G, Welinder KG, Gajhede M (2001) Structure of soybean seed coat peroxidase: a plant peroxidase with unusual stability and haem-apoprotein interactions. Protein Sci **10:** 108-115

Hirai H, Kashima Y, Hayashi K, Sugiura T, Yamagishi K, Kawagishi H, Nishida T (2008) Efficient expression of laccase gene from white-rot fungus Schizophyllum commune in a transgenic tobacco plant. FEMS Microbiol Lett

Hoff T, Liu SY, Bollag JM (1985) Transformation of halogen-, alkyl-, and alkoxy-substituted anilines by a laccase of *Trametes versicolor*. Appl Environ Microbiol **49:** 1040-1045

Hood EE (2004) Bioindustrial and biopharmaceutical products from plants. New directions for a diverse planet. Proceedings of the 4 th International Crop Science Congress. 26 Sep - 1 Oct. Brisbane, Austrialia

Hood EE, Bailey MR, Beifuss K, Magallanes-Lundback M, Horn ME, Callaway E, Drees C, Delaney DE, Clough R, Howard JA (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. Plant Biotechnol J **1:** 129-140

Hood EE, Howard JA (2002) Plants as factories for protein production. . Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hood EE, Jilka JM (1999) Plant-based production of xenogenic proteins. Curr Opin Biotechnol **10:** 382-386

Hood EE, Love R, Lane J, Bray J, Clough R, Pappu K, Drees C, Hood KR, Yoon S, Ahmad A, Howard JA (2007) Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. Plant Biotechnol J **5:** 709-719

Horvath H, Jensen LG, Wong OT, Kohl E, Ullrich SE, Cochran J, Kannangara CG, von Wettstein D (2001) Stability of transgene expression, field performance and recombination breeding of transformed barley lines. TAG Theoretical and Applied Genetics **102:** 1-11

Hu S, Li L, Qiao J, Guo Y, Cheng L, Liu J (2006) Codon optimization, expression, and characterization of an internalizing anti-ErbB2 single-chain antibody in *Pichia pastoris*. Protein Expr Purif **47:** 249-257

Ishimizu T, Sasaki A, Okutani S, Maeda M, Yamagishi M, Hase S (2004) Endo-beta-mannosidase, a plant enzyme acting on N-glycan: purification, molecular cloning, and characterization. J. Biol. Chem. **279:** 38555-38562

Jönsson LJ, Saloheimo M, Penttila M (1997) Laccase from the white-rot fungus *Trametes versicolor*: cDNA cloning of lcc1 and expression in *Pichia pastoris*. Curr Genet **32:** 425-430

Jakowitsch J, Papp I, Moscone EA, van der Winden J, Matzke M, Matzke AJ (1999) Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans. Plant J **17:** 131-140

Jefferson RA (1989) The GUS reporter gene system. Nature **342:** 837-838

Jiang L, Sun SS (2002) Membrane anchors for vacuolar targeting: application in plant bioreactors. Trends Biotechnol **20:** 99-102

Jin Z (2008) Hourglass cell development in the soybean seed coat. University of Ottawa, Ottawa

Joersbo M, Brunstedt J, Marcussen J, Okkels FT (1999) Transformation of the endospermous legume guar (*Cyamopsis tetragonoloba* L.) and analysis of transgene transmission. Molecular Breeding **5:** 521-529

Johannes C, Majcherczyk A (2000) Natural mediators in the oxidation of polycyclic aromatic hydrocarbons by laccase mediator systems. Appl. Environ. Microbiol. **66:** 524-528

Jolliffe NA, Craddock CP, Frigerio L (2005) Pathways for protein transport to seed storage vacuoles. Biochem Soc Trans **33:** 1016-1018

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol **43:** 179-188

Kamal JK, Behere DV (2003) Activity, stability and conformational flexibility of seed coat soybean peroxidase. J Inorg Biochem **94:** 236-242

Kamal JKA, Behere DV (2008) Kinetic stabilities of soybean and horseradish peroxidases. Biochemical Engineering Journal **38:** 110-114

Kawaoka A, Matsunaga E, Endo S, Kondo S, Yoshida K, Shinmyo A, Ebinuma H (2003) Ectopic expression of a horseradish peroxidase enhances growth rate and increases oxidative stress resistance in hybrid aspen. Plant Physiol **132:** 1177-1185

Kawaoka A, Sato S, Nakahara K, Matsushima N, Okada N, Sekine M, Shinmyo A, Takano M (1992) Expression and promoter activity of genes for isozymes of horseradish peroxidase. Plant Cell Physiol. **33:** 1143-1150

Kehlenbach RH, Dickmanns A, Kehlenbach A, Guan T, Gerace L (1999) A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export. J Cell Biol **145:** 645-657

Kerley MS, Allee GL (2003) Modifications in soybean seed composition to enhance animal feed use and value: Moving from a dietary ingredient to a functional dietary component. AgBioForum **6:** 14-17

Kiiskinen L (2004) Characterization and heterologous production of a novel laccase from *Melanocarpus albomyces*. VTT Publication **566**

Kiiskinen LL, Kruus K, Bailey M, Ylosmaki E, Siika-Aho M, Saloheimo M (2004) Expression of *Melanocarpus albomyces* laccase in *Trichoderma reesei* and characterization of the purified enzyme. Microbiology **150:** 3065-3074

Kim YH, Kim CY, Song WK, Park DS, Kwon SY, Lee HS, Bang JW, Kwak SS (2008) Overexpression of sweet potato *swpa*4 peroxidase results in increased hydrogen peroxide production and enhances stress tolerance in tobacco. Planta **227:** 867-881

Kis M, Burbridge E, Brock IW, Heggie L, Dix PJ, Kavanagh TA (2004) An N-terminal peptide extension results in efficient expression, but not secretion, of a synthetic horseradish peroxidase gene in transgenic tobacco. Ann Bot (Lond) **93:** 303-310

Klein TM, Arentzen R, Lewis PA, Fitzpatrick-McElligott S (1992) Transformation of microbes, plants and animals by particle bombardment. Biotechnology (N Y) **10:** 286-291

Klein TM, Fromm M, Weissinger A, Tomes D, Schaaf S, Sletten M, Sanford JC (1988) Transfer of foreign genes into intact maize cells with high-velocity microprojectiles. Proc Natl Acad Sci U S A **85:** 4305-4309

Ko K, Koprowski H (2005) Plant biopharming of monoclonal antibodies. Virus Res **111:** 93-100

Kojima Y, Tsukuda Y, Kawai Y, Tsukamoto A, Sugiura J, Sakaino M, Kita Y (1990) Cloning, sequence analysis, and expression of ligninolytic phenoloxidase genes of the white-rot basidiomycete *Coriolus hirsutus*. J Biol Chem **265:** 15224-15230

Koschorreck K, Richter SM, Swierczek A, Beifuss U, Schmid RD, Urlacher VB (2008) Comparative characterization of four laccases from *Trametes versicolor* concerning phenolic C-C coupling and oxidation of PAHs. Arch Biochem Biophys **474:** 213-219

Kotula L, Curtis PJ (1991) Evaluation of foreign gene codon optimization in yeast: expression of a mouse IG kappa chain. Biotechnology (N Y) **9:** 1386-1389

Koziel MG, Carozzi NB, Desai N (1996) Optimizing expression of transgenes with an emphasis on post-transcriptional events. Plant Mol Biol **32:** 393-405

Kruys V, Wathelet M, Poupart P, Contreras R, Fiers W, Content J, Huez G (1987) The 3' untranslated region of the human interferon-beta mRNA has an inhibitory effect on translation. Proc Natl Acad Sci U S A **84:** 6030-6034

Kumar SV, Phale PS, Durani S, Wangikar PP (2003) Combined sequence and structure analysis of the fungal laccase family. Biotechnol Bioeng **83:** 386-394

Kushad MM, Guidera M, Bratsch AD (1999) Distribution of horseradish peroxidase activity in horseradish plants. HortScience **34:** 127-129

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature **227:** 680-685

Latham JR, Wilson AK, Steinbrecher RA (2006) The mutational consequences of plant transformation. J Biomed Biotechnol **2006:** 25376

Li L, Steffens JC (2002) Overexpression of polyphenol oxidase in transgenic tomato plants results in enhanced bacterial disease resistance. Planta **215:** 239-247

Li SC, Squires CL, Squires C (1984) Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda nut-like sequences. Cell **38:** 851-860

Li X, Zeng W, Zhang Y (2002) Cloning, purification and biological activity of human vascular endothelial growth factor fragment in *E. coli*. Zhonghua Zhong Liu Za Zhi (Cancer Journal of China) **24:** 448-450

Lissemore JL, Jankowski JT, Thomas CB, Mascotti DP, deHaseth PL (2000) Green fluorescent protein as a quantitative reporter of relative promoter activity in *E. coli*. Biotechniques **28:** 82-84, 86, 88-89

Liu J, Zhang Y, Qiu L, Ye L, Xia Y, Su Z (2005) A novel process of purifying soybean hull peroxidase. Chemical and Biochemical Engineering Quarterly **19:** 199-205

Loustau MN, Romero LV, Levin GJ, Magri ML, Lopez MG, Taboga O, Cascone O, Miranda MV (2008) Expression and purification of horseradish peroxidase in insect larvae. Process Biochemistry **43:** 103-107

Lu J, Sivamani E, Azhakanandam K, Samadder P, Li X, Qu R (2008) Gene expression enhancement mediated by the 5' UTR intron of the rice *rubi*3 gene varied remarkably among tissues in transgenic rice plants. Mol Genet Genomics **279:** 563-572

Ma JKC, Drake PMW, Christou P (2003) The production of recombinant pharmaceutical proteins in plants. Nat Rev Genet **4:** 794-805

Macauley-Patrick S, Fazenda M, McNeil B, and Harvey L (2005) Heterologous protein production using the *Pichia pastoris* expression system. Yeast **22:** 249-270

Mahan JR, Dotray PA, Light GG (2004) Thermal dependence of enzyme function and inhibition; implications for, herbicide efficacy and tolerance. Physiologia Plantarum **120:** 187-195

Malik K, Wu K, Li XQ, Martin-Heller T, Hu M, Foster E, Tian L, Wang C, Ward K, Jordan M, Brown D, Gleddie S, Simmonds D, Zheng S, Simmonds J, Miki B (2002) A constitutive gene expression system derived from the tCUP cryptic promoter elements. Theor Appl Genet **105:** 505-514

Mander GJ, Wang H, Bodie E, Wagner J, Vienken K, Vinuesa C, Foster C, Leeder AC, Allen G, Hamill V, Janssen GG, Dunn-Coleman N, Karos M, Lemaire HG, Subkowski T, Bollschweiler C, Turner G, Nusslein B, Fischer R (2006) Use of laccase as a novel, versatile reporter system in filamentous fungi. Appl Environ Microbiol. **72:** 5020–5026

Mardanova ES, Zamchuk LA, Ravin NV (2007) The 5' untranslated region of the maize alcohol dehydrogenase gene provides efficient translation of mRNA in plants under stress conditions. Mol Biol (Mosk) **41:** 1002-1008

Maruyama N, Mun LC, Tatsuhara M, Sawada M, Ishimoto M, Utsumi S (2006) Multiple vacuolar sorting determinants exist in soybean 11S globulin. Plant Cell **18:** 1253-1273

Matsui T, Hori M, Shizawa N, Nakayama H, Shinmyo A, Yoshida K (2006) High-efficiency secretory production of peroxidase C1a using vesicular transport engineering in transgenic tobacco. J Biosci Bioeng **102:** 102-109

Matsui T, Nakayama H, Yoshida K, Shinmyo A (2003) Vesicular transport route of horseradish C1a peroxidase is regulated by N- and C-terminal propeptides in tobacco cells. Appl Microbiol Biotechnol **62:** 517-522

Matzke AJ, Matzke MA (1998) Position effects and epigenetic silencing of plant transgenes. Curr Opin Plant Biol **1:** 142-148

Matzke MA, Neuhuber F, Matzke AJ (1993) A variety of epistatic interactions can occur between partially homologous transgene loci brought together by sexual crossing. Mol Gen Genet **236:** 379-386

Melander M, Kamnert I, Happstadius I, Liljeroth E, Bryngelsson T (2006) Stability of transgene integration and expression in subsequent generations of doubled haploid oilseed rape transformed with chitinase and beta-1,3-glucanase genes in a double-gene construct. Plant Cell Rep **25:** 942-952

Mello-Farias PCd, Chaves ALS (2008) Advances in *Agrobacterium*-mediated plant transformation with enphasys on soybean. Scientia Agricola **65:** 95-106

Miklas P, Townsend C, Ladd S (1987) Seed coat anatomy and the scarification of *Cicer Milkvetch* seed. Crop Sci **27:** 766-772

Miller S, Bowman L, Gijzen M, and Miki B (1999) Early development of the seed coat of soybean (*Glycine max*). Annals of Botany **84:** 297-304

Miranda MV, Fernández-lahore HM, Osvaldo Cascone D, Dobrecky J (1998) The extractive purification of peroxidase from plant raw materials in aqueous two-phase systems. Acta Biotechnologica **18:** 179-188

Mlynarova L, Hricova A, Loonen A, Nap JP (2003) The presence of a chromatin boundary appears to shield a transgene in tobacco from RNA silencing. Plant Cell **15:** 2203-2217

Mlynarova L, Loonen A, Heldens J, Jansen RC, Keizer P, Stiekema WJ, Nap JP (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell **6:** 417-426

Moïse JA, Han S, Gudynaite-Savitch L, Johnson DA, Miki BLA (2005) Seed coats: structure, development, composition, and biotechnology. In Vitro Cellular and Developmental Biology - Plant **41:** 620-644

Moravec T, Schmidt MA, Herman EM, Woodford-Thomas T (2007) Production of *Escherichia coli* heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine. Vaccine **25:** 1647-1657

Morawski B, Lin Z, Cirino P, Joo H, Bandara G, Arnold FH (2000) Functional expression of horseradish peroxidase in *Saccharomyces cerevisiae* and *Pichia pastoris*. Protein Eng **13:** 377-384

Morcillo F, Gagneur C, Adam H, Richaud F, Singh R, Cheah SC, Rival A, Duval Y, Tregear JW (2006) Somaclonal variation in micropropagated oil palm. Characterization of two novel genes with enhanced expression in epigenetically abnormal cell lines and in response to auxin. Tree Physiol **26:** 585-594

Mortensen K, Larsson LI (2001) Quantitative and qualitative immunofluorescence studies of neoplastic cells transfected with a construct encoding p53-EGFP. J Histochem Cytochem **49:** 1363-1367

Muntz K (2007) Protein dynamics and proteolysis in plant vacuoles. J Exp Bot **58:** 2391-2407

Muntz K, Christov V, Saalbach G, Saalbach I, Waddell D, Pickardt T, Schieder O, Wustenhagen T (1998) Genetic engineering for high methionine grain legumes. Nahrung **42:** 125-127

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiologia Plantarum **15:** 473-497

Necochea R, Valderrama B, Diaz-Sandoval S, Folch-Mallol JL, Vazquez-Duhalt R, Iturriaga G (2005) Phylogenetic and biochemical characterisation of a recombinant laccase from *Trametes versicolor*. FEMS Microbiol Lett **244:** 235-241

Neuhaus J, Boevink P (2001) The green fluorescent protein (GFP) as a reporter in plant cells. Oxford University Press, Oxford, UK

Norris S (2005) Molecular farming. Science and Technology Division, Library of Parliament

Nuttall J, Vine N, Hadlington JL, Drake P, Frigerio L, Ma JKC (2002) ER - resident chaperone interactions with recombinant antibodies in transgenic plants. European Journal of Biochemistry **269:** 6042-6051

Ogasawara T, Chikagawa Y, Arakawa F, Nozaki A, Itoh Y, Sasaki K, Umetsu H, Watanabe T, Akiyama H, Maitani T, Toyoda M, Kamada H, Goda Y, Ozeki Y (2005) Frequency of mutations of the transgene, which might result in the loss of the Glyphosate-tolerant phenotype, was lowered in Roundup Ready® soybeans. Journal of Health Science **51:** 197-201

Ohkawa J, Okada N, Shinmyo A, Takano M (1989) Primary structure of cucumber (*Cucumis sativus*) ascorbate oxidase deduced from cDNA sequence: homology with blue copper proteins and tissue-specific expression. Proc Natl Acad Sci U S A **86:** 1239-1243

Olins PO, Devine CS, Rangwala SH, Kavka KS (1988) The T_7 phage gene 10 leader RNA, a ribosome-binding site that dramatically enhances the expression of foreign genes in *Escherichia coli*. Gene **73:** 227-235

Ong E, Pollock WB, Smith M (1997) Cloning and sequence analysis of two laccase complementary DNAs from the ligninolytic basidiomycete *Trametes versicolor*. Gene **196:** 113-119

Ortega JL, Moguel-Esponda S, Potenza C, Conklin CF, Quintana A, Sengupta-Gopalan C (2006) The 3' untranslated region of a soybean cytosolic glutamine synthetase (GS1) affects transcript stability and protein accumulation in transgenic alfalfa. Plant J **45:** 832-846

Outchkourov NS, Rogelj B, Strukelj B, Jongsma MA (2003) Expression of sea anemone equistatin in potato: effects of plant proteases on heterologous protein production. Plant Physiol. **133:** 379-390

Pütter J, Becker R (1983) Methods of Enzymatic Analysis (Bergmeyer, HU, ed.) 3rd ed., . *In* Verlug Chemie, Vol III. Deerfield Beach, FL., pp 286-293

Passardi F, Theiler G, Zamocky M, Cosio C, Rouhier N, Teixera F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L, Dunand C (2007) PeroxiBase: the peroxidase database. Phytochemistry **68:** 1605-1611

Patnaik D, Khurana P (2001) Wheat biotechnology: A minireview Journal of Biotechnology **4:** 4

Pawlowski WP, Somers DA (1998) Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. Proc Natl Acad Sci U S A **95:** 12106-12110

Peach C, Velten J (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol **17:** 49-60

Pellegrineschi A, Kis M, Dix I, Kavanagh TA, Dix PJ (1995) Expression of horseradish peroxidase in transgenic tobacco. Biochem Soc Trans **23:** 247-250

Piller KJ, Clemente TE, Jun SM, Petty CC, Sato S, Pascual DW, Bost KL (2005) Expression and immunogenicity of an *Escherichia coli* K99 fimbriae subunit antigen in soybean. Planta **222:** 6-18

Piontek K, Antorini M, Choinowski T (2002) Crystal structure of a laccase from the fungus *Trametes versicolor* at 1.90-A resolution containing a full complement of coppers. J Biol Chem **277:** 37663-37669

Plikaytis BD, Carlone GM, Edmonds P, Mayer LW (1986) Robust estimation of standard curves for protein molecular weight and linear-duplex DNA base-pair number after gel electrophoresis. Anal Biochem **152:** 346-364

Potenza C, Aleman L, Sengupta-Gopalan C (2004) Invited review: targeting transgene expression in research, agricultural, and environmental applications: promoters used in plant transformation. *In Vitro* Cellular and Developmental Biology - Plant **40:** 1-22

Praitis V (2006) Creation of transgenic lines using microparticle bombardment methods. *In* C. elegans, pp 93-107

Quemada H, Roth EJ, Lark KG (1987) Changes in methylation of tissue cultured soybean cells detected by digestion with the restriction enzymes *Hpa*II and *Msp*I. Plant Cell Reports **6:** 63-66

Rai M, Padh H (2001) Expression systems for production of heterologous proteins. Current Science **80:** 1121-1128

Rakotondrafara AM, Miller WA (2008) *In vitro* analysis of translation enhancers. Methods Mol Biol **451:** 113-124

Ranocha P, Chabannes M, Chamayou S, Danoun S, Jauneau A, Boudet AM, Goffner D (2002) Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. Plant Physiol **129:** 145-155

Rao MS, Bhagsari AS, Mohamed AI (1998) Yield, protein, and oil quality of soybean genotypes selected for tofu production. Plant Foods Hum Nutr **52:** 241-251

Reddy MS, Dinkins RD, Collins GB (2003) Gene silencing in transgenic soybean plants transformed via particle bombardment. Plant Cell Rep **21:** 676-683

Reed DG, Nopo-Olazabal LH, Funk V, Woffenden BJ, Reidy MJ, Dolan MC, Cramer CL, Medina-Bolivar F (2005) Expression of functional hexahistidine-tagged ricin B in tobacco. Plant Cell Reports **24:** 15-24

Regalado C, Asenjo JA, Pyle DL (1996) Studies on the purification of peroxidase from horseradish roots using reverse micelles. Enzyme Microb Technol **18:** 332-339

Robic G, Farinas CS, Rech EL, Bueno SMA, Miranda EA (2006) Downstream process engineering evaluation of transgenic soybean seeds as host for recombinant protein production. Biochemical Engineering Journal **32:** 7-12

Rodriguez-Pontes M (2007) Development of megagametophyte, embryo, and seed in *Senna corymbosa* (Lam.) H.S. Irwin & Barneby (*Leguminosae - Caesalpinioideae*). *In*, Vol 153, pp 169-179

Romano E, Soares A, Proite K, Neiva S, Grossi M, Faria JC, Rech EL, Aragao FJ (2005) Transgene elimination in genetically modified dry bean and soybean lines. Genet Mol Res **4:** 177-184

Rose AB, Last RL (1997) Introns act post-transcriptionally to increase expression of the *Arabidopsis thaliana* tryptophan pathway gene *PAT*1. Plant J **11:** 455-464

Russell DA, Spatola LA, Dian T, Paradkar VM, Dufield DR, Carroll JA, Schlittler MR (2005) Host limits to accurate human growth hormone production in multiple plant systems. Biotechnology and Bioengineering **89:** 775-782

Sailaja M, Tarakeswari M, Sujatha M (2008) Stable genetic transformation of castor (*Ricinus communis* L.) via particle gun-mediated gene transfer using embryo axes from mature seeds. Plant Cell Rep **27:** 1509-1519

Samadder P, Sivamani E, Lu J, Li X, Qu R (2008) Transcriptional and post-transcriptional enhancement of gene expression by the 5' UTR intron of rice *rubi*3 gene in transgenic rice cells. Mol Genet Genomics **279:** 429-439

Sambrook J, Fritsch E, Maniatis T (1989) Molecular cloning: a laboratory manual, Vol 3. Cold Spring Harbor Laboratory Press

Santagata S, Gomez CA, Sobacchi C, Bozzi F, Abinun M, Pasic S, Cortes P, Vezzoni P, Villa A (2000) N-terminal RAG1 frame shift mutations in Omenn's syndrome: internal methionine usage leads to partial V(D)J recombination activity and reveals a fundamental role *in vivo* for the N-terminal domains. Proc Natl Acad Sci U S A **97:** 14572-14577

Santos PM, Di Bartolo I, Blatny JM, Zennaro E, Valla S (2001) New broad-host-range promoter probe vectors based on the plasmid RK2 replicon. FEMS Microbiol Lett **195:** 91-96

Santos PO, Santos IS, Gomes VM, Machado OLT, Fernandes KVS, Xavier-Filho J, Oliveira AEA (2008) *In vitro* evaluation of antifungal activity of soybean (*Glycine max*) seed coat proteins. Journal of Stored Products Research **pubished online**

Sardana R, Dudani AK, Tackaberry E, Alli Z, Porter S, Rowlandson K, Ganz P, Altosaar I (2007) Biologically active human GM-CSF produced in the seeds of transgenic rice plants. Transgenic Res **16:** 713-721

Sato Y, Whetten R (2006) Characterization of two laccases of loblolly pine (*Pinus taeda*) expressed in tobacco BY-2 cells. Journal of Plant Research **119:** 581-588

Scheid OM, Paszkowski J, Potrykus I (1991) Reversible inactivation of a transgene in *Arabidopsis thaliana*. Molecular and General Genetics **228:** 104-112

Schillberg S, Zimmermann S, Voss A, Fischer R (1999) Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in *Nicotiana tabacum*. Transgenic Res **8:** 255-263

Serrato-Valenti G, De Vries M, Cornara L (1995) The hilar region in *Leucaena leucocephala Lam.* (De Wit) seed: structure, histochemistry and the role of the lens in germination. Ann Bot **75:** 569-574

Sessa D, Wolf W (2001) Bowman-Birk inhibitors in soybean seed coats. Industrial Crops and Products **14:** 73-83.

Shannon LM, Kay E, Lew JY (1966) Peroxidase isozymes from horseradish roots. I. Isolation and physical properties. J Biol Chem **241:** 2166-2172

Shpaer EG (1989) Amino acid composition is correlated with protein abundance in *Escherichia coli*: can this be due to optimization of translational efficiency? Protein Seq Data Anal **2:** 107-110

Siemering KR, Golbik R, Sever R, Haseloff J (1996) Mutations that suppress the thermosensitivity of green fluorescent protein. Current Biology **6:** 1653-1663

Singh J, Freeling M, Lisch D (2008) A position effect on the heritability of epigenetic silencing. PLoS Genet **4:** e1000216

Smith ML, Mason HS, Shuler ML (2002) Hepatitis B surface antigen (HBsAg) expression in plant cell culture: kinetics of antigen accumulation in batch culture and its intracellular form. Biotechnol Bioeng **80:** 812-822

Sonoki T, Kajita S, Ikeda S, Uesugi M, Tatsumi K, Katayama Y, Iimura Y (2005) Transgenic tobacco expressing fungal laccase promotes the detoxification of environmental pollutants. Applied Microbiology & Biotechnology **67:** 138-142

Sorensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli.* J Biotechnol **115:** 113-128

Souza F, Marcos-Filho J (2001) The seed coat as a modulator of seed-environment relationships in *Fabaceae.* **24:** 365-375

Stewart JR CN (2008) Plant biotechnology and genetics: principles, techniques, and applications. WILEY A John Wiley & Sons, Inc., Publication, Knoxville

Streatfield SJ (2007) Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnology Journal **5:** 2-15

Streatfield SJ, Jilka JM, Hood EE, Turner DD, Bailey MR, Mayor JM, Woodard SL, Beifuss KK, Horn ME, Delaney DE, Tizard IR, Howard JA (2001) Plant-based vaccines: unique advantages. Vaccine **19:** 2742-2748

Streatfield SJ, Lane JR, Brooks CA, Barker DK, Poage ML, Mayor JM, Lamphear BJ, Drees CF, Jilka JM, Hood EE, Howard JA (2003) Corn as a production system for human and animal vaccines. Vaccine **21:** 812-815

Sunilkumar G, Connell JP, Smith CW, Reddy AS, Rathore KS (2002) Cotton alpha-globulin promoter: isolation and functional characterization in transgenic cotton, *Arabidopsis*, and tobacco. Transgenic Res **11:** 347-359

Takamatsu K, Tachibana N, Matsumoto I, Abe K (2004) Soy protein functionality and nutrigenomic analysis. Biofactors **21:** 49-53

Tamura K, Shimada T, Ono E, Tanaka Y, Nagatani A, Higashi SI, Watanabe M, Nishimura M, Hara-Nishimura I (2003) Why green fluorescent fusion proteins have not been observed in the vacuoles of higher plants. Plant J **35:** 545-555

Tang W, Leisner S (1998) Methylation of nonintegrated multiple copy DNA in plants. Biochemical and Biophysical Research Communications **245:** 403-406

Taylor CB (1997) Promoter fusion analysis: an insufficient measure of gene expression. Plant Cell **9:** 273-275

Thorne JH (1981) Morphology and ultrastructure of maternal seed tissues of soybean in relation to the import of photosynthate. Plant Physiol. **67:** 1016-1025

Thurston P (1994) The structure and function of fungal laccases. Microbiology **140:** 19-26.

Travella S, Ross SM, Harden J, Everett C, Snape JW, Harwood WA (2005) A comparison of transgenic barley lines produced by particle bombardment and *Agrobacterium*-mediated techniques. Plant Cell Rep **23:** 780-789

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R (2003) Molecular farming in plants: host systems and expression technology. Trends Biotechnol **21:** 570-578

Tzfira T, Jensen CS, Wang W, Zuker A, Vinocur B, Altman A, Vainstein A (1997) Transgenic *Populus tremula*: a step-by-step protocol for its *Agrobacterium*-mediated transformation. Plant Molecular Biology Reporter **15:** 219-235

Van Dongen JT, Ammerlaan AM, Wouterlood M, Van Aelst AC, Borstlap AC (2003) Structure of the developing pea seed coat and the post-phloem transport pathway of nutrients. Ann Bot (Lond) **91:** 729-737

Velasquez MT, Bhathena SJ (2007) Role of dietary soy protein in obesity. Int J Med Sci **4:** 72-82

Verma R, Boleti E, George AJ (1998) Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems. J Immunol Methods **216:** 165-181

Verweire D, Verleyen K, De Buck S, Claeys M, Angenon G (2007) Marker-free transgenic plants through genetically programmed auto-excision. Plant Physiol. **145:** 1220-1231

Villaverde A, Carrio MM (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnol Lett **25:** 1385-1395

Vitale A, Denecke J (1999) The endoplasmic reticulum-gateway of the secretory pathway. Plant Cell **11:** 615-628

Vitale A, Hinz G (2005) Sorting of proteins to storage vacuoles: how many mechanisms? Trends Plant Sci **10:** 316-323

Vitale A, Pedrazzini E (2005) Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. Mol Interv **5:** 216-225

Vitale A, Raikhel NV (1999) What do proteins need to reach different vacuoles? Trends Plant Sci **4:** 149-155

Wang G, Xu Y (2008) Hypocotyl-based *Agrobacterium*-mediated transformation of soybean (*Glycine max*) and application for RNA interference. Plant Cell Rep **27:** 1177-1184

Wang GD, Li QJ, Luo B, Chen XY (2004) Ex planta phytoremediation of trichlorophenol and phenolic allelochemicals via an engineered secretory laccase. Nat Biotechnol **22:** 893-897

Wang HL, Grusak MA (2005) Structure and development of *Medicago truncatula* pod wall and seed coat. Ann Bot (Lond) **95:** 737-747

Wang J, Wang C, Zhu M, Yu Y, Zhang Y, Wei Z (2008) Generation and characterization of transgenic poplar plants overexpressing a cotton laccase gene Plant Cell Tiss Cult

Wang L, Webster DE, Campbell AE, Dry IB, Wesselingh SL, Coppel RL (2008) Immunogenicity of *Plasmodium yoelii* merozoite surface protein 4/5 produced in transgenic plants. International Journal for Parasitology **38:** 103-110

Ward TH, Brandizzi F (2004) Dynamics of proteins in Golgi membranes: comparisons between mammalian and plant cells highlighted by photobleaching techniques. Cell Mol Life Sci **61:** 172-185

Welinder KG (1979) Amino acid sequence studies of horseradish peroxidase: amino and carboxyl termini, cyanogen bromide and tryptic fragments, the complete sequence, and some structural characteristics of horseradish peroxidase C. Eur J Biochem **96:** 483-502

Welinder KG, Justesen AF, Kjaersgard IV, Jensen RB, Rasmussen SK, Jespersen HM, Duroux L (2002) Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. Eur J Biochem **269:** 6063-6081

Welinder KG, Larsen YB (2004) Covalent structure of soybean seed coat peroxidase. Biochim Biophys Acta **1698:** 121-126

Widsten P, Kandelbauer A (2008) Laccase applications in the forest products industry: A review. Enzyme and Microbial Technology **42:** 293-307

Wiley PR, Tosi P, Jones HD, Shewry PR (2008) The molecular basis for grain texture in wheat. 14th Australian Agronomy Conference: http://www.cropscience.org.au/icsc2004/poster/2003/2002/2001/2598_wileypr.htm

Windels P, Taverniers I, Depicker A, Van Bockstaele E, De Loose M (2001) Characterisation of the Roundup Ready soybean insert. European Food Research and Technology **213:** 107-112

Wu S, Druka A, Horvath H, Kleinhofs A, Gamini Kannangara C, von Wettstein D (2000) Functional characterization of seed coat-specific members of the barley germin gene family. Plant Physiology and Biochemistry **38:** 685-698

Xiao YZ, Tu XM, Wang J, Zhang M, Cheng Q, Zeng WY, Shi YY (2003) Purification, molecular characterization and reactivity with aromatic compounds of a laccase from basidiomycete *Trametes sp.* strain AH28-2. Appl Microbiol Biotechnol **60:** 700-707

Yaklich R, Vigil E, Wergin W (1998) Developmental changes of antipit cell ultrastructure in the soybean seed coat during seed maturation. Scanning Microscopy **12:** 523-532

Yamada Y, Nishizawa K, Yokoo M, Zhao H, Onishi K, Teraishi M, Utsumi S, Ishimoto M, Yoshikawa M (2008) Anti-hypertensive activity of genetically modified soybean seeds accumulating novokinin. Peptides **29:** 331-337

Yang L, Kajiura H, Suzuki K, Hirose S, Fujiyama K, Takaiwa F (2008) Generation of a transgenic rice seed-based edible vaccine against house dust mite allergy. Biochemical and Biophysical Research Communications **365:** 334-339

Yao Q, Cong L, Chang JL, Li KX, Yang GX, He GY (2006) Low copy number gene transfer and stable expression in a commercial wheat cultivar via particle bombardment. J Exp Bot **57:** 3737-3746

Yasuda H, Hayashi Y, Jomori T, Takaiwa F (2006) The correlation between expression and localization of a foreign gene product in rice endosperm. Plant Cell Physiol **47:** 756-763

Yaver DS, Xu F, Golightly EJ, Brown KM, Brown SH, Rey MW, Schneider P, Halkier T, Mondorf K, Dalboge H (1996) Purification, characterization, molecular cloning, and expression of two laccase genes from the white rot basidiomycete *Trametes villosa*. Appl Environ Microbiol **62:** 834-841

Yeung E (1990) Adhesion of endosperm cells to the inner surface of the bean seed coat. Journal of Structural Biology **105:** 103-110

Yoshida K, Kaothien P, Matsui T, Kawaoka A, Shinmyo A (2003) Molecular biology and application of plant peroxidase genes. Appl Microbiol Biotechnol **60:** 665-670

Yoshida K, Shinmyo A (2000) Transgene expression systems in plant, a natural bioreactor. J Biosci Bioeng **90:** 353-362

Yukawa K, Kaku H, Tanaka H, Koga-Ban Y, Fukuda M (2008) Enhanced soybean infection by the legume "super-virulent" Agrobacterium tumefaciens strain KAT23. Biosci Biotechnol Biochem **72:** 1809-1816

Zhao Y, Qian Q, Wang H, Huang D (2007) Hereditary behavior of bar gene cassette is complex in rice mediated by particle bombardment. J Genet Genomics **34:** 824-835

Zheng YY, Ling JQ, Mai S (2007) Expression of saliva-binding region of *Streptococcus* mutans pac in transgenic tomatoes. Hua Xi Kou Qiang Yi Xue Za Zhi **25:** 180-183

Zhou C, Labbe H, Sridha S, Wang L, Tian L, Latoszek-Green M, Yang Z, Brown D, Miki B, Wu K (2004) Expression and function of HD2-type histone deacetylases in *Arabidopsis* development. Plant J **38:** 715-724

Zubieta C, Krishna SS, Kapoor M, Kozbial P, McMullan D, Axelrod HL, Miller MD, Abdubek P, Ambing E, Astakhova T, Carlton D, Chiu HJ, Clayton T, Deller MC, Duan L, Elsliger MA, Feuerhelm J, Grzechnik SK, Hale J, Hampton E, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kumar A, Marciano D, Morse AT, Nigoghossian E, Okach L, Oommachen S, Reyes R, Rife CL, Schimmel P, van den Bedem H, Weekes D, White A, Xu Q, Hodgson KO, Wooley J, Deacon AM, Godzik A, Lesley SA, Wilson IA (2007) Crystal structures of two novel dye-decolorizing peroxidases reveal a beta-barrel fold with a conserved heme-binding motif. Proteins **69:** 223-233

Appendix 1 There Synthetic Genes of HRPC, LCC1, and LCC4 for Soybean Transformation Constructs

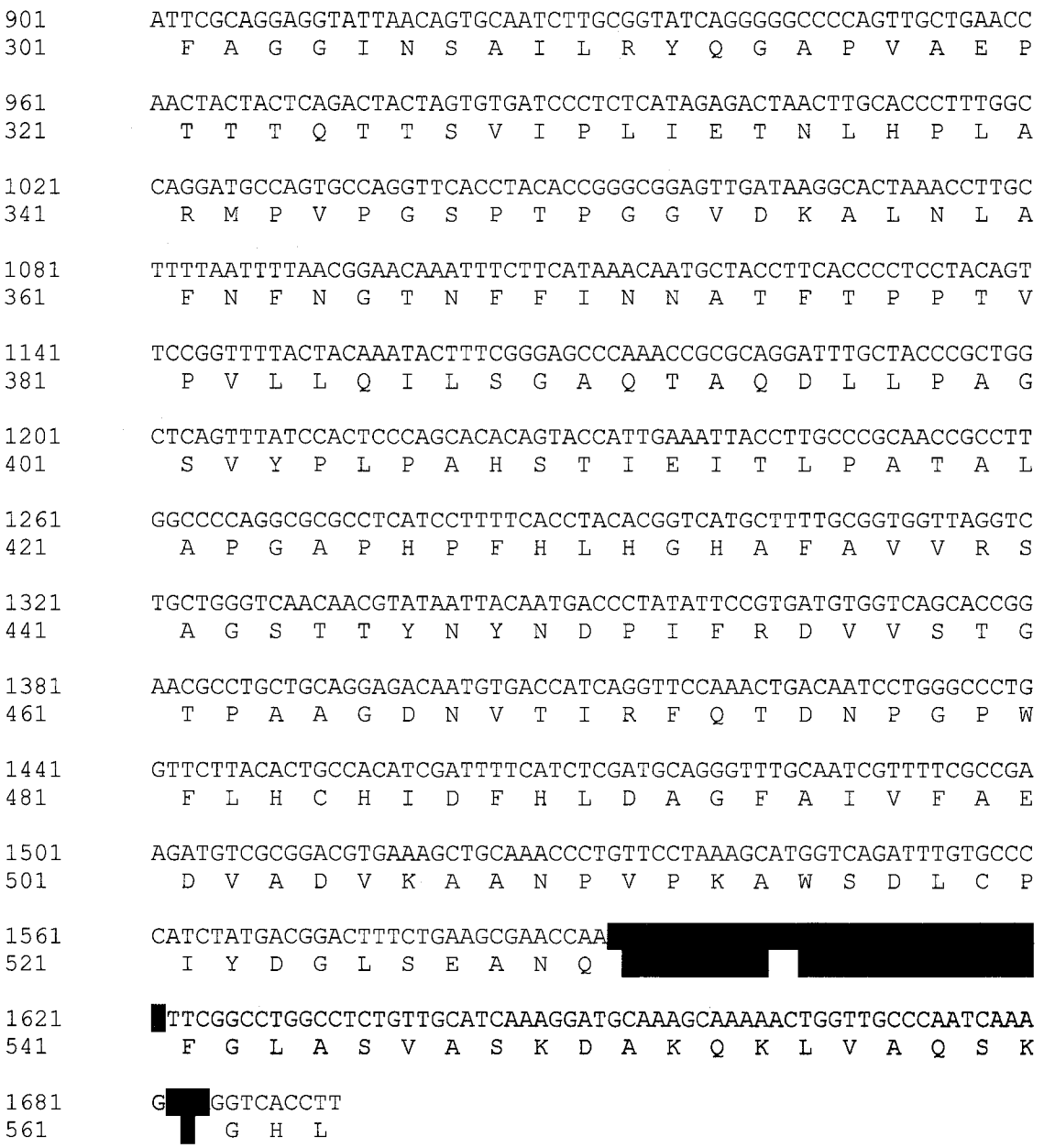
HRPC Gene Construct: The synthetic cDNA and derived protein sequence of HRPC codon optimized for soybean. The start codon is highlighted in green and the stop codon is highlighted in red. Both NTPP and CTPP of SBP gene are indicated in yellow. The adaptor GGAA, located in-between of signaling peptides and mature HRPC, is shown in pink. HIS tag is labeled in blue.

```
1          TTCC[redacted]GGTTCTATGAGGCTTTTGGTGGTTGCTTTGCTTTGTGCTTTTGCTATGCACGC
1            S [redacted] G  S  M  R  L  L  V  V  A  L  L  C  A  F  A  M  H  A

61         TGGTTTTTCTGTTTCATATGCT[redacted]CAATTGACTCCTACTTTTATTGATAA
21           G  F  S  V  S  Y  A [redacted] Q  L  T  P  T  F  I  D  N

121        CTCTTGCCCAAATGTTTCAAATATTGTTAGGGATACTATTGTGAACGAGTTGAGGTCTGA
41           S  C  P  N  V  S  N  I  V  R  D  T  I  V  N  E  L  R  S  D

181        TCCTAGGATTGCTGCTTCTATTCTTAGATTGCATTTCCATGATTGCTTCGTGAATGGTTG
61           P  R  I  A  A  S  I  L  R  L  H  F  H  D  C  F  V  N  G  C

241        CGATGCTTCTATTTTGCTTGATAACACAACTTCTTTCAGAACTGAGAAGGATGCTTTCGG
81           D  A  S  I  L  L  D  N  T  T  S  F  R  T  E  K  D  A  F  G

301        AAACGCTAATTCTGCTAGAGGATTCCCTGTTATTGATAGGATGAAGGCTGCTGTGGAGTC
101          N  A  N  S  A  R  G  F  P  V  I  D  R  M  K  A  A  V  E  S

361        AGCTTGCCCAAGAACTGTGTCATGTGCTGATCTTCTTACTATTGCTGCTCAACAATCTGT
121          A  C  P  R  T  V  S  C  A  D  L  L  T  I  A  A  Q  Q  S  V

421        GACATTGGCTGGAGGTCCTTCTTGGAGAGTTCCTTTGGGAAGGAGAGATTCACTTCAGGC
141          T  L  A  G  G  P  S  W  R  V  P  L  G  R  R  D  S  L  Q  A

481        TTTTCTTGATCTTGCTAATGCTAACTTGCCAGCTCCTTTCTTTACTCTTCCTCAACTTAA
161          F  L  D  L  A  N  A  N  L  P  A  P  F  F  T  L  P  Q  L  K

541        GGATTCTTTTAGAAACGTGGGTCTTAACAGATCATCTGATTTGGTGGCTTTGTCAGGTGG
181          D  S  F  R  N  V  G  L  N  R  S  S  D  L  V  A  L  S  G  G

601        ACATACATTTGGAAAGAATCAATGTAGATTCATTATGGATAGGCTTTATAATTTCTCTAA
201          H  T  F  G  K  N  Q  C  R  F  I  M  D  R  L  Y  N  F  S  N

661        TACAGGTTTGCCTGATCCAACTTTGAACACTACTTACTTGCAAACTCTTAGAGGATTGTG
221          T  G  L  P  D  P  T  L  N  T  T  Y  L  Q  T  L  R  G  L  C

721        CCCTCTTAACGGTAACTTGTCTGCTTTGGTGGATTTTGATCTTAGAACACCAACTATTTT
```


241 P L N G N L S A L V D F D L R T P T I F
781 TGATAACAAGTACTACGTTAATTTGGAAGAACAAAAGGGTTTGATTCAATCTGATCAAGA
261 D N K Y Y V N L E E Q K G L I Q S D Q E
841 ATTGTTCTCTTCACCAAATGCTACTGATACTATTCCTTTGGTTAGATCATTTGCTAACTC
281 L F S S P N A T D T I P L V R S F A N S
901 TACTCAAACATTCTTCAATGCTTTCGTTGAAGCTATGGATAGGATGGGTAATATTACTCC
301 T Q T F F N A F V E A M D R M G N I T P
961 TCTTACAGGTACTCAAGGTCAAATTAGATTGAATTGTAGAGTTGTTAATTCTAACTCT
321 L T G T Q G Q I R L N C R V V N S N S
1021 TTCGGACTTGCTTCTGTGGCTTCAAAGGATGC
341 F G L A S V A S K D A
1081 TAAGCAAAAGTTGGTGGCTCAATCTAAG GGTGACCTT
361 K Q K L V A Q S K G D L

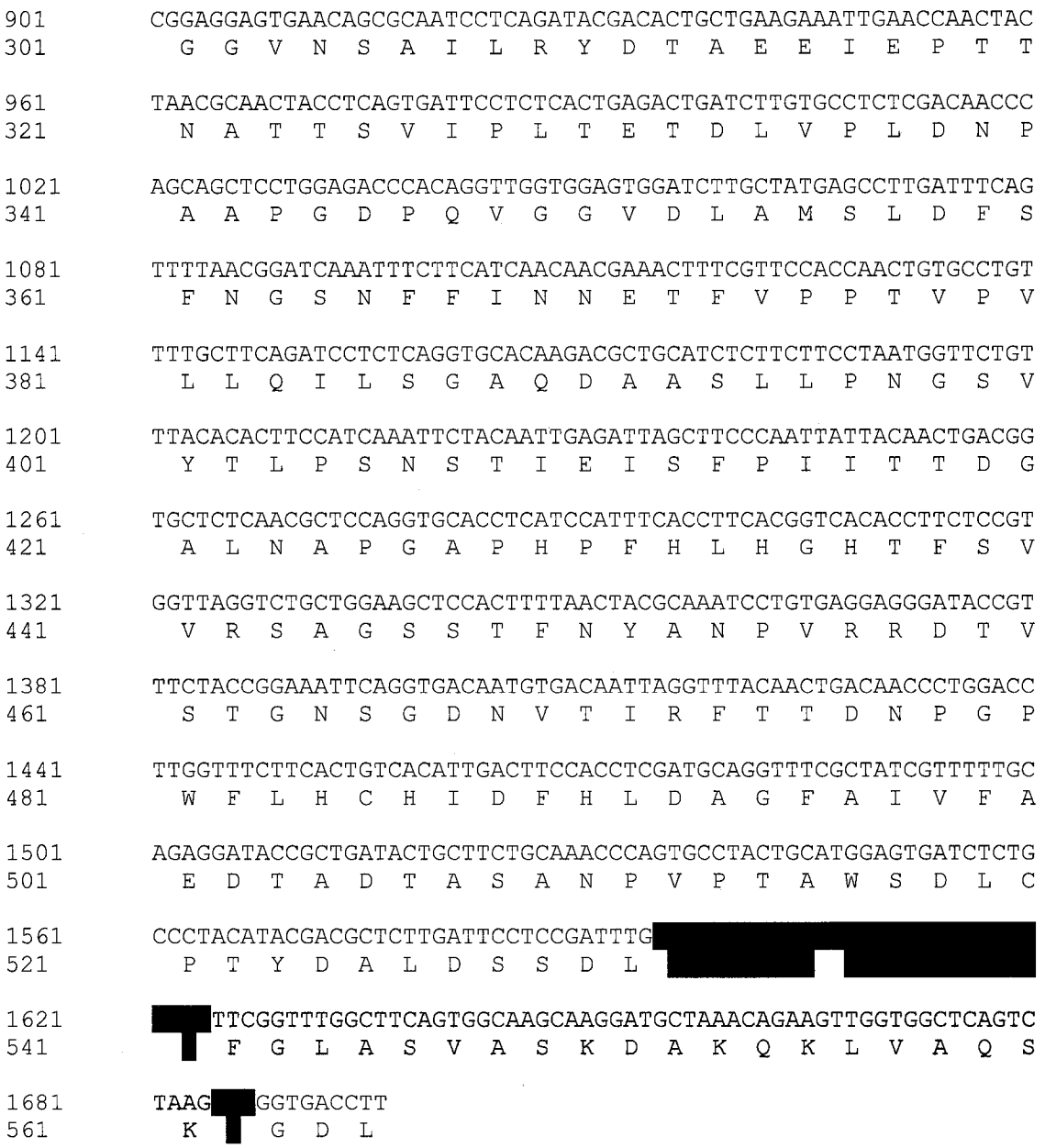
LCC1 Gene Construct: The synthetic cDNA and derived protein sequence of LCC1 codon optimized for soybean. The start codon is highlighted in green and the stop codon is highlighted in red. Both NTPP and CTPP of SBP gene are indicated in yellow. The adaptor GGAA, located in-between of signaling peptides and mature HRPC, is shown in pink. HIS tag is labeled in blue.

```
1          TTCC[redacted]GGTTCCATGCGCCTGCTGGTGGTGGCTTTGCTTTGTGCATTTGCCATGCACGC
1            S [redacted] G  S  M  R  L  L  V  V  A  L  L  C  A  F  A  M  H  A

61         TGGTTTCAGTGTTTCGTACGCC[redacted]GCTATCGGACCTGTGGCAAGCCTGGT
21           G  F  S  V  S  Y  A [redacted] A  I  G  P  V  A  S  L  V

121        TGTTGCTAATGCTCCAGTGTCCCCAGATGGGTTTCTTCGCGATGCAATAGTTGTAAATGG
41           V  A  N  A  P  V  S  P  D  G  F  L  R  D  A  I  V  V  N  G

181        TGTTGTTCCATCGCCACTCATTACAGGAAAGAAAGGAGATCGCTTTCAACTAAACGTTGT
61           V  V  P  S  P  L  I  T  G  K  K  G  D  R  F  Q  L  N  V  V

241        TGATACTCTTACAAACCATTCTATGCTTAAGTCAACCTCTATACACTGGCACGGGTTTTT
81           D  T  L  T  N  H  S  M  L  K  S  T  S  I  H  W  H  G  F  F

301        CCAAGCCGGAACTAACTGGGCAGATGGACCTGCGTTTGTTAACCAATGCCCTATTGCATC
101          Q  A  G  T  N  W  A  D  G  P  A  F  V  N  Q  C  P  I  A  S

361        AGGACATAGCTTTCTGTATGACTTCCATGTGCCAGATCAAGCAGGTACTTTTTGGTATCA
121          G  H  S  F  L  Y  D  F  H  V  P  D  Q  A  G  T  F  W  Y  H

421        CTCACATCTATCCACACAATATTGCGACGGACTACGCGGGCCATTTGTTGTATACGACCC
141          S  H  L  S  T  Q  Y  C  D  G  L  R  G  P  F  V  V  Y  D  P

481        AAAGGACCCTCATGCTTCCAGGTATGACGTGGACAATGAGTCAACTGTCATTACTCTTAC
161          K  D  P  H  A  S  R  Y  D  V  D  N  E  S  T  V  I  T  L  T

541        AGATTGGTATCATACTGCAGCAAGATTGGGGCCCCGCTTCCCTCTTGGCGCTGATGCCAC
181          D  W  Y  H  T  A  A  R  L  G  P  R  F  P  L  G  A  D  A  T

601        TCTTATTAATGGCTTAGGAAGGTCAGCTTCTACTCCAACTGCGGCTCTGGCCGTGATCAA
201          L  I  N  G  L  G  R  S  A  S  T  P  T  A  A  L  A  V  I  N

661        TGTGCAGCATGGAAAGAGATATAGATTCAGGTTGGTTTCAATTTCCTGTGATCCAAATTA
221          V  Q  H  G  K  R  Y  R  F  R  L  V  S  I  S  C  D  P  N  Y

721        TACATTCAGCATTGACGGGCATAACCTCACTGTTATAGAGGTTGACGGTATTAACTCACA
241          T  F  S  I  D  G  H  N  L  T  V  I  E  V  D  G  I  N  S  Q

781        GCCCTTACTTGTTGACTCAATACAGATTTTTGCAGCTCAACGCTATAGTTTTGTTCTCAA
261          P  L  L  V  D  S  I  Q  I  F  A  A  Q  R  Y  S  F  V  L  N

841        CGCTAATCAAACTGTTGGAAATTACTGGGTTCGAGCCAATCCAAACTTCGGTACGGTTGG
281          A  N  Q  T  V  G  N  Y  W  V  R  A  N  P  N  F  G  T  V  G
```


901 ATTCGCAGGAGGTATTAACAGTGCAATCTTGCGGTATCAGGGGGCCCCAGTTGCTGAACC
301 F A G G I N S A I L R Y Q G A P V A E P
961 AACTACTACTCAGACTACTAGTGTGATCCCTCTCATAGAGACTAACTTGCACCCTTTGGC
321 T T T Q T T S V I P L I E T N L H P L A
1021 CAGGATGCCAGTGCCAGGTTCACCTACACCGGGCGGAGTTGATAAGGCACTAAACCTTGC
341 R M P V P G S P T P G G V D K A L N L A
1081 TTTTAATTTTAACGGAACAAATTTCTTCATAAACAATGCTACCTTCACCCCTCCTACAGT
361 F N F N G T N F F I N N A T F T P P T V
1141 TCCGGTTTTACTACAAATACTTTCGGGAGCCCAAACCGCGCAGGATTTGCTACCCGCTGG
381 P V L L Q I L S G A Q T A Q D L L P A G
1201 CTCAGTTTATCCACTCCCAGCACACAGTACCATTGAAATTACCTTGCCCGCAACCGCCTT
401 S V Y P L P A H S T I E I T L P A T A L
1261 GGCCCCAGGCGCGCCTCATCCTTTTCACCTACACGGTCATGCTTTTGCGGTGGTTAGGTC
421 A P G A P H P F H L H G H A F A V V R S
1321 TGCTGGGTCAACAACGTATAATTACAATGACCCTATATTCCGTGATGTGGTCAGCACCGG
441 A G S T T Y N Y N D P I F R D V V S T G
1381 AACGCCTGCTGCAGGAGACAATGTGACCATCAGGTTCCAAACTGACAATCCTGGGCCCTG
461 T P A A G D N V T I R F Q T D N P G P W
1441 GTTCTTACACTGCCACATCGATTTTCATCTCGATGCAGGGTTTGCAATCGTTTTCGCCGA
481 F L H C H I D F H L D A G F A I V F A E
1501 AGATGTCGCGGACGTGAAAGCTGCAAACCCTGTTCCTAAAGCATGGTCAGATTTGTGCCC
501 D V A D V K A A N P V P K A W S D L C P
1561 CATCTATGACGGACTTTCTGAAGCGAACCAA
521 I Y D G L S E A N Q
1621 TTCGGCCTGGCCTCTGTTGCATCAAAGGATGCAAAGCAAAAACTGGTTGCCCAATCAAA
541 F G L A S V A S K D A K Q K L V A Q S K
1681 GGGTCACCTT
561 G H L

LCC4 Gene Construct: The synthetic cDNA and derived protein sequence of LCC4 codon optimized for soybean. The start codon is highlighted in green and the stop codon is highlighted in red. Both NTPP and CTPP of SBP gene are indicated in yellow. The adaptor GGAA, located in-between of signaling peptides and mature HRPC, is shown in pink. HIS tag is labeled in blue.

```
1          TTCC[redacted]GGTAGCATGAGACTTCTCGTTGTGGCTTTGTTGTGCGCTTTTGCTATGCACGC
1            S [redacted] G  S  M  R  L  L  V  V  A  L  L  C  A  F  A  M  H  A

61         TGGTTTCTCAGTTAGCTATGCT[redacted]GTTACTGATCTTACAATTTCCAATGC
21           G  F  S  V  S  Y  A [redacted] V  T  D  L  T  I  S  N  A

121        TGATGTGACACCTGATGGTATTACAAGGGCTGCTGTTTTGGCTGGTGGAGTGTTCCCAGG
41           D  V  T  P  D  G  I  T  R  A  A  V  L  A  G  G  V  F  P  G

181        TCCTTTGATTACCGGAAACAAAGGTGACGAGTTCCAAATTAATGTGATTGACAACCTCAC
61           P  L  I  T  G  N  K  G  D  E  F  Q  I  N  V  I  D  N  L  T

241        AAACGAGACCATGTTGAAATCAACTACAATTCATTGGCATGGAATTTTCCAAGCAGGAAC
81           N  E  T  M  L  K  S  T  T  I  H  W  H  G  I  F  Q  A  G  T

301        TAACTGGGCAGACGGAGCTGCATTTGTTAACCAATGCCCAATTGCTACAGGAAATAGTTT
101          N  W  A  D  G  A  A  F  V  N  Q  C  P  I  A  T  G  N  S  F

361        TCTTTACGATTTCACAGTGCCAGATCAAGCTGGAACTTTCTGGTATCATTCTCATCTCTC
121          L  Y  D  F  T  V  P  D  Q  A  G  T  F  W  Y  H  S  H  L  S

421        AACACAGTATTGCGATGGTTTGAGGGGTCCTCTTGTGGTTTATGATCCAGATGACGCTAA
141          T  Q  Y  C  D  G  L  R  G  P  L  V  V  Y  D  P  D  D  A  N

481        CGCTTCCCTCTACGACGTGGACGATGACACCACTGTTATCACTCTCGCTGATTGGTATCA
161          A  S  L  Y  D  V  D  D  D  T  T  V  I  T  L  A  D  W  Y  H

541        TACAGCAGCTAAGTTGGGACCTGCTTTCCCAGCTGGACCAGACTCAGTTCTTATTAATGG
181          T  A  A  K  L  G  P  A  F  P  A  G  P  D  S  V  L  I  N  G

601        TTTGGGAAGGTTTTCAGGAGATGGAGGTGGTGCAACTAATCTTACCGTTATCACTGTTAC
201          L  G  R  F  S  G  D  G  G  G  A  T  N  L  T  V  I  T  V  T

661        CCAGGGAAAAAGATATAGGTTCAGACTCGTGTCAATCTCTTGCGACCCAAATTTCACTTT
221          Q  G  K  R  Y  R  F  R  L  V  S  I  S  C  D  P  N  F  T  F

721        CTCAATTGATGGACATAATATGACCATTATTGAGGTTGGAGGAGTGAATCACGAAGCTCT
241          S  I  D  G  H  N  M  T  I  I  E  V  G  G  V  N  H  E  A  L

781        TGACGTGGATAGCATTCAGATTTTCGCAGGTCAGAGATACTCATTCATTTTGAATGCTAA
261          D  V  D  S  I  Q  I  F  A  G  Q  R  Y  S  F  I  L  N  A  N

841        CCAGTCAATTGATAATTACTGGATCAGGGCTATCCCTAATACTGGAACCACCGATACCAC
281          Q  S  I  D  N  Y  W  I  R  A  I  P  N  T  G  T  T  D  T  T
```


901 CGGAGGAGTGAACAGCGCAATCCTCAGATACGACACTGCTGAAGAAATTGAACCAACTAC
301 G G V N S A I L R Y D T A E E I E P T T
961 TAACGCAACTACCTCAGTGATTCCTCTCACTGAGACTGATCTTGTGCCTCTCGACAACCC
321 N A T T S V I P L T E T D L V P L D N P
1021 AGCAGCTCCTGGAGACCCACAGGTTGGTGGAGTGGATCTTGCTATGAGCCTTGATTTCAG
341 A A P G D P Q V G G V D L A M S L D F S
1081 TTTTAACGGATCAAATTTCTTCATCAACAACGAAACTTTCGTTCCACCAACTGTGCCTGT
361 F N G S N F F I N N E T F V P P T V P V
1141 TTTGCTTCAGATCCTCTCAGGTGCACAAGACGCTGCATCTCTTCTTCCTAATGGTTCTGT
381 L L Q I L S G A Q D A A S L L P N G S V
1201 TTACACACTTCCATCAAATTCTACAATTGAGATTAGCTTCCCAATTATTACAACTGACGG
401 Y T L P S N S T I E I S F P I I T T D G
1261 TGCTCTCAACGCTCCAGGTGCACCTCATCCATTTCACCTTCACGGTCACACCTTCTCCGT
421 A L N A P G A P H P F H L H G H T F S V
1321 GGTTAGGTCTGCTGGAAGCTCCACTTTTAACTACGCAAATCCTGTGAGGAGGGATACCGT
441 V R S A G S S T F N Y A N P V R R D T V
1381 TTCTACCGGAAATTCAGGTGACAATGTGACAATTAGGTTTACAACTGACAACCCTGGACC
461 S T G N S G D N V T I R F T T D N P G P
1441 TTGGTTTCTTCACTGTCACATTGACTTCCACCTCGATGCAGGTTTCGCTATCGTTTTTGC
481 W F L H C H I D F H L D A G F A I V F A
1501 AGAGGATACCGCTGATACTGCTTCTGCAAACCCAGTGCCTACTGCATGGAGTGATCTCTG
501 E D T A D T A S A N P V P T A W S D L C
1561 CCCTACATACGACGCTCTTGATTCCTCCGATTTG
521 P T Y D A L D S S D L
1621 TTCGGTTTGGCTTCAGTGGCAAGCAAGGATGCTAAACAGAAGTTGGTGGCTCAGTC
541 F G L A S V A S K D A K Q K L V A Q S
1681 TAAG GGTGACCTT
561 K G D L

Appendix 2 HRPC Enzyme Activity Assay Protocol

Protein extracts were prepared from plant material (seed coats, embryos, leaves) by using 0.067 M phosphate buffer (0.144 g $Na_2HPO_4 \cdot 2H_2O$ + 0.798 g KH_2PO_4 in 100 ml deionized water, pH 6.0) as described in Section 2.8.2.

The reaction conditions were as follows: temperature 25°C, pH 6.0, wavelength 405 nm, and light path, 10 mm. The reagents were prepared fresh daily, which included phosphate buffer 0.067 M, ABTS 2 x 10^{-2} M (one tablet dissolved in 912µl of ddH_2O), and H_2O_2 10^{-2} M (1µl of 30% H_2O_2 was added into 999µl of ddH_2O), and protein sample (0.20-0.80 U/ml). The spectrophotometer used for measuring the continuous absorbance at 405 nm for 5 min was Spectronic BioMate 3.

Temperature of solutions was adjusted to 25°C just prior to beginning. At all other times, solutions were kept at approximately 4°C. The recipes for blank control and test cuvettes were as follows. Blank control cuvette contained 0.05 ml of 2 x 10^{-2} M ABTS, 0.05 ml of 10^{-2} M H_2O_2, and 0.5 ml of 0.067 M phosphate buffer; while sample cuvette contained 0.05 ml of 2 x 10^{-2} M ABTS, 0.05 ml of 10^{-2} M H_2O_2, and 0.5 ml of sample in phosphate buffer. The reaction solution was mixed by pipetting. The changes in A_{405} over the next approximately 6 minutes was monitored and recorded from the addition of the sample. For the calculations, the ΔA_{405} was taken during the maximum linear rate, which occurred shortly (usually 20 sec) after the addition of the sample. The ΔA_{405} for the blank was subtracted from that of the test. The specific activity was calculated using the equation below.

$$\text{Specific Activity} \quad (\mu\text{mol min}^{-1}\,\text{mg}^{-1}\text{ or U mg}^{-1}) = \frac{\Delta A_{405} \cdot V_A \cdot 1000 \cdot df}{\varepsilon \cdot d \cdot \Delta t \cdot v_E \cdot \rho_{protein} \cdot 2^*}$$

ΔA_{405} was the change in absorbance of the test minus the change in the blank during the maximum linear rate. V_A was the total volume of the assay in litres, which was 0.6×10^{-3} L. df was the dilution factor of the enzyme sample. ε was the linear millimolar absorption coefficient. For ABTS under these temperature and pH conditions, it was 1.86 L $mmol^{-1}$ mm^{-1} (Pütter and Becker, 1983). d was the path length in millimetres, standard cuvettes were 10 mm. Δt was the change in 5 min during which $\Delta A405$ was recorded. v_E is the volume of the sample in litres, which was 0.5×10^{-3} L. $\rho_{protein}$ was the mass concentration of the enzyme sample, in mg/L.

*The denominator was multiplied by two because 1 mole of H_2O_2 yielded 2 moles of oxidized ABTS.

Based on the above procedure, the equation was reduced to:

$$\text{Specific Activity} \quad (\mu\text{mol min}^{-1}\,\text{mg}^{-1}\text{ or U mg}^{-1}) = 32.26 \quad \frac{\Delta A_{405} \cdot df}{\Delta t \cdot \rho_{protein}}$$

Appendix 3 Laccase Enzyme Activity Assay Protocol

Protein extracts were prepared from plant material (seed coats, embryos, leaves) by using 0.1 M sodium acetate pH 5.0 (1.36 g sodium acetate in 100 ml deionized water, pH 5.0), detailed protocol for plant protein extraction is in section 2.8.2.

Copper treatment was carried out on clarified supernatant of seed coat protein by adding 10 mM $CuSO_4$ from concentrated 1M stock solution in H_2O. Samples were incubated in the treatment solution, typically at either room temperature, for about 1 h. Incubation with metal salts causes up to 90% of the total soluble protein to precipitate (Hood et al., 2003). Therefore, activated solutions were centrifuged for 10 min at 10, 000 g to remove this precipitate, and then analysed for laccase activity.

The reaction conditions were as follows: temperature 25°C, pH 6.0, wavelength 420 nm, and light path, 1 cm. The reagents were prepared fresh daily, which included 0.1 M sodium acetate pH 5.0, 5 mM ABTS (one tablet dissolved in 3.65 ml of ddH_2O), and protein sample. The spectrophotometer used for measuring the continuous absorbance at 420 nm for 5 min was Spectronic BioMate 3.

Temperature of solutions was adjusted to 25°C just prior to beginning. At all other times, solutions were kept at approximately 4°C. The recipes for blank control and test cuvettes were as follows. Blank control cuvette contained 0.1 ml of 5 mM ABTS and 0.4 ml of 0.1 M sodium acetate pH 5.0; while sample cuvette contained 0.1 ml of 5 mM ABTS, and 0.4 ml of sample in phosphate buffer. The reaction solution was mixed by pipetting, and then the reaction solution was incubated at room temperature until the color changed into light green. The changes in A_{420} over the next approximately 5 minutes was monitored and recorded from the addition of the sample. For the

calculations, the ΔA_{420} was taken during the maximum linear rate. The ΔA_{420} for the blank was subtracted from that of the test. The specific activity was calculated using the equations below.

$$\text{Specific Activity } (\mu\text{mol min}^{-1}\text{ mg}^{-1}\text{ or U mg}^{-1}) = \frac{\Delta A_{420} \bullet \text{Vtotal}}{\varepsilon \bullet d \bullet \Delta t \bullet v_{sample} \bullet \rho_{protein}}$$

ΔA_{420} was the change in absorbance of the test minus the change in the blank during the maximum linear rate. V_{total} was the total volume of the assay in litres, which was 0.5 mL. ε was the linear millimolar absorption coefficient. For ABTS under these temperature and pH conditions, the extinction coefficient of ABTS was 36.8 L $mMol^{-1}$ cm^{-1} (Mander et al., 2006). d was the path length in centimetre, standard cuvettes were 10 cm. Δt was the change in 5 min during which ΔA_{420} was recorded. v_{sample} is the volume of the sample in litres, which was 0.05 mL. $\rho_{protein}$ was the mass concentration of the enzyme sample, in mg/L..

Based on the above procedure, the equation was reduced to:

$$\text{Specific Activity } (\mu\text{mol min}^{-1}\text{ mg}^{-1}\text{ or U mg}^{-1}) = 271.73 \frac{\Delta A_{420}}{\Delta t \bullet \rho_{protein}}$$

Appendix 4, HRPC Transgenic Soybean Seed Screening by Enzyme Assay

		HRPC Activity of Individual Seed Coats (U/mg)																
Lines	#Seeds tested	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Average**	**SD**
Jack	10	0.08	0.05	0.09	0.05	0.04	0.04	0.05	0.03	0.04	0.03						0.05	0.01
S8 HRPC #1	4	0.09	0.23	0.26	0.16												0.19	0.04
S8 HRPC #1-1	5	1.3	0.88	1.64	2.66	0.82											1.46	0.33
S8 HRPC #1-1	10	0.16	0.16	0.09	0.27	0.24	0.14	0.18	0.06	0.09	0.05						0.14	0.02
S8 HRPC #1-1	9	0.06	0.15	0.11	0.22	0.03	0.05	0.14	0.18	0.21							0.13	0.02
S8 HRPC #1-1	14	0.33	0.23	0.29	0.72	0.27	0.17	0.37	0.81	0.31	1.95	0.08	0.35	0.38	0.15		0.46	0.13
S8 HRPC #1-2	5	1.29	0.78	0.18	0.54	1.08											0.77	0.2
S8 HRPC #3-2	2	0.68	0.25														0.47	---
S8 HRPC #3-2-1	6	0.48	0.29	0.22	0.17	0.11	0.65										0.32	0.08
S8 HRPC #3-2-1	5	0.68	2.74	0.25	3.09	1.32											1.62	0.56
S8 HRPC #3-2-2	3	1.11	1.54	1.12													1.26	0.14
S8 HRPC #3-4	4	0.45	0.66	0.43	0.36												0.48	0.06
S8 HRPC 3-4-1	5	0.22	0.24	0.14	0.26	0.23											0.22	0.02
S8 HRPC 3-4-2	3	0.45	0.6	0.64	0.18												0.47	0.1
S8 HRPC 3-4-3	2	0.61	0.42														0.52	---
S8 HRPC #3-5-1	3	0.31	0.25	0.39													0.32	0.04
S8 HRPC #3-5-2	5	0.1	0.35	0.21	0.36	0.27											0.26	0.05
S8 HRPC #3-6	15	0.33	0.27	0.29	0.09	0.02	0.67	0.32	0.56	0.45	0.04	0.05	0.47	0.23	0.55	0.36	0.31	0.05
S8 HRPC #3-6-1	4	0.26	0.28	0.31	0.48												0.33	0.05
S8 HRPC #3-6-2	3	0.73	0.43	0.49													0.55	0.09
S8 HRPC #3-6-3	3	0.72	1.86	0.5													1.03	0.42
S8 HRPC #3-8	1	0.55															0.55	---
S8 HRPC #3-8-1	5	1.02	0.41	0.46	0.48	0.28											0.53	0.13
S8 HRPC #3-8-2	5	0.36	0.47	1.33	0.95	0.36											0.69	0.19
S8 HRPC #3-8-2-4	7	0.647	0.561	0.376	0.354	0.295	0.603	0.209									0.44	0.06
S8 HRPC #3-10-2	4	0.62	0.25	0.24	0.22												0.33	0.1
S8 HRPC #3-11	4	0.25	0.68	0.33	0.04												0.33	0.13
S8 HRPC #3-11-1	7	1.04	0.13	0.67	1.41	0.72	0.87	2.2									1.01	0.25
S8 HRPC #3-11-1-4	6	1.34	0.77	0.62	0.80	0.83	0.56										0.62	1.34
S8 HRPC #3-11-2	2	1.01	0.28														0.65	---
S8 HRPC #5	13	0.29	0.33	0.45	0.46	0.22	0.05	0.06	0.03	0.35	0.29	0.44	0.04	0.02			0.23	0.05
S8 HRPC #5-1	3	0.39	0.68	0.56													0.54	0.08
S8 HRPC #13	15	0.12	0.09	0.25	0.2	0.36	0.35	0.4	0.22	0.16	0.05	0.03	0.34	0.28	0.27	0.12	0.22	0.03
S8 HRPC #14	6	0.18	0.35	0.3	0.14	0.13	0.1										0.2	0.04
S8 HRPC #15-1	5	0.05	0.06	0.03	0.07	0.08											0.06	0.01
S8 HRPC #16	2	0.45	0.12														0.29	---
Total seed # tested	**202**																	

Appendix 5 LCC Transgenic Soybean Seed Screening by Enzyme Assay

Lines	# seed tested	LCC Activity Individual Seed Coats (U/mg) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Average	SD
WT (Jack)	10	0	0.015	0.015	0.01	0	0.001	0.002	0.02	0.021	0.016						0.01	0.003
S8 Vector																		
S8LCC4 # 1	10	0.02	0.03	0.01	0	0.008	0.04	0.03	0.023	0.025	0.019	0.021	0.018				0.02	0.003
S8LCC4 # 2	8	0.001	0	0.008	0.011	0.019	0	0.005	0.023								0.008	0.003
S8LCC4 # 3	1	0.03															0.03	---
S8LCC4 # 5	10	0.041	0.04	0.037	0.009	0.03	0.04	0.034	0.023	0.043	0.034						0.033	0.003
S8LCC4 # 9-2	2	0.02	0.031														0.026	---
S8LCC4 # 11	2	0.04	0.002														0.021	---
S8LCC4 # 13	3	0	0	0.006													0.002	0.002
S8LCC4 # 14	3	0	0	0.005													0.002	0.002
S8LCC4 # 15-2	2	0.04	0.03														0.035	0.005
S8LCC4 # 15-4	1	0.04															0.04	---
S8LCC4 # 15-6	1	0.005															0.005	---
S8LCC4 # 16	2	0.03	0.04														0.035	---
S8LCC4 # 16-1	13	0.099	0.03	0.02	0.03	0	0	0.01	0.03	0.05	0.04	0.034	0.05	0.005			0.031	0.007
S8LCC4 # 16-1-1	11	0.019	0.035	0.009	0.021	0.014	0.050	0.013	0.014	0.016	0.048	0.031					0.025	0.004
S8LCC4 # 17-1	12	0	0.03	0.02	0.006	0.025	0.03	0.03	0.02	0	0.003	0.003	0.032				0.017	0.004
S8LCC4 # 17-2	5	0.036	0.02	0.008	0.001	0.008											0.015	0.006
S8LCC4 # 18	13	0.009	0.007	0.001	0	0.01	0.01	0.02	0.023	0.021	0.012	0.01	0.009	0.012			0.011	0.002
S8LCC4 # 19-1	10	0	0	0.004	0.012	0.01	0	0	0.015	0.015	0.019						0.008	0.002
S8LCC4 # 19-2	10	0	0.003	0.004	0.005	0.001	0	0	0.009	0.014	0.016						0.005	0.002
S8LCC4 # 21	9	0.03	0.02	0.01	0.018	0.02	0.031	0.032	0.009	0							0.019	0.004
S8LCC4 # 24	8	0.051	0.02	0.01	0.066	0.002	0.008	0.02	0.01								0.023	0.008
S8LCC4 # 26	7	0	0.03	0.04	0.007	0.02	0.03	0.032									0.023	0.005
S8LCC4 # 27	15	0	0	0	0.008	0.003	0.04	0	0	0.01	0.008	0.005	0.013	0.02	0	0.023	0.009	0.003
S8LCC4 # 30-1	4	0.02	0.04	0.01	0.023												0.023	0.006
S8LCC4 # 30-1-4	4	0.04	0.01	0.034	0.028												0.028	0.006
S8LCC4 # 30-1-4-2	10	0.039	0.032	0.023	0.034	0.033	0.027	0.031	0.046	0.047	0.063						0.037	0.004
S8LCC4 # 31	10	0.03	0	0.02	0.008	0.005	0.005	0.008	0.03	0.02	0.03						0.016	0.004
S8LCC4 # 32-1	6	0.05	0.008	0.03	0	0.04	0.034										0.027	0.008

S8LCC4 # 32-2	15	0.052	0.03	0.002	0.005	0.03	0.042	0.034	0.009	0	0	0.032	0.029	0.001	0.008	0	0.018	0.005
S8LCC4 # 33	1	0.01															0.01	---
S8LCC1 # 4-1	8	0.033	0.035	0.022	0.028	0.01	0.031	0.034	0.031								0.028	0.003
S8LCC1 # 4-1-1	3	0.012	0.009	0.011													0.011	0.001
S8LCC1 # 4-2	7	0.035	0.023	0.033	0.035	0.029	0.029	0.032									0.031	0.002
S8LCC1 # 4-2-2	5	0.038	0.04	0.026	0.037	0.021											0.032	0.004
S8LCC1 # 4-2-2-1	10	0.018	0.014	0.026	0.022	0.020	0.050	0.028	0.027	0.044	0.017						0.027	0.004
S8LCC1 # 5	15	0.02	0.03	0.004	0.033	0.022	0.03	0.04	0.045	0.031	0.033	0.028	0.022	0.04	0.033	0.029	0.029	0.003
S8LCC1 # 5-1	4	0.01	0.027	0.007	0.009												0.014	0.005
S8LCC1 # 5-1-2	13	0.017	0.044	0.047	0.046	0.049	0.049	0.022	0.020	0.048	0.011	0.044	0.003	0.012			0.032	0.005
S8LCC1 # 10-2	11	0	0	0	0	0	0.004	0	0.005	0.004	0.001	0.002					0.001	0.001
S8LCC1 # 12-2	8	0	0	0	0.005	0.001	0	0	0								0.001	0.001
S18 vector																		
S8LCC1#3A	15	0.015	0	0.03	0.04	0.045	0	0.033	0.028	0.035	0	0	0.034	0.024	0.034	0.028	0.023	0.004
S8LCC1#2B	5	0	0.005	0.007	0.004	0.004											0.004	0.001
S8LCC1#3B	11	0	0	0	0.005	0	0	0.002	0	0	0	0					0.001	0
S8LCC1#2C	5	0.037	0.04	0.037	0.018	0.041											0.035	0.004
S8LCC1#3C	15	0.001	0.003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S8LCC1#2E	3	0	0	0													0	0
S8LCC1#2F	15	0.004	0.006	0	0	0	0	0.001	0	0	0	0	0	0	0	0	0.001	0
S8LCC1#2G	15	0.035	0.023	0.004	0.015	0.04	0.05	0.034	0.022	0	0.035	0.022	0.03	0.022	0.025	0.021	0.025	0.003
S8LCC1#2J	14	0	0	0	0	0.001	0.001	0.006	0.007	0.006	0.008	0.01	0.009	0.008	0.009		0.005	0.001
S8LCC1#2H	15	0	0	0	0	0	0	0.005	0.001	0.001	0.002	0	0	0	0	0	0.001	0
S8LCC1#2I	12	0.034	0.033	0.028	0.03	0.022	0.034	0.028	0.033	0.023	0.033	0.03	0.032				0.03	0.001
S8LCC4#1A	15	0.001	0.03	0.05	0	0.005	0.004	0	0	0	0	0	0	0	0.001	0	0.006	0.004
S8LCC4#1B	6	0.004	0.005	0.005	0.006	0.005	0.005										0.005	0
S8LCC4#1C	11	0.02	0.022	0.01	0	0.005	0.004	0	0.02	0.005	0.001	0.023					0.01	0.003
S8LCC4#1D	15	0.034	0.033	0.022	0.034	0.04	0	0.03	0.032	0.036	0.031	0.028	0.033	0.037	0.031	0.029	0.03	0.002
Total seed tested	414																	