Characterization and heterologous expression of *cis*-abienol synthase from balsam fir (*Abies balsamea*) towards high-end fragrance components

by

Baillie Redfern

B.Sc. (Honors), University of Ottawa, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2014

Abstract

Background

The *cis*-abienol synthase (*Ab*CAS) gene of balsam fir (*Abies balsamea*) encodes a diterpene synthase (diTPS), which produces a compound, *cis*-abienol, of value for the fragrance industry. The same compound is also produced in *Abies lasiocarpa. cis*-Abienol can be used as a starting material for semisynthesis of the commercial substance Ambrox® which is used in the formulation of high end perfumes. *Ab*CAS is a bifuncational diTPS which produces *cis*-abienol in a two-step process involving two active sites responsible for the class I and class II activities. Class-II activity generates the hydroxylated labda-13-en-8-ol diphosphate (LPP) intermediate, followed by the class-I activity which produces the *cis*-abienol product.

Results

We investigated the effects that amino acid residue changes have on catalytic activity of *Ab*CAS and related diTPSs and the product profiles of these protein variants. We identified critical amino acid residues (D348H and S345S) that are responsible for product specificity in *Ab*CAS. In addition, in order to determine if the mutations in *Ab*CAS were also relevant to other diTPSs, this study extended investigation into the *Ab*LAS enzyme, a multiproduct levopimaradiene/abietadiene synthase. The *Ab*LAS:C345S:H346D protein variant showed a new activity and produced the oxygenated diterpene manoyl oxide (100% product profile), where as *Ab*LAS (WT) produces exclusively olefins.

To develop a metabolic engineering platform for *cis*-abienol production we investigated yeast *(Saccharomyces cerevisiae)* and in the bryophyte *Physomitrella patens* as hosts for

expression of *Ab*CAS. We examined expression of *Ab*CAS with three different *S. cerevisiae* strains (Am94, BY4741 and KE Strain) and successfully produced *cis*-abienol in all three strains with up to 12.8 mg x L^{-1} in the AM94 strain. *Physcomitrella* transformed with *Ab*CAS did not produce detectable levels of *cis*-abienol.

Conclusions

Product specificity of the diTPS enzymes *Ab*CAS and *Ab*LAS can be altered by changing as few as one or two amino acid residues. Expression of *Ab*CAS in yeast resulted in the formation of *cis*-abienol indicating opportunities for metabolic engineering of a recombinant production system for this valuable diterpene compound.

Preface

This dissertation is original, unpublished work by Baillie Redfern. I conceived and designed the experiments with the help of Dr. Philipp Zerbe and Dr. Jörg Bohlmann. I performed the experiments and analyzed the data for the entirety of this project with the exception of data analysis for the *Abies lasiocarpa* experiment (performed by Dr Philipp Zerbe) as I had a personal family emergency. *Cis*-abienol synthase homology modeling was done prior to this project and is published elsewhere (Zerbe 2012) .

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DMAPP: Dimethylallyl diphosphate
diTPS: Diterpene synthase
GA: Gibberellins
ER: Endoplasmic reticulum
FPP: Farnesyl diphosphate
GPP: Geranyl diphosphate
GGPP: Geranylgeranyl diphosphate
IPP: Isopentenyl diphosphate
MVA: Mevalonic acid
NPP: Neryl diphosphate
LPP: Labda-13 en-8-ol diphosphate
SDM: Site-directed mutagenesis
TPS: Terpene synthase

Acknowledgments

This project would not have been possible without the assistance of many outstanding people. First and foremost I would like to thank Dr. Jörg Bohlmann and Karen Reid for their patience, support, guidance and expertise specifically in the field of plant genomics. Jörg and Karen have been great mentors, both academically and outside of the lab. Jörg took a chance on a new student with limited experience, guided me and supported me financially through research assistantships for which I am truly grateful. Karen and Angela Chiang, our lab managers, thank you for your organizational prowess and commitment to ensuring our lab affairs run smoothly.

I would like to extend my gratitude to every member of the Bohlmann Lab. Their assistance and friendship has been indispensable. Thank you all for providing me for an opportunity to work here and be apart of your team, the group interview was delightful. It was a pleasure to work with all of you, and an unforgettable experience to be carried to other life stages. Special thanks to Kate Storey and Chris Roach for being wonderful friends, best of luck with your own writing in the near future.

To my project advisor Dr. Philipp Zerbe and my committee members Drs. Phil Hieter and Nobuhiko Tokuriki, I would like to extend the sincerest thank you for their advice for my project and corrections to my final thesis. Dr. P. Zerbe was an essential part to this project, for it is his discovery of *Ab*CAS that my work is based upon. Without your vast knowledge and kind support this project would have never come to pass.

As the first in my family to attend post-secondary and complete graduate school I hope this encourages my younger cousins (Kyle Baillie, Steven Maxim, Hermione Clarke, Tommy Baillie, Pheobe Clarke, Hedi Baillie, Robbie Willis, Kaylee Willis, Quinn Coker, Jack Weir and Dane Redfern among others) to choose the road less traveled, wherever that road may lead you.

Dedication

To my brother and sister, Austin Redfern and Mindy Willis

Chapter 1: Introduction and Research Objectives

Plant Secondary Metabolism and Diterpenoids

Plants require only carbon dioxide, minerals, water, and sunlight to synthesize all of their metabolites. Metabolites that are necessary for their growth and development are often referred to as 'general' (i.e. primary) metabolites. These compounds are found ubiquitously in almost all plant species. However, the majority of compounds found in plants can be classified as 'specialized' (i.e., secondary). These specialized metabolites are the most diverse chemicals of the plant kingdom. Specialized metabolites have functions in interaction of plants with their biotic environments, such as protection against pests and pathogens, referred to as chemical defense in the form of repellents, anti-feedants, toxins, ovipositor deterrents, or anti-biotic's. For example in Sitka spruce (*Picea sitchensis*), the accumulation of the diterpene resin acid dehydroabietic acid is associated with resistance against spruce shoot weevil (*Pissodes strobi*) (Robert *et al*. 2010).

Ten of thousands of different specialized metabolites have been found in the plant kingdom, and many of them are of narrow taxonomic distribution in a single species, genus or family. Humans have found the chemical diversity of specialized metabolism to be of large value as pharmaceuticals and other industrial products such as flavours, fragrances, food supplements, and pesticides. The three largest groups of plant specialized metabolites are the phenolics, alkaloids and terpenoids. Most plant phenolics are derived from the amino acid phenylalanine (Taiz and Zeiger 2006). Most alkaloids are derived from various different amino acids, and contain nitrogen. Terpenoids (also known

as terpenes or isoprenoids) are made of 5-carbon building blocks produced through the mevalonic acid (MVA) pathway or the methylerythritol phosphate (MEP) pathway.

Gibberellins (GA) are a group of tetracyclic diterpenoids and an example of a primary metabolite. The bioactive GAs are important as phytohormones involved in the reproduction of plant growth and development, including seed germination, stem elongation, leaf expansion, trichome development, and flower development (Olszewski et al. 2002). This introduction chapter will focus on ecologically and economically important diterpenoids that occur as natural defense compounds in the oleoresin (pitch) of coniferous trees, such as species of pine (*Pinus*), fir (*Abies*) and spruce (*Picea*) (Zulak et al. 2010, Keeling and Bohlmann 2006). The terpenoid-containing oleoresin accumulates in large amounts in the tree's wood, bark, and needles as a chemical defense system. Conifers respond to invading insect pest and pathogens by secreting this resin at the wound site. Complex blends of mono-, di- and to lesser extent sesqui-terpenes are the primary components of oleoresin which is toxic to many potential insect pests and pathogens (Kolosova and Bohlmann 2012). The volatile mono-terpenoids evaporate from the pitch which results in the remaining diterpenes forming a long-lasting chemical and physical barrier at the site of insect attack. Oleoresin is also of economic value to humans because of its chemical properties and associated uses, such as the production of coatings, varnishes, constituents of perfume and food glazing agents (Bohlmann and Keeling 2008, Gershenzon and Dudareva 2007). Historically the oleoresin was used extensively in the naval stores industry.

Diterpene Biosynthesis and Diterpene Synthases

In plants, the MEP pathaway and the MVA pathway produce the isomeric 5-carbon compounds isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (Lichtenthaler 1999). A head to tail condensation of one IPP molecule to one DMAPP molecule gives rise to geranyl diphosphate (GPP) the precursor to 10-carbon monoterpenes. Adding another IPP will give rise to farnesyl diphosphate (FPP) the precursor to 15-carbon sesquiterpenes and another IPP will give geranylgeranyl diphosphate (GGPP) the precursor to 20-carbon diterpenes. GPP, FPP and GGPP are substrates of monoterpene synthases (monoTPS), sesquiterpene synthases (sesquiTPS) and diterpene synthases (diTPS), respectively (Chen *et al*. 2011).

Often, plant genomes contain sets of related genes (i.e. gene families) that encode enzymes that use similar substrates and give similar products, but have members that have diverged in different lineages. The TPS family in plants is such a multi-gene family (Bohlmann *et al.* 1998, Chen *et al.* 2011). TPS form a wide range of structurally diverse cyclic and acyclic mono-, sesqui- and di-terpenes. The monophyletic plant TPS family can be divided into subfamilies (TPS-a through TPS-g) based on sequence identify and functional relationships (Bohlmann *et al*. 1998; Chen *et al.* 2011). The TPS of gymnosperm secondary metabolites are found specifically in the TPS-d subfamily (Keeling *et al.* 2011, Martin *et al.* 2004, Zerbe and Bohlmann 2014). Apparently, these TPS-d enzymes arose from an ancestral di-TSP gene through repeated gene duplication events. The proposed ancestral diTPS gene was likely involved in the cyclization of GGDP to produce *ent*-kaurene, a precursor to the GA phytohormones. In the evolution of the gymnosperms lineage of plants, this ancestral diTPS is thought to have undergone a

series of gene duplications, followed by sub- or neo- functionalization to give rise to many TPS-d genes found in extant conifer species (Zerbe and Bohlmann, 2014).

My thesis deals with enzymatic activity of select bifunctional class I/II diTPSs of the TPS-d subfamily. Bifunctional refers to the enzyme having two functional active sites, class II and class I, respectively. The class II activity resides in the **β** and **γ** domain and contains, in its active form, a 'DXDD' motif for the protonation-initiated cyclization of GGPP to CPP. The class I active site, resides in the C-terminal domain and includes the 'DDXXD' and 'NSE/DTE' motifs and **α** domain for metal-dependent ionization of the prenyl diphosphate substrate. Bifunctional diTPSs of the TPS-d subfamiliy catalyze the cyclization of linear GGPP precursors to form the complex hydrocarbon skeletons of all the cyclic diterpenoids discussed in this paper. In the biosynthesis of conifer diterpenoids resin acids, there are P450-mediated pathway steps that subsequently follow diTPS activity, and decorate the terpene molecule with oxygen functionality. The P450 CYP720B4 from Sitka Spruce (*Picea sitchemsis*) is one such example that catalyzes multiple oxidations in the biosynthesis of resin acids (Hamberger *et al.* 2011).

The Diterpene Alcohol *cis*-Abienol and *cis*-Abienol Synthase

The diterpene alcohol *cis*-abienol is found in balsam fir (*Abies balsamea*) with the highest abundance of *cis*-abienol in the bark (up to 40% of the total oleoresin that accumulates in resin ducts and blisters) (Zerbe *et al*. 2012). The closely related conifer tree, *Abies lasiocarpa,* also produces *cis*-abienol (appendix 1). In addition to these two conifer (i.e. gymnosperm) species, *cis*-abienol has been reported in a few angiosperms specifically

tobacco (*Nicotiana tabacum*; family *Solanaceae)* in the trichomes (Guo et al. 1994, Guo and Wagner 1995) and the tuberous roots of Bolivian sunroot (*Polymnia sonchifolia*; family *Asteraceae*) (Miyazawa *et al.* 2008).

The diTPS genes responsible for *cis*-abienol formation have been cloned as cDNAs from balsam fir and tobacco and the encoded enzymes characterized (Sallaud *et al*. 2012, Zerbe *et al*. 2012). The balsam fir *cis*-abienol synthases (*Ab*CAS) is a bifunctional class I/II diTPS enzyme (Zerbe *et al*. 2012). In contrast, formation of *cis*-abienol in tobacco requires a pair of monofunctional class I and class II diTPSs (Sallaud *et al*. 2012). The bifunctional structure the *Ab*CAS may provide advantages for engineering *cis*-abienol formation in a recombinant host as discussed in Zerbe *et al*. (2012).

Compared to the biosynthesis of other conifer diterpenes, the *cis*-abienol pathway is unique with regards to the reaction mechanism of the *Ab*CAS diTPS enzyme (Zerbe *et al*. 2012; Figure 1). Activity of *Ab*CAS begins with the protonation of GGPP in the class-II active site and proceeds through water capture at C-8 carbocation to form the hydroxylated intermediate labda-13-en-8-ol diphosphate (LPP) (Figure 2). LPP can freely diffuse to the class-I active site where it is converted via ionization of the allylic diphosphate group to give rise to *cis*-abienol. In contrast, the diTPS LAS enzyme that produces epimeric forms of 13-hydroxy-8(14)-abietene (Keeling *et al*. 2011) starts with the protonation of GGPP at the class-II active site, resulting in copalyl diphosphate (CPP) of (9S,10S)-configuration (i.e. normal or (+)-configuration). Following translocation to active

site-I, CPP undergoes a secondary ionization dependent cyclization and enzyme specific rearrangement of intermediate carbocations.

Applications of *cis*-Abienol

Cis-abienol and other oxygen-containing diterpenoids of plant origin (e.g. sclarerol and manool) are used as precursors for synthesizing ambroxides (Ambrox® is trademarked by Firmenich, Geneva) for applications as fixatives in high-end perfumes. Ambroxides are a sustainable replacement for the use of ambergris (grey amber), an animal product secreted from the intestines of sperm whales, which are listed as endangered species (Barrero *et al*. 1996). *Salvia sclarea* (clary sage) is cultivated and harvested for the commercial production of labdanoid diterpenediol sclareol, which is used in the industry for the synthesis of Ambrox ® (Schmiderer *et al.* 2008). However, there are limitations when using sage and fir trees to produce sclareol and *cis*-abienol, respectively. For instance, there are issues with the purity of the compound of interested that can be isolated and extracted from the plant species. For example, in the fir tree there is a mixture of terpenoids, of which cis-abienol accounts for 40%. In addition, we investigated microbial hosts for production of *cis*-abienol because of the ability to increase production to industrial scale.

Heterologous Hosts for *cis*-Abienol Production

There are challenges associated with the supply of *cis*-abienol from conifers because the compound of interest (*cis*-abienol) is one among many secondary metabolites and therefore purity of the extraction is a concern. As a substitute for harvesting *cis*-abienol

from its natural source, metabolic engineering of *cis*-abienol production in a heterologous host may provide an alternative. Two host systems were investigated in this thesis, yeast (*Saccharomyces cerevisiae*) and the moss *Physcomitrella patens*. Each host offers potential unique advantages for terpenoid pathway engineering toward the possible production of *cis*-abienol.

The bryophyte *Physcomitrella patens* (velvet mud moss) is a non-vascular plant that is being explored as a system for production of plant natural products such as pharmaceuticals and other industrially interesting compounds (Prigge and Bezanilla 2010, Decker and Reski 2008). The moss contains only two endogenous diterpenes which are non-essential. The diTPSs responsible for these diterpenes can be knocked out, therefore leaving an essentially clear background devoid of endogenous diterpenes (Hayashi *et al.* 2010). *P. patens* can be propagated on solid or liquid minimal media (without complex nutrients) on petri dishes and as suspension cultures, respectively. *P. patens* is potentially advantageous to yeast because as a plant it offers a more native environment for the expression of *cis*-abienol. For example the codon usage, protein targeting and post-translational modifications that occurs in *P. patens* will be more similar to the native *A. balsamea* plant environment than that of *S. cerevisiae.*

The eukaryotic model organism yeast (*S. cerevisiae*) is a well established system for metabolic engineering of plant chemicals. Yeast has been successfully used for the biosynthesis of several other diterpenes such as tanshinone and taxol (Kai *et al.* 2011,

Engels *et al*. 2008). Compared to *E.coli*, many of the cellular pathways in yeast share a high degree of similarity to those of plant cells.

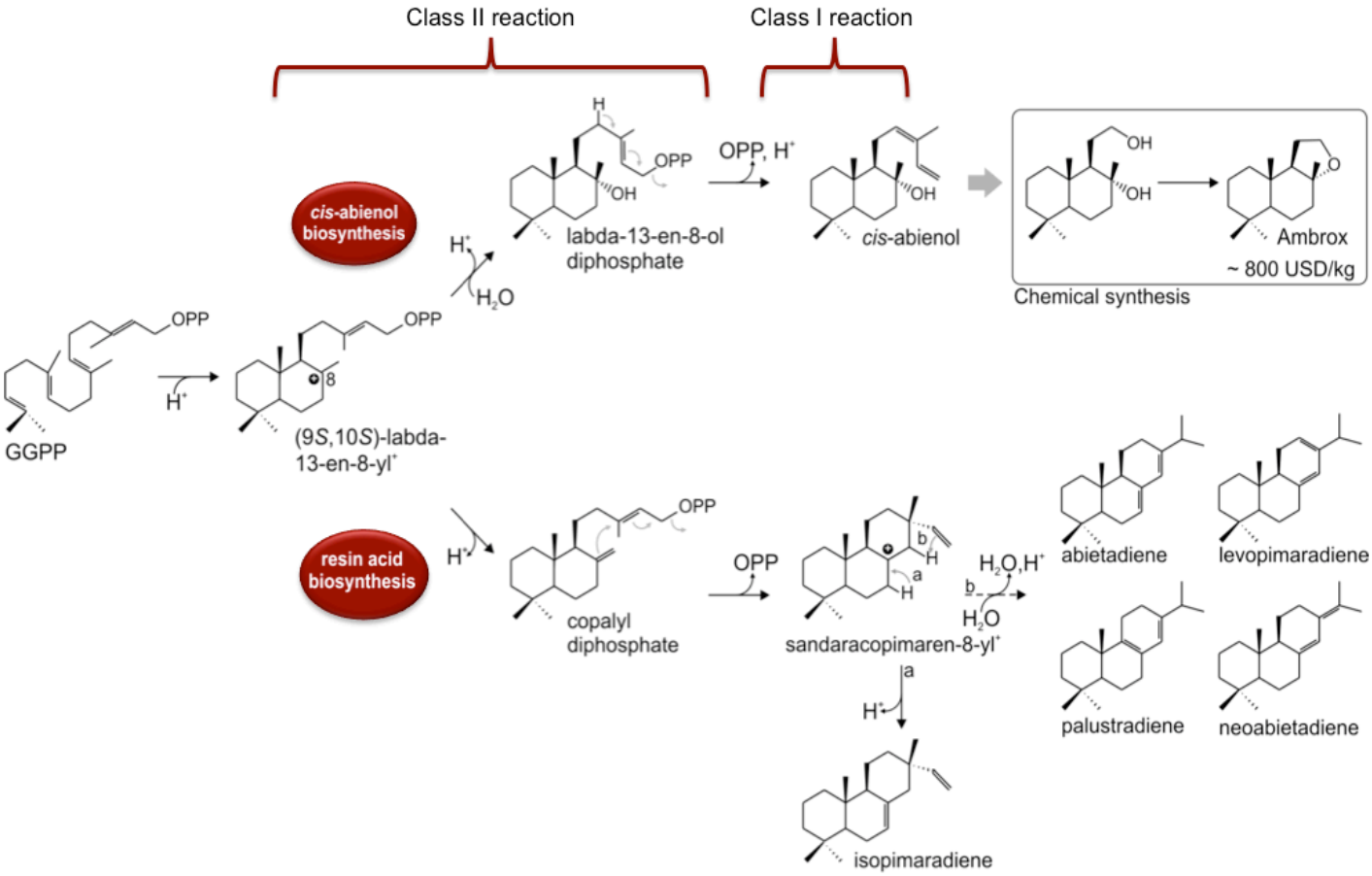


Figure 1. **Biosynthesis of specialized diterpenes in conifers.** Bifunctional diterpene synthases (diTPS) control the first steps in the formation of oleoresin diterpenes. The reaction is initiated by the protonation-dependent cycloisomerization of GGPP in the enzyme's class II active site, followed by cleavage of the diphosphate group and rearrangement of the resulting carbocations in the class I active site. While the common diterpene resin acid precursors are formed via a bicyclic copalyl diphosphate (CPP) intermediate and subsequent secondary cyclization, *cis*-abienol is formed by a bifunctional diTPS (*Ab*CAS) that produces a hydroxylated diphosphate intermediate, labda-13-en-8-ol diphosphate (LPP), and transforms it directly into *cis*-abienol without any further cyclization. This figure was revised from Zerbe *et al* 2012.

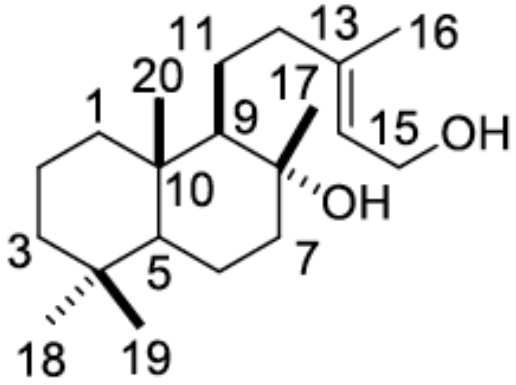


Figure 2. Carbon numbering of LPP. The biosynthesis of *cis*-abienol proceeds through the hydroxylated intermediate labda-13-en-8-ol diphosphate (LPP) via water capture at C-8 of the labda-13-en-8-yl^{+} carbocation and subsequent ionization of the allylic diphosphate group without further cycloisomerization.

Research Goals and Objectives

We can use the mechanistic information about CAS and LAS in conjunction with sequence comparison between *Ab*LAS and *Ab*CAS to predict and test unique residues that may be important in product and reaction specificity. Previous work performed structure modeling and molecular docking of GGPP and LPP in the *Ab*CAS enzyme in the class II and class I active sites, respectively (Zerbe *et al.* 2012) . Work on *Ab*CAS in my thesis was based on this previous computer-based modeling that suggested several important residues important for functionality. My work investigated functions of these specific residues by performing characterization of the mutated *Ab*LAS and *Ab*CAS proteins. These experiments were conducted in order to determine the catalytic activity and record observed changes in product profile. As my thesis was in progress the following new information regarding three solved diTPS structures was published, (1) the class-I diTPS of taxadiene synthase from *Taxus brevifolia* (TbTS), (2) the class-II diTPS copalyl diphosphate synthase from *Arabidopsis thaliana* (AtCPS), and (3) the bifuncation (class I/II) abietadiene synthase from *Abies grandis* (AgAS) (Koksal *et al.* 2010, Koksal 2011, Zhou *et al*. 2012).

In summary, given the above information the goal of my thesis was to better understand the metabolic pathways that support the expression of the *Ab*CAS enzyme for the production of *cis*-abienol and to explore *Ab*CAS' functionality through a series of site-directed mutagenesis experiments. The specific objectives of my thesis are;

(1) What are the critical amino acid residues responsible for product specificity that occur in *Ab*CAS?

(2) How can the pathway in *Saccharomyces cerevisiae* be bioengineered and manipulated for production of *cis*-abienol.

(3) Can *Ab*CAS be transformed and expressed in *Physomitrella patens* for the production of *cis*-abienol.

Chapter 2. Structural and functional analysis of balsam fir (*Abies balsema*) *cis*-abienol synthase reveals active site residues that control product specificity

2.1 Introduction

The TPS-d subfamily of gymnosperm terpene synthases

The TPS-d subfamily is specific to gymnosperms. Members of the TPS-d subfamily contributes to the chemical complexity and diversity of conifer specialized terpenoid metabolism (Keeling and Bohlmann 2006, Keeling *et al.* 2011, Martin *et al.* 2004). The functional diversity of mono-, sesqui- and diTPS within the TPS-d subfamily is thought to be the evolutionary result of repeated gene duplications followed by sub- and neo-functionalization (Chen *et al.* 2011). The CAS enzyme of balsam fir is functionally unique compared to other known TPS-d enzymes. CAS is a bifunctional class I/II diTPS producing *cis*-abienol (Figure 3). All previously characterized bifunctional diTPSs involved in conifer specialized metabolism produce either isopimaradiene and minor amounts of sandaracopimaradiene (Iso-type diTPS) or epimeric forms of 13-hydroxy-8(14)-abietene (LAS-type diTPS) (Vogel *et al.* 1996, Peters *et al.* 2000, Peters *et al.* 2001, Martin *et al.* 2004, Ro and Bohlmann 2006, Keeling *et al.* 2008, Keeling *et al.* 2011). The diTPSs LAS, ISO and CAS have a conserved structure comprised of three domains, the β and γ-domains forming the class-II active site and the α-domain forming the class-I active site (Figure 4) (Zerbe *et al.* 2012). Previous work with LAS- and ISO-type diTPS showed that changes of only a few amino acids in the active sites can alter their product specificity (Zerbe 2011, Keeling 2008, Leonard 2010, Ghao 2010).

Reaction Mechanism of ISO- and LAS-type diTPS compared to CAS

Despite their different product profiles, all conifer diTPS use a similar reaction mechanism that can be divided into two major steps: first, in the class-II reaction, protonation-dependent cyclo-isomerization of GGPP affords an intermediate carbocation (Figure1). Then, in in the class-I reaction, the carbocation is neutralized through deprotonation or aqueous quenching to form a bicyclic prenyl diphosphate (Wendt and Schulz 1998). More details are discussed in the paragraphs below.

The reaction mechanisms of ISO- and LAS-type diTPSs are well characterized (Keeling 2008, Martin 2004, Keeling 2011, Peters 2001). They begin with the protonation-initiated transformation of GGPP into copalyl diphosphate (CPP) of (9*S*,10*S*)-configuration (i.e. normal or (+)-configuration) at the class-II active site. CPP can freely diffuse to the class-I active site, where it undergoes a secondary ionization-dependent cyclization through cleavage of the diphosphate group and further rearrangement reactions of the resulting carbocation. The LAS-type diTPSs catalyze rearrangement and water capture at C-13, resulting in 13-hydroxy-8(14)-abietene as the initial product (Keeling *et al.* 2011). In contrast to all known ISO and LAS diTPSs that form (+)-CPP as an intermediate, the biosynthesis of *cis*-abienol proceeds through the hydroxylated intermediate labda-13-en-8-ol diphosphate (LPP) via water capture at C-8 of the labda-13-en-8-yl$^+$ carbocation and subsequent ionization of the allylic diphosphate group without further cycloisomerization (Zerbe *et al.* 2012) (Figure 1). This reaction results in the bicyclic *cis*-abienol, in contrast to the tricyclic diterpene products of the ISO- and LAS diTPSs.

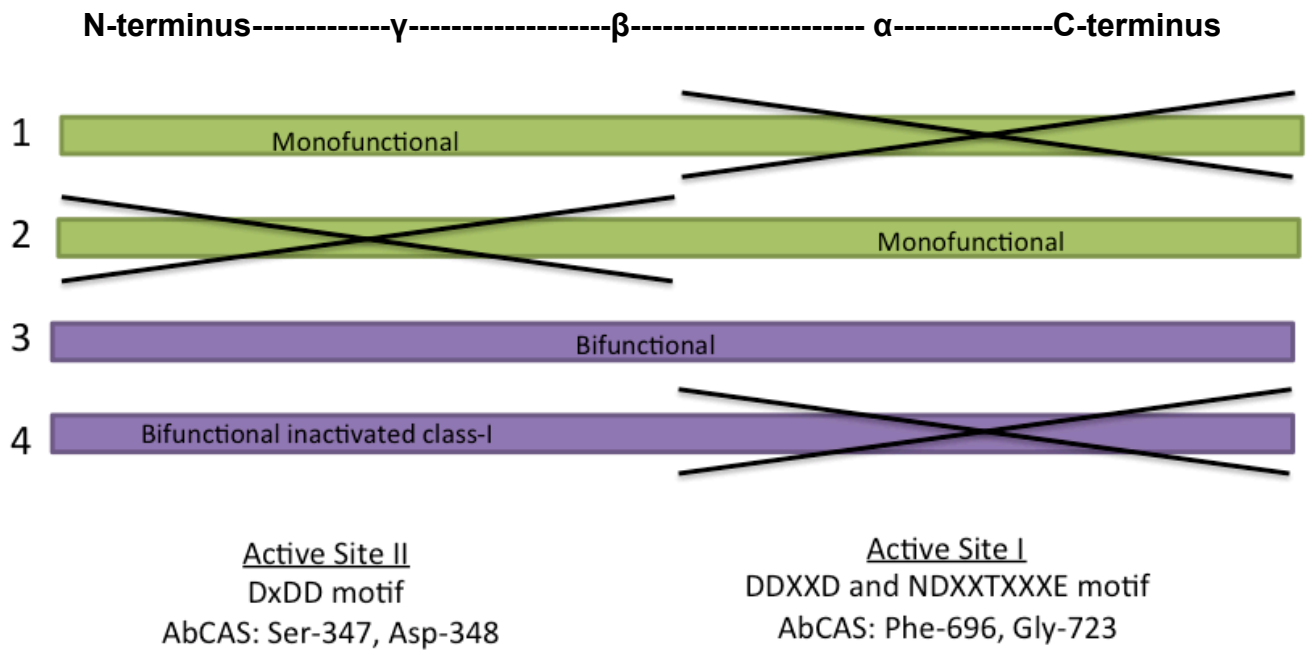


Figure 3. Schematic representation of monofunctional and bifunctional (e.g. *Ab*CAS) terpene synthases conserved active site motifs and location of predicted important residues. The class I active site resides in the α domain of the protein and includes the 'DDXXD' and 'NSE/DTE' motifs for metal-dependent ionization of the prenyl diphosphate substrate. The class II activity resides in the β and γ domain and contains, in its active form, a 'DXDD' motif for the protonation-initiated cyclization of GGPP to CPP. In angiosperms, all diterpene synthases that have been characterized to date are monofunctional, with loss of activity in one of the active sites (indicated by the x). In contrast, gymnosperm diTPSs have both active class-II and class-I active sited and are referred to as bifuncational. (1) Monofunctional enzyme with a nonfunctional class-I active site. (2) Monofunctional enzyme with a nonfunctional class-II active site. (3) Bifunctional enzyme. (4) Bifunctional enzyme with an inactivated class-I active site by mutagenesis.

Mutational analysis of *Ab*CAS and *Ab*LAS

Class-II and class-I catalysis are controlled by characteristic functional motifs (Wendt and Schulz 1998). The class-II domain harbors a DxDD motif, where in the central aspartic acid residue acts as the catalytic acid (Prisic *et al*. 2007). In contrast, the class I active site is characterized by a DDxxD and a N(N/D)xx(S/T/Q)xxxE motif for binding and cleaving of the diphosphate linkage. Mutagenesis of these catalytic motifs can be used to obtain protein variants that contain non-functional class-II or class-I domains, respectively. Using this approach, we prepared the variants *Ab*CAS(D621A) and *Ab*LAS(D611A) which render the class-I active site non-functional and therefore are useful for further analysis of the class-II reactions and the intermediate products that the enzymes produce. To further investigate the catalytic specificity of *Ab*CAS, previous work in our group used homology modeling combined with molecular substrate docking of GGPP and LPP in the class-II and class-I active sites to predict unique residues which may affect product specificity (Zerbe *et al.* 2012). To enable a comparative analysis between *Ab*CAS and *Ab*LAS, reciprocal amino acid exchanges were carried out in both enzymes. Variants with nonfunctional class-I active sites, specifically *Ab*CAS(D621A) and *Ab*LAS(D611A), to identify the intermediate class-II products.

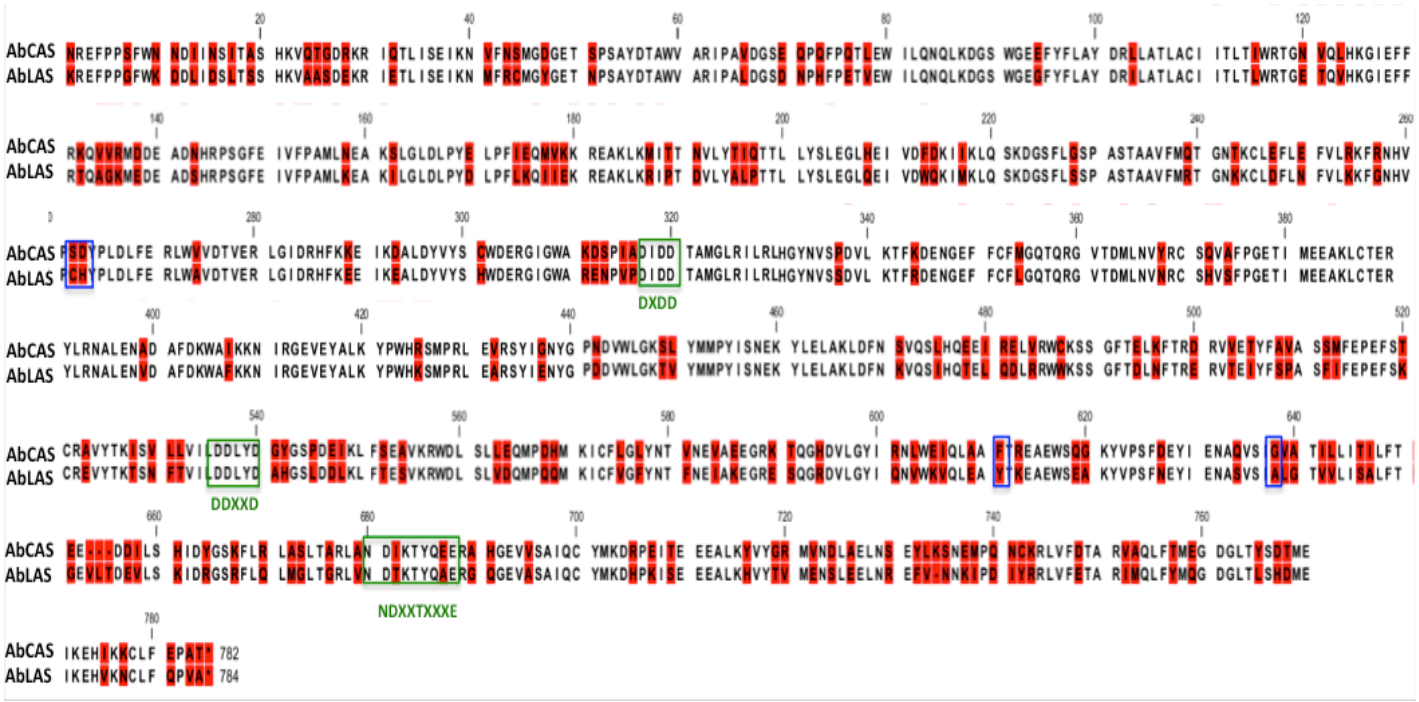


Figure 4. Sequence alignment *Ab*CAS and *Ab*LAS. Amino acid sequence alignment of *Abies balsamea cis*-abienol synthase (AbCAS) and labda-13-en-8-ol diphosphate (AbLAS) that have 75% sequence identity. Shown in green are the conserved motifs DXDD in active site-II and DDXXD and NDXXTXXXE in active site-I.

2.2 Materials and Methods

Diterpene standards

Authentic diterpene resin acid standards were purchased from Orchid Celmark (New Westminster, Canada). The corresponding diterpene olefins were from our laboratory collection of diterpenes synthesized from the acids at Best West Laboratories Inc. (Salt Lake City, UT) as previously described (Ro *et al.* 2005).

Generation of Site-directed Protein Variants

Mutant cDNA clones were constructed by whole plasmid PCR amplification with overlapping mutagenesis primers. All mutant cDNAs were verified by complete sequencing prior to directional ligation into expression vector pET28b(+). Site directed mutagenesis was performed using the QuickChange protocol (Stratagene, Mississauga, Canada) with cDNA in pET28b(+) as a template and the following PCR program: initial denaturation at 98°C for 90 sec followed by 30 cycles of denaturation at 98°C for 50 sec, annealing at 60°C for 30 sec and extension at 72°C for 7 min, and a final extension at 72°C for 10 min. Oligonucleotides used for mutagenesis are listed in the Table 1.

Table 1. Primer identities used to prepare site-directed mutagenesis constructs

AbCAS:S347C-fwd	AAC CAT GTG CCT TGC GAC TAT CCC CTC GAT CTA TTT G
AbCAS:S347C-rev	CAA ATA GAT CGA GGG GAT AGT CGC AAG GCA CAT GGT T
AbCAS:S347C:D348H-fwd-new	AAC CAT GTG CCT TGC CAC TAT CCC CTC GAT CTA TTT G
AbCAS:S347C:D348H rev-new	CAA ATA GAT CGA GGG GAT AGT GGC AAG GCA CAT GGT T
AbCAS:D348H	AAA CAT GTG CCT AGC CAC TAT CCC CTC GAT C
AbCAS:D348H antisense	GAT CGA GGG GAT AGT GGC TAG GCA CAT GGT T

AbLAS:C345S:H346D	AAA TTC GGA AAC CAT GTG CCT AGT GAC TAT CCG CTT GAT C
AbLAS:C345S:H346D antisense	GAT CAA GCG GAT AGT CAC TAG GCA CAT GGT TTC CGA ATT T
AbLAS:H346D	CGG AAA CCA TGT GCC TTG TGA CTA TCC GCT TGA TC
AbLAS:H346D antisense	GAT CAA GCG GAT AGT CACA AG GCA CAT GGT TTC CG
AbCAS:G723A	GAA CGC CCA AGT ATC AAT AGC TGT AGC AAC TAT ACT TCT TA
AbCAS:G723A antisense	TAA GAA GTA TAG TTG CTA CAG CTA TTG ATA CTT GGG CGT TC
AbCAS:F696Y	ATC CAG CTC GCA GCT TAC ACC AGA GAA GCA G
AbCAS:F696Y antisense	CTG CTT CTC TGG TGT AAG CTG CGA GCT GGA T
Ab CAS-fw-Not1-HF	TAG CGG CCG CAT GCG AGA ATT TCC TCC TTC ATT T
Ab CAS- rev-Cla1	TAA TCG ATT TAG GTA GCC GGC TCG AAG

Constructs were transformed into *E.coli* BL21DE3-C41 cells and grown on plates containing LB Medium with Kananmycin overnight. Individual colonies were selected to inoculate 5 mL LB cultures with 50 µg mL^{-1} Kanamycin and grown overnight at 37°C and 220 rpm. 500 µL of each culture were then transferred into 50 mL of autoclaved TB medium (12 g tryptone, 24 g yeast extract, 8 mL glycerol, 0.17 M KH_2PO_4 and 0.72 M $KHPO_4$ per L medium, pH 7.0) with 50 µg mL^{-1} Kanamycin. Cultures were grown at 37°C and 180 rpm in 250 mL unbaffled culture flasks until reaching an optical density at 600 nm (OD_{600}) of ~0.3. At this time cultures were cooled to 16°C before induction with 0.2 mM IPTG (end concentration) at an OD_{600} of ~0.6 and protein expression overnight at 16°C and 180 rpm. Cells were then harvested at 3750 rpm, 4°C for 20 min. The supernatant was discarded and pellets were resuspended in lysis buffer (50 mL Binding buffer, 1/2 protease inhibitor tablet (Roche) and 0.1 mg mL^{-1} lyzozyme) and incubated at 4°C for 30

mins. Sonicated for 30 sec, followed by centrifugation at 4°C for 25 min at 12000 rpm. Recombinant proteins were affinity-purified over His-spin trap columns (Healthcare G. GST Sprin Trap Purification Module) , and desalted over Sephadex G-25 medium beads in a column with desalting buffer (350 mM NaCl, 1 mM $MgCl_2$).

Activity Assays of diTPS expressed from pET28b(+) in *E.coli* BL21DE3-C41

Activity assays were performed in 1.5 mL glass vials (Agilent) with 50 µg protein and 15 µM GGPP in a total volume of 500 µl of buffer (50 mM HEPES, 100 mM KCl, 7.5 mM $MgCl_2$, 20 µM $MnCl_2$, 5% glycerol, 5 mM DTT at a pH of 7.2). 500 µL of pentane spiked with internal standard eicoscene (2 µM) was added as the organic layer, then incubated at 30°C for 60 min. Terpenes were extracted by vortexing twice for 15 sec. Reaction products of *Ab*LAS:D611A and *Ab*CAS:D621A variants required dephosphorylation prior to GC-MS analysis by adding 10 U alkaline phosphatase (Invitrogen) to the aqueous phase and incubation at 37°C overnight.

Gas chromatography–mass spectrometry (GC-MS) Conditions for Analysis of Assay Products

GC-MS analysis was preformed on an Agilent 6890N GC (Agilent Technologies Inc., Mississauga, Canada), 7683B series autosampler and a 5975 Inert XL MS Detector at 70 eV and 1mLmin^{-1} helium as a carrier gas using a HP-5MS Column with the following GC temperature program: injection volume of 1uL, initial temperature 40°C for 2 min then 20°C/min to 300°C for 2 mins. Front inlet conditions 250°C with an injection pulse pressure 50 psi until 0.5 min, pulsed splitless with a flow rate of 1.2 mL/min. Compound

identification was achieved by comparison of mass spectra with those of authentic standards and reference mass spectral databases of the National Institute of Standards Technology MS Search Libraries (Wiley W9N08).

Kinetics Analysis of *Ab*LAS (WT) and *Ab*LAS:C345S:H246D

Single-vial assays were used as described above. Assays were completed in triplicate with 1 to 60 mM GGPP for 20 min at 37°C. Assays products were analyzed on the GC-MS by using selected ion monitoring of m/z 83 (for internal standard) and both 257, 272 and 275 for diterpene products and manoyl oxide. The response of the GC mass spectrometer at each substrate concentration was determined relative to the internal standard by using 20 ug of enzyme in the single-vial assay. Kinetic parameters were determined by nonlinear regression using the Michaelis-Menten model in ANEMONA (Hern and Ruiz 1998) .

SDS-PAGE

I poured a two layered gel, the bottom layer is a separating gel (for 4 gels, 3.8 ml H_20, 6.25 ml 1 M Tris-HCl, 4.7 ml 40% Acrylamide/Bisacrylamide, 150 ul 0% SDS, 90 ul, 10% AMPS and 15ul TEMED) and the top layer is a stacking gel (for 4 gels, 8 ml H_20, 700 ul M Tris-HCl, 1.65 ml 40% Acrylamide/Bisacrylamide, 100 ul 0% SDS, 50 ul, 10% AMPS and 15 ul TEMED). The protein sample (10 ml) and 2x sample buffer (5x, 25 mLTris-HCl, 2 g SDS, 5 mL Beta-mercaptothanol, 5 g sucrose, 0.002 g bromphenolblue) was heated to 98 °C for 10 mins then loaded into wells. The gel was submerge in running buffer (10x, 30.3 g Tris, 183.3 g glycine, 10 g SDS, fill to 1 L H_20) and run at 150 V for 1 hour, then I

stained it with Coomasie Blue Stain (100ml methanol, 50 mL acetic acid, 350 ml H_20 and coomasie brilliant blue) for 1 hour at room temperature then de-stain for 24 hours.

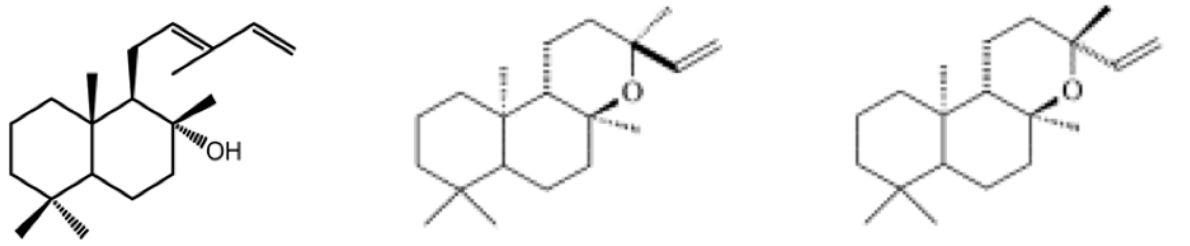
2.3 Results

The work described here investigates the effects that amino acid residue changes have on catalytic activity of diTPSs and the products that these protein variants generate. In order to probe the function of residues that could influence the product profile of diTPSs, I prepared the following class-II active mutants of CAS and LAS enzymes: *Ab*CAS:S347C, *Ab*CAS:D348H, *Ab*CAS:S347C:D348H, *Ab*LAS:H346D, and *Ab*LAS:C345S:H346D.

Since diTPSs have a two step (class-II then class-I) reaction mechanism, the effects of the above mentioned mutations in the class-II active site could only be detected after completion of both reactions. In order to identify the intermediate products of the mutated class-II reaction we generated additional class-I mutations in each diTPSs inactivating the class-I active site to prevent the reaction from going to completion. Mutagenesis of the DXXDD motif allows us to analyze the class-II reactions. To this end, I prepared protein variants *Ab*CAS(D621A) and *Pa*LAS(D611A). The D621A and D611A mutations within the catalytic motifs of *Ab*CAS and *Pa*LAS, respectively, result in non-functional class-I domains and therefore accumulate the reaction intermediate LPP or CPP (Figure 7).

Following these product profile assays, and dephosphorylation of the intermediate reaction product I observed an additional product from the *Ab*LAS:C345S:H346D protein variant which was identified as an epimer of manoyl oxide (either *epi*-manoyl oxide or *ent*-

manoyl oxide) (Figure 5). In order to rapidly identify which manoyl oxide epimer was present, I determined the ratio of product mass fragments because it is a direct characteristic of the epimeric composition of manoyl oxide (Demetzos, Kolocouris and Anastasaki 2002). The ratio of the mass fragments 275:257 is lower for *epi*-manoyl oxide compared to *ent*-manoyl oxide. Based on the results in figure 7, the reaction intermediate of *Ab*LAS:C345S:H346D was identified as *ent*-manoyl oxide.



(A) *cis*-abienol (B) *ent*-manoyl oxide (C) *epi*-manoyl oxide

Figure 5. Structure of (A) *cis*-abienol, (B) *ent*-manoyl oxide and (C) *epi*-manoyl oxide.

CAS Class II Active Site Variants

Based on the homology modeling of *Ab*CAS and GGPP substrate docking described by Zerbe *et al.* 2012 two residues were identified in the class-II domain, S347 and D348 (Figure 6). These residues were identified as targets for mutagenesis because they are positioned in proximity to the docked substrate therefore it was hypothesized by *Zerbe et al.* 2012 that they may impact catalysis.

Protein variant *Ab*CAS(D611A):S347C resulted in only minor changes of the class-II active site product profile with a decrease in LPP and an increase in both *ent*-manoyl oxide and *epi*-manoyl oxide. No changes to the final *cis*-abienol product profile were detected when analyzing *Ab*CAS:S347C (Figure 7). In contrast, protein variant *Ab*CAS(D611A):D348H showed a change in product profile with (+)-CPP (84.2%), which was not present in the product profile of *Ab*CAS(D611A). This change in the class II active site product translated into trace amounts of olefins produced by A*b*CAS:D348H containing a functional class I active site. Protein variant *Ab*CAS(D621A):S347C:H346D containing two mutations in the class II active site showed a reversion back to a more native intermediate product profile (58.9% LPP, 11.2% CPP, 16.8% *ent*-manyol oxide and 13% *epi*-manyol oxide). Olefins were detected in trace amounts the product profile of *Ab*CAS:S347C:H346D

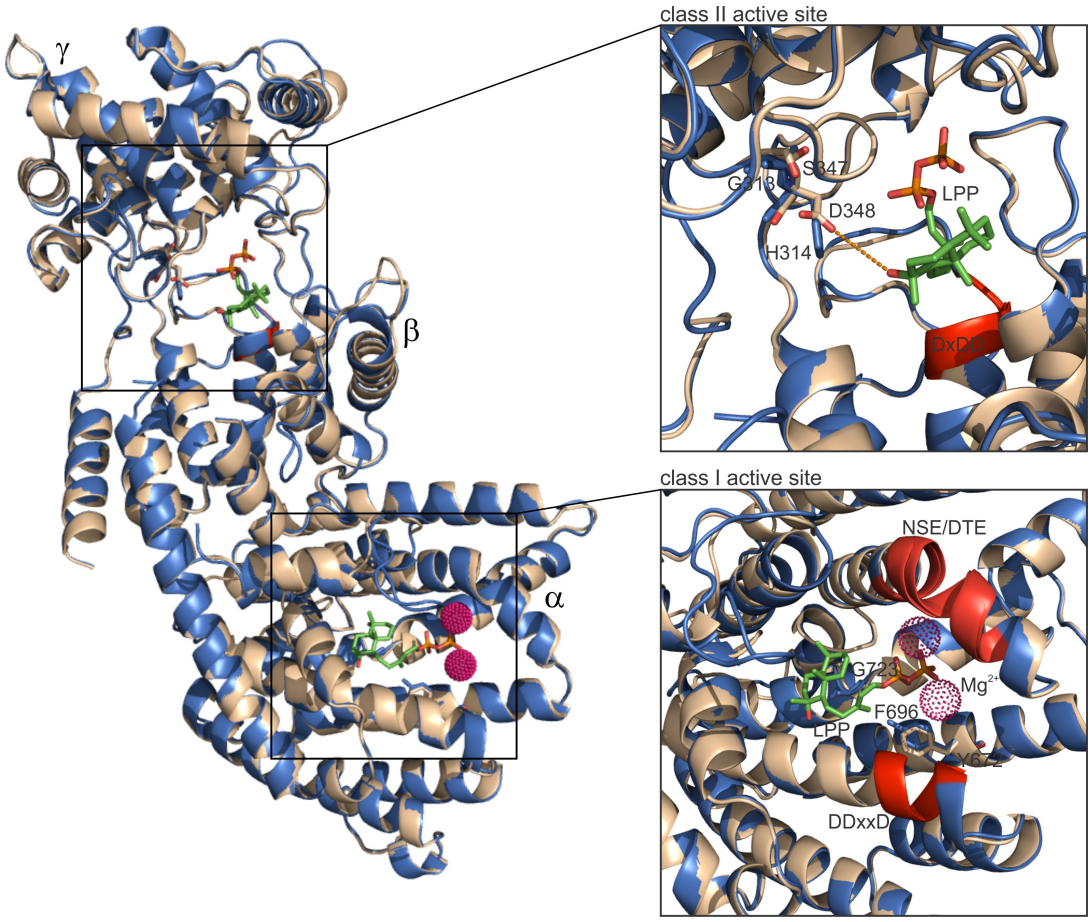


Figure 6. Homology model of diTPS (*Ab*CAS and *Ab*LAS reveals) unique active site residues for *cis*-abienol formation. (A) *Ab*CAS adopts the conserved alpha-helical TPS folding pattern, comprised of a N-terminal β-domain, γ-domain and a C-terminal α-domain. Within the class II active site of *Ab*CAS, D348 is ideally positioned to stabilize the intermediate carbocation for labda-13-en-8-ol diphosphate to be formed. Similarly, Gly723 and Phe696 may impact the enzyme's class I activity. This figure was drawn by Philipp Zerbe 2014.

Table 2. Product Profile of *Ab*CAS, *Ab*LAS and *Pa*LAS, and the respective protein variants. Red is CAS and blue is LAS. Variants *Ab*CAS(D621A) and *Ab*LAS(D611A) have a non-functional class-I active site. (*) represents a mutation in the class I active site. Experiments were performed in triplicate, ± indicates standard error. Chromatograms displaying the product profile from enzyme variants is in Appendix 3.

Product Profile Composition of *Ab*CAS, *Ab*LAS and *Pa*LAS, and their Mutants

		Percent of Product Profile			
Enzyme Variant		LPP	CPP	*ent*-manoyl oxide	*epi*-manoyl oxide
AbCAS(D621A)		**75.7±0.4**	-	13.4±0.03	10.9±0.04
PaLAS(D611A)		-	**100**	-	-
AbCAS(D621A): S347C		**60.1±0.13**	-	23.6±0.01	16.3±0.01
AbCAS(D621A): D348H		15.8±0.31	**84.2±0.05**	-	-
AbCAS(D621A): S347C:D348H		**58.9±0.29**	11.2±0.03	16.8±0.03	13±0.01
PaLAS(D611A): H346D		**100**	-	-	-
PaLAS(D611A): C345S:H346D		-	-	-	-
		cis-abienol	olefins	*ent*-manoyl oxide	*epi*-manoyl oxide
AbCAS (WT)		**100**	-	-	-
AbCAS: S347C		**100**	-	-	-
AbCAS: D348H		-	trace	-	-
AbCAS: S347C:D348H		-	trace	-	-
AbLAS (WT)		-	**100**	-	-
AbLAS: H346D		-	-	-	-
AbLAS: C345S:H346D		-	-	**100**	-

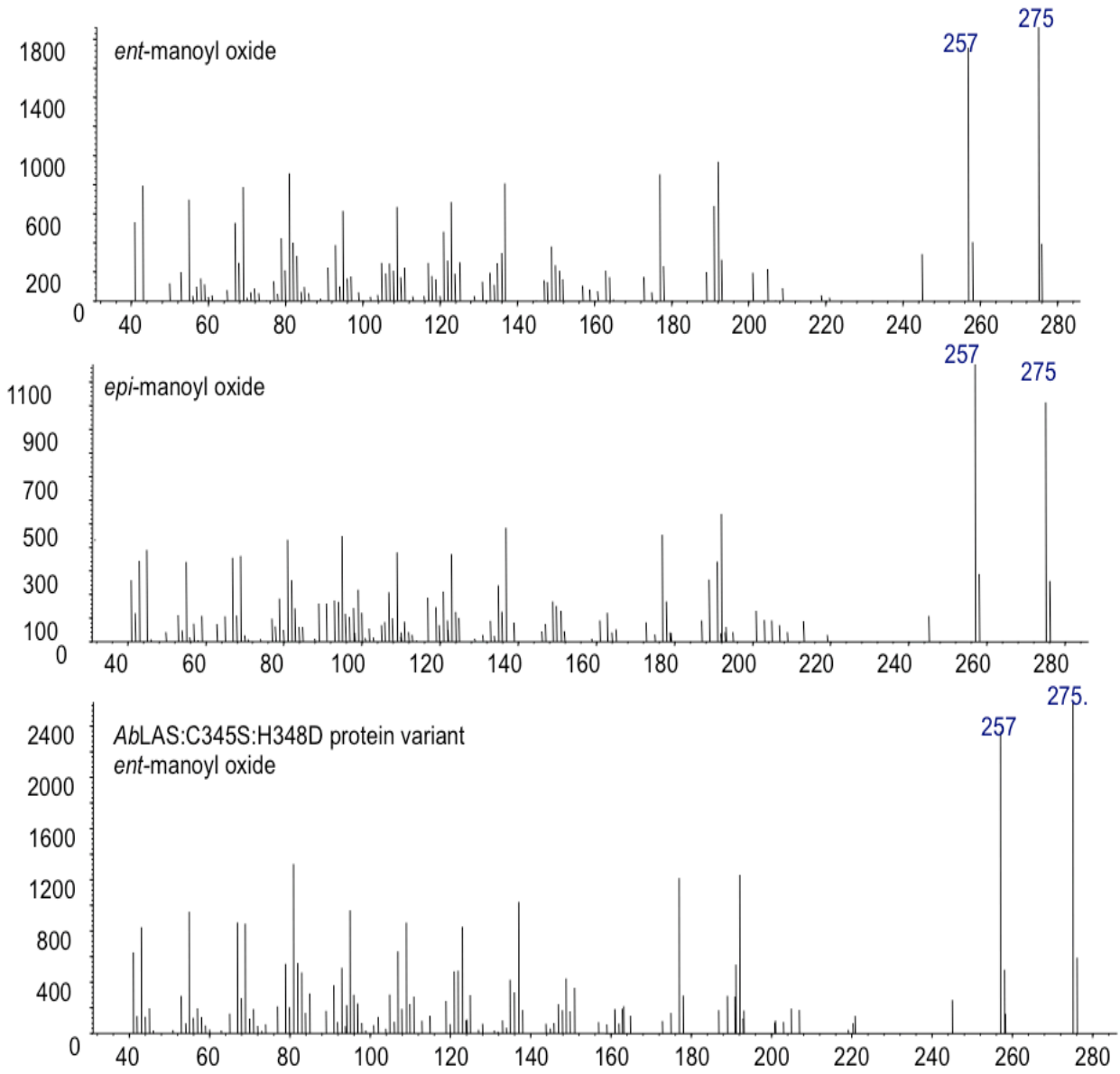


Figure 7. *Ab*CAS(D621A) and *Ab*LAS:C345S:H348D protein variants products, displaying experimental mass spec for *ent*-manoyl oxide and *epi*-manoyl oxide. GC-MS analysis was performed on an Agilent HP5ms column with electronic ionization at 70 eV.

LAS Protein Variants

In order to determine if the mutations (S347C and D348H) in *Ab*CAS also affect product profiles of the class-II activity of other diTPSs, I extended my study into the LAS enzyme. The LAS enzyme was used from two sources, one from balsam fir (*Ab)* and the other is Norway Spruce (*Pa)*. The *Pa*LAS enzyme was used in all tests with inactivated class-I enzyme because *Pa*LAS(D611A) was previously developed in our lab (unpublished results) and available for my project. *Pa*LAS(D611A) produces CPP as the only product (Table 2). Mutation H346D in protein variant *Pa*LAS(D611A):H346D results in a complete change of product profile to LPP. I was unable to successful generate the individual mutation C345S in an LAS enzyme. I generated the double mutant enzyme *Pa*LAS(D611A):C345S:H346D which produced no traceable amounts of CPP or LPP products.

In the *Ab*LAS background, no product was detected for the mutant *Ab*LAS:H346D. However, the double mutant *Ab*LAS:C345S:H346D showed a new activity and produces ent-manoyl oxide (100% product profile) (figure 8 and 9).

Kinetic properties of *Ab*LAS:C345S:H346D

I preformed kinetic analysis of *Ab*LAS:S347C:H346D (Figure 10, Table 2). Since *Ab*LAS:S347C:H346D is a bifuncational enzyme, the K*cat* value of 0.57min^{-1} calculated in this experiment encompasses both class II and I active sites. Based on the results it is evident that *Ab*LAS (WT) has the capacity to generate 15x more product per unit of time than *Ab*LAS:C345S:H346D. However, the Km of *Ab*LAS:C345S:H346D is lower than that of *Ab*LAS (WT) indicating that *Ab*LAS:C345S:H346D is more efficient enzyme at binding to the GGPP substrate. Despite the greater affinity *Ab*LAS:C345S:H346D enzyme has for the substrate, *Ab*LAS (WT) produces more product. These differences could be attributed to the reaction within the class-II active site (deprotonation is a quicker reaction than hydroxylation) or to the rate at which the intermediate and final products are released from the active sites. While the structural consequences of this mutation must be fully evaluated in future X-ray crystal structure determinations to confirm that no unanticipated structural changes are triggered by the mutations.

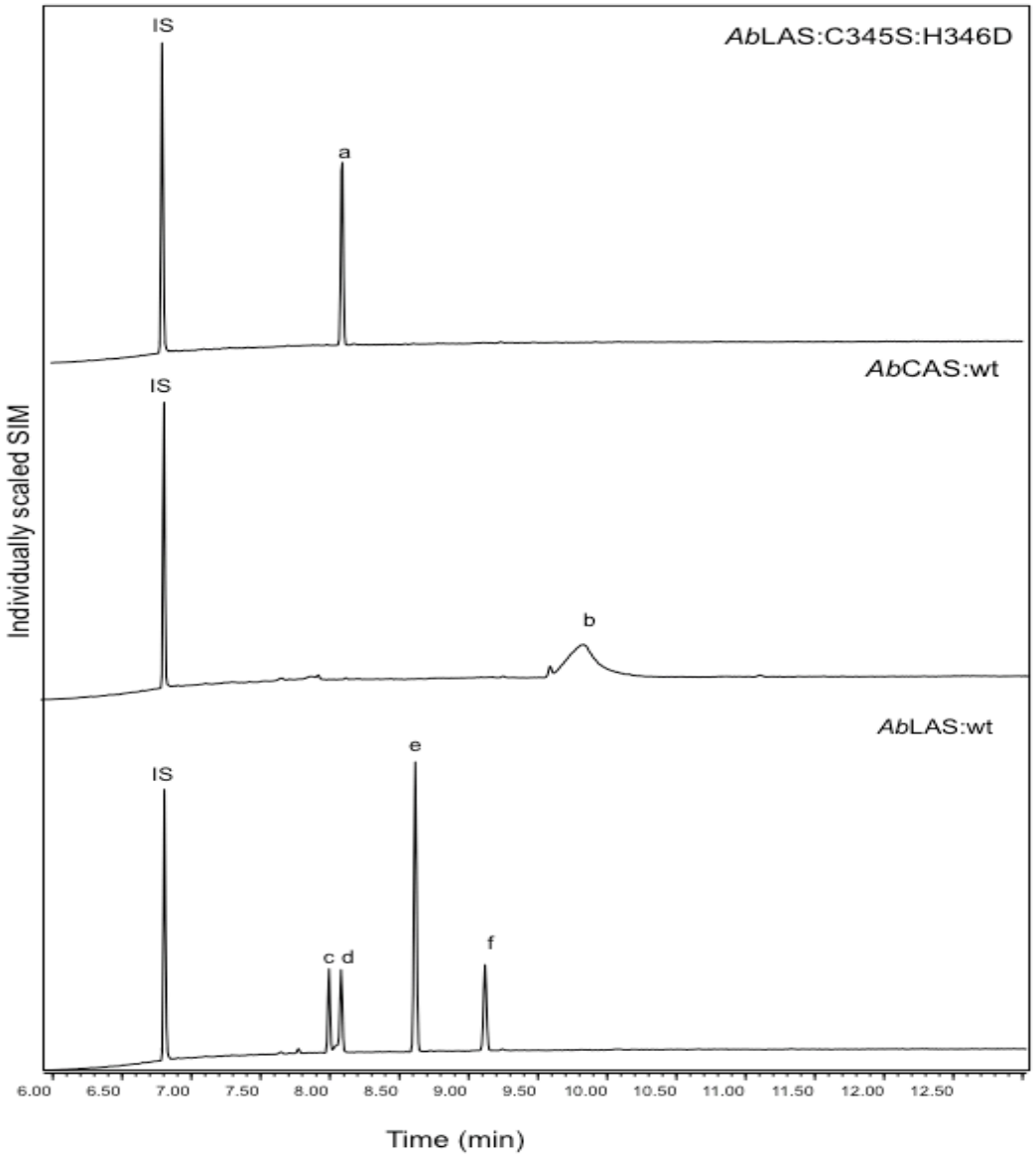


Figure 8: Activity of balsam fir diTSPs *Ab*CAS (WT), *Ab*LAS (WT), *Ab*LAS:C345S:H346D. Individually scaled SIM chromatograms of reaction products from in vitro assays with purified recombinant enzymes using GGPP as a substrate. GC-MS analysis was performed on an Agilent HP5ms column with electronic ionization at 70 eV. Enzymatic activity assays were confirmed with three independent experiments, Peak IS, internal standard 2 uM 1-eicosene, peak (a) Manoyl oxide, peak (b) *trans*-abienol, peak (c) Palustradiene, peak (d) Levopimaradiene, peak (e) Abietadiene, peak (f) Neoabietadiene.

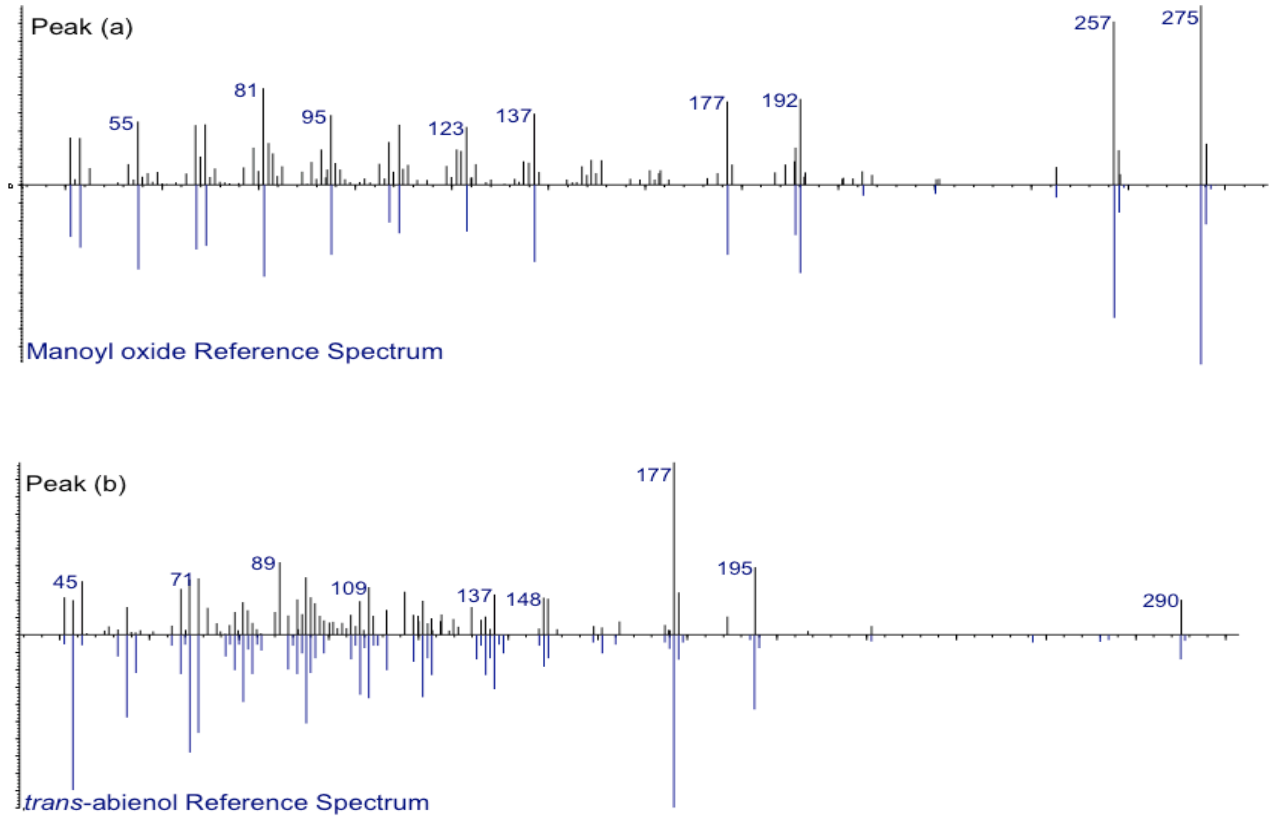


Figure 9: Experimental (black) and reference spectrum (blue) shown for manoyl oxide and *trans*-abienol. Chromatogram for Peak (a) is shown in figure 8. Manoyl oxide is the product of *Ab*LAS:C345S:H346D**.** Peak (b) is trans-abienol from *Ab*CAS (WT). GC-MS analysis was performed on an Agilent HP5ms column with electronic ionization at 70 eV. The heat from the HP5 column isomerizes the *cis*-abienol to *trans*-abienol.

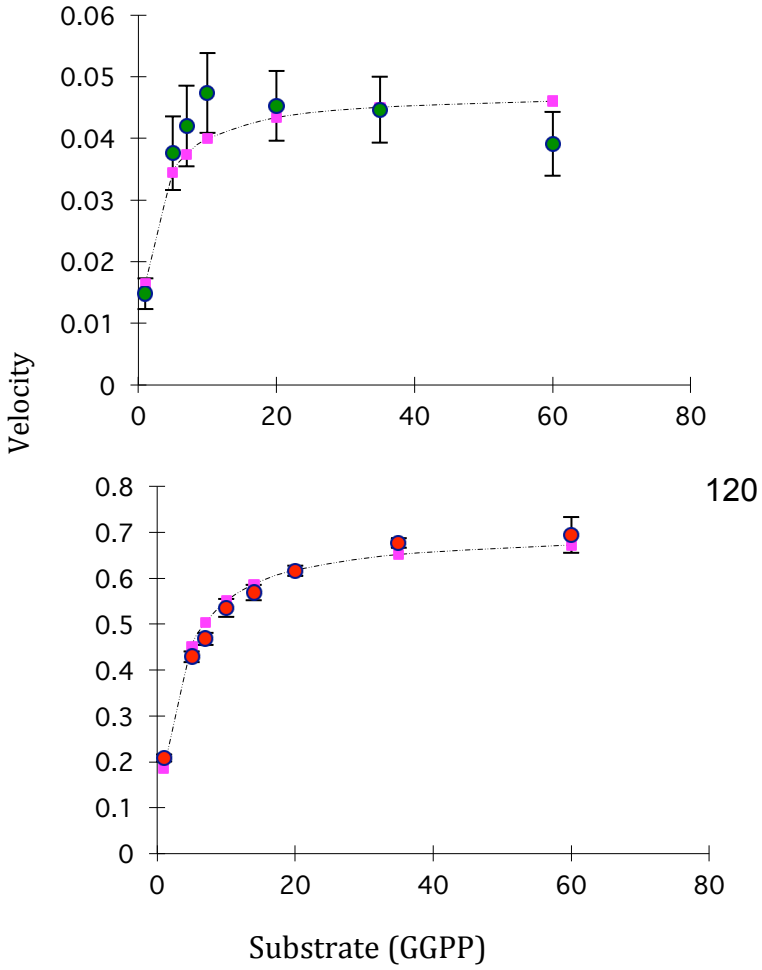


Figure 10: Rate of product formation of olefins by *Ab*LAS (WT) and manoyl oxide by *Ab*LAS:C345S:H346D at increasing concentration of GGPP substrate. Error bars represent standard deviation of three independent experiments. Pink squares are calculated to the Michaelis-Menten model. **(A)** Green circles are observed data for *Ab*LAS:S347C:D348H, **(B)** red circles are observed data for *Ab*LAS (WT).

Class-I protein variants in *Ab*CAS

Previous mutational and modeling studies by Zerbe *et al.* (2012) and Keeling *et al*. (2008) identified catalytic residues in the class-I active site of ISO- and LAS diTPS that impact class-I activity. Given the close similarity between diTPSs of the TPS-d subfamily, I investigated corresponding mutations the class-I reaction of *Ab*CAS. I introduced two different single residue substitutions (F696Y and G723A) into the *Ab*CAS (WT) background. F696 is located in the class-I active site cavity upstream of the DDxxD motif, creating an expansion of the hydrophobic pocket relative to LAS and ISO, which contain a tyrosine and histidine in these positions (Figure 6). G723 contributes to a change in the hinge region between helix G1 and G2. These residues were hypothesized by Zerbe *et al.* 2012 to contribute to the release of a bicyclic product. In my work with CAS, *Ab*CASF696Y and *Ab*CASG723A did not show a change of product profile compared to the wild type enzyme (Appendix 2). However, these amino acid exchanges reduced *Ab*CAS enzyme activity. *Ab*CAS:F696Y variant and the *Ab*CAS:G723A variants were only 33% and 9.75%, respectively, as effective at producing *cis*-abienol as the *Ab*CAS (WT) enzyme.

2.4 Discussion

In order to probe the function of residues that could affect catalytic specificity and product profiles of *Ab*LAS and *Ab*CAS enzymes, I prepared protein variants with site-specific mutations. Based on the LPP substrate position in the *Ab*CAS model (Figure 6) it was hypothesized that the unique residue D348 (which is a H346 in LAS and ISO), is found in close proximity to the docked LPP substrate and may contribute to the hydroxylation of the primary labda-13-en-8-yl (LPP) carbocation by stabilizing the positive charge at C-8 for

water quenching to occur. This ability to stabilize the charge does not occur in *Ab*LAS which contain a conserved potentially positively charged His^+ at this same potition. In contrast to the formation of the hydoxylated LPP produced at the class-II active site of *Ab*CAS, *Ab*LAS produces (+)-CPP at the class-II active site. Confirming this hypothesis, the D348H mutation in the *Ab*CAS(D621A) background showed a shift of product profile from 75.7% LPP in *Ab*CAS(D621A) to 84.2% CPP in *Ab*CAS(D621A):D348H. While this work was in progress, Criswell *et al*. (2012) found that mutating the relevant His in both *At*CPS and *Ag*AS changes the product outcomes of both of these enzymes with the mutant versions catalyzing the production of a novel hydroxylated form of CPP. Criswell *et al*. (2012) research independently confirmed our work in *Ab*CAS. A D348H amino acid mutation may have occurred in the evolution of *Ab*CAS from a common ancestor with *Ab*LAS and may have given rise to the *cis*-abienol synthase function and the cis-abienol metabolite that is abundant in balsam fir. Indeed, in the presence of a functional class-I active site the *Ab*CAS:D348H (as well as *Ab*CAS:S347C:D348H) produced trace amounts of olefins which are not found in the CAS product profile, but are typical for LAS.

The *Ab*LAS:H346D protein variant showed no enzyme activity; however, *Ab*LAS:S347C:H346D catalyzed the formation of a product, manoyl oxide, not previously shown to be formed by a diTPS of the TPS-d family. Manoyl oxide is a diterpene intermediate in the biosynthesis of the pharmaceutical compound forskolin in *Coleus forskohlii* (Pateraki et al. 2014). In *C. forskohlii*, (13R)-manoyl oxide is produced by the combined activities of two monofunctional diTPS, class II CfTPS2 combined with class I CfTPS3 or CfTPS4.

Overall, the results of this chapter demonstrate the inherent plasticity of diTPSs and the impact of minor mutations on product profiles which affect the metabolite profiles of plants. Screens of mutated diTPS also have potential to discover new and unexpected enzyme functions as demonstrated with the formation of manoyl oxide by *Ab*LAS:S347C:H346D.

Chapter 3: Exploring engineered yeast (*Saccharomyces cerevisiae*) strains for the production of *cis*-abienol

3.1 Introduction

Progress in metabolic engineering has given rise to a number of microbial platforms available for research for the expression of a variety of interesting natural products (Nielsen 2013, Kirby and Keasling 2009, Ro *et al.* 2006, Westfall 2012). This field of biotechnology has made possible the industrial scale up of previously rare compounds that were costly and difficult to synthesize such as the anti-cancer drug, taxol (Engels 2008, Ajkumat *et al.* 2010). Recent developments in synthetic biology promise to expand the metabolic engineering toolbox further through the development of 'plug-and-play' vectors and reliable microbial platforms, which are all vital biological components for designing pathways that are user-friendly and cost-effective. To date more than 12,400 patents using *Saccharomyces cerevisiae* (yeast) have been filed for the production of pharmaceuticals, food ingredients, chemicals and fuels (Hong and Nielsen 2012). One prominent example is the successful generation of a yeast strain producing the artemisinin precursor amorphadiene (Paddon and Keesling 2014). In other systems for diterpene production, Caniard *et al*. (2012) cloned two diTPS enzymes for the complete biosynthesis of sclareol in yeast; Hamburger *et al.* (2011) produced a suite of diterpene resin acids and other diterpenoids in metabolically engineered yeast; and Ghou *et al.* 2013 demonstrated the production of tanshinones in yeast, which are the main lipophilic bioactive components of the Chinese medicinal herb danshen (*Salvia miltiorrhiza*).

The use of microbial platforms for terpene production may have advantages compared to extracting biochemicals from their natural sources such as the availability of data on metabolic networks, the ease of reconstructing pathways and the ability to scale up the fermentation process. One of the most attractive hosts for de-novo production of terpenoids is baker's yeast (*Saccharomyces cerevisiae)*. Many of the cellular pathways in yeast share a high degree of similarity to those of plant cells. In yeast (and with all organisms used for microbial engineering) some of the critical elements for stable production of terpenoids are: choice of host strain, type of transformation, expression level, codon-bias and N-terminal modifications.

The objective of this thesis chapter is to investigate the potential production of *cis*-abienol in *S. cerevisiae.* The total synthesis of the perfume component Ambrox is difficult and costly therefore we are interested in other sources. The semi-synthesis of Ambrox from *cis*-abienol and other oxygen-containing diterpenoids of plant origin (e.g. sclarerol and manool) or such compounds produced in a microbial host could be a cost effective and reliable source of the perfume fixative agent. As a eukaryote yeast has the advantage over prokaryote organisms (i.e. *E.coli*) of being able to support post-translational modifications and proper localization of host membrane-anchored and organelle-specific enzymes. However, compared to the native environment of plant cell where the *Ab*CAS originates, yeast lacks the plastid organelles which are the location of diTPS activity in plants. Therefore, I have investigated the microbial engineering of *S. cerevisiae* for the ability to produce *cis*-abienol.

Central to any metabolic engineering effort is the ability to redirect appreciable carbon flux from central metabolism towards desired heterologous end products (Ignea et al. 2012). In the present work, we used three different yeast strains. Each yeast strain has its own characteristics and potential benefits to using it as a platform for diterpene production. The genotypes of the three yeast strains used in this experiment are presented in table 3 and their modifications are discussed in more detail below.

Table 3: Genotypes of *Saccharomyces cerevisiae strains* used for engineering and production of *cis*-abienol

Strain	Genotype	Source
AM94	Mat a/α, Galp-(K6R)HMG2::HOX2,ura3, trp1, PTDH3-HMG2(K6R)X2-::leu2 ERG9/erg9	Ignea et al. 2012
BY4741	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Standard Lab strain
KE	CENPKERG20K197E: mat alpha, haploid(ΔERG20::ERG20K197E;KanMX; Ura-, Leu-, His-, Trp-)	Dr. Vince Martin, Concordia University, unpublished

The mevalonate pathway is the isoprenoid pathway for the production of terpenoids in yeast. It begins with acetyl-CoA and is responsible for the production of GGPP, which *Ab*CAS converts into *cis*-abienol. Initially the pathway is linear, (in the order of: Acetyl-CoA, acetoacetyl-CoA, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), mevalonic acid (MVA), mevalonate-5-phosphate, mevalonate-5-pyrophosphate, isopentenyl-5-pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP)) followed by a branch matrix for the production of geranyl- pyrophosphate (GPP), farnesyl- pyrophosphate (FPP) and geranylgeranyl- pyrophosphate (GGPP) (Figure 11).

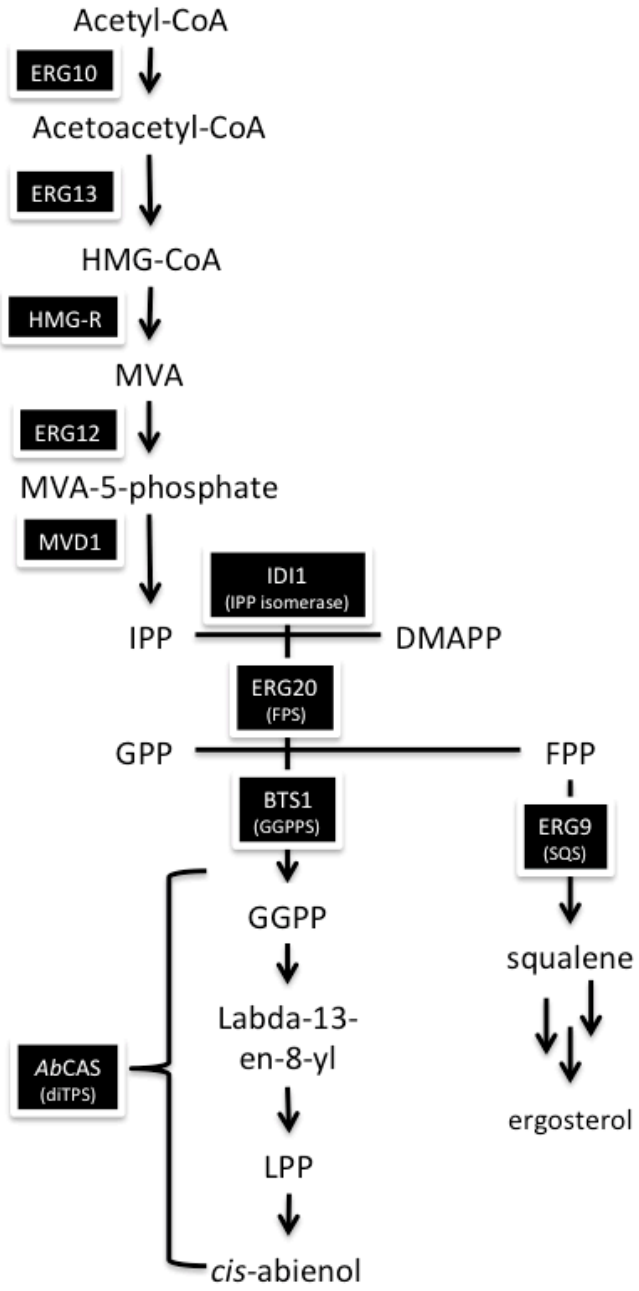


Figure 11. The mevalonate pathway is the isoprenoid pathway for the production of terpenoids, and eventually *Ab*CAS in yeast. The solid arrows indicate the one-step conversions in the biosynthesis, and the multiple arrows indicate the several steps.

3-Hydroxy-3-methylglutaryl (HMG)-CoA reductase (HMGR) is a key enzyme in isoprenoid biosynthesis. The degradation of HMGR is subject to feedback signals from the mevalonate pathway in yeast. Yeast has two isozymes of HMGR termed hmg1p and hmg2p and encoded by HMG1 and HMG2. Hmg1 is more stable, whereas Hmg2 undergoes feedback regulated ubiquitination that in turn leads to proteasomal degradation when sterol pathway activity is high (Hampton and Rine 1994, Hampton and Bhakta 1997). GGPP is a potent endogenous regulator of Hmg2p and can cause degradation by directly altering the structure of the protein therefore limiting flux through the mevalonic acid pathway to produce GGPP as a substrate available to *Ab*CAS for *cis*-abienol production. The HMG2(K6R) variant is known to decrease protein degradation therefore stabilizing the protein while maintaining the specific localization of the enzyme in the perinuclear (Garza *et al.* 2009) . Proper localization of the Hmg2p enzyme could be of importance to toxicity, as over expression was reported to result in growth properties reduction in yeast strains (Demetzos *et al.* 2002). The variant HMG2 (K6R) was inserted into AM94 to increase isoprenoid flux and does not alter the yeast growth properties (Ignea *et al.* 2012).

Yeast also has several regulatory mechanisms to maintain tight control over the intracellular level of farnesyl diphosphate (FPP), the central precursor to nearly all yeast isoprenoid products and a building block for the formation of GGPP (Paradise *et al.* 2008). The ERG20 enzyme is an essential gene in yeast primarily responsible for FPP biosynthesis, but also produces GPP. The ERG20 (K197E) non-native gene contains a galactose inducible promoter and was previously characterized to show it produced

GPP:FPP at a ratio of 2:1 and overall causes a decrease in FPP production. This specific mutation is present in the KE-yeast strain used in experiments for the production of *cis*-abienol in this chapter. The KE strain was provided by Dr. Vince Martin, Concordia University, Canada.

In contrast to the KE strain's modification of the ratio of FPP produced by the ERG20 (K197E) mutant, the AM94 yeast strain has a different approach to controlling flux through the mevalonic acid pathway. AM94 is engineered to produce high levels of FPP, which is used for production of endogenous sterols and non-native terpenoids. Ergosterol is the main sterol produced in yeast and is an important constituent of secretory vesicles and for mitochondrial respiration, therefore rendering it essential to yeast viability. In order to produce non-endogenous terpenoids, it is required that FPP flux is diverted from production of sterols to the heterologous metabolic pathway. To do so, expression of the gene encoding squalene synthase (ERG9), the first committed step in sterol biosynthesis (the enzyme that acts downstream of ERG20), can be down-regulated. Squalene synthase (SQS) receives FPP for its immediate conversion into squalene and therefore plays a key role in determining the flux through the FPP branch point. To decrease the amount of FPP routed to sterol synthesis, flux to squalene was limited by down-regulating expression of ERG9. As ERG9 deletion is lethal without exogenous supplementation of sterols, only one of the two ERG9 alleles was deleted in the AM94 yeast strain. Single allele deletions in yeast can lead to 50% reduction of gene expression. In summary, this genetic change can in theory increase non-endogenous terpene production and decrease the amount of squalene and thus ergosterol production. In other terpenoid biotechnology

research the production of amorphadiene was increased five-fold by ERG9 repression (Paradise *et al.* 2008).

The three different yeast strains (AM94, BY4741 and KE strain) and there respective genetic changes were used in continuations with the goal to test them for production of *cis*-abienol.

Table 4. Experimental overview of yeast engineering for the production of *cis*-abienol.

Host Strain	Yeast Strain used for experiments
AM94 (Mat a/α, Galp-(K6R)HMG2::HOX2,ura3, trp1, PTDH3-HMG2(K6R)X2-::leu2 ERG9/erg9)	pESC^{His} AbCAS/BTS1 pESC^{His} AbCAS/BTS1 + pGREG 505^{Leu}:ERG2.0 pESC^{His} AbCAS/BTS1 + pGREG 505^{Leu}:ERG2.0 + pESC^{Trp} Pg prenyl-transferase pESC^{His} AbCAS/BTS1 + pESC^{Trp} Pg prenyl-transferase No vector
BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0)	pESC^{His} AbCAS/BTS1 pESC^{His} AbCAS/BTS1 + pGREG 505^{Leu}:ERG2.0 pESC^{His} AbCAS/BTS1 + pGREG 505^{Leu}:ERG2.0 + pESC^{Trp} Pg prenyl-transferase pESC^{His} AbCAS/BTS1 + pESC^{Trp} Pg prenyl-transferase No vector
KE strain (CENPKERG20K197E: mat alpha, haploid(ΔERG20::ERG20K197E;KanMX; Ura-, Leu-, His-, Trp-)	pESC^{His} AbCAS/BTS1 pESC^{His} AbCAS/BTS1 + pESC^{Trp} Pg prenyl-transferase No vector

3.2 Materials and Methods

Table 5. Vector constructs used for *S. cerevisiae* transformation.

Vector construct	Details
pESC^{His} AbLAS/BTS1	AbLAS is a 2.3 kb cDNA encoding vector. Co-expression of GGPP synthase encoded by the BTS1 gene (1kb).
pESC^{His} AbCAS/BTS1	AbCAS is a 2.3 kb cDNA encoding vector flanked by restriction sites Not1 and Cla1. Co-expression of GGPP synthase encoded by the BTS1 gene (1kb).
pESC^{Trp}Pg prenyl-transferase	PgcDNA encoded from Pholem, untreated with methyl jasmonate (MJ). Primers "PgGGPP1 fw" and "PgGGPP1 rev" containing restriction sites Eco RI and Not 1. PgGGPP1 is 1kb in length. Predicted target peptide using Cholorp Program.
pGREG505 ERG2.0 Plasmid	This plasmid is part of the "pGREG-SET" (Jansen 2005). The specific genotype of this plasmid is pGREG505 ERG2.0= ERG8-ERG12-ERG13-tHMGR kindly provided by Dr. Vince Martins of Concordia University, Canada.

S. cerevisiae strain construction

The *S. cerevisiae* yeast host strains used in this study were AM94 (Mat a/α, Galp-(K6R)HMG2::HOX2,ura3, trp1, PTDH3-HMG2(K6R)X2-::leu2 ERG9/erg9), BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and KE Strain (CENPKERG20K197E: mat alpha, (ΔERG20::ERG20K197E;KanMX; Ura-, Leu-, His-, Trp-). *Escherichia coli* α-Select Chemically Competent Cells (Bioline) were used for routine cloning and plasmid propagation. Plasmid transformation of all strains of *S.cerevisiae* was preformed by the standard LiCl method according to Gietz *et al.* 1992. Transformed yeast strains were

selected on plates with appropriate synthetic complete drop-out selection medium and grown at 30°C for 48 h.

Media and Growth Conditions

Recombinant yeast was initially grown overnight at 30°C in 5 mL of 2% dextrose in minimal selective media. The next day, a 50 mL culture was initiated at a starting OD_{600} of 0.2 and grown at 30°C with shaking at 180 rpm until the culture reached an OD_{600} of 0.6–0.8. All optical density at 600 nm (OD600) measurements were taken using a Biochrom Ultrospec 3000 pro spectrophotometer. Expression was initiated by transfer into minimal selective media with 2% galactose and grown for 14–16 h. 50 mL cultures were grown at 30°C to OD600 in 250 mL unbaffled culture flasks containing SD media. Yeast cells were harvested by centrifugation at 1,000×*g* for 10 min and washed once with 5 mL sterile ddH_2O. Cells were extracted twice by vortexing for 1 min with 2 mL hexane and 250 µL acid-washed glass beads (425–600 µm, Sigma). Pooled extracts were transferred to a clean test-tube containing anhydrous Na_2SO_4 and evaporated under a gentle stream of N_2 gas to about 200 µL. The samples were transferred to a GC glass vial for GC-MS analysis or stored at −80°C.

GC-MS Analysis

GC-MS analysis was carried out on an Agilent 7890A/5975C GC-MS system operating in electron ionization selected ion monitoring (SIM)-scan mode. Samples were analyzed on an HP5 (non-polar; 30 m, 0.25 mm ID, 0.25 µm film thickness), the injector was operated in pulsed splitless mode with the injector temperature maintained at 250°C.

Helium was used as the carrier gas with a flow rate of 0.8 mL min^{-1} and pulsed pressure set at 25 psi for 0.25 min. Scan range: m/z 40–400; SIM: m/z 83, 151, 177, 257, 272 and 275. The oven program for the HP5 column was: 40°C for 3 min; ramp of 10°C min^{-1} to 130°C, 2°C min^{-1} to 180°C, 50°C min^{-1} to 300°C; 300°C for 10 min. Chemstation software was used for data acquisition and processing. Compound identification was achieved by comparison of mass spectra with those of authentic standards and reference mass spectral databases of the National Institute of Standards Technology MS Search Libraries (Wiley W9N08)

3.3 Results

Results for production of *cis*-abienol in three different *S. cerevisiae* strains (AM94, BY4741 and KE Strain) are shown in Figure 12. Following galactose induction, the ether-extractable fraction of the yeast culture medium and cell pellet were analyzed by GC-MS. It was discovered that the product *cis*-abienol was only in the lysed cell pellets, therefore all results discussed refer to only the lysed cell fraction. A single chromatographic peak (12.95min) corresponding to *cis*-abienol was observed in all three strains expressing [$pESC^{His}$ AbCAS/BTS1] (Figure 13). The mass fragment pattern of this peak had the same diagnostic ions (*m/z* 290, 257, 191, 134) as that reported for *cis*-abienol. These results showed that the production of *cis*-abienol is achievable in yeast. Figure 13 displays the chromatograms of the three yeast strains engineered with a variety of vectors ($pESC^{His}$ AbLAS/BTS1, $pESC^{His}$ AbCAS/BTS1, $pESC^{Trp}$Pg prenyl-transferase, and pGREG505 ERG2.0 Plasmid). Co-expression of [$pESC^{His}$ AbCAS/BTS1] and [pGREG 505^{Leu}:ERG2.0] produced the highest levels of *cis*-abienol in yeast strains AM94 (7.2 mg/L) and BY4741

(1.9 mg/mL) (Figure 14). In contrast, KE yeast strain was not viable when co-transformed with [pESC^{His} AbCAS/BTS1] and [pGREG 505^{Leu}:ERG2.0] (Figure 14). The highest level (1.1 mg/mL) of *cis*-abienol produced in the KE strain occurs when it is expressing [pESC^{His} AbCAS/BTS1]+ [pESC^{Trp}Pg prenyl-transferase] vector.

Figure 15 displays the chromatogram overlay of the *cis*-abienol peak from the AM94, BY4741 and KE yeast strains that produced the highest levels (7.2 mg/L, 1.9 mg/mL and 1.1 mg/mL, respectively) of *cis*-abienol. *S. cerevisiae* AM94 ([pESC^{His} AbCAS/BTS1] + [pGREG 505^{Leu}:ERG2.0]) is the highest producer of *cis*-abienol, BY4741 ([pESC^{His} AbCAS/BTS1] + [pGREG 505^{Leu}:ERG2.0]) produces less *cis*-abienol but a similar amount of squalene to AM94 and, finally, KE Strain ([pESC^{His} AbCAS/BTS1]+ [pESC^{Trp}*Pg* prenyl-transferase]) is the lowest producer of *cis*-abienol but produces less side products therefore rendering it a much purer *cis*-abienol producer, comparatively.

To address the question of localization of *cis*-abienol following expression in yeast, we tested for *cis*-abienol in the yeast cells and in the culture medium. *cis*-Abienol was only detected in the cytosolic cell fraction and not in the growth media or attached to cell debris (Figure 16), and was also not present in untransformed control yeast (strains transformed with water or expressing only pGREG 505^{Leu}:ERG2.0). Squalene was present at varying amounts in the cell pellets of the engineered yeast strains (data not shown).

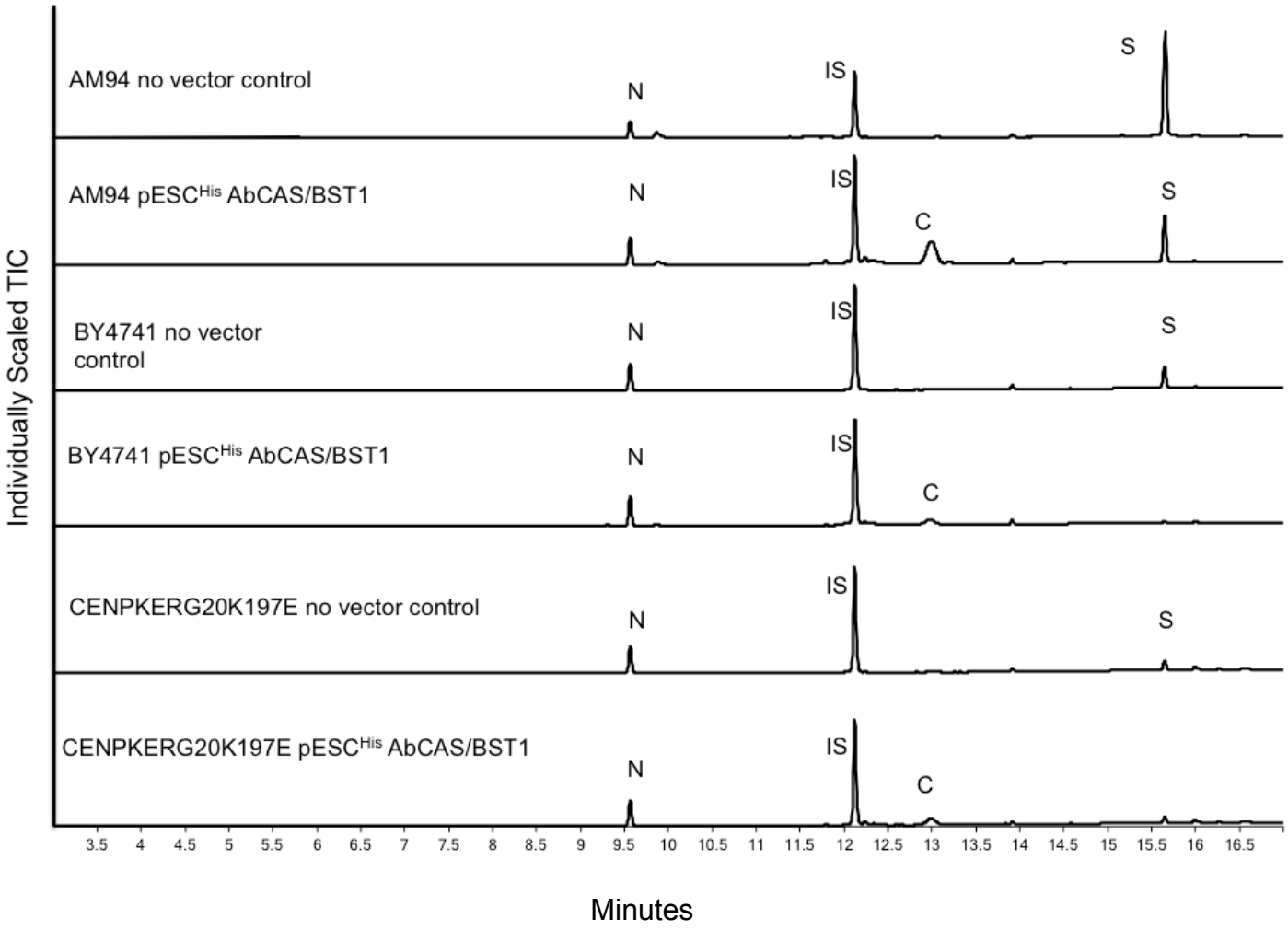


Figure 12. Induced expression of *cis*-abienol in *Saccharomyces cerevisiae*. Shown are total ion chromatograms (TIC) of extracts from three lysed cell pellets from yeast strains expressing *Ab*CAS as compared to control strains. Peak IS, internal standard 0.1 nmol/µl eicosene; peak N, nerolidol; peak C, cis-abienol; peak S, squalene.

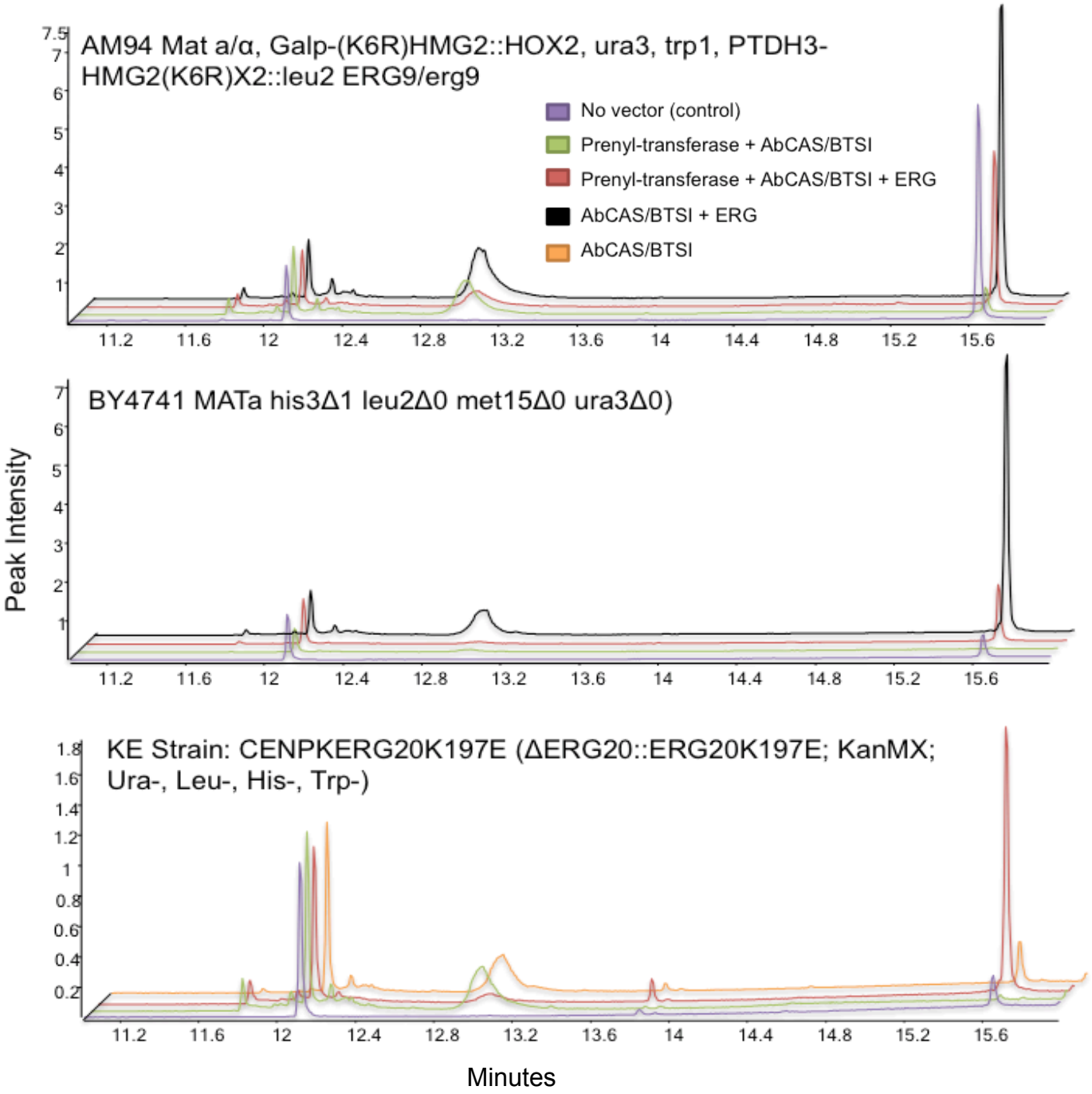


Figure 13. *cis*-abienol production by engineered baker's yeast strains (*Saccharomyces cerevisiae*) over expressing multicopy vectors. **(A)** Shown are total ion chromatograms (TIC) of extracts from three yeast strains expressing vectors for the production of cis-abienol. Vectors include pESC^{His} *Ab*CAS/BTS1, pGREG 505^{Leu}:ERG2.0, pESC^{Trp}Pg prenyl-transferase and a no vector control.

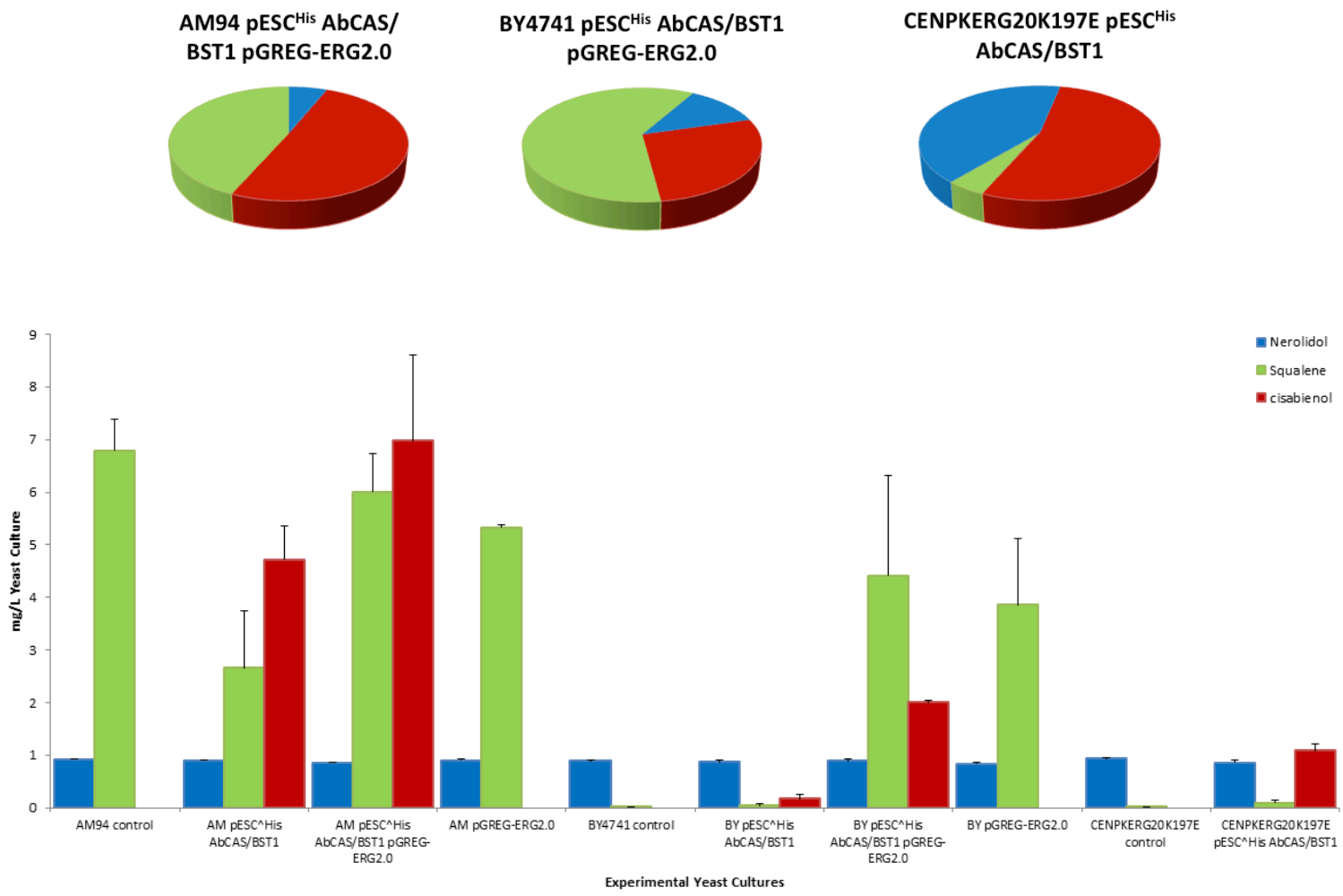


Figure 14. Induced expression of *cis*-abienol in baker's yeast (*Saccharomyces cerevisiae*). (A) Compounds extracted from three unique engineered yeast strains displayed as a percentage of total profile expression. **(B)** Amounts of yeast terpene metabolites produced in yeast expressed in mg/L yeast culture.

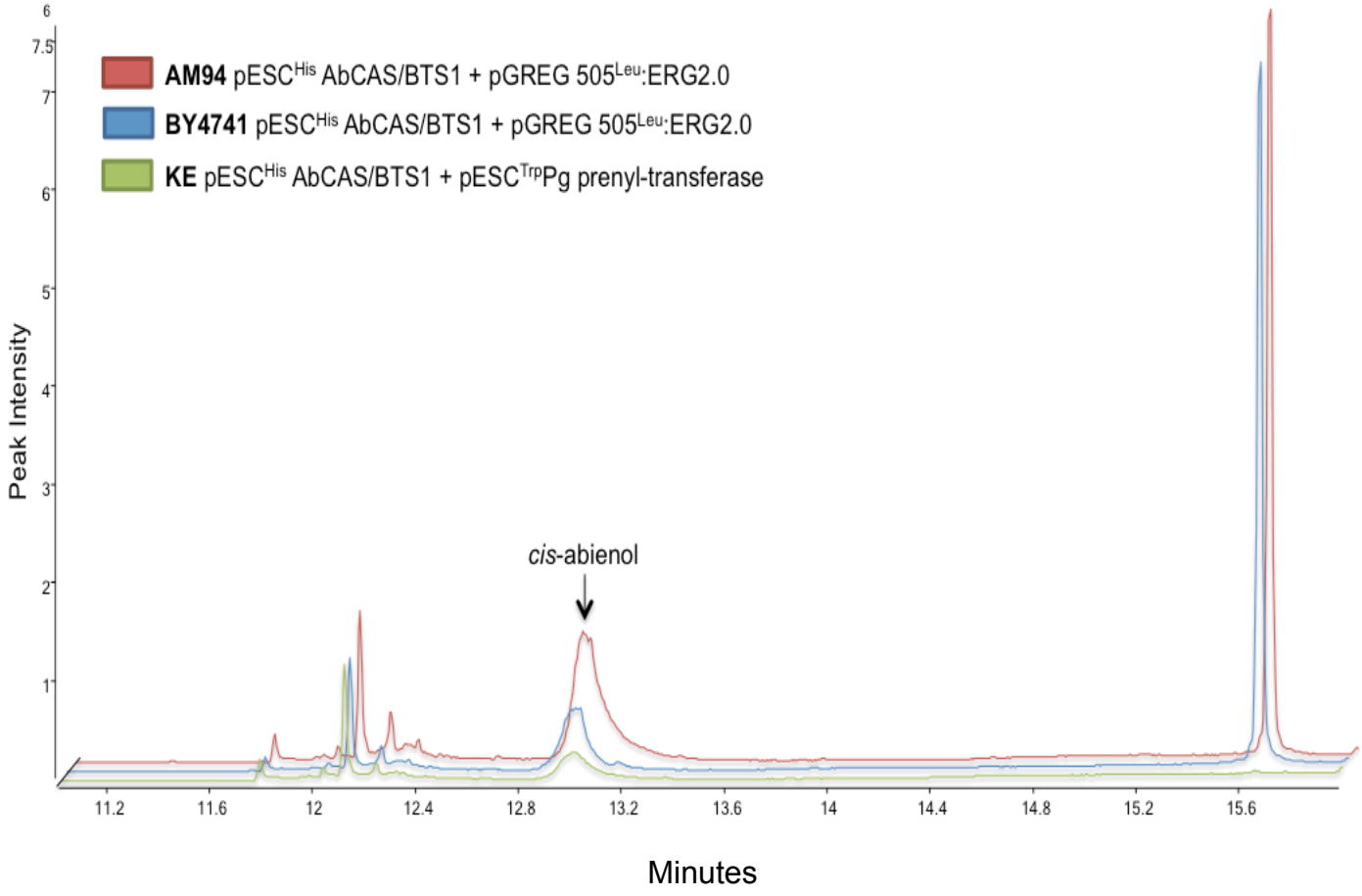


Figure 15. Induced expression of *cis*-abienol in baker's yeast (*Saccharomyces cerevisiae*). **(A)** Shown are total ion chromatograms (TIC) of extracts from three yeast strains expressing vectors for the production of cis-abienol. Shown are the strains with the vector combinations which resulted in the highest production of cis-abienol. For AM94 pESC^{His} *Ab*CAS/BTS1 + pGREG 505^{Leu}:ERG2.0, for BY4741 pESC^{His} *Ab*CAS/BTS1 + pGREG 505^{Leu}:ERG2.0 and KE strain pESC^{His} *Ab*CAS/BTS1 + pESC^{Trp}*Pg* prenyl-transferase.

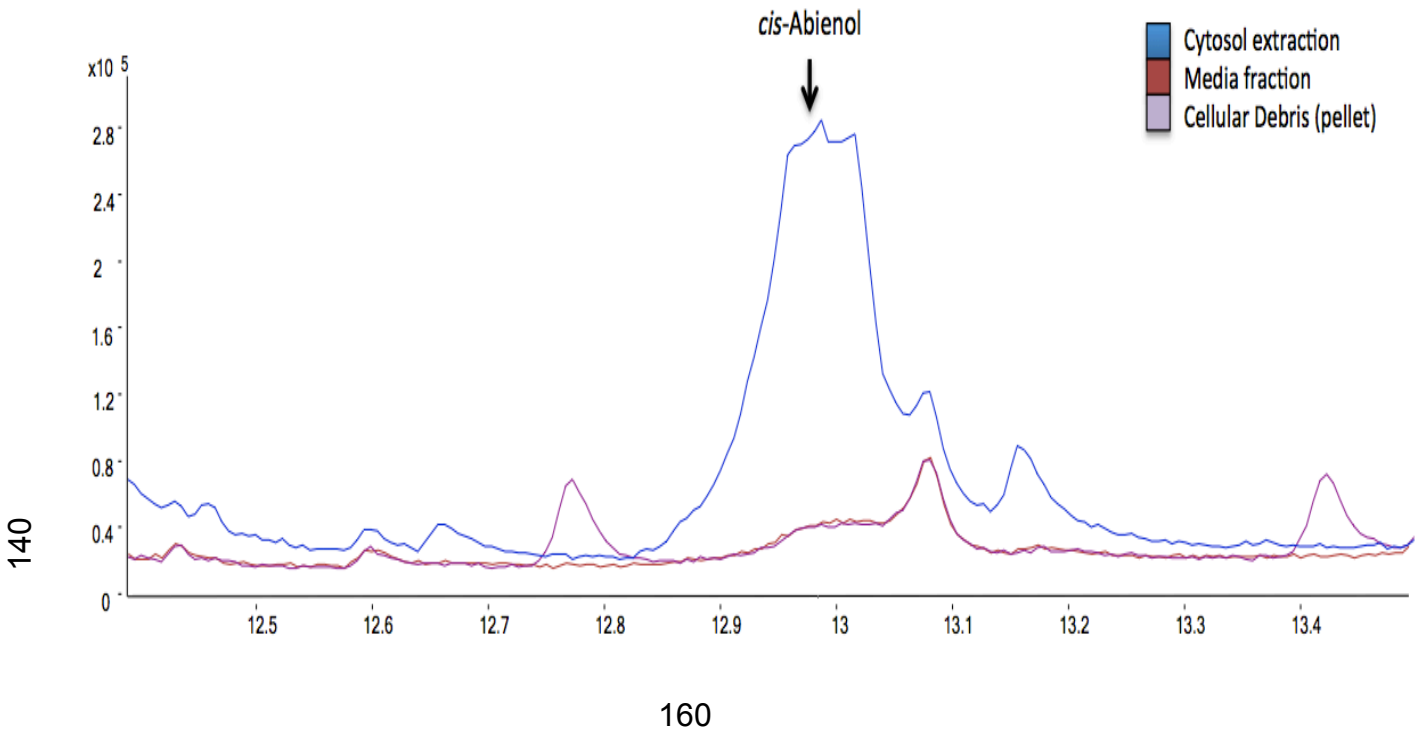


Figure 16. Localization of *cis*-abienol product following expression of transformed AM94 *S. cerevisiae* strain with vector *pESC^{His} AbCAS/BTS1*. The mass fragment pattern of this peak had the same diagnostic ions (*m/z* 290, 257, 191, 134) as that reported for *cis*-abienol. Blue line cytosol extraction, red line is liquid media fraction and purple is extraction from cellular debris.

3.4 Discussion

One of the potential benefits of producing *cis*-abienol in *S. cerevisiae* by expression of *Ab*CAS is that *cis*-abienol can be recovered from this system in a much purer form compared to extraction from balsam fir (*Abies balsamea)*. In addition, the productivity of the yeast systems can potentially be improved by strain engineering and eventually a production system could be scale up.

The three yeast strains used in these experiments were previously engineered for enhanced flux into and through the mevalonate pathway for increased terpenoid production. Here I specifically performed transformation with *Ab*CAS for the production of *cis*-abienol. I report the production of *cis*-abienol in three different strains of *S. cerevisiae* to produce titers (up to 12.8 mg L^{-1} in AM94 strain). *Cis*-abienol synthase gene was sourced from *A. balsamea* and was transformed on $pESC^{HIS}$ vector into the hosts, in some instances in the presence of other MVA pathway genes located on other selectable vectors such as $pESC^{Leu2d}$, $pESC^{Trp}$ and pGREG 505^{Leu}:ERG2.0. This resulted in unstable yeast strains that produced varying levels of *cis*-abienol. The synthesized *cis*-abienol is retained within the yeast (Figure 12) and could be recovered using a simple extraction and purification protocol. This confirms our hypothesis that *Ab*CAS can be expressed in yeast for the possible production of *cis*-abienol.

Out of the three engineered yeast strains, the highest producer of *cis*-abienol was AM94 strain, co-transformed with [$pESC^{His}$ *Ab*CAS/BTS1] + [pGREG-ERG2.0] (averaging 7 mgL^{-1}). However, this engineered strain also contained high levels of other terpenoid products, squalene and nerolidol. This engineered strain has a strongly up-regulated

terpenoid biosynthetic pathway including five copies of a gene that codes for the rate-limiting enzyme HMG-R. These multiple gene copies in conjunction with other up-regulated enzymes (ERG12 and ERG 13) are hypothesized to result in appreciable amount of GGPP precursor for the *Ab*CAS enzyme.

The KE (CENPKERG20K197E) was previously engineered for production of monoterpenes which utilize the high levels of GPP precursor that this strain produces (Vince Martin, unpublished). The ERG20 (k197E) mutant produces GPP:FPP at a 2:1 ratio. The KE strain came in as the lowest producer of *cis*-abienol ($1mgL^{-1}$), however, this strain possessed a much cleaner terpenoid background and *cis*-abienol was over 50% of the total products produced.

Future work to improve diterpene production in yeast should include engineering the *Ab*CAS enzyme to improve the substrate binding affinity and enzyme efficiency. Other similar compounds such as sclareol have been reported with up to ~1.5g/L in high-cell-density fermentation of *E.coli* (Schalk *et al.* 2012). Therefore, exploring other avenues of engineering could increase the *cis*-abienol to similar amounts as sclarerol.

Figure 16 shows the retention of *cis*-abienol within the cell which could elute to potential toxicity issues and therefore could result in cell death and overall decreased production of *cis*-abienol. To address this issue, one solution could be an investigation of potential transporters to move *cis*-abienol out of the cell and into the extracellular environment. For example, the *Nt*PDR1 transporter (from *Nicotiana tabacum*) is a plasma

membrane ABC transporter and has previously demonstrated the ability to move diterpenes across membranes (Crouzet et al. 1013) . An additional benefit to using diterpene transporters is that cell growth can continue as *cis*-abienol is being harvested. In contrast to our current system that requires the disruption of the cell in order to collect *cis*-abienol, and therefore renders the cell lysed and dead.

No host optimization was performed in these experiments, and considerable variability was observed in terms of diterpene production (~1 mg/L to ~13 mg/L in AM94 strain). Previous work has shown that one of the most beneficial ways to increase the levels of terpene production in yeast has been the optimization of yeast cell growth factors. For example, amorpha-4,11-diene strain led to production of > 40 g/L product (Westfall et. al 2012). Initially Westfall *et al.* 2012 reported 153 mg/L of amorpha-4, 11 – diene in shake flasks, but following the elimination of the utilization of galactose, codon optimization, changing the yeast feeding strategies and controlling oxygen and pH levels they were able to increase production 250-fold.

Techniques used in this study consisted of batch cultures that contained all the nutrients available to the yeast cells for the duration of the experiment. Future work could experiment with fed-batch cultures which are advantageous because one can control concentration of fed-substrate in the culture liquid at arbitrarily desired levels. In general, fed-batch culture may be superior to conventional batch culture when controlling concentrations of a nutrient (or nutrients) that affect the yield or productivity of the desired metabolite. In addition, using more controlled parameters such as monitoring pH and

oxygen levels, and cell density could all be important factors in determining the levels of terpenes produced from yeast.

Thus far all *Ab*CAS expression experiments have resulted in positive unstable yeast lines (yeast strains that are able to produce *cis*-abienol but the gene that encodes *Ab*CAS is not integrated into the yeast genome). Unstable lines possess a free plasmid and have the potential to loose the vector containing *Ab*CAS if the selection is removed. In the interest of creating a stable yeast strain (where the transformed DNA is integrated into the *S. cerevisiae* genome) the *Ab*CAS gene fused with BTS1 gene was cloned into the SB221 vector (figure 17). A colony PCR experiment confirmed the integration of *Ab*CAS/BTS1 into the SB221 vector between restriction sites AatII and SalI. The linear CAS/BTS1/URA3 was amplified using primer set “Recombo Fw AHD1” (GACCATCAAAGAAAATTAATGTGGCTGTGGTTTCAGGGTC ACATTTCCCCGAAAAGTGCCACCTG) and “Recombo rev URA”. (TTTTTTC GTCATTATAGAAATCATTACGACCGAGATTCCCAACTGATATAATTAAATTGA) in order to produce overhanging sequences that could be used for homolgous recombination into the URA3 site of the yeast genome. Yeast strains AM94, BY4741 and KE strain were transformed with the “FwRecombo-CAS/BTS1/URA3-Rev recombo” linear DNA fragment, then growth at 30°C for 2 days. Following transformation no growth was observed in transformed yeast strains but controls grew normally (data not shown). These results suggest two possibilities for unsuccessful stable yeast lines (1) the recombination of the linear fragment into the genome was unsuccessful, or (2) there were issues with the URA3

marker following transformation and therefore the yeast are unable to grow on the -URA3 plates.

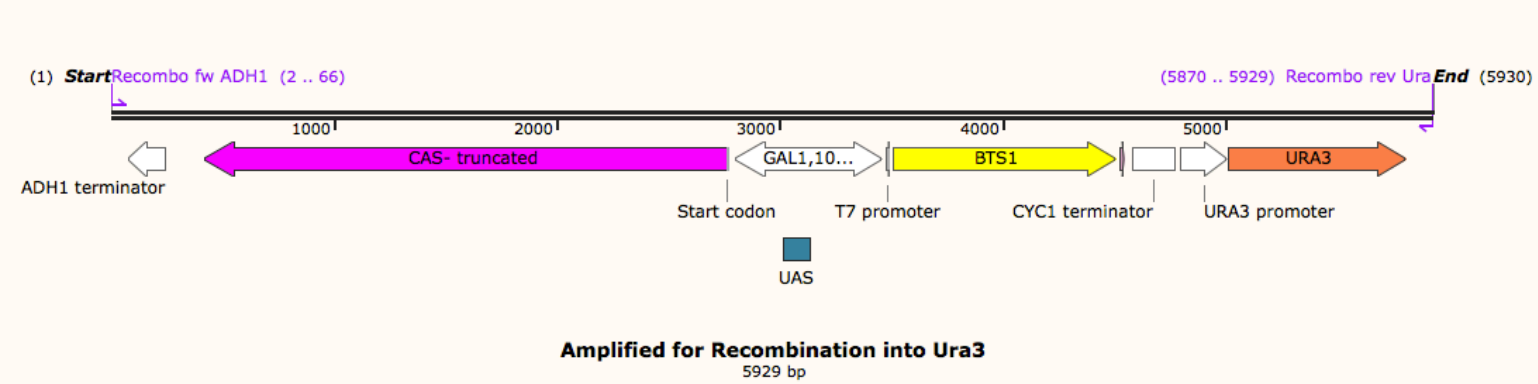


Figure 17. Schematic representation of the construction of the SB221 vector with AbCAS and BTS1 gene

Chapter 4: Expression of *cis*-Abienol Synthase in *Physcomitrella patens*

4.1 Introduction

Given the difficulties associated with the purity of *cis*-abienol from conifers and the expense of chemical synthesis, there is interest in the exploration of alternative production platforms. Plant cell factories may offer several advantages compared to engineered yeast strains such as the presence of plastids and the universal isoprenoid building blocks for *cis*-abienol biosynthesis. Additionally, a plant based expression host offers native codon usage, protein targeting and post-translational modifications more similar than other common expression hosts such as *E. coli* or yeast. The advantage of having post-translational modifications that are more similar to the *planta* kingdom is that some plant enzymes require modifications for activity or proper protein folding. It is for the reasons mentioned above that we chose to explore the transformation of *Physcomitrella patens* with the CAS gene.

The bryophyte *P. patens* (velvet mud moss) is a non-vascular plant that is being explored as a plant system for production of plant natural products such as pharmaceuticals and other industrially interesting compounds (Weissman and Leadlay 2005). *P. patens* can be propagated on solid or liquid minimal medium (without complex nutrients) on petri dishes and as suspension cultures, respectively. Figure 18 shows *P. patens* on petri dishes. The moss has a simple morphology and its life cycle is dominated by the photoautrophic haploid gametophytic generation, while the diploid sporophyte is short-lived and completely dependent on the gametophyte (Reski 1998). Under laboratory conditions it takes three months to complete a life cycle, but if the gametophores are

disrupted one can prevent the differentiation into adult tissue and *P.patens* can be kept in the haploid stage indefinitely (Frank et al. 2005).

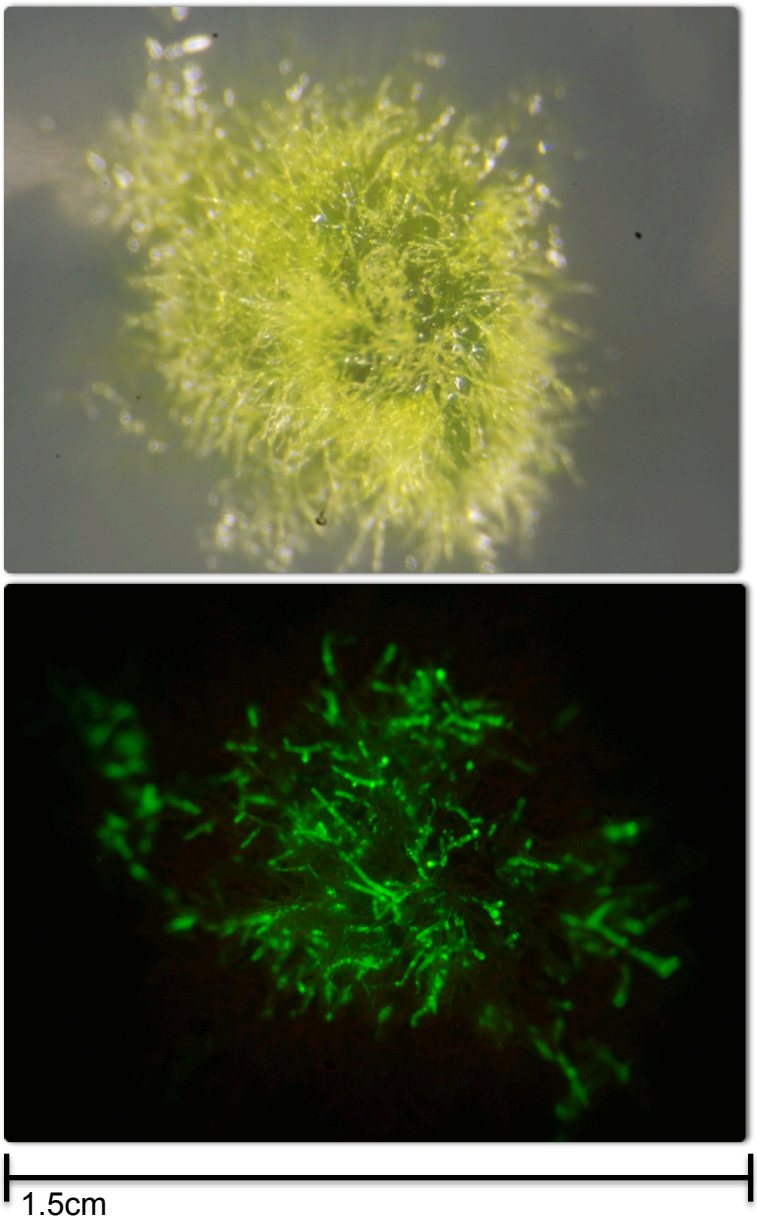


Figure 18: Fluorescent microscope image of *Physcomitrella patens* transformed with YFP. (A) in natural light and (B) UV light. The figure is 1.5cm x 1.5cm.

A low background of terpene metabolites in biological production hosts is considered a desirable characteristic. By decreasing the number of other biological compounds found within a host, the desired product is found in greater purity. *P. patens* has only two diterpenoid compounds, the secondary metabolites *ent*-kaurene and 16-hydroxykaurane, a common precursor for gibberellins (GAs). However, moss do not require GAs for growth. *ent*-Kaurene is synthesized via sequential cyclization steps of GGPP by *ent*-copalyl diphosphate synthase and *ent*-kaurene synthase (Sun and Kamiya 1994, Yamaguchi et al. 1996, and Yamaguchi et al. 1998). Previous research has generated a PpCPS/KS disruption mutant that does not accumulate *ent*-kaurene (Hayashi 2010). By knocking out the only diTPS gene (PpCPS/KS) in the *P. patens* genome, a strain is obtained that may allow for more GGPP substrate to be channeled to heterologous diTPSs. The particular PpCPS/KS knockout strain that was used in this work is identified as strain #29 provided by Bjorn Hamberger from The University of Copenhagen. The low-terpenoid background creates a unique in *planta* platform for production of *cis*-abienol.

To date, the best examples of metabolic engineering in *P. patens* have been performed in the field of glycoprotein production (Decker and Reski 2012). More recently, *Pa*LAS expression has been established in *P. patens* (S. Spanner-Bach PhD thesis University of Copenhagen). My objective is to establish expression of *cis*-abienol within the *P. patens* host.

4.2 Materials and Methods

USER Cloning Technique

pUNI33 is a USER compatible cloning vector (Nour-Eldin 2006). Using CAS_FL_USER_fw and CAS_USER_REV primers and AbCAS as a template the following PCR program: was used to clone the gene, initial denaturation at 98°C for 90sec followed by 29 cycles of denaturation at 98°C for 30 sec, annealing at 50°C for 30 sec and extension at 72°C for 2 min, and a final extension at 72°C for 8 min. Plasmid pUNI33-FL-AbCAS was then transformed into α-Select Competent Cells. Plasmids were recovered by QIAprep Spin Miniprep Kit as described in the manufactures handbook. The *Ab*CAS cDNA sequence was verified by complete sequencing using primers E7, E8 and internalCASi2. Prior to transformation into *Physcomitrella patents,* pUNI33-FL-AbCAS was linearized with Not-HF in NEB buffer #4 overnight at 37°C, followed by heat inactivation at 60°C for 10 min, then precipitated with isopropanol (0.7volumns IPA, centrifuge 30min at 15k rcf), wash with 70% EtOH (15min, centrifuge 30 min at 15k rcf), air dry then dissolve in 50 µl H_2O overnight.

Table 6: Primers used to construct pUNI33 vector for transformation into *P.patens*

CAS_FL_USER_fw	GGC TTA AUA TGG CCC TGC CTG TCT ATT C
CAS_USER_REV	GGT TTA AUT TAG GTA GCC GGC TCG AA
E7	AGA GGA CG ACCT GCA GGC
E8	GCA TGC GTC GAC GAG C
internalCASi2	CTG TTC TCG GAA GCA GTC AAA A

P.patens Tissue Culture and Gene Transformation

Protonemal *P. patens was* grown for 7 days on PHY B media (f0.8 g $Ca(NO_3)*4H_2O$, 0.25 g $MgSO_4*7H_2O$, 0.0125 g $FeSO_4*7H_2O$, 0.5 mL micro stam, 10mL $KH_2\ PO_4$, 0.5g NH_4 tartrate, 7 g agar and 0.6 mL hygromicine). *P. patens* grown on two separate petri dishes and was harvested 5 mL driselase (2 g driselace and top up to 100

mL with 8.5% D-mannitol). Next, the plant material was filtered using a wire mesh grate (100 uM, then 50 um filters). I then, centrifuged the protoplasts at 200 g for 5 min starting at a slow speed without a break at end of centrifugation to prevent the protoplast from rupturing. After discarding the supernatant, I resuspended the protoplasts in 5 mL ProtoWash (for 500 mL solution, 1M $CaCl_2$*$6H_2O$ and 8.5% D-mannitol. The pUNI33-*Ab*CAS DNA construct is then incubated with moss protoplasts in the presence of polyethylene glycol (PEG) and resuspend in 5 mL (8.5%) Mannitol-solution, then centrifuged at 300 g for 5 min. Again, I discarded the supernatant and resuspend in MMM solution. Next I transformed the moss by adding 30 ug of linearlized DNA (pUNI33-FL-AbCAS or pUNI33-t-AbCAS) and an additional 300 uL of protoplast regeneration medium for the bottom layer (PRMB) (3.5 g agar, 10 mM $CaCl_2$, 460 mg diammonium tartrate, 30 g D-mannitol, and top up to 500 mL BCD medium) suspension and 300 uL PEG. This mixture was Incubated in a 45°C water bath for 5 min, then incubated at room temperature for another 5 min. Next I added 300 uL 8.5% mannitol solution x4, gently mixed and inverted the tube, then, added 1 mL 8.5% mannitol solution x4. Finally I centrifuged at 200 g for 5 mins and discarded supernatant and resuspend in 500 uL 8.5% mannitol solution. I added 2.5 mL of molten PRMT (protoplast regeneration medium for top layer, 2 g agar, 10 mM $CaCl_2$, 460 mg diammonium tartrate, 40 g D-mannitol, and top up to 500 mL BCD medium) and dispensed 1 mL per petri dish with PRMB with cellophane. The transformed *P. patens* was then incubated for 7 days at room temperature and transfer on a selective media.

Metabolite Analysis

Terpenoids were extracted from samples of crushed tissue with 1 mL of diethyl ether for 16 hr at room temperature. Water was removed by the addition of anhydrous Na_2SO_4. GC-MS analysis was preformed on Agilent 6890N GC (Agilent Technologies Inc., Mississauga, Canada), 7683B series autosampler and a 5975 Inert XL MS Detector at 70 eV and 1 mL min^{-1} helium as a carrier gas using a DB-Wax Column with the following GC temperature program: initial temperature 250°C back inlet, pulsed splitless with a pulse pressure 25 psi for 0.25 min and constant flow of 1 mL min^{-1}. Scan (40-400 mass) and SIM mode (m/z 83, 151, 177, 257, 272, 275) were run simultaneously. Compound identification was achieved by comparison of mass spectra with those of authentic standards and reference mass spectral databases of the National Institute of Standards Technology MS Search Libraries (Wiley W9N08).

4.3 Results

*Ab*CAS was transformed into wildtype *P. patens* and into line #29 which lacks the endogenous diTPS. Hygromycin resistant *P. patens Ab*CAS transformants were isolated and grown as described previously (Nour-Eldin 2006). No phenotypic differences were observed between transgenic wild type *P. patens* and the untransformed wild type when examined with a microscope at 10-fold to 40-fold magnification (data not shown). I compared the untransformed wild-type *P. patens* and PpCPS/KS knockout *P. patens* lines (strain #29) by GC-MS analysis and confirmed that *P.patens* (WT) produces *ent*-kaurenic acid and that the #29 *P.patens* strain did not.

Following transformation with *Ab*CAS terpenoid extraction samples were collected from the WT moss and #29 lines then analyzed GC-MS, (Figure 19). The top GC trace in

figure 19 is displaying the chromatogram reading from *P.patens* (WT) *Ab*CAS which contains the internal standard and three additional products. As previously mentioned, *P.patens* produces *ent*-kaurene and 16-hydroxykaurane naturally. It is hypothesized that conditions within the GC-MS machine or environmental pH conditions have resulted in the addition of a third product (*ent-iso*kaurene, refered to as iso-kaurene). This third product is an isomer of ent-kaurene and is not a product of *Ab*CAS. The mass fragment pattern of the peak (2) had the same diagnostic ions (m/z *257, 272, 137, 229*) as that reported for kaurene and was therefore concluded that it was not *cis*-abienol. This hypothesis is further supported by the *Ab*CAS gene transformed into *P. patens* #29 line. The #29 moss line has the endogenous diterpene synthase gene knockout and no products (*ent*-kaurene, iso-kaurene and 16-hydroxykaurane) are observed in the GC-MS. trace.

In conclusion, the GC-MS analysis showed no presence of *cis*-abienol in either WT or #29. To verify the transformation process, yellow- fluorescent protein (YFP) was used and was observed under the microscope (figure 18).

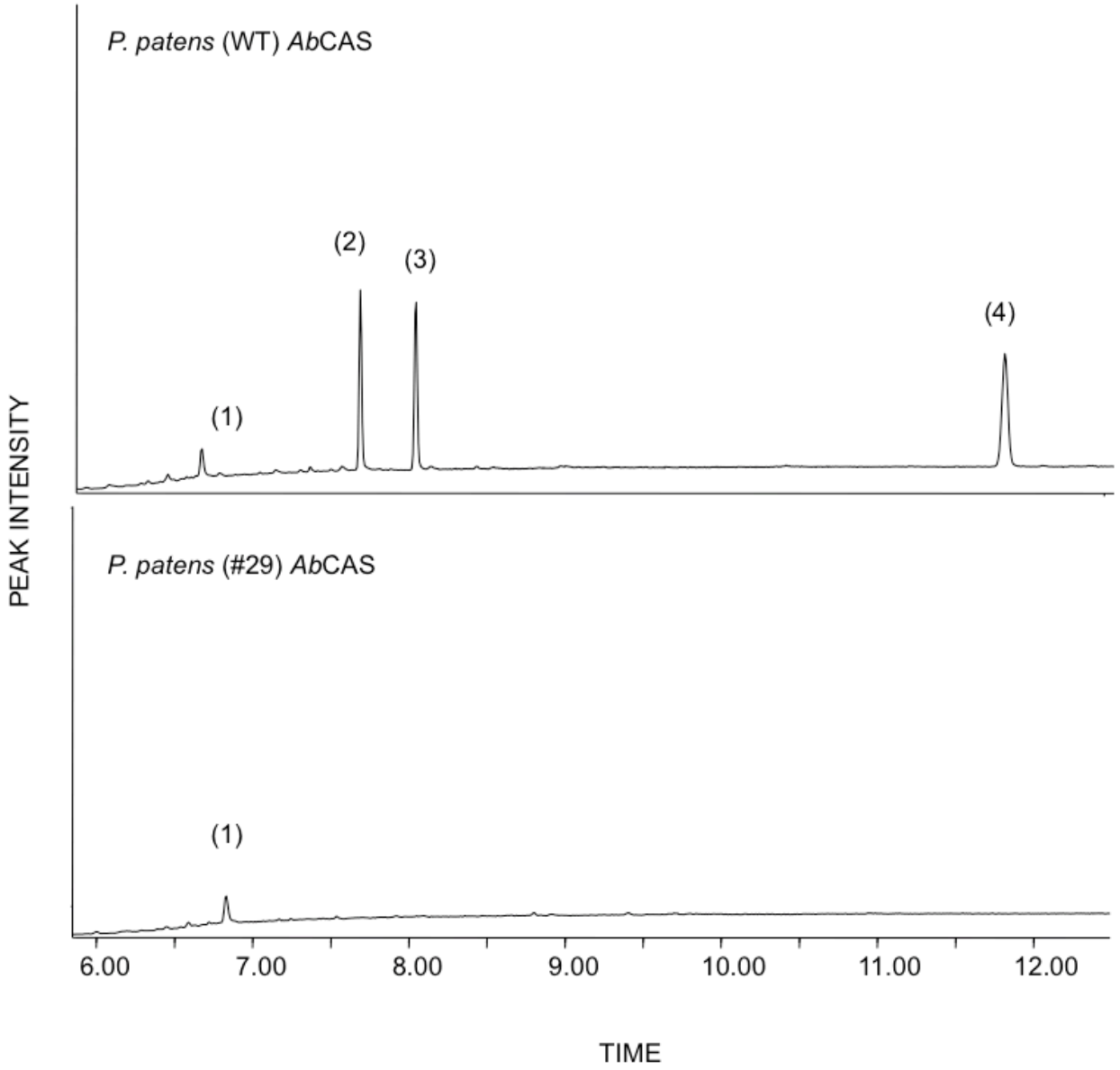


Figure 19. GC-MS trances for products from *P. patens* (WT) and (#29) strains transformed with *Ab*CAS. (1) Eicosene IS, (2) *iso*-kaurens, (3) *ent*-kaurene and (4) 16-hydroxykaurene. Shown is the total ion chromatograms (TIC) from an HP5 column.

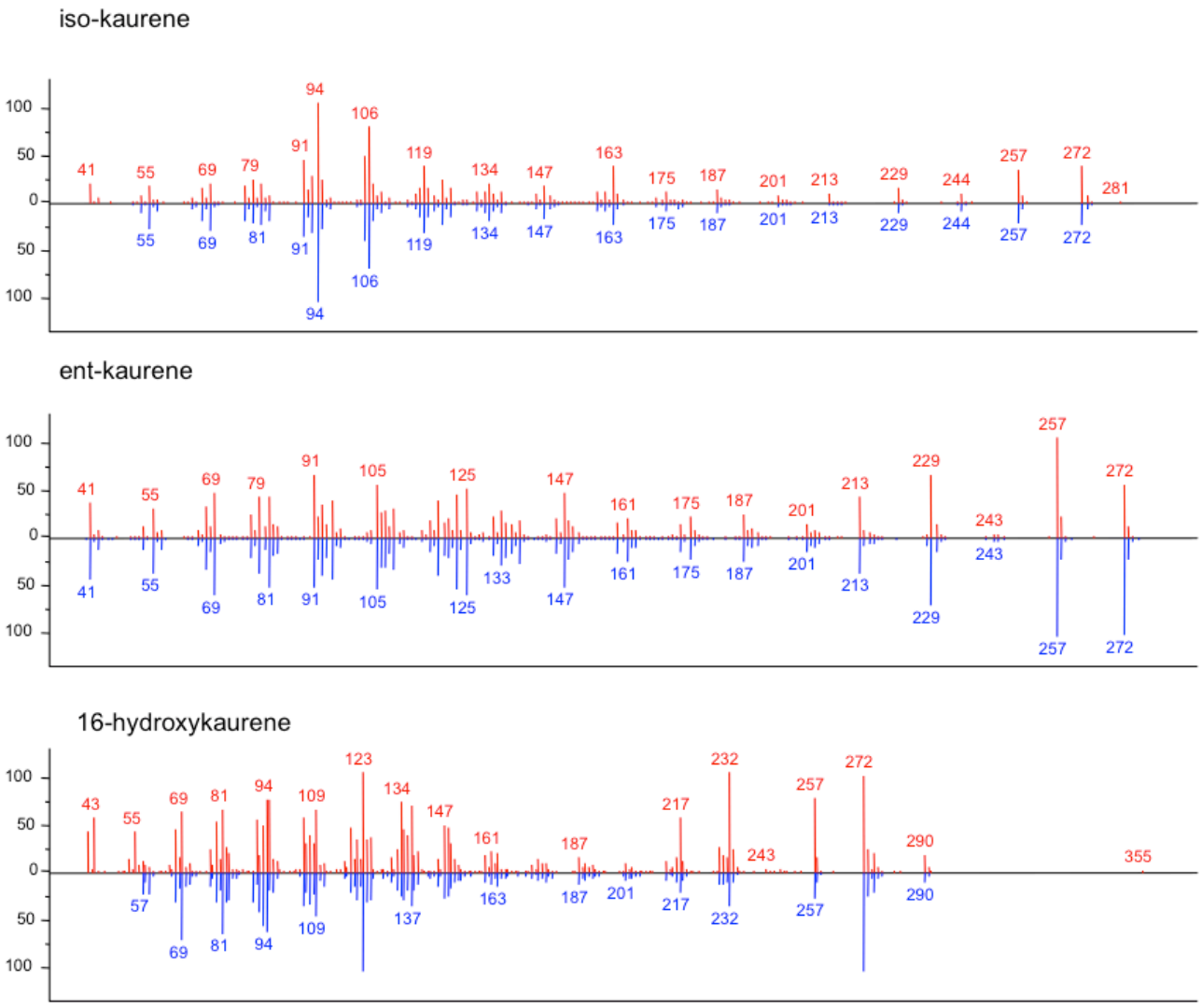


Figure 20. Experimental (red) and reference spectrum (blue) shown for iso-kaurene, ent-kaurene and 16-hydroxykaurene. Kaurene-type compounds are produced in *P. patens.* GC-MS analysis was performed on an Agilent HP5ms column with electronic ionization at 70 eV.

4.4 Discussion

The objective of this chapter was to explore the possibilities of producing *cis*-abienol in *P. patens*. Previous work by Anterola et al. (2009) demonstrated the production of the diterpenoid taxadiene in the *P. patents* system. The taxadiene synthase gene from *Taxus brevifolia* was used to produce taxa-4(5),11(12)-diene, the precursor of the anticancer drug paclitaxel (Taxol) (Anterola *et al.* 2009). This project was successful at producing 0.05% of the terpenoid product per mg of dry weight. On these grounds, *P. patens* has been described as an alternative platform for biotechnological production of plant terpenes. However our results suggest that this system needs much more attention in terms of developing an efficient platform for discovery and production of these secondary metabolites. We performed experiments to assess the feasibility of expressing *Ab*CAS in *P. patens*. Transformation was successful, but failed to produce the *cis*-abienol product. In addition we experienced lengthy moss regeneration times (6-8 weeks), high levels of contamination and unsuccessful expression of *cis*-abienol despite a total of seven independent attempts of *Ab*CAS transformation.

The DNA used in these experiments was a 7 kb plasmid (pUNI33) carrying a hygromycin resistance cassette, either supercoiled or linearized depending on the desired host mechanism; either stable or unstable lines. A stable line refers to the transformed DNA integrated into the *P. patens* genome, whereas unstable lines posses a free plasmid that is retain as long as there is selectable pressure (i.e. from the antibiotic). DNA can be integrated by homologous recombination or randomly if the transformed DNA lacks any sequences homologous to the genome. However, the efficiency of generating non-

targeted stable transformants is 1/10 that of the achieved when mediated by homologous recombination (Perroud and Quatrano 2006). Regardless of the stability of the lines, positive transformants can be selected within 6-8 weeks, which is consider fast compared to other more conventional plant systems.

To confirm the presence of *Ab*CAS in the pUNI33 cloning vector, universal primers (E7 and E8) were used to verify the correct gene sequence and orientation. To confirm the transformation process, yellow fluorescent protein (YFP) was used as a positive control. Since *Ab*CAS gene of interest is transformed on the same vector as the antibiotic (hygromycin), it is presumed that the ability of moss to grow in the presence of hygromycin demonstrates a successful transformation. However, despite the conformation of all these check point, no *cis*-abienol was detected by GC-MS analysis. We hypothesized that the failure was a result of expressing a truncated version of the *Ab*CAS gene (*Ab*CAS). This truncation (86 amino acids from the N-terminus of *Ab*CAS) was originally made for proper expression and folding of *Ab*CAS within the *E.coli* and yeast systems as described in Chapter 2 and 3. We hypothesized that the first 86 amino acids may be important for the correct localization of the *AbCAS* enzyme. Following trouble shooting a full length *Ab*CAS cDNA was transformed into *P. patens.* However, despite six separate transformations, positive controls and multiple check points, *cis*-abienol still did not appeared on the GC-MS readings. Therefore this portion of the project was completed and it was concluded with the unsuccessful production of *cis*-abienol in *P. patens*.

Based on the outcomes of this study it is apparent that *P. patens* and its vectors

would benefit from having more selectable markers and tightly controlled inducible promotes. The lack of selectable controls led to high levels of contamination during the project, something that is not observed when working with antibiotics in bacteria or essential amino acids in yeast. Although *P. patens* has a short regeneration period (6-8weeks) compared to other *in planta* systems, the regeneration time is still of considerable length and was a major bottleneck of the entire research project.

Chapter 5: Conclusions and Future Directions

Here we have established the successful biosynthesis of *cis*-abienol within three yeast strains. Future work should concentrate on the optimization of yeast growth conditions such as maintaining a constant pH, cell density and carbon source. In addition, we used site directed mutations of diTPSs to elucidate protein function and gain insight into the functional evolution of diTPSs of the conifer TPS-d subfamily. Lastly, our results suggest that other emerging model organisms (such as *P. patens*) should be approached with caution as to the practicality within the laboratory setting.

We also showed the presence of *cis*-abienol in *Abies lasiocarpa* and presented the work as a comparative diterpene profile of *Abies lasiocarpa* and *Abies balsamea* extracted from bark tissue (appendix 1). Future work should encompass gene mining for the enzyme responsible for producing the *cis*-abienol compound, followed by kinetic analysis of catalytic activity comparatively between the potentially new *Al*CAS enzyme and the kinetic results we reported here for *Ab*CAS.

Bibliography

Anterola A, Shanle E, Perroud P, Quatrano R. (2009) Production of taxa-4 (5), 11 (12)-diene by transgenic Physcomitrella patens. *Transgenic Research*. **18,** 655-660.

Ajikumar PK, Xiao W, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science*. **330,** 70-74.

Barrero AF, Alvarez-Manzaneda EJ, Altarejos J, Salido S, Ramos JM. (1993) Synthesis of< i> Ambrox® from (−)-sclareol and ( )-< i> cis-abienol. *Tetrahedron*. **49,** 10405-10412.

Barrero AF, Altarejos J, Alvarez-Manzaneda EJ, Ramos JM, Salido S. (1996) Synthesis of (±)-Ambrox from (E)-nerolidol and β-ionone via allylic alcohol [2, 3] sigmatropic rearrangement. *Journal of Organic Chemistry*. **61,** 2215-2218.

Bohlmann J, Meyer-Gauen G, Croteau R. (1998) Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proceedings of the National Academy of Sciences*. **95,** 4126-4133.

Bohlmann J, Keeling CI. (2008) Terpenoid biomaterials. *The Plant Journal*. **54,** 656-669.

Caniard A, Zerbe P, Legrand S, Cohade A, Valot N, Magnard JL, Bohlmann J, (2012) Discovery and functional characterization of two diterpene synthases for Sclareol biosynthesis in Salvia sclarea (L.) and their relevance for perfume manufacture. *BMC Plant Biology.* **12,** 119

Chen F, Tholl D, Bohlmann J, Pichersky E. (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *The Plant Journal*. **66,** 212-229.

Criswell J, Potter K, Shephard F, Beale MH, Peters RJ. (2012) A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase. *Organic Letters*. 14, 5828-5831.

Crouzet J, Roland J, Peeters E. (2013) NtPDR1, a plasma membrane ABC transporter from Nicotiana tabacum, is involved in diterpene transport. *Plant Molecular Biology*.1-12.

Decker EL, Reski R. (2008) Current achievements in the production of complex biopharmaceuticals with moss bioreactors. *Bioprocess and biosystems engineering*. **31**, 3-9.

Decker EL, Reski R. (2012) Glycoprotein production in moss bioreactors. *Plant Cell Reports*. **31,** 453-460.

Demetzos C, Kolocouris A, Anastasaki T. (2002) A simple and rapid method for the differentiation of C-13 manoyl oxide epimers in biologically important samples using GC–MS analysis supported with NMR spectroscopy and computational chemistry results. *Bioorganic and Medicinal Chemistry Letters*.**12,** 3605-3609.

Engels B, Dahm P, Jennewein S. (2008) Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic Eng*ineering. **10,** 201-206.

Frank W, Decker E, Reski R. (2005) Molecular tools to study Physcomitrella patens. *Plant Biology*. **7,** 220-227.

Garza RM, Tran PN, Hampton RY. (2009) Geranylgeranyl pyrophosphate is a potent regulator of HRD-dependent 3-Hydroxy-3-methylglutaryl-CoA reductase degradation in yeast. *Journal Biology Chemistry*. **284,** 35368-35380.

Gershenzon J, Dudareva N. (2007) The function of terpene natural products in the natural world. *Nature chemical biology*. **3,** 408-414.

Gietz D, St Jean A, Woods RA, Schiestl RH. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res*. **120,** 1425.

Guo Z, Severson R, Wagner G. (1994) Biosynthesis of the Diterpene< i> cis-Abienol in Cell-Free Extracts of Tobacco Trichomes. *Archives of Biochemistry and Biophysics*. 308, 103-108.

Guo Z, Wagner GJ. (1995) Biosynthesis of labdenediol and sclareol in cell-free extracts from trichomes of Nicotiana glutinosa. *Planta*. **197,** 627-632.

Guo J, Zhou YJ, Hillwig ML, Shen Y (2013) CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proceedings of the National Academy of Science.* **110,** 12108-13.

Hamberger B, Ohnishi T, Hamberger B, Séguin A, Bohlmann J. (2011) Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. *Plant Physiol*ogy. **157**, 1677-1695.

Hamberger B, Ohnishi T, Hamberger B, Seguin A, Bohlmann J. (2011) Evolution of Diterpene Metabolism: Sitka Spruce CYP720B4 Catalyzes Multiple Oxidations in Resin Acid Biosynthesis of Conifer Defense against Insects. *American Society of Plant Biologists.* 4, 1677

Hampton, R. and J. Rine. (1994) Regulated degradation of HMG-CoA reductase, an integral membrane protein of the endoplasmic reticulum, in yeast. *Journal of Cell Biology.* **125**, 299-312.

Hampton, R. and H. Bhakta. (1997) Ubiquitin-mediated regulation of 3-hydroxy-3-methylglutaryl-CoA reductase. *Proceedings of the National Academy of Science USA* **94**, 12944-12948.

Hayashi K, Horie K, Hiwatashi Y, Kawaide H, Yamaguchi S, Hanada A, Nakashima T, Nakajima M, Mander L, Yamane H, Hasebe M, Nozaki H. (2010) Endogenous diterpenes derived from ent-kaurene, a common gibberellin precursor, regulate protonema differentiation of the moss Physcomitrella patens. *Plant Physiology*. **153,** 1085-1097.

Healthcare G. GST SpinTrap Purification Module.

Hern A, Ruiz MT. (2008) An EXCEL template for calculation of enzyme kinetic parameters by non-linear regression. *Bioinformatics*. 14, 227-228.

Hong K, Nielsen J. (2012) Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. *Cellular and Molecular Life Sciences*. **69,** 2671-2690.

Ignea C, Trikka FA, Kourtzelis I, Argiriou A, Kanellis AK, Kampranis SC, Makris AM. (2012) Positive genetic interactors of HMG2 identify a new set of genetic perturbations for improving sesquiterpene production in Saccharomyces cerevisiae. *Microbial cell factories*. **11,** 1-16.

Jansen G, Wu C, Schade B, Thomas DY, Whiteway M. (2005) Drag&Drop cloning in yeast. *Gene*. **344,** 43-51.

Kai G, Xu H, Zhou C, Liao P, Xiao J, Luo X, You L, Zhang L. (2011) Metabolic engineering tanshinone biosynthetic pathway in< i> Salvia miltiorrhiza hairy root cultures. *Metabolic Eng*ineering. **13,** 319-327.

KirbyJ, Keasling J. (2009) Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. *Annual Review of Plant Biology.* 60, 335-355.

Keeling CI, Bohlmann J. (2006) Diterpene resin acids in conifers. *Phytochemistry*. **67,** 2415-2423.

Keeling CI, Weisshaar S, Lin RP, Bohlmann J. (2008) Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proceedings of the National Academy of Sciences*. 105, 1085-1090.

Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J. (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (Picea spp.). *BMC plant biology*. **11,** 43.

Keeling CI, Bohlmann J. (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defense of conifers against insects and pathogens. *New Phytologist.* **4,** 657-75.

Keeling CI, Madilao LL, Zerbe P, Dullat HK, Bohlmann J. (2011) The primary diterpene synthase products of Picea abies levopimaradiene/abietadiene synthase (PaLAS) are epimers of a thermally unstable diterpenol. *Journal of Biological Chemistry*. **286,** 21145-21153.

Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW. (2010) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature*. **469,** 116-120.

Köksal M, Hu H, Coates RM, Peters RJ, Christianson DW. (2011) Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. *Nature Chemical Biology*. **7,** 431-433.

Kolesnikova MD, Obermeyer AC, Wilson WK, Lynch DA, Xiong Q, Matsuda SP. (2007) Stereochemistry of water addition in triterpene synthesis: the structure of arabidiol. *Organic Letters*. **9,** 2183-2186.

Kolosova N and Bohlmann J. (2012) Conifer defense against insects and fungal pathogens. Submitted for publication in: "Growth and Defence in Plants: Resource Allocation at Multiple Scales, Editor R. Matyssek, Springer.

Leonard E, Ajikumar PK, Thayer K, Xiao WH, Mo JD, Tidor B, Stephanopoulos G, Prather KL. (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proceedings of the National Academy of Science of USA.* **107,** 13654- 135649

Lichtenthaler HK. (1999) The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annual review of plant biology*. **50**, 47-65.

Martin DM, Fäldt J, Bohlmann J. (2004) Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol*ogy. **135,** 1908-1927.

Miyazawa M, Horiuchi E, Kawata J. (2008) Components of the essential oil from Adenophora triphylla var. japonica. *Journal of Essential Oil Research*. **20,** 125-127.

Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ. (2010) Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of

MEV and MEP isoprenoid precursor pathway engineering. *Applied Microbiology and Biotechnology*. **85,** 1893-1906.

Nielsen J, Larsson C, van Maris A, Pronk J. (2013) Metabolic engineering of yeast for production of fuels and chemicals. *Current Opinion in Biotechnology.* **3,** 398-404

Nour-Eldin HH, Hansen BG, Nørholm MH, Jensen JK, Halkier BA. (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research*. 34, 122-122.

Olszewski N, Sun T, Gubler F. (2002) Gibberellin signaling biosynthesis, catabolism, and response pathways. *The Plant Cell Online*. **14**, 61-80.

Paddon C. and Keasling J. (2014) Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature Reviews Microbiology

Paradise EM, Kirby J, Chan R, Keasling JD. (2008) Redirection of flux through the FPP branch-point in Saccharomyces cerevisiae by down-regulating squalene synthase. *Biotechnology Bioengineering*. **100,** 371-378.

Perroud P, Quatrano RS. (2006) The role of ARPC4 in tip growth and alignment of the polar axis in filaments of Physcomitrella patens. *Cell Motility Cytoskeleton*. **63,** 162-171.

Peters, R.J. (2010) Two rings in them all: The labdane-related diterpenoids. *Natural Product Report*. **27,** 1521-1530

Peters RJ, Flory JE, Jetter R, Ravn MM, Lee H, Coates RM, Croteau RB. (2000) Abietadiene synthase from grand fir (Abies grandis): characterization and mechanism of action of the “pseudomature” recombinant enzyme. *Biochemistry.* **39,** 15592-15602.

Peters R, Ravn M, Coates R, Croteau R. (2001) Bifuncational Abietadiene Synthase: Free Diffusice Transfer of the (+) Copalyl Diphosphate Intermediate between Two Distinct Active Site. Journal of the American Chemicals Society. **123,** 8974.

Prigge MJ, Bezanilla M. (2010) Evolutionary crossroads in developmental biology: Physcomitrella patens. *Development*. **137,** 3535-3543.

Prisic S, Xu J, Coates RM, Peters RJ. (2007) Probing the role of the DXDD motif in class II diterpene cyclases. *ChemBioChem*. 8, 869-874.

Reski R. (1998) Development, genetics and molecular biology of mosses. *Botanica Acta*. **111,** 1-15.

Ro D, Bohlmann J. (2006) Diterpene resin acid biosynthesis in loblolly pine (< i> Pinus taeda): Functional characterization of abietadiene/levopimaradiene synthase (< i> PtTPS-< i> LAS) cDNA and subcellular targeting of PtTPS-LAS and abietadienol/abietadienal oxidase (PtAO, CYP720B1). *Phytochemistry*. **67,** 1572-1578.

Ro D, Arimura G, Lau SY, Piers E, Bohlmann J. (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proceedings of the National Academy of Sciences*. **102,** 8060-8065.

Ro D, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Change MC, Withers ST, Shiba Y, Sarpong R, Keasling JD. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. **440,** 940-943.

Robert JA, Madilao LL, White R, Yanchuk A, King J, Bohlmann J. (2010) Terpenoid metabolite profiling in Sitka spruce identifies association of dehydroabietic acid,(-)-3-carene, and terpinolene with resistance against white pine weevil. *Botany*. **88,** 810-820.

Sallaud C, Giacalone C, Topfer R, Goepfert S, Bakaher N, Rosti S, Tissier. (2012) Characterization of two genes for the biosynthesis of the labdane diterpene z-abienol in tobacco (Nicotiana tabacum) glandular trichomes. *The Plant Journal.* **74,** 713

Schaefer DG, Zrÿd J. (1997) Efficient gene targeting in the moss Physcomitrella patens. *The Plant Journal*. **11,** 1195-1206.

Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L. (2012) Toward a biosynthetic route to sclareol and amber odorants. *Journal American Chemical Society*. **134,** 18900-18903.

Schiff PB, Fant J, Horwitz SB. (1979) Promotion of microtubule assembly in vitro by taxol. *Nature.* **277,** 665-667.

Schmiderer C, Grassi P, Novak J, Weber M, Franz C. (2008) Diversity of essential oil glands of clary sage (Salvia sclarea L., Lamiaceae). *Plant Biology*. **10,** 433-440.

Schaefer DG, Zrÿd J. (2001) The moss Physcomitrella patens, now and then. *Plant Physiology*. **127,** 1430-1438.

Sun T, Kamiya Y. (1994) The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. *The Plant Cell Online*. **6,** 1509-1518.

Taiz L and Zeiger E. (2006) Plant Physiology 4th edition. Sinauer Associates Inc.

Valdes L, Mislankar S, Paul A.(1987) Coleus barbatus (C. forskohlii)(Lamiaceae) and the potential new drug forskolin (Coleonol). *Economic Botany*. **41,** 474-483.

Vogel BS, Wildung MR, Vogel G, Croteau R. (1996) Abietadiene synthase from grand fir (*Abies grandis*) cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *Journal of Biological Chemistry*. **271,** 23262-23268.

Weissman KJ, Leadlay PF. (2005) Combinatorial biosynthesis of reduced polyketides. *Nature reviews microbiology*. **3,** 925-936.

Wendt KU, Schulz GE. (1998) Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes. *Structure*. 6, 127-133.

Westfall PJ, Pitera DJ, Lenihan JR, Eng D, Woolard FX, Regentin R, Horning T, Tsurutw H, Melis DJ, Owens A, Fickes S, Diola D, Benjamin KR, Keasling JD, Leavell MD, McPhee DJ, Renninger NS, Newman JD, Paddon CJ (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proceedings of the National Academy of Sciences*. **109,** 111-118.

Yamaguchi S, Saito T, Abe H, Yamane H, Murofushi N, Kamiya Y. (1996) Molecular cloning and characterization of a cDNA encoding the gibberellin biosynthetic enzyme ent-kaurene synthase B from pumpkin (Cucurbita maxima L.). *The Plant Journal*. **10,** 203-213.

Yamaguchi S, Sun T, Kawaide H, Kamiya Y. (1998) The GA2 Locus of Arabidopsis thalianaEncodes ent-Kaurene Synthase of Gibberellin Biosynthesis. *Plant Physiol*ogy. **116,** 1271-1278.

Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Britton R, Bohlmann J. (2012) Bifunctional cis-abienol synthase from Abies balsamea discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *Journal of Biological Chemistry*. **287,** 12121-12131.

Zhou K, Gao Y, Hoy JA, Mann FM, Honzatko RB, Peters RJ. (2012) Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from Abies grandis. *Journal of Biological Chemistry*. **287,** 6840-6850.

Zulak KG, Dullat HK, Keeling CI, Lippert D, Bohlmann J**.** (2010) Immunofluorescence localization of levopimaradiene/abietadiene synthase in methyl jasmonate treated stems of Sitka spruce (Picea sitchensis) shows activation of diterpenoid biosynthesis in cortical and developing traumatic resin ducts. *Phytochemistry*. **71,** 1695-1699.

Appendix 1

Diterpene	Metabolite abundance [µg/g dry weight]	
	Abies lasiocarpa	*Abies balsamea*
cis-abienol	14035 ± 9340	7343 ± 2105
Pimaric acid	1666 ± 741	86 ± 39
Sandaraco pimaric acid	947 ± 244	51 ± 8
Levopimaric acid	876 ± 946	986 ± 289
Palustric acid	20765 ± 7349	642 ± 121
Isopimaric acid	3652 ± 740	589 ± 585
(Dehydro)-Abietic acid (pooled)	26647 ± 12558	604 ± 162
Neoabietic acid	12882 ± 4521	144 ± 40

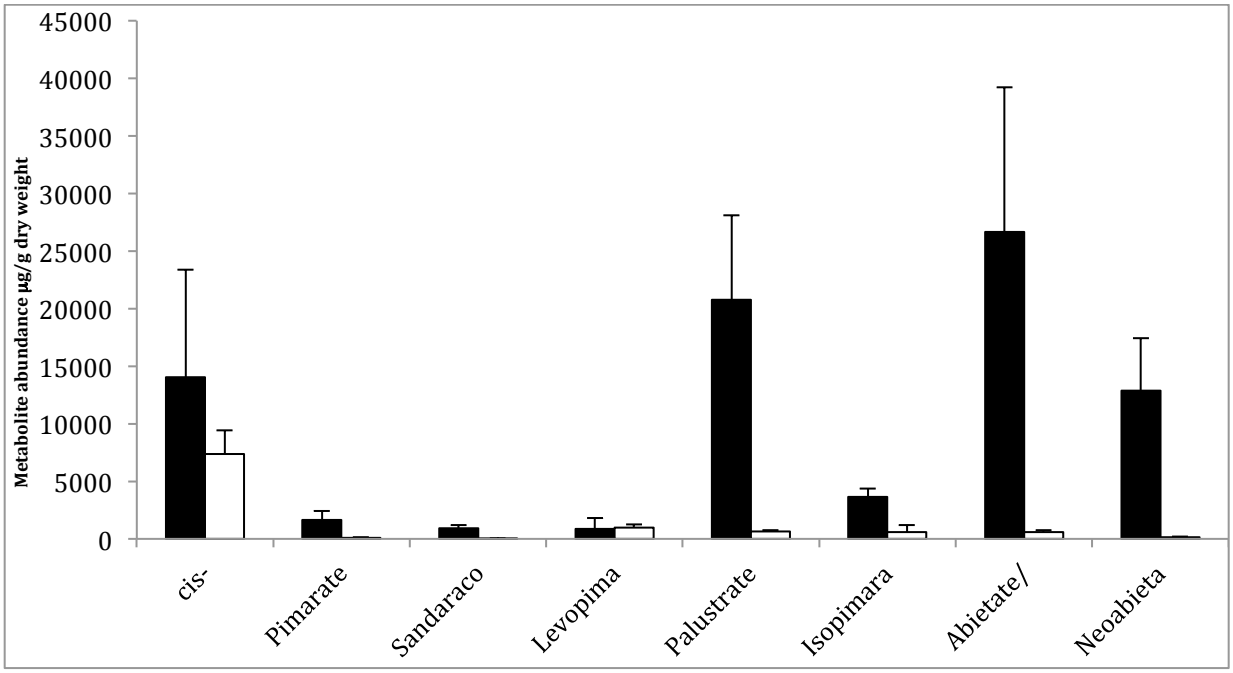


Figure 21. Comparative diterpene profile of *Abies lasiocarpa* and *Abies balsamea* extracted from bark tissue.

Appendix 2

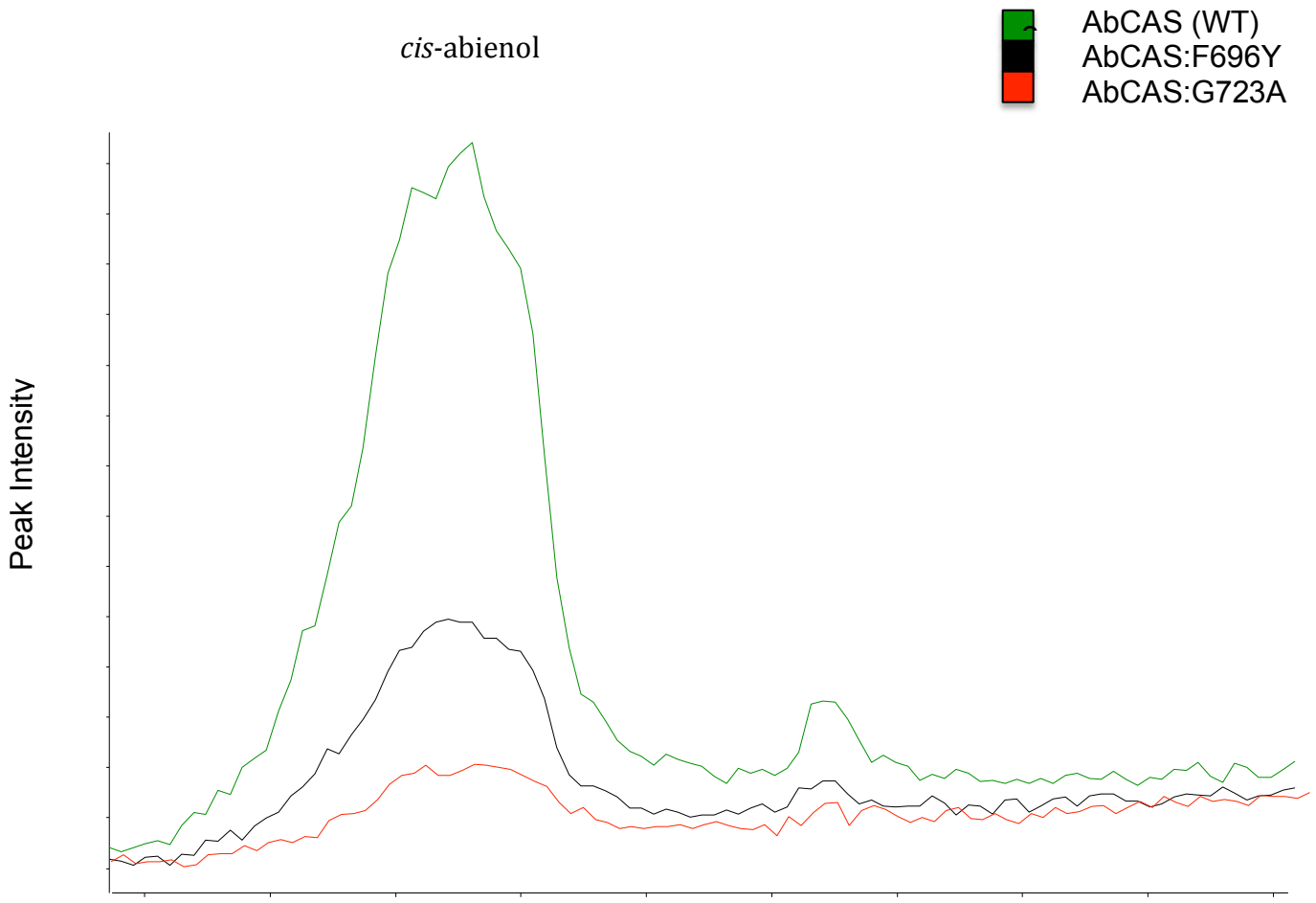
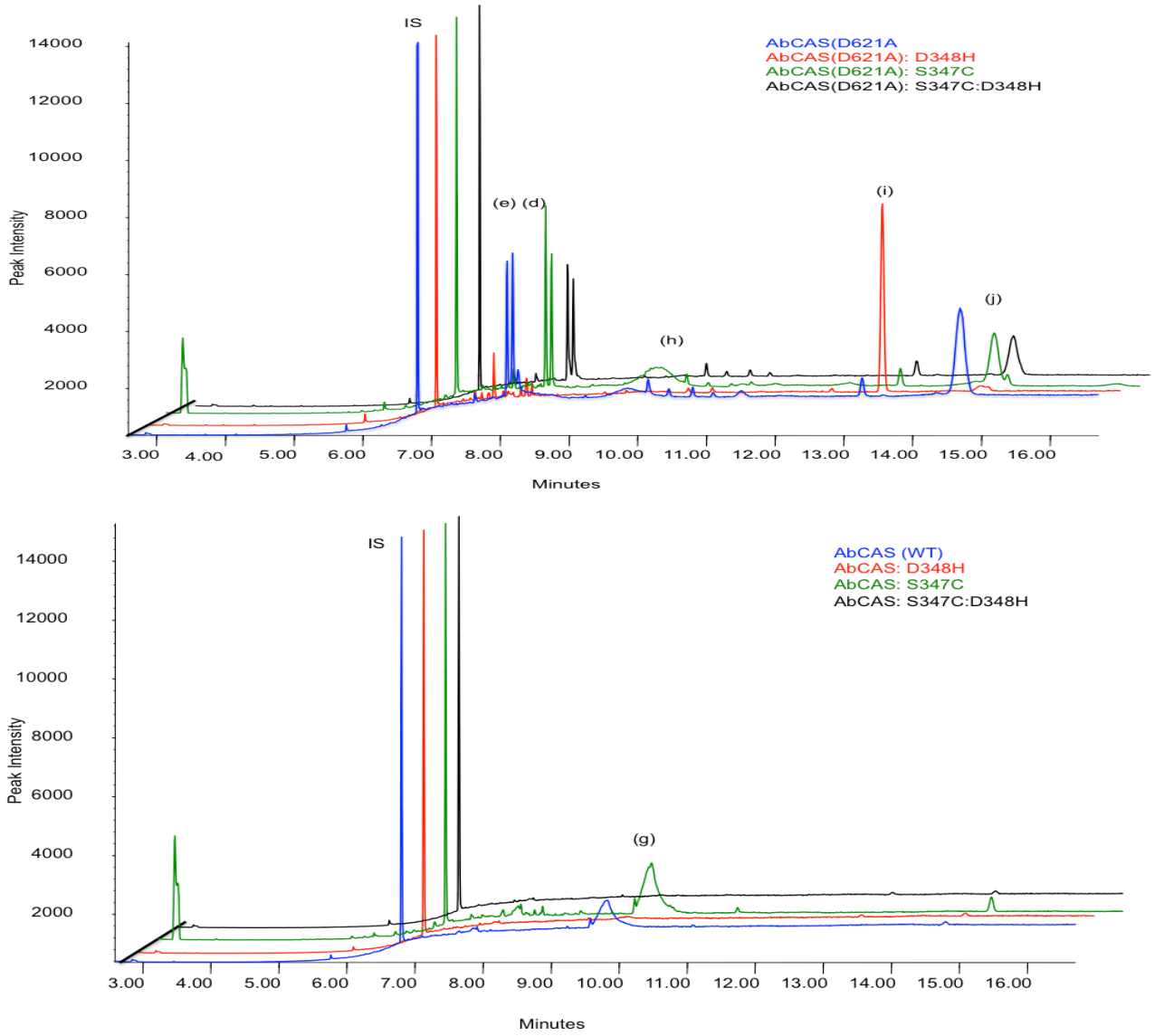


Figure 22. Site-directed mutagenesis in the class-I active site of *Ab*CAS: Product Profile Composition of *Ab*CAS Phe-696 and Gly-723.Green *Ab*CAS wildtype (WT), black *Ab*CAS:F696Y, red *Ab*CAS:G723A. Eicosene internal standard (IS) 2uM.

Appendix 3


IS
(e) (d)
(h)
(i)
(j)
AbCAS(D621A
AbCAS(D621A): D348H
AbCAS(D621A): S347C
AbCAS(D621A): S347C:D348H
Peak Intensity
Minutes
IS
(g)
AbCAS (WT)
AbCAS: D348H
AbCAS: S347C
AbCAS: S347C:D348H
Peak Intensity
Minutes

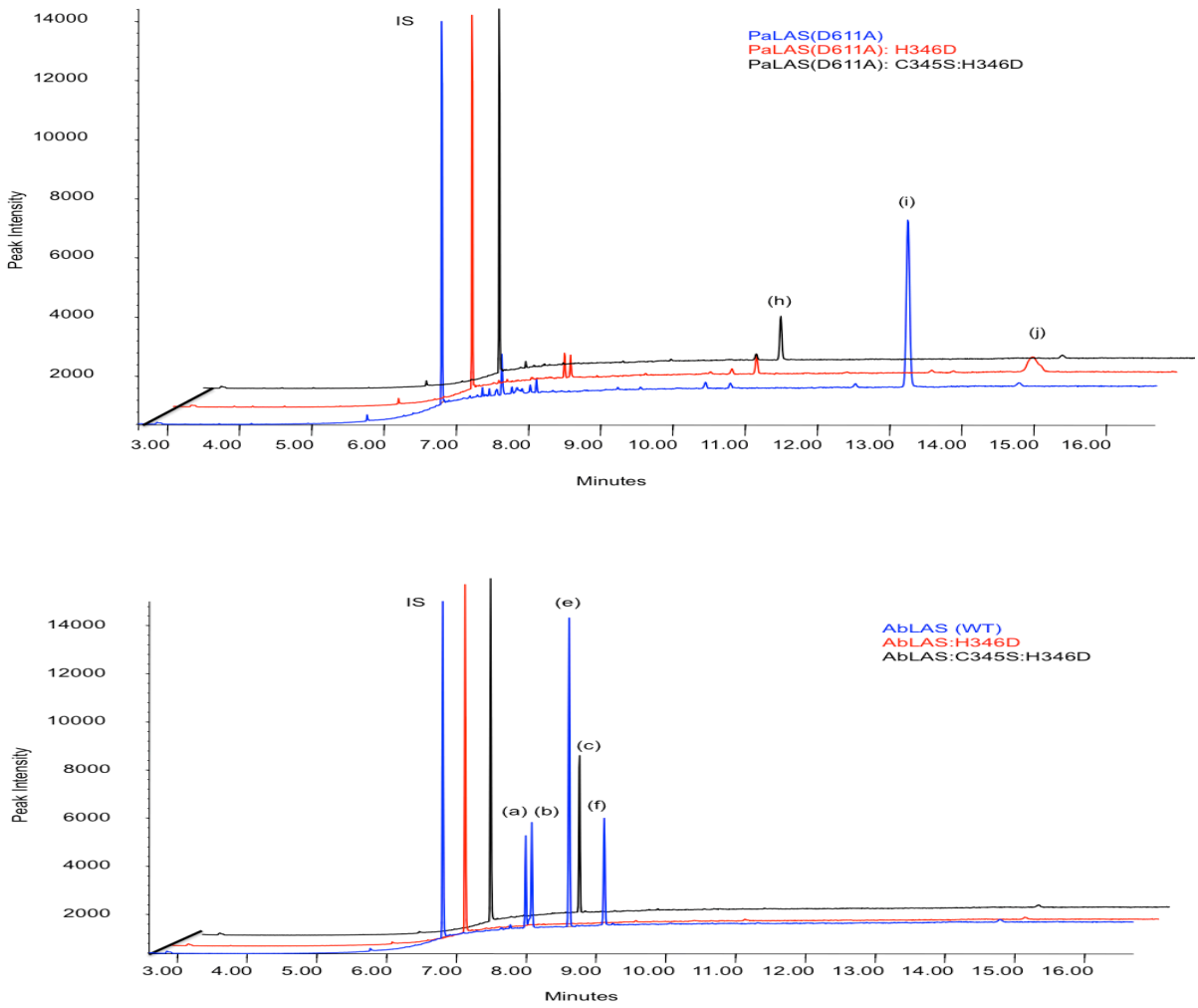


Figure 21. Activity of diTSPs *Ab*CAS (WT), *Ab*LAS (WT), and their protein variants. Individually scaled SIM chromatograms of reaction products from in vitro assays with purified recombinant enzymes using GGPP as a substrate. GC-MS analysis was performed on an Agilent HP5ms column with electronic ionization at 70 eV. Enzymatic activity assays were confirmed with three independent experiments, Peak IS, internal standard 2 uM 1-eicosene, peak (a) Palustradiene. peak (b) Levopimaradiene, peak (c) ent-manoyl oxide, peak (d) epi-manoyl oxide, peak (e) Abietadiene, peak (f) Neoabietadiene, peak (g) *cis*-abienol synthase, peak (h) GGPP, peak (i) CPP, and peak (j) LPP