FUNCTIONAL EXPRESSION OF *IPS PARACONFUSUS* CYTOCHROMES P450 EXHIBITING CONSPICUOUS UP-REGULATION WITH FEEDING ON HOST PHLOEM

by

Caitlin Pitt

B.Sc. University of Northern British Columbia, 2007

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
NATURAL RESOURCES AND ENVIRONMENTAL STUDIES
IN BIOLOGY

THE UNIVERSITY OF NORTHERN BRITISH COLUMBIA

May 2010

Your file Votre référence
ISBN: 978-0-494-60843-2
Our file Notre référence
ISBN: 978-0-494-60843-2

Abstract

Owing to the possible roles of insect P450s in detoxification of host material, and the biosynthesis or degradation of aggregation pheromones and hormones, these enzymes are ecologically important. Several *Ips paraconfusus* cytochromes P450 (P450s) were up-regulated with feeding on host phloem. Beetle P450s have previously been functionally characterized via heterologous expression in baculovirus/insect cell systems. *Ips paraconfusus* P450s were similarly expressed as proteins, and co-expressed with a P450 reductase. Capillary electrophoresis quantified recombinant enzymes during optimization efforts, augmenting spectral measurements. Various host secondary metabolites were tested as substrates, but functions were not determined, which may mean that the enzymes are not specific to those compounds. Although these P450s were microsomal, cellular fraction evaluation revealed possible mitochondrial and cytosolic targeting and two distinct size groups. These may represent the first insect examples of chimeric P450 signaling. The utility of capillary electrophoresis to streamline expression optimization and alternate targeting are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgement

I would first like to thank my mom, Shawn, and my dad, Tim, as well as Jeff for their support and encouragement over the years while I have pursued my education. I am grateful to my supervisor (Dr. Dezene Huber) and committee members (Dr. Daniel Erasmus and Dr. Hugues Massicotte) for their invaluable support, understanding, contributions, and perspectives. I appreciate the efforts of the other students in the lab, particularly Amy Thommasen and Sophie McClair, who have assisted me along the way. Finally, I would like to thank those of you who have patiently listened to me talk through problems involving proteins, especially the hypothetical ones.

I would like to acknowledge the following groups for their financial contributions: the Canada Research Chairs Program, the Canada Foundation for Innovation, the BC Knowledge Development Fund, the Natural Sciences and Engineering Research Council of Canada, the University of Northern British Columbia, and the Canfor Corporation.

1.0 Introduction

1.1 Bark Beetle Ecology

There are over 6000 species of bark beetles worldwide; virtually all trees and some woody and non-woody plants are susceptible to at least one species (Wood & Storer, 2002). In addition to their roles in plant decomposition and nutrient cycling, some bark beetles influence the life cycles of some *Pinus* species, such as lodgepole pine (*P. contorta* Dougl. ex Loud.). At the stand level, in lodgepole pine forests, bark beetles can influence fire disturbances; contribute to clearing out older, more susceptible trees so that a new generation of seedlings may repopulate the area (Wood & Storer, 2002; Logan & Powell, 2001); or help to prevent other competitive species from dominating the stand. Fire may contribute to the establishment of differently aged patches of lodgepole pine (Kauffman *et al.*, 2008). When bark beetle populations are low, dead, wounded, or otherwise compromised trees are targeted (Paine *et al.,* 1997). At outbreak levels, populations are capable of a more concentrated mass attack, utilizing larger, stronger hosts and devastating entire stands (Paine *et al.,* 1997). Populations may even spill over into less suitable hosts, such as mountain pine beetle (*Dendroctonus ponderosae* Hopkins) has done in British Columbia, utilizing interior hybrid spruce (*Picea engelmannii* (Parry) ex Engelm. x *glauca* (Moench) Voss) (Huber *et al.,* 2009).

Bark beetles, specifically mountain pine beetle, are major forest pests in British Columbia, which highlights the need to better understand the chemical ecology of this group of insects. Wood loss in British Columbia, due to the current mountain pine beetle outbreak, is expected to be in the range of 10.1 million hectares (Westfall & Ebata, 2008). Most bark beetles will take advantage of logging remnants and freshly wind-thrown trees (Byers, 1989a). The mountain pine beetle is a primary bark beetle, meaning that it attacks a living

host and kills it in order to successfully reproduce (Paine *et al.,* 1997). This may be due to an inability to effectively compete with other species aggregating in more compromised hosts, unless the opportunity presents itself. Aggressive bark beetle species, such as *Dendroctonus* spp., can cope with the higher resin flow of vigorous hosts, and some even rely on resin monoterpenes as precursors for pheromone synthesis (Borden, 1982). Less aggressive *Ips* species tend to locate plant tissue areas with lower resin volumes, like weaker trees, the crown, or cut boles (Borden, 1982). Secondary bark beetles such as the pine engraver (*Ips pini* Say) can both compete for and productively utilize recently killed pine trees and also use live standing trees as hosts (Paine *et al.,* 1997). The California fivespined ips (*Ips paraconfusus* Lanier) can inhabit recently killed or fallen pine trees, but will colonize pole-sized, young hosts (Wood & Storer, 2002). Considering the extensive stands of recently killed lodgepole pine trees in British Columbia, it is possible that secondary bark beetles may be able to take advantage of that resource while the dead trees are still nutritionally valuable as hosts. This might allow them to build up their populations and invade the smaller lodgepole pines that were unsuitable for mountain pine beetle and those younger trees that have recently grown to fill in the canopies of dead stands. So not only is the current outbreak a concern, but there may be repercussions involving other pest and host species, stressing the importance of better understanding how they interact. *Ips paraconfusus* is one of the most-studied bark beetles and has been established as a model organism in terms of molecular biology (Huber *et al.*, 2007; Seybold *et al.*, 1995; Tittiger *et al.*, 1999).

California Fivespined Ips

The California fivespined ips engraver (Coleoptera: Scolytidae) is found in California and Oregon (Byers & Birgersson, 1990) and is an extensively researched model species for

bark beetle ecology. It feeds on virtually all species of California pine, including ponderosa (*Pinus ponderosa* Dougl. ex Laws), sugar (*P. lambertiana* Dougl.), Jeffrey (*P. jeffreyi* Grev. & Balf.), grey (*P. sabiniana* Dougl.), and lodgepole pine (Byers & Birgersson, 1990). *Ips paraconfusus* has been intensively characterized in terms of chemical ecology (Byers *et al.*, 1979; Byers, 1981, 1983; Seybold *et al.*, 1995; Huber *et al.*, 2007). The male initially attacks and constructs a nuptial gallery, producing a pheromone blend of (S)-(-)-ipsenol, (S)-(+)-ipsdienol, and (4S)-*cis*-verbenol in order to attract additional conspecifics (Byers & Birgersson, 1990; Seybold *et al.*, 2006). Phenylethanol has also been reported as a possible pheromone component of this species (Borden, 1982). Up to three females join each male and construct their own galleries after mating, resulting in a star-like gallery pattern with eggs laid along the maternal galleries (Wood & Storer, 2002). Larvae feed on the phloem tissue, constructing their own pupal galleries, maturing under the bark, and emerging to colonize new hosts. Multi-voltinism (more than two broods per year) can occur with this bark beetle, but a univoltine (one year) lifecycle is representative of many species, and there are examples of semivontinism (more than one year to produce brood) in others (summarized by Stark, 1982).

Pine Defence Systems

Bark beetles utilize almost all trees and some woody plants; certain species may infest cones, small or large branches, the bole, or even roots (Wood & Storer, 2002). Trees have both mechanical and chemical defences to cope with bark beetles and their symbiotic fungi (Wood & Storer, 2002). Conifer bark is the first line of defence against mechanical wounding and pathogens. Additionally, thick bark protects the cambium from fires that remove other plant species (Agee, 1998). Lignification also plays an important role in the

physical defence against herbivory. It involves the thickening of cell walls, resulting in sclerenchyma whose form can vary from irregular sclereids to organized fibers (Franceschi *et al.,* 2005). Increased presence of lignin in Norway spruce (*Picea abies* (Linnaeus) Karsten) and Sitka spruce (*Picea sitchensis* (Bong.) Carr.) was found to negatively impact larval survivorship, growth, gallery construction, and oviposition of the spruce bark beetle *Dendroctonus micans* Kugelann (Wainhouse *et al.,* 1990).

Besides bark thickness, conifers utilize resin ducts to mechanically reduce damage. Resin is produced by resin cells, which depending upon the conifer species, may be dispersed within the stem, or organized into blisters, passages or duct systems as in *Pinus* species (Trapp & Croteau, 2001). Once a breach occurs, resin flows from ducts in an effort to displace or envelop invading pathogens. This also fills the wounded area with sticky, thick resin. Resin components seal the wound once the resin dries out, providing a barrier (Trapp & Croteau, 2001). However, the resin contained in such ducts is not simply for encasing pests; it consists of a range of toxic chemical components. If mechanical defences – such as bark and resin flow – fail, conifers have chemical-based defense system to cope with invaders.

Resin defence systems include specialized chemical properties. Resin contains a variety of hydrocarbons known as terpenes. Terpenes have many roles in plant physiology, including as metabolic precursors, and as vital components in growth and development (Bohlmann & Keeling, 2008). However, in general, they are thought to be important in defence against herbivores and pathogens. Monoterpenes and sesquiterpenes are well-studied terpenoid molecules that often act as toxins (Phillips & Croteau, 1999). Experiments on bark beetles with high or prolonged monoterpene exposure, including volatile phases, have

resulted in decreased feeding and boring activity as well as mortality (reviewed by Seybold *et al.*, 2006). Resin, containing terpenes, plugs the injury site and the more volatile monoterpenoid and sesquiterpenoid fractions are released into the air (Phillips & Croteau, 1999). These chemicals are present at varying constitutive levels in different plant species, and are produced in the absence of pests or pathogens despite being energetically expensive to maintain. The stationary nature of plants requires some sort of pre-emptive defensive strategy, and various chemical components serve this purpose. Attack triggers induced resistance, which in addition to responses like cellular necrosis, can result in a shift towards producing more toxic or repellent monoterpenes in the resin (Paine *et al.*, 1997).

Levels of particular terpenoids differ both between and within species. Although many species of conifer have the same general monoterpene profiles qualitatively, quantitatively the individual monoterpenes differ greatly in their principal components (Pureswaran *et al.*, 2004a). This makes sense as the range of parasites and pathogens will vary with host species, resulting in variant defensive strategies. A study of the oviposition behaviour of defoliating pine processionary caterpillars (*Thaumetopoea pityocampa* Schiffermüller) investigated the monoterpene profiles of preferred and avoided host species (Tiberi *et al.*, 1999). An avoided tree species had significantly more limonene in its needles than other species. When four other preferred species were sprayed with limonene, they were protected from oviposition. In terms of monoterpene composition, ponderosa pine has five observable chemotypes (reviewed by Latta *et al.*, 2003), which seem to survive herbivory differentially. A study of host selection in an area with several of the chemotypes found that trees killed by *Ips lecontei* Swaine had relative proportions of α-pinene, 3-carene, and limonene that were average for trees in the area (Thoss & Byers, 2006). Half of the living

pines were missing one of these compounds in their monoterpene complements, and were perhaps not selected because they were perceived as a non-host species by searching beetles. The authors also suggested that beetles may choose the trees to which their detoxification systems evolved and those trees with fewer types of monoterpenes could have amounts of those compounds too high for the detoxification systems to metabolize (Thoss & Byers, 2006). Such findings have suggested diversifying selection as a strategy to overcome insect pest pressure in particular. Although the species as a whole remains similar qualitatively, shifts towards different relative quantities of components allow some individuals to better adapt. The presence of generalist pests that are less phased by these shifts ensures that directional selection does not reduce the variability in the population (Latta *et al.*, 2003).

The terpenoids are a diverse class of molecules, with varying known effects on other organisms and I will focus on monoterpenes. Monoterpenes are perhaps the best-studied of the resin components; those commonly found in conifer resin include α-pinene, β-pinene, myrcene, limonene, camphene, β-phellandrene, terpinolene, and 3-carene (Phillips & Croteau, 1999). An experiment involving two symbiotic fungi of bark beetles, *Leptographium procerum* (Kendrick) Wingfield and *L. terebrantis* (Barras & Perry), showed that spore germination was negatively impacted by increasing doses of α-pinene and β-pinene (Klepzig *et al.*, 1996). Mycelial growth was inhibited not only by the pinenes, but also camphene, limonene, γ-terpinene and 3-carene. Doses of monoterpenes are known to be toxic to tunnelling bark beetles (reviewed by Seybold *et al.*, 2006).

Elevated monoterpene release rates can repel beetles searching for hosts, but they have also been implicated in other roles (reviewed by Seybold *et al.*, 2006). Host volatiles and bark beetle chemical signals augment bark beetle predator presence, indicating trees that

are currently infested with potential prey (Reddy & Guerrero, 2004; Seybold *et al.,* 2006). Some monoterpenes, such as myrcene and α-pinene, have been identified as attractant molecules and potential pheromone precursors (Byers & Birgersson, 1990; Seybold *et al.,* 2006). Monoterpenes of Norway spruce were found to increase the number of six-spined spruce bark beetles (*Pityogenes chalcographus* Linnaeus) entering artificial hosts, thus acting as kairomones, which synergize with pheromone components to increase attraction (Byers *et al.,* 1988). Douglas-fir beetle (*D. pseudotsugae* Hopkins), mountain pine beetle, spruce beetle (*D. rufipennis* Kirby), and western balsam bark beetle (*Dryocoetes confusus* Swaine) antennal electrophysical responses were measured with exposure to monoterpene volatiles (Pureswaran *et al.*, 2004b). All species detected 3-carene, myrcene, α-pinene, β-pinene, limonene, β-phellandrene and terpinolene. As pine trees in general contain similar monoterpene components, but at different relative levels, it was suggested that bark beetles must be equipped to detect particular monoterpene blends to find their ideal host. The relationship between bark beetles and the defensive response of their hosts is clearly more complex than mere toxin production and neutralization.

Host Colonization

In order for successful host colonization to occur, resin flow must first be countered and, in many cases, the host killed (Paine *et al.,* 1997). The pioneering sex, which varies by species, emits aggregation pheromones to attract not only mates, but also conspecifics. The very production of aggregation pheromones may indicate to other bark beetles that the host tree is not capable of defending itself against the invaders (Byers, 1989a). The resulting mass attack response to aggregation signals overwhelms the host's resin flow capabilities, preventing invaders from being flushed out. Symbiotic fungi carried by bark beetles also

infect the phloem tissue, effectively girdling the tree and contributing to the death of the host (Paine *et al.,* 1997). This is necessary for the larvae to survive to maturity, feeding on host tissue to grow and complete their life cycle.

To persist in a host long enough to kill it, bark beetles must also be equipped to detoxify host oleoresin monoterpenes, which although not typically lethal, do negatively impact bark beetles (Raffa & Berryman, 1987). Beetles and associated fungi are exposed to these toxins in their diet and in the volatiles within the galleries (Paine *et al.,* 1997). *Dendroctonus* and *Ips* species cope with such compounds via oxidative reactions (Pierce *et al.,* 1987). Oxidation renders the monoterpenes more soluble and increases the ability of the beetle to excrete them.

1.2 Bark Beetle Molecular Biology

The needs to produce pheromones and detoxify resin components represent important shifts in the physiology of these insects as they go from flying in search of a host to attracting or finding a mate and producing offspring. Feeding on the host elicits endocrine responses that alter the physiology of the beetle. Flight muscles are energetically expensive, and once they are no longer needed, that energy can be rerouted elsewhere. For example, within four days of landing, *I. paraconfusus* flight muscles are degraded (Borden & Slater, 1969). Females especially might require this energy for oogenesis. Related to this, vitellogenins from the fat body are recruited for incorporation into the eggs (reviewed by Bellés *et al.*, 2005). Vitellogenins bind lipids and are taken up by developing oocytes (Hartfelder, 2000). These processes are controlled by the endocrine system and all are related to the mevalonate pathway, which is responsible for the construction of isoprenoids from additions of acetyl-

coenzyme A (Bellés *et al.,* 2005). Normally, the mevalonate pathway leads to the synthesis of cholesterol, but this branch has been lost in insects. However, insects have a sesquiterpenoid branch that produces juvenile hormone, whose presence or absence dictates many physiological processes.

Pheromones as Signals

Bark beetle pheromones do not simply signal aggregation, although that function is integral to successful colonization. Several species, including *I. paraconfusus,* may utilize verbenone in order to signal that a host is fully occupied (reviewed by Byers, 1989a). This signals searching beetles to colonize elsewhere to reduce larval competition. Treatment with verbenone and nonhost-specific volatiles reduced attack of lodgepole pine by mountain pine beetle, and ponderosa pine by western pine beetle (*Dendroctonus brevicomis* LeConte) and red turpentine beetle (*D. valens* LeConte) (Lindgren *et al.*, 1989; Fettig *et al.,* 2008). Another strategy bark beetles employ to prevent over-aggregation may be in how the sexes respond differently to aggregation cues. *Ips paraconfusus* females were observed directing themselves right to the source of an aggregation cue, whereas pioneering males landed in an adjacent, presumably vacant area on the host to construct their gallery (Byers, 1983).

Pheromone components are shared between species, but in such a way that allows differentiation. Some *I. pini* populations predominantly use the opposite enantiomer of ipsdienol as *I. paraconfusus*, but are sensitive to the other version and are repelled by it (Light & Birch, 1979). Different ratios can even occur within a species in different populations; *I. pini* ipsdienol chiral ratios were found to vary regionally (Miller *et al.,* 1989). *Ips pini* females have even been found to prefer the same ratios of (-)-ipsdienol and (+)-

ipsdienol that were released by the males they were found paired with, indicating specific courtship properties that were not correlated to body size (Teale *et al.,* 1994).

Bark beetle-produced volatiles have also been found to reduce the occurrence of some types of insect predators that feed on bark beetles at certain life stages (Lindgren & Miller, 2002). It was suspected that since verbenone is present at later stages of attack, it repelled predators that did not consume late-stage bark beetles instars. Essentially, predators may have used volatiles to gauge when to forage certain trees. However, it seems bark beetles are capable of partial escape from predator detection through shifting pheromone component preference. Predators of *I. pini* were attracted to either (+) or (-) stereoisomers of ipsdienol, whereas the bark beetle preferred racemic blends of the stereoisomers in conjunction with a component not detectable by the predators (Raffa *et al.*, 2007). It was found that *I. pini* preferred different blends of the pheromone stereoisomers, perhaps enabling them to keep a step ahead of their predators.

Host Compounds as Pheromone Precursors

Bark beetles produce aggregation pheromones in the anterior midgut or hindgut, as indicated by the localized presence of such semiochemicals (Byers, 1989b; Nardi *et al.,* 2002; Eigenheer *et al.,* 2003). Pheromone components are thus deposited in the frass that the beetles deposit as they excavate galleries, thus exposing pheromones to the air (Nardi *et al.,* 2002). The pheromone components of *I. paraconfusus* were the first semiochemicals to be identified, and structural similarity to the host defence volatile myrcene was immediately observed, implicating the monoterpene as a possible precursor (Byers *et al.,* 1979). Byers *et al.* (1979, 1981) found an increasing presence of ipsenol and ipsdienol in the hindguts of

male *I. paraconfusus* with exposure to myrcene vapours and verbenol production in both sexes, albeit more in males, with (-)-α-pinene exposure. (+)-Ipsdienol was produced in adult male *D. brevicomis* exposed to myrcene vapors, but for a different purpose (Byers, 1982). Males are not the pioneering sex in this species, and (+)-ipsdienol may have been produced to repel co-occuring *I. pini,* which utilize (-)-ipsdienol for aggregation. Male *D. ponderosae* exposed to myrcene vapors also produced (+)-ipsdienol, perhaps to repel *I. pini* (Hunt *et al.*, 1986). These responses to monoterpene exposure indicated that bark beetles utilize host components to construct pheromone constituents. It has been proposed that oxidizing host resin compounds may have first occurred for detoxification purposes before evolving into a means for producing volatiles for communication (Franke & Vité, 1983).

Studies with varying levels of supplied myrcene have indicated that dietary supplementation alone is not enough to account for the amount of ipsdienol produced (Byers & Birgersson, 1990). The amount of phloem that would have to be ingested in 48 hours to supplement enough myrcene would have made up half the nuptial gallery for *I. paraconfusus*, although ingesting resin could decrease that amount substantially (Byers, 1981). Beetles devoid of any myrcene exposure still produced necessary pheromone components at normal levels. *De novo* biosynthesis has been investigated in *I. paraconfusus* and *I. pini* with radiolabelled substrates of the mevalonate pathway (Seybold *et al.,* 1995). Only males converted the labelled mevalonate pathway constituent, acetate, into the expected radiolabelled pheromone products ipsenol and ipsdienol. Keeling *et al.* (2004) investigated the gene expression levels related to components of the mevalonate pathway of *I. pini* before and after feeding. Although females exhibited up-regulation of early mevalonate pathway genes, the entire set was stimulated in males, linking increased production of mevalonate

components temporally to both the detoxification of host monoterpenes and pheromone producing processes. Biological symbionts have been suggested as a source of monoterpene conversion (Seybold *et al.,* 1995). However, such a symbiont, reacting to hormones, maturation, and the needs of only males, would be exceedingly unique. Female mountain pine beetle and male California fivespined ips (the pioneering sexes of their species), reared under anexic conditions, still produced pheromones when fed aseptic host phloem (Conn *et al.*, 1984). Most likely, the beetles are capable of producing monoterpenes *de novo* without such symbionts.

Juvenile Hormone and de novo Pheromone Biosynthesis

Juvenile hormones (JHs) are known to impact insect metamorphosis, female fertility, and polymorphism (in social insects); the variant JH0 has been detected in embryos of lepidopterans, JHI and JHII in lepidopterans, and JHIII in all orders of insects (reviewed by Hartfelder, 2000). When a juvenile hormone analog, fenoxycarb, was applied to male *I. paraconfusus,* their abdominal extracts attracted females (Chen *et al.*, 1988). These extracts were found to contain ipsenol and ipsdienol. Treatment with JHIII in the absence of dietary monoterpenes elicited a similar pheromonal response to myrcene feeding in both *I. paraconfusus* and *I. pini* (Seybold *et al.,* 1995; Byers & Birgersson, 1990). Certain mevalonate pathway genes, including 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-R), have exhibited up regulation in *I. pini* males with feeding (Keeling *et al.,* 2004). Tittiger and colleagues (1999) noted that in *I. paraconfusus,* JHIII treatment increased the presence of acyclic isoprenoids by increasing expression of HMG-R, supporting a *de novo* pheromone biosynthesis pathway. Hall *et al.* (2002) found that JHIII induced HMG-R expression in the anterior midguts of male Jeffrey pine beetles (*Dendroctonus jeffreyi* Hopkins). These males

also converted radiolabelled acetate into the pheromone component frontalin in the anterior midgut. This induction of HMG-R with JHIII application has also been found in *I. paraconfusus* (Tittiger *et al.*, 1999). A study of *I. pini*, found that *de novo* pheromone synthesis was both induced by feeding and JHIII, and that feeding increased JHIII production in the corpora allata (glands attached to the brain) (Tillman *et al.*, 1998). The *de novo* biosynthesis of pheromone had to do with stimulatory factors acting on the mevalonate pathway.

JHIII's influence on the mevalonate pathway implicates it with other processes occurring as bark beetles land on a tree. Treatment with a juvenile hormone analogue leads to flight muscle degradation in female *I. paraconfusus*, which is linked to energy use for oogenesis (Unnithan & Nair, 1977). The mevalonate pathway in insects has been linked to the release of vitellogenins during the reproductive cycle by expression of HMG-R and 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-S), which peak before oviposition in a cockroach, *Blattella germanica* Linnaeus (Casals *et al.*, 1997). By considering JHIII's involvement in several important processes, it may be possible to narrow the search for physiological requirements, enzymes or substrates, related to successful host colonization.

1.3 The Cytochromes P450 Superfamily

Cytochromes P450 (P450s) are related to, or directly involved with, the multiple physiological processes mentioned above (Feyereisen, 1999). The P450 superfamily is ancient and extensive; genes are found throughout the genomes of most organisms in clusters. Related genes in this family seem to have originated through duplication and subsequent mutation events (Feyereisen, 1999). There are 57 known P450s in humans, 102

in mice (*Mus musculus* Linnaeus), 84 in the fruit fly (*Drosophila melanogaster* Meigen), and 134 in the red flour beetle (*Tribolium castaneum* Herbst) (Scott, 2008). Protein products of these genes all have characteristic motifs and a haeme group, which aid in identification from sequence information and are thus integral to the discovery of new genes (Feyereisen, 1999). The soret peak (the point of most absorption) at 450 nm, for which the group is named, is observed when carbon monoxide binds to P450 haeme groups (Omura & Sato, 1964). Functions range from biosynthesis and degradation of hormones, fatty acids, and pheromones, to metabolism of xenobiotics and anthropogenic toxins (Feyereisen, 1999). These opposing functions, biosynthesis and degradation, are integral to maintaining the meaning of a signal, with turnover ensuring sensitivity to new signals and clearing out those that have already conveyed the message. Insects may have over a hundred different P450 genes in their genome, each equipped to oxidize one or several substrates to one or more products (Feyereisen, 1999). A "highly speculative estimation" of the percentages of insect P450s involved in certain functions, based on known human P450s, extrapolates that 65% could be involved with "biosynthetic processes" (35-50% hormone/pheromone biosynthesis, 5-15% reproduction, 5-10% pigmentation, 10% other), 30% with xenobiotic metabolism, and 5% with sensory physiology (Scott, 2008, p. 121).

P450s are prokaryotic in origin; mitochondrial P450s arose through erroneous targeting of microsomal ancestors, which localize to the endoplasmic reticulum (Werk-Reichhart & Feyereisen, 2000). Denoted first with the prefix 'CYP', they are categorized into families (over 40% amino acid sequence identity), and subfamilies (over 55% identity), denoted by a number and a letter (or combination of letters), respectively, before their specific gene number (Werk-Reichhart & Feyereisen, 2000). Owing to evolution through

unpredictable duplication events and mutations, and with nomenclature following overall amino acid identity, there exists no clear correlation between classification and function (Scott, 2008). Representative members exist within each major insect CYP clade that oxidize xenobiotics, insecticides and pheromones (Feyereisen, 2006). The insect CYP clade containing the CYP4 family is recognized as being particularly numerous and diverse in terms of functional range. It is also recognized as possibly being among the least studied of the insect P450 families to date (Feyereisen, 2006).

Although not solely responsible for pesticide resistance, P450s have been implicated as major factors in resistant insect strains (reviewed by Bergé *et al.,* 1998). The proteins derived from this gene group have been researched during pesticide development due to their detoxifying roles. One example would be when a P450 activates a compound that in turn inactivates the P450 so it cannot detoxify another compound (Feyereisen, 1999). Some plants employ this mechanism to render insects sensitive to their toxins. Several furanocoumarins, another group of plant defence compounds, synergize to impact the black swallowtail, *Papilio polyxenes* Fabricius, by interfering with P450s (Berenbaum & Zangerl, 1993). By interfering with P450-mediated detoxification, overall furanocoumarin detoxification was negatively impacted more than if only one compound was involved. Alternately, some P450s activate otherwise inert compounds into active toxins themselves. Besides inhibitor binding, P450 expression itself could be targeted, perhaps by using interfering RNAs to prevent detoxifying P450s from being transcribed. When fed host material that was expressing specific interfering dsRNAs designed for complementary binding to a particular P450 gene, the cotton bollworm (*Helicoverpa armigera* Hubner) exhibited repressed growth (Mao *et al.,* 2007). The P450 targeted for silencing was normally expressed in the midgut

during a growth phase, and seemed to confer tolerance to a specific host defensive compound. Most research of P450s found to be responsible for detoxification were characterized by studying microsomes, which are derived from the endoplasmic reticulum, where many P450s are targeted (Omura, 1999).

Microsomal P450s are integrated with the endoplasmic reticulum (ER), with a hydrophobic anchor residing at the N-terminus (Monier *et al.,* 1988). These P450s are processed similarly to those using the signal recognition particle (SRP) receptor. As the protein is translated at the ribosome, the N-terminal signal sequence is recognized and interacts with the SRP, halting translation until the complex binds to a SRP receptor on the targeted membrane (reviewed by Keenan *et al.*, 2001). After interaction with the receptor, the SRP is released and the peptide is translated into the selected membrane. The recognized signal peptide typically includes a cleavage site, which is the target of a signal peptidase (Paetzel *et al.*, 2002). Unlike other SRP-assisted proteins, P450 peptides generally lack cleavable signal sequences. Instead, the hydrophobic N-terminal signal is of sufficient length to act as an anchor and contains some sort of stop-transfer signal (Sato *et al.,* 1990). Several models have been proposed based on research on mammalian P450s, including a maintained transmembrane loop consisting of a secondary anchor (Black, 1992), or a mechanism by which a loop stretches back out into the cytosol, leaving only one N-terminal anchor to interact with the membrane and any signal peptide excluded from proteases in the interior of the ER (Monier *et al.*, 1988). It is unclear how some P450s lacking the N-terminal anchor still associate with the ER (Szczesna-Skorupa *et al.*, 1995). Some anchors can be converted to permit cleavage by the addition of positive charges at the N-terminus, which causes the formation of a loop structure that exposes a cleavage signal to proteases (Szczesna-Skorupa

et al., 1988). Alternately, there are some P450s that localize to the mitochondria. In this case, transport occurs after translation and involves multiple receptors that respond to either N-terminal or internal signal sequences (reviewed by Addya *et al.*, 1997). Addya *et al.* (1997) found that charged residues, residing C-terminal to the anchor, represented chimeric signals that permitted alternate mitochondrial targeting of otherwise microsomal targeted peptides. Generally, N-termini will encode a signal for one location or the other; microsomal P450s generally have hydrophobic N-terminal residues followed by a basic region and several proline residues, whereas mitochondrial P450s have N-termini with interspersed positive residues and a cleavable site (reviewed by Omura, 1999; Neve & Ingelman-Sundberg, 2008). The above two peptide sequence patterns can be sufficiently recognizable to predict localization of some P450s (Sandstrom *et al.*, 2006, 2008), but motif-recognizing software can also be employed (e.g. Huber *et al.*, 2007).

Cytochrome P450 Oxidation

A generalized P450 reaction mechanism involves binding of molecular oxygen to the haeme group, followed by the release of water. Depending on the P450's subcellular location, reduction occurs differently (Werk-Reichhart & Feyereisen, 2000). P450s are first reduced with electrons from NADPH or NADH by a cytochrome P450 reductase (CPR), in the case of microsomal P450s (reviewed by Omura, 1999). Mitochondrial P450s are reduced by a ferredoxin and a ferredoxin reductase, but reduction by CPR has also been observed, albeit inefficient (Guzov *et al.*, 1998). The first donated electron reduces the haeme, while the second allows for the splitting of molecular oxygen; one oxygen atom is released as water and the other is passed to the substrate (Feyereisen, 1999). Another coenzyme that is occasionally involved in P450-based reaction is cytochrome *b5*, which may improve the

transfer of the second electron (Guzov *et al.*, 1996). In addition to its own reductase (cytochrome *b5* reductase), it also accepts electrons from CPR (Enoch & Strittimatter, 1979).

The Diverse Roles of Insect Cytochromes P450

The vast number of enzymes and substrate possibilities makes direct characterization difficult. The following are examples of functionally characterized P450 enzymes, but there are several cases where bark beetle representatives have yet to be elucidated beyond conspicuous expression levels with varying treatments. However, a few bark beetle P450s have been either functionally characterized or otherwise implicated in pathways related to colonization success (Sandstrom *et al.*, 2006, 2008; Huber *et al.*, 2007).

Effective signalling not only requires the synthesis of signalling molecules, but also their catabolism to refresh the system. To ensure that the presence of a given signal retains meaning, it must be removed from a system to convey the opposing message, the absence of the signal. JHIII is synthesized by the epoxidation of methyl farnesoate by a P450, and is catabolized by a mechanism also involving a P450 (Feyereisen, 1999). The female cockroach (*Diploptera punctata* Eschscholtz) exhibits a decline in juvenile hormone levels leading to oviposition; it was found that CYP4C7 expression increased during this decline (Sutherland *et al.*, 1998). When the cloned gene was expressed in bacteria, its enzyme product hydroxylated the hormone, which implicated it in the suppression required to finish the gonotrophic cycle (oocyte production and oviposition). Similarly, in a converse manner to the way that P450s have been implicated in pheromone production, antennal P450s break down pheromones. In the pale-brown chafer (*Phyllopertha diversa* Waterhouse), CYP4AW1 was proposed as a pheromone-degrading enzyme because expression was limited to the

antennae of males and pheromone degradation activity was compromised with the application of a P450 inhibitor (Maïbèche-Coisne *et al.*, 2004). This would allow for rapid signal turnover by a male following a female-emitted pheromone plume, which would be important for direction adjustments in flight. As mentioned, insects lack the ability to synthesize cholesterol. Dietary cholesterol is a precursor of ecdysone, which is an insect moulting hormone. In this case, cholesterol is the precursor, the last step is the oxidation of ecdysone by a P450 to 20-hydroecdysone (Petryk *et al.*, 2003). Two other hydroxylations occur prior to that step, also facilitated by P450s (Gilbert, 2004). P450s have also been implicated in the breakdown of ecdysteroids (reviewed by Feyereisen, 1999). Insect P450s have been shown to ω-hydroxylate fatty acids and have thus been suggested as possible eicosanoid metabolizing candidates (Feyereisen, 1999).

A P450 that has been functionally characterized, *I. pini* CYP9T2, was selected for investigation based on mRNA transcript levels in the midguts of males and the aforementioned increase with both feeding on host phloem and JHIII application (Sandstrom *et al.*, 2006). In order to elucidate function, the gene and the necessary reductase were expressed in a baculovirus-mediated *Spodoptera frugiperda* Smith (SF9) insect cell system. Endoplasmic reticulum-derived microsomes were isolated and exposed to myrcene, and the expressed protein catalyzed the production of ipsdienol. This is the first functional study positively linking a bark beetle P450 enzyme to the conversion of a host monoterpene into a pheromone component. More recently, an analogue of this gene in *Ips confusus* LeConte, IcCYP9T1, was expressed in SF9 cells. The resulting enzyme performed the same function (Sandstrom *et al.*, 2008). Considering how many bark beetles not only utilize identical or similar aggregation pheromones, but also colonize hosts with similar monoterpene

complements, it stands to reason that similar genes will be found in other species. Combined, this evidence suggests P450s that are up-regulated with feeding in the anterior midguts of the pioneering sex are primary candidates for research into pheromone production. Since pheromone production possibly acts in concert with detoxification pathways, P450s are prime targets for studies of the relationship between bark beetles and their hosts.

Several *I. paraconfusus* cytochrome P450s were isolated for qPCR expression analysis in both sexes, with feeding as the factor eliciting response (Huber *et al.,* 2007). Several genes showed increased expression with feeding on *P. ponderosa* phloem; one (CYP9T1) was expressed 85,000x above baseline in only males. This response occurred 8 to 24 hours from initial attack, which would be indicative of genes that the male expresses to deal with the associated monoterpenes or to attract more individuals. Those expressed in both sexes could be general detoxifiers or related to endocrine regulation, and those specific to females may be involved with oogenesis.

1.4 Heterologous Expression of Cytochromes P450 and P450 Reductase

In order to study the function of P450s, heterologous expression systems have been employed (Sandstrom *et al.*, 2006, 2009; Duan *et al.*, 2004; Wen *et al.*, 2003). By producing a protein with a cell line, enough enzyme can be synthesized to study it in isolation. Mammalian (Chinese hamster ovary - CHO), bacterial (*Escherichia coli* (Migula) Castellani & Chalmers), yeast (*Saccharomyces cerevisiae* Meyen), and insect cell (SF9) cultures can be used. Eukaryotic cells have an obvious advantage in regards to the functionality of some enzymes, as they are capable of post-translational processing – such as glycosylation and phosphorylation (Geisse *et al.*, 1996). Although yeast cells are capable of producing P450s,

there could be targeting issues, and insect cells can effectively cleave signal peptides during transcription (Geisse *et al.*, 1996). Uptake of recombinant virus in insect cells is also more effective than plasmid vectors of yeast or *E. coli.*

Conditions must be optimized to synthesize active P450s in an insect cell system. The multiplicity of infection (MOI) must be optimized to get the most functional protein without overwhelming the cells with virus. Generally, P450 protein levels increase with MOI until a threshold is reached, after which that amount will level off and increasing amounts of nonfunctional protein will be present (Duan *et al.,* 2004; Wen *et al.,* 2003). A virus encoding the necessary CPR must also be optimized in the system by testing a range of ratios in a cotransfection with the P450-containing virus. An *Arabidopsis* P450 was optimized with reductase in insect cells; P450 content clearly decreased with increased levels of reductase-containing virus (Duan *et al.,* 2004). Enzyme activity increases with reductase levels, so a balance must be struck between expression of target P450s and expression of required CPR (Wen *et al.,* 2003).

Optimizations are complicated by issues of quantification. Virus concentration itself is titered based on its ability to infect a uniform plaque of cells, and cell preparations can exhibit batch-to-batch variance in terms of infectivity and production. Also, the concentration of infective virus decays over time, which can increase margins of error. P450s themselves are measured by an assay involving monitoring spectra when the proteins' haeme groups are bound to carbon monoxide. Common issues with this assay include insufficient protein levels and turbidity that scatters the light (Humphreys & Chapple, 2004). This can be a significant problem, but even undetectable levels of P450s can still produce positive substrate oxidation under the right conditions (D.-K. Ro, pers. comm.). CPR protein expression is also quantified

via an assay, where the ability to reduce cytochrome *c* with NADPH equivalents is measured over time. Alternately, recombinant virus that includes a polyhistidine tag or epitope can be constructed and microsomal preparations can be subjected to immunoblot analysis to monitor expression levels (Humphreys & Chapple, 2004). However, unlike the spectral assay, antibody analysis does not evaluate functionality. Immunoblotting requires gel electrophoresis to separate proteins, which can take over an hour to run before blotting can begin (which itself can take a working day). Capillary electrophoresis using a microfluidic gel can speed this process up substantially, taking a half an hour for ten samples and requiring a tenth of the sample volume (Dolnník *et al.*, 2000; Brobacher *et al.*, 2008). Although electrophoresis-based techniques (e.g. coomassie blue staining) detect proteins besides the recombinant products, the automated Experion capillary electrophoresis setup provides concentration values for each protein size (Bio-Rad Laboratories, Hercules, CA). Comparison to a control sample from cells not exposed to virus allows for identification of recombinant proteins. The drawback to this method is that active and inactive P450s are not distinguishable, but this is compensated for by low sample volumes, faster analysis, and the more specific size and concentration estimates.

1.5 Investigation of Recombinant *I. paraconfusus* CYP4 Enzymes

Several *I. paraconfusus* P450 genes share some expression characteristics with CYP9T2 (Huber *et al.,* 2007). Based on their over-expression following feeding in one or both sexes, these genes are prime targets for functional investigation regarding host xenobiotic detoxification and other processes. By using an SF9 insect cell expression system and isolating the subsequent microsomes, genes involved in the detoxification and pheromone biosynthesis pathways may be positively identified. In my research, I measured

enzymes by their spectral properties, but also employed capillary electrophoresis with a digital gel to gain both size and accurate concentration data. This refined the results by providing protein expression data that was more reliable than the difficult spectral analyses. This allowed for effective optimization for functional recombinant P450 co-expression with a cytochrome P450 reductase. Microsomes were exposed to an array of potential host monoterpene substrates and analyzed for products. Signal peptide-detecting software programs were employed to investigate alternate organelle targeting of the *I. paraconfusus* CYP4 P450s. The automated gel electrophoresis system was also used to investigate possible peptide processing events that manifested themselves as detectable changes in the apparent molecular mass. Histidine-tagged CYP4 baculoviral constructs were similarly employed to detect possible isozyme variants with alternate cellular localization, but by the specific antibody blotting of tagged recombinant proteins. Taken together, these analyses will aid in the characterization of these particular *I. paraconfusus* CYP4 subfamily P450s.

OBJECTIVES:

(I) Optimize expression of recombinant *Ips paraconfusus* P450s with an SF9/baculovirus system, employing both spectral and capillary electrophoresis methods

(II) Test recombinant enzymes with potential substrates based on a prior DNA expression study and available bioinformatics data

(III) Investigate alternate subcellular targeting events by using peptide-analyzing software and observing samples for differently-sized recombinant protein variants

2.0 Expression Optimization of Recombinant CYP4s with MPB-CPR

2.1 Introduction

An expression system must possess the ability to accurately produce and modify a given recombinant protein and it must also produce adequate amounts of protein for use in assays. As noted by several reviews, baculovirus systems employing insect cells have been developed to fulfill both requirements (Possee, 1997; Giesse *et al.*, 1996; Kost *et al.*, 2005). As such systems involve eukaryotic cell lines, modifications such as glycosylation, phosphorylation, and the addition of fatty acid chains can take place, although insect and mammalian glycosylation processes vary (Possee, 1997; Giesse *et al.*, 1996). Baculovirus constructs can also be used to transfect insect larvae and mammalian cells but insect cell lines are easier to use in many applications because they can be quickly scaled-up (Kost *et al.*, 2005). A comparative study of heterologous systems for the expression of a bovine P450 using SF9 insect, yeast, and mammalian cell lines, found that the insect cell line produced the most molecules of P450 per cell (Barnes *et al.*, 1994).

Optimization of a baculovirus expression system relies upon the accurate measurements of various components of the system. Virus measurement by plaque assay is based on infectivity, where an affixed layer of cells exhibit infective loci as cell lysis spreads to adjacent cells. The epicenters of infection are then counted and divided by the volume of inoculum used. Immunoassays also detect signs of infection on a well of adherent cells, but an antibody is used to tag a baculoviral protein that is expressed prior to lysis (Clontech, Mountain View, CA). These measurements are critical for optimizing transfection conditions and obtaining protein. The multiplicity of infection (MOI), which is essentially the number of viral particles per cell in a transfection, resulting in the most recombinant protein, must be

determined. Traditionally, P450 and reductase measurements are based on catalytic activity. Cytochrome P450 reductase (CPR) activity is measured by tracking its transfer of electrons from NADPH to cytochrome *c*. The rate of reduction is measured, associated with an absorbance change at 550 nm (described by Guengerich *et al.*, 2009). P450s are measured and named for their propensity to bind carbon monoxide (CO), which occurs when the haeme prosthetic group is coordinated correctly (Omura & Sato, 1964; Guengerich *et al.*, 2009). Properly formed enzyme, artificially reduced with sodium dithionite, will bind CO and exhibit a characteristic peak at 450 nm (or a 420 nm peak if malformed).

There are several problems with spectral measurement of P450 content, which together can create a "significant obstacle" to recombinant protein production (Humphreys & Chapple, 2004). It can be difficult to obtain enough P450 to be detected by this method, and samples themselves may be too turbid to allow for a good baseline estimate. General UV/VIS microbiology spectrometers may not be accurate or precise enough to resolve turbid samples that scatter light; detergents can solubilize samples and counter turbidity, but this manipulation may compromise the applicability of the measurement to the original sample (Humphreys & Chapple, 2004). Instruments with photomultiplier tubes closer to the cuvettes may accommodate turbidity better, but other haemoproteins, or the interference of a large P420 peak trough, can still obscure the P450 peak (Guengrich *et al.*, 2009).

Immunoblotting techniques have been employed to circumvent some of the problems with spectral quantification. Although it does not distinguish between nonfunctional and functional forms of P450s, the system is designed to specifically measure only recombinant protein. N-terminal epitope tagging was used to monitor expression of a P450 and it was found to have a minimal detection limit 400x less than CO-difference analysis (Humphreys

& Chapple, 2004). This method also has the benefit of conserving sample. One CO-difference spectrum typically requires 200 μL of sample, whereas immunoblotting assays generally require less than 10 μL. The Experion automated electrophoresis system (Bio-Rad Laboratories, Hercules, CA), which does not require tagged proteins, uses only 4 μL. It can take a week to amplify enough SF9 cells for transfection, and three days of incubation to express protein. Valuable viral stock is consumed during optimization, as are microsomal isolates that could have been put to use in downstream substrate assays. Additional virus production also requires more cells and an incubation time of up to 5 days. Viral titering procedures are either time-consuming (plaque assay, up to 10 days), or expensive (antibody kits). Virus titer also decays over time, thus it is very important, given the time it takes to amplify cells and viral stocks, and express protein, to conduct optimizations in a timely and efficient manner. Should an optimization method using a fast, volume-saving electrophoresis system be successful, it could save time and money during expression optimization. Traditional methods that test catalytic function would still be necessary, but perhaps they would be used as check-points. Larger volume transfections could be produced to verify functional enzyme in a smaller range of conditions, already suspected to be near-optimal.

In this study, I aimed to use spectral measurements, but supplemented that information with concentration data derived from an automated electrophoresis system capable of distinguishing individual size classes of protein in a given sample. Recombinant baculoviruses encoding *Ips paraconfusus* P450s (CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and CYP4G27) were constructed for expression in SF9 cells. The amounts of the viruses used for the CYP4AYs, CYP4BD1 and CYP4BG1 were first optimized individually. Co-optimizations of CYP4AY1 and CYP4AY2 with an insect CPR were then conducted.

2.2 Materials and Methods

2.2.1 Materials

Oligonucleotides for sequencing and cloning (Table 2.1) were from Integrated DNA Technologies (San Diego, CA). Plasmids, CONCERT Rapid Gel purification columns and PCR reagents were from Invitrogen (Carlsbad, CA), and restriction enzymes were from New England BioLabs (Ipswich, MA). Plasmid DNA and PCR products were isolated and purified with QIAprep Spin Miniprep, and QiaQuick and MinElute PCR Purification kits (Qiagen, Germantown, MD). Viral DNA was isolated with the Easy DNA kit (Invitrogen). The shuttle vector, C-term Baculodirect Expression Kit, SF9 cells, Grace's Insect Media, antibiotics and fetal bovine serum were from Invitrogen. Bovine cytochrome *c*, and NADPH were from Sigma-Aldrich (St. Louis, MO). Capillary electrophoresis was conducted on Pro260 chips with the Experion automated electrophoresis station and accompanying software, version 3.0 (Bio-Rad Laboratories).

Table 2.1 - Restriction end cloning primers. Primers were designed with restriction enzyme sites adjacent to the start or stop codon. Tagged versions had either one base deleted to nullify the stop codon, or the whole stop codon plus one more base to ensure the downstream viral histidine tag would be in-frame.

Primer	Sequence (5'-3')
AY1F	CTGGGATCCATGATTCTGTTTGAAGTAGTCGTGG
AY2F	CTGGGATCCATGTTGTTTTGGTTGTTTCAAGTAGT
BD1F	CGTGGATCCATGTTATTCGTGTTGTTAGCGTTG
BG1F	CTGGGATCCATGAAAATGTTCATTGGTGTTTTG
G27F	ATAGGTACCAAACATGTCGACTGCCA
REDF	TACTAGGTACCATGGCTTCGAGTGTTGTGTCT
AY1R	CTGGCGGCCGCCTAGTTTGAAGAACCCAACTTGTTC
AY2R	CTGGCGGCCGCTTAGGACCTTTTCTCCAAACGA
BD1R	CTGCTCGAGTTAACGGGCTTGTAAGGAAATTT
BG1R	CTGCTCGAGTTACTCTAGGGGTTGAAAATTAATTTT
G27R	ATTCTCGAGTTAGGCTTTCACTGAAACGATT
REDR	GCATACAGCGCGGCCGCTTAACTCCAGACATCCATCGAGA
TAY1R	CGTGCGGCCGCTTTGAAGAACCCAACTTGTTCAA
TAY2R	CGTGCGGCCGCGACCTTTTCTCCAAACGAACTTT
TBD1R	CGTCTCGAGTAACGGGCTTGTAAGGAAATTTTA
TBG1R	CGTCTCGAGTACTCTAGGGGTTGAAAATTAATTTTAA
TG27R	ATACTCGAGACGATTGGTTTCCTTTTCTC
TREDR	GCATACGTAGCGGCCGCACTCCAGACATCCATCGAGAT

2.2.2 Entry Clone Construction

Plasmids containing *Ips paraconfusus* P450s (Huber *et al.*, 2007) were fully sequenced, using gene-specific internal primers, to confirm identity and quality of the full-length P450 inserts. *Dendroctonus ponderosae* cytochrome P450 reductase-containing vector (MPB-CPR) was generously supplied by J. Bohlmann, C. Keeling and colleagues (Michael Smith Laboratories, University of British Columbia). An internal shine-dalgarno sequence previously present in the MPB-CPR had been removed to allow for expression in *E. coli* (Keeling *et al.*, in prep.). Open reading frames (ORFs) of P450s and MPB-CPR were amplified using complementary primers with restriction enzyme cutting sites designed adjacent to start and stop codons (Table 2.1). For the CYP4AYs, CYP4BD1 and CYP4BG1, a BamHI site was included at the 5' end; CYP4G27 and MPB-CPR were amplified with a KpnI site at the 5' end. All constructs had either NotI, or XhoI sites incorporated at the 3' end, adjacent to the stop codon. Histidine tagged versions of the constructs were also created by deleting the stop codon and one base so that a downstream virally encoded polyhistidine tag would be in frame. As per Sandstrom *et al.* (2006), PCR was used to amplify 5-10ng of ORF-containing plasmid with high fidelity Taq polymerase (Invitrogen). Reactions (100 µl) were held at 95°C for 1 minute; followed by four cycles of 94°C for 4 minutes, 41°C for 1 minute, and 72°C for 4 minutes; 36 cycles of 94°C for 40 seconds, 57°C for 1 minute, and 72°C for 4 minutes; and concluded with 10 minutes at 72°C. PCR products were purified and then digested for 4-8 hours at 37°C with the appropriate restriction enzymes, as was the shuttle vector, pENTr1A. After digestion, samples were gel purified in 1% agarose (1hr, 100V) and resuspended in water. Inserts were ligated at a 3:1 ratio into pENTr1A for 14 hours at 16°C and transformed into competent DH5α *E. coli* that were produced in-lab

(protocol available at www.neb.com, adapted by J. Utermohlen from John Innes Institute, Norwich, England). The plasmids that were produced were isolated and sequenced with gene and vector-specific primers to ensure proper orientation of the gene and PCR fidelity. Aliquots were concentrated via ethanol precipitation and resuspended to 50-150 ng/μl in TE (pH 8.0).

2.2.3 SF9 Cell Maintenance

SF9 insect cells were grown at 27°C in adherent culture in Supplemented Grace's Insect Media with 10% Certified Fetal Bovine Serum (Gibco/Invitrogen). Media was sterilized through a 0.2 μm filter and gentamycin (10 μg/mL) and amphoceterin B (0.25 μg/mL) were added to prevent bacterial and fungal contamination, respectively. Adherent culture was initiated and maintained in standard 25 cm^2 and 75 cm^2, tissue culture treated, cell culture flasks with vented caps (Corning, Lowell, MA). The tissue culture treatment oxidizes the polystyrene surface, rendering the surface hydrophobic and negatively charged, to permit cell attachment and spread. Cell counts were conducted on a haemocytometer with trypan blue stain (0.036% final concentration). Once cells could be passaged regularly at 72 hours (i.e. harvested once cells completely covered a flask's surface to initiate a new flask) with at least 90% viability, they were maintained in adherent culture or used to initiate suspension culture in Erlenmeyer flasks topped with aluminum foil (to permit air exchange). Each suspension line was maintained for up to a month in complete media with 0.01% pluronic F68 polyol (VWR, West Chester, PA) to prevent shearing (27°C, 105 rpm). Maintained suspension cultures were grown to a maximum of 2×10^6 cells/mL and diluted to 1×10^6 cells/mL when passaged. When glassware was reused, it was treated with 10% acetic

acid and autoclaved with water (121°C), followed by a dry cycle (121°C), and then dried for 1 hour in a 120°C oven. All other materials were certified sterile and not reused.

2.2.4 Recombinant Baculovirus Construction

For each construct, the ORF/pENTr1A vectors were recombined with linear viral DNA as per the manufacturer's directions (Invitrogen). Cells were transfected with the recombination product and media, containing the virus, was harvested after 96 hours (centrifuged at 3,000xg for 5 minutes to remove cells and debris). Second generation (P2) viral stock was attained through a subsequent transfection, at which point cells were also analyzed for protein expression via capillary electrophoresis and western blot of tagged constructs. P2 virus was subjected to a plaque assay to screen for non-recombinant virus presence and to obtain a titer. For each plaque assay of a construct, 0.8 x 10^6 cells were seeded in each of 12 wells in 2 mL media (enough to test six viral dilutions in duplicate). After attachment, they were exposed to dilutions of viral stock for one hour and overlaid with 2X Grace's Insect Media/20%FBS in low melting point agarose (1% final concentration) with Bluo-gal (150 μg/mL final concentration). Plates were kept in sealed plastic bags with moist paper towel, incubated at 27°C and protected from light. Once a complete layer of cells formed under the solid agarose mixture and plaques were observable (~96 hours), a neutral red overlay was added to the wells (as per manufacturer's directions). If non-recombinant (blue) plaques were present, the next amplification included ganciclovir to select against non-recombinant virus. After a maximum of 10 days, titer could be calculated by counting the infective foci (plaques) and using the formula below:

Titer (pfu/mL) = plaque forming units (pfu)/ volume of inoculum (mL) x dilution factor

P2 or P3 viral DNA was isolated and amplified with viral-specific primers flanking the inserted genes using high fidelity Taq polymerase. Reactions (100 μL) were held at 95°C for 5 minutes; followed by 30 cycles of 94°C for 45 seconds, 52°C for 1 minute, and 68°C for 3 minutes and 45 seconds; and concluded with 10 minutes at 68°C. Products were fully sequenced with viral and gene-specific primers to confirm orientation and fidelity. High-titer stocks were created by infecting at a low (0.1) multiplicity of infection (MOI) and incubating for 96 hours, or just until cell lysis was confirmed. Cells and debris were spun out at 3,000xg for five minutes. Viral stocks (supernatant) were stored at 4°C and protected from light, while a small aliquot was stored at -80°C. Titers of P3 or later generation stock were estimated with the BakPak Baculovirus Rapid Titer Kit (Clontech) as per the manufacturer's directions. The calculation to determine volume of virus to add to obtain a selected MOI is as follows (Invitrogen):

Volume of viral stock (mL) = (number of cells x MOI)/ viral titer (pfu/mL)

2.2.5 Optimization of P450 Transfections

To determine the best MOI to express each protein, CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 were individually optimized via single transfections. Typically, 3 or 7.5 x 10^6 cells were exposed to a range MOIs (e.g. 0.5, 1, 5, 7.5, 10, and 20). Hemin (MP Biomedicals, Solon, OH) was added to a final concentration of 5 μg/mL at 24 hours post-transfection and cells were harvested for microsomes at 72 hours as follows: cells were centrifuged at 3,000xg for 15 minutes at 4°C to remove media. The pellet was rinsed three times in phosphate buffered saline (1/3 cell volume) and resuspended in the same volume of lysis buffer (100 mM Na_2HPO_4/NaH_2PO (pH 7.8), 0.1 mM DTT, 0.5 mM PMSF, 1.1 mM EDTA, 5 μg/mL leupeptin, 20% glycerol). Samples were sonicated on ice with an ultrasonic

dismembrator (Fisher Scientific, Model 150T) for 30 seconds, 3 times, output 3. Cell debris and mitochondria were spun out at 800xg and 10,000xg, respectively, for 10 minutes at 4°C, if needed for analysis. Otherwise, secondary pellets were resuspended in 1/10 volume cell lysis buffer after supernatant isolation for an additional 30 seconds of sonication to break up any remaining cells, before re-pelleting debris at 3,000xg. Supernatant was ultracentrifuged at 100,000xg for 1 hour at 4°C to isolate and concentrate microsomes (this step was necessarily circumvented for a period of time when the ultracentrifuge was out of order - resulting samples were called 'cell lysate'). These microsomal pellets were mechanically broken up and suspended in 1/10 to 1/20 volume of cell lysis buffer (sonicated for 10 seconds if clumping deemed it necessary), depending on the observed density of the pellet. Samples were either analyzed or flash-frozen in liquid nitrogen and stored at -80°C, for a maximum of three weeks, prior to use.

Optimizations were analyzed by capillary gel electrophoresis and CO-difference spectra. If samples were suspected to be too turbid for spectral analysis, they were first treated with 1% CHAPS detergent (Fisher Scientific, Waltham, MA) at 4°C for 1 hour with light shaking, and debris was removed by centrifugation at 15,000xg (4°C). Sodium dithionite crystals were added and cuvettes were scanned to develop a baseline before bubbling carbon monoxide into a cuvette for 1 minute (2 bubbles/second). Absorbance difference was measured between 400 and 500 nm with a Perkin-Elmer, UV-Vis spectrometer (Lambda 2S model), with PECSS software (version 4.2). The concentrations of active enzyme (nmol P450/mL) in microsomal and cell lysate preparations were estimated using an extinction coefficient of 91mM cm^{-1} as follows (Omura & Sato, 1964):

P450 (nmol/mL) = Abs at 450nm - Abs at 490nm / 0.091 x dilution factor

If the representative baseline scan was compromised such that a decline is absorbance towards 500 nm could not be obtained, a conservative estimate to replace the absorbance at 490 nm was taken at 480 nm. If even that was not amenable, which was only the case twice, an estimate was obtained by averaging the absorbance values of the P450 peak's ascent (~430 nm) and the value at 460 nm (similar to Johnsten *et al.*, 2008).

2.2.6 Optimizations of Co-transfections including MPB-CPR

To test whether SF9 cells produced active MPB-CPR, 3 x 10^6 cells were infected with that viral construct at a MOI of 0.25 and microsomes were tested by cytochrome *c* assay (as per Guengerich *et al.*, 2009). Briefly, 25 µl of MPB-CPR expressing sample and 80 µl of 0.5 mM bovine cytochrome *c* (in 10mM Na_2HPO_4/NaH_2PO, pH 7.7) was diluted in 0.3 M Na_2HPO_4/NaH_2PO, pH 7.7) in a cuvette to a final volume of 990 µl. A baseline scan was conducted for 2 minutes. Next 10 ul of 10 mM NADPH was added, the cuvette was covered with parafilm and inverted to mix, and the change in absorbance was immediately measured at 550 nm over 3 minutes. Activity was calculated by dividing rate (absorbance change/minute) during intervals of linear increase by the constant 0.021 nmol cytochrome *c* reduced/λ. This above protocol was repeated during co-optimization of P450 with CPR. When the optimal P450 MOI was selected, transfections containing MPB-CPR were tested. P450 MOI was held constant and MPB-CPR:P450 ratios of 1:1-1:40 were tested in 5 mL adherent culture as described previously. Cells were treated and harvested as before; P450 and CPR contents were also measured via Experion Pro 260 analysis, as per the manufacturer's directions (Bio-Rad Laboratories).

2.2.7 Confirmation of Optimization Applicability and Harvest Time

To verify the applicability of optimizations done in adherent culture to expression in suspension culture (used for scale-up), both culture techniques were compared directly with transfections of 15 x 10^6 cells (one suspension shaker flask and two 75 cm^2 adherent flasks). CYP4AY1 virus was used at an MOI of 10. Microsomes were harvested as before and were resuspended in the same final volume of lysis buffer (1 mL). As the two culture types used different cell densities, which might have influenced the overall productivity of the cells, the concentration of recombinant enzyme was calculated as a percentage of the total protein isolated.

The harvest time of 3 days recommended by the manufacturer (and the convention in the literature: Wen *et al.*, 2003; Duan *et al.*, 2004; Mao *et al.*, 2008), was analyzed by testing harvest at days 2, 3, 4, and 5 post-transfection in 5 mL adherent culture. In this case, the co-optimization conditions selected for CYP4AY1/CPR co-expression were utilized (CYP4AY1 MOI:10, MPB-CPR MOI: 0.25). Microsomal samples were collected and evaluated by Experion Pro260 chip analysis, as per the manufacturer's directions (Bio-Rad Laboratories).

2.3 Results

2.3.1 Single Transfection Optimization of CYP4AY1

Experion analysis indicated that the optimal MOI for CYP4AY1 was 10, as this was before recombinant protein expression levels reached a plateau or maximum. Figure 2.1a shows an increase in band intensity (~48 kDa) with increasing MOI. This is represented graphically in Figure 2.1b, where a maximum CYP4AY1 protein concentration was observed at an MOI of 15, followed by a decline. Corresponding carbon monoxide spectra are illustrated in Figure 2.1c,d. These were the only two MOI tested for CYP4AY1 expression that had detectable P450 peaks. An MOI of 7.5 was the lowest amount of virus resolving a shallow peak corresponding to 0.0824 nmol/mL of microsome (Figure 2.1c); the baseline measurement had to be estimated due to high background noise after ~460 nm. For an MOI of 10, there was a clear peak at 450 nm representing 0.165 nmol/mL of active recombinant P450 (Figure 2.1d). The accompanying P420 peak represents recombinant enzyme in which the haeme was not properly bound. This result was in agreement with the prediction that the MOI resulting in the most functional P450 would be the next-highest to the MOI associated with the maximum concentration by Experion analysis. Higher MOIs were thus more likely to have non-functional CYP4AY1 present due to the negative impacts of higher viral concentration on cell health.

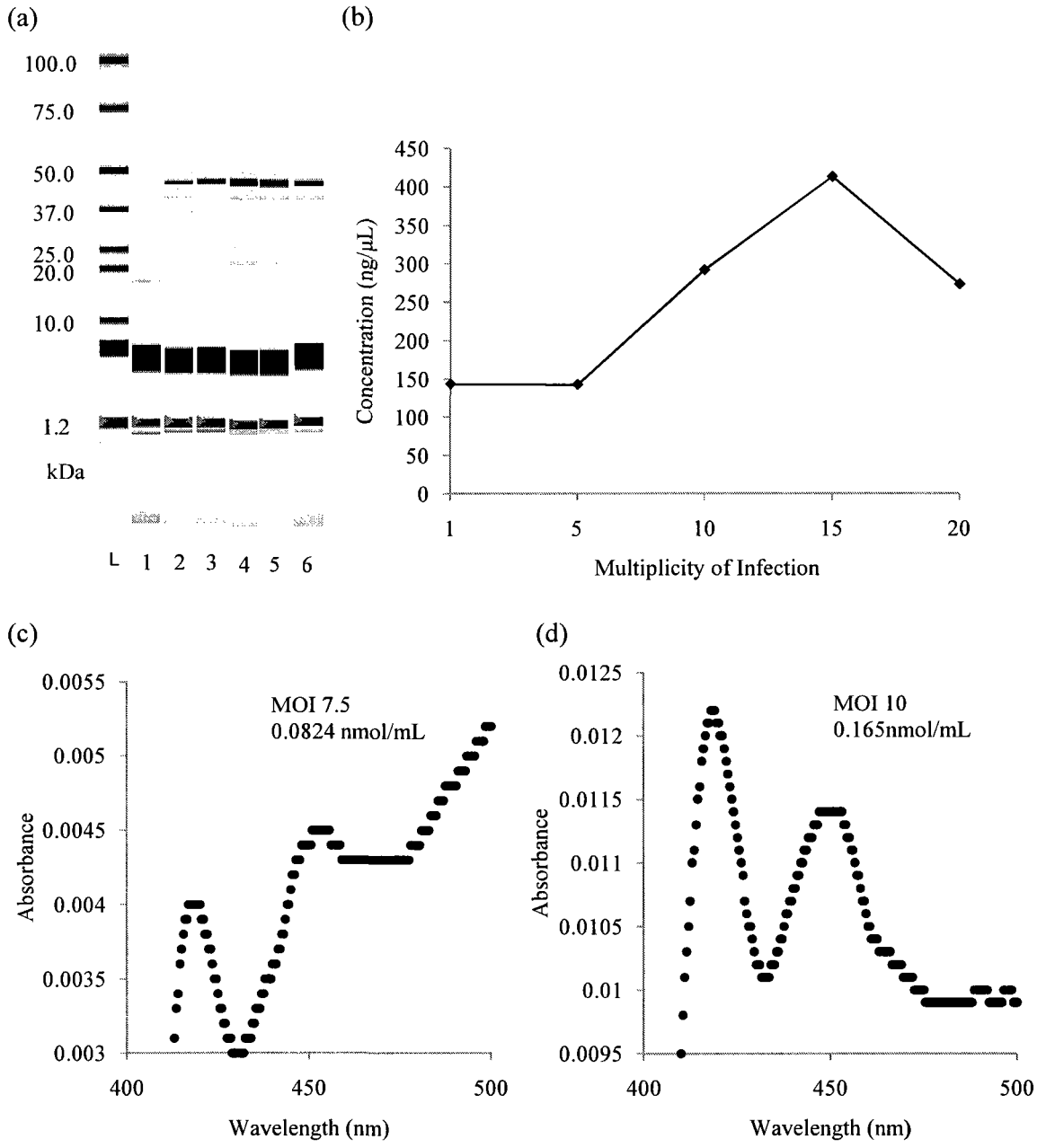


Figure 2.1 – Optimization of CYP4AY1 transfections. (a) The digital gel output ladder (L) is shown in kDa. Lane 1 is cell control microsome followed by MOIs of 1, 5, 10, 15, and 20 with CYP4AY1 resolving at ~48 kDa. (b) Concentrations of CYP4AY1 are maximized at a MOI of 15. (c) The CO-difference spectrum of CYP4AY1 at an MOI of 7.5, corresponding to 0.0824 nmol/mL, had a baseline that was an average of the absorbance values at the beginning (~430 nm) and end (~460 nm) of the P450 peak. (d) The CO-difference spectrum of CYP4AY1 at an MOI of 10 is shown with corresponding concentration, calculated with an alternate baseline at 480 nm, of 0.165 nmol/mL microsomal isolate.

2.3.2 Single Transfection Optimization of CYP4AY2

CYP4AY2 was optimized in a manner similar to CYP4AY1 optimization. Equipment failure (ultracentrifuge) lead to this optimization being conducted with cell lysate as opposed to an ultracentrifuged microsomal sample, so cytosolic proteins may also have been present. This may have accounted for some of the additional bands that resolved on the digital gel (Figure 2.2a). Figure 2.2b, the graphical representation of the concentrations, revealed an earlier protein productivity plateau, at an MOI of 5, for CYP4AY2 optimization than was seen in CYP4AY1 optimization.

Figure 2.3 illustrates the CO-difference spectra across the range of MOI tested for CYP4AY2 (these were the same samples that were tested by Experion analysis). The P450 peak corresponding to the most active recombinant protein was at an MOI of 5 (0.275 nmol/mL microsomal isolate). The same concentration was obtained at an MOI of 15, but this required three times as much virus. The spectrum for the MOI 7.5 test required a baseline to be estimated by averaging the absorbancies at ~430 nm and ~460 nm as opposed to taking the value at 480 or 490 nm. All spectra exhibited an upswing in the curve approaching 490 nm, which was not ideal for accurate P450 measurement, but had clear P450 peaks. Again, the CO-difference analysis was congruent with the pattern seen in terms of recombinant CYP4AY2 concentration changes with MOI, detected by the automated electrophoresis setup.

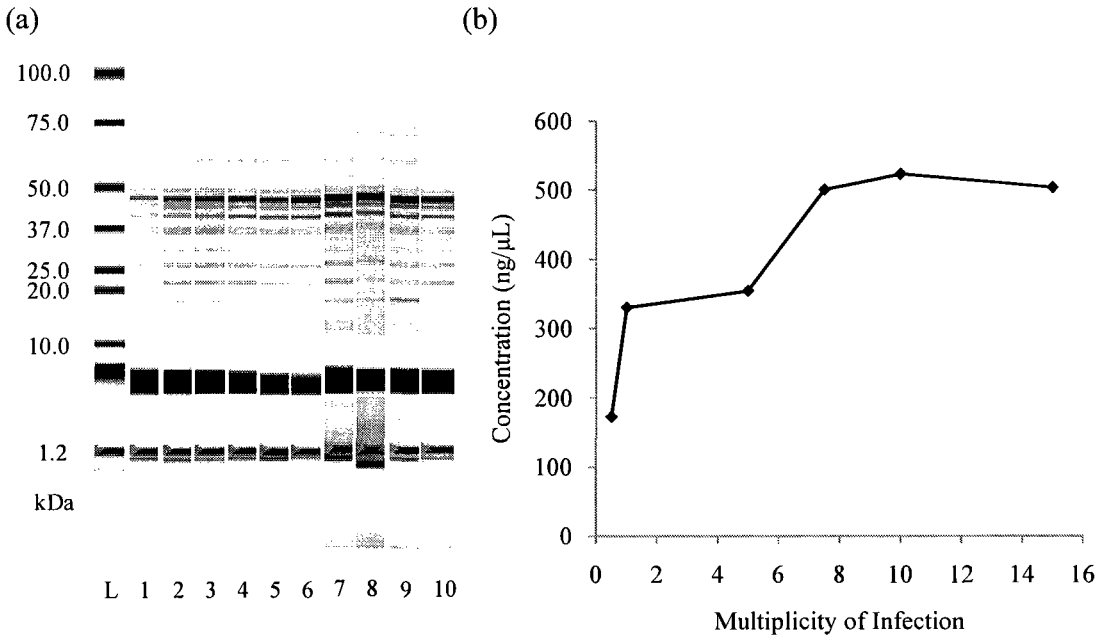


Figure 2.2 – Optimization of CYP4AY2 transfections. (a) Digital gel analysis showed expression across MOI. Lane 1 is MOI 0.5, and lane 2 is MOI 1, followed by pairs of lanes for MOI 5, 7.5, 10, and 15. Recombinant CYP4AY2 resolved at ~48 kDa. Higher isolate concentrations lead to more microsomal bands resolving. A cell control lane was run on a separate Pro260 chip and did not contain recombinant protein (not shown, similar to lane 1 of Figure 2.1a). (b) Graphical representation of Experion-derived CYP4AY2 concentrations across MOI indicated an optimal MOI of 5.

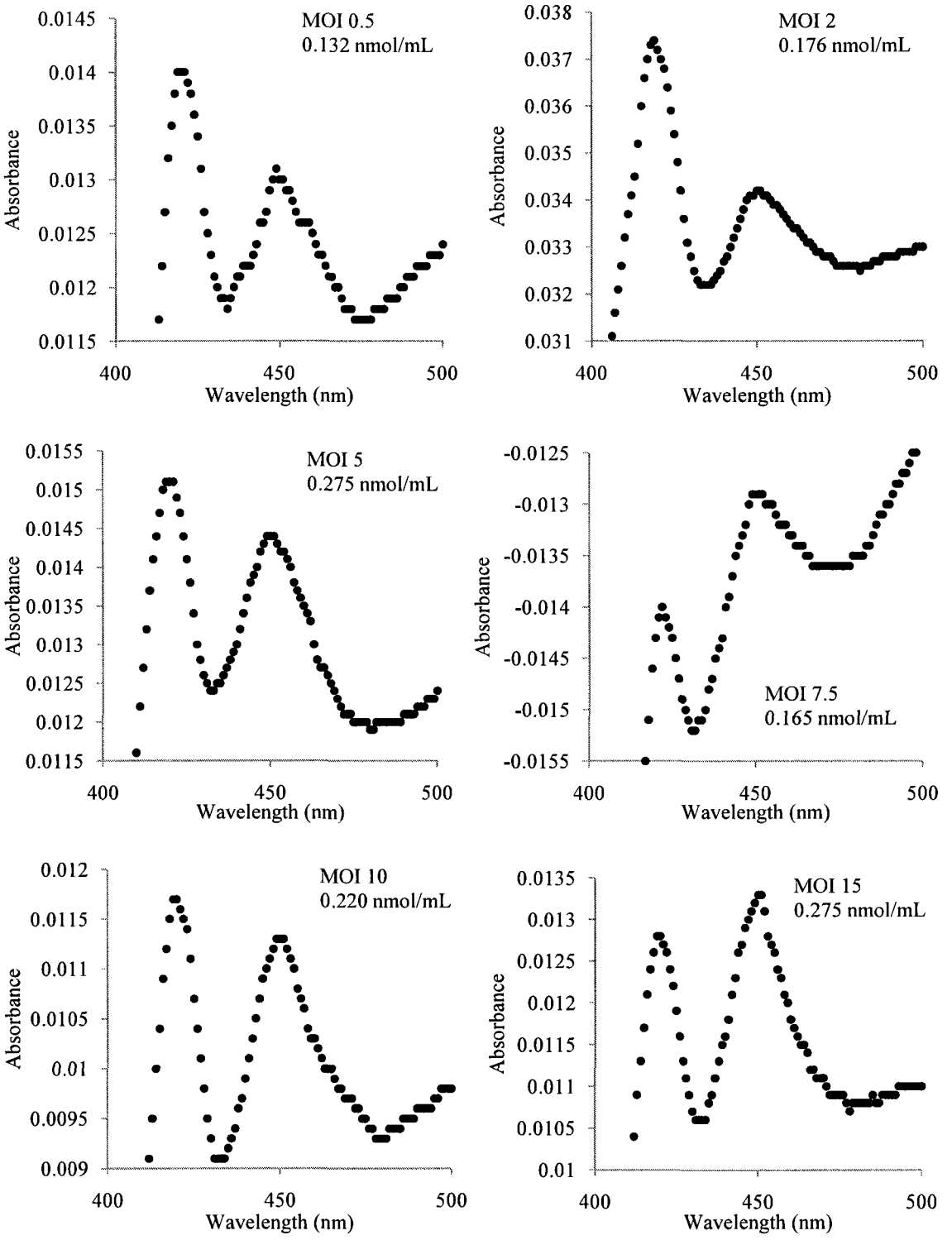
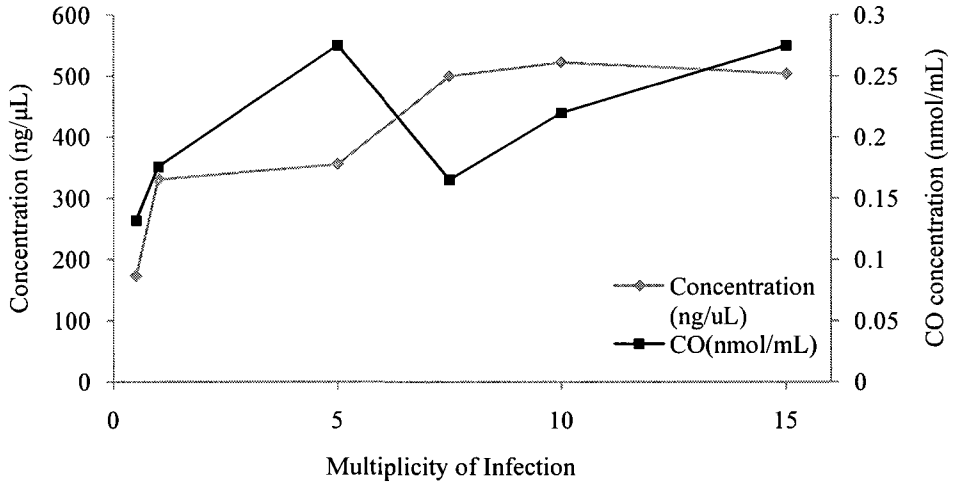


Figure 2.3 - CO-difference spectrum of CYP4AY2 across selected MOI. All spectra are shown with their MOI and associated active CYP4AY2 concentrations. An MOI of 5 has the highest concentration at the lowest MOI at 0.275 nmol/mL (the same as an MOI of 15).

2.3.3 Comparative Analysis of P450 Measuring Techniques

Given that all of the CYP4AY2 CO-difference spectra produced peaks that could be utilized for active P450 concentration calculations, the correlation between the two methods was investigated. The quantities of recombinant CYP4AY2 determined by CO-difference analysis was superimposed onto a graph of the Experion concentration values across MOI (Figure 2.4a). It can be seen that CYP4AY2 protein production, measured by CO-difference analysis, exhibited a sharp decline at an MOI 7.5, but this spectrum had a sharply ascending baseline and is likely an underestimation of the true concentration. Leaving that value out of the analysis, the linear regression of the two measurement techniques (Figure 2.4b) had an R^2 of 0.8123 (calculated in Microsoft Excel 2007). The R^2 value is convincing evidence for a correlation, especially given that there are always going to be variations in overall productivity for each flask of cells; however, a low number of comparisons was used.

(a)



(b)

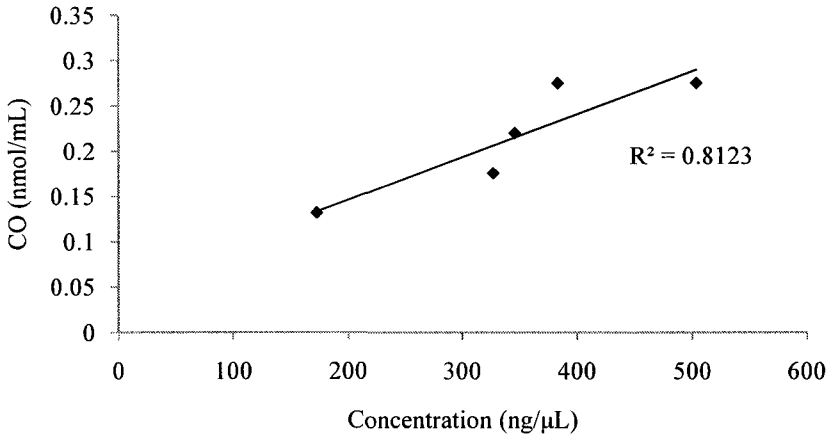
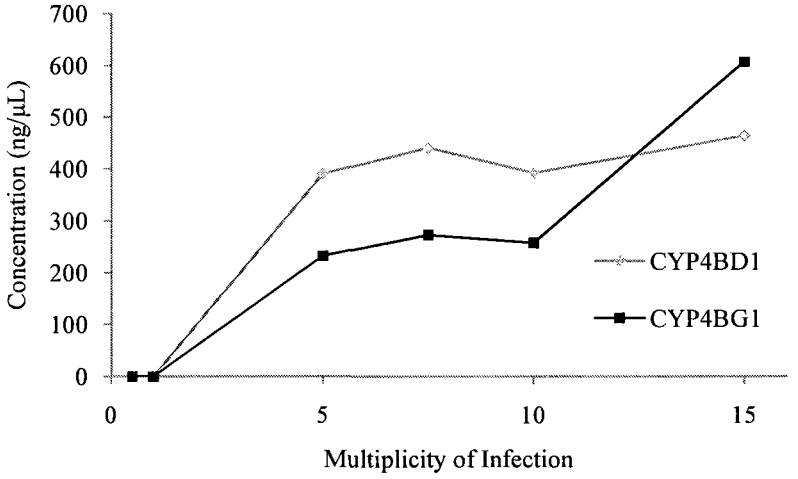


Figure 2.4 - Comparative analysis of CYP4AY2 measurement techniques. (a) Graphical representation of P450 content across MOI with Experion and CO-difference spectral analyses. (b) Regression of P450 measurement techniques, omitting the MOI of 7.5 data pair.

2.3.4 CYP4BD1 and CYP4BG1 Individual Optimizations

Preliminary optimizations of CYP4BD1 and CYP4BG1 were conducted (Figure 2.5a). Time restraints and limited viral stocks prevented further analysis, but it was clear from the CO-difference spectra, which resolved a low P450 peak for CYP4BD1 (possibly representing 0.137 nmol/mL, not shown) and a clear peak (with a normal descent) for CYP4BG1 (0.736 nmol/mL, Figure 2.5b), both at MOIs of 15, that these two constructs would require higher MOIs than the CYP4AY constructs did for optimized expression. It was unclear where the plateau or expression limit was for either of these constructs as the highest MOI tested for both could have been either the beginning or the limit of the ascent in protein production. It is of note that the CO-difference spectrum obtained for CYP4BG1 (Figure 2.5a) is what is normally expected for P450 spectra, with a continuing descent in absorbance after 450 nm. This was likely reflective of a sample with a sufficiently high recombinant protein level for the assay and a decreased turbidity due to solubilization with CHAPS buffer, which subsequently removed interfering proteins. However, the same protocol was employed for the other spectra, so individual differences in samples may have played a large role.

(a)



(b)

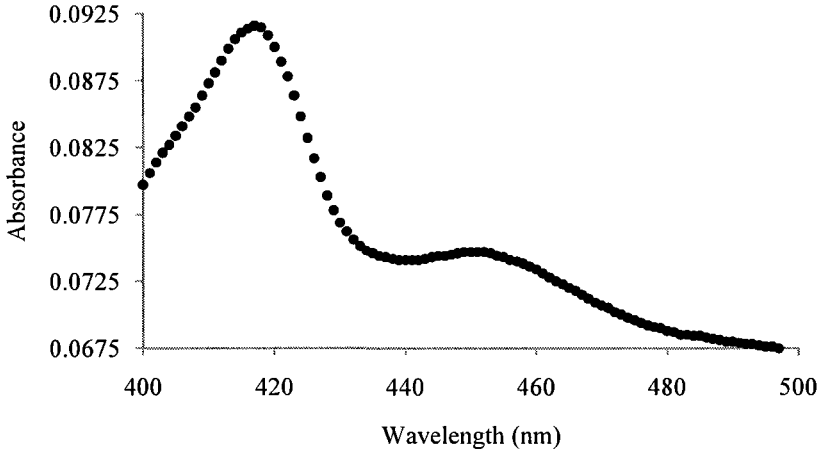


Figure 2.5 – Optimizations of CYP4BD1 and CYP4BG1 transfections. (a) Experion analyses across MOI for both genes revealed an increase at a MOI of 15. (b) The CO-difference spectrum for CYP4BG1 at a MOI of 15 corresponds to an active P450 concentration of 0.736 nmol/mL.

2.3.5 Recombinant MPB-CPR Activity

Activity was calculated from the rate (during linear increase) of absorbance change with the addition of NADPH to a reductase and cytochrome *c* solution, as described in the methods. An MOI of 0.25 was sufficient for expression levels that exhibited electron transfer activity. This indicated that MPB-CPR MOIs of much lower magnitude than those of the recombinant P450s would likely be required for downstream co-expression efforts. The rate calculated from the curve in Figure 2.6 represented 946 nmol cytochrome *c* reduced/ minute per mL of microsomal sample.

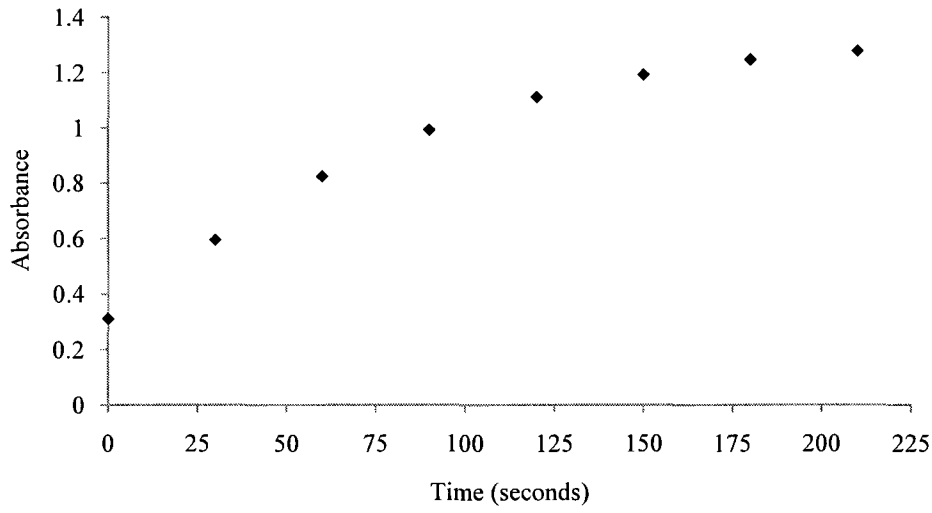


Figure 2.6 - Recombinant MPB-CPR activity. Values obtained during the first 10-20 seconds were utilized to calculate the reduction rate of cytochrome *c* (946 nmol cytochrome *c* reduced/ minute).

2.3.6 Co-expression of CYP4AY1 with MPB-CPR

Experion analyses of CYP4AY1/CPR co-transfections from 1:1 to 1:40 clearly indicated that the less CPR (~69 kDa) virus that was added, the more CYP4AY1 (~48 kDa) was expressed (Figure 2.7a). This was accompanied by a reduced presence of recombinant MPB-CPR. This response is graphically represented in Figure 2.7b. It was also evident that, despite drastically reducing the volume of MPB-CPR virus, the response in recombinant MPB-CPR level was relatively modest. That is, CYP4AY1 levels were much more influenced by changes in ratio than MPB-CPR levels were. The two enzymes clearly have different needs in terms of expression. P450s need to correctly coordinate a haeme group to ensure functionality, which may be why it was first important to optimize the conditions of the P450 before adding the reductase to the system.

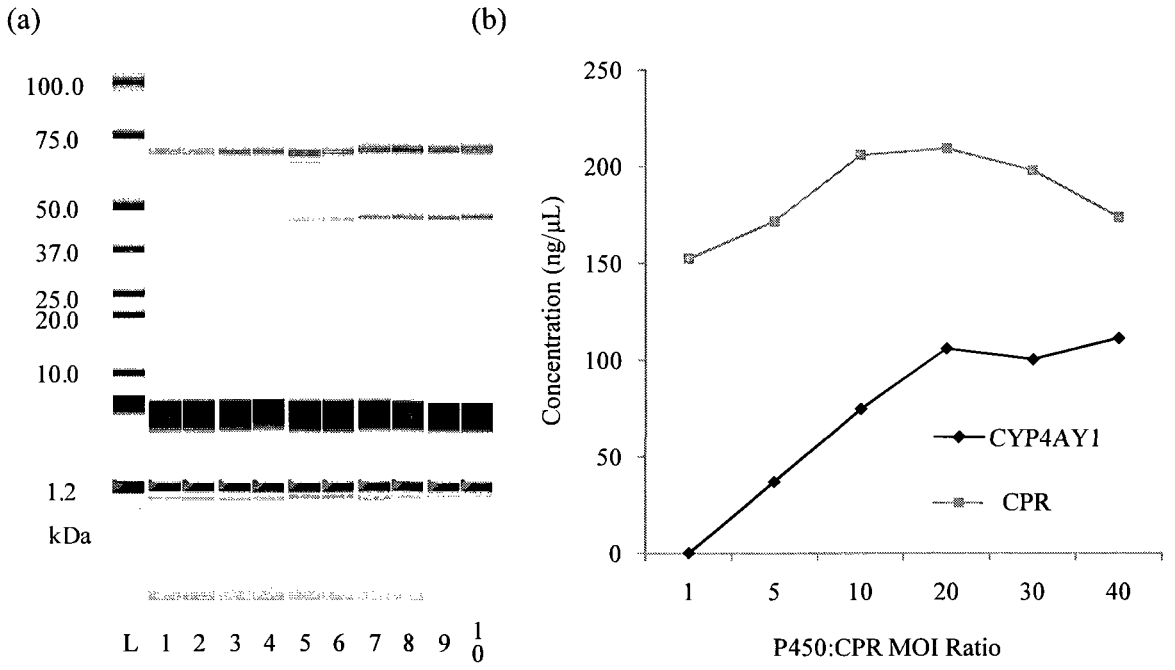


Figure 2.7 - Co-expression of CYP4AY1 with MPB-CPR. (a) Experion analyses across different CPR:P450 co-transfection ratios is shown (CPR resolves at ~68 kDa). Lane 1 is 1:1, and lane 2 is 1:5, followed by pairs of lanes for 1:10, 1:20, 1:30, and 1:40. Cell controls were run separately (not shown) and did not contain recombinant protein. (b) The concentrations for the constituents are graphically represented across ratios.

2.3.7 Co-expression of CYP4AY2 with MPB-CPR

CYP4AY2/CPR co-transfections exhibited a similar outcome to that of the CYP4AY1/CPR co-transfections, with a ratio of 1:40 being selected as the best candidate (Figure 2.8a,b). Again, the additional bands to the microsomal proteins, resulting from the whole cell lysate isolation procedure, are evident. In this case, MPB-CPR protein expression had a much greater response to changes in the MOI of its recombinant virus. That said, one must still consider the magnitudes that the ratios differed by; the sheer amount of CYP4AY2 virus being employed compared to that of the MPB-CPR is still drastically larger, highlighting the efficiency of MPB-CPR production by SF9 cells, compared to that of the recombinant P450s.

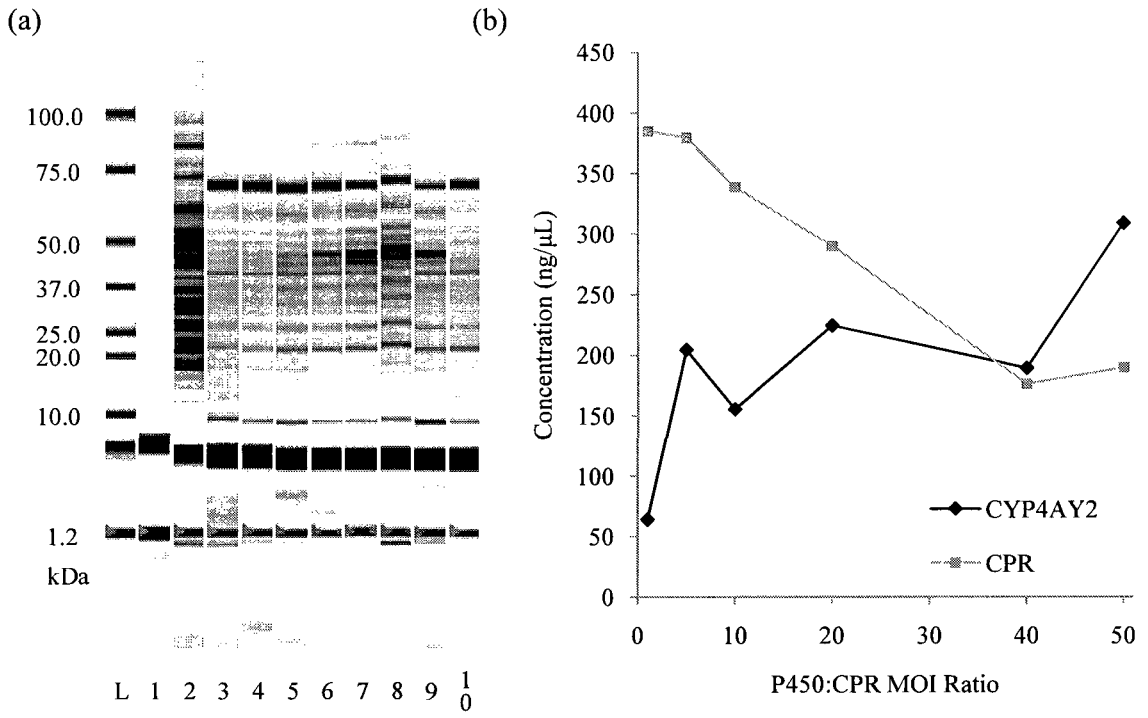


Figure 2.8 - Co-expression of CYP4AY2 with MPB-CPR. (a) Lane 1 is blank, lane 2 is untransfected cell lysate, and lane 3 is 1:1 ratio. Following lanes are 1:5, 1:10, 1:20, 1:30, 1:40, 1:50, and lane 10 is CPR only (MOI 0.25). (b) Concentrations are graphically represented across the MOI ratios.

2.3.8 Functional Confirmation of Co-optimized Constructs

Co-optimization conditions were used to transfect 15 mL suspension cultures (at 1-1.5 x 10^6 cells/mL) and CO-difference spectra were conducted to ensure that functionally folded recombinant P450s were present at detectable levels. The CYP4AY1 spectrum was not smooth, but a peak with calculable values was obtained, corresponding to 0.0660 nmol/mL microsomal isolate (Figure 2.9a). The background was clearly not ideal and the estimate was likely exceedingly conservative, but it was a good sign that the P420 peak was of lower amplitude than the P450 peak. Figure 2.9b for the CYP4AY2 co-optimization corresponded to a concentration of only 0.0879 nmol/mL, when the absorbance at 480 nm was used to estimate background. When an adjustment was made based on the values at ~430 nm and ~460 nm, the calculated amount was then 0.297 nmol/mL.

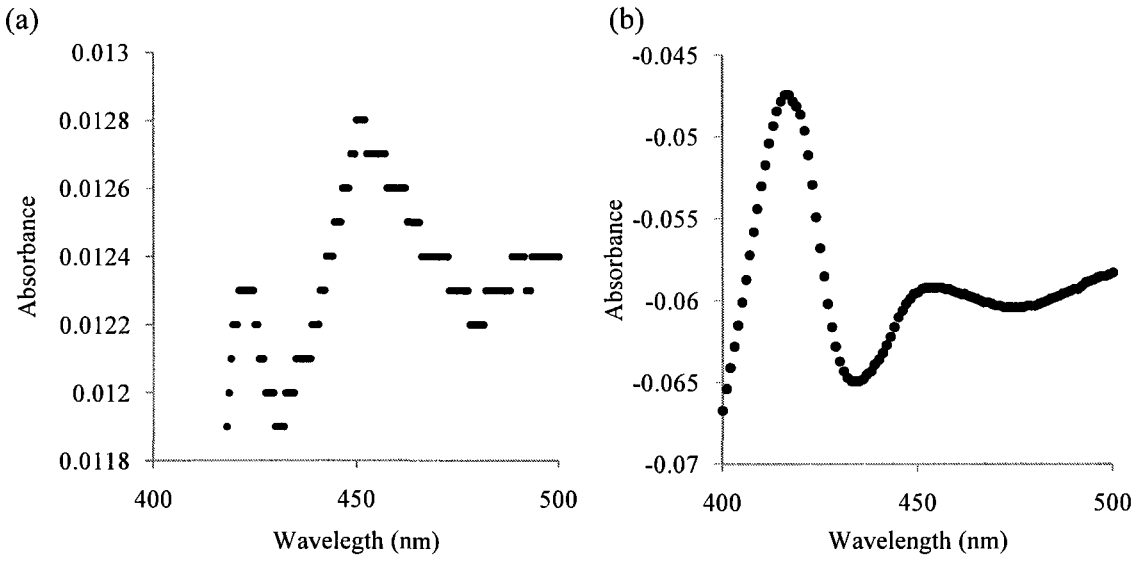


Figure 2.9 - Functional confirmation of co-optimized constructs. (a) The CO-difference spectrum of CYP4AY1/CPR (CPR:P450 MOI ratio of 1:40) exhibited a P450 peak higher than the P420 peak. (b) The CYP4AY2/CPR CO-difference spectrum (also at a ratio of 1:40) exhibited an ascending background, but did have a clear P450 peak.

2.3.9 Comparative Analysis of Reductase Quantification

Reductase activity was also measured for co-optimizations by cytochrome *c* reduction assays. Ratios of 1:40 were chosen for both CYP4AY1/CPR and CYP4AY2/CPR co-transfections because of not only the continued P450 expression, but because CPR rates were 209.52 and 176.3 nmol cytochrome *c* reduced/minute. Again, direct comparisons between systems were considered cautiously, but CPR activity at about 180 nmol cytochrome *c* reduced/minute was optimal for CYP6B1 catalytic activity (Wen *et al.*, 2003).

Calculated activities for each ratio (for both co-transfections with CYP4AY1 and CYP4AY2) were plotted against their concentrations as determined by Experion automated electrophoresis analyses (Figure 2.10). Two CPR measurements were excluded because Experion lanes were shifted and concentrations could not be estimated and two because PECSS software froze and the linear rate could not be measured (thereby underestimating reduction rates). Of the remaining samples, there was an obvious linear relationship between the activities of the MPB-CPR enzyme and its capillary-electrophoresis-measured concentration. The R^2 value (0.7533) indicated a modestly positive correlation. The sample size for comparison was small, but a relationship was evident.

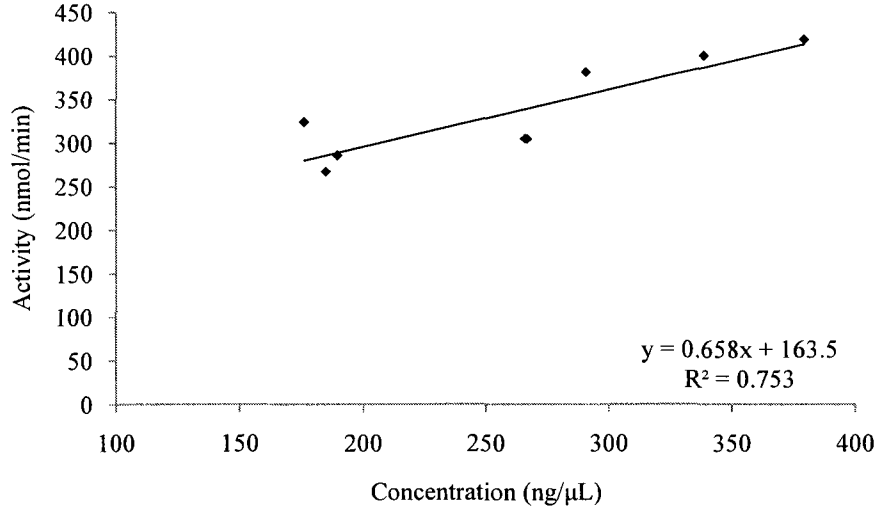


Figure 2.10 - Comparative analysis of reductase quantification. Experion and cytochrome *c* catalytic activity quantification methods for CPR, for both CYP4AY1/CPR and CYP4AY2/CPR co-expressions, are correlated with an R^2 of 0.7533.

2.3.10 Applicability of Adherent Optimizations to Suspension Cultures

Productivity in similarly transfected adherent and suspension cultures of the same number of SF9 cells revealed that they were highly similar in terms of recombinant protein content (Figure 2.11a,b). Percent content was measured because the same amount of virus was used, so the proportion of recombinant protein should have been the same as well (when comparing different MOI, different numbers of cells may be uninfected, so the average amount of protein per cell would be expected to differ). It should be noted that error bars represent replicates for the measurement of percent content, not experimental replicates; only one adherent and one suspension transfection were used. All available lane wells were used to get the best concentration estimate possible, given the possibilities of lane variability and the occasional lane failure. Cell lysate and microsome samples tended to not completely solubilize, despite manually breaking up pellets and an additional 10 seconds of sonication after microsome isolation. The digital gel result of such samples typically manifested as a shifting of bands (as in lane 7, Figure 2.11), which would often be associated with a complete lack of concentration estimates or obviously compromised values (when compared to the others in the run). The two conditions were being directly compared, not an overall trend, so a gauge of variability at a finer scale was desired.

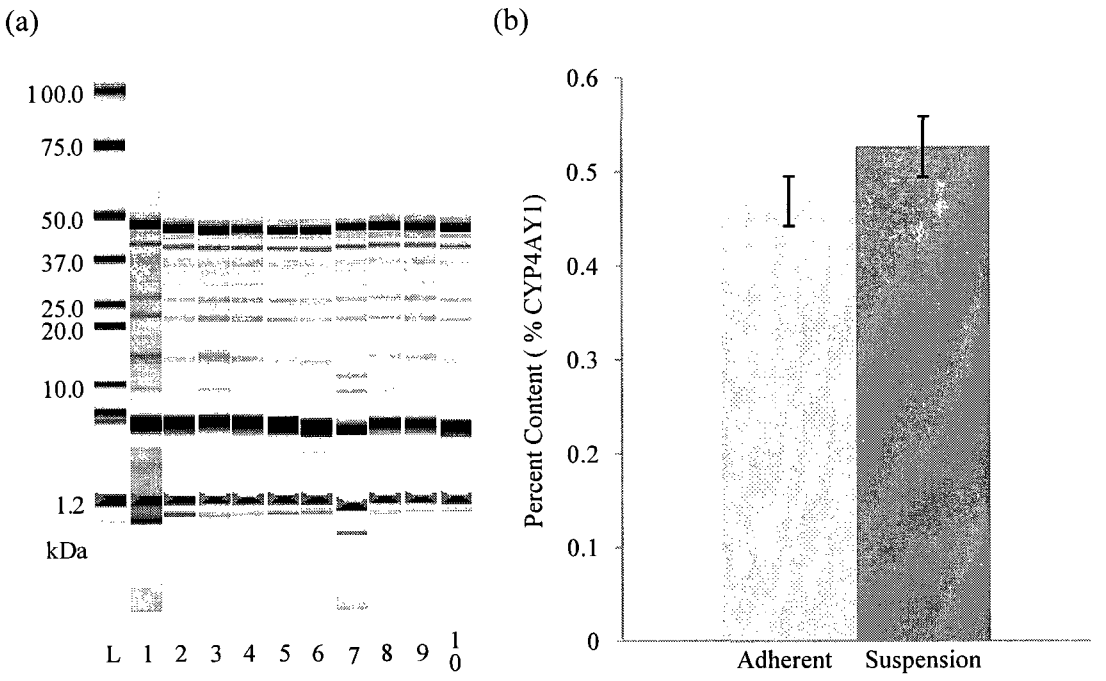


Figure 2.11 - Applicability of adherent optimizations to suspension cultures. (a) Experion analyses show adherent isolates in lanes 1-5 and suspension in lanes 6-10. (b) A comparison of percent content of recombinant protein was used to account for general variability between transfections. One replicate of each was measured five times, with lanes 1 and 7 being eliminated as sample buffer standards were not detected and measurements were compromised. Samples were the same within standard error.

2.3.11 Harvest Time Trial Analysis

Based on concentration estimates from the Experion digital gel output, harvesting at 72 hours resulted in the most protein for both CYP4AY1 and MPB-CPR (Figure 2.12a,b). As before, CYP4AY1 resolved at approximately 48 kDa and MPB-CPR at approximately 69 kDa. A slight increase is exhibited from 96 to 128 hours. This could have been a result of the next generation of virus infecting the rest of the cells in the flask. It is likely that more of the enzyme at later stages is nonfunctional, owing to the progression of cell lysis and increased presence of proteases. The 72 hours post-transfection harvest time was recommended by the manufacturer and was the convention in literature related to recombinant P450 production in SF9/baculovirus expression systems (Wen *et al.*, 2003; Duan *et al.*, 2004; Mao *et al.*, 2008). The purpose of investigating harvest time was not necessarily to modify it, but to ensure that the cell line being maintained was expressing protein as expected and not deviating from established guidelines.

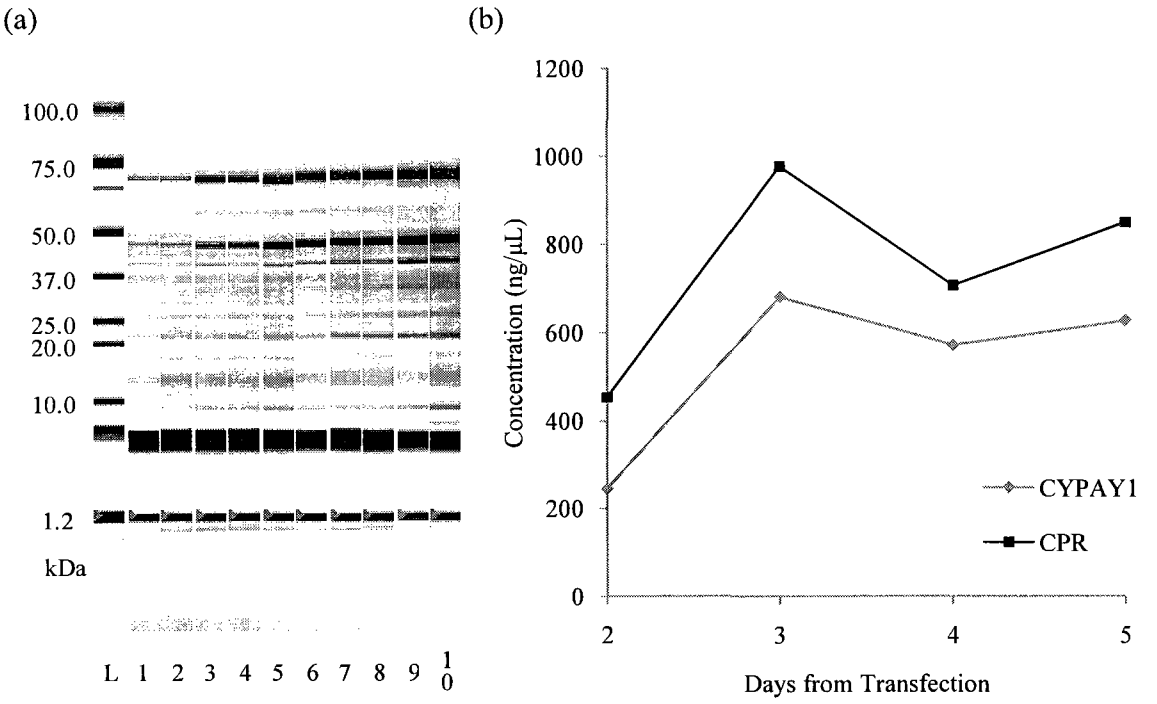


Figure 2.12 - Time trial analysis. (a) Experion analyses of time trials is shown with lanes 1-2 being harvested at day 2, lanes 3-6 being day 3, 7-8 day 4, and 9-10 day 5. (b) A representative concentration was chosen from the replicates and values are graphically represented, where day 3 was confirmed as the best in terms of CYP4AY1 and CPR concentrations.

2.4 Discussion

2.4.1 Viral Pathology

Baculovirus systems offer many advantages for the production of recombinant proteins, but optimization of those systems for maximum protein production is vital. This was demonstrated by the varying optimal MOIs of the constructs. CYP4AY1 was optimally expressed at an MOI of 10, CYP4AY2 at an MOI of 5, and CYP4BD1 and CYP4BG1 at MOIs of at least 15. These differences were present despite all of the P450 ORFs inserted into the baculoviruses being similar in terms of base composition, length, and gene family. Thus, presumably, cellular processing was similar as well. One factor causing these differences may be the pathological effects of the nuclear polyhedrosis viruses employed as baculoviral vectors on the cells (outlined by O'Reilly *et al.*, 1992). In cell culture, enveloped nucleocapsids enter the cell through absorptive endocytosis. Once cellular machinery is commandeered and replication begins, more viruses in nucleocapsids are exocytosed from the plasma membrane. Generally nucleocapsids are thus packaged individually, but multicapsid envelopes are also possible. During late-stage infection, other viral particles assemble in the nucleus and form multicapsid polyhedra, encapsulated in a different envelope. But these occlusion bodies are not considered infective in cell culture. In larvae, it is these occlusion bodies that are released from the polyhedra, when they reach the gut, that begin the infective process; cell culture transfection occurs via the budded virus, although both presumably contain the same DNA (O'Reilly *et al.*, 1992). Nonetheless, there remains the possibility that a variable number of nucleocapsids are in each budding virus, which means, depending on their composition, that similarly constructed and amplified viral stocks

may have varying infectivity. This could be why a simple measurement of viral concentration cannot be taken and the actual activity of each stock must be individually titered.

2.4.2 Measuring Active P450 by CO-Difference Analyses

P450 levels are measured by binding the reduced enzyme to CO and evaluating the amplitude of the peak at 450 nm relative to a baseline measurement. As noted, problems with obtaining a clear baseline, due to turbidity or equipment limitations, can make estimations difficult. A good CO-difference spectrum should exhibit a smooth decline after the peak at 450 nm that allows for accurate measurement of the baseline at 490 nm (as in Figure 2.5b). Very few CO spectra were obtained that had this clear baseline, despite having obvious P450 peaks and despite using CHAPS detergent to reduce turbidity. The dual-beam spectrophotometer available did not seem capable of resolving the samples after 460 nm, to the level required. Many of the spectra obtained had a decline following the top of the P450 peak until around 480 nm, after which the baseline would again begin to climb. Occasionally, baselines began ascending even sooner (e.g. Figures 2.1c and 2.3, MOI 7.5), so using the absorbance at 490 nm would not provide the intended estimate. In cases where there was a substantial decline, the absorbance at 480 nm was used as a conservative (and higher) estimate of what the 490 nm absorbance would be (as in Figure 2.3, MOI 7.5). A more aggressive, but still conservative, estimate that may be used when the baseline is obviously compromised has been proposed based on whole-cell CO-difference spectra with *E. coli* (Johnston *et al.*, 2008). In this case, a value is estimated by taking the average of the absorbance values at the beginning and end of the P450 peak (about 430 nm and 460 nm). That value represented approximately where a normal decline would have fallen and thus replaced the baseline measurement at 490 nm in the calculation for P450 content. This was

done only if the baseline was especially compromised, such that the 480 nm absorbance could not be used. Use of this method resulted in an increased amount of enzyme present in co-expressing CYP4AY2/CPR lysate (Figure 2.9b). That said, if one considered a spectrum lacking a continuing descent after 450 nm and extrapolated how a normal decline would fall, both methods were sufficiently conservative to avoid overestimation. For example, CYP4AY2 at an MOI of 7.5 resulted in a spectrum where the 480 nm estimate could not be used. Once modified, the value did increase, but that data point still fell below the overall trend indicated by the rest of the data (Figure 2.4a). When data manipulation such as this had to take place, I preferred to error towards lower estimates for these optimizations to avoid false confidence in the levels of functionally active recombinant P450. Thus, this manipulation was reasonable.

2.4.3 Optimizations of Individual Transfections Using the Experion System

The Experion capillary electrophoresis system with Pro260 chips provided size estimates with concentrations for each band on a digital gel output. This information was used for both P450 optimizations and co-optimizations including MPB-CPR. Although specific concentration values were provided by the instrument, these values were not considered useful when measuring protein content in cell culture. That is to say, I did not intend to aim for a certain concentration range when optimizing. Protein concentrations will vary with each set of cells transfected. Also, the final volume that a microsomal pellet will be resuspended in depends upon the general size of that pellet. Because of this, concentration data from separate trials, and especially from separate viral constructs, should not be directly compared in terms of finite values. The concentration data were simply viewed across a range of treatments to observe general trends during expression of the recombinant proteins.

An MOI of 10 was chosen as optimal for CYP4AY1 based on comparative Experion concentration data (Figure 2.1a,b). This MOI was chosen as opposed to an MOI of 15 because the highest possible MOI may be too close to the expression capacity of the cells and improperly folded P450 may comprise a portion of the isolate. The higher viral infection level could have lead to premature cell lysis, releasing recombinant proteins into the media and exposing them to proteases. Harvest should thus occur prior to visible cell lysis. Additionally, since a recombinant MPB-CPR virus will also be added to the system, cells need to be below viral capacity at this point to allow for the second viral construct addition. CO-difference analysis revealed the best result at an MOI 10, corresponding to 0.165 nmol/mL (Figure 2.1d); the only other sample exhibiting a peak had an MOI of 7.5, and it was of very low amplitude, corresponding to 0.0824 nmol/mL of microsome (Figure 2.1c). Despite it having a P420 peak of lower amplitude than the P450 peak, the confirmed amount of active P450 was considered too low to carry through to co-transfection. If the resulting active P450 concentrations had been the same for both of these MOI, 7.5 would have been chosen because of the lower P420 peak, which indicated less inactive enzyme. However, it was suspected that those levels were low because overall production itself was low.

CYP4AY2 had an optimal MOI of 5, but showed a different optimization pattern than that of CYP4AY1. Instead of sharply dropping off, Experion analyses showed increasing amplitude of the corresponding P450 peak with increasing MOI that reached a plateau at 0.275 nmol/mL (Figure 2.2a,b). For the same reasons as with CYP4AY1, the lowest MOI required for active P450 protein production was chosen. CO-difference spectra had detectable P450 peaks for all samples (Figure 2.3), with an MOI of 5 exhibiting a clear P450 peak representing 0.275 nmol/mL. The P420 peak was reduced in height compared to other

spectra in all samples, except for MOIs 7.5 and 15, which had higher P450 peaks relative to P420, but lower or equal concentrations of active P450 (Figure 2.3). An MOI of 15 resulted in the same active P450 concentration as an MOI of 5. This pattern agreed with the Experion data in that an MOI of 5 was about as productive as the higher MOIs (Figure 2.2a,b). Experion data was clearly useful for P450 optimization and reflected the findings of the CO-difference analyses, as long as it was considered that P450 functionality could be compromised at higher MOI because of the negative impact on cell health.

Experion and CO-difference data for CYP4AY2 optimization seemed to follow the same pattern (Figure 2.4a). Since all CYP4AY2 protein optimization samples exhibited measureable P450 peaks and the same samples had representative Experion-derived concentrations, an opportunity to compare them directly was presented. A simple linear regression of the two measurement techniques (Figure 2.4b) had an R^2 value of 0.8123, which supported the use of Experion analyses to infer recombinant P450 presence. Considering the variability inherent in a cell system, this is strong evidence for the use of this system. So long as the pathology of viral infection and the subsequent requirement for the addition of CPR is kept in mind, informed decisions can be made regarding the appropriate MOI to use even with the absence of clear CO-difference spectral measurements.

CYP4BD1 and CYP4BG1 seemed to have required higher MOIs for the expression of adequate amounts of active enzyme. An MOI of 15 lead to an increase in recombinant protein production for both (Figure 2.5a), but since a higher MOI was not tested, it is unknown whether this is the beginning of an upward trend or close to a productive plateau. An MOI of 15 produced a discernable P450 peak for CYP4BG1, representing 0.736 nmol/mL and exhibiting a normal spectrum with the expected decline following the P450

peak (Figure 2.5b). The spectrum for CYP4BD1, however, had an interfering peak at 460 nm and high background noise, so positive identification of the P450 peak was occluded. With background adjustments, it is possible that this sample contained 0.137 nmol/mL.

2.4.4 Activity of MPB-CPR Expressed in SF9 Cells

MPB-CPR was supplied in an *E. coli* vector, however the full-length cDNA was moved into a virus for SF9 expression instead. Although functional CPR can be expressed in *E. coli*, P450-CPR interactions in functional assays could conceivably be compromised because the two proteins would not necessarily be in as close proximity. Like microsomal P450s, CPR is also a membrane protein, and both are required for proper activity (reviewed by Omura, 1999). When CPR and P450 are co-expressed, they should be located in the same endoplasmic reticulum and thus should be isolated under the same buffer conditions. Another benefit of this method is that the maintenance of an additional cell line for CPR production is not required. As demonstrated by the curve in Figure 2.6, MPB-CPR was capable of reducing cytochrome *c* when produced in SF9 cells at an MOI of 0.25. A reduction rate of 946 nmol of cytochrome *c* reduced/minute was calculated using absorbance values from the first 10-20 seconds during linear increase (Guengerich *et al.*, 2009). This confirmed the activity of this recombinant enzyme when expressed in SF9 cells. This amount was comparable to higher expression levels. For example, activity around 250 nmol cytochrome *c* reduced/minute per mL was associated with a higher catalytic rate for recombinant CYP6B1 activity, also produced in SF9 cells (Wen *et al.*, 2003).

2.4.5 Co-expression Optimizations of CYP4AY1 and CYP4AY2 with MPB-CPR

Optimal P450 MOIs for CYP4AY1 and CYP4AY2 were used in conjunction with varying amounts of CPR virus. Addition of CPR virus is known to reduce the expression of

P450 proteins in insect cell culture (Wen *et al.*, 2003; Duan *et al.*, 2004). CYP4AY1 expression improved when CPR MOI was reduced (Figure 2.7a,b). A MPB-CPR:P450 ratio of 1:40 was selected based on the presence of more P450 at that amount. This was also the ratio chosen for CYP4AY2, but a higher ratio (1:50) was also tested (Figure 2.8a,b). Although Experion analyses indicated that a ratio of 1:50 lead to more P450 protein production overall, experiments by Wen *et al.* (2003), with *Papilio polyxenes* Fabricius (black swallowtail butterfly) CYP6B1, indicated that higher amounts of CPR resulted in more enzymatic turnover than higher amounts of the P450. Obviously active P450 was integral to substrate turnover, but given they were both present and functional, increasing reductase levels increased the reaction rate; thus, it seemed reasonable to choose the less extreme ratio.

Reductase activities for co-optimizations were also measured by cytochrome *c* reduction assay. Specifically, in the 1:40 ratio systems selected as optimal, the activities of the MPB-CPR were calculated to be 209.52 and 176.30 nmol cytochrome *c* reduced/min, for CYP4AY1/CPR and CYP4AY2/CPR co-optimizations, respectively as this was comparable to the optimal amount for CYP6B1 activity, ~180 nmol cytochrome *c* reduced/min (Wen *et al.*, 2003). There was a linear correlation between concentration and activity measurements, with an R^2 value of 0.7533, which supported the use of Experion analysis to measure CPR presence. Considering the variability in cell productivity with each transfection, the digital gel output offered a reasonable estimate that used a tenth of the sample volume. Valuable microsomal samples and expensive NADPH are needed for conducting substrate assays; due to the limited amount of virus and cells for transfection, this made economical sense.

Once optimized, the cotransfection conditions were tested using 15 mL suspension cultures to ensure that active P450 was still present despite the addition of MPB-CPR virus to the system. I wanted to ensure that recombinant enzyme produced under optimized conditions would be suitable for assays by testing the selected MOI in suspension culture (Figure 2.9a,b). CO-difference analysis of these samples revealed CYP4AY1 was present at 0.0660 nmol/mL and CYP4AY2 at 0.297 nmol/mL, and considering the poor background resolution, these values were likely underestimations. The amounts of recombinant P450 and should be sufficient for catalysis. In terms of functionality, the highest turnover rates for CYP6B1 were attained with about 0.06 to 0.20 nmol P450/mL (Wen *et al.*, 2003). To further investigate the applicability of adherent optimizations to scaled-up suspension culture, a direct comparison was also used.

2.4.5 Applicability of Optimizations to Transfections in Suspension Cultures

As per the manufacturer's directions, and literature related to recombinant P450 expression (for an example, see Wen *et al.*, 2003), optimizations were carried out in smaller-scale adherent culture and protein was harvested at 72 hours post-transfection. Optimizations were conducted in 2, 5, or 15 mL volumes in adherent culture (2, 3, and 8 x 10^6 cells), although 15 mL volume transfections were generally required for CO-difference analysis. Large-scale production to obtain enough P450 for substrate analyses required the use of suspension culture. To ensure that optimizations with adherent culture were applicable to suspension-culture, especially since shaker cultures had just been initiated in the lab and quality was not assumed, the two were compared directly by transfecting the same number of cells with CYP4AY1 virus (MOI 10). The two methods did not differ in their overall outcomes. Samples were run in multiple lanes during Experion analysis to account for

instrument variation and the viscosity of the samples, and the average CYP4AY1 percent content of adherent and suspension transfections were the same within standard error (Figure 2.11a,b).

2.4.6 Confirmation of Recommended Harvest Time

Time trial analyses, based on Experion measurement only, confirmed that 72 hours was optimal for transfection harvest (Figure 2.12a,b). These results are in agreement with the manufacturer's directions and other recombinant P450 research utilizing SF9/baculoviral systems (Wen *et al.*, 2003; Mao *et al.*, 2008; Duan *et al.*, 2004). Thus, cell lines were healthy and uniform enough for analysis. Scale-up to suspension culture with the given protocol should accommodate the amounts of virus indicated by the optimization and co-optimization results and produce microsomes containing active forms of both recombinant P450s and CPR.

2.5 Conclusions

Even expressing CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 alone required individual optimization conditions, despite the overall similarities between the viral constructs. The effective co-expression of CYP4AY1 and CYP4AY2 with MPB-CPR in a SF9/baculoviral system was further complicated. A balance needed to be met in order to produce adequate amounts of both P450 and CPR enzymes, while still ensuring their functionality and maintaining the integrity of the cells. Although traditional measurements based on spectral properties can effectively measure enzymes in this system, there are inherent difficulties related to sample composition and equipment capability. These difficulties can be circumvented with the use of a capillary electrophoresis system, as long as the pathology of transfection and progression of expression is factored into decisions regarding MOI selection. Although it is tempting to relate optimizations to absolute concentration values, such correlations are not necessarily precise as they do not reflect the presence of active enzymes. SF9 cells are variable in their production levels with each transfection and a specific concentration does not imply functionality. Although capillary electrophoresis did not distinguish active P450s from the inactive P420 form, the overall trends in productivity provided the information needed to optimize the MOIs in this expression system. P450 expression increases towards an optimal MOI and then plateaus or declines as the cells become limited at higher MOI. MPB-CPR expression seems to be more robust as very low MOI produced measureable amounts of active enzyme. Cell cultures infected with viruses are living systems that are impacted by numerous, immeasurable factors. Thus, any given system will not necessarily fit into derived parameters.

There are several practical advantages associated with the Experion system. A CO-difference spectrum analysis requires 50 times more sample to conduct an assay, and a maximum of approximately six can be conducted in one working day (per researcher). Use of the dangerous CO gas requires supervision, so work can only be conducted during normal laboratory hours and if someone else can attend. If samples are too turbid to resolve the P450 peak, or an accurate baseline, they are effectively wasted, no useful information is gained, and analysis must be repeated. Six CO-difference analyses roughly use the same volume as about two downstream substrate tests, which can limit the amount of substrates tested in a given batch. Each batch of cells for use in protein expression requires several days to a week of cell amplification (assuming consistent growth rate). Transfections require three days of incubation and a minimum of four hours to isolate the microsomes (assuming low volume scale-up). Additional fresh lysis buffer must also be prepared daily for sample dilution. CPR analysis is not as involved, but still requires about six times the volume of microsomal sample to conduct and NADPH is expensive. Experion Pro260 analysis takes 15 minutes to prepare, requires 4 µl of sample, and takes roughly 30 minutes to run ten samples; analysis is fast enough to take place on the same day as microsome isolation and does not substantially impact the number of possible substrate tests on that batch. Both total protein concentration and individual estimates are simultaneously obtained. The savings in terms of cell culture volume, media, buffer components and technician hours are substantial.

Spectral analyses are integral to confirming activity of P450s, but this automated electrophoresis system augments and complements those traditional methods. In particular, in the case of CYP4AY1, where CO-difference spectra were not initially attainable for any samples, the capillary electrophoresis system allowed for optimization decisions to be made

pending other measurements. The expected MOI for both CYP4AY1 alone, and the CPR:CYP4AY1 ratio that were predicted without CO-difference analysis were the same. As viral titer decreases with time, and SF9 cell lines can be delicate, taking time and expense to maintain or amplify, it is important to be able to make effective and rapid decisions to move a project forward. In other words, working with a cell line and virus requires one to optimize the system as quickly as possible, lest the cell line becomes less viable or optimizations are skewed by changing viral titers. The optimization decisions outlined here were the same as if CO-difference analysis had been lacking. These data both support current optimization protocols and provide a time and money-saving alternative. With such optimizations attained, scale-up can take place to produce active recombinant P450 and CPR-expressing microsomes for functional analysis with potential substrates.

3.0 *Ips paraconfusus* CYP4 Substrate Predictions and Functional Assays

3.1 Introduction

A typical bark beetle has several physiological concerns related to host colonization. To ensure their survival, individuals of the pioneering sex must attract other conspecifics, including potential mates (Borden, 1982; Byers *et al.*, 1979). Aggregation pheromone production is the key to ensuring that the host is sufficiently overwhelmed and weakened to permit both colonization and the maturation of potential offspring (Paine *et al.*, 1997). Colonizing a host denotes the end of the need for flight muscles for further dispersal, which can release energy for other uses (Unithan & Nair, 1977). The pioneering sex must construct a nuptial gallery in the host tree phloem, where it is exposed to toxins contained in the host resin (Byers, 1981). Ultimately, however, both sexes need to withstand the onslaught of host defences as both sexes reside in the resin-saturated phloem following host colonization. Female insects must produce eggs- energy and components for which may be drawn from their fat bodies (Casals *et al.*, 1997). There are a multitude of pathways involved in these processes, including the synthesis and degradation of pheromones, hormones, fatty acids, and the neutralization of xenobiotics. Insect cytochromes P450 (P450s) have been implicated, in some capacity, in all of those pathways (Feyereisen, 1999).

The extraordinary range of possible functions carried out by P450s makes them interesting subjects of research, but makes predicting possible substrates a challenging endeavor. Quantitative analyses of gene expression patterns with varying stimuli can provide hypotheses pertaining to the ecological needs of the organism and thus give hints as to likely substrates for the gene products. The ability to hypothesize is clearer in those genes that have

a sharp increase in expression with a treatment. For example, real-time PCR revealed that *Ips pini* (Say) CYP9T2 was induced in males and predominately expressed in the midgut (Sandstrom *et al.*, 2006). Functional characterization of recombinant CYP9T2 revealed that the gene product hydroxylated the monoterpene myrcene to the pheromone component ipsdienol. *Ips confusus* (LeConte) CYP9T1 is 94% similar to CYP9T2, with a similar inductive response to feeding and it also converted myrcene to ipsdienol when expressed in SF9 cells (Sandstrom *et al.*, 2009). Transcripts of *Ips paraconfusus* (Lanier) CYP9T1 increased in fed males 85,030x at 8 hours and 25,938x at 24 hours (only 152x in fed females) (Huber *et al.*, 2007). Although providing the largest transcript response in that work, CYP9T1 was not the only novel P450 induced with feeding; several others were identified.

Other CYP4s that exhibited differential transcript levels with feeding have been identified (Huber *et al.*, 2007). Those with available full open reading frame (ORF) clones- CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1- were selected for heterologous expression and functional characterization due to transcript levels higher than baseline with feeding. Two increased only in males to levels of 28.6x and 39.6x (CYP4AY1), and 5x and 2.8x (CYP4BD1), at 8 and 24 hours, respectively. At the same time points, CYP4AY2 transcript accumulation in males were 2.5x and 1.5x, respectively, while expression only increased 4.5x at 8 hours in females. CYP4BG1 expression for both sexes increased 2.5x and 6.3x at 8 and 24 hours respectively after feeding, relative to unfed individuals. The high, constitutive expression of CYP4G27 in both sexes, despite treatment, warranted construction of its recombinant baculovirus for heterologous expression and functional analysis. Given the gene expression responses of four of the CYP4 P450s mentioned above to feeding stimuli, monoterpenes were considered a logical substrate group to test. In addition to myrcene being

converted to ipsdienol by the action of P450s in some bark beetle species, it is also known that α-pinene is converted to verbenol, which is a pheromone component of *I. paraconfusus* (Byers, 1981). The conversion of α-pinene to verbenol is an oxidation reaction similar to the myrcene-to-ipsdienol conversion. 3-Carene, β-phellandrene, and limonene are also present in resin, among other monoterpenes, and are known insect toxins (Phillips & Croteau, 1999) that should require processing by the insect, potentially by P450s.

Based on the expression analyses by Huber *et al.* (2007), and the role of chemical defense in a host's response to bark beetles, monoterpenes in pine resin were investigated as possible substrates. A review of antennal responses of tree-killing bark beetle species revealed those host components that the insects could detect, which could indicate compounds that were ecologically relevant for host choice. Pureswaran *et al.* (2004b) found that *D. ponderosae, D. rufipennis, D. pseudotsugae,* and *Dryocoetes confusus* had antennal responses to α-pinene, β-pinene, 3-carene, limonene, myrcene, and terpinolene. All but *D. ponderosae* also exhibited an antennal response to bornyl acetate. A review of the literature related to monoterpene toxicity was provided by Raffa and Berryman (1987). For *Scolytus* and *Dendroctonus* genera, death was induced by α-pinene, β-pinene, 3-carene, limonene, and myrcene exposure. α-Pinene, β-pinene, 3-carene, limonene, and myrcene have toxic effects on *D. frontalis* adults, and α-pinene, limonene and 3-carene are ovicidal to *D. ponderosae* (reviewed by Raffa & Berryman, 1987). Although not a bark beetle, a study utilizing monotepene fumigants to control a kidney bean pest, *Acanthoscelides obtectus* (Say), found linalool to be the most toxic and inhibitory in terms of oviposition and larval development (Regnault-Roger & Hamraoui, 1995). Also related to the defence mechanisms of host trees is lignification. *Dendroctonus micans* adults and larvae were found to be less successful in

terms of gallery construction, survival and growth, respectively, in more lignified bark (Wainhouse *et al.*, 1990). *I. paraconfusus* has been cited as producing phenylethanol as a pheromone component (Renwick *et al.*, 1976). Male *I. pini* boring into host material released toluene and 2-phenylethanol (Gries *et al.*, 1990). Males also performed this conversion when treated with L-phenylalanine in the lab; L-phenylalanine is a lignin precursor. Any of these compounds could be oxidized by a P450, either to detoxify, metabolize, or produce pheromone components. Table 3.1 summarizes these findings.

Table 3.1 - Ecological relevance of selected compounds to bark beetles. All monoterpenes are host pine resin constituents.

Compound	Ecological Relevance
α-pinene	pheromone precursor, antennal response in tree-killing bark beetles, toxic to adults, ovicidal, fatal at high concentrations
β-pinene	antennal response in tree-killing bark beetles, toxic to adults, fatal at high concentrations
3-carene	antennal response in tree-killing bark beetles,ovicidal, fatal at high concentrations
limonene	antennal response in tree-killing bark beetles, toxic to adults, ovicidal, fatal at high concentrations
linalool	toxic fumigant properties in *A. obectus*
myrcene	pheromone precursor, antennal response in tree-killing bark beetles, toxic to adults, fatal at high concentrations
terpinolene	antennal response in tree-killing bark beetles
bornyl acetate	antennal response in tree-killing bark beetles
L-phenylalanine	possible pheromone precursor, host precursor to lignin which inhibits colonization and brood success

At this point, only one pheromone-producing P450 has been confirmed (CYP9T2 and its ortholog IcCYP9T1) (Sandstrom *et al.*, 2006, 2008). Given that host-colonizing bark beetles are exposed to many xenobiotics, that they produce different enantiomeric ratios of pheromones, and the diversity of the pheromone components themselves, there are many detoxifying and pheromone-producing P450s that have yet to be elucidated. Furthermore, besides those processes, there are other roles in fatty acid, hormone, and ecdysteroid pathways in which P450s such as CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and CYP4G27 might be found. As such, the number of potential substrates for P450s in bark beetles is extremely large.

Substrate candidates for the CYP4 P450s mentioned were selected by narrowing down potential functions as much as possible. Peptide sequences for five selected CYP4s were entered into the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov) to search for other P450s with amino acid identity. Although this was done when these P450s were first characterized, new entries into the database may provide clues to augment the transcript accumulation data (Huber *et al.*, 2007). A recently developed mountain pine beetle (*Dendroctonus ponderosae* Hopkins, MPB) expressed sequence tag (EST) database was also searched for open reading frames (ORFs) of orthologous genes (Keeling *et al.*, in prep.). Huber *et al.* (2007) utilized whole body extracts for analysis. The source libraries of ther MPB reads most similar to those CYP4s may provide some tissue localization analogies. Finding similar P450s with known or suspected functions and tissue localization extrapolations may augment expression data, and support or refute the hypothesis of monoterpenes as potential substrate candidates. Even low identity

matches may provide some information if one assumes the relationship derived from evolving similarly.

Based on expression profiles and the aforementioned *in silico* analysis, CYP4AY1 and CYP4AY2 were co-expressed with MPB-CPR in scaled-up suspension cultures for functional characterization. Isolated microsomes were subjected to functional assays with α-pinene, β-pinene, limonene, linalool, 3-carene, myrcene, terpinolene, bornyl acetate, and L-phenylalanine. Despite the indications that monoterpenes or L-phenylananine could be substrates, CYP4AY1 and CYP4AY2 did not oxidize any of the tested compounds.

3.2 Materials and Methods

3.2.1 Similarity Searches and Localization Estimates

In order to identify similarities to other P450s, and to help predict possible substrates based on the suspected functions of the matches, the NCBI database was searched. Peptide sequences corresponding to the ORFs of CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and CYP4G27 were entered into pBLAST on NCBI (Altschul *et al.*, 1990). The search was limited to reference sequences and repetitive matches were left out if uninformative. This was done to focus on P450s that had been assigned to a family or at least given a suspected grouping. Names or noted similarities to named P450s were recorded, along with the percent amino acid identities of each match to the searched *I. paraconfusus* CYP4.

3.2.2 Tissue Localization Analogies Based on Identities to *Dendroctonus ponderosae* ESTs

To supplement expression data with tissue localization information to improve substrate selection, *I. paraconfusus* gene names were entered into the MPB EST database (Keeling *et al.*, in prep.) to conduct keyword searches for matches to contigs derived from several EST libraries (Table 3.2). Each contig is composed of individual DNA sequencing reads that overlap, and represents a segment of genomic DNA. The number of reads and their source libraries were noted for each contig, along with the associated E-value of the contig match. Libraries with no representative reads for the searches included those derived from antennae and normalized pupae and adult, as well as both un-normalized midgut and fat body (MGFB) libraries. Briefly, normalized libraries account for reads that occur more often than others and reduce the redundancy of common transcripts in EST libraries.

Table 3.2 - Treatment of mountain pine beetles used for construction of EST libraries. *Ips paraconfusus* P450 peptides were searched against the EST libraries. Libraries derived from larvae, antennae, and adults with juvenile hormone (JH) treatment did not have normalized counterparts, and adult head and cold-hardening larvae libraries did not have un-normalized counterparts.

Library	Tissue Description
Larvae	whole, mixed instar larvae extracted from stored host bolts
Larvae/cold	whole, mixed instar larvae extracted from host October/November 2008
Pupae	whole, untreated pupae extracted from host
Adult	whole teneral adults extracted from host
Adult/head	heads of untreated teneral, JH treated, and fed adults
Adult/JH	whole adults, abdominal venter treated with 2 μg JH, incubated 24 hours (RT)
Adult/Terp	whole adults exposed to (+)-3-carene, myrcene, (racemic)-α-pinene, (-)-β-pinene, or (-)-verbenone vapours for 24 hours
Antenna	antennae removed from adults extracted from host
MGFB/JH	MGFB removed 1, 5,or 15 hours after treatment with 10 μg JH to abdominal venter
MGFB/fed	MGFB removed 24 and 40 hours, respectively, from females and males fed on host

3.2.3 Scaled-up Suspension Culture Protein Production

Suspension cultures for scaled-up protein production were maintained as in Chapter 2 (2.2.3 SF9 Cell Maintenance). SF9 shaker cultures were initiated from a 70% confluent, 75 cm^2 flask of adherent cells (~48 hours since passage). The side of the flask was impacted against the palm and adherent cells were gently sloughed off with a 2 mL plastic pipette. Cells/media were transferred to a 25 mL Erlenmeyer flask and media added to a 15 mL total volume with 0.01% pluronic F68 to reduce shearing. Cells were incubated (27°C) on a rotary shaker (105 rpm), and flasks were topped with tin foil to permit air-exchange. Cells were grown to 2 x 10^6 cells/mL in 48 hours, before being transferred to a 50 mL Erlenmeyer flask, and media was added to a 30 mL total volume. Suspension lines were maintained for up to one month before a new line was initiated from adherent culture.

Once transfection conditions were optimized and viral MOI selected, suspension cultures of 15-30 mL at 1 to 2 x 10^6 cells/mL were transfected with the viral volumes calculated from optimizations with recombinant P450s and MPB-CPR (see Chapter 2, sections 2.3.6, 2.3.7). CYP4AY1 was transfected at an MOI of 10 and CYP4AY2 at an MOI of 5 (both with MPB-CPR at MOIs corresponding to the CPR:P450 ratio of 1:40). Shaker cultures were capped and completely covered with aluminum foil to protect virus from light (105 rpm, 27°C). As before (Chapter 2, section 2.2.5), hemin was added at 24 hours post-transfection to 5 ug/mL. Cells were harvested at 72 hours as before (see Chapter 2, section 2.2.5). Experion Pro260 chip analyses, in conjunction with CO-difference spectra, were used to confirm P450 presence in batches prior to conducting assays with possible substrates.

3.2.4 Substrate Assays

Substrate assays were performed similarly to Sandstrom *et al.* (2006) except in 300 uL total volume reactions instead of 500 uL, based on recommendations that the lesser volume was sufficient for assays with icCYP9T1 (Michael Smith Laboratories, University of British Columbia). A concentration of 0.02 mg/mL of a selected test substrate was added to microsomal isolates. Tested substrates included myrcene, 3-carene, β-pinene, (-)-α-pinene, (+)-α-pinene, limonene, linalool, bornyl acetate, terpinolene, and L-phenylalanine. Negative controls were conducted with microsomes isolated from MPB-CPR transfections (MOI 0.25). Reactions were conducted in 2 mL silanized amber gas chromatography (GC) vials kept on ice until they were initiated with NADPH (1 mM final concentration). Vials were incubated at 30°C in a rotary shaker (60 rpm) for a maximum 30 or 60 minutes.

Assays were extracted three times in 300 µL hexane using glass pipettes and tubes. Briefly, after the addition of 10 µg of n-octanol standard, ice-cold hexane (300 µL) was added to the reaction and the vial was gently shaken. The entire volume was then transferred to a glass tube and centrifuged at 3,000xg for 15 minutes at 4°C to separate the organic and aqueous phases. The solvent layer was collected and held on ice, while the remaining portion was extracted twice more as above. The pooled extracts were gently concentrated under clean nitrogen gas to approximately 100 µL. Extractions were also conducted in pentane for solvent comparison. A silanized glass insert held the sample within a 2 mL amber GC vial, which was then stored at -20°C pending analysis.

A Varian CP-3800 Gas Chromatograph equipped with CP-8400 auto-sampler and Saturn 2200 GC/MS was used for the analysis. A 30 m x 0.25 mm ID with 0.25 µm film

thickness HP-5MS capillary column (Agilent, Santa Clara, CA) was used for the separation of the extracted samples. Typically, 1 μL of the sample was injected into the GC system using a Varian CP-8400 auto-sampler. Splitless injection mode was performed on the 1079 PTV injector and after 0.7 min, the split mode was activated at split ratio 10:1. The injector temperature was kept at 280°C during the analysis. The capillary column temperature was initially held at 40°C for 2 min, then increased at 5.0°C min^{-1} to 80°C and further increased at 20.0°C min^{-1} to 280°C and held at 300°C for 15 min. The total run time was 35 min for each sample. The carrier gas (helium) was maintained at a constant flow rate of 1.0 mL min^{-1} for the whole analysis and no pressure pulse was used for the injection.

3.3 Results

3.3.1 NCBI Protein BLAST Searches

Many matches from the searches were from predicted or unnamed P450s, but often similarity to a known P450 was indicated. Protein sequence search (pBLAST) results indicated a similarity of 56% between CYP4AY1 and CYP4AY2 (Table 3.3). A P450 similar to CYP4C3 was one of the matches for both CYP4AY1 and CYP4AY2. CYP4BD1 and CYP4BG1 also had matches to P450s that were considered similar to CYP4C3 (comments for CYP4Q7 indicated that it was also considered similar to CYP4C3). CYP4BN1 had 44% identity with CYP4AY2 and CYP4BD1 (45% identity), but CYP4BG1 seemed to share unique identity to various CYP4Q subfamily enzymes. All four had matches with pheromone degrading enzymes (PDE) or similar enzymes, and beyond CYP4BG1's identity with various CYP4Qs, there seemed to be little difference between the overall search results of these four *I. paraconfusus* P450s. CYP4G27 had identities with other CYP4G subfamily P450s and did not share identity to the others at its higher level matches. Not only did the matches to CYP4G27 have higher identities (beginning at 76%), but more types of insects were involved, including bee and mosquito representatives. As functions are not indicated by family designation, it is unlikely that identities at this level indicate definitive functional similarities.

These results were similar to analysis by Huber *et al.* (2007), except I found that CYP4AY1 was similar to different CYP4Cs, and CYP4AY2 had high identity with CYP4AW1 and CYP4AW2. Huber *et al.* (2007) found that CYP4AY2 and CYP4BD1 shared identity with CYP4C1, and the CYP4AB2 had identity to CYP4BG1, instead of the current

match to CYP4AB5. A match to a jewel wasp (*Nasonia vitripennis* Walker) P450, CYP4AB5, was detected during this analysis (which was noted as similar to *N. vitripennis* CYP4AB2). The match Huber *et al.* (2007) found represented CYP4AB2 from *Solenopsis invicta* Buren (red imported fire ant, GenBank accession number AAQ90477). However, overall the families and subfamilies found were generally alike; the changes in representatives likely stemmed from new matches that share slightly more identity.

Table 3.3 - NCBI pBLAST results. Genes with an asterisk (*) are probable or similar to a named P450. For simplicity, superfluous *Tribolium* hits are omitted, as are most at or below 40% identity, with the exception of CYP4G27, where the top 10 matches are noted.

Gene	Match	Organism	Accession #	Identity (%)	Comments
CYP4AY1	CYP4AY2	*Ips paraconfusus*	ABF06545	56	
	CYP4C3*	*Tribolium castaneum*	XP966858	43	similar to PDE
	CYP4AW1	*Phyllopertha diversa*	AAT38512	41	PDE
	CYP4D14*	*Tribolium castaneum*	XP972624	42	similar to probable P450
CYP4AY2	CYP4AY1	*Ips paraconfusus*	ABF06544	56	
	CYP4BN1	*Tribolium castaneum*	NP001123993	44	predicted protein
	CYP4C3*	*Tribolium castaneum*	XP966858	42	similar to PDE
	CYP4BH1	*Ips paraconfusus*	ABF06551	40	
CYP4BD1	CYP4C3*	*Tribolium castaneum*	XP971963	54	predicted similar PDE
	CYP4D14*	*Tribolium castaneum*	XP972624	50	similar to known P450
	CYP4BN1	*Tribolium castaneum*	NP001123993	45	predicted protein
		Tribolium castaneum	XP971612	44	similar to PDE
		Culex quinquefasciatus	XP001851423	43	
	CYP4BN11	*Tribolium castaneum*	NP001123994	42	
		Aedes aegypti	XP001652929	40	
CYP4BG1	CYP4Q4	*Tribolium castaneum*	NP001034529	46	predicted protein
	CYP4Q7	*Tribolium castaneum*	XP001809180	45	similar to CYP4C3
	CYP4AB5	*Nasonia vitripennis*	XP001606257	40	similar to CYP4AB2
	CYP4M5	*Bombyx mori*	NP001103833	40	
CYP4G27	CYP4G14	*Tribolium castaneum*	NP001107791	76	
	CYP4G15*	*Anopheles gambiae*	XP558699	68	similar to known P450
	CYP4G15	*Culex quinquefasciatus*	XP001851084	67	
		Aedes aegypti	XP001648376	67	
		Nasonia vitripennis	XP001606417	67	
	CYP4G11	*Apis mellifera*	NP001035323	66	
	CYP4G25	*Bombyx mori*	NP001106223	65	
		Drosophila erecta	XP001977118	63	
	CYP4G15	*Drosophila sechellia*	XP002044080	62	
		Acyrthpsiphon pisom	XP001944205	63	

3.3.2 MPB EST Database Contig Matches

The MPB representative P450s that best matched to *I. paraconfusus* CYP4AY1, CYP4AY2, CYP4BD1, and CYP4G27 share 55%, 51%, 56%, and 81% identity with their *I. paraconfusus* counterparts, respectively. These identities were all higher than the matches obtained using the NCBI database (Table 3.3). The best matches with the lowest E-values were placed first in each grouping (and cover more of the ORF than the next contigs, Table 3.4). CYP4AY1 lacked a full ORF contig. The CYP4AY2 match in the MPB database favored the midgut and fat bodies (MGFB) of adults treated with JH or feeding. The same was generally true for CYP4BD1 and reads from the adult library could be explained by this as whole insects that may have been feeding were used. CYP4BG1 had no representatives in the database. CYP4G27, which matched to a contig with an E-value of zero, had representatives from all of the MPB cDNA libraries shown. Antennae and un-normalized MGFB libraries did not contain contigs that showed substantial identity to any of the *I. paraconfusus* CYP4s that were analyzed.

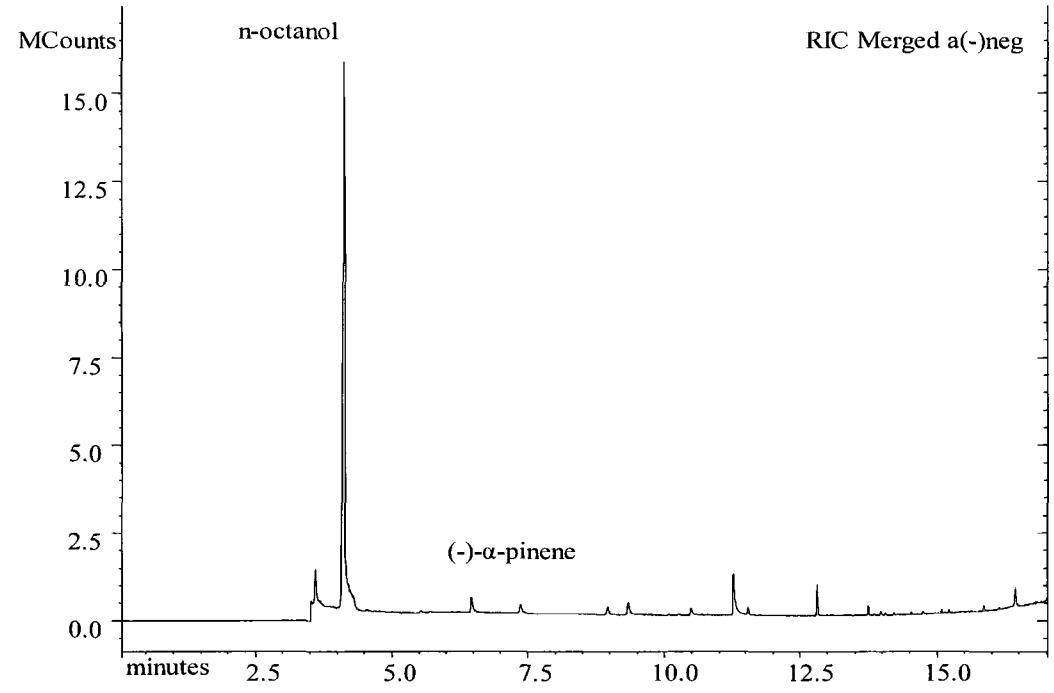
Table 3.4 - Mountain pine beetle EST library search. All contigs that resulted from a keyword search of the libraries are shown with the number of reads each contig is composed of. Notation is as follows: N, normalized; MGFB, midgut and fat body; cold, cold-hardened; JH, juvenile hormone treatment; Terp, exposed to monoterpenes; fed, fed on host phloem.

			Source Library								
Gene	Reads/ Contig	E-value	larvae	larvae (cold)	pupae	adult	adult (head,N)	adult (JH)	adult (Terp)	MGFB (JH,N)	MGFB (fed,N)
CYP4AY1	3	1E-77					2				1
CYP4AY2	17	1E-139								9	8
CYP4AY2	3	1E-96		2							1
CYP4AY2	1	2E-70								1	
CYP4BD1	13	1E-151	1							3	9
CYP4BD1	4	2E-84				3				1	
CYP4BD1	3	2E-76	1							1	1
CYP4BD1	1	4E-72								1	
CYP4BD1	1	1E-62				1					
CYP4BD1	1	8E-32							1		
CYP4G27	54	0	14	10	2	14	2	4	4	2	2
CYP4G27	2	1E-151			2						
CYP4G27	6	1E-104	3						3		
CYP4G27	1	6E-70	1								

3.3.3 Substrate Assay GC-MS

The following substrates were tested: (+)-α-pinene, (-)-α-pinene, β-pinene, myrcene, bornyl acetate, limonene, linalool, terpinolene, 3-carene, and L-phenylalanine. When no possible products were detected compared to negative controls, incubation time was increased to 60 minutes to no effect. Reactions were re-tested using pentane as the extraction solvent, also to no effect. All co-occurring peaks (those besides n-octanol and the substrate of that run) were always present. Oxidized monoterpenoids should have eluted near their substrate peaks. Not one was detected in an assay. Representative GC results for a substrate assay and its associated negative control, and the MS result for the substrate peak is provided (Figures 3.1a,b, Figure 3.2). Figure 3.1a is the negative control for (-)-α-pinene. The substrate peak elutes at approximately 7.5 minutes. The tallest peak is that of the n-octanol standard. Hexane eluted prior to the detector being activated (before 3 minutes). Figure 3.1b is the substrate assay with (-)-α-pinene and CYP4AY1/CPR-expressing microsome. It does not differ from the negative control. The mass spectrometry output for the substrate peak is shown in Figure 3.2, the top section being the MS pattern for the sample and the bottom that of the database match. It can be seen that they are identical, confirming the identity of (-)-α-pinene in the extraction. The presence of each substrate was confirmed, except for that of L-phenylalanine. Only an oxidized product of L-phenylalanine would have been extracted in solvent, and neither was detected.

(a)



(b)

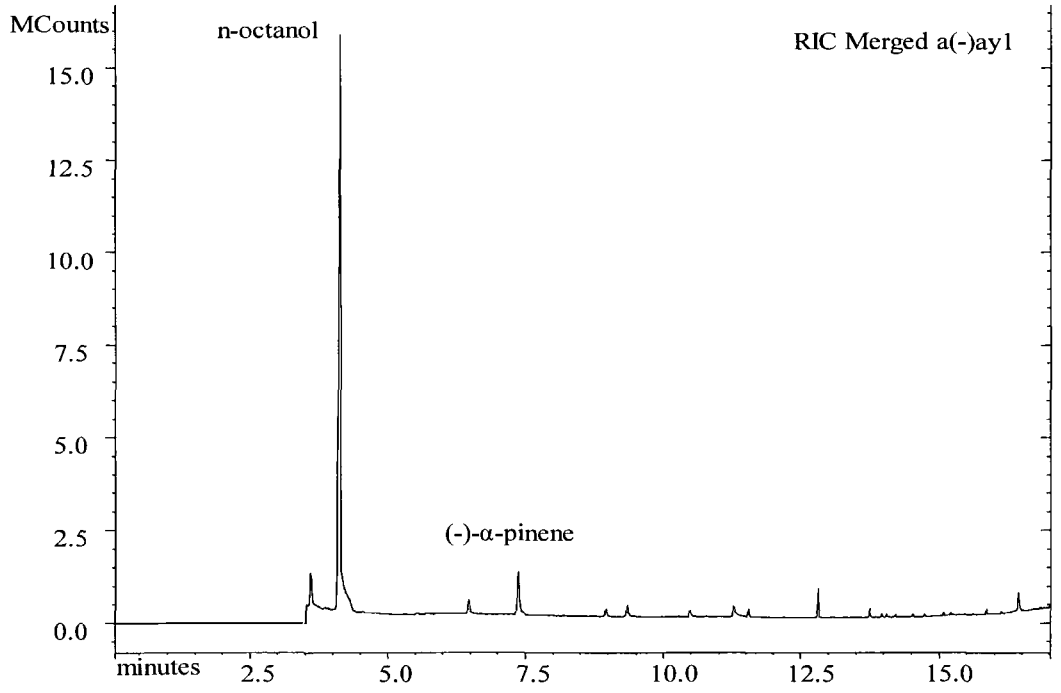


Figure 3.1 – GC Analysis for CYP4AY1/CPR assay with (-)-α-pinene. (a) Negative control assay of (-)-α-pinene with CPR-expressing microsome. (b) CYP4AY1/CPR- expressing microsome assay with (-)-α-pinene.

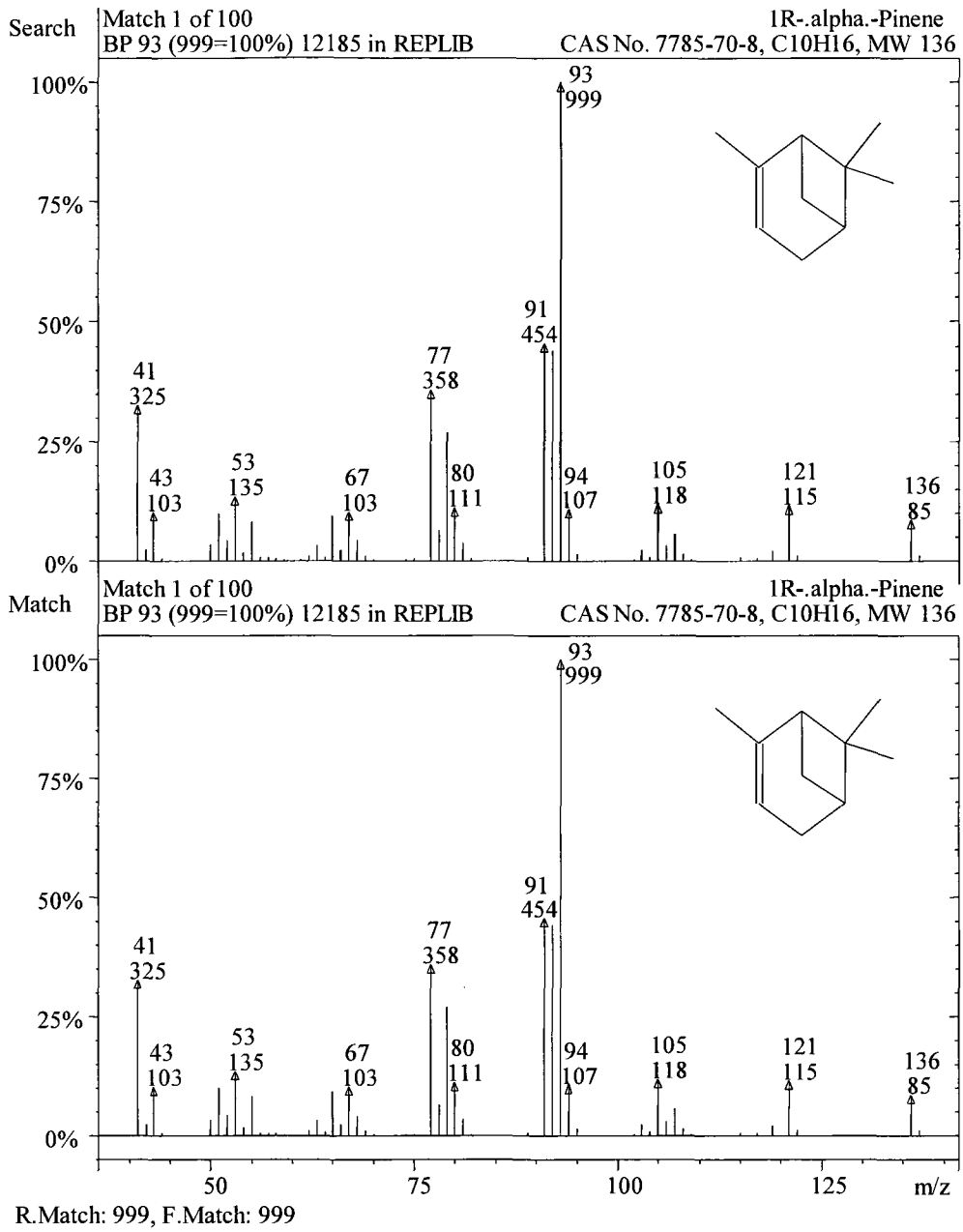


Figure 3.2 – MS Analysis of suspected (-)-α-pinene substrate peak. The pattern it the top box is that of the suspected substrate peak, the pattern below is for the database entry for (-)-α-pinene.

3.4 Discussion

3.4.1 Overview of *Ips paraconfusus* CYP4 Database Matches

Based upon the diverse matches generated by pBLAST analysis, there exists a range of possible functions for the *I. paraconfusus* CYP4 enzymes in this study. However, the values of the identities make this information highly speculative. For example, over 40% amino acid identity places two P450s in the same family, and over 55% places them in the same subfamily (Werk-Reichhart & Feyereisen, 2000). Functional similarities are certainly not assumed at such a level, and the highest identities from the pBLAST matches for CYP4AY1, CYP4AY2, CYP4BD1 and CYP4BG1 are 56% at the very most. Thus it is very unlikely that matches actually catalyze similar reactions, but the types of matches may provide insight into the general possibilities that should be considered. CYP4G27 and its highest match, CYP4G14, have an identity of 76%, which may make that match more functionally informative, but they are still very different peptides. Also, much of the information pertaining to the matching P450s is based upon hypotheses related to analyses of expression, however, although such studies provide context for gene induction, actual functional determination relies upon investigation of the catalytic activity of the protein products themselves.

Gene expression data was for whole body extracts (Huber *et al.*, 2007), so the specific P450-expressing tissues were not defined. In order to attempt to infer tissue localization, comparisons with the MPB EST database (Keeling *et al.*, in prep.) were made (Table 3.4). Keyword searches provided matches for CYP4AY1, CYP4AY2, CYP4BD1, and CYP4G27, with 55%, 51%, 56%, and 81% identity to the MPB representative, respectively. However, as

noted above, this level of identity cannot predict function. Nonetheless, the searches are of particular interest because the two bark beetle species have similar physiological challenges as they are both exposed to monoterpenes in host resin during colonization. The different tissue and treatment libraries used to construct the contigs may provide surrogate localization data for comparison, as speculative as it may be. Taken cautiously, the predictions made with these searches may augment the gene expression data. The compiled search and expression data is discussed for each CYP4 in an attempt to hypothesize possible functions.

3.4.2 Possible Functions of CYP4AY1

CYP4AY1 had protein sequence identity to the known pheromone-degrading enzyme (PDE) CYP4AW1, and to a predicted P450 noted as similar to a PDE and CYP4C3 (Table 3.3). CYP4C3 (*Drosophila melanogaster*) was expressed in the hindgut of feeding larvae and RNA interference of its expression caused death in larvae and pupae (Chung *et al.*, 2009). CYP4AW1 (*P. diversa*, pale-brown chafer beetle) was proposed as a PDE because expression was limited to the antennae of males, which cue into pheromones to locate a female (Maïbèche-Coisne *et al.*, 2004). P450 inhibitors were found to interfere with the insect's ability to detect pheromone, so CYP4AW1 enzyme was thus characterized by the process of elimination rather than by functional characterization. The similarity to a known PDE could implicate CYP4AY1 in the turnover of pheromones at the antennae, but then one would expect *I. paraconfusus* female expression levels to be similar to the expression in males because the female also responds to pheromone components. CYP4AY1 transcript levels are higher in feeding *I. paraconfusus* males than in females (Huber *et al.*, 2007). It seems possible that CYP4AY1 could instead be related to monoterpene detoxification for aggregation pheromone production. As some pheromone components in this species are

derived from monoterpenes, it is possible that the peptide similarity to PDEs stems from the active site binding of similarly shaped molecules. A common evolutionary ancestor in terms of gene family is also possible, especially given only 41% identity (Table 3.3). CYP4AY1 also shared identity with a *Tribolium castaneum* P450 similar to CYP4D14, which in third instar larvae of *Drosophila melanogaster* exhibited increased expression with pheonbarbital exposure (Sun *et al.*, 2006). Pheonbarbital exposure induces some of the same genes as pesticide exposure does in resistant strains, so is applied to find genes related to pesticide resistance. Pesticide resistance is the realm of much insect P450 research (Feyereisen, 1999).

MPB database results for CYP4AY1 matched to one contig with 55% identity, but the contig did not contain a full ORF. The contig consisted of reads from the normalized adult head library and one from the MGFB fed library. Although interesting, the low number of reads in the contig and short overall sequence length preclude any reasonable predictions based on analysis of contig source libraries.

3.4.3 Possible Functions of CYP4AY2

CYP4AY2 shared 42% amino acid identity with the P450 similar to CYP4C3, and exhibited an increase in expression with feeding in both sexes, which could be indicative of its role as a detoxification enzyme (Huber *et al.*, 2007). A P450 that showed 44% identity to CYP4AY2 was CYP4BN1, for which there was little literature available. A member of the subfamily, CYP4BN13v1, however, was either up- or down-regulated in conjunction with moving *Leptinotarsa decemlineata* (Say) to different host material (Zhang *et al.*, 2008). Although this particular enzyme was not one of the matches, it was thought to be related to xenobiotic detoxification activity necessary for feeding.

CYP4AY2 had a 51% match in the MPB database to a high quality contig with 17 reads. All reads in the contig were either from the MGFB of JH treated adults, or the MGFB of fed adults. The next best contig had three reads, two of which were derived from the cold-hardened larvae library, but the only suitable reading frame was interrupted by a stop codon after the probable N-terminal signal. This could represent a sequencing error, but may also be a pseudogene that has lost function, but is located next to an expressing gene, which is why there is an EST. If contig represents a functional P450, it is an unusual one, and not very similar to CYP4AY2. The best contig thus indicated a role in detoxifying ingested monoterpenes. Since both sexes express CYP4AY1, cis-verbenol production from α-pinene is also a possible catalytic function for this gene. This hypothesis agrees with the MPB match's presence in the MGFB of fed and JH treated adults owing to those treatments causing pheromone biosynthesis and the midgut being the location of that activity.

3.4.4 Possible Functions of CYP4BD1

CYP4BD1 pBLAST searches had similar hits to those for the CYP4AY subfamily P450s, with matches with CYP4C3, and CYP4BN1, and CYP4D14. Owing to the suspected functions of the matching genes, described previously, this indicated possible roles in detoxification of host material and possibly pheromone production. The CYP4BD1-matching (56%) contig had 13 reads, which were also derived from the two MGFB libraries, with the exception of one larval library read. If CYP4BD1 is also expressed primarily in the MGFB region, detoxification and pheromone production as functions are further supported. Aggregation pheromone production in particular may be catalyzed by CYP4BD1 owing to its increased expression specific to male *I. paraconfusus* (Huber *et al.*, 2007).

3.4.5 Possible Functions of CYP4BG1

Protein BLAST search results for CYP4BG1 were more directed towards a particular subfamily than were the analyses of the three protein sequences already discussed. The top two matches with CYP4BG1 were with CYP4Q subfamily P450s, one of which, CYP4Q7 (46% identity), was noted as being similar to the aforementioned CYP4C3. Although scant literature is available pertaining to the functions of CYP4Q subfamily P450s, CYP4Q11 (which was not a listed match) is expressed after the movement of *L. decemlineata* (Say) to a different host (Zhang *et al.*, 2008). In addition, *T. castaneum* pyrethoid insecticide resistance seemed to be related to the expression of CYP4Q4 (Ray, 1998). A 40% match to CYP4BG1, CYP4AB5, was noted as being similar to CYP4AB2 of the same species. CYP4AB2 was found to have increased expression specific to the workers (not the queen) of the red imported fire ant, but little information was available about this P450 subfamily in social insects (Liu & Zhang, 2004). CYP4M5 also shared only 40% identity to CYP4BG1. In *Bombyx mori* Linnaeus (silk moth), CYP4M5 exhibited high expression in the midgut and fat body, in addition to its presence in the ovaries and testis (Yamamoto *et al.*, 2010). Expression of CYP4M5 in the fat bodies of larvae increased with diazinon exposure, which the authors interpreted as evidence for a role in insecticide resistance. So the pBLAST results seemed to favor CYP4BG1's association with detoxification pathways. Owing to how P450s evolve through gene duplication events (Feyereisen, 1999) and the presence of conserved regions, similarities should be expected. Cytochromes P450 are numerous, especially in insects, and they are highly variable in terms of function. However, it is of interest how CYP4BG1 seems to deviate from CYP4AY1, CYP4AY2, and CYP4BD1, in terms of matches with other

P450s. Where the other three all had representatives from the same various subfamilies, CYP4BG1's matches to CYP4Q subfamily P450s made it unique.

A keyword search of the MPB database did not elicit a result for CYP4BG1. Perhaps CYP4BG1 provides a function that separates the more aggressive *Dendroctonus* from *Ips* species. More aggressive tree-killing bark beetles species are thought to be less competitive compared to other bark beetle species that utilize compromised host material (wind-thrown or otherwise weakened trees), which forced the move to more vigorous hosts (Borden, 1982). CYP4BG1 could be involved in the pathways that give less aggressive species a competitive edge, perhaps by metabolizing lower quality host material more efficiently.

3.4.6 Possible Functions of CYP4G27

Perhaps the most interesting thing about the CYP4G27 database search was not just that matches deviated completely from the searches involving the other CYP4 family P450s, but that such different species had higher identity matches. Identities in general were much higher than the matches generated by the other CYP4s, which also tended to be only with other beetles (*Tribolium castaneum* Herbst; *Phyllopertha diversa* Waterhouse). Mosquito (CYP4G15, *Anopheles gambiae* Giles), wasp (*Nasonia vitripennis*), honey bee (CYP4G11, *Apis mellifera ligustica* Spinola), and fruit fly (*Drosophila erecta* Tsacas & Lachaise) proteins had identities with CYP4G27 of 65% or greater. This was indicative of a more conserved gene subfamily. This does not point as strongly towards an enzyme that synthesizes specialized pheromones, as much more variability in pheromone biosynthetic P450s between such different species with vastly different life histories would be expected. Also, CYP4G27's stable expression in adult *I. paraconfusus*, despite feeding stimulus,

indicated a potential role in some sort of metabolic housekeeping function (Huber *et al.*, 2007). *D. melanogaster* larvae, pupae and adults were found to express CYP4G15 in the nervous system (brain and/or thoracic ganglion, respectively); the authors suggested possible roles in metabolizing endogenous compounds like ecdysteroids (Maïbèche-Coisne *et al.*, 2000). Oak silkmoth larvae (*Antheraea yamamai* Guerin-Meneville) were found to express CYP4G25 when initiating diapause, perhaps in response to environmental factors such as cold temperature (Yang *et al.*, 2008). It would be interesting to see if CYP4G27 expression responded differently to physiological states or developmental factors such as diapause or metamorphosis. Similarity obviously does not imply function, however, a great deal of expression data is now emerging and that information will help to elucidate P450 roles. Taken together, the BLAST results reinforce the diversity of the P450s and to the importance but difficulty of functionally characterizing more of its members.

The best MPB contig match to CYP4G27 (81%) had a markedly different library profile consisting of 54 reads and an error value of zero. The majority of the contigs matching CYP4G27 reads were derived from larval (normal and cold) and untreated adult tissues, but other tissues and treatments were also represented. Of note, like the other CYP4 searches, there were no hits to the antennal library, making pheromone degradation an unlikely function. It is possible that this gene has some role in ecdysteroid, JH, or fatty acid pathways. In order for any signaling system to function, whether it involves pheromones or hormones, the presence and absence of the signaling molecules must have context. That is to say, they must be both created and broken down to have meaning; a signal that is not broken down quickly enough loses meaning. P450s are involved in the construction of insect hormones, but their roles in breaking them down are just as important. Given its constitutive

expression profile, until further work pinpoints more particular information related to this gene's expression, it will remain difficult to narrow down a list of possible substrates.

3.4.7 Substrate Candidate Selection

Several possible functions for various *I. paraconfusus* P450s were mentioned by Huber *et al.* (2007) based upon the expression profiles that were elicited with feeding. CYP4AY1 and CYP4BD1 were hypothesized to have possible roles in pheromone production or related JH biosynthesis. Owing to their expression in both sexes, CYP4AY2 and CYP4BG1 could have roles in the synthesis or catabolism of reproductive hormones, or in the detoxification of host compounds. Males and females also produce cis-verbenol, potentially via oxidation of host-derived α-pinene as an aggregation pheromone component (Seybold *et al.*, 2006). Therefore, α-pinene was considered to be a possible monoterpene substrate for both CYP4AY2 and CYP4BG1. These predictions are in good agreement with the MPB EST database findings for CYP4AY2, CYP4BD1, and to some extent CYP4AY1, with midgut and fat body being the predominant locations. This is where detoxification and pheromone biosynthesis occurs, so monoterpenes present in host tissue were considered the best place start testing substrates (Table 3.4).

Detoxification of host secondary metabolites may also be a function of the constitutively expressed CYP4G27 and degradation of flight muscles following host colonization was also mentioned (Huber *et al.*, 2007). Untreated adult and larval tissue were the main contributors to the MPB EST contig that best matched CYP4G27. If it were specific to flight muscle degradation processes, the presence of transcripts in larvae would be unexpected. However, JH is involved in controlling processes ranging from pheromone

biosynthesis to flight muscle degradation (Tillman *et al.*, 1998, 2004; Unnithan & Nair, 1977). Perhaps CYP4G27 represents a constitutively produced enzyme that is involved in signal turnover so that systemic responses can be fine-tuned to immediate temporal changes in JH production.

3.4.8 Substrate Assays

Substrate assays with CYP4AY1/CPR and CYP4AY2/CPR microsomes and monoterpenes and L-phenylalanine did not yield products. Thirty minutes is the generally accepted incubation time for such assays (Wen *et al.*, 2003; Sandstrom *et al.* 2006, 2008), but one hour was also tested in case of inefficient electron transfer or a slow substrate turnover rate. Results did not differ with increasing incubation time. Hexane was the primary solvent used for extraction; due to its slower evaporation rate it was easier to handle. Following the extraction solvent in the GC column, octanol standard was still detected, so hexane moved through the column at an appropriate rate (i.e. slow enough to be within the detection time allowed). Nonetheless, pentane was also tested as a solvent, but potential products were still not detected. Pentane extraction resulted in fewer co-occurring GC peaks, representing other hydrocarbons in the sample, but as these occurred out of the range of the substrates and expected products, hexane was still preferred as pentane was difficult to pipette. Through comparisons between runs, I was able to identify the location of each of the possible substrate monoterpenes on the GC output. The aforementioned co-occurring peaks were consistently present in each run. Assay and substrate conditions were consistent with those in the IcCYP9T1 assays demonstrating the conversion of myrcene to ipsdienol. It seems that CYP4AY1/CPR and CYP4AY1/CPR expressing microsomes simply did not act on the chosen substrates. Considering how even undetectable levels of recombinant P450 (by CO-

difference spectra) still turnover substrate (D.-K. Ro, pers. comm.), this is the logical conclusion.

It is possible that activity was not observed because the right substrates were not selected, and other substrate possibilities certainly remain. The monoterpenes that were chosen were among those that the insect is most likely to encounter in host resin, and some (α-pinene, myrcene) are structurally similar to *I. paraconfusus* pheromone components. However, many other endogenous pathways and related physiological factors are active during host selection. For instance, JH exposure causes pheromone production, but also flight muscle degradation (Unnithan & Nair, 1977), which may be an important energy source for reproductive activities. Ecdysteroids signal the release of vitellogenins from the fat body for egg production (reviewed by Bellés *et al.*, 2005). Fatty acid metabolism may be taking place. Any and all of the range of substrates P450s are known to act on could be involved in the physiological transitions between finding a host and reproducing.

It is possible that a suitable substrate was selected, but CPR was unable to transfer reducing equivalents of electrons to CYP4AY1 and CYP4AY2. Cytochrome *b5* (CYTb5) is sometimes coexpressed in recombinant P450/reductase systems. This is thought to improve the transfer of electrons from CPR to the P450, but in some cases its presence is necessary for functionality (Wen *et al.*, 2003). CYTb5's role is context dependent (by P450 and substrate), sometimes involving its own reductase (cytochrome *b5* reductase), and not always transferring an electron (Feyereisen, 1999). Hannemann *et al.* (2007) reviewed the three ways that CYTb5, which can act in both microsomal and mitochondrial P450 systems, was involved in P450 function. CYTb5 can transfer both electrons necessary for complete P450 reduction from NADH-cytochrome *b5* reductase. Alternately, it can transfer only the second

electron from either its reductase or CPR. Finally, it can allosterically assist a P450 without actually transferring any electrons (Hannemann *et al.*, 2007). In the last case, CYTb5 would induce a conformational change that would allow the P450 to be catalytically active. It is possible that the substrates and enzymes were present, but this additional enzymatic component was missing and the recombinant P450s were thus not reduced or otherwise capable of functioning.

3.5 Conclusions

The functions of CYP4AY1 and CYP4AY2 remain unknown, despite efforts to identify possible substrates. Pheromone biosynthesis and detoxification are not the only processes vital to the survival of a host tree-colonizing bark beetle. It would not be surprising if these enzymes were instead part of JH or ecdysteroid biosynthetic pathways. Fatty acid breakdown is another possibility that would link these enzymes to energy processes that could certainly be active when adults get to a host tree. Despite the MPB EST library results, we cannot be certain that CYP4AY1, CYP4AY2 or CYP4BD1 are really expressed in the midgut or fat body without investigating that specifically in *I. paraconfusus*.

It is also possible that the right substrate was tested, but the assay system conditions were not conducive to functionality. Although the P450 and CPR were present, those are not always the only important components of the system. Cytochrome *b5* may be required to facilitate electron transfer so that the recombinant P450 enzymes can be reduced and bind molecular oxygen (Wen *et al.*, 2003). Further work with these recombinant cytochromes P450 still needs to done if functionality is to be determined. P450s are not only a diverse family, but individual enzymes can have several possible substrates and products. Unfortunately, the diversity that makes them so interesting also makes it difficult to elucidate function. Another aspect of P450 diversity to consider is subcellular localization. Some P450s, which are described as chimeric, have the ability to localize to multiple cellular compartments, including the mitochondria. Although CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and CYP4G27 all localize to the ER, as functional recombinant P450 was detected in microsomal fractions, there is the possibility that they have mitochondrial counterparts. To further characterize these proteins, this aspect must also be explored.

4.0 N-terminal Processing and Localization of *Ips paraconfusus* CYP4 P450s

4.1 Introduction

In previous work, Huber *et al.* (2007) found several novel *I. paraconfusus* cytochrome P450 genes. Heterologous expression of CYP4AY1, CYP4AY2, CYP4BG1, and CYP4BD1 proteins revealed molecular weights of approximately 48 kDa for all recombinant proteins (e.g. see Chapter 2, Figures 2.1a, 2.2a), despite the expected sizes, based on amino acid sequences, of 56.1 kDa, 56.3 kDa, 57.7 kDa, and 58.2 kDa, respectively. The molecular weights were 14.4%, 14.7%, 16.8% and 17.5% lower than expected by *in silico* molecular weight prediction. These differences were greater than the 10% error inherent in both sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and capillary electrophoresis via the Experion automated setup (Invitrogen; Bio-Rad). However, increased migration rates for P450 proteins has been observed, possibly as a result of the hydrophobic regions binding more SDS, thus becoming more electronegative per mass unit (Black and Coon, 1986 in Brown *et al.*, 1998). For example, a CYP30 protein discovered in the clam *Mercenaria mercenaria* (Linnaeus) had a molecular weight 14.5% lower than predicted (Brown *et al.*, 1998). This faster migration hypothesis seemed reasonable, especially since carbon monoxide difference spectra confirmed that the four P450 proteins were properly incorporating the haeme prosthetic group (see Chapter 2, sections 2.3.1, 2.3.2, 2.3.4). The proteins were suspected to be both functionally viable and correctly processed, but the faster migration hypothesis did not satisfy all of the findings. CYP4G27, which appeared to have a longer N-terminal hydrophobic region than the other CYP4s investigated, had a predicted molecular weight of 63.9 kDa within 3.0% of the Experion-derived value of ~62 kDa (not shown). Either something about the amino acid compositions or SDS-binding lead to

increased migration rates, or some other mechanism was acting on CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 to cause the recombinant proteins to migrate faster than predicted.

P450 processing involves N-terminal signaling, and previous investigation of the *I. paraconfusus* CYP4 amino acid sequences suggested localization to the ER, as would be the case with the normal microsomal incorporation expected of most P450s (Huber *et al.*, 2007). The predominant model for the insertion of microsomal P450s involves a hydrophobic, un-cleaved N-terminal peptide sequence that is co-translationally recognized by the signal recognition particle (SRP), which transfers the protein to the ER where the protein anchors (Bar-Nun *et al.*, 1980). The processing of P450 N-terminal peptides are related to their length and charges; additions of positively charged residues to N-termini lead to translocation via conversion into cleavable signal peptides (Szczesna-Skorupa *et al.*, 1988; Sato *et al.*, 1990). Interestingly, a study of chimeric proteins showed P450s to have intrinsic abilities to associate with the ER without the N-terminal anchor-containing signal (Szczesna-Skorupa *et al.*, 1995); thus the targeting system is still not completely understood. This is also the case with mitochondrial P450s, but in general, the N-termini will have several positive residues to facilitate interaction with mitochondrial processing elements; this signal is only active when at the N-terminus (Omura, 2006). Addya *et al.* (1997) have proposed a model by which P450 peptides could be cleaved at the ER and then further modified or escorted by cytosolic factors, before reaching the mitochondrial processing elements. This involved a chimeric N-terminus that had both ER and mitochondrial signaling regions, resulting in different isozymes from the same initial peptide chain being associated with different organelles (Bhagwat *et al.*, 1999; Robin *et al.*, 2001; Neve & Ingelman-Sundberg, 2000). Isozymes at

alternate organelles preferred electron donors specific to those locales (Robin *et al.*, 2001; Bhagwat *et al.*, 1999), further emphasizing the versatility of the cytochrome P450 superfamily, even at the individual peptide level.

The presence of several conspicuous clusters of positive amino acids interspersed near the N-termini of some of the CYP4s indicated that the size discrepancies could be the result of processing events different than those of normal microsomal targeting. An exploratory study of the proteins was undertaken. The possibility of alternate subcellular targeting was first investigated based on proteolytic cleavage sites predicted *in silico*, and then by isolating different cellular fractions of SF9 cells transfected with recombinant baculovirus. *Ips paraconfusus* CYP4 P450s were investigated using programs designed to detect proteolytic cleavage and targeting. However, because such programs are not necessarily proficient at detecting internal signals and may not resolve signals for the inner mitochondrial compartment or cell membrane (Small *et al.*, 2004), multiple programs were used to analyze P450s with the systematic *in silico* removal of sections that might occlude such secondary signals. Histidine-tagged recombinant P450s were expressed so that the presence of lower levels of alternate isoforms could be validated by western blots. Several cellular fractions were isolated for Experion and western blot analyses. The objective of this work was to detect an overall pattern that might identify processes related to the observed size discrepancies, assuming that those deviations from expected molecular weights were real and not artifacts of differential SDS binding to peptide chains. If processing by proteases was occurring, different sizes of recombinant protein should have been present. This information was used to hypothesize possible subcellular trafficking mechanisms acting on these particular *I. paraconfusus* CYP4 P450s.

4.2 Materials and Methods

4.2.1 *In silico* Analyses of N-termini

All P450s characterized by Huber *et al.* (2007) that had complete N-terminal sequence data available were analyzed. This included CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, CYP4G27, CYP4BE1, CYP4BF1, CYP4BH1 and CYP4BJ1 (GenBank accession numbers DQ471874, DQ471875, DQ471876, DQ471880, DQ471883, DQ471877, DQ471879, DQ471881, and DQ471882). IcCYP9T1 (*Ips confusus*) and CYP9T2 (*Ips pini*), were also examined for cleavage or alternate targeting signals (EU915209 and DQ676820). N-terminal deduced amino acid sequences were entered into the program SignalP 3.0 with eukaryote selected as the source organism type and both neural networks and Hidden Markov Models (HMM) selected (Bendtsen *et al.*, 2004). A truncation limit of 70 (standard setting) was chosen. In the majority of instances, both models agreed in terms of cleavage location. Sometimes the neural networks model did not identify a site, so HMM probabilities for the presence of signals, anchors, and cut sites were recorded (Table 4.1). When a cleavage site was detected, the pre-cleavage sequence was removed and the new N-terminus was re-evaluated. Approximate molecular weights resulting from cleavage events, for those genes for which recombinant viruses were constructed, were estimated with Compute pI/Mw tool on the Expasy Proteomics Server (Gasteiger *et al.*, 2003). TargetP 1.1 (Emanuelsson *et al.*, 2000) was used to evaluate the localization potentials of the signal sequences resulting from detected cleavage sites. The "non-plant" setting and the standard "winner-takes-all" default setting with no cutoffs were selected. The program was not set to find cleavage sites as TargetP simply uses SignalP to perform that function. The program Predotar 1.03 was utilized similarly, with the "animal/fungal" option being selected (Small *et al.*, 2004).

Putative peptide sequences were also analyzed with PSORT II (Nakai & Horton, 1999), which is the version recommended for fungal and animal-derived sequences. The probability of mitochondrial import was assessed with the program MITOPROT II v1.101, which like PSORT II had no setting options (Claros & Vincens, 1990). Any cleavage sites or import signals detected by the above programs were compiled on maps of the N-terminal amino acids to observe any possible patterns.

4.2.2 Localization Analyses of Recombinant *Ips paraconfusus* CYP4s

Histidine-tagged viral constructs were developed and SF9 cells transfected at MOIs of 10 (see Chapter 2, sections 2.2.2-2.2.5). The microsome isolation protocol was modified in order to isolate extracellular media, cytosol, cell debris, mitochondria, and microsomes by differential centrifugation (centrifuge forces based on a protocol by Thomas *et al.* (1999)). Cells were centrifuged at 3,000xg for 15 minutes at 4°C to remove media. The pellet was rinsed three times in phosphate buffered saline (1/3 cell volume) and resuspended in the same volume of lysis buffer (Chapter 2, section 2.2.5). Samples were sonicated as before, on ice for 30 seconds, three times, on the output 3 setting. Cell debris and mitochondria were spun out at 800xg and 10,000xg, respectively, for 10 minutes each at 4°C. Mitochondrial pellets were rinsed twice by resuspending and pelleting samples twice in lysis buffer. Supernatant was spun down at 100,000xg for 1 hour at 4°C to isolate the cytosolic fraction and concentrate microsomes. Microsomal pellets, like the other fractions, were mechanically broken up and suspended in 1/25 cell volume of lysis buffer to ensure they were sufficiently concentrated to increase the likelihood of detecting recombinant protein variants occurring at low levels. The Experion capillary electrophoresis system, with Pro260 chips, was used to evaluate protein content of fractions as per the user manual (Bio-Rad Laboratories).

To see if the uniquely migrating proteins at lower concentrations than the predominant bands were recombinant proteins, the same samples run on the Experion were also analyzed by western blotting. Isolates were subjected to SDS-PAGE on 4-12% Bis/Tris polyacrylamide gels for 1 hour at 200V, under reducing conditions (NuPAGE Novex Bis-Tris Mini Gel, Invitrogen). Proteins were transferred to PVDF membranes and recombinant proteins bound the anti-histidine/alkaline phosphatase-conjugated antibody (IBlot, Western Breeze Chromogenic Kit, Anti-His C-term Antibody; Invitrogen). The chromogenic substrate BCIP/NBT (5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine salt and Nitro-Blue Tetrazolium Chloride) reacted with the alkaline phosphatase, and thus recombinant proteins resolved as dark purple bands. Standard curves were estimated using the Magic Mark protein standard (Invitrogen), and the approximate molecular weights represented by bands of interest were calculated from them.

4.3 Results

4.3.1 *In silico* analyses of N-termini

Of those CYP4s expressed as recombinant proteins, all five had an early cleavage signal that was detected by SignalP. CYP4AY2 and CYP4BD1 had relatively high cleavage probabilities (0.580 and 0.445, respectively), and CYP4G27 had the greatest cleavage probability value at 0.9 (Table 4.1). CYP4BG1 is the only representative with a strong initial anchoring probability (0.934). Those with only sequence data, including two CYP9s and the additional CYP4s, all exhibited cleavage probabilities in a site near their N-termini. Consequently, many of the first sites were predicted to be signal peptides. Sites evaluated by deleting the pre-cleavage peptides decrease in their signal and cleavage scores. The CYP9s had no predictable third site, and the third site of CYP4G27 was estimated to be after the WXXXR functional motif (which is the most N-terminal of the conserved P450 motifs), so a cleavage event there would likely compromise function of the enzyme product. There is no clear pattern for the second or third cleavage sites that would indicate a specific targeting role. The second CYP4BD1 site and third CYP4AY2 site had scores of zero for all categories, and CYP4G27 only had values different from zero at the first site. Despite these probabilities, SignalP still provided a location for the most likely cleavage site, which was kept consistent in subsequent analyses. The first three sites obtained by consecutive *in silico* cleavage were used for the remaining analyses when cleavage sites are noted.

Table 4.1 - SignalP analyses of N-termini. Cleavage sites are numbered sequentially based on proximity to the putative N-terminus. Values are the estimated probabilities that the terminus contains a signal peptide, membrane anchor, or cleavage site.

Gene/site#	prediction	p(signal)	p(anchor)	p(cleave)	*(in silico)* Kda prediction
CYP4AY1/1	nonsecretory	0.030	0.003	0.029	53.769
CYP4AY1/2	nonsecretory	0.038	0.010	0.009	51.313
CYP4AY1/3	nonsecretory	0.001	0.001	0.001	47.136
CYP4AY2/1	signal peptide	0.969	0.031	0.580	54.153
CYP4AY2/2	nonsecretory	0.055	0.047	0.029	51.079
CYP4AY2/3	nonsecretory	0.000	0.000	0.000	46.859
CYP4BD1/1	signal peptide	0.900	0.099	0.445	55.764
CYP4BD1/2	nonsecretory	0.000	0.000	0.000	52.794
CYP4BD1/3	nonsecretory	0.015	0.000	0.006	48.092
CYP4BG1/1	signal anchor	0.580	0.934	0.033	56.375
CYP4BG1/2	nonsecretory	0.014	0.000	0.100	52.204
CYP4BG1/3	nonsecretory	0.000	0.000	0.000	48.065
CYP4G27/1	signal peptide	0.996	0.004	0.900	60.993
CYP4G27/2	nonsecretory	0.000	0.000	0.000	57.402
CYP4G27/3	nonsecretory	0.000	0.000	0.000	52.770
CYP4BE1/1	nonsecretory	0.463	0.000	0.365	
CYP4BE1/2	nonsecretory	0.000	0.000	0.000	
CYP4BE1/3	nonsecretory	0.000	0.000	0.000	
CYP4BF1/1	signal peptide	0.972	0.076	0.541	
CYP4BF1/2	nonsecretory	0.006	0.015	0.005	
CYP4BF1/3	nonsecretory	0.000	0.015	0.000	
CYP4BH1/1	signal peptide	0.998	0.000	0.987	
CYP4BH1/2	nonsecretory	0.021	0.000	0.012	
CYP4BH1/3	nonsecretory	0.000	0.000	0.000	
CYP4BJ1/1	signal peptide	0.951	0.098	0.507	
CYP4BJ1/2	nonsecretory	0.012	0.007	0.006	
CYP4BJ1/2	nonsecretory	0.000	0.000	0.000	
CYP9T1/1	signal peptide	0.996	0.002	0.969	
CYP9T1/2	nonsecretory	0.000	0.000	0.000	
CYP9T2/1	signal peptide	0.979	0.013	0.963	
CYP9T2/1	nonsecretory	0.000	0.000	0.000	

TargetP provided P-values not considered probabilities, but location scores, although they are interpreted like probabilities (Emanuelsson *et al.*, 2000). Reliability class (RC) values represent the difference between the two most likely locations (1= >0.8, 2=0.8-0.6, 3=0.6-0.4, 4=0.4-0.2, 5= <0.2), where 1 is the class most reliable, and 5 the least. Prior to the *in silico* removal of amino acid residues, TargetP localized all of the tested P450 proteins to the ER through predicted signal peptides, with CYP4BE1 having the lowest ER targeting probability (0.764, Table 4.2). Removal of a series of amino acids at the N-terminal ends to emulate the first cleavage site seemed to favor an unknown intracellular target location, although CYP4BF1, CYP4BG1 and both CYP9s had scores that lent some support to mitochondrial targets. However, their reliability classes (RC) were too high to indicate a single destination. CYP4BH1 and CYP4BJ1 seemed to be targeted to the mitochondria after the first cleavage, but targeting following the second and third cleavages were unclear.

Table 4.2 - TargetP analyses of N-termini. Sequential deletions are denoted by the cleavage sites used; amino acids N-terminal to the sites were considered the presequence that was removed by processing.

Gene/Cut	P(MIT)	P(SIG)	P(0THER)	prediction	RC
CYP4AY1/0	0.036	0.928	0.159	signal pep.	2
CYP4AY1/1	0.158	0.188	0.412	na	4
CYP4AY1/2	0.075	0.356	0.759	na	3
CYP4AY1/3	0.004	0.283	0.832	na	3
CYP4AY2/0	0.117	0.935	0.030	signal pep.	1
CYP4AY2/1	0.122	0.098	0.575	na	4
CYP4AY2/2	0.123	0.102	0.858	na	2
CYP4AY2/3	0.087	0.109	0.800	na	2
CYP4BD1/0	0.019	0.972	0.067	signal pep.	1
CYP4BD1/1	0.066	0.119	0.874	na	2
CYP4BD1/2	0.115	0.417	0.392	signal pep.	5
CYP4BD1/3	0.052	0.12	0.935	na	1
CYP4BG1/0	0.163	0.864	0.031	signal pep.	1
CYP4BG1/1	0.281	0.131	0.356	na	5
CYP4BG1/2	0.176	0.163	0.717	na	3
CYP4BG1/3	0.246	0.055	0.607	na	4
CYP4G27/0	0.029	0.842	0.186	signal pep.	2
CYP4G27/1	0.155	0.209	0.514	na	4
CYP4G27/2	0.102	0.118	0.848	na	2
CYP4G27/3	0.081	0.087	0.880	na	2
CYP4BE1/0	0.046	0.764	0.391	signal pep.	4
CYP4BE1/1	0.094	0.076	0.873	na	2
CYP4BE1/2	0.102	0.075	0.845	na	2
CYP4BE1/3	0.289	0.085	0.555	na	4
CYP4BF1/0	0.076	0.956	0.014	signal pep.	1
CYP4BF1/1	0.269	0.206	0.300	na	5
CYP4BF1/2	0.075	0.083	0.913	na	1
CYP4BF1/3	0.038	0.118	0.924	na	1
CYP4BH1/0	0.313	0.819	0.021	signal pep.	3
CYP4BH1/1	0.418	0.079	0.412	mitochondria	5
CYP4BH1/2	0.114	0.128	0.822	na	2
CYP4BH1/3	0.044	0.543	0.787	na	4
CYP4BJ1/0	0.226	0.871	0.007	signal pep.	2
CYP4BJ1/1	0.415	0.177	0.150	mitochondria	4
CYP4BJ1/2	0.090	0.046	0.928	na	1
CYP4BJ1/3	0.081	0.161	0.861	na	2
CYP9T1/0	0.049	0.938	0.037	signal pep.	1
CYP9T1/1	0.252	0.067	0.703	na	3
CYP9T1/2	0.108	0.041	0.909	na	1
CYP9T2/0	0.043	0.935	0.042	signal pep.	1
CYP9T2/1	0.434	0.051	0.543	na	5
CYP9T2/2	0.108	0.040	0.909	na	1

Predotar program analysis also predicted ER localization prior to the first cleavage event at site one, with the exception of CYP4G27, whose location was 'other' (Table 4.3). CYP4G27 exhibited 'possible' mitochondrial targeting after the first cleavage site, as did CYP4AY1and CYP4BD1. CYP4AY2, CYP4BG1, CYP4BF1, and CYP4BH1 possessed mitochondrial-targeting peptides after the first cleavage event. There was no clear pattern when the subsequent two cleavage sites were evaluated, although some sites seem to be recognized as ER or mitochondrial signal peptides (CYP4AY1, CYP4BD1, and CYP4BG1).

Table 4.3 - Predotar analyses of N-termini. Sequentially cleaved peptides were denoted as before. P-values are probabilities associated with mitochondria (MIT), ER, or other (O) target sequences where 0.2 is considered above background when evaluating possible locations. Predictions denoted 'poss' are possible, and 'na' means no prediction.

Gene/cut	p(MIT)	p(ER)	p(O)	prediction
CYP4AY1/0	0.08	0.99	0.01	ER
CYP4AY1/1	0.41	0.25	0.46	poss MIT
CYP4AY1/2	0.00	0.01	0.46	na
CYP4AY1/3	0.01	0.42	0.58	poss ER
CYP4AY2/0	0.02	0.99	0.01	ER
CYP4AY2/1	0.50	0.01	0.49	MIT
CYP4AY2/2	0.01	0.01	0.98	na
CYP4AY2/3	0.00	0.03	0.97	na
CYP4BD1/0	0.02	0.99	0.01	ER
CYP4BD1/1	0.24	0.01	0.75	poss MIT
CYP4BD1/2	0.01	0.40	0.60	poss ER
CYP4BD1/3	0.00	0.01	0.99	na
CYP4BG1/0	0.02	0.99	0.01	ER
CYP4BG1/1	0.50	0.01	0.49	MIT
CYP4BG1/2	0.00	0.01	0.99	na
CYP4BG1/3	0.27	0.01	0.72	poss MIT
CYP4G27/0	0.00	0.13	0.87	na
CYP4G27/1	0.39	0.01	0.60	poss MIT
CYP4G27/2	0.00	0.01	0.99	na
CYP4G27/3	0.00	0.01	0.99	na
CYP4BE1/0	0.02	0.98	0.02	ER
CYP4BE1/1	0.01	0.01	0.99	na
CYP4BE1/2	0.00	0.01	0.99	na
CYP4BE1/3	0.01	0.01	0.98	na
CYP4BF1/0	0.01	0.99	0.01	ER
CYP4BF1/1	0.54	0.01	0.46	MIT
CYP4BF1/2	0.04	0.01	0.95	na
CYP4BF1/3	0.00	0.01	0.99	na
CYP4BH1/0	0.02	0.99	0.01	ER
CYP4BH1/1	0.59	0.04	0.39	MIT
CYP4BH1/2	0.00	0.01	0.99	na
CYP4BH1/3	0.00	0.01	0.99	na
CYP4BJ1/0	0.01	0.99	0.01	ER
CYP4BJ1/1	0.14	0.01	0.86	na
CYP4BJ1/2	0.00	0.01	0.99	na
CYP4BJ1/3	0.00	0.01	0.99	na
CYP9T1/0	0.01	0.99	0.01	ER
CYP9T1/1	0.18	0.01	0.82	na
CYP9T1/2	0.02	0.01	0.97	na
CYP9T2/0	0.00	0.99	0.01	ER
CYP9T2/1	0.26	0.01	0.73	na
CYP9T2/2	0.03	0.01	0.97	na

According to MITOPROT II analysis, CYP4BF1, CYP4BH1, and CYP4BJ1 were particularly likely to be imported into the mitochondria with probabilities of 0.9359, 0.9393, and 0.8068, respectively (Table 4.4). Only CYP4AY1, CYP4BD1 and CYP4BE1 had import probabilities below 0.3. Estimated mitochondrial presequence cleavage sites, which are where proteolytic cleavage removes a portion of the N-terminal peptide chain to reveal a signal for the next import step, were represented (if detected) on Figure 4.1 as an underlined amino acid. Possible mitochondrial import presequence cleavage sites were recognized in CYP4AY1, CYP4BF1, CYP4BJ1, and CYP9T2.

Table 4.4 - MITOPROT II analysis of N-termini to predict mitochondrial import. Probability values were provided for mitochondrial import based on un-cleaved, putative N-terminal peptides. Probabilities are based on (a) the targeting properties of the sequence, and (b) the presence of a cleavage site. The locations of any MITOPROT II-predicted cleavage sites are underlined on Figure 4.1.

Gene	p(MITtimport)
CYP4AY1	0.0641
CYP4AY2	0.3263
CYP4BD1	0.1093
CYP4BG1	0.3306
CYP4G27	0.3870
CYP4BE1	0.0430
CYP4BF1	0.9359
CYP4BH1	0.9393
CYP4BJ1	0.8068
CYP9T1	0.4777
CYP9T2	0.4638

A map of N-terminal peptide sequences showed the relative positions of the proposed cleavage sites, related positive residues, and locations of the possible mitochondrial presequences (Figure 4.1). Most were shown to be fairly similar in terms of relative positions of the three possible cleavage sites, but CYP4G27 had a longer hydrophobic segment before the first cleavage site and the CYP9s had only two sites detected by SignalP analysis. CYP4BE1 seemed to have sites that were closely clustered and the third cleavage site was immediately followed by the first functional motif (WXXXR), which in the others is farther downstream (not shown). Positive residue clusters follow the first cleavage sites in the other enzymes and also the second position in CYP4AY2 and third of CYP4BE1 (the location is similar to the second site of the others). The third site was often followed by one or two serine residues, except in the case of CYP4G27, which has several serines close to the N-terminus. Possible mitochondrial presequence sites tended to occur near cleavage site one (PSORT II) and were identified in CYP4AY1, CYP4BF1, CYP4BJ1, and CYP49T2. Those detected by MITOPROT II (underlined) occurred between cleavage sites one and two.

> **CYP4AY1**
MILFEVVVVILTIWYIRKCR**FREKV**KWVATVPGHPILGVALDLTDPTKTLDRMTEYLTKYNGMCYTEFMLCPTLVVSDLSFLKWFLTSNIPIHKGDIYAVL

> **CYP4AY2**
MLFWLFQVVFALAAAFCIAFCLKKYNFKKKISWVPFVSGYPVIGAALELRDRKKILENIENHLNKHNGLCYMEVGTIPMLMASNPDFLKWFLSSNLPLAKS

> **CYP4BD1**
MLFVLLALVTCILLGLFTLYIVAKIKKDTSKYLTNVPCSEQTFPFGNTLPFLKGSVVYLDLIMQGVKELGRTSLFHDGPLSWVVITADPEFFIELIYSSST

> **CYP4BG1**
MKMFIGVLCGLVAVIFLFLLRKYNKLRKKR**CRIPN**PPGHWLLGNLTLNNLNSGELFDQLRNFARDYGPIYRISIPFLDIVNFFHPADLEIILSQKKHMKKS

> **CYP4G27**
MSTATLSSAAAPGLLTSTNLFLFLLAPALALLYVYWKV**SRKH**MVELAERIPGPSGLPILGNALEFIGTPNQIFNTIYQKSFEFGRTIKVWVGPRLLIFLTD

> **CYP4BE1**
MIKAWLGPVPILFVVDPDSAKVILESNTLITKSSFYDKVADWIGTGLLISTNEKWQSRRKLLTPAFHFNILKGYTEVMVKEGEVFVDQLDKLADTGREFNL

> **CYP4BF1**
MGFWVPVVVLLLFVYATLRIWPWILKRKKRLIQMVDRIPGPTAIPILGCAYQ**FRPKI**EDFSYELLEYARLHKDSEVVRFWLGPIPIVCAFGPESVKTVLESN

> **CYP4BH1**
MFFLTVFTALATVLAYYVGRWWMKLHN**ARNKL**ASWMPTAPGHPLIGSVFEFGDNTVFTTNLHRLTDNPCKSCYVEVMGEWMILTRDHDLLDLVLSSNKILT

> **CYP4BJ1**
MNPTVWLLVVFLTFLFLYWKRLWLAFRERRR**RFTH**MINQLPGPMGLPLVGSALNFSPDTEKCTYQMELYFRTYTEDIDSSGIMRVWIGPKPLVFIYKPETAK

> **CYP9T1**
MLVGLVLVAVLALLFYYQF**VRPLN**HFTKMGVKQTNTALPIFGDRWGVELRLDKSYFDLIKRVYFSCDKDDRFVGLYNFTRPILFIRDPQLIKELGIKHFDS

> **CYP9T2**
MLVELVLVAILALLFYYQFVRPLNHF**TRMGV**KQTNTALPIFGDRWGVELRLDKSYFDLIKRVYFSCDKDDRYVGLYNFTRPILFIRDPQLIKELAIKHFDS

Figure 4.1 - Characteristics of P450 N-termini. SignalP-derived cleavage sites are in boxes and positively charged residues are shaded. Smaller boxes indicate where SignalP detected another possible location for the first site based on neural networks. Mitochondrial presequence motifs recognized by PSORTII are in boldface type, and any presequence cleavage sites predicted by MITOPROT II are underlined where peptide removal would have occurred.

4.3.2 Experion Analyses of Cellular Fractions

Experion analyses of different cellular fractions collected from transfections revealed that recombinant proteins were not only located in the ER. As cell debris would include whole cells not disrupted by sonication (and thus ER), recombinant protein appeared to be in that fraction (Figure 4.2a). As indicated (arrows) in lane 4 of Figure 4.2a, ~53 kDa and ~60 kDa bands were present for all but CYP4G27 (which was bigger than the other four recombinant proteins and resolved at ~68 kDa). Once 5 kDa was subtracted for the histidine tag, the ~53 kDa and ~60 kDa bands represented the ~48 kDa size detected earlier during optimizations, and larger ~55 kDa isoforms. The unique band in lanes 9 and 10 would then be close to the expected size of CYP4G27 (~62 kDa) at ~63 kDa. A smaller band at ~65 kDa was detected for all tagged CYP4G27 transfections, at a low concentration, except for the cytosolic fraction. Most surprising was the obvious presence of recombinant protein in the mitochondria (Figure 4.2b). The intensities of the bands indicated relatively high amounts of recombinant protein. The cytosolic fraction did not seem to have the lower molecular weight (LMW) version of the proteins as the ~53 kDa band was indistinguishable from a co-migrating protein in the cell control, except for CYP4BD1, but did have higher molecular weight (HMW) proteins. CYP4G27 levels in the cytosolic fraction were barely detectable (Figure 4.2c). In all of the digital gels, CYP4BD1 lanes seemed to have another lower molecular weight band below the ~53 kDa representative, but this also occurs in Figure 4.2c, associated with the HMW protein (indicated), so that size might be a result of cleavage or breakdown occurring during electrophoresis. This character of CYP4BD1 lends support to the other ~60 kDa digital gel bands in the cytosolic fractions being recombinant protein.

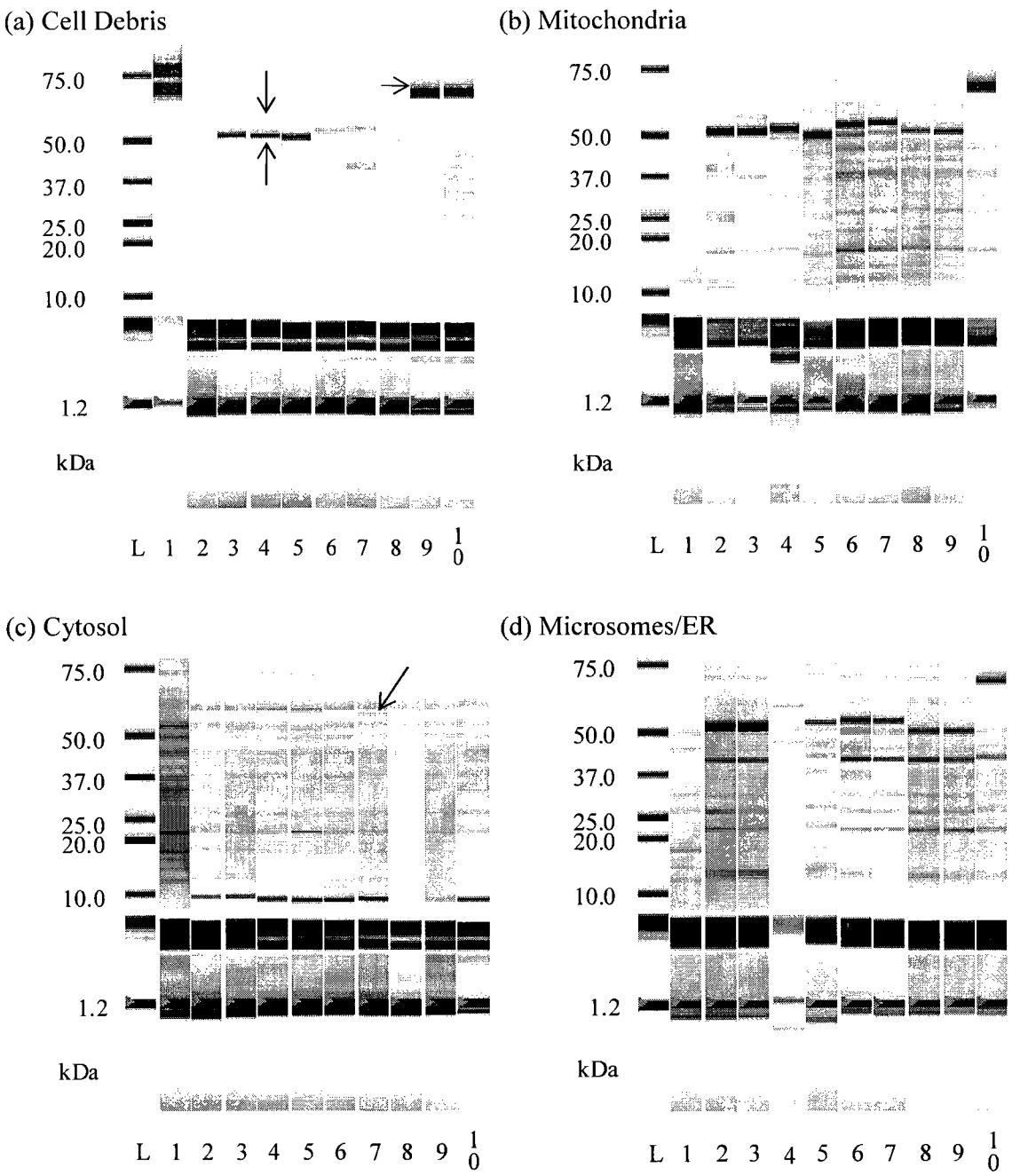


Figure 4.2 – Experion analyses of subcellular fractions. All lanes marked 'L' are protein ladder (in kDa). (a) Cell Debris: Lane 1 is representative of a secreted protein isolate, for comparison; lane 2: uninfected cell control; lanes 3-4: CYP4AY1; lane 5: CYP4AY2; lanes 6-7: CYP4BD1; lane 8: CYP4BG1; lanes 9-10: CYP4G27. (b) Mitochondria: Lane 1 is uninfected cell control; lanes 2-3: CYP4AY1; lanes 4-5: CYP4AY2; lanes 6-7: CYP4BD1; lanes 8-9: CYP4BG1; lane 10: CYP4G27. (c) Cytosol: Lanes are as in (b). (d) Microsomes/Endoplasmic Reticulum: Lanes are as in (b), except lane 4 failed.

4.3.3 Western Blots of Tagged Recombinant CYP4 Protein

To investigate whether the faint HMW bands detected by the Experion were recombinant protein, the maximum allowable volume of sample was next evaluated by western blot. Unfortunately, this volume resulted in nonspecific binding in samples from membrane fractions. It is difficult to reconcile the two sets of results, especially given that both systems have a 5-10% margin of error and the different size groups are within that error of eachother. Western blots, however, did indicate that membrane-associated fractions had detectable faster- and slower-migrating forms for all but CYP4BG1 (~60 kDa) and CYP4G27 (~68 kDa) (Figure 4.3). This mostly agreed with the findings from the Experion analyses, except for CYP4BG1. The LMW size was clearly detected as the more prevalent form by the Experion (and during optimization) and yet it did not seem to resolve, unless the only resolving band is the LMW version (it is only 3 kDa higher than the expected ~53 kDa from Experion analysis). CYP4G27 resolved as one band (bands from nonspecific binding corresponded to molecular weights too large to be recombinant protein, at over 74 kDa), but it was such an intense signal that it could have occluded another similarly-sized protein.

CYP4AY2 membrane-containing fractions (cell debris mitochondria, microsomes, Figure 4.3) had several bands in close proximity, and considering the error margin, it was difficult to resolve them. The 62.3 kDa band was faint, but the size was still calculated for comparison. The most intense band should represent the LMW version detected by the Experion, as it was the most concentrated. Based on that assumption, there was also a faster-migrating form of CYP4AY2. However, the SDS-PAGE gel fell prematurely onto a lower portion of the PVDF and had to be readjusted prior to transfer. Blotting detected bands where the gel had been (not shown), so it is possible some intensity was lost to that error, so

intensities may not have been as correlated to actual concentrations. CYP4BD1 and CYP4AY1 membrane-containing fractions seemed to have two bands of the expected sizes, agreeing the best with the Experion findings of two size types per recombinant protein.

Also in agreement with the digital gels, cytosolic fractions seemed to contain only one size, which may have been HMW proteins. CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 resolved at ~56 kDa, which did not match up with the other HMW predictions from western blots of other fractions. However, Experion digital analysis indicated that the unique HMW bands were approximately the same size across fractions, and the western blot of the cytosolic CYP4G27 matched up in terms of size with the other blots. It may be that the SDS-PAGE gel used for the other cytosolic fractions had a slightly different composition that impacted migration. That said, the 10% error margin easily encompassed this slight discrepancy. CYP4BD1 did have a detectable unique band at ~ 53 kDa on the digital gel, in addition to the ~60 kDa version, with also matches the blotting findings. The lack of the smaller version could have been due to these fractions being less concentrated, but in any case, it is obvious that the cytosolic fraction, unlike all of the others, seemed to favor HMW protein. CYP4BD1 was the only sample that still resolved as two size types and this provided a good comparison for the other proteins, further indicating that they were likely the HMW versions. Although the error margins precluded absolute identification of specific bands, it was clear that recombinant proteins were in multiple fractions and sometimes more than one size could exist. This supported the digital gel findings where all but CYP4G27 resolved primarily at ~53 kDa (~48 kDa without the tag), with a larger version present at lower concentrations. Based on this, digital gel bands distinct from the cell controls were considered possible P450 isoforms.

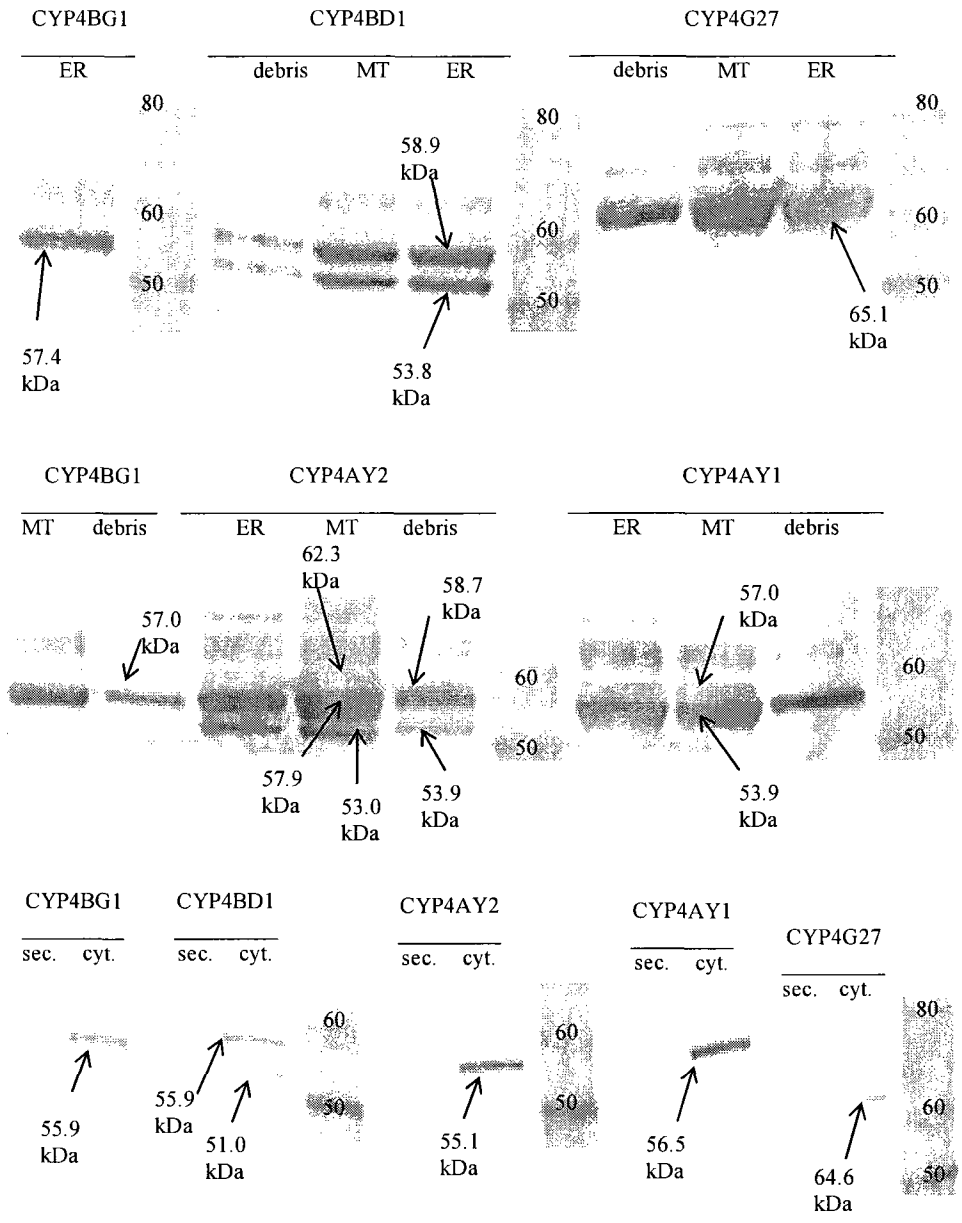


Figure 4.3 – Western blots of histidine-tagged recombinant proteins. Tagged CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and CYP4G27 protein samples were analyzed. Western blots of recombinant protein are shown for SF9 cellular fractions (debris – cell debris; ER – microsomes; MT- mitochondria; cyt. – cytosol; sec. - secreted). Unmarked lanes are protein ladder. Cell controls are not shown. Despite non-specific binding, it can be seen that membrane-containing fractions seem to have contained two bands, with the exception of CYP4BG1. The sizes of these bands seemed to correspond to the sizes estimated on the digital gels. Cytosolic fractions, with the exception of CYP4BD1, only contained one size of recombinant peptide, and there were no detectable secreted recombinant proteins.

4.3.4 Relative Proportions of HMW and LMW P450 proteins by Cellular Compartment

Relative concentrations of the designated LMW and HMW proteins were compared in isolates representing different cellular compartments, based on the concentrations derived from the Experion digital gels (Figure 4.4a,b). CYP4G27 was represented by ~65 kDa and ~68 kDa forms, and the other four CYP4s by ~53 and ~60 kDa forms (5 kDa for histidine tag included). Concentrations of recombinant proteins detected in cytosolic fractions were multiplied by a factor of 25/3 to compensate for its suspension in 1/3 cell culture volume of lysis buffer. All other samples (pellets) were resuspended in 1/25 cell culture volume. One size was predominantly observed for CYP4G27, agreeing with the western-blot findings, but CYP4BG1 resolved as two sizes like the others, opposing the western-blot result. Surprisingly, CYP4AY2 and CYP4BD1 had higher mitochondrial levels of protein than the microsomal fraction (Figure 4.3). There was also an apparent preference of LMW protein over HMW protein in mitochondrial isolates. HMW proteins were often found in the cytosol, but an additional (cell control) protein at ~53 kDa could have occluded detection of LMW protein occurring at very low concentrations from being detected. However, a general lack of LMW proteins would be in agreement with the western blot results.

CYP4G27's relative proportions of HMW and LMW proteins opposed to that of the other CYP4s, with HMW being the predominant protein isoform by far. Mitochondria had a higher concentration than the ER, which is more than would be expected with cross-contamination alone. The high presence in cell debris compared to the other proteins could be the result of more whole cells left intact after sonication.

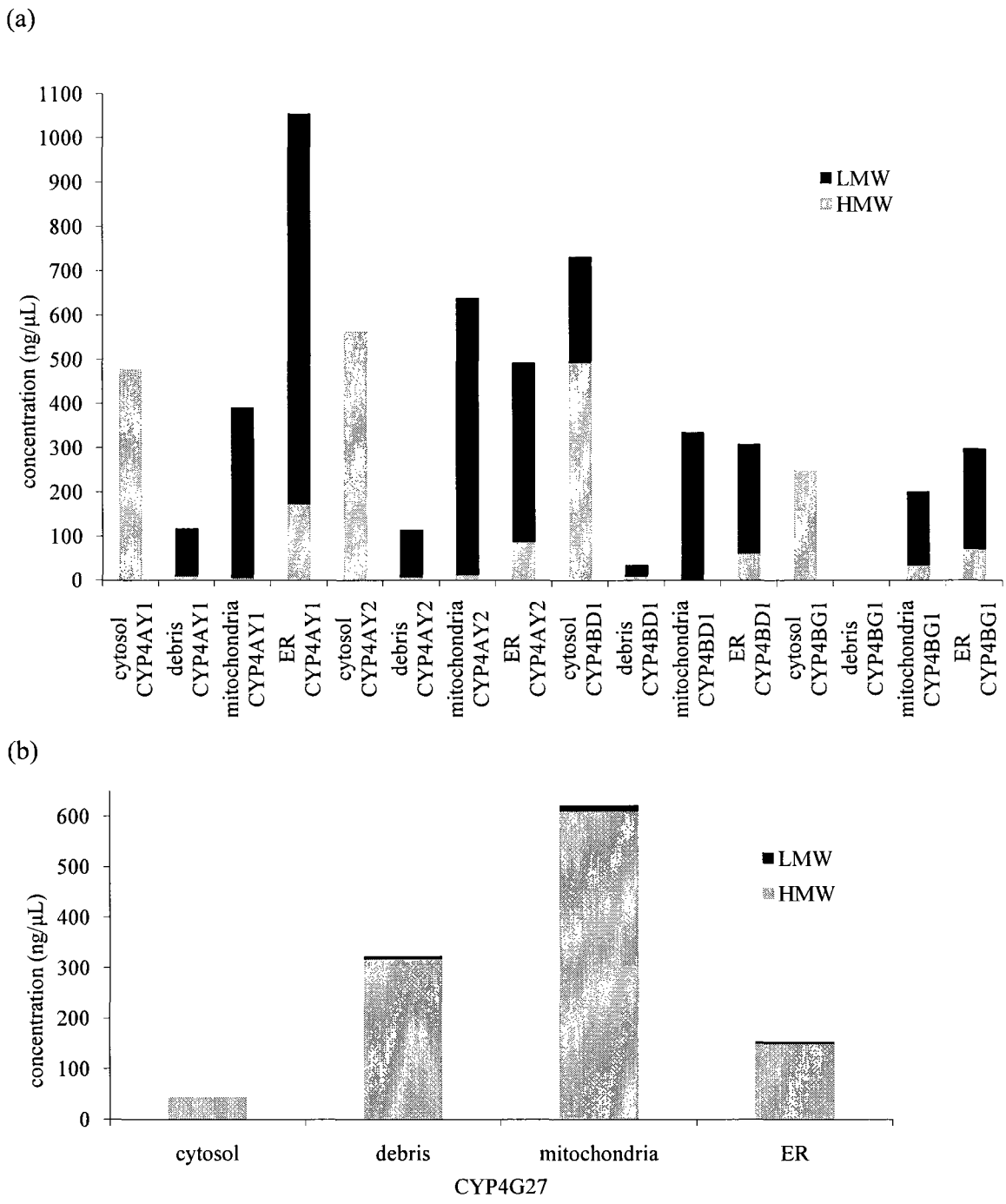


Figure 4.4 - Relative contributions of designated lower and higher molecular weight proteins. Sampes were evaluated by capillary electrophoresis. (a) LMW (~48 kDa) and HMW (~53 kDa) content of CYP4s were determined by capillary electrophoresis and values from the resulting digital gel are represented in ng/µL. Comparisons between recombinant enzymes cannot be made as each construct has variant productivity. (b) Capillary electrophoresis evaluation of CYP4G27 localization. As before, concentration values were derived from bands distinct from the cell controls on digital gels.

4.4 Discussion

4.4.1 *In silico* Analyses of N-termini

SignalP analysis indicated several cleavage sites upon which much of the subsequent analyses were based. These should be considered with caution given that the intended utility of the software was to analyze putative peptide sequences. Normally one does not artificially truncate the peptide chain; this was done to emulate processes that might have been occurring in the cell. The cleavage sites themselves were predicted by not just the probability of cleavage, but also by evaluating characteristics of the signal peptide, so errors may be further compounded if an undefined type of signal is involved (e.g. mitochondrial import). Also, transport pathways outside of the classical signal recognition particle (SRP) and simple N-terminal signal models are far from being completely understood, especially those involved with mitochondrial import. Thus, one should look at the results as a broad estimate of possible processing. Individual findings do not stand alone however; all of the output, across the numerous peptide sequences, should be considered together.

Based on the probabilities, SignalP analysis supported the presence of the first cleavage site (Table 4.1). The very detection of cleavable signals early in the N-termini, and lack of clear support for anchors, supported the idea that 'normal' microsomal processing may not have been taking place. That is not to say that the proteins could not end up in the microsome. This potential processing may be a mechanism for some of it to be liberated from the ER and thus exposed to cytosolic elements responsible for targeting to alternate locations. TargetP analysis (Table 4.2) also supported the notion that un-cleaved protein could be staying in the ER in that N-termini may be associating there and somehow become excluded from proteases – perhaps through reorientation that eliminates loop formation of the peptide

chain required to expose cleavage sites to proteases inside the ER. All of the P450s in this study were presumably targeted to the ER prior to any cleavage events owing to the detection of probable signal peptides (Table 4.2). ER retention and retrieval signals played a role in the localization of some membrane proteins, including some plant P450s (McCartney *et al.*, 2004). A mammalian P450, CYP2C2, which lacks a recognizable retrieval signal, is able to associate with the ER despite truncation to remove the N-terminal anchor, while still retaining functionality (Szczesna-Skorupa *et al.*, 1995). So the occurrence of cleavage events would not eliminate the possibility of ER targeting.

TargetP and Predotar results did not provide conclusive subcellular locations for peptides after the second and third *in silico* cleavage site. However, they indicated that after the first portion of peptide was removed at cleavage site one, mitochondrial targeting was a possibility (0.418 and 0.415 probabilities) for CYP4BH1 and CYP4BJ1 (Table 4.2). Predotar analysis found that mitochondrial targeting was possible after cleavage at the first site in all but CYP4BE1, CYP4BJ1, and the CYP9s (Table 4.3). In regards to cleavage sites two and three, there are several reasons why the programs often failed to predict either ER or mitochondria as the location. First, the programs were not intended for internal targeting sequence prediction. Sequences needed the addition of an N-terminal methionine to even be processed by Predotar – which cites in the manual that additional targets are difficult to elucidate with such programs (Small *et al.*, 2004). Only classical import pathways like the ER or the outer mitochondria are recognized. In fact, cytochrome *c*, which is a mitochondrial enzyme, does not have its targeting properties recognized by Predotar software (Small *et al.*, 2004). Secondly, there are many other targeting possibilities. For example, mammalian

CYP2E1 has representatives in the ER, mitochondria, lysosomes, peroxisomes, and the outer membrane (reviewed by Neve & Ingelman-Sundberg, 2001).

Analyses of putative N-terminal peptide sequences supported mitochondrial targeting in some of these P450s. MITOPROT II analysis (Table 4.4) clearly indicated that mitochondrial targeting was highly probable for CYPBF1 and CYP4BH1 (0.9359, 0.9393) and likely for CYP4BJ1 (0.8068), based on the putative peptide sequences. Other probability levels were not as high, but the analyzed CYP9s had probabilities approaching 0.5. The program also detected mitochondrial presequence motifs in CYP4AY1, CYP4BG1, and the two CYP9s. PSORT II detected areas where presequence may be cleaved in CYP4AY1, CYP4BF1, CYP4BJ1, and CYP9T2. The existence of these sites is interesting, but taken alone do not really provide strong evidence for mitochondrial import. They are short consensus regions based on known processing sites, but that does not that mean they are actually exposed to the processing elements that would carry out the cleavage events. The results indicated that the processing events occurring at the mitochondria were possible.

Figure 4.1 illustrated the relative locations of the possible cleavage sites, targeting features, and the nature of the surrounding residues. The locations of the cleavage sites for CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 seemed to correspond; this supported the existence of true cleavage sites. They also corresponded with some of the predicted mitochondrial presequences (boldface and underlined). The first cleavage site of CYP4AY1 occurred where a mitochondrial presequence was predicted by two programs. This also occurs in CYP9T2 and seemingly in CYP4BJ1. If the cleavage sites or presequence estimates were random, one would not expect the observed pattern across different peptide sequences.

There seemed to be positive residues after the first predicted cleavage site in most P450s, which hinted toward more processing. Positive residues have been implicated with mitochondrial targeting, because it is thought they are needed to interact with the transport proteins on the surface of the organelle. It is possible that, after the first cleavage occurs, these residues were available for binding to the mitochondria. This has been proposed for the mitochondrial 'P450MT2', which is an alternately processed form of CYP1A1. Addya *et al.* (1997) studied this system in rats and proposed that some protein escaped microsomal insertion and that cytosolic cleavage revealed a mitochondrial target sequence which was partially composed of positively charged residues. Two cleavage sites for cytosolic proteases were proposed at the 4^{th} and 33^{rd} amino acids. The cleavage at the latter site seemed to be dependent on cleavage at the 4^{th} site (Addya *et al.*, 1997). A study of known mitochondrial target sequences found that positively charged residues were implicated in the formation of amphiphilic helices, where positive residues reside on one side of the helix (von Heijne, 1986). CYP2E1, truncated to remove the SRP-signal at the N-terminus, was targeted to the mitochondria (Neve & Ingelman-Sundberg, 2000). A model was proposed where cytosolic chaperones target the peptides to the mitochondria, where binding to the negatively charged TOM (transport outer membrane) complex is facilitated by the aforementioned positively charged residues. Once across the outer membrane, interaction with the TIM (transport inner membrane) complex may take place, with CYP2E1 finally residing in the matrix and the mitochondrial presequence removed (Neve & Ingelman-Sundberg, 2000). The observed positive residues that were next to some of the proposed cleavage sites further support the possibility of mitochondrial targeting for these particular P450s (Figure 4.1).

4.4.2 Western Blots of Cellular Fractions

Western blots of histidine-tagged constructs (Figure 4.2) supported the presence of cleavage sites as most variants occurred as differentially migrating pairs. These generally were evident in the ~60 kDa and ~53 kDa size range, which, when the ~5 kDa histidine-tag was removed, corresponded to molecular weights of ~55 kDa and ~48 kDa (Table 4.1). Of note, these sizes seemed to correspond to the predicted molecular masses resulting from cleavage at sites one and three (Table 4.1). In the case of the slower migrating species, the likely candidate is the result of the first cleavage site or even no initial cleavage at all. However, due to variable factors inherent in interactions between the gel matrix and the migrating proteins, it is possible that this represents the second cleavage site, as there is only a ~2-4 kDa difference between the protein sizes resulting from cleavage sites one and two. The next size possible could be the result of the third cleavage site, which would have resulted in a product about ~4 kDa smaller than the result of cleavage at site two. There is an estimated ~6 kDa difference between the products of cleavage sites one and three, which is approximately the difference between the resolved bands on both Experion digital gels and western blots. Taken together, this is evidence for the first and third processing sites. The lack of a clear midrange size for the proteins might have been a result of rapid turnover. The intermediate protein product may be immediately processed. Assuming sites one and three are cleaved, site two may be part of a two-step cleavage that ends at site three, or site two may not be real. The sizes are too close and within error, so positive identification of sites is not possible by the methods of the current study; peptide sequencing would be needed.

4.4.3 Comparative Analyses of Recombinant CYP4 Abundance in Cellular Fractions

Experion analysis provided some insight into the processes that may have occurred. However, it should be noted that cross-contamination was likely in these membrane fractions. To begin with, the cell debris fraction included nuclei and cell membrane components. Whole cells not disrupted by sonication would also be present, and so the fraction thus contained both ER and mitochondria. The next differential centrifugation fraction, containing mitochondria, is known to have "significant" contamination from other cellular compartments when isolated in yeast, particularly from the ER, which lead to the development of sucrose gradients for mitochondrial purification (Miesinger *et al.*, 2000, p. 339). That said, even with cross-contamination in crude mitochondrial samples, some interesting patterns were still gleaned from the comparative concentration data derived from the digital gels (Figure 4.3). First, the cytosol is lacking LMW P450 variants, with the exception of CYP4BD1. This indicated that LMW P450 variants were either retained completely in other structures and processed there, or immediately carried to other compartments by chaperones following cleavage by cytosolic proteases. The mitochondria, on the other hand, seem to predominantly contain the LMW P450 protein variants. This could mean that only completely processed (cut) protein was targeted there, or that HMW proteins were efficiently processed to LMW proteins at the mitochondria. The ER had mostly LMW proteins, but HMW proteins were present. This could be the location of HMW processing to LMW, but that would be contrary to the presence of mitochondrial-targeting sequences occurring after cleavage site one, unless there is another way that the potential mitochondrial signals can be present. It could be that some of the HMW is simply retained in

the ER or retrieved. Models are proposed here that attempt to explain the observed patterns, beginning with what seemed to be the simpler of the observed patterns, CYP4G27.

CYP4G27 seemed to have higher concentrations in the mitochondria while maintaining a population in the ER (Figure 4.4b). Mitochondrial targeting of CYP4G27 was also supported by the presence of two positive amino acids at the beginning of the L-helix (not shown), which is a character that has been implicated in proper interaction with mitochondrial electron donors, and is indicative of P450s targeted to that organelle (Werck-Reichhart & Feyereisen, 2000). The HMW protein variant was by far the predominant version, with only low levels of LMW protein (the presence of which was not corroborated by western blotting) (Figure 4.4b). It is possible that the HMW variant detected was the result of cleavage at the first proposed site, or no cleavage at all. There is only a ~3 kDa difference between those predicted sizes, and both western blot and Experion analyses had intense main bands that could have occluded smaller nearby ones. If a cleavage event did occur, it could be explained by a model proposed by Addya *et al.* (1997), where some P450s escape insertion into the ER, allowing proteolytic cleavage that permitted processing of a mitochondrial target sequence, while un-cleaved protein remained in the ER unaffected. Bhagwat *et al.* (1999) found that residues 1-44 targeted both ER and mitochondria, but that 33-44 were mitochondria-specific. Some CYP4G27 may be translated into the ER normally, but another portion may be cleaved in the cytosol, and thus targeted to the mitochondria. The lack of a probable third cleavage site could indicate that CYP4G27 or the CYP9s simply have shorter regions with signaling capacities, so the variants would be difficult to distinguish by size. However, to my knowledge, mitochondrial targeting has not been investigated for these CYP9s, so they may not share that character with CYP4G27.

Given that the predicted molecular weights are close to those of the expressed proteins for CYP4G27 and the CYP9s (Sandstrom *et al.*, 2006, 2008), it is also possible that proteolytic cleavage does not occur at their N-termini. Un-truncated peptides have exhibited chimeric signaling properties. 'P450MT5' (alternately processed microsomal CYP2E1) phosphorylated at a particular serine was targeted to the mitochondria while unphosphorylated peptide was not (Robin *et al.*, 2001). N-terminal peptide sequences of both microsomal and mitochondrial species were found to be identical. Phosphorylation may activate cryptic mitochondria-targeting signals that are not at the N-termini (Omura, 2006). Some of the N-terminal sequences evaluated here had conspicuous serine residues C-terminal to the third cleavage site. Alternately, CYP4G27 had serines interspersed throughout the first 20 amino acids (Figure 4.1). CYP4G27 manifesting as predominantly one size, contrary to the findings for the other CYP4s investigated, could indicate that they represent different targeting mechanisms.

Of note in the dual targeting of P450s is the regulatory control exhibited by cytosolic endoproteases, which could have played a role in regulating which P450 isoforms were processed. Boopathi *et al.* (2000) found that the brain tissue of rats differentially expressed mitochondrial 'P450MT2' relative to the microsomal CYP1A1 with chronic β-naphthoflavone exposure. These two P450 isoforms were derived from the same putative peptide. At 72 hours, the resulting microsomal fraction was drastically reduced, accompanied by the reciprocal response in the mitochondrial isoform. They proposed that a surge of cytosolic endoproteases, reacting to the β-naphthoflavone exposure, could have been influencing localization by cleaving the putative peptide, leading to alternate targeting. If such targeting can be attributed to a cellular stress response, extended exposure to

baculovirus in the heterologous protein expression system could be a factor influencing targeting in the recombinant P450s that I tested. The possible roles of cytosolic proteases and those on the luminal side of the ER could be involved with the potential processing events that were observed for the other CYP4s.

There was evidence to suggest there were at least two differently-migrating versions of CYP4AY1, CYP4AY2, CYP4BD1, and CYP4BG1 proteins. A possible model for targeting begins similarly to that described by Addya and colleagues (1997). The P450s are being translated by a ribosome in the cytosol 'normally', and when the SRP signal emerges from the nascent chains and SRP binds to it, translation stops until the complex binds to a receptor at the ER. The P450s may span the membrane once or twice with an anchor that encodes a stop-transfer signal, as is proposed for most microsomal P450s (Bar-Nun *et al.*, 1980). Perhaps there is a cleavage site downstream of the anchor, exposed to the cytosol (and its proteases). When mitochondrial targeting is required, proteases cleave some of the P450s, facilitating their 'escape' from the ER into the cytosol – this is the HMW protein detected there and if the anchor is not very long, the uncleaved protein left in the ER membrane may not be sufficiently different in size so as to be differentiated by the methods used here. This cleavage event may reveal or activate a mitochondrial-targeting signal. Once translocated across one or both membranes of the mitochondria, a mitochondrial presequence is removed by proteases in the matrix and predominately LMW protein is thus detected at that organelle. Cleavage may reveal peptides that signal incorporation into the inner mitochondrial membrane, anchoring the P450, but there are also cases where P450s enclosed by the mitochondrial inner membrane were soluble (Omura, 2006).

The presence of LMW protein at the ER can also be explained by proteases. Perhaps to prevent release from the ER by residual cytosolic proteases, the first cleavage site is removed. It has been proposed that P450s span the membrane twice with a loop structure and that loop later stretches out (Monier *et al.*, 1988); this could account for the lack of cleavage generally observed in P450 translocation. Classic SRP-processing involves a cleavage event that removes the N-terminal signal sequence, resulting in translocation. Perhaps that loop structure forms on the luminal side of the ER and thus exposes another cleavage site that would normally be on the cytosolic side if the loop stretched back out. The loop may expose this cleavage site to luminal ER-proteases. This would remove the N-terminus and thus the first cleavage site that permitted escape to the cytosol. A secondary anchor at the new N-termini (resulting from cleavage) would then keep the protein associated with the ER. This could explain why predominantly LMW protein was associated with the ER and fits with the hypothesis that mitochondrial P450s are derived from mis-targeted microsomal P450s. However, this particular model has not been proposed previously.

As to the actual locations of potential cleavage sites, SignalP analysis indicated suspect regions, but these cannot be confirmed without sequencing the recombinant proteins. However, the validity of the program prediction method was tested by evaluating a known mitochondrial P450 that likely required processing (not shown). The *Drosophila* gene *shade*, a Halloween gene, is one of several mitochondrial P450s mediating the synthesis of 20-hydroecdysone from cholesterol (Gilbert, 2004). Although they found the enzyme only in the mitochondria, the author indicated it had an N-terminal peptide sequence that could facilitate microsomal targeting (hydrophobic residue stretch). SignalP analysis indicated a cleavage site (P=0.796) at the 17^{th} base from the N-terminus, and two other sites (P=0). TargetP and

Predotar analyses indicated the uncut peptide targeted the ER (0.975 and 0.99, respectively). After cleavage at site 17, the Predotar program detected possible mitochondrial targeting (0.51). Although only recognized by one program, the correct localization of a known mitochondrial P450 shows that under the right circumstances, the program can detect the signal. Additionally, PSORT II detected a mitochondrial presequence motif at site 40, which is one amino acid away from where SignalP detected the second possible cleavage site.

CYP1A1, which is known to have an N-terminus with chimeric signaling properties (Addya *et al.*, 1997), was also evaluated as above. Addya *et al.* (1997) suspected N-terminal cleavage events to be occurring at amino acids 4 and 33, resulting in alternative mitochondrial targeting. SignalP did not detect the cleavage at site 4, but a possible site was recognized at amino acid 35 (0.342). The first site may not have been detected because the 'best' site in a stretch of 70 amino acids is selected by the program. TargetP predicted the uncut peptide and peptide cleaved at site 4 went to the ER (0.960). Cleavage at the authors' predicted site (33) resulted in mitochondria being the predicted location (RC=5); cleavage at amino acid 35 did not have a predictable location (MT= 0.284, ER= 0.231, other= 0.350). So although not definitive, my approach did detect one of the cleavage sites within 2 amino acids and a difference in predicted targeting was observed with *in silico* peptide cleavage.

4.4.4 Possible Chimeric Signals in Insect CYP4 N-termini

It is possible that the N-terminal sequences evaluated here are part of complex, chimeric targeting signals. Animal models have revealed that several human P450s thought to be microsomal were also present in the mitochondria (Nirajan *et al.*, 1984; Addya *et al.*, 1997; Neve & Ingelman-Sundberg, 1999; Anandatheerthavarada *et al.*, 1997). Research of microsomal P450s with mitochondrial counterparts in liver and brain tissue has shown that

there can even be two mitochondrial isoforms in addition to the version in the ER (Addya *et al.*, 1997; Boopathi *et al.*, 2007). These variants from the same initial peptide sequence can adopt different configurations, allowing them to interact with alternate electron donating systems and substrates (Robin *et al.*, 2001).

In terms of functional prediction, these targeting mechanisms could aid in determining potential substrates. It has already been mentioned that there are mitochondrial P450s involved in the conversion of cholesterol to ecdysone and 20-hydroecdysone (Gilbert, 2004). Mammalian P450s have been characterized that act in the biosynthesis of steroid hormones from cholesterol, the metabolism of sterols, and vitamin D3 conversion (Omura, 2006). It is unknown whether mitochondrial P450s have a large role in the breakdown of xenobiotics (Omura, 2006). The possibility that insect P450s may have chimeric signaling properties warrants further investigation as this could indicate general functional groups.

Subcellular targeting mechanisms have been detected by this exploratory study, but not to the degree necessary to draw further conclusions. Mitochondria could instead be isolated with a sucrose gradient, as opposed to differential centrifugation, to ensure a more pure sample (e.g. Neve & Ingelman-Sundberg, 1999; Anandatheerthavarada *et al.*, 1997). Also, amino acid sequencing of the N-terminus of each version would confirm whether or not the proposed proteolytic cleavage events are occurring (e.g. Addya *et al.*, 1997). As mitochondrial P450s resulted from mis-targeting of microsomal P450s (Werck-Reichhart & Feyereisen, 2000), we should not be surprised to find representatives intermediate to either targeting pathway. Evidence for chimeric signaling properties of CYP4AY1, CYP4AY2, CYP4BD1, CYP4BG1, and to some extent CYP4G27, warrants further investigation of the subcellular targeting of these P450s.

4.5 Conclusions

Collectively, bioinformatics and targeting analyses indicated that there was some form of processing occurring at the N-termini of these CYP4s. Limitations of understanding and predictive tools, in regards to protein targeting to certain subcellular compartments, made it difficult to specificy locations. Evidently, mitochondrial targeting and processing by proteases were possible. Confounding resolution was the possibility of chimeric signals, which by their nature might produce conflicting results. Given that there are examples of P450s that localize to the cell membrane, peroxisomes, ER, and mitochondria, possible targeting mechanisms are numerous. Different isoforms can even exist in single compartments, as with 'P450MT2a' and 'P450MT2b' (Addya *et al.*, 1997). It was noted that results implicating functional enzyme in multiple compartments have been largely ignored (Feyereisen, 1999), which could be why multiple targeting has not been reported in the literature more often.

A simple un-cleaved signal recognition peptide-anchor model does not adequately represent the diversity of P450s. N-termini with chimeric properties represent yet another tier of diversity that, to my knowledge, has yet to be exemplified in insects. It is of note that mitochondrial and microsomal P450s can have roles in the same pathways; such is the case with the steroidogenic pathway in humans (Guengerich, 2004). A review of mammalian mitochondrial P450s noted their involvement in synthesizing hormones and bile acids from cholesterol, as well as metabolizing Vitamin B3 (Omura, 1999). Perhaps pathways involving hormone synthesis or other endogenous compounds should be investigated for substrates of these CYP4s. However, unlike other animals, insect mitochondrial P450s have also been implicated in xenobiotic-metabolizing roles, based on studies of *Drosophila*, so that substrate

type cannot be excluded (Gillam & Hunter, 2007; Guzov *et al.*, 1998). It is unclear what purpose dual targeting would serve in terms of function, but an interesting question would be how such a system could have developed. There may be mitochondrial counterparts to other microsomal P450s that have as yet remained undetected that could help elucidate such pathways.

5.0 Conclusions and Future Perspectives

Mechanisms by which bark beetles detoxify host compounds, produce aggregation pheromones, and physiologically adjust to the changes associated with host colonization and reproduction all likely involve cytochrome P450 enzymes. A better understanding of the dynamics of this insect-host interaction could be obtained by elucidating the molecular basis of such activities. Knowledge of detoxification capabilities of one pest could potentially be extended to other similar pests, as was the case of functional characterization of *Ips pini* CYP9T2 and *Ips confusus* IcCYP9T1 (Sandstrom *et al.*, 2006, 2008), in which understanding of one protein led to the characterization of the second. The identification of P450 enzymes vital to host-colonization success could be used to predict the outbreak capabilities of bark beetle populations by surveying for variation in the activity of these enzymes.

To study specific P450s in isolation, heterologous expression with SF9 cells was used in this thesis. To optimize baculovirus systems, a balance must be struck between getting the most recombinant protein possible and keeping the cells healthy enough to produce protein effectively. As spectral measurements of P450s consume more sample and virus and are prone to several problems, I employed a capillary electrophoresis setup to gain specific protein concentration estimates in microsomal samples (Chapter 2). Like immunoblotting, this system was capable of distinguishing particular proteins in the sample and provided the information necessary to track the productivity of cells exposed to different amounts of virus. This method operated under the same hypothesis as the spectral technique, but measured the proteins in a different way. The method did not specifically measure functional forms of the P450s, but allowed for both the optimization of transfections to synthesize P450s, and of co-

transfections incorporating CPR. This technique offered significant savings in terms of time and resources while producing adequate amounts of the two enzymes for functional analyses.

Functional assays of CYP4AY1/CPR and CYP4AY2/CPR-expressing microsomes with multiple host-derived substrates did not yield products (Chapter 3). It is likely that this means that the tested substrates were not the correct substrates for the enzymes, but there were other factors. Cytochrome *b5* may have also been necessary in the system, which sometimes acts to transfer electrons or provide steric assistance (reviewed by Hannemann *et al.*, 2007). This enzyme, sometimes included in microsomal P450 systems, is not always necessary and is used to improve turnover, but some insect P450s have been found to require it for function (Gillam & Hunter, 2007). The CYP4 family of insect P450s is involved in a variety of pathways, so many other substrates are also possible, including, but not limited to, hormones, their intermediates, and fatty acids. Optimizations of CYP4BD1 and CYP4BG1 for coexpression with MPB-CPR and functional assays are pending. The MPB EST database is currently being mined for more bark beetle cytochrome P450 genes; full-length cDNA clones of several contigs (Chapter 3, Table 3.4) have already been obtained.

The diversity of P450s is also reflected in their processing. Although many P450s are microsomal (inserted into the ER through SRP-like mechanisms) or mitochondrial (probably utilizing TOM and TIM complex mechanisms), studies of rat P450s has revealed chimeric signaling properties of N-terminal peptides (e.g. Addya *et al.*, 1997). Also of note, conspicuous enzymatic activity in mitochondrial fractions has sometimes been observed in insect P450s that have been investigated as ER proteins (Feyereisen, 1999). *In silico* investigation of multiple bark beetle CYP4s and two CYP9s, and protein gel electrophoresis of different subcellular fractions, revealed that several *I. paraconfusus* CYP4s were possibly

being targeted to the mitochondria (Chapter 4). These results were supported by western blotting, although the exact locations of cleavage events were unclear. It is possible that targeting was being mediated by cytosolic elements such as proteases or chaperones, an unprecedented observation for insect P450s. However, more thorough isolation and identification techniques will be needed to corroborate the exploratory results of this work.

P450 enzymes likely interact with proteins besides their electron-transfer partners. It has been proposed that toxic products of some plant P450s, which if released into the cytosol could be lethal, are enclosed in enzyme aggregates (Werck-Reichhart & Feyereisen, 2000). If P450s are involved, they may also play an anchoring role if associated with certain membranes. Metabolic channeling in plants was shown to involve two P450s in a pathway where the toxic intermediates for the production of a cyanide precursor were sequestered by the enzymes mediating synthesis (Winkel, 2004). The aggregation of necessary enzymes in a pathway means that the intermediates are immediately available to the active sites of the next enzyme; when a toxic intermediate is involved, its release into the cytosol is also prevented to the benefit of that cell. Since P450s process toxins and other xenobiotics, this may be another facet to P450 localization that warrants further investigation, especially when it is considered that several pheromone precursors in bark beetles are toxic host compounds that may negatively impact cells if not sequestered. Although not investigated here, the idea of metabolic channeling highlights another intricacy of P450 localization that is still unclear.

Proteomics research provides more than just information about the cellular environment. To liken it to language, genomics data are like having a dictionary for an organism, where each gene is a word that describes something that organism must do. Although interesting, having a collection of words does not provide context. Expression

analyses then reveal which genes are involved at a given time or with a given stimulus, which would be like taking those words and making sentences. That can be informative, but only if one can read that language. Functional characterization of gene products would be the translation of those words. Once one has enough of the words translated, the sentences can begin to have meaning. The dictionary is then useful when another sentence is discovered that incorporates some of those translated words. If another organism uses similar words, something can be said about the sentences describing what it does to survive. Similarly, if they do not share certain words, that could say something about their differing capabilities or needs. In this way, genetic information is extended to the ecology of organisms.

Although the specific functions of these CYP4s were not determined, preliminary steps were made towards a useful optimization method that saved time and conserved valuable viral stock, cell line resources, and recombinant protein samples. Investigation of alternate targeting mechanisms revealed that these CYP4 family P450s may be capable of chimeric targeting, which has only been observed thus far in mammalian representatives. Considering how insect P450s are far more numerous than those of mammals, it stands to reason that insect systems encompass some alternate targeting mechanisms, if not more than what has been observed in other organisms. Future investigation of these and similar bark beetle P450s may require investigation of tissue localization to better narrow down substrate possibilities. Should alternate subcellular targeting be confirmed, it would also be a factor in predicting possible substrates. A more in-depth study of subcellular targeting will also be necessary to ensure that supplementation with the correct electron donors occurs. There remains much to be discovered in regards to the roles, synthesis, and protein-protein interactions associated with insect cytochromes P450.

Works Cited

Addya, S., Anandatheerthavarada, H.K., Biswas, G., Bhagwat, S.V., Mullick, J., and Avadhani, N.G. 1997. Targeting of NH_2-terminal-processed microsomal protein to mitochondria: A novel pathway for the biogenesis of hepatic mitochondrial P450MT2. *The Journal of Cell Biology* 139: 589-599.

Agee, J.K. 1998. The landscape ecology of western forest fire regimes. *Northwest Science* 72: 24-34.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 251: 403-410.

Anandatheerthavarada, H.K., Addya, S., Dwivedi, R.S., Biswas, G., Mullick, J., and Avadhani, N.G. 1997. Localization of multiple forms of inducible cytochromes P450 in rat liver mitochondria: Immunological characteristics and patterns of xenobiotic substrate metabolism. *Archives of Biochemistry and Biophysics* 339: 136-150.

Barnes, H.J., Jenkins, C.M., and Waterman, M.R. 1994. Baculovirus expression of bovine cytochrome P450c17 in Sf9 cells and comparison with expression in yeast, mammalian cells, and *E. coli*. *Archives of Biochemistry and Biophysics* 315: 489-494.

Bar-Nun, S., Kreibich, G., Adesnik, M., Alterman, L., Negishi, M., and Sabatini, D.D. 1980. Synthesis and insertion of cytochrome P-450 into endoplasmic reticulum membranes. *Proceedings of the National Academy of Sciences USA* 77: 965-969.

Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. 2004. Improved prediction of signal peptides: SignalP3.0. *Journal of Molecular Biology* 340: 783-795.

Bellés, X., Martín, D., and Piulachs, M. 2005. The mevalonate pathway and the synthesis of juvenile hormone in insects. *Annual Review of Entomology* 50: 181-199.

Berenbaum, M.R., and Zangerl, A.R. 1993. Furanocoumarin metabolism in *Papilio polyxenes*: Biochemistry, genetic variability, and ecological significance. *Oecologia* 95: 370-375.

Berenbaum, M.R., and Zangerl, A.R. 1994. Costs of inducible defense: Protein limitation, growth, and detoxification in parsnip webworm. *Ecology* 75: 2311-2317.

Bergé, J.-B., Feyereisen, R., and Amichot M. 1998. Cytochrome P450 monooxygenases and insecticide resistance in insects. *Philosophical Transactions of the Royal Society B: Biological Sciences* 353: 1701-1705.

Bhagwat, S.V., Biswas, G., Anandatheerthavarada, H.K., Addya, S., Pandak, W., and Avadhani, G. 1999. Dual targeting property of the N-terminal signal sequence of P4501A1. *The Journal of Biological Chemistry* 274: 24010-24022.

Bio-Rad. 2009. Experion^{TM} Pro260 analysis kit instruction manual (rev. B). Pp. 1-35.

Black, S.D. 1992. Membrane topology of the mammalian P450 cytochromes. *The Federation of American Societies for Experimental Biology Journal* 6: 680-685.

Bohlmann, J., and Keeling, C.I. 2008. Terpenoid biomaterials. *The Plant Journal* 54: 656-669.

Boopathi, E., Anandatheerthavarada, H.K., Bhagwat, S.V., Biswas, G., Fang, J., and Avadhani, N.G. 2000. Accumulation of mitochondrial P450MT2, NH_2-terminal truncated cytochrome P4501A1 in rat brain during chronic treatment with β-naphthoflavone. *The Journal of Biological Chemistry* 275: 34415-34423.

Borden, J.H. 1982. Chapter 4: Aggregation pheromones. In: Mitton, J.B., Sturgeon, K.B. Bark Beetles in North American Conifers – A System for the Study of Evolutionary Biology (1st ed.) (Pp. 74-139). Austin, Texas: University of Texas Press.

Borden, J.H., and Slater, C.E. 1968. Induction of flight muscle degeneration by synthetic juvenile hormone in *Ips confusus* (Coleoptera: Scolytidae). *Zeitschrift für vergleichende Physiologie* 61: 366-368.

Brobacher, M.G. 2008. The new face of electrophoresis: Modernization of a workhorse technology. *BioTechniques* 44: 568-570.

Brown, D.J., Clark, G.C., and van Beneden, R.J. 1998. A new cytochrome P450 (CYP30) identified in the clam, *Mercenaria mercenaria. Comparative Biochemistry and Physiology Part C* 121: 351-360.

Byers, J.A., Wood, D.L., Browne, L.E., Fish, R.H., Piatek, B., and Hendry, L.B. 1979. Relationship between a host plant compound, myrcene and pheromone production in the bark beetle, *Ips paraconfusus. Journal of Insect Physiology* 25: 477-482.

Byers, J.A. 1981. Pheromone biosynthesis in the bark beetle, *Ips paraconfusus,* during feeding or exposure to vapours of host plant precursors. *Insect Biochemistry* 11: 563-569.

Byers, J.A. 1982. Male-specific conversion of the host plant compound, myrcene, to the pheromone, (+)-ipsdienol, in the bark beetle *Dendroctonus brevicomis. Journal of Chemical Ecology* 8: 363-371.

Byers, J.A. 1983. Sex-specific responses to aggregation pheromone: Regulation of colonization density in the bark beetle *Ips paraconfusus. Journal of Chemical Ecology* 9: 129-142.

Byers, J.A., Birgersson, G., Löfqvist, and Bergström, G. 1988. Synergistic pheromones and monoterpenes enable aggregation and host recognition by a bark beetle. *Naturwissenschaften* 75: 153-155.

Byers, J.A. 1989a. Behavioral mechanisms involved in reducing competition in bark beetles. *Holarctic Ecology* 12: 466-476.

Byers, J.A. 1989b. Chemical ecology of bark beetles. *Experientia* 45: 271-283.

Byers, J.A., and Birgersson, G. 1990. Pheromone production in a bark beetle independent of myrcene precursor in host pine species. *Naturwissenschaften* 77: 385-387.

Casals, N., Martín, D., Buesa, C., Piulachs, M.-D., Hegardt, F.G., and Bellés, X. 1997. Expression and activity of 3-hydroxy-3-methylglutaryl-CoA synthase and reductase in the fat body of ovariectomized and allatectomized *Blattella germanica*. *Physiological Entomology* 22: 6-12.

Chen, N.-M., Borden, J.H., and Pierce, H.D., Jr. Effect of juvenile hormone analog, fenoxycarb, on pheromone production by *Ips paraconfusus* (Coleoptera: Scolytidae). *Journal of Chemical Ecology* 14: 1087-1098.

Chung, H., Szral, T., Pasricha, M., Batterham, P., and Daborn, P. 2009. Characterization of *Drosophila melanogaster* cytochrome P450 genes. *Proceedings of the National Academy of Sciences* 106: 5731-5736.

Claros, M.G., and Vincens, P. 1996. Computational method to predict mitochondrially imported proteins and their targeting sequence. *European Journal of Biochemistry* 241: 779-786.

Clontech Laboratories, Inc. 2008. BacPAK^{TM} baculovirus rapid titer kit user manual. Pp. 1-13.

Conn, J.E., Borden, J.H., Hunt, D.W.A., Holman, J., Whitney, H.S., Spanier, O.J., Pierce, H.D., Jr., and Oehlschlager, A.C. 1984. Pheromone production by anexically reared *Dendroctonus ponderosae* and *Ips paraconfusus* (Coleoptera: Scolytidae). *Journal of Chemical Ecology* 10: 281-289.

Dolník, V., Liu, S., and Jovanovich, S. 2000. Capillary electrophoresis on microchip. *Electrophoresis* 21: 41-54.

Duan, H., Civjan, N.R., Sligar, S.G., and Schuler, M.A. 2004. Co-incorporation of heterologously expressed *Arabidopsis* cytochrome P450 and P450 reductase into soluble nanoscale lipid bilayers. *Archives of Biochemistry and Biophysics* 424: 141-153.

Eigenheer, A.L., Keeling, C.I., Young, S., and Tittiger, C. 2003. Comparison of gene representation in the midguts from two polyphagous insects, *Bombyx mori* and *Ips pini*, using expressed sequence tags. *Gene* 316: 127-136.

Emanuelsson, O., Nielson, H., Brunak, S., and von Heijne, G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology* 300: 1005-1016.

Enoch, H.G., and Strittimatter, P. 1979. Cytochrome b_5 reduction by NADPH-cytochrome P-450 reductase. *The Journal of Biological Chemistry* 254: 8976-8981.

Fettig, C.J., Dabney, C.P., McKelvey, S.R., and Huber, D.P.W. 2008. Nonhost angiosperm volatiles and verbenone protect individual ponderosa pines from attack by western pine beetle and red turpentine beetle (Coleoptera: Curculionidae, Scolytinae). *Western Journal of Applied Forestry* 23: 40-45.

Feyereisen, R. 1999. Insect P450 enzymes. *Annual Review of Entomology* 44: 507-533.

Feyereisen, R. 2006. Evolution of insect P450. *Biochemical Society Transactions* 34: 1252-1255.

Franceschi, V.R., Krokene, P., Christiansen, E., and Krekling, T. 2005. Anatomical and chemical defences of conifer bark against bark beetles and other pests. *New Phytologist* 167: 353-376.

Franke, W., and Vité, J.P. 1983. Oxygenated terpenes in pheromone systems of bark beetles. *Zeitschrift für Angewandte Entomologie* 96: 146-156.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivyani, I., Appel, R.D., Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and alanlysis. *Nucleic Acids Research* 31: 3784-3788.

Geisse, S., Gram, H., Kleuser, B., and Kocher, H.P. 1996. Eukaryotic expression systems: A comparison. *Protein Expression and Purification* 8: 271-282.

Gilbert, L.I. 2004. Halloween genes encode P450 enzymes that mediate steroid hormone biosynthesis in *Drosophila melanogaster. Molecular and Cellular Endocrinology* 215: 1-10.

Gillam, E.M.J., and Hunter, D.J.B. 2007. Chemical defense and exploitation: Biotransformation of xenobiotics by cytochrome P450 enzymes. In: Sigel, A., Sigel, H., and Sigel, R.K.O. Metal Ions in Life Sciences, Vol III: The Ubiquitous Roles of Cytochrome P450 Proteins (Pp. 498-507). Hoboken, New Jersey: John Wiley & Sons, Ltd.

Gries, G., Smirle, M.J., Leufvén, Miller, D.R., Borden, J.H., and Whitney, H.S. 1990. Conversion of phenylalanine to toluene and 2-phenylethanol by the pine engraver *Ips pini* (Say) (Coleoptera, Scolytidae). *Experientia* 46: 329-331.

Guengerich, F.P. 2004. Human cytochrome P450 enzymes. In: Ortiz de Montellano, P.R. CYTOCHROME P450: Structure, mechanism, and biochemistry (3rd Ed.) (Pp. 443-445). New York: Kluwer Academic/Plenum Publishers.

Guengerich, F.P., Martin, M.V., Sohl, C.D., and Cheng, Q. 2009. Measurement of cytochromes P450 and NADPH-cytochrome P450 reductase. *Nature Protocols* 4: 1245-1251.

Guzov, V.M., Houston, H.L., Murataliev, M.B., Walker, F.A., and Feyereisen, R. 1996. Molecular cloning, overexpression in *Escherichia coli,* structural and functional characterization of house fly cytochrome b_5. *The Journal of Biological Chemistry* 271: 26637-26645.

Guzov, V.M, Unnithan, G.C., Chernogolov, A.A., and Feyereisen, R. 1998. CYP12A1, a mitochondrial cytochrome P450 from the house fly. *Archives of Biochemistry and Biophysics* 359: 231-240.

Hall, G.M., Tittiger, C., Blomquist, G.J., Andrews, G.L., Mastick, G.S., Barkawi, L.S., Bengoa, C., and Seybold, S.J. 2002. Male Jeffrey pine beetle, *Dendroctonus jeffreyi,* synthesizes the pheromone component frontalin in anterior midgut tissue. *Insect Biochemistry and Molecular Biology* 32: 1525-1532.

Hannemann, F., Bichet, A., Ewen, K.M., and Bernhardt, R. 2007. Cytochrome P450 systems-biological variations of electron transport chains. *Biochemica et Biophysica Acta* 1770: 330-344.

Hartfelder, K. 2000. Insect juvenile hormone: from "status quo" to high society. *Brazilian Journal of Medical and Biological Research* 33: 157-177.

Nakai, K., and Horton, P. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends in Biochemical Sciences* 24: 34-35.

Huber, D.P.W., Erickson, M.L., Leutenegger, C.M., Bohlmann, J., and Seybold, S.J. 2007. Isolation and extreme sex-specific expression of cytochrome P450 genes in the bark beetle, *Ips paraconfusus,* following feeding on the phloem of host ponderosa pine, *Pinus ponderosa. Insect Molecular Biology* 16: 335-349.

Huber, D.P.W., Aukema, B.H., Hodgkinson, R.S., and Lindgren, B.S. 2009. Successful colonization, reproduction, and new generation emergence in live interior hybrid spruce *Picea engelmannii* x *glauca*, by mountain pine beetle *Dendroctonus ponderosae. Agricultural and Forest Entomology* 11: 83-89.

Humphreys, J.M., and Chapple, C. 2004. Immunodetection and quantification of cytochromes P450 using epitope tagging: immunological, spectroscopic, and kinetic analysis of cinnimate 4-hydroxylase. *Journal of Immunological Methods* 292: 97-107.

Hunt, D.W.A., Borden, J.H., Pierce, H.D., Jr., Slessor, K.N., King, G.G.S., and Czyzewska, E.K. 1986. Sex-specific production of ipsdienol and myrcenol by *Dendroctonus ponderosae* (Coleoptera: Scolytidae) exposed to myrcene vapors. *Journal of Chemical Ecology* 12: 1579-1586.

Invitrogen. 2002. Growth and maintenance of insect cell lines – user manual (ver. K). Pp.1-34.

Invitrogen. 2003. Easy-DNA™ kit for genomic DNA isolation – instruction manual (ver. F). Pp. 10.

Invitrogen. 2003. Western Breeze® chromogenic western blot immunodetection kit – user manual (ver. F). Pp.1-13.

Invitrogen. 2007. iBlot™ dry blotting system for dry, electrophoresis of proteins from mini, midi, and E-PAGE™ gels – user manual. Pp. 18-31.

Invitrogen. 2009. BaculoDirect™ baculovirus expression system – user manual (ver. I). Pp.1-50.

Invitrogen (Life Technologies). 2004. CONCERT™ gel extraction systems. Part No. 999-00704 04/98A.

Johnston, W.A., Huang, W., De Voss, J.J., Hayes, M.A., and Gillam, E.M.J. 2008. Quantitative whole-cell cytochrome P450 measurement suitable for high-throughput application. *Journal of Biomolecular Screening* 13: 135-141.

Kaufmann, M.R., Aplet, G.H. , Babler, M., Baker, W.L., Bentz, B., Harrington, M., Hawkes, B.C., Stroh Huckaby, L., Jenkins, M.J., Kashian, D.M., Keane, R.E., Kulakowski, D., McHugh, C., Negron, J., Popp, J., Romme, W.H., Schoennagel, T., Shepperd, W., Smith, F.W., Kennedy Sutherland, E., Tinker, D., and Veblen, T.T. 2008. The status of our scientific understanding of lodgepole pine and mountain pine beetles – a focus on forest ecology and fire behavior. *The Nature Conservancy*, Arlington, VA. GFI technical report 2008-2.

Keeling, C.I., Blomquist, G.J., and Tittiger, C. 2004. Coordinated gene expression for pheromone biosynthesis in the pine engraver beetle, *Ips pini* (Coleoptera: Scolytidae). *Naturwissenschafen* 91: 324-328.

Keeling, C.I., Li, M., Henderson, H., Huber, D.P.W., Clark, E.L., Fraser, J.D., Ott, D.S., Liao, N., Docking, R., Chan, S., Taylor, G., Moore, R., Munro, S., Mayo, M., Jefferson, K., Lee, H.W., Leung, A., Thorne, K., Trinh, E., Matsuo, C., Chand, S., Brown-John, M., McMurtrie, H., Cruz, K., Smith, J., Holt, R., Jones, S., Marra, M., Cooke, J.E.K., and Bohlmann, J. 2009. Expressed sequence tags from the mountain pine beetle, *Dendroctonus ponderosae*. In prep.

Keenan, R.J., Freymann, D.M., Stroud, R.M, and Walter, P. 2001. The signal recognition particle. *Annual Review of Biochemistry* 70: 755-775.

Klepzig, K.D., Smalley, E.B., and Raffa, K.F. 1996. Combined chemical defenses against an insect-fungal complex. *Journal of Chemical Ecology* 22: 1367-1388.

Kost, T.A., Condreay, J.P., and Jarvis, D.L. 2005. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology* 23: 567-575.

Latta, R.G., Linhart, Y.B., Lundquist, L., and Snyder, M.A. 2000. Patterns of monoterpene variation within individual trees in ponderosa pine. *Journal of Chemical Ecology* 26: 1341-1357.

Light, D.M., and Birch, M.C. 1979. Inhibition of the attractive pheromone response in *Ips paraconfusus* by (R)-(-)-ipsdienol. *Naturwissenschaften* 66: 159-160.

Lindgren, B. S., Borden, J. H., Cushon, G. H., Chong, L. J., and Higgins, C. J. 1989. Reduction of mountain pine beetle attacks by verbenone in lodgepole pine stands in British Columbia. *Canadian Journal of Forest Research* 19: 65–68.

Lindgren, B.S., and Miller, D.R. 2002. Effect of verbenone on attraction of predatory and woodboring beetles (Coleoptera) to kairomones in lodgepole pine forests. *Physiological and Chemical Ecology* 31: 766-773.

Liu, N., and Zhang, L. 2004. *CYP4AB1*, *CYP4AB2*, and *Gp-9* gene overexpression associated with workers of the red imported fire ant, *Solenopsis invicta* Buren. *Gene* 327: 81-87.

Logan, J.A., and Powell, J.A. 2001. Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). *American Entomologist* 47: 160-172.

McCartney, A.W., Dyer, J.M., Dhanoa, P.K., Kim, P.K., Andrews, D.W., McNew, J.A., and Mullen, R.T. 2004. Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *The Plant Journal* 37: 156-173.

Maïbèche-Coisne, M., Monti-Dedieu, L, Aragon, S., and Dauphin-Villemant, C. 2000. A new cytochrome P450 from *Drosophila melanogaster*, CYPG15, expressed in the nervous system. *Biochemical and Biophysical Research Communications* 273: 1132-1137.

Maïbèche-Coisne, M., Nikonov, A.A., Ishida, Y., Jacquin-Joly, E., and Leal, W.S. 2004. Pheromone anosmia in a scarab beetle induced by *in vivo* inhibition of a pheromone-degrading enzyme. *Proceedings of the National Academy of Sciences* 101: 11459-11464.

Mao, W., Berenbaum, M.R., and Schuler, M.A. 2008. Modifications in the N-terminus of an insect cytochrome P450 enhance production of catalytically active protein in baculovirus-Sf9 cell expression systems. *Insect Biochemistry and Molecular Biology* 38: 66-75.

Mao, Y.-B, Cai, W.-J., Wang, J.-W., Hong, G.-J., Tao, X.-Y., Wang, L.-J., Huang, Y.-P., and Chen, X.-Y. 2007. Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nature Biotechnology* 25: 1307-1313.

Meisinger, C., Sommer, T., and Pfanner, N. 2000. Purification of *Saccharomyces cerevisiae* mitochondria devoid of microsomal and cytosolic contaminations. *Analytical Biochemistry* 287: 339-342.

Miller, D.R., Borden, J.H., and Slessor, K.N. 1989. Inter- and intrapopulation variation of the pheromone, ipsdienol produced by male pine engravers, *Ips pini* (Say) (Coleoptera: Scolytidae). *Journal of Chemical Ecology* 15: 233-247.

Monier, S., van Luc, P., Kreibich, G., Sabatini, D.D., and Adesnik, M. 1988. Signals for the incorporation and orientation of cytochrome P450 in the endoplasmic reticulum membrane. *The Journal of Cell Biology* 107: 457-470.

Nardi, J.B., Young, A.G., Ujhelyi, E., Tittiger, C., Lehane, M.J., and Blomquist, G.J. 2002. Specialization of midgut cells for synthesis of male isoprenoid pheromone components in two scolytid beetles, *Dendroctonus jeffreyi* and *Ips pini*. *Tissue & Cell* 34: 221-231.

Neve, E.P.A., and Ingelman-Sundberg, M. 1999. A soluble NH_2-terminally truncated catalytically active form of rat cytochrome P450 2E1 targeted to liver mitochondria. *Federation of European Biochemical Societies Letters* 460: 309-314.

Neve, E.P.A., and Ingelman-Sundberg, M. 2000. Molecular basis for the transport of cytochrome P450 2E1 to the plasma membrane. *The Journal of Biological Chemistry* 275: 17130-17135.

Neve, E.P.A., and Ingelman-Sundberg, M. 2001. Identification and characterization of a mitochondrial targeting signal in the rat cytochrome P450 2E1 (CYP2E1). *The Journal of Biological Chemistry* 276: 11317-11322.

Neve, E.P.A., and Ingelman-Sundberg, M. 2008. Intracellular transport and localization of microsomal cytochrome P450. *Analytical and Bioanalytical Chemistry* 392: 1075-1084.

Nirajan, B.G., Wilson, N.M., Jefcoate, C.R., and Avadhani, N.G. 1984. Hepatic mitochondrial cytochrome P-450 System. *The Journal of Biological Chemistry* 259: 12495-12501.

Omura, T., and Sato R. 1964. The carbon monoxide-binding pigment of liver microsomes. *The Journal of Biological Chemistry* 239: 2370-2378.

Omura, T. 1999. Forty years of cytochrome P450. *Biochemical and Biophysical Research Communications* 266: 690-698.

O'Reilly, D.R., Miller, L.K., and Lucklow, V.A. 1992. Baculovirus Expression Vectors: A Laboratory Manual (Pp. 3-10). New York: W.H. Freeman and Company.

Paetzel, M, Karla, A., Strynadka, N.C.J., and Dalbey, R.E. 2002. Signal peptidases. *Chemical Reviews* 102: 4549-4579.

Paine, T.D., Raffa, K.F., and Harrington, T.C. 1997. Interactions among scolytid bark beetles, their associated fungi, and live host conifers. *Annual Review of Entomology* 42: 179-206.

Perkin Elmer. 1993. PECSS software package for UV/VIS/NIR spectroscopy – running PECSS user documentation (Vol. 1). Pp. 4.3-4.35.

Petryk, A., Warren, J.T., Marqués, G., Jarcho, M.P., Gilbert, L.I., Kahler, J., Parvy, J., Li, Y., Dauphin-Villemant, C., and O'connor, M.B. 2003. Shade is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect moulting hormone 20-hydroecdysone. *Proceedings of the National Academy of Sciences* 100: 13773-13778.

Phillips, M.A., and Croteau, R.B. 1999. Resin-based defences in conifers. *Trends in Plant Science* 4:184-190.

Pierce, H.D, Jr., Conn, J.E., Oehlschlager, A.C., and Borden, J.H. 1987. Monoterpene metabolism in female mountain pine beetles, *Dendroctonus ponderosae* Hopkins, attacking ponderosa pine. *Journal of Chemical Ecology* 13: 1455-1480.

Possee, R.D. 1997. Baculoviruses as expression vectors. *Current Opinion in Biotechnology* 8: 569-572.

Pureswaran, D.S., Gries, R., and Borden, J.H. 2004a. Quantitative variation in monoterpenes in four species of conifers. *Biochemical Systematics and Ecology* 32: 1109-1136.

Pureswaran, D.S., Gries, R., and Borden, J.H. 2004b. Antennal responses of four species of tree-killing bark beetles (Coleoptera: Scolytidae) to volatiles collected from beetles, and their host and nonhost conifers. *Chemoecology* 14: 59-66.

Qiagen. 2004. MinElute® handbook (2nd Ed.). Pp. 1-35.

Qiagen. 2005. QIAprep® Miniprep handbook (2nd Ed.). Pp. 1-51.

Qiagen. 2006. QIAquick® Spin handbook. Pp. 1-39.

Raffa, K.F., and Berryman, A.A. 1987. Interacting selective pressures in conifer-bark beetle systems: A basis for reciprocal adaptations? *The American Naturalist* 129: 234-262.

Raffa, K.F., Hobson, K.R., LaFontaine, S., and Aukema, B.H. 2007. Can chemical communication be cryptic? Adaptations by herbivores to natural enemies exploiting prey semiochemicals. *Oecologia* 153: 1009-1019.

Reddy, G.V.P., and Guerrero, A. 2004. Interactions of insect pheromones and plant semiochemicals. *Trends in Plant Science* 9: 253-261.

Regnault-Roger, C., and Hamraoui, A. 1995. Fumigant toxic activity and reproductive inhibition induced by monoterpenes on *Acanthoscelides obectus* (Say) (Coleoptera), a bruchid of kidney bean (*Phaseolus vulgaris* L.). *Journal of Stored Products Research* 31: 291-299.

Renwick, J.A., Hughes, P.R., and Krull, I.S. 1976. Selective production of cis- and trans-verbenol from (-) and (+) alpha-pinene by a bark beetle. *Science* 191: 199-201.

Robin, M., Anandatheerthavarada, H.K., Fang, J., Cudic, M., Otvos, L., and Avadhani, N.G. 2001. Mitochondrial targeted cytochrome P450 2E1 (P450 MT5) contains an intact N terminus and requires mitochondrial specific electron transfer protein for activity. The *Journal of Biological Chemistry* 276: 24680-24689.

Roy, S.R. 1998. Genetic and molecular analysis of a cytochrome P450 based pyrethroid resistance in the red flour beetle, *Tribolium castaneum* (Herbst). Doctoral Dissertation. Purdue University, West Lafayette, IN.

Sandstrom, P., Welch, W.H., Blomquist, G.J., and Tittiger, C. 2006. Functional expression of a bark beetle cytochrome P450 that hydroxylates myrcene to ipsdienol. *Insect Biochemistry and Molecular Biology* 36: 835-845.

Sandstrom, P, Ginzel, M.D., Bearfield, J.C., Welch, W.H., Blomquist, G.J., and Tittiger, C. 2008. Myrcene hydroxylases do not determine enantiomeric composition of pheromonal ipsdienol in *Ips* spp. *Journal of Chemical Ecology* 34: 1584-1592.

Sato, T., Sakaguchi, M., Mihara, K., and Omura, T. 1990. The amino-terminal structures that determine topological orientation of cytochrome P-450 in microsomal membrane. *The European Molecular Biology Organization Journal* 9: 2391-2397.

Scott, J.G. 2008. Insect cytochrome P450s: Thinking beyond detoxification. In: Recent Advances in Insect Physiology, Toxicology and Molecular Biology, N. Liu, ed. Research Signpost, Kerala, India: 117-124.

Seybold, S.J., Quilici, D.R., Tillman, J.A., Vanderwel, D., Wood, D.L., and Blomquist, G.J. 1995. *De novo* biosynthesis of the aggregation pheromone components ipsenol and ipsdienol by the pine bark beetles *Ips paraconfusus* Lanier and *Ips pini* (Say) (Coleoptera: Scolytidae). *Proceedings of the National Academy of Sciences USA* 92: 8393-8397.

Seybold, S.J., Huber, D.P.W., Lee, J.C., Graves, A.D., and Bohlmann, J. 2006. Pine monoterpenes and pine bark beetles: A marriage of convenience for defense and chemical communication. *Phytochemistry Reviews* 5: 143-178.

Small, I., Peeters, N., Legeai, F., and Lurin, C. 2004. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4: 1581-1590.

Stark, R.W. 1982. Chapter 2: Generalized ecology and life cycle of bark beetles. In: Mitton, J.B., Sturgeon, K.B., Bark Beetles in North American Conifers – A System for the Study of Evolutionary Biology (1st ed.) (Pp. 21-45). Austin, Texas: University of Texas Press.

Sun, W, Margam, V.M., Sun, L., Buczkowski, G., Bennett, G.W., Schemerhorn, B., Muir, W.M., and Pittendrigh, B.R. 2006. Genome-wide analysis of phenobarbital-inducible genes in *Drosophila melanogaster*. *Insect Molecular Biology* 15: 455-464.

Sutherland, T.D., Unnithan, G.C., Andersen, J.F., Evans, P.H., Murataliev, M.B., Szabo, L.Z., Mash, E.A., Bowers, W.S., and Feyereisen, R. 1998. A cytochrome P450 terpenoid hydroxylase linked to the suppression of insect juvenile hormone synthesis. *Proceedings of the National Academy of Sciences* 95: 12884-12889.

Szczesna-Skorupa, E., Browne, N., Mead, D., and Kemper, B. 1988. Positive charges at the NH_2 terminus convert the membrane-anchor signal peptide of cytochrome P-450 to a secretory signal peptide. *Proceedings of the National Academy of Sciences USA* 85: 738-742.

Szczesna-Skorupa, E., Ahn, K., Chen, C.-D., Doray, B., and Kemper, B. 1995. The cytoplasmic and N-terminal transmembrane domains of cytochrome P450 contain independent signals for retention in the endoplasmic reticulum. *The Journal of Biological Chemistry* 270: 24327-24333.

Szczesna-Skorupa, E., Chen, C., Rogers, S., and Kemper, B. 1998. Mobility of cytochrome P450 in the endoplasmic reticulum membrane. *Proceedings of the National Academy of Sciences* 95: 14793-14798.

Teale, S.A., Hager, B.J., and Webster, F.X. 1994. Pheromone-based assortive mating in a bark beetle. *Animal Behaviour* 48: 569-578.

Thomas, J.L., Evans, B.W., Blanco, G., Mason, J.I., and Strickler, R.C. 1999. Creation of a fully active, cytosolic form of human type I3β-hydroxysteroid dehydrogenase/isomerase by the deletion of a membrane-spanning domain. *Journal of Molecular Endocrinology* 23: 231-239.

Thoss, V., and Byers, J.A. 2006. Monoterpene chemodiversity of ponderosa pine in relation to herbivory and bark beetle colonization. *Chemoecology* 16: 51-58.

Tibieri, R., Niccoli, A., Curini, M., Epifano, F., Marcotullio, M.C., and Rosati, O. 1999. The role of the monoterpene composition in *Pinus* spp. needles, in host selection by the pine processionary caterpillar, *Thaumetopoea pityocampa. Phytoparasitica* 27: 263-272.

Tillman, J.A., Holbrook, G.L., Dallara, P.L., Schal, C., Wood, D.L., Blomquist, G.J., and Seybold, S.J. 1998. Endocrine regulation of de novo aggregation pheromone biosynthesis in the pine engraver, *Ips pini* (Say) (Coleoptera: Scolytidae). *Insect Biochemistry and Molecular Biology* 28: 705-715.

Tittiger, C., Blomquist, G.J., Ivarsson, P., Borgeson, C.E., and Seybold, S.J. 1999. Juvenile hormone regulation of HMG-R gene expression in the bark beetle *Ips paraconfusus* (Coleoptera: Scolytidae): implications for male aggregation pheromone biosynthesis. *Cellular Molecular Life Sciences* 55: 121-127.

Trapp, S., and Croteau, R. 2001. Defensive resin biosynthesis in conifers. *Annual Review of Plant Physiology and Plant Molecular Biology* 52:689-724.

Unnithan, G.C., and Nair, K.K. 1977. Ultrastructure of juvenile hormone-induced degenerating flight muscles in a bark beetle, *Ips paraconfusus. Cell and Tissue Research* 185: 481-490.

von Heijne, G. 1986. Mitochondrial targeting sequences may form amphiphilic helices. *The European Molecular Biology Organization Journal* 5: 1335-1342.

Wainhouse, D., Cross, D.J., and Howell, R.S. 1990. The role of lignin as a defence against the spruce bark beetle *Dendroctonus micans*: effect on larvae and adults. *Oecologia* 85: 257-265.

Wen, Z., Pan, L., Berenbaum, M.R., and Schuler, M.A. 2003. Metabolism of linear and angular furanocoumarins by *Papilio polyxenes* CYP6B1 co-expressed with NADPH cytochrome P450 reductase. *Insect Biochemistry and Molecular Biology* 33: 937-947.

Werck-Reichhart, D., and Feyereisen, R. 2000. Cytochromes P450: A success story. *Genome Biology* 1: reviews 3003.1-3003.9.

Westfall, J., and Ebata, T. 2008. 2007 Summary of Forest Health Conditions in British Columbia. British Columbia Forest Service: Forest Health – British Columbia, Pest Management Report Number 15: 1-81.

Winkel, B.S.J. 2004. Metabolic channeling in plants. *Annual Review of Plant Biology* 55: 85-107.

Wood, D.L., and Storer, A.J. 2002. Bark beetles infesting California's conifers. *Fremontia* 30: 19-25.

Yamamoto, K., Ichinose, H., Aso, Y., and Fujii, H. 2010. Expression analysis of cytochrome P450s in the silkmoth, *Bombyx mori*. *Pesticide Biochemistry and Physiology* doi:10.1016/i.pestbp.2009.11.06.

Yang, P., Tanaka, H., Kuwano, E., and Suzuki, K. 2008. A novel cytochrome P450 gene (*CYP4G25*) of the silkmoth *Antheraea yamamai*: Cloning and expression pattern in pharate first instar larvae in relation to diapauses. *Journal of Insect Physiology* 54: 636-643.

Zhang, J., Yvan, P., and Goyer, C. 2008. Identification of potential detoxification enzyme genes in *Leptinotarsa decemlineata* (Say) and study of their expression in insects reared on different plants. *Canadian Journal of Plant Science* 88: 621-629.