2011

ELUCIDATING THE BIOCHEMICAL WIZARDRY OF TRITERPENE METABOLISM IN *BOTROYCOCCUS BRAUNII*

Thomas Daniel Niehaus
University of Kentucky, tomniehaus@uky.edu

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Thomas Daniel Niehaus, Student

Dr. Joe Chappell, Major Professor

Dr. Arthur G. Hunt, Director of Graduate Studies

ELUCIDATING THE BIOCHEMICAL WIZARDRY OF TRITERPENE METABOLISM IN *BOTROYCOCCUS BRAUNII*

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Agriculture
at the University of Kentucky

By
Thomas Daniel Niehaus

Lexington, Kentucky

Director: Dr. Joe Chappell, Professor of Plant Physiology

Lexington, Kentucky

2011

ABSTRACT OF DISSERTATION

ELUCIDATING THE BIOCHEMICAL WIZARDRY OF
TRITERPENE METABOLISM IN *BOTROYCOCCUS BRAUNII*

B. braunii is a green alga that has attracted attention as a potential renewable fuel source due to its high oil content and the archeological record of its unique contribution to oil and coal shales. Three extant chemotypes of *B. braunii* have been described, namely race A, race B, and race L, which accumulate alkadienes and alkatrienes, botryococcene and squalene and their methylated derivatives, and lycopadiene, respectively. The methylated triterpenes, particularly botryococcenes, produced by race B can be efficiently converted to high quality combustible fuels and other petrochemicals; however, botryococcene biosynthesis has remained enigmatic.

It has been suggested that botryococcene biosynthesis could resemble that of squalene, arising from an initial condensation of two molecules of farnesyl diphosphate (FPP) to form pre-squalene diphosphate (PSPP), which then undergoes a reductive rearrangement to form squalene, or in an alternative reductive rearrangement, botryococcene. Based on the proposed similarities, we predicted that a botryococcene synthase would resemble squalene synthase and hence, isolated squalene synthase-like genes from *B. braunii* race B. While *B. braunii* does harbor at least one typical squalene synthase, none of the other three squalene synthase-like (SSL) genes encode for botryococcene biosynthesis directly. SSL-1 catalyzes the biosynthesis of PSPP and SSL-2 the biosynthesis of bisfarnesyl ether and to a lesser extent squalene, while SSL-3 does not appear able to directly utilize FPP as a substrate. However, when SSL-1 is combined with either SSL-2 or SSL-3, *in vivo* and *in vitro*, robust squalene or botryococcene biosynthesis was observed, respectively. These findings were unexpected because squalene synthase, an ancient and likely progenitor to the other *Botryococcus* triterpene synthases, catalyzes a two-step reaction within a single enzyme unit without intermediate release, yet in *B. braunii,* these activities appear to have separated and evolved inter-dependently for specialized triterpene production. Expression of various configurations of the SSL genes in TN-7 yeast demonstrates that botryococcene can be efficiently produced in a heterologous host.

Additionally, three triterpene methyltransferase (TMTs) were isolated which efficiently catalyze the transfer of a methyl group from S-adenosyl methionine (SAM) to either squalene (TMT-1 and TMT-2) or botryococcene (TMT-3) *in vivo* and *in vitro*. Co-

expression of the various TMT genes with either squalene synthase or botryococcene synthase in TN-7 yeast resulted in the accumulation of C_{31} and C_{32} methyl derivatives of squalene or botryococcene, demonstrating their potential for heterologous production. The methylation sites were determined by NMR spectroscopy to be identical to C_{31} and C_{32} methyl-derivatives of squalene or botryococcene observed in *B. braunii* race B.

Expression studies of various heterologous squalene synthase genes in *S. cerevisiae* corroborated an earlier but surprising observation reported in the literature. While the squalene synthase gene of *S. cerevisiae* was able to complement an erg9 (squalene synthase) knockout in yeast, squalene synthase genes from plants and animals were not. Chemical profiles revealed that squalene accumulated to significant levels in yeast expressing the squalene synthase of plant, animal, or *S. cerevisiae.* This suggested that it was not the ability of these heterologous synthase enzymes to produce squalene, but their inability to feed squalene into the native sterol biosynthetic pathway that prevented them from restoring normal ergosterol biosynthesis in *S. cerevisiae.* By examining the ability of chimera squalene synthase enzymes to complement the erg9 mutation, a discrete sequence of amino acids near the C-terminus of the enzyme was identified which is necessary and sufficient for allowing any squalene synthase to restore normal sterol metabolism.

KEYWORDS: *Botryococcus braunii*, Botryococcene, Squalene, Bisfarnesyl Ether, Triterpene Methyl-Transferase, erg9 complementation

Thomas D Niehaus

Student's Signature

November 16, 2011

Date

ELUCIDATING THE BIOCHEMICAL WIZARDRY OF
TRITERPENE METABOLISM IN *BOTROYCOCCUS BRAUNII*

By

Thomas Daniel Niehaus

Joe Chappell

Director of Dissertation

Arthur G Hunt

Director of Graduate Studies

November 16, 2011

ACKNOWLEDGMENTS

The following dissertation, while an individual effort, benefited from the direction and insights of several people. First, my dissertation chair, Joe Chappell, acted as a mentor and gave insight and advice throughout my entire graduate career. In addition, Shigeru Okada performed much of the pioneering work that made some of the findings described in this dissertation possible. Next I wish to thank the entire dissertation committee, and outside reader, respectively: Joe Chappell, George Wagner, Robert Houtz, Dave Watt, Todd Porter, Charles Waechter, and Michael Montross. Each individual provided insights, experimental advice, and research collaborations that substantially contributed to my work and enhanced this dissertation. Finally, Scott Kinison provided technical support throughout my graduate career, especially on work involving triterpene methyltransferases.

In addition to the technical and instrumental assistance above, I received equally important assistance from family and friends. My Mother and Father, Debbie and Karl Niehaus, instilled in me an interest in science and fostered my understanding of the natural world from a young age. This greatly impacted my desire and ability to obtain the Ph. D. Finally I want to acknowledge my fiancée, Mary Beth Lassiter, for the ongoing support she provided throughout most of my graduate studies and dissertation process.

ATTRIBUTIONS

This dissertation reflects my own conceptual and intellectual body of work; however, contributions were made by others that deserve recognition. Shigeru Okada created the *B. braunii* phage cDNA library and originally cloned SSL-1 in the Chappell lab. In addition, Shigeru performed NMR and solved the structure of bisfarnesyl ether as well as methylated derivatives of botryococcene and squalene, and contributed reagents and algal strains that were crucial for a number of experiments. Dave Watt concurrently solved the bisfarnesyl ether structure from NMR data supplied by Shigeru, contributed to the proposed reaction scheme for bisfarnesyl ether, and along with Vitaliy Sviripa, prepared chemically synthesized bisfarnesyl ether and $NADPH_3$ which aided in several experiments. Scott Kinison originally identified TMT-3 and performed much of the initial work characterizing its triterpene methyltransferase activity and worked closely with me on the triterpene methyltransferase investigations. Tim Devarenne contributed to the general cloning effort of triterpene biosynthetic genes from *B. braunii*.

Table of Contents

List of Tables

List of Figures

Chapter 1 Background and Introduction

Background and Introduction

Botryococcus braunii is an interesting algal species that dates back at least 500 million years ago (Figure 1.1). The earliest evidence of *B. braunii* are found in sedimentary basins in Australia and date to the Precambrian era (1), and its fossil and chemical remains have been observed in oil shales around the world dating from the Precambrian to as recently as the Pliocene era (2). Aside from oil shales, *B. braunii* remains have been discovered in Kentucky coal beds produced during the Carboniferous era (3). In some cases, *B. braunii* oils appear to have contributed as much as 1.4% to the total hydrocarbons in oil shales (4); thus, *B. braunii* is recognized as a major contributor to oil and coal deposits throughout geological time.

Another interesting aspect of *B. braunii* is that three extant races (chemotypes) of this alga have been documented that differ in the type of oil that each accumulates (Figure 1.2). Race A accumulates C_{23}-C_{27} alkadienes and alkatrienes which are fatty acid or polyketide derived compounds (5). Race B accumulates the C_{30} triterpenes, squalene and botryococcene, and their methylated derivates (6). Race L accumulates the tetraterpene, lycopadiene (7). All races are morphologically similar, consisting of single oblong-shaped cells that form colonies ranging from four to several tens of cells held together by an extracellular matrix (Figure 1.3). Oil accumulates in both intracellular oil bodies and in association with the extracellular matrix (Figure 1.3) (8). The major metabolites of each race of *B. braunii* are unusual compounds in that they are only known to be produced by their respective race of *B. braunii.* Thus, these chemicals have been used as markers indicating the presence of *B. braunii* in oil and coal shale deposits around the world (9). While the different races share physical and morphological similarities to one another, little is actually known about their biochemical and genetic similarities and differences. For instance, is each of the races equally capable of producing all three kinds of oils? And, if so, is there a genetic, biochemical, environmental or some other factor that determines what class of oils a particular race will accumulate?

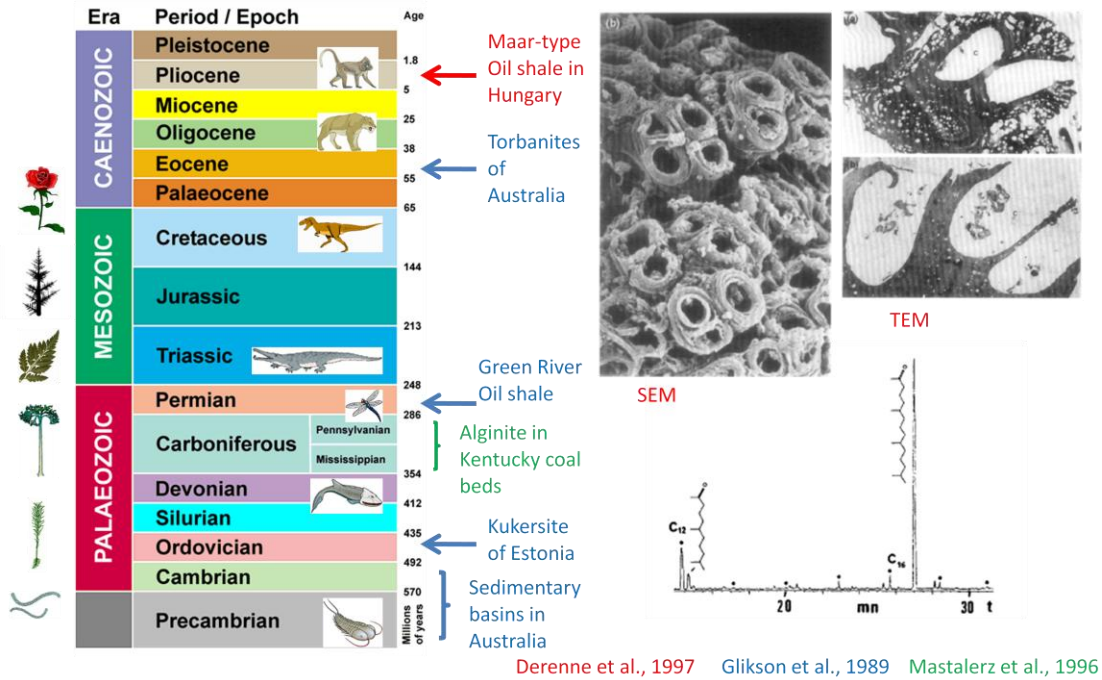


Figure 1.1 B. braunii is a major contributor to coal and petroleum shale deposits throughout geological time.

The left-hand panel reflects a timescale for geological events and biological evolution (plants, far left; animals on the right) on Earth. The timeline is further annotated (red, blue and green) for reports of petroleum and coal deposits sampled, dated, and analyzed for fossil evidence (scanning and transmission micrographs) and chemical profiling (GC chromatograph) document the prior existence of Botryococcus in these sediments.

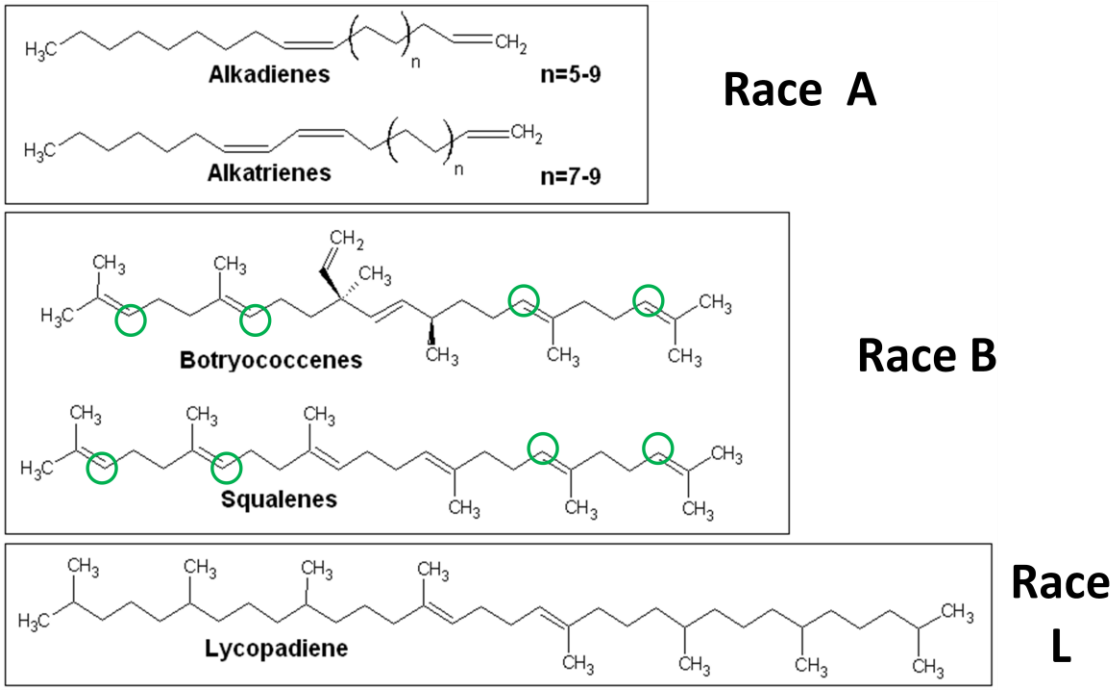


Figure 1.2 Three extant chemotypes of *B. braunii* are described.

Race A accumulates alkadienes and alkatrienes; race B accumulates botryococcene, squalene, and their methylated derivatives (methylation sites indicated with green circles); and race L accumulates lycopadiene.

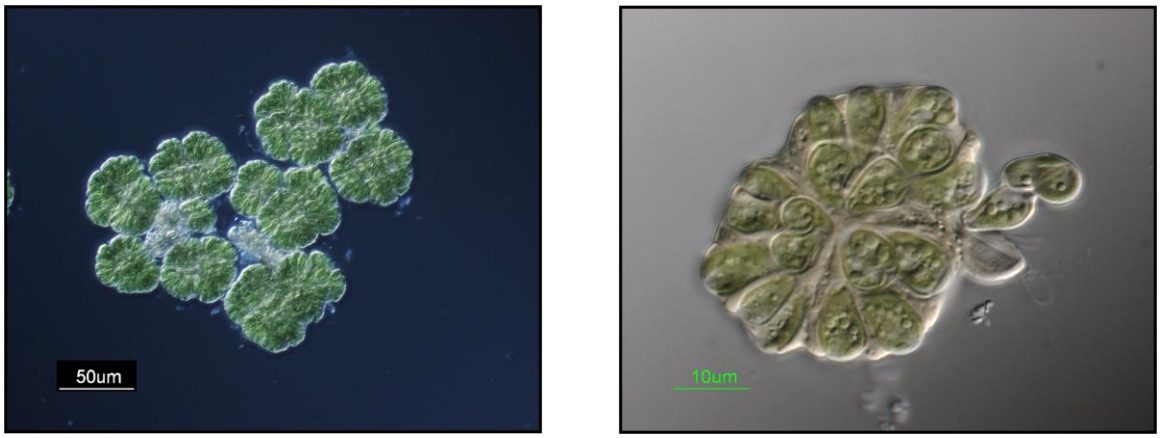


Figure 1.3 Real-color light micrographs of *B. braunii* Race B.

The micrographs illustrate the clustering of cells held together by an extracellular matrix. Photos courtesy of Tim Devarrene, Texas A&M University.

Another reason *B. braunii* has attracted significant attention is because all three races accumulate excessive amounts of oil ranging from 30 to 86% of their dry weight (10). Oil levels of 30-40% dry weight are more commonly observed in races A and B grown in the Chappell laboratory. The high levels of oil that accumulate in the algae cause it to float in calm water (Figure 1.4), which might represent an ecological adaptation. Unlike most algae, *B. braunii* do not have flagella, which enable other algae to swim towards light or nutrients. The accumulating oil may serve to help *B. braunii* float to the top of ponds or lakes where it can compete for resources such as light. A practical consequence of *B. braunii*'s ability to float is that it would facilitate collection of the algae if grown on an industrial scale. However, *B. braunii* grows considerably slower than most algae, and the doubling time for *B. braunii* is about six days in comparison to *C. reinhardtii* that doubles in as little as 5-8 hours (11). Figure 1.5 shows a side-by-side comparison of these two algal species grown under identical conditions and clearly indicates the difficulty of growing *Botryococcus braunii* rapidly on a large scale.

The oils of *B. braunii*, particularly Race B, are excellent feed stocks for the generation of high octane fuels for combustion engines (12) as well as the building materials for the industrial chemistry sector; however, the slow growing nature of *B. braunii* limits its utility as a production platform for the biosynthesis of such compounds. Capturing the genetic blueprints for botryococcene biosynthesis would allow this pathway to be engineered into faster growing organisms more suitable for large scale oil production.

Squalene biosynthesis is well described (13, 14, 15) whereas botryococcene biosynthesis has remained elusive. It has been shown that botryococcene is derived from the precursor, FPP (16), and it was proposed that botryococcene biosynthesis could mirror that of squalene in that 2 FPPs are condensed into a pre-squalene diphosphate intermediate that is reductively rearranged into either squalene or botryococcene (17, 18). Botryococcene and squalene are further metabolized to C_{31}-C_{37} methyl derivates, putatively through the action of SAM-dependent methyltransferases (19). Botryococcene and its methyl-derivatives accumulate as hydrocarbon oil in both intra- and extracellular compartments. Squalene and its methyl-derivatives accumulate as a small proportion of the total hydrocarbon oil (less than 5% w/w) but are more commonly incorporated into a variety of natural products such as carotenoid-like compounds (20) or polymerized into an extracellular matrix (21). Figure 1.6 shows a schematic representation of some of the unique biochemistry in *B. braunii* race B.

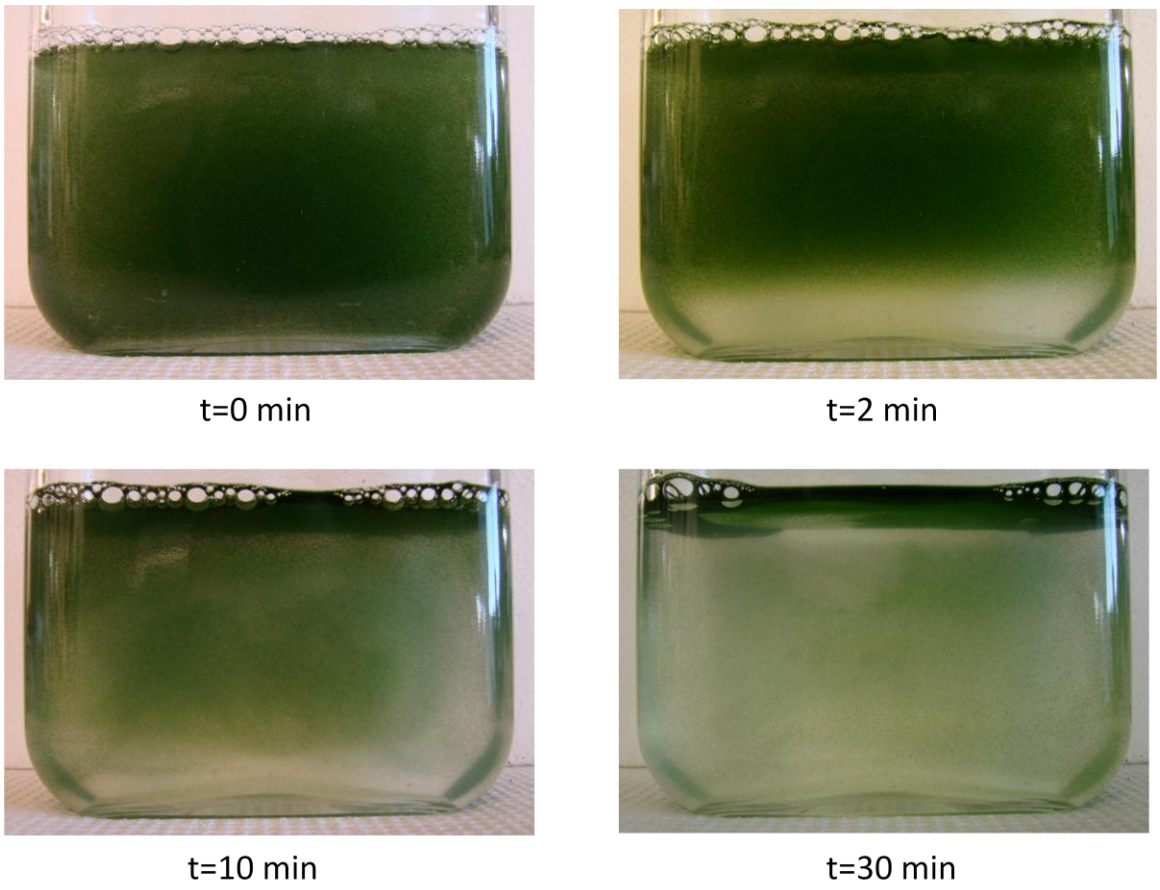


Figure 1.4 *B. braunii* floats in calm water.

B. braunii race B was grown in 1 liter Roux flasks aerated with 2.5% CO2 in air. The air was turned off and photographs were taken at the indicated times.

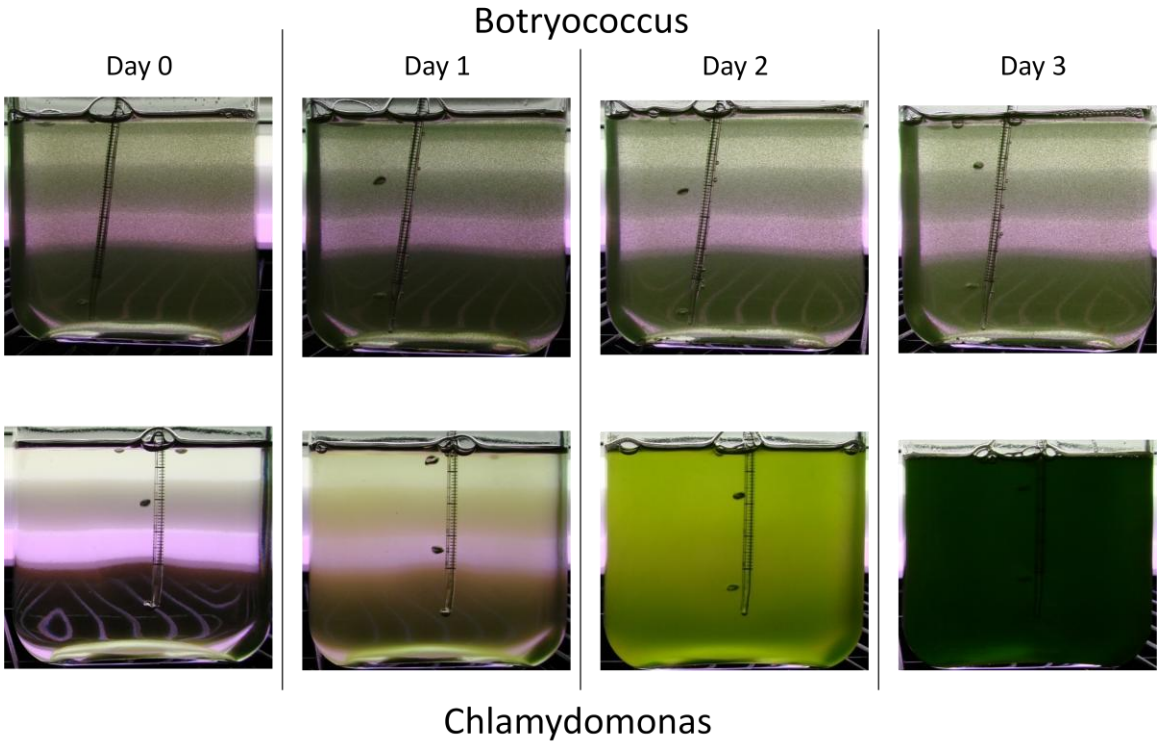


Figure 1.5 Growth comparision of *B. braunii* and C. reinhardtii.

Cultures of *B. braunii* and C. *reinhardtii* were used to inoculate 500 mL of Chu13 media and the algae was grown side-by-side in a fluorescent light bank and aerated with 2.5% CO2 in air. Photographs were taken every 24 h to compare the growth of the two algal species.

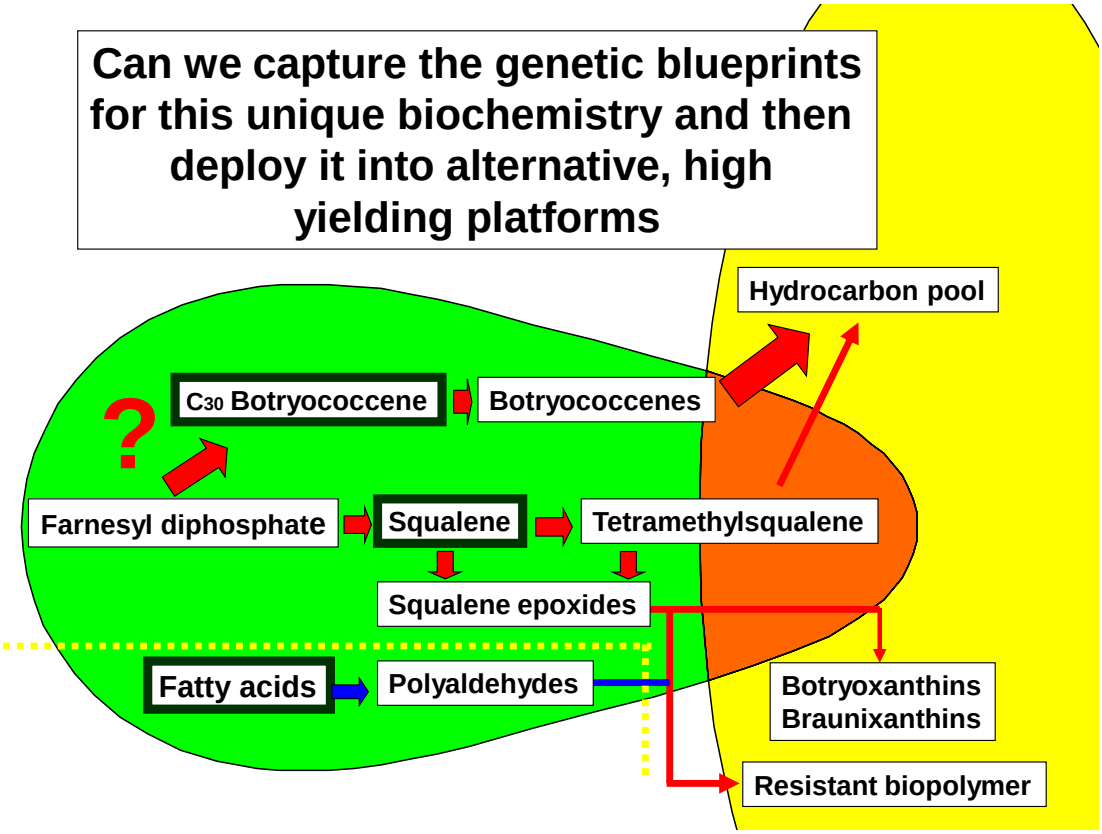


Figure 1.6 Predicted triterpene metabolism in *B. braunii*.

FPP is the precursor to the triterpenes, botryococcene and squalene, which are subsequently methylated to primarily C_{32} and C_{34} derivatives. Triterpene oils accumulate in intracellular oil bodies (cell colored as green and orange) and in an extracellular matrix (colored yellow). Squalene and its methyl-derivatives are also incorporated into a variety of natural products, such as the carotenoid-like compounds, botryoxanthins, and braunixanthins, and also polymerized along with polyaldehydes into a resistant extracellular matrix.

One difficulty in characterizing triterpene synthase genes, such as squalene synthase, is that many of these enzymes contain membrane-spanning domains that limit their expression in *E. coli*. An alternative to bacterial expression is characterization of these enzymes in a eukaryotic host that contains membrane systems, such as yeast. The Chappell laboratory has used a yeast line, Cali-7, to characterize sesquiterpene synthases (22). Cali-7 is an engineered yeast line that overproduces FPP, the precursor to sesquiterpenes and triterpenes, by having its endogenous erg9 (squalene synthase) knocked out and overexpression of HMGR, the rate limiting step in the MVA pathway. To create a yeast line specific for triterpene expression, Cali-7 was modified by knocking out its endogenous erg1 (squalene epoxidase), resulting in yeast line TN-7. TN-7 has the ability to overproduce FPP, but is unable to metabolize squalene and is well suited for triterpene synthase characterization. Not only does characterization of triterpene synthase enzymes in TN-7 supplant bacterial expression systems, but also it allows for the production and purification of mg quantities of product which enables structural determination of the enzyme catalyzed molecule(s).

Experimental Aims of this Thesis

B. braunii has been studied extensively for several decades, but the biosynthesis of botryococcene has remained enigmatic (16). Chapter Two describes efforts to elucidate this biosynthetic pathway and provides an example of how the resulting genetic blueprints can be used to engineer this unique metabolism into other host species. The discovery that botryococcene and squalene biosynthesis occur by unique mechanisms not observed in any other organisms to date, and chapter 4 provides preliminary enzymological studies defining reaction conditions and co-factor requirements for several of these interesting enzymes.

Chapter 2 explores the mechanism responsible for botryococcene biosynthesis in *B. braunii*. The majority of the oil is composed of its C_{31} to C_{34} methylated-derivatives. Methylated squalenes are also found into a variety of *B. braunii* natural products such as those polymerized into the extracellular matrix, as a small component of the oil itself, and carotenoid-like compounds. The methylation of botryococcene and squalene are important physiological processes in *B. braunii*. Methylated botryococcene and squalene are also superior starting materials for the hydrocracking process and provide for superior yields of combustible fuels (23). Chapter 4 describes efforts along with Scott Kinison in the Chappell laboratory to identify and to functionally characterize the triterpene methyltransferase enzymes (TMTs) of *B. braunii*.

Because these studies of algal triterpene synthases involved contrast and comparative studies with other squalene synthases, this work included co-expression studies of heterologous squalene synthases genes in *S. cerevisiae* and corroborated an earlier, but surprising observation reported in the literature. While the squalene synthase gene of *S. cerevisiae* was able to complement an erg9 (squalene synthase) knockout in yeast,

squalene synthases genes from plants and animals were not. A surprising observation was that while expression of these other squalene squalene genes in yeast were not able to complement the erg9 knockout (*i.e*., restore normal sterol metabolism), the yeast efficiently produced and accumulated squalene. This suggested that it was not the ability of these heterologous synthase enzymes to produce squalene, but their inability to feed squalene into the native sterol biosynthetic pathway that prevented them from restoring normal ergosterol biosynthesis in *S. cerevisiae*. Chapter 5 describes the work which defines a discrete sequence of amino acids within the yeast squalene synthase enzyme that are necessary and sufficient for allowing any squalene synthase to restore normal sterol metabolism in the erg9 genetic background.

Chapter 2 Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*

Introduction

Botryococcene biosynthesis resembles that of squalene, a metabolite essential for sterol metabolism in all eukaryotes. Squalene arises from an initial condensation of two molecules of farnesyl diphosphate (FPP) to form pre-squalene diphosphate (PSPP), which undergoes a reductive rearrangement to form squalene. In principle, botryococcene could arise from an alternative rearrangement of the pre-squalene intermediate. Because of these similarities, we predicted that a botryococcene synthase would resemble squalene synthase and hence isolated squalene synthase-like genes from *B. braunii* race B. While *B. braunii* does harbor at least one typical squalene synthase, none of the other three squalene synthase-like (SSL) genes encode for botryococcene biosynthesis directly. SSL-1 catalyzes the biosynthesis of PSPP and SSL-2 the biosynthesis of bisfarnesyl ether, while SSL-3 does not appear able to directly utilize FPP as a substrate. When combinations of the synthase-like enzymes were mixed together, *in vivo* and *in vitro*, robust botryococcene (SSL-1+SSL-3) or squalene biosynthesis (SSL1+SSL-2) was observed. These findings were unexpected because squalene synthase, an ancient and likely progenitor to the other *Botryococcus* triterpene synthases, catalyzes a two-step reaction within a single enzyme unit without intermediate release, yet in *B. braunii,* these activities appear to have separated and evolved interdependently for specialized triterpene oil production greater than 500 MYA. Co-expression of the SSL-1 and SSL-3 genes in different configurations, as independent genes, as gene fusions, or targeted to intracellular membranes, demonstrates the potential for engineering even greater efficiencies of botryococcene biosynthesis.

Background

Botryococcus braunii is a colony-forming, freshwater green algae reported to accumulate 30 to 86% of its dry weight as hydrocarbon oils (10). Three distinct races of *B. braunii* have been described based on the types of hydrocarbons that each accumulates (24). Race A accumulates fatty acid-derived alkadienes and alkatrienes (5); race L accumulates the tetraterpene lycopadiene (7); and race B accumulates triterpenes, predominately botryococcene, squalene, and their methylated derivatives (6). The oils accumulate both in intracellular oil bodies and in association with an extracellular matrix (8), which in race B consists largely of long-chain, cross-linked biopolymers formed in part from acetalization of polymethylsqualene diols (21). Di- and tetra-methylated botryococcenes are generally the most abundant triterpenes accumulating in race B with reduced amounts of tetramethylated-squalene (25) and other structural derivatives of squalene and botryococcene that range from C_{31} to C_{37} accumulating to various levels in different strains and in response to variable culture conditions (26). Other polymethylated derivatives such as diepoxy-tetramethylsqualene (27), botryolins (28), and brauixanthins (20) have also been reported.

B. braunii race B has received significant attention because it is considered an ancient algal species dating back at least 500 MYA and is one of the few organisms known to have directly contributed to the existing oil and coal shale deposits found on Earth (1-3), accounting for up to 1.4% of the total hydrocarbon content in oil shales (29). Secondly, because the hydrocarbon oils of *B. braunii* race B are readily converted to starting materials for industrial chemical manufacturing and high quality fuels under standard hydrocracking/distillation conditions in yields approaching 97% (Figure 2.1*A*) (12), race B has been considered a potential production host for renewable petrochemicals and biofuels. Unfortunately, the slow growth habit of *B. braunii* poses serious limitations to its suitability as a robust biofuel production system. Capture of the genes coding for this unique oil biosynthetic capacity would provide opportunities to engineer this metabolism into other faster growing and potentially higher yielding organisms (19) than *B. braunii*.

Our approach for identifying the triterpene biosynthetic genes in *B. braunii* relied in large part on the putative similarities in the biosynthetic mechanisms for squalene and botryococcene (17-18, 30). Squalene biosynthesis has been extensively investigated because it is positioned at a putative branch point in the isoprenoid biosynthetic pathway directing carbon flux to sterol metabolism, and thus represents a potential control point for cholesterol biosynthesis in man (31). Evidence for a two-step reaction mechanism catalyzed by squalene synthase has been described (13) (Figure 2.1*B*). The initial reaction step consists of a head-to-head condensation of two farnesyl diphosphate (FPP) molecules to form a stable cyclopropyl intermediate, pre-squalene diphosphate (PSPP) (14, 32). In the second step, PSPP undergoes a reductive rearrangement in the presence of NADPH to yield squalene possessing a C1-C1' linkage between the two farnesyl substituents (15, 33) (Figure 2.1*B*). Poulter (13) suggested that botryococcene biosynthesis occurs *via* an analogous reaction mechanism with the initial reaction

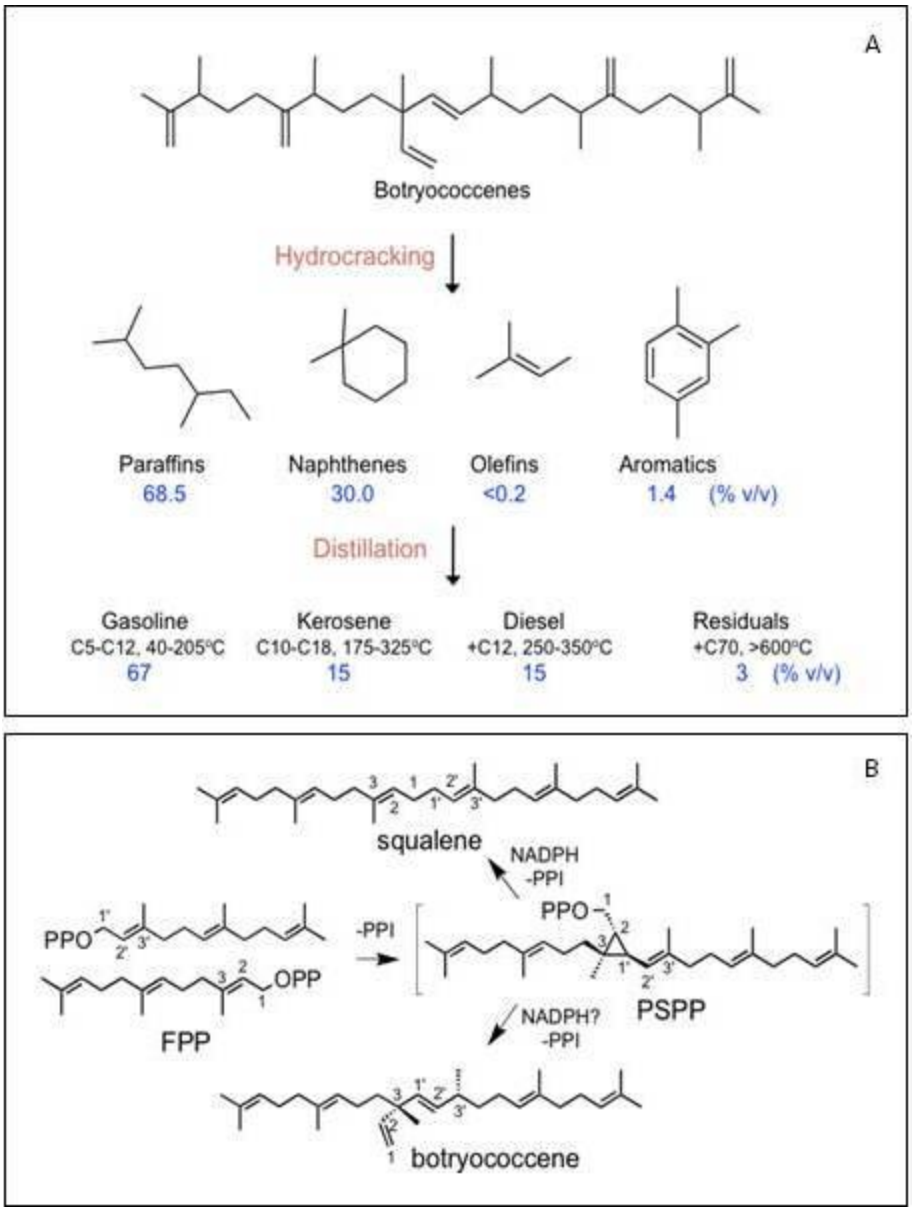


Figure 2.1 The triterpene oils of *B. braunii* Race B and their putative biosynthetic mechanism.

The triterpene oils of *B. braunii* race B (illustrated as tetramethyl-botryococcene) are likely progenitors to existing coal and oil shale deposits according to geochemical and fossil records (54). These oils are readily converted under standard hydrocracking processes to molecular species of direct utility in industrial chemical manufacturing or can be distilled in high yields to all classes of combustible fuels, including gasoline (67%), a*via*tion fuels (15%) and diesel (15%) (carbon chain-length, distillation temperature, % volume conversion) (12) (A). The biosynthetic origin of the *B. braunii* triterpene oils has remained enigmatic. Poulter (13) suggested that the biosynthesis of the botryococcene scaffold could arise from a mechanism similar to that for squalene, a key intermediate in sterol and cyclized triterpene metabolism (B). Squalene biosynthesis

occurs from an initial head-to-head condensation of two farnesyl diphosphate molecules (FPP) into the stable intermediate pre-squalene diphosphate (PSPP), followed by a reductive rearrangement to form squalene catalyzed by a single enzyme without release of the PSPP intermediate (40). Botryococcene biosynthesis is suggested to parallel that of squalene in the first half-reaction, differing only in the reductive rearrangement of PSPP to yield the methyl/ethyl branched, 1'-3 linked botryococcene product.

proceeding through PSPP, followed by a reductive rearrangement yielding a C3-C1' linkage between the two farnesyl precursors and possessing an ethyl as well as a methyl group at C3 in the final product.

Extensive investigations of squalene synthase including site-direct mutagenesis (34) and structural elucidation of 3-dimensional structure (35) have focused on five highly conserved domains (domains I-V) thought associated with catalysis (36). Many studies have also utilized these highly conserved domains as a means for isolating the corresponding genes from a diverse range of organisms. For instance, the Cappell laboratory previously described the functional characterization of a squalene synthase gene from *B. braunii* race B (37). In that work, degenerate oligonucleotide primers complementary to several of the conserved domains were used to amplify a small region of a putative squalene synthase gene, and that gene fragment was then used to isolate a full-length cDNA from a cDNA library. Heterologous expression of that cDNA in bacteria and *in vitro* characterization of the encoded enzyme validated that the cDNA encoded for a squalene synthase enzyme, but lacked any detectable botryococcene synthase activity.

The current results represent efforts to define the botryococcene biosynthetic pathway, to capture the genes coding for these unique enzymological transformations, and to reconstruct the initial steps of these unusual triterpene pathways in a heterologous host.

Results

Functional Identification of Genes for Triterpene Biosynthesis

Because we surmised that a botryococcene synthase enzyme might possess amino acid domains in common with squalene synthase, the *B. braunii* squalene synthase cDNA was used to re-screen the *B. braunii* cDNA library under low stringency hybridization conditions. A unique squalene synthase-like gene (SSL-1) was isolated and characterized. The SSL-1 gene predicted a squalene synthase-like protein exhibiting some resemblance to other squalene synthase enzymes within domains I-V, but missing a carboxy-terminal, membrane-spanning domain (Figure 2.2). Surprisingly, purified bacterial expressed SSL-1 protein did not exhibit either squalene nor botryococcene biosynthesis when assayed *in vitro* even in the presence of a variety of reducing co-factors like NADPH (Figure 2.3a and 2.4), ferredoxin or cytochrome B5 systems. However, when SSL-1 was expressed in a yeast engineered for high-level production of FPP and having its endogenous squalene synthase and squalene epoxidase genes inactivated, pre-squalene alcohol (PSOH), the dephosphorylated form of PSPP, accumulated to significant levels (Figure 2.5*B*). Subsequent incubations of SSL-1 with radiolabeled FPP confirmed robust *in vitro* production of PSPP as the sole reaction product with a K_m for FPP of 12.8 µM and catalytic turnover rate (k_{cat}) equal to 2.7 x 10^{-2}/sec with no stimulation of activity by NADPH addition (Figs. 2.4 and 2.5J). This suggested that SSL-1 was catalytically competent for the first half-reaction of squalene synthase, but perhaps required additional conditions or algal factors for complete catalytic activity. Mixing the purified SSL-1 enzyme with algal cell-free lysate did indeed enhance NAD(P)H-dependent botryococcene biosynthesis up to 10-fold, which was also proportional to the amount of the purified SSL-1 protein or the algal lysate added (Figure 2.3*A*-*C*). The mechanism for botryococcene biosynthesis thus appeared to be similar to squalene synthase in its first half-reaction, catalysis of PSPP formation, but differed in requiring another algal co-factor that either shuttled reducing equivalents to the reaction mechanism of SSL-1 or participated directly in the conversion of PSPP to botryococcene.

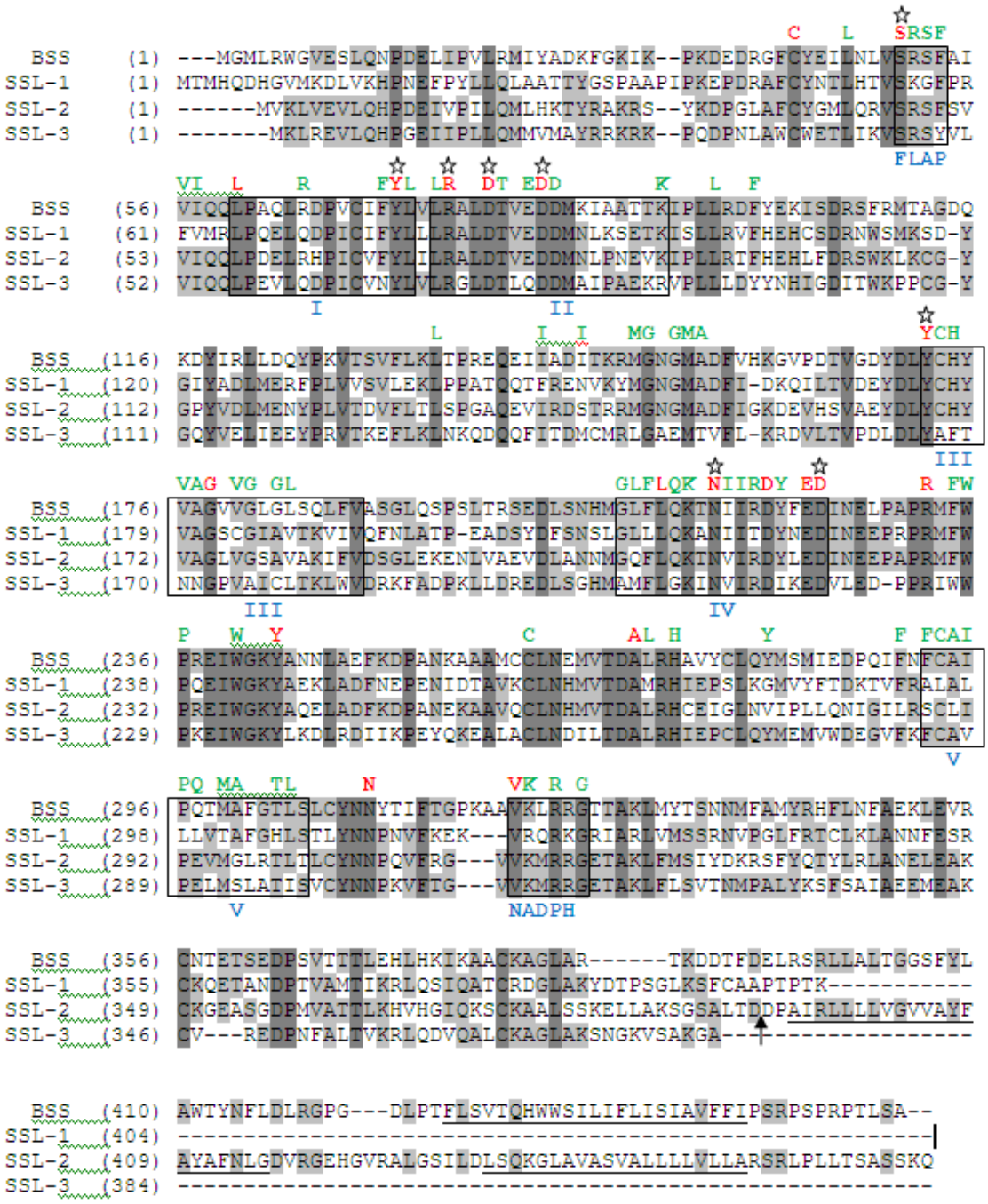


Figure 2.2 Amino acid alignment of SSL enzymes.

Amino acid alignment of BSS, SSL-1, SSL-2, and SSL-3 from *B. braunii* race B. Five highly conserved domains among squalene synthase's identified by Robinson *et al.* (36), and the "FLAP" and putative NADPH-binding site identified by Gu *et al.* (34) are boxed and labeled in blue. Amino acids completely conserved in the squalene synthase of *B. braunii* (AF205791), *C. reinhardtii* (XM_001703395), *A. thaliana* (NM_119630), *N. tabacum* (U60057), *H. sapiens* (NM_004462), *R. norvegicus* (NM_019238), *S. cerevisiae* (X59959), *S. pombe* (NM_001021271), and *Y. lipolytica* (AF092497) are

labeled above the alignment in green, and those residues also conserved with these squalene synthases and dehydrosqualene synthase (CrtM) from *S. aureus* (AM920687) are labeled in red. Residues of CrtM identified by Lin *et al.* (52) as important for the first and second half-reactions are labeled with a star. Possible membrane-spanning regions of BSS and SSL2 as predicted by TMpred are underlined. In the subsequent studies, a carboxy terminal truncated form of SSL2 (tr2) missing the putative membrane-spanning domain after D392 (indicated by arrow) was heterologously expressed in bacteria.

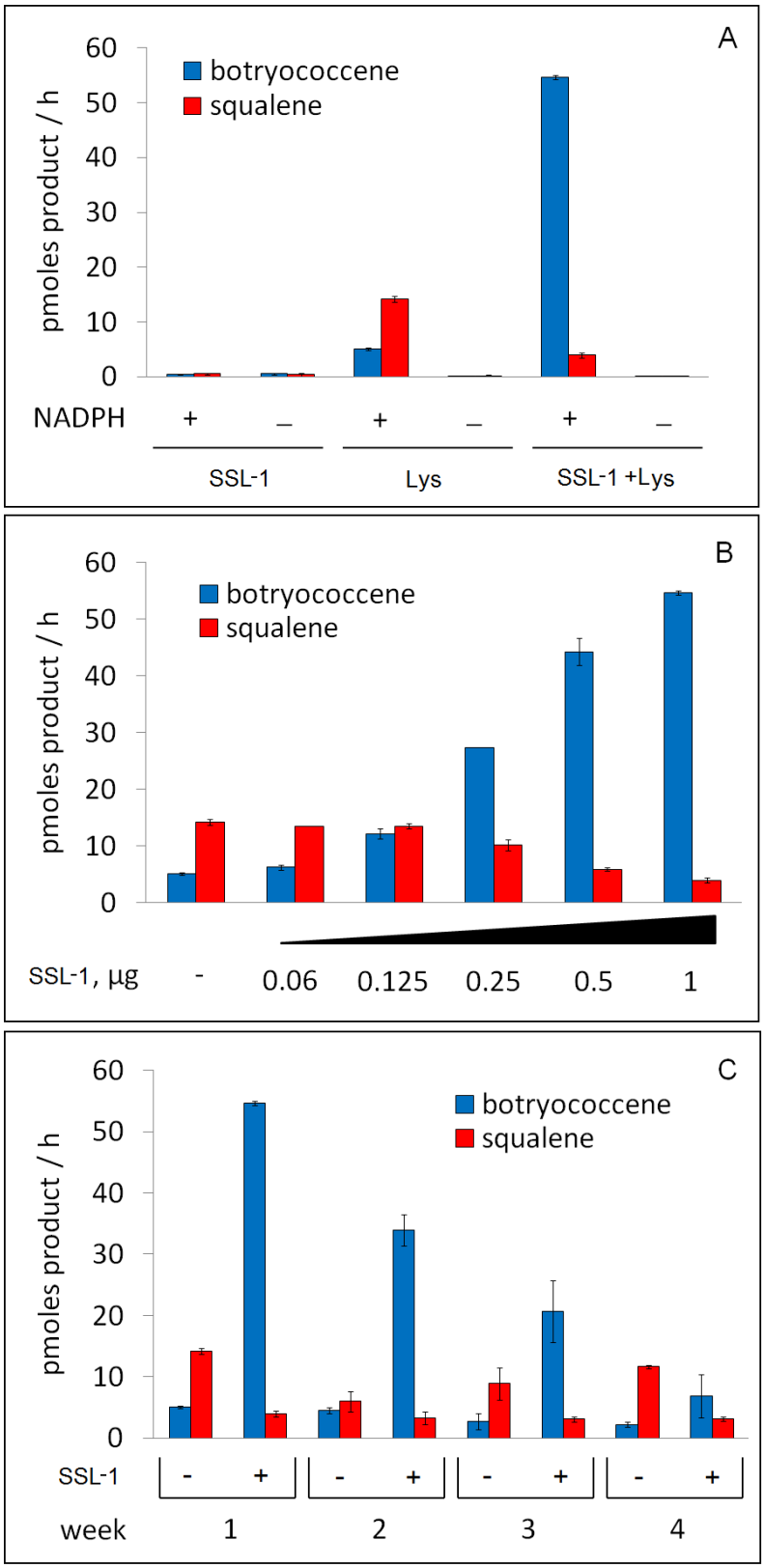


Figure 2.3 Dependence of the SSL-1 enzyme on algal lysate for botryococcene biosynthesis.

Purified SSL-1 enzyme (1 µg) (SSL-1), *B. braunii* 2000g whole-cell lysate (10 µg protein) (Lys), and equal aliquots of both SSL-1 and lysate were incubated with radiolabeled FPP, with (+) or without (-) 2mM NADPH and the incorporation into squalene and botryococcene determined by TLC separation of the reaction products followed by scintillation counting of the corresponding zones (panel a). Increasing amounts of purified SSL-1 were incubated with 10 µg of *B. braunii* 2000g whole-cell lysate and the incorporation of radiolabeled FPP into squalene and botryococcene determined by TLC separation/scintillation counting (panel b). *B. braunii* 2000g whole-cell lysates were prepared from cells collected at the indicated times (weeks) after subculturing, and aliquots containing 10 µg of protein were incubated without (-) or with (+) 1 µg of purified SSL-1 protein, and incorporation of radiolabeled FPP into squalene and botryococcene determined. The SSL-1 gene containing a 5' terminal sequence coding for a hexa-histidine tag was over-expressed in *E. coli,* and the corresponding enzyme purified by nickel affinity chromatography according to the manufacturer (Novagen). *B. braunii* lysate was prepared from cells collected at the indicated stages of culture development according to the procedure described by Okada *et al.*, (2004), and 10 µg of lysate protein was per assay. Assays were incubated at 37°C for 1h, then the reaction products extracted with hexane. Aliquots of the hexane extracts was separated by silica TLC and the radioactivity migrating to zones corresponding to authentic standards of botryococcene and squalene determined by scintillation counting.

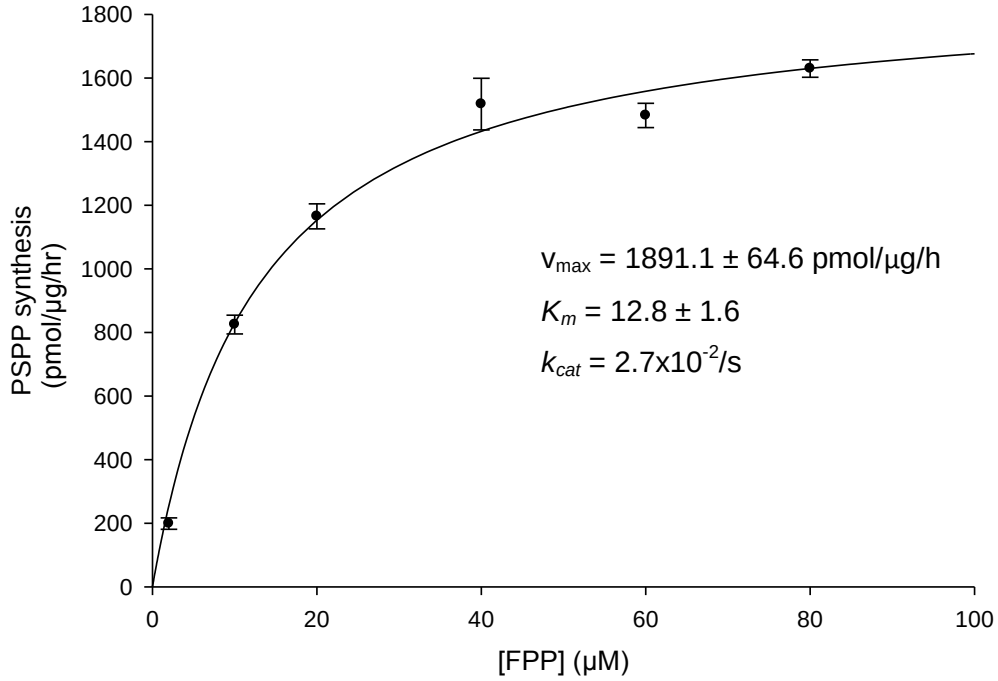


Figure 2.4 Michaelis-Menten enzyme kinetics of SSL-1.

Enzyme assays (50 µl) were set up as described in the methods section with purified SSL-1 (0.2 – 1.0 µg) and the indicated concentration of ^{3}H-FPP. Assays also contained 0.1% Triton X-100 and ± NADPH. Assays were incubated for 15 min at 37°C and stopped by addition of 50 µl 0.5 M EDTA. The reactions were then extracted three times with 200 µl water saturated 1-butanol and pooled in a 4 mL glass screw cap *via*l. The butanol was evaporated with a stream of nitrogen gas, and the white residue was resuspended in 2 mL of acid phosphatase solution (20% 1-propanol (v/v), 100 mM sodium acetate pH 4.7, 0.1% Triton X-100, 10 U sweet potato acid phosphatase) and incubated 12-16 h in a 28°C shaker. Dephosphorylated products were extracted three times with 1 mL n-hexane, pooled, dried with a stream of nitrogen gas, and resuspended in 200 µl of n-hexane. Aliquots of the hexane extract were spotted on reverse-phase TLC plates along with standards of FOH and PSOH and developed with methanol:acetone (8:2). The standards were visualized with iodine vapors, and the zones corresponding to FOH (rf= 0.65) and PSOH (rf= 0.45) were scraped and analyzed by scintillation spectroscopy. Addition of NADPH had no significant effect on enzyme activity and greater than 90% of the expected radioactivity was recovered in the FOH and PSOH zones indicating that PSPP is the dominant reaction product formed from FPP. The data was analyzed using the SigmaPlot Enzyme Kinetics 1.3 module. Data represents mean ± S.E.M.

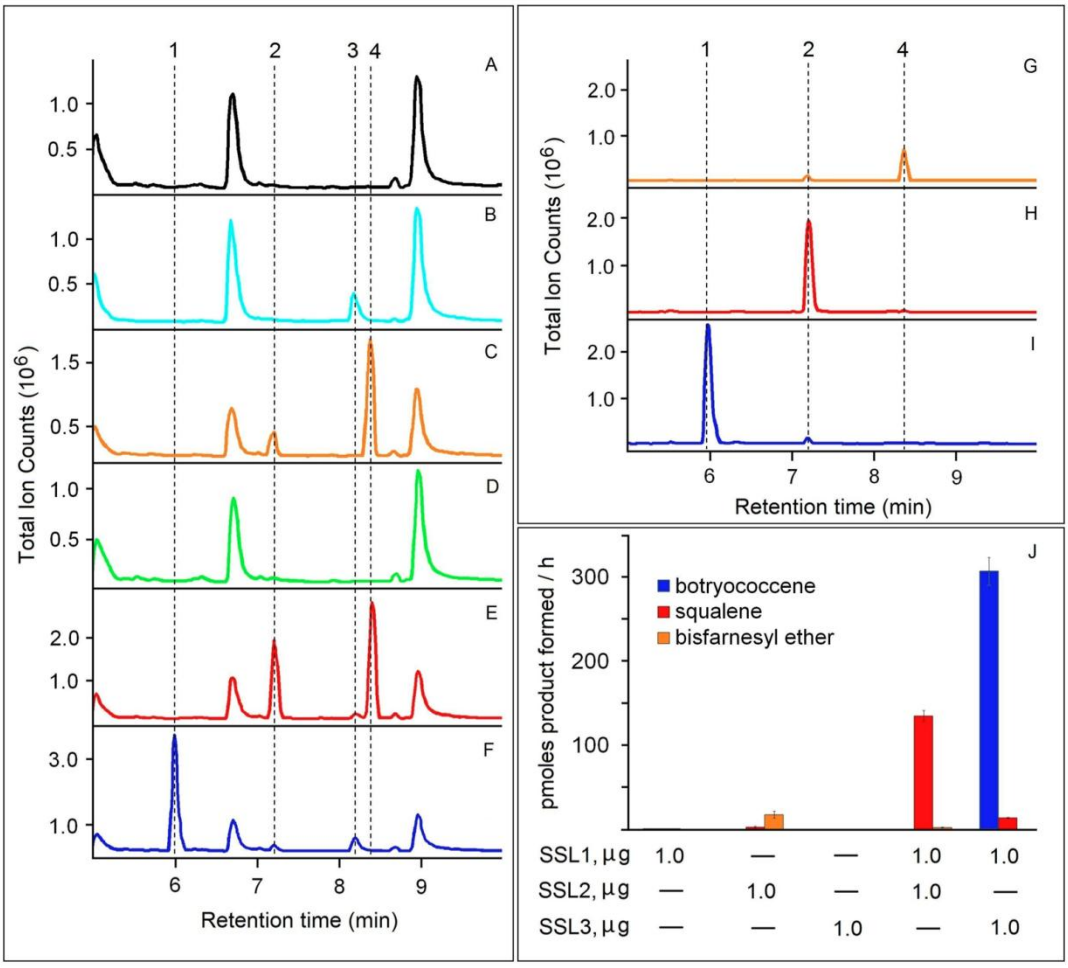


Figure 2.5 Functional characterization of the squalene synthase-like genes of *Botryococcus braunii* race B.

The squalene synthase-like genes, SSL-1, SSL-2 and SSL-3, were expressed in yeast separately (SSL-1(B), SSL-2(C) or SSL-3(D)) or in combinations (SSL-1 + SSL-2(E), SSL-1 + SSL-3(F)) and the hexane extractable metabolites profiled by GC-MS. The chemical profile of yeast not engineered with any gene constructs serves as the background control (A). The SSL genes were also expressed in bacteria, the affinity-tagged proteins purified and assayed separately (SSL-2 (G)) or in combinations (SSL-1 + SSL-2 (H); SSL-1 + SSL-3 (I)) for the reaction products generated upon incubation with FPP and profiled by GC-MS (G-I), or for quantitative determination of radiolabeled FPP incorporated into specific reaction products separated by TLC (J). Data (J) represents mean ± S.E.M. obtained from three independent experiments (n=3). The chromatograms (A-I) are also annotated for the elution behavior of botryococcene (1), squalene (2), presqualene alcohol (3) and bisfarnesyl ether (4).

Because no natural occurring squalene synthase catalyzing only the first or second half-reactions has been reported, we reasoned that other squalene synthase-like cDNAs for botryococcene biosynthesis might exist. We undertook an exhaustive assessment of the SSL genes expressed in the *Botryococcus braunii* race B cells. The transcriptomic data from two independent sequencing efforts were assembled and screened computationally for additional squalene synthase-like genes. Two additional SSL genes were uncovered and labeled SSL-2 and SSL-3 (Figure 2.2). Although both of the predicted proteins showed amino acid sequence similarity to other squalene synthases in excess of 62%, neither bacterial-expressed, purified enzymes exhibited any botryococcene biosynthesis and only SSL-2 showed a low capacity for squalene biosynthesis when incubated with FPP as substrate (Figure 2.5*G*). When expressed in yeast, SSL-3 also did not cause the accumulation of any distinct products (Figure 2.5*D*), but SSL-2 resulted in the accumulation of a small amount of squalene (~10% of the total) and a terpene of unknown structure (Figure 2.5*C*). The dominant terpene accumulating in the SSL-2 expressing yeast was subsequently identified by NMR as bisfarnesyl ether and confirmed by comparative analysis of corresponding ether prepared by chemical synthesis (Figure 2.6). Subsequent analysis of the reaction products generated by *in vitro* incubation of SSL-2 with FPP also verified this enzyme as the source of this unique terpene (Figure 2.5*G* & *J*).

The observations of unique terpene products from squalene synthase-like enzymes in *Botryococcus*, namely PSPP by SSL-1 and bisfarnesyl ether by SSL-2, suggested that triterpene metabolism in this alga operates differently from that in other organisms. We considered the possibility that multiple SSL proteins might be required in botryococcene biosynthesis. To evaluate this possibility, the different SSL genes were co-expressed in yeast, or the heterologous expressed and purified proteins were incubated in various combinations. When SSL-1 was co-expressed with SSL-2, the amount of squalene accumulating increased about 30-fold (Figure 2.5*E*) along with a significant accumulation of bisfarnesyl ether. When purified SSL-1 and SSL-2 enzymes were incubated in a 1:1 stoichiometric ratio, squalene accumulation predominated (Figure 2.5*H*), suggesting that something different mechanistically might be occurring when the SSL-1 and 2 genes were co-expressed in yeast (see below). Most surprising, when SSL-1 and SSL-3 were co-expressed, botryococcene accumulation became readily apparent and accumulated to levels of 20 mg/L along with 0.7 mg/L of squalene (Figure 2.5*F*). *In vitro* incubations of the purified SSL-1 and SSL-3 proteins confirmed botryococcene as the predominant reaction product with squalene representing only 3-4% of the total reaction products (Figure 2.5*I* & *J*). Additional *in vitro* studies have also confirmed that both SSL-2 and SSL-3 are able to efficiently catalyze the biosynthesis of squalene and botryococcene, respectively, from PSPP but not FPP, and these activities of SSL-2 and SSL-3 are sufficient to account for all the squalene and botryococcene biosynthesized in combined assays with SSL-1 (Figure 2.7).

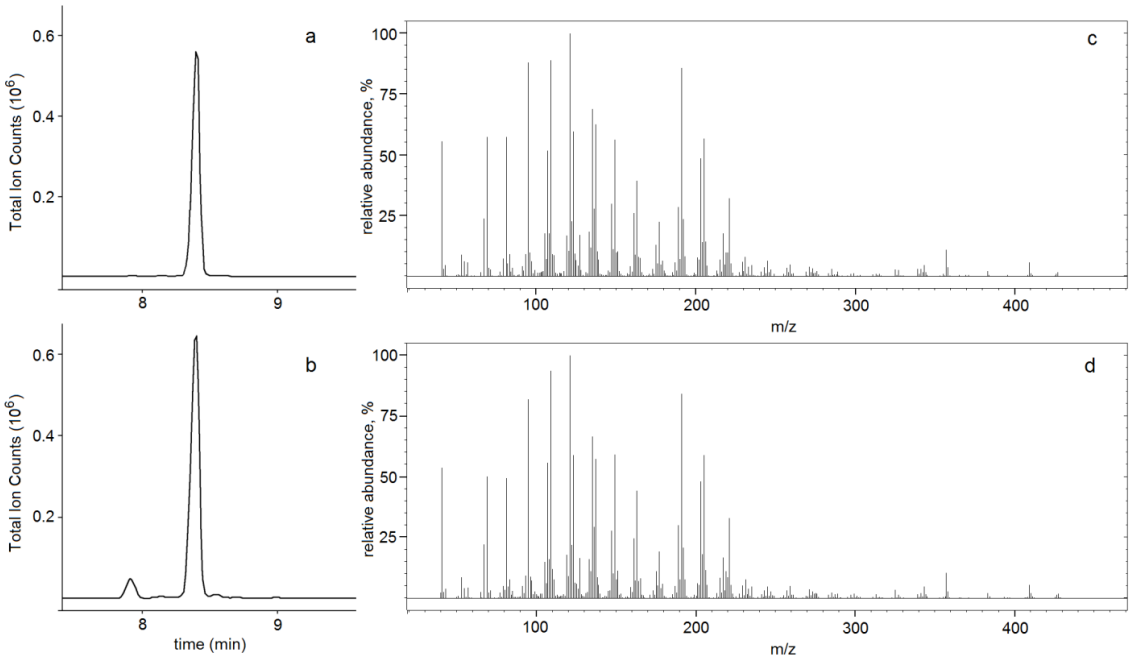


Figure 2.6 GC chromatographs of unknown triterpene purified from yeast over-expressing the SSL-2 gene (a) in comparison to chemically synthesized bisfarnesyl ether (b).

The MS for the dominant peak compounds with retention time of 8.38 min in panels a and b are shown in panels c and d, respectively. Chemically synthesized bisfarnesyl ether produced an identical NMR spectrum to the unknown triterpene purified from yeast over-expressing the SSL-2 gene.

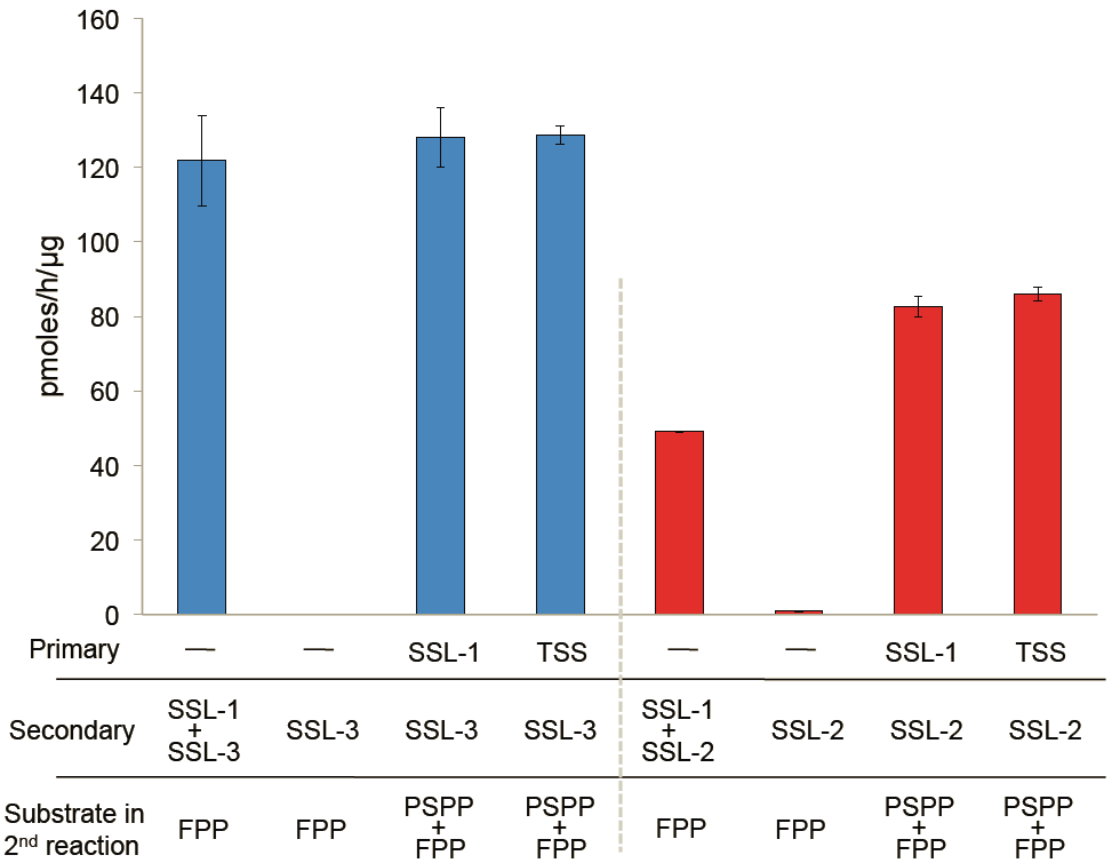


Figure 2.7 Enzyme assays showing PSPP utilization by SSL-2 and SSL-3.

Two primary enzyme assays (500 µL) were set up with 20 µg of purified SSL-1 or *Nicotiana benthamiana* squalene synthase (TSS) and containing 20 µM ^{3}H-FPP. The assays were incubated at 37˚C for 1 h, then extracted three times with 500 µL water saturated 1-butanol, pooled, and evaporated under a stream of nitrogen gas. The white residue was resuspended in 50 µL of 25 mM NH_4HCO_3 in 70% ethanol (v/v). An aliquot of the resuspension was analyzed by the acid phosphatase assay (described in Figure 2.4) and shown to consist of 45% and 40% PSPP for the SSL-1 and TSS primary incubations, respectively (the remaining 55% and 60% consisting of FPP). It was calculated that 1 µL of each resuspension contained ~150 pmoles of ^{3}H-PSPP. Secondary enzyme assays (50 µL) were set up containing 1 µL of resuspension (3 µM PSPP), 2 mM NADPH, and 1 µg of SSL-2 or SSL-3. For comparison, enzyme assays were set up containing 3 µM FPP, 2 mM NADPH and either 1 µg each of SSL-1 and SSL-2 or SSL-1 and SSL-3, and 5 µM FPP, 2 mM NADPH, and 1 µg of SSL-2 or SSL-3. Assays were incubated at 37˚C for 15 min, stopped by addition of 50 µL 0.5 M EDTA, and extracted with 100 µL n-hexane. Aliquots of the hexane extract were spotted on silica TLC plates with standards of squalene and botryococcene and developed with n-hexane. Standards were visualized with iodine vapors and the corresponding zones were scraped and analyzed by scintillation spectroscopy. The rate of botryococcene synthesis by SSL-3 or SSL-1 + SSL-3 is shown in blue and squalene synthesis by SSL-2 or SSL-1 + SSL-2 is shown in red. Data represents mean ± S.E.M.

Mechanistic Considerations for Bisfarnesyl Ether Biosynthesis

When incubated by itself, SSL-2 catalyzes the NADPH-dependent biosynthesis of approximately ~90% of all-*trans* bisfarnesyl ether and 10% of squalene (Figure 2.5*G*). This suggests that SSL-2 does have the ability to generate PSPP, but at a lower efficiency relative to ether formation. Based on a consideration of the detailed carbocation mechanism elucidated for the biosynthesis of squalene from FPP (Figure 2.1*B*) (13), one might not expect bisfarnesyl ether biosynthesis to involve a PSPP intermediate. Instead, if the initial carbocation generated on one of the two SSL-2 bound FPP molecules were quenched by reaction with an available water molecule, and if the farnesol (FOH) were positioned in the correct orientation and proximity to a second FPP molecule, displacement of the pyrophosphate group *via* a S_N2 Williamson ether synthesis-type reaction (38) could yield the bisfarnesyl ether (Figure 2.8). Support for such a mechanism comes from the incorporation of radiolabeled FOH directly into the bisfarnesyl ether product, but only when SSL-2 is incubated with both FOH and FPP (Table 2.1).

To determine if the mechanism of NADPH dependence for bisfarnesyl ether formation by SSL-2 was catalytic or structural, the quantitative yield of reaction product and NADPH oxidation were determined. While the biosynthesis of 1072 pmoles of squalene was correlated with an equal stoichiometric oxidation of 1098 pmoles of NADPH by the *Nicotiana benthamiana* squalene synthase enzyme, greater than 21 pmoles of bisfarnesyl ether were formed by SSL-2 when only 4.6 pmoles of NADPH were oxidized (Table 2.2). Approximately half of the NADPH oxidation by SSL-2 under these conditions is associated with the biosynthesis of 2.2 pmoles of squalene (Table 2.2). About 10 times more bisfarnesyl ether is formed per mole equivalent of NADPH oxidation (21 pmoles versus 2.4 pmoles), consistent with an allosteric or structural role for NADPH in the SSL-2 bisfarnesyl ether reaction rather than a catalytic one. A similar role for NADPH in stimulating PSPP formation by squalene synthases was reported earlier (33, 39, 40).

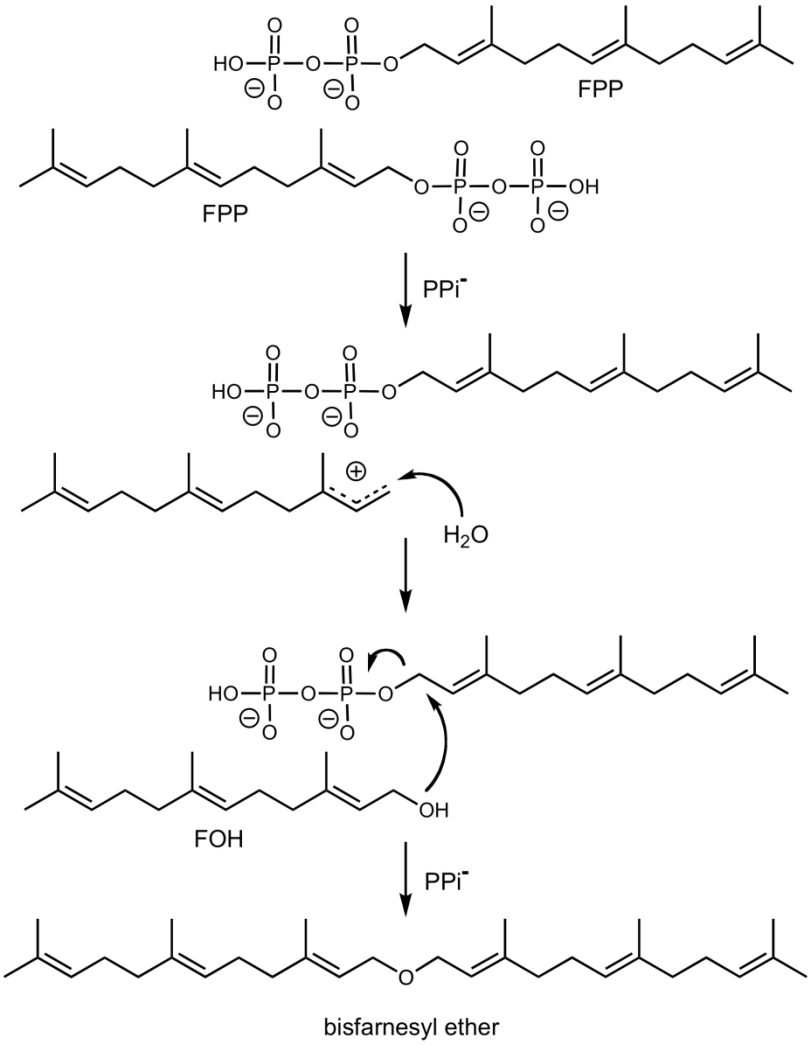


Figure 2.8 Proposed mechanism for bisfarnesyl ether biosynthesis by SSL-2.

When two molecules of FPP are bound by the SSL-2 enzyme, dissociation of the diphosphate substituent from one creates a carbocation, which can react with a water molecule in close proximity to generate farnesol (FOH). If the FOH is appropriately positioned relative to the second FPP molecule, a reaction akin to a Williamson ether synthesis (38) reaction could occur to yield bisfarnesyl ether.

Table 2.1 Substrate specificity of SSL-2

Substrate	**Bisfarnesyl ether formed (pmoles /h•µg protein)**
^{3}H-farnesol	0
FPP + ^{3}H-farnesol	16.3 +/- 0.4
^{3}H-FPP	13.5 +/- 3.3

Purified SSL-2 protein (2 µg) was incubated with either 10 µM 1-^{3}H-farnesol, 10 µM 1-^{3}H-farnesol and 10 µM FPP, or 10 µM 1-^{3}H-FPP in a 50 µL reaction at 37^{o}C for 1 h, the reaction products extracted with MTBE, and aliquots separated on silica TLC plates. The radioactivity incorporated in the zones corresponding to bisfarnesyl ether were determined by scintillation counting. Data represents mean ± S.E.M.

Table 2.2 NADPH oxidation by SSL-2

	pmoles NADPH oxidized per µg protein	pmoles squalene recovered per µg protein	pmoles bisfarnesyl ether recovered per µg protein
TSS	1098.6 ± 59.6	1072.6 ± 63.6	-
SSL-2	4.6 ± 2.6	2.2 ± 0.4	21.7 ± 1.8

Assays were performed as described above in the Methods except that an alternative reaction buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM $MgCl_2$) was used to minimize spontaneous NADPH oxidation and prevent protein precipitation at high concentrations. All the assays also contained 100 µM NADPH and 40 µM FPP in 300 µL final reaction volumes. The oxidation of NADPH was monitored at 340 nm with a Biorad SmartSpec Plus at room temperature (23°C). An extinction coefficient of 6220/M/cm was used to calculate the amount of NADPH oxidized. No NADPH oxidation could be measured in complete reaction buffer without protein over a 1 h incubation. When either *N. benthamiana* squalene synthase (TSS) or SSL-2 was incubated in reaction buffer without FPP, NADPH was oxidized at a rate of 24 and 22 pmoles/µg protein/h, respectively, suggesting that both enzymes cause a slight oxidation of NADPH that is uncoupled with squalene or bisfarnesyl ether biosynthesis. This background NADPH oxidation was subtracted from the experimentally determined rates. The experimental enzyme assays contained either purified TSS (0.75 or 1.5 µg) or SSL-2 (8 or 16 µg) and 40 µM 3H-FPP. Absorbance at 430 nm was recorded every 2.5 min. for 15 minutes, after which the reaction was stopped by adding an equal volume of 0.5 M EDTA. The reaction mixture was collected, extracted with 200 µL of hexane and aliquots were separated by silica TLC along with standards of squalene and bisfarnesyl ether using hexane:MTBE (25:1) as the developing solvent. The standards were subsequently visualized with iodine vapor, and corresponding zones were analyzed by scintillation spectroscopy. Data represents mean ± S.E.M of duplicate samples.

Improving the Efficiency of Botryococcene Biosynthesis

Production of botryococcene by yeast was improved by engineering different configurations of the SSL-1 and 3 genes (Figure 2.9). While co-expression of SSL-1 and SSL-3 yielded significant botryococcene, peptide fusions of SSL-1 and SSL-3 connected by a triplet repeat linker of GGSG improved production capacity greater than two-fold to upwards of 50 mg/L. Further enhancement to over 70 mg/L was observed by appending the carboxy-terminal 63 or 71 amino acids of the *Botryococcus* squalene synthase on to the carboxy-termini of SSL-1 and SSL-3 enzymes, respectively. These terminal amino acids serve to tether squalene synthase, and by inference SSL-1 and 3, to the yeast's endo-membrane system, which presumably brings the enzymes in closer proximity to one another or give the enzymes greater access to endogenous FPP pools. Further support for this notion has been the observation of greater than 100 mg/L of botryococcene by yeast over-expressing gene fusions of SSL-1 and SSL-3 harboring the putative ER membrane targeting sequence of the botryococcus squalene synthase.

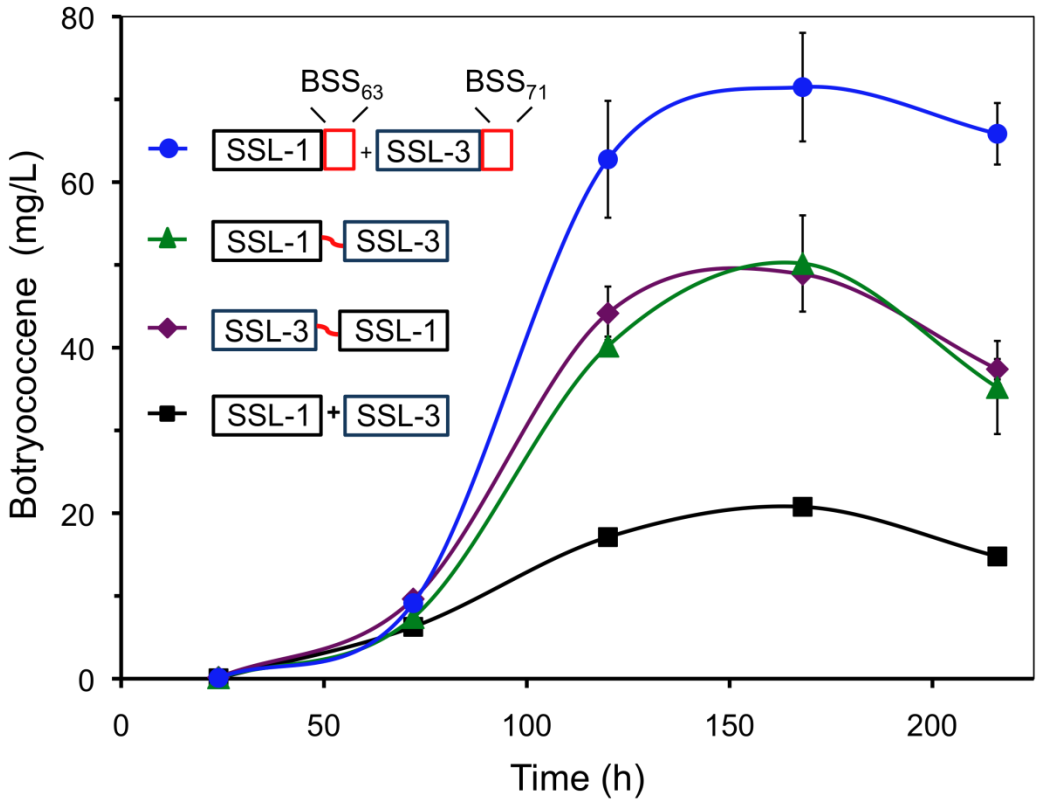


Figure 2.9 Comparison of botryococcene production in yeast engineered with different configurations of SSL-1 and SSL-3.

Yeast line TN7 was engineered with the SSL-1 and SSL-3 genes on separate plasmids (squares), with gene fusions (SSL-1 fused to SSL-3 *via* a triplet repeat of GGSG (triangles), or vice versa (diamonds)), or with 63 or 71 amino acids of the carboxy terminus of the *Botryococcus* squalene synthase, sequences containing a membrane-spanning domain, appended to the carboxy termini of the SSL-1 and SSL-3 enzymes, respectively (circles). The data represents mean ± S.E.M.

Discussion

The results were unexpected because squalene biosynthesis is known as a two-step process catalyzed by a single enzyme (Figure 2.10). FPP is first converted to the intermediate PSPP, followed by its reductive rearrangement to squalene (14). PSPP is not evident in these reactions unless NADPH, the reducing agent, is omitted from the incubations (17). Under conditions of adequate NADPH, it is unlikely that PSPP is released from the squalene synthase enzyme, then re-bound and reduced to squalene (40). Regardless, a single enzyme is responsible for the entire conversion process, and this mechanism appears highly conserved from yeast to man, including algae like *Botryococcus* (37). In contrast, botryococcene biosynthesis requires the successive action of two distinct enzymes. First, SSL-1 catalyzes the biosynthesis of PSPP as a separate and distinct product, which the second enzyme, SSL-3, efficiently converts to botryococcene in a NADPH-dependent manner. Whatever the evolutionary forces driving this division of labor might have been, it appears to have occurred twice within the life history of *Botryococcus*. When SSL-1 is co-expressed with SSL-2, squalene accumulates, which we speculate might represent a distinct pool of squalene in *Botryococcus* destined to specialized roles like the biogenesis of the extracellular matrix and other squalene derivatives.

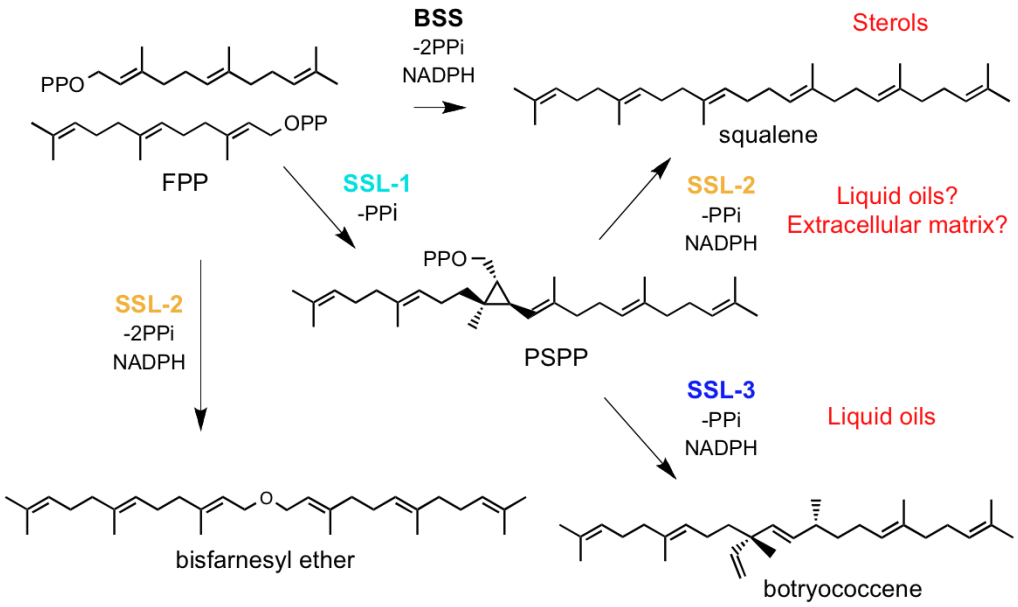


Figure 2.10 A cartoon depiction of the catalytic roles of the squalene synthase-like enzymes in *Botryococcus braunii* race B and their putative contributions to the triterpene constituents that accumulate.

The previously identified squalene synthase gene (BSS) (37) is thought to provide squalene essential for sterol metabolism, whereas the squalene synthase-like genes SSL-1, SSL-2, and SSL-3 provide the triterpene oils serving specialized functions for the algae. In combination with SSL-1, SSL-2 could provide squalene for extracellular matrix and methylated squalene derivatives, while SSL-1 plus SSL-3 generates botryococcene, which along with its methyl derivatives, accounts for the majority of the triterpene oil.

Support for the neofunctionalization of these unusual binary systems for triterpene biosynthesis is provided by the distinctive biosynthetic activities associated with SSL-2 (Figure 2.10). First, this enzyme catalyzes the NADPH-dependent biosynthesis of an unusual terpene ether. There are no reports of bisfarnesyl ether accumulation in *Botryococcus* or any other organism, but it could be incorporated into other more complex matrix polymers which would have masked its detection. One possible means for bisfarnesyl ether biosynthesis does not involve a PSPP intermediate, but instead an alternative reactivity of two bound farnesyl moieties *via* a S_N2 Williamson ether synthesis-type reaction (Figure 2.3) (38). Support for such a mechanism comes from the incorporation of radiolabeled FOH directly into the bisfarnesyl ether product, but only when SSL-2 is incubated with both FOH and FPP (Table 2.1). Second, the accumulation of both squalene and bisfarnesyl ether in yeast co-expressing SSL-1 and SSL-2 is also consistent with this proposed mechanism. The yeast line used for these studies is engineered for high FPP production, but tends to accumulate FOH as a consequence of FPP dephosphorylation catalyzed by endogenous phosphatases (41, 22). Hence, the yeast co-expressing SSL-1 and SSL-2 has significant pools of FOH and FPP, which will compete with any PSPP generated by SSL-1 for binding and catalysis by SSL-2. Third, while there is no obvious or direct chemical requirement for reducing equivalents in the biosynthesis of the bisfarnesyl ether from FPP and FOH, the significance of the NADPH dependence might relate to a structural role rather than a catalytic one. Pandit *et al.* (35) suggested that NADPH binding to its putative bind site in the human squalene synthase might stabilize a region of the enzyme not well resolved in the crystal structure, and thus positioning a domain into close association with the active site. NADPH binding to the SSL-2 enzyme could evoke a similar conformational change that renders the SSL-2 enzyme competent for either bisfarnesyl ether or squalene biosynthesis dependent on available substrates (FPP, FOH and PSPP). Hence, not only has SSL-2 maintained its catalytic ability to convert PSPP to squalene, it has evolved a novel catalytic activity yielding a bisprenyl ether from prenyl diphosphates.

One possibility for how these unique triterpene synthases arose is that a progenitor squalene synthase gene could have duplicated to yield multiple gene copies. While one copy (BSS) maintained its coding capacity for squalene synthase activity, essential for sterol metabolism, the other copies (SSL-1, SSL-2, and SSL-3) would have afforded opportunities for evolutionary diversification. Alternatively, *Botryococcus* could have acquired multiple copies of SSL genes by a horizontal gene transfer process, and those genes may have evolved specialized synthase-like activities. For example, one of the acquired squalene synthase like genes could have evolved the capacity for botryococcene biosynthesis and a subsequent gene duplication event could have resulted in loss of function for either the first half-reaction or the second (45). No matter

the specific mechanism, what makes the possible events associated with the neofunctionalization of the SSL enzymes particularly intriguing is that specialized triterpene oil accumulation, like botryococcene, could not have occurred without both SSL-1 and SSL-3 evolving in concert with one another.

There are other examples of similar division and diversification of enzymological capacities within key genes for pyrimidine (42), diterpene (43), and triterpene (44) metabolism. For instance, biosynthesis of the diterpene, kaurene, in many fungi relies on a single, multifunctional enzyme (45) that catalyzes the conversion of the linear isoprenoid intermediate, geranylgeranyl diphosphate, to the bicyclic copalyl diphosphate (CPP) product. CPP then undergoes a second cyclization reaction initiated at a separate binding site on the same enzyme to yield kaurene. In higher plants, the enzymes for CPP and kaurene biosynthesis are encoded by separate and distinct genes (43). Specific CPP synthases within rice catalyze the biosynthesis of either *ent*-CPP or *syn*-CPP isomers (46, 47). These are complemented with equally distinct diterpene synthases that can utilize one or the other CPP isomer for hormone or defense compound biosynthesis (48, 49). Yet, there are other diterpene synthases that have retained these two enzyme functions, but have evolved whole new catalytic outcomes (50). Osbourn and co-workers (44, 51) have provided evidence that the genes encoding for the enzymes catalyzing the cyclization of oxidosqualene to distinct tetra- and penta-cyclic classes of triterpenes, primarily sterols and defense related saponins, respectively, likely arose from common ancestor genes evolving novel catalytic functions dedicated to primary and specialized metabolism. Microbial forms of dihydrosqualene synthase, like CrtM, might also be considered an example of squalene synthase-like enzyme diversification (52, 53). CrtM relies on PSPP biosynthesis, but does not utilize NADPH for the second half-reaction. CrtM instead yields dehydrosqualene, a reaction product with much in common with phytoene, the tetraterpene equivalent of dehydrosqualene, and by inference shares catalytic features of the second half-reaction in common with phytoene synthase. Nonetheless, what distinguishes the current results from all the others is there are no other known examples where the half-reaction specificity of squalene synthases appear separated from one another and subject to evolutionary diversification, except for that reported here for *Botryococcus*.

The family of squalene synthase-like enzymes in *Botryococcus* is also informative relative to the recent elucidation of the crystal structure of dehydrosqualene synthase (CrtM) of *Staphylococcus aureus,* a target enzyme for a new generation of anti-infective reagents, along with refinements in the human squalene synthase structure (52, 53). Those studies detailed how two FPP molecules bind to CrtM and human squalene synthase, are converted to the PSPP intermediate, and then re-positioned in the active site pocket in preparation for the second half-reaction. Key residues identified include those that coordinate magnesium ions for their interactions with the diphosphate substituents of the FPPs and PSPP, which are involved in both half-reactions. Based on sequence alignments (Figure 2.2), many of these residues (S19, Y41, R45, D48, D52, Y129, N168, and D177, numbering according to CrtM and annotated by a star above the

residue in Figure 2.2) appear conserved in the *Botryococcus* squalene synthase and all three of the SSL enzymes. Because SSL-2 and SSL-3 are deficient in PSPP biosynthesis, these particular residues are not by themselves sufficient for PSPP biosynthesis. Conversely, since SSL-1 can only catalyze the formation of PSPP, these same residues do not appear sufficient to initiate the second half-reaction. Amino acids at other positions are undoubtedly important for PSPP formation and the catalytic specificity of the second half-reaction, squalene *versus* botryococcene biosynthesis. Experiments to functionally define which amino acids at which positions are responsible for the enzymological specificity of these triterpene synthases will be significantly advantaged by having these unique *Botryococcus* SSL enzymes which are specialized to either the first half-reaction or the second.

Altogether, our results establish that botryococcene and squalene oils are synthesized in *Botryococcus braunii* race B by the combined action of separate and distinct squalene synthase-like enzymes. These findings have opened up new avenues for understanding the chemical specificity and diversification within this class of enzymes and provide a demonstration for the bioengineering and production of a potentially valuable petrochemical replacement.

Methods

Reagents

[1-^{3}H]Trans, trans farnesol, and [1-^{3}H]FPP were purchased from ARC (St. Louis, MO). All other reagents were purchased from Sigma (St. Louis, MO) unless stated otherwise.

Culturing of *B. braunii*

Botryococcus braunii Berkeley (Showa) strain was grown as previously described (16), except that cultures were aerated with filter sterilized air containing 2.5% CO_2. Algal cells were collected by vacuum filtration using a 20 µM nylon mesh, then scrapped into collection tubes, snap frozen in liquid N_2 and stored at -80ºC until further use.

Cloning SSL-1

Plaque lifts of a *B. braunii* cDNA library previously described by Okada *et al.* (37) were prepared and hybridized with the full-length *B. braunii* squalene synthase cDNA radiolabeled with [α-^{32}P]dCTP using a Prime-It kit (Stratagene). Hybridization was performed at 30°C in hybridization buffer consisting of 5X SSPE, 2X Denhardt's solution, 0.2% SDS, 100 µg mL $^{-1}$ salmon sperm DNA and 40% formamide (34). The plaque lifts were washed three times at room temperature for 5 min with 2X SSC, 0.1% SDS and hybridization detected by autoradiography. After 2 rounds of plaque purification, isolated plaques were converted to their plasmid forms according to the manufacturer's instructions (Stratagene), restriction digestions of the isolated plasmids compared, and only those showing distinctive differences to that for the *Botryococcus* squalene synthase examined further by automated DNA sequencing. DNA sequence of the squalene synthase-like (SSL-1) cDNA clone yielded a putative full-length open reading frame (ORF) coding for a 402 amino acid protein having a predicted molecular size of 45,692 daltons.

SSL-1 was cloned into the pET28a vector *via* the cloning site BamHI/XhoI in order to generate a SSL-1 protein with an amino terminal hexa-histidine extension to aid in purification. The SSL-1 gene was also inserted into two standard yeast expression vectors, YEp352 harboring an ADH1 promoter and Ura3 selection *via* the cloning site EcoRI/HindIII, and pESC harboring an AHD1 promoter and Leu2 selection *via* the cloning site BamHI/NotI (22).

Cloning SSL-2

Transcriptomic sequencing was performed using RNA pooled from *B. braunii* cultures ranging from 1-4 weeks after subculturing. cDNA samples were prepared for de novo transcriptome sequencing, and the resulting cDNA samples processed for DNA sequencing according to the emPCR Method Manual (Roche). Sequencing was performed on a Roche Genome Sequencer FLX and the data assembled using Newbler (Roche). The assembled DNA sequence data was computationally screened using the NCBI blast search function with the *Botryococcus* squalene synthase cDNA sequence as the query, which revealed a partial ORF with strong similarity to the amino terminal halves of BSS and SSL-1. To determine the full length sequence for this SS - like gene (termed SSL-2), an aliquot of plasmid DNA derived from the phage cDNA library described above (using the mass excision protocol as described by Stratagene) was used in PCR reactions with primers specific to SSL-2 and primers specific to the pBluescript SK- vector in attempt to amplify the missing 3' sequence of SSL-2. An 800 bp band was isolated, ligated into the pGEM T-Easy vector (Promega), and sequenced. This sequence information was used to deduce the full length sequence for SSL-2, which encodes for a predicted protein consisting of 465 amino acids and having a molecular size of 52,149 daltons.

The full length SSL-2 cDNA was cloned into the Yep352 yeast expression vector *via* the cloning site EcoRI/NotI. Because SSL-2 encodes for a protein with at least one predicted membrane-spanning region at its C-terminus, we designed a truncated form of SSL-2 in which 73 C-terminal amino acids were deleted (SSL-2-tr2). SSL-2-tr2 was cloned into the pET28a vector *via* the cloning site EcoRI/NotI.

Cloning SSL-3

The DNA sequence data obtained from a second transcriptomic profiling effort (www.jgi.doe.gov/sequencing/why/bbraunii.html) was combined with the first, and the combined dataset assembled with CLC Genomics Workbench (CLC Bio). Screening of this dataset with the Botryococcus squalene synthase revealed another ORF encoding a 383 amino acid squalene synthase-like (SSL-3) protein with a predicted molecular size of 44,127 daltons. SSL-3 was cloned into the pET28a and Yep352 vectors *via* the cloning sites EcoRI/NotI.

Protein Expression, Purification, and Enzyme Assays

The recombinant vectors were transformed into *E. coli* strain BL21(DE3) according to the manufacturer's recommendations (Novagen). The selected lines were grown with kanamycin selection at 37ºC with vigorous shaking until the cultures reached an optical density of ~0.8 (OD_{600} nm), then expression of the corresponding SSL gene induced by addition of 0.5 mM isopropylthio-□-D-galactoside (IPTG) and the cultures incubate for an additional 3 to 20 h with shaking at room temperature. One hundred mL of the culture was subject to centrifugation at 4,000 *g* for 10 min, the pelleted cells resuspended in 10 mL of lysis buffer containing 50 mM NaH_2PO_4, pH 7.8, 300 mM NaCl, 10 mM imidazole, 1 mM $MgCl_2$, 1 mM PMSF, 1% glycerol (v/v), then sonicated 4 x for 10 sec with a microprobe sonicator at 60% maximum power. The samples were cooled on ice for 2 min between sonication treatments. The sonicate was centrifuged at 16,000 *g* for 10 min at 4°C and the supernatant used for purification of the hexa-histidine tagged enzymes.

Purification of the bacterial expressed enzymes was afforded by the amino-terminal hexa-histidine tag using His-Select Cobalt affinity gel (Sigma) columns according the manufacturer's recommendations. Recovery of proteins with the expected molecular sizes was determined by SDS-PAGE. The purified protein fractions were concentrated using Amicon Ultra (0.5 ml, 10K) centrifugation filter units and stored in 300 mM NaCl, 20 mM Tris-HCl pH 7.5, 5 mM dithiothreitol, 2 mM $MgCL_2$, 50% glycerol (v/v) at -20°C for 2-3 weeks without noticeable loss of activity.

Typical enzyme assays were initiated by mixing aliquots of purified enzyme with 50 mM Mops, pH 7.3, 20 mM $MgCl_2$, 2.5 mM 2-mercaptoethanol, 10 µM [1-^{3}H]-FPP (~$2x10^5$ dpm total), and 2 mM NADPH in total reaction volume of 50 µL. Reactions were incubated at 37ºC for 1 h and then extracted with 100 µl n-hexane or 100 ul MTBE. Forty µL of the n-hexane or MTBE extract was then spotted onto silica TLC plates with authentic standards of botryococcene and squalene and developed with n-hexane, or standards of bisfarnesyl ether and developed with n-hexane:MTBE, 25:1. The standards were visualized with iodine vapors and TLC zones corresponding to the standards were scrapped and analyzed by scintillation spectrometry. If *Botryococcus* lysate was added to enzyme assays, typically 5 µL of lysate (corresponding to 10 µg total protein) prepared from *B. braunii* cells according to Okada *et al.* (50) was added. Cold assays were scaled up to 0.5 mL total volume and contained 10 µM FPP. Assays were extracted once with 1 mL n-hexane, then with 1 mL MTBE, the organic extracts pooled, and solvent evaporated under a stream of N_2. Extracts were resuspended in 50 µL hexane and an aliquot analyzed by GC-MS with a Varian CP-3800 GC coupled to a Varian Saturn 2200 MS/MS (Varian Medical Systems) using a Supelco SLB-5ms fused silica capillary column (30 m x 0.25 mm x 0.25 µM film thickness, Supelco). Initial oven

temperature was set at 220°C for 1 min., ramped to 280°C at 20°C/min., then ramped to 298°C at 3°C/min.

Expression in Yeast

A yeast line, CALI-7, developed previously for the generation of high intracellular concentrations of FPP was used for these purposes (41, 22). One further modification was an insertional inactivation of the ERG 1 gene. The insertional mutation of this gene was created by introducing the TRP1 gene flanked by DNA sequences of the 5' and 3' region of the ERG 1 gene into the CALI-7 cells and subsequent selection for reversion of tryptophan auxotrophic growth according to the method of Wang *et al.* (55). This modified yeast line capable of accumulating high levels of FPP but not metabolizing squalene is referred to as TN7.

The various recombinant yeast expression vectors were introduced into the TN7 yeast line *via* lithium acetate transformation, followed by selection for uracil and leucine auoxtrophic growth (22). Yeast lines were confirmed to possess the various expression vectors by colony PCR. Individual colonies of TN7 and the various TN7-transformants were subsequently grown in 25 mL of YPDE (nutrient rich) or Yeast Synthetic Drop-out medium (selection) for the indicated time at 25°C before analyzing the cultures for production of novel triterpene components. In brief, 1 mL aliquots of the culture were combined with 1 mL of acetone, vigorously mixed, and incubated at room temperature for 10 min. One mL of hexane was added and mixed vigorously for 60 sec. The mixture was then centrifuged briefly at 500 *g* to separate the phases, and the organic phase removed and concentrated to dryness under a nitrogen stream. The dried extract was resuspended in 50-500 µL of n-hexane and a 1 µL aliquot analyzed by GC-MS as described above.

SSL-1 and SSL-3 Yeast Expression Constructs

Fusion constructs were created by employing an assembly PCR strategy as described by Sun *et al.* (56). For creation of the SSL-1-SSL-3-fusion construct, oligonucleotide primers (5'-ccg**GAATTC**aaaacaatgactatgcaccaagaccacgg, EcoR1 restriction site in bold, and 5'-**ACCAGAACCACCACCAGAACCACCACCAGAACCACC**cttggtgggagttg gggctgcgc, (GGSG)x3 linker in bold) were used with SSL-1 as the template to amplify SSL-1 with a 3'-extension, and oligonucleotide primers (5'-**GGTGGTTCTGGTGGTGG TTCTGGTGGTGGTTCTGGT**atgaaacttcgggaagtcttgc, (GGSG)x3 linker in bold, and 5'-ataagaat**GCGGCCGC**ctaagcacccttagctgaaacc, NotI restriction site in bold) were used with SSL-3 as the template to amplify SSL-3 with a 5'-extension. The two PCR amplification products were purified and both used in a subsequent PCR reaction with the flanking oligonucleotide primers (5'-ccg**GAATTC**aaaacaatgactatgcaccaagaccacgg, EcoR1 restriction site in bold, and 5'-ataagaat**GCGGCCGC**ctaagcacccttagctgaaacc, NotI restriction site in bold). The fused amplification product was purified, digested with

EcoRI and NotI, and ligated into YEp352 (YEp352-SSL-1-3-fus). The SSL-3-SSL-1-fusion construct was created similarly (YEP352-SSL-3-1-fus), except that the oligonucleotide primers (5'-cgg**GAATTC**aaaacaatgaaacttcgggaagtcttgcagc, EcoRI restriction site in bold, and 5'-**ACCAGAACCACCACCAGAACCACCACCAG AACCACC**agcacccttagctgaaacctttcc, (GGSG)x3 linker in bold) were used with SSL-3 as the template and oligonucleotide primers (5'-**GGTGGTTCTGGTGGTG GTTCTGGTGGTGGTTCTGGT**atgactatgcaccaaga ccacgg, (GGSG)x3 linker in bold, and 5'-ataagaat**GCGGCCGC**ttacttggtgggagttggggctg cgc, NotI restriction site in bold) were used with SSL1 as the template.

The SSL1-BSS_{63} construct was created using the same assembly PCR methodology with oligonucleotide primers (5'-cgc**GGATCC**aaaacaatgactatgcaccaagaccacgg, BamHI restriction site in bold, and 5'-gcgctaacaacttggtgggagttggggctgcgcagaaagatttc) with SSL-1 as the template to amplify SSL-1 with a 3'-extension, and oligonucleotide primers 5'-ctcccaccaagttgttagcgctgacgggaggcagcttctacc, and 5'-ataaagaat**GCGGCCGC**ttaggc gctgagtgtgggtctagg, NotI site in bold) with BSS as the template to amplify the c-terminus of BSS with a 5'-extension. Following completion of the assembly PCR protocol, the amplification product was digested with BamHI and NotI and ligated into pESC (pESC-SSL-1-BSS_{63}). The SSL-3-BSS_{71} construct was created in the same manner except that oligonucleotide primers (5'-cgg**GAATTC**aaaacaatgaaacttcgggaagtcttgcagc, EcoRI restriction site in bold, and 5'-cgtcaaaggtagcacccttagctgaaacctttccatttgattttg) were used with SSL-3 as the template and (5'-gctaagggtgctacctttgacgaattgaggagcaggttgttagcg, and 5'-ataaagaat**GCGGCCGC**ttaggcgctgagtgtgggtctagg, NotI site in bold) were used with BSS as the template. The assembly PCR amplicon was ligated into YEp352 (YEP352-SSL3-BSS_{71}).

Purification of Bisfarnesyl Ether

TN7 yeast containing YEp352-SSL2 was grown in 1L YPDE media at 25°C for 8 days, after which hexane extracts were prepared. The raw yeast extracts were then subject to HPLC separation on a Waters 2695 HPLC with a Waters 2996 Photodiode Array detector (Waters Corporation) and a Develosil 60-3, 250 mm x 20 mm column (Nomura Chemical), run with an isocratic solvent (n-hexane:MTBE, 50:1) at 8 ml/min. Under these conditions, bisfarnesyl ether eluted at ~16 min. Further purification of the bisfarnesyl ether was afforded by successive chromatographic runs.

Synthesis of (2*E*,6*E*)-3,7,11-trimethyl-1-((2*E*,6*E*)-3,7,11trimethyldodeca-2,6,10-trienyloxy)dodeca-2,6,10-triene or Bisfarnesyl Ether

To 81 mg (2.03 mmole, 2 eq) of 60% sodium hydride (washed with anhydrous hexanes to remove mineral oil) in 1 mL of anhydrous THF under an argon atmosphere at 0^{o}C, 228 mg (1.02 mmol) of *trans,trans*-farnesol in 1 mL of anhydrous THF was added. The

mixture was stirred for 10 min, then 438 mg (1.53 mmol, 1.5 eq) of *trans,trans*-farnesyl bromide was added in 0.5 mL of anhydrous THF. The mixture was stirred for 20 h and allowed to warm to 25°C. The reaction was quenched with 2 mL of 1M HCl, diluted with EtOAc, washed successively with water and brine and dried over anhydrous $MgSO_4$. The crude product was chromatographed on silica gel F254 preparative TLC plates (Merck) in 1:10 EtOAc:hexanes to afford 174 mg (40%) of bisfarnesyl ether that had NMR and mass spectral data identical to that of material produced in enzymatic reactions (Figure 2.6).

NMR of Bisfarnesyl Ether

^{1}H and ^{13}C NMR spectra were recorded on a JEOL alpha 600 NMR spectrometer at 300K. Chemical shifts were referenced relative to solvent peaks, namely δ_H 7.24 and δ_C 77.00 for $CDCl_3$.

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.58 (br s, 6), 1.64-1.66 (m, 18), 1.92-2.14 (m, 16), 3.96 (d, J=6.8Hz, 4), 5.04-5.12 (m, 4), 5.32-5.38 (m, 2);

^{13}C NMR (100 MHz, $CDCl_3$) δ 16.2 ($CH_2C(\underline{C}H_3)=CHCH_2$), 16.7 ($CH_2C(\underline{C}H_3)=CHCH_2O$), 17.9 (*Z*-$CH_3$ of $(\underline{C}H_3)_2C=CHCH_2$), 25.9 (*E*-$CH_3$ of $(\underline{C}H_3)_2C=CHCH_2$), 26.5 ($((CH_3)_2C=CH\underline{C}H_2$), 26.9 ($CH_2C(CH_3)=CH\underline{C}H_2$), 39.8 ($\underline{C}H_2C(CH_3)=CHCH_2$), 39.9 ($\underline{C}H_2C(CH_3)=CHCH_2O$), 66.6 ($CH_2C(CH_3)=CH\underline{C}H_2O$), 121.3 ($CH_2C(CH_3)=\underline{C}HCH_2O$), 124.1($(CH_3)_2C=\underline{C}HCH_2$), 124.5 ($CH_2C(CH_3)=\underline{C}HCH_2$), 131.5 ($(CH_3)_2\underline{C}=CHCH_2$), 135.4 ($CH_2\underline{C}(CH_3)=CHCH_2$), and 140.2 ($CH_2\underline{C}(CH_3)=CHCH_2O$).

Chapter 3 Funtional Identification of Triterpene Methyltransferases from *Botryococcus braunii* race B

Introduction

Botryococcus braunii race B is a colony-forming, green algae that accumulates triterpene oils to levels in excess of 30% of its dry weight. The composition of the triterpene oils is dominated by di-methylated and tetra-methylated forms of botryococcene and squalene. While unusual mechanisms for the biosynthesis of the botryococcene and squalene were recently described, the enzyme(s) responsible for decorating these triterpene scaffolds with methyl substituents were unknown. Assuming that the triterpene methyltransferases (TMTs) might resemble the SAM-dependent enzymes described for methylating the side chain of sterols, a transcriptome of *B. braunii* was screened computationally and six sterol methyltransferase-like genes were isolated and functionally characterized. Three of these genes when co-expressed in yeast with complementary squalene synthase or botryococcene synthase expression cassettes resulted in the accumulation of mono- and di-methylated forms of both triterpene scaffolds. Unexpectedly, TMT-1 and TMT-2 exhibited preference for squalene as the methyl acceptor substrate, while TMT-3 showed a preference for the botryococcene as its methyl acceptor substrate. These *in vivo* preferences were confirmed with *in vitro* assays utilizing microsomal preparations from yeast over-expressing the respective genes, which encode for membrane associate enzymes. Structural examination of the *in vivo* yeast generated mono- and di-methylated products by NMR identified terminal carbons, C-3 and C-22/C-20, as the atomic acceptor sites for the methyl additions to squalene and botryococcene, respectively. These sites were identical to those previously reported for the triterpenes extracted from the algae itself. The availability of closely related triterpene methyltransferases exhibiting distinct substrate specificities and successive catalytic activities provides an important tool for investigating the molecular mechanisms responsible for the specificities embedded within these unique enzymes.

Background

Botryococcus braunii is a colony-forming, freshwater green alga that has attracted considerable interest because it reportedly accumulates hydrocarbon oils from 30 to 86% (10) of its dry weight and because these oils are considered progenitors to oil and coal shale deposits (1, 2, 3). While all *B. braunii* are morphologically similar, three distinct chemotypes of *B. braunii* have been reported depending on the type of hydrocarbons each accumulates (24). Race A accumulates fatty acid-derived alkadienes and alkatrienes (5); race L accumulates the tetraterpene lycopadiene (7); and race B amasses the linear triterpenes, botryococcene, squalene, and their methylated derivatives (6). Di- and tetra-methylated botryococcenes are generally the most abundant oil components accumulating in race B (57). Lower amounts of tetramethylated-squalene (25) and other structural derivatives of botryococcene and squalene ranging from C_{31} to C_{37} accumulate to various levels in different race B strains and in response to variable culture conditions (26). The oils accumulate both in intracellular oil bodies and in association with an extracellular matrix (8), which in race B consists mainly of long-chain, cross-linked polyacetals formed in large part from acetalization of polymethylsqualene diols that account for approximately 10% of the dry weight (21). Other polymethylsqualene derivatives have been described in race B, such as diepoxy-tetramethylsqualene (27), botryolins (28), and brauixanthins (20). The linear triterpenes, botryococcene, squalene, and their methylated derivatives, are hence common components of *B. braunii* race B and make up a large proportion of its total biomass.

A unique mechanism for botryococcene biosynthesis was recently described by Niehaus *et al.* (58), in which two squalene synthase-like (SSL) enzymes perform the successive half-reactions that are normally catalyzed by a single enzyme in the case of squalene synthase. SSL-1 uses farnesyl diphosphate (FPP) as a substrate to catalyze the production of pre-squalene diphosphate (PSPP), which a second enzyme, SSL-3, converts to botryococcene in an NADPH-dependent manner. A third enzyme, SSL-2, catalyzes the biosynthesis of squalene from PSPP produced by SSL-1 but cannot efficiently use FPP as a substrate. Overall, it was suggested that the squalene and botryococcene produced by the SSL enzymes were channeled into the production of the liquid oils and the biosynthesis of squalene derivatives, such as the extracellular matrix (58), while the conventional *B. braunii* squalene synthase (37) synthesized squalene destined for sterol biosynthesis.

It is not botryococcene and squalene, however, which accumulate to substantial levels in this alga, but the methylated forms of these tripterpenes. For instance, while the liquid oil content of *B. braunii* race B is composed primarily of botryococcenes, generally less than 1% is in the non-methylated C_{30} form and the majority is dominated by dimethylated and tetramethylated forms, depending on the strain or culture conditions (26, 59). Essentially all the squalene that accumulates is in methylated forms, accumulating in the oil fraction (less than 5% of the total oil (60)) or incorporated into a variety of other *B.*

braunii natural products (20, 21, 27, 28). Because *B. braunii* race B accumulates 40% or more of its dry weight as these triterpene components, we estimate that the methylated triterpenes account for at least 30 to 35% of the total algal biomass dry weight and contribute directly to the buoyancy that distinguishes these algal colonies. Unlike many green algae that are flagellated and phototaxic (61), the buoyance characteristic of *Botryococcus* provides a means for it to float in its normal aqueous habitats and have greater access to light for photosynthesis. In addition to these purported physiological roles, the methylated forms of botryococcene and squalene enhance their utility as feedstocks for petrochemical processing and chemical manufacturing. The increased branching evident in the methylated triterpenes improves their hydrocracking to chemical species of value for the synthesis of industrial polymers and other commodity based chemicals (23) and yields high quality gasoline, kerosene and diesel fuels upon distillation (12).

While the unique mechanisms for C30 botryococcene and squalene biosynthesis in *Botryococcus braunii* have been elucidated (58), the specific mechanism(s) by which these triterpenes are methylated was unclear at the start of this work. Small-molecule methylation has been extensively characterized for many diverse compounds and typically consists of a methyltransferase (MT) that utilizes the universal methyl-donor, S-adenosyl methionine (SAM), and exhibits variable degrees of selectivity for a wide range of methyl acceptor molecules (62). MTs are also distinguished as C-, O-, N-, S- or halide methyltransferases, an indication of the methylation target within the acceptor substrate (63). While MTs may only share limited overall amino acid sequence similarities, domains responsible for SAM binding (64) appear to be broadly conserved and conserved structural folds have served to associate MTs into five distinct Classes (65). Most of the small molecule MTs fall into Class one but do not appear to cluster phylogenetically based on their target site (*i.e*. methylation of carbon versus nitrogen) or the particular chemical class of the methyl acceptor substrate (66). An indole alkaloid MT, for instance, shows closer sequence similarity to a tocopherol MT rather than any other alkaloid specific MTs. Clustering in this instance appears more related to the evolutionary origins of the MTs and the propensity of MTs to undergo neofunctionalization.

In the current work, we supposed that the methylation of botryococcene and squalene to their tetra-methylated forms was catalyzed by a single multi-functional MT or multiple MTs with each responsible for one particular methylation in the catalytic cascade of each triterpene. We also assumed that because squalene and botryococcene are chemically and biochemically related to sterols, the methylated forms of triterpenes in *B braunii* possibly arose within the same evolutionary timeframe as sterols and the mechanisms for sterol biosynthesis. Consequently, we searched for homologs of sterol MTs within a transcriptome of *B. braunii*, and we provide here the first functional identification of several interesting triterpene methyltransferases.

Results

Identification of triterpene methyltransferase candidate genes

We predicted that a methyltransferase that could act on squalene or botryococcene might resemble a C-24 sterol methyltransferase (SMT) because these enzymes act on the linear isoprenoid side chain of sterols. A *B. braunii* transcriptomic database (58) was screened computationally for cDNAs showing amino acid sequence similarities to the *A. thaliana and C. reinhardtii* SMT-1 enzymes. The BLAST search revealed six candidate genes that were greater than 42% identical and 59% similar to the *C. reinhardtii* SMT-1 (Figure 3.1). For comparison, the *A. thaliana* genome contains three predicted SMT genes (67), and the *C. reinhardtii* genome contains only one SMT gene (68). These particular genes appear overrepresented in *B. braunii* compared to other plant and alga and enhanced the prospects these could be triterpene methyl-transferases (TMTs). Amino acid alignments revealed that all six candidate genes share three conserved SAM binding sites as identified by (69); however, the sterol binding domain II, which is highly conserved in all known plant SMTs, is also conserved in three of the candidates (SMT-1,-2, and -3), but not in the other three (TMT-1,-2, and -3) (Figure 3.1).

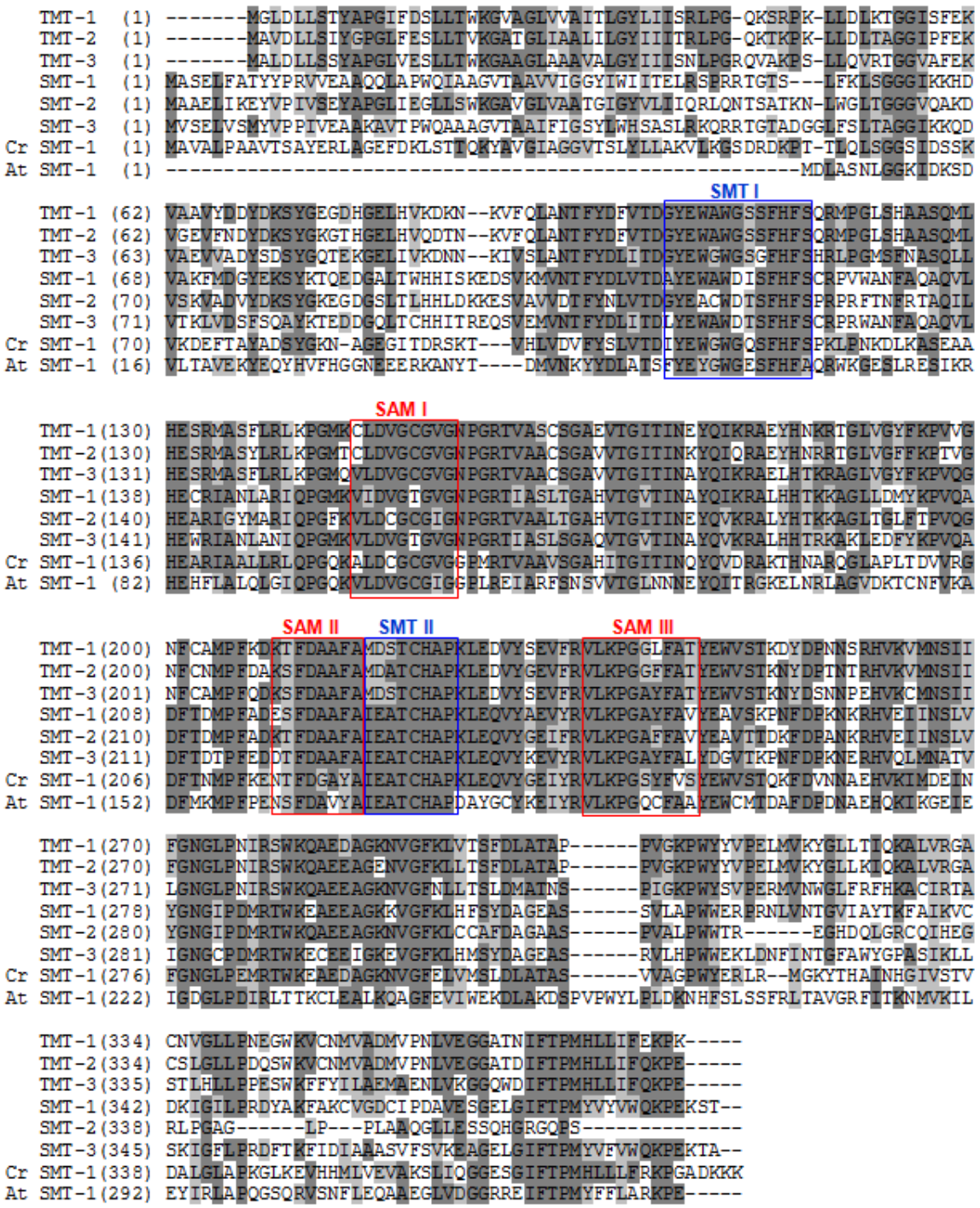


Figure 3.1 Amino acid alignment of the six sterol C-24 methyltransferases-like genes from *B. braunii* Race B.

Amino acid alignment of the six sterol C-24 methyltransferases-like genes from *B. braunii* Race B along with those of *C. reinhardtii* (EDP05221) and *A. thaliana* (AAG28462). Conserved sterol-binding domains (SMT) and S-adenosyl methionine-binding domains (SAM) as identified by (69) are boxed and labeled in blue or red, respectively.

In vivo functional characterization of MT activities

To screen the six candidates for TMT capabilities, we coexpressed the various SMT-like genes in TN7 yeast engineered with either *B. braunii* squalene synthase (BSS) or a construct in which SSL-1 and SSL-3 are fused with a $(GSGG)_3$ amino acid linker and also contains the 73 C-terminal amino acids of BSS fused to its C-terminus (SSL-1-3m). TN7 yeast engineered with BSS or SSL-1-3m can accumulate squalene or botryococcene, respectively, to levels above 100 mg/L (Figure 3.2, Figure 3.3B and C). When SMT-1, -2, or -3 were coexpressed with either BSS or SSL-1-3m, no distinct products could be detected in organic extracts by GC-MS analysis; however, coexpression of TMT-1, -2, or -3 all resulted in the accumulation of several unique products (Figure 3.3D-I). Analysis of the mass spectra of the unique peaks showed parent ions of 424 and 438 amu (Figure 3.4), suggesting mono- and di-methyltated triterpenes, respectively.

When TMT-1 was co-expressed with BSS, 63% of the total squalenes accumulated as methyl-derivatives with 43% accumulating as dimethyl-squalene (Figure 3.3D and 21A). However, when coexpressed with SSL-1-3m, only 3% of the total botryococcenes accumulated as monomethyl-botryococcene and no dimethyl-botryococcene was detected (Figure 3.3G and 21D). Similarly, when TMT-2 was co-expressed with BSS, 40% of squalenes accumulated as methyl-derivatives with 31% accumulating as monomethyl-squalene (Figure 3.3E and 21B). Only 2% of total botryococcenes accumulated as monomethyl-botryococcene when co-expressed with SSL-1-3m (Figure 3.3H and 21E). When TMT-3 was co-expressed with BSS, approximately 18% of the total accumulating squalene was converted to its methyl-derivatives, with 11% of that was as monomethyl-squalene (Figure 3.3F and 21C). When TMT-3 was co-expressed with SSL-1-3m, 33% of the accumulating botryococcene was methylated with greater than half of that in the dimethyl-botryococcene form (Figure 3.3I and 21F).

While the conversion of botryococcene and squalene to their mono- and di-methyl derivatives were readily detected, no further methylated products (tri- and tetra-methylated) accumulated. We considered the possibility that multiple methyltransferases might act successively and cooperatively in the formation of C34 triterpenes, with one methyltransferase catalyzing the C30 to C32 conversion and another using C32 as a substrate to form a C34 triterpene. To test this possibility, yeast expressing either BSS or SSL-1-3m with TMT-1, TMT-2 or TMT-3 as well as one of the remaining five other SMT-like *B. braunii* genes were evaluated for their triterpene content. No unique products other than the C31 and C32 triterpenes observed in the yeast lines expressing only TMT-1, -2 or -3 (Figure 3.3) were detected by GC-MS analysis.

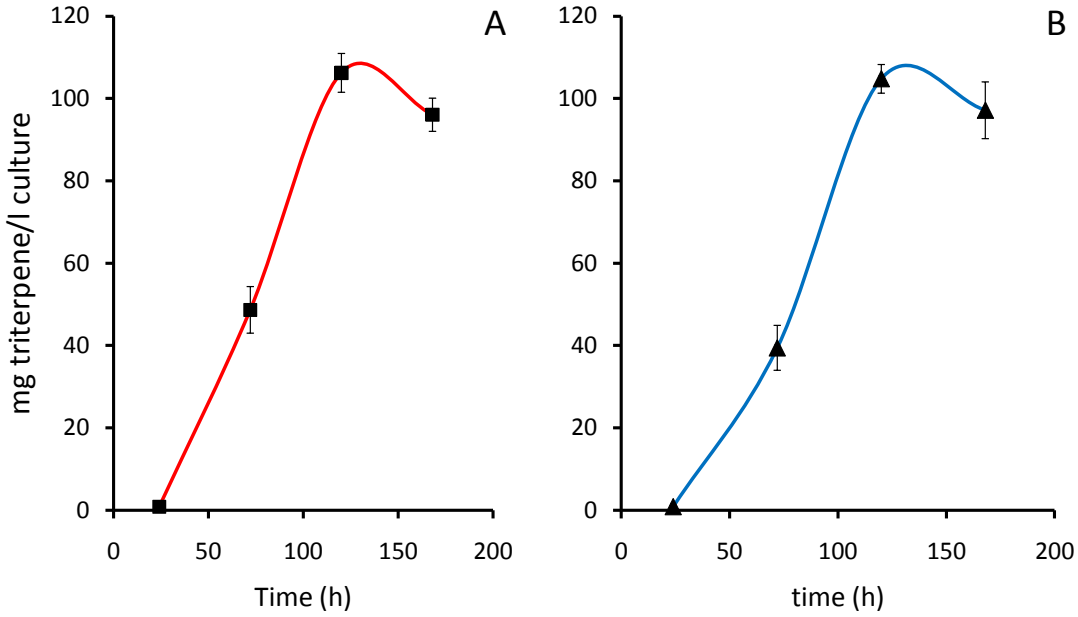


Figure 3.2 Accumulation of triterpenes in yeast engineered with various triterpene synthases.

Yeast line TN-7 was engineered with either *B. braunii* squalene synthase (A) or SSL-1-3m (B) and accumulation of squalene (squares), or botryococcene (triangles) measured. Yeast was grown in selection media in shake flasks at 30°C for the indicated time and organic extracts analyzed by GC-MS. Data represents mean ± S.E.M.

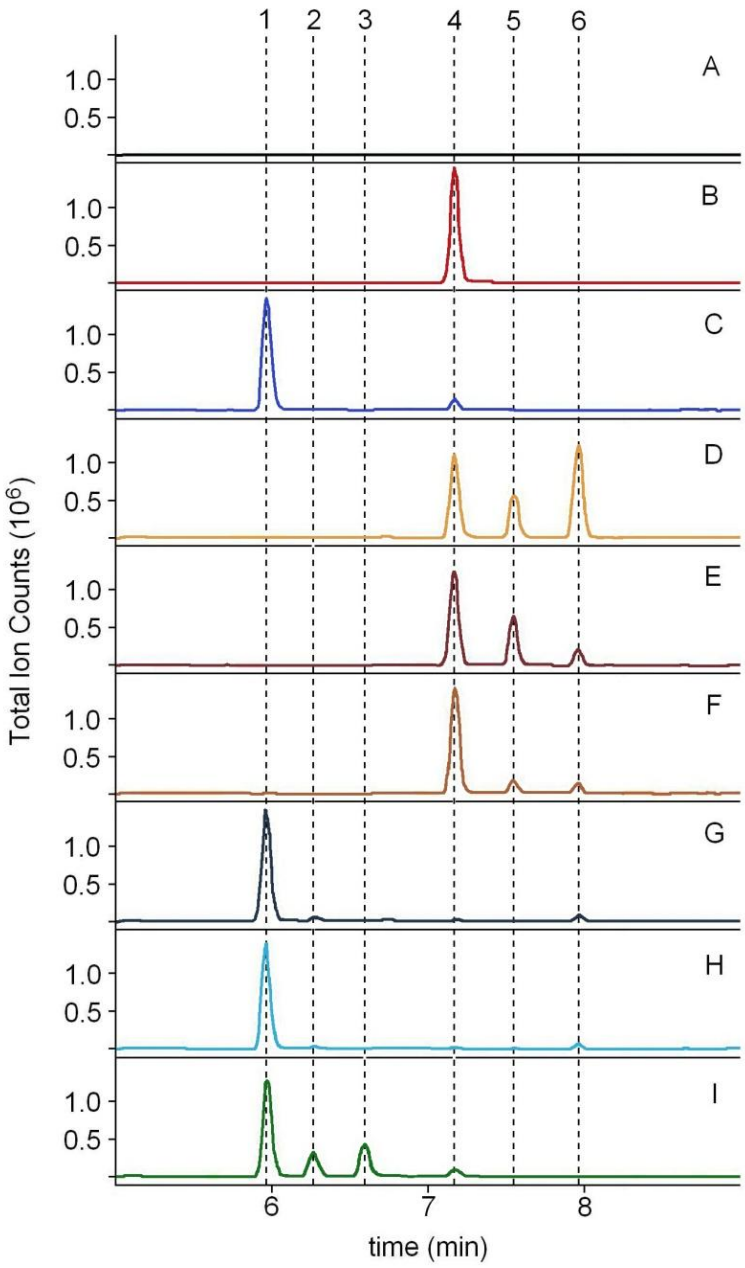


Figure 3.3 Functional characterization of *B. braunii* race B TMT genes.

Yeast expressing various combinations of triterpene synthase and TMTs were grown in shake flask for five days and organic extracts analyzed by GC-MS (chromatograms shown). TMT genes were co-expressed with BSS (squalene synthase) [TMT-1 (D), TMT-2 (E), and TMT-3 (F)], or SSL-1-3m (botryococcene synthase) [TMT-1 (G), TMT-2 (H), and TMT-3 (I)]. Yeast expressing only BSS (B) or SSL-1-3m (C) or only harboring empty expression vectors (A) serve as background controls. The chromatograms are annotated for the elution behavior of botryococcene (1), C_{31}-botryococcene (2), C_{32}-botryococcene (3), squalene (4), C_{31}-squalene (5), and C_{32}-squalene (6).

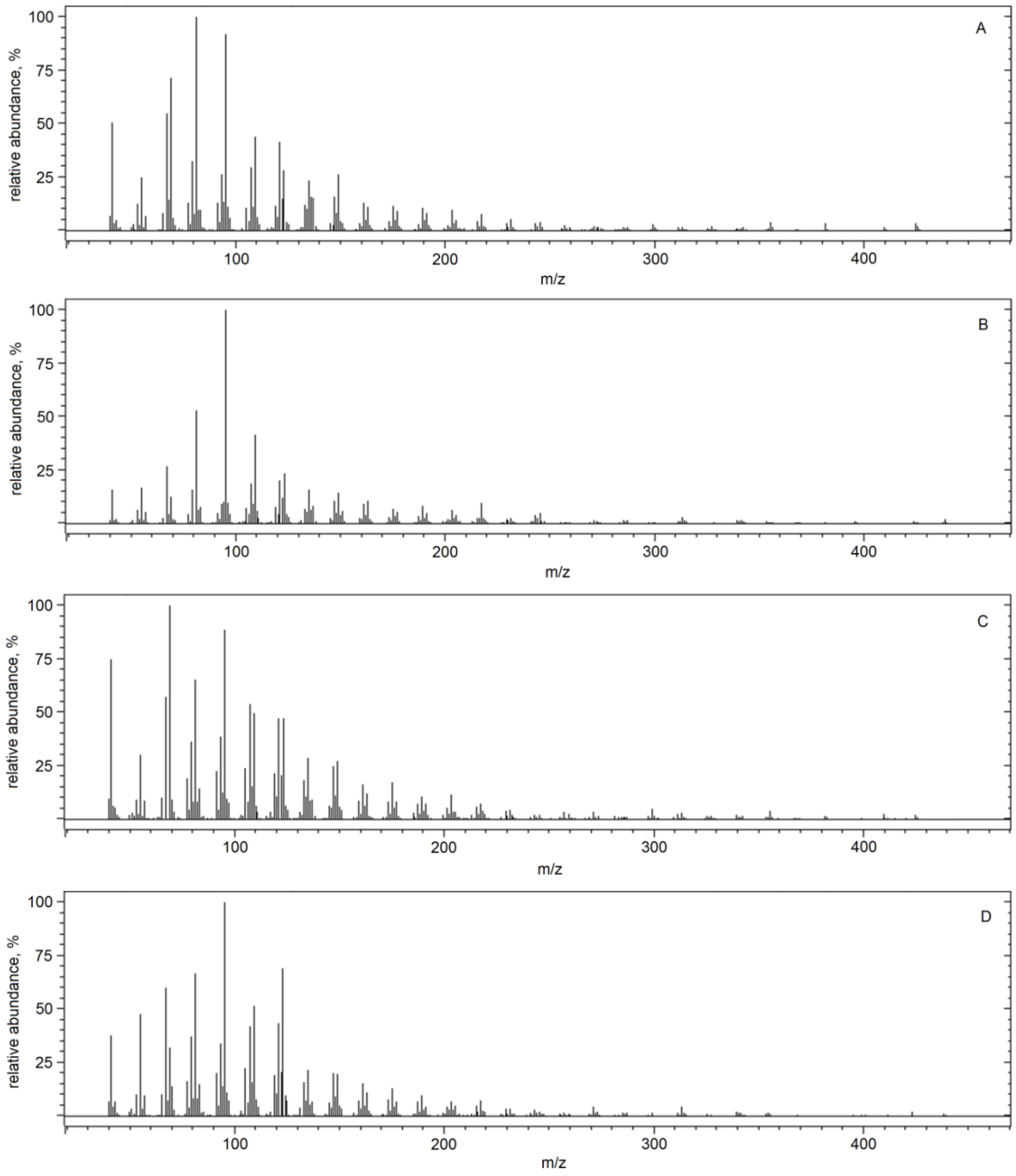


Figure 3.4 Mass spectra of methyl-triterpenes.

Mass Spectra of C31 squalene (A) and C32 squalene (B) produced in TN7 yeast expressing BSS and TMT-1, TMT-2 and TMT-3, C31 botryococcene (C) produced in TN7 yeast expressing SSL-1-3m and TMT-1, TMT-2 and TMT-3, and C32 botryococcene (D) produced in TN7 yeast expressing SSL-1-3m and TMT-3.

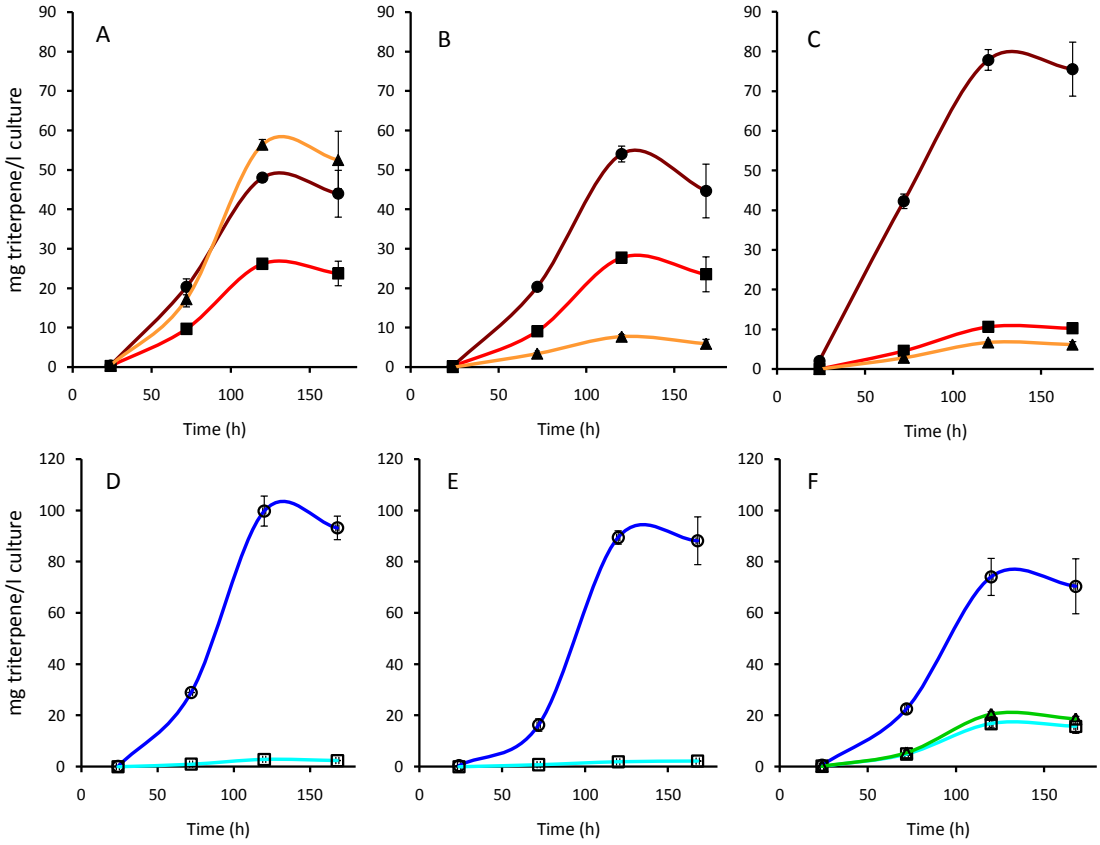


Figure 3.5 Accumulation of triterpenes in yeast engineered with various triterpene synthases and triterpene methyltransferases (TMT's).

Yeast were engineered with *B. braunii* squalene synthase (BSS) and either TMT-1 (A), TMT-2 (B), or TMT-3 (C) on separate plasmids and accumulation of squalene (closed circles), C_{31} squalene (closed squares), and C_{32} squalene (closed triangles) measured, or with the botryococcene SSL-1-3m and either TMT-1 (D), TMT-2 (E), or TMT-3 (F) on separate plasmids and accumulation of botryococcene (open circles), C_{31} botryococcene (open squares), and C_{32} botryococcene (open triangles) measured. Yeast was grown in shake flasks at 30°C for the indicated time and organic extracts analyzed by GC-MS. Data represents mean ± S.E.M.

In vitro biochemical confirmation

To verify the *in vivo* results with *in vitro* determinations, the six SMT-like genes were expressed in yeast and microsomal preparations used as the source of the enzymes in assays containing ^{3}H-SAM and either botryococcene or squalene as substrates. TMT-1 and TMT-2 readily catalyzed the transfer of a methyl group from SAM to squalene, but showed less than 1/100 those levels of activity with botryococcene as the acceptor (Table 3.1). In contrast, TMT-3 favored botryococcene as the methyl acceptor and exhibited only very modest activity with squalene. None of the other 3 SMT-like genes showed any measurable methyltransferase activity with botryococcene or squalene as substrates. None of the six enzymes was able to methylate C32 botryococcene or C32 squalene, possible intermediates to the tetramethylated forms (see below). Equally surprising, none of the six *B. braunii* SMT-like genes methylated cycloartenol, zymosterol, or lanosterol, which suggested we had yet to find the proper substrate(s) for these MTs, that we have not been able to provide these hydrophobic substrates in a form available for catalytic turnover, or that the MTs were not catalytic competent under these *in vitro* conditions.

Table 3.1 Substrate preference of the various *B. braunii* SMT-like enzymes.

	Enzyme activity (pmoles/h/µg)						
substrate	empty	TMT-1	TMT-2	TMT-3	SMT-1	SMT-2	SMT-3
squalene	0	513.7±8.6	862.2±59.9	35.4±3.0	0	0	0
botryococcene	0	3.3±1.3	4.5±1.3	434.9±31.8	0	0	0
C_{32} squalene	0	0	0	0	0	0	0
C_{32} botryococcene	0	0	0	0	0	0	0

The various *B. braunii* SMT-like genes or empty vector control were constitutively expressed in yeast for 3 days, after which microsomes were prepared according to the methods of Pompon *et al*. (*77*). Enzyme assays contained 50 mM HEPES, pH 7.5, either 0.01% (TMT-3) or 0.1% (all other samples) DHPC, 2 mM acceptor substrate (botryococcene, squalene, C_{32} botryococcene, C_{32} squalene), 50 µM ^{3}H-SAM (~150 dpm/pmole), 2 µL aliquots of microsomes (~ µg protein) in 100 µL final reaction volume. Assays were set up by first combining everything except ^{3}H-SAM and microsomes, and sonicating the mixture until the solution became cloudy due to micelle formation. The ^{3}H-SAM and microsomes were added, and the reaction incubated at 37°C for 5 min. Reactions were stopped by adding an equal volume of 10% KOH in methanol, followed by extraction of hydrocarbon products with 400 µL n-hexane. Aliquots of the hexane extract were separated by TLC and radioactivity incorporated into the triterpene fractions determined by scintillation counting. Data is reported as pmoles of methyl groups transferred to acceptor substrate per unit time and per µg of microsomal protein. Data represents mean ± S.E.M., n=3.

Chemical identification of the reaction products

To determine the specific methylation sites on squalene and botryococcene, the mono- and di-methylated squalenes and botryococcens produced *in vivo* by the engineered yeast were purified and subjected to ^{1}H and ^{13}C NMR analyses. The mono-methylated squalenes produced by yeast expressing TMT-1, -2, or -3 were all identical based on their NMR signals and methylated at the C-3 position of squalene (Figure 3.6, compound 2). Similarly, all the di-methylated squalenes produced by all three yeast lines gave identical NMR signals indicative of methylation at the C-3 and C-22 positions (Figure 3.6, compound 3). In contrast, the mono-methylated botryococcene produced by yeast expressing TMT-3 occurred at two positions, either the C-20 position yielding showacene (Figure 3.6, compound 5) or the C-3 position yielding isoshowacene (Figure 3.6, compound 6). Based on the relative intensity of the NMR signals for the methyl substituent at C-20 in showacene (compound 5 in Figure 3.6) *versus* that for C-3 in isoshowacene (compound 6 in Figure 3.6), showacene accounted for 63% of the total mono-methylated products. Di-methylated botryococcene produced by TMT-3 was methylated at the C-3 and C-20 positions (Figure 3.6, compound 7). The very small amounts of methylated botryococcenes produced in yeast expressing TMT-1 or TMT-2 were not sufficient for NMR analysis; however, the GC-MS patterns of mono-methylated botryococcene produced by TMT-1, -2, and -3 were all identical (Figure 3.4). These findings suggest that all the TMT's methylate botryococcene at identical positions.

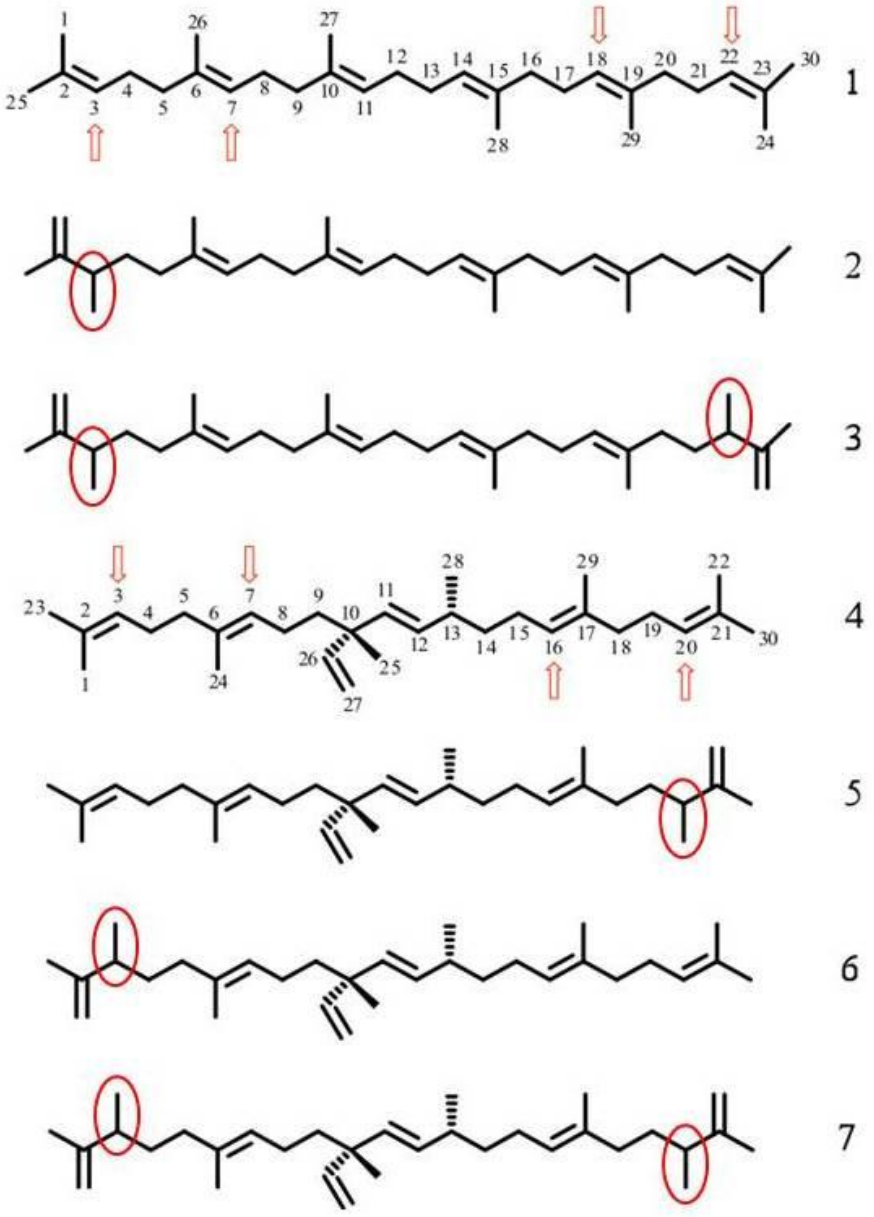


Figure 3.6 Structures of the various triterpenes accumulating in yeast expressing squalene synthase or botryococcene synthase in combination with TMT-1, -2 or -3.

Yeast expressing the squalene synthase (BSS) gene accumulates squalene (1), and C_{31} mono-methylated squalene (2) and C_{32} di-methylated squalene (3) when co-expressed with the TMT-1, TMT-2, or TMT-3 genes. Yeast expressing the botryococcene synthase expression cassette (SSL-1-3m) accumulates botryococcene (4), but a mixture of C_{31} mono-methylated isomers, showacene (5) and isoshowacene (6), and C_{32} di-methylated botryococcene (7) when co-expressed with TMT-3. Squalene and botryococcene have their carbons labeled, and the common sites of methylation are indicated with red arrows. The mono- and di-methylation sites with the triterpenes that accumulate in the respective yeast lines are highlighted with red circles. Methylation sites were assigned according to NMR signatures of the isolated compounds with reference to those previous reported (20, 60, 74, 75, 76).

Discussion

The large accumulation of triterpene oils by *Botryococcus braunii* race B has provided the impetus for considerable interest in elucidating the biosynthesis of these seemingly simple molecules. The oil is composed largely of linear, branched-chain triterpenes resembling squalene, yet the triterpene scaffold, botryococcene, is synthesized by the successive action of two enzymes rather than a single enzyme like that typical for squalene biosynthesis (58). While small amounts of the botryococcene and squalene triterpenes do accumulate, methylated forms of these molecules predominate and accumulate upwards of 40% of the total algal dry weight. These algae must possess a robust mechanism(s) for converting the triterpene scaffolds to their methylated forms, which also lend these molecules to a variety of industrial applications (70).

In the current effort, we identified three triterpene methyltransferases contributing to the methylation status of botryococcene and squalene, and we were surprised by several of the biochemical properties uncovered during this investigation. While we identified three triterpene MTs genes exhibiting sequence similarity to sterol methyltransferases (Figure 3.1), two of these encoded enzymes showed activity and specificity for squalene methylation. The third TMT appears to have specificity for botryococcene methylation. The specificity for squalene or botryococcene was unexpected because these molecules have very similar physical features. Nonetheless, the *B. braunii* TMTs were found to discriminate between the two methyl acceptors, as depicted in Figure 3.7, and this must arise from the ability of the respective enzymes to recognize differences of the internal linkages within squalene and botryococcene. TMT-3 must be able to recognize the internal ethyl, methyl substituents at C-10 of botryococcene, while TMT-1 and -2 must prefer the straight-chain linkage across C-11, -12, -13 and -14 of squalene.

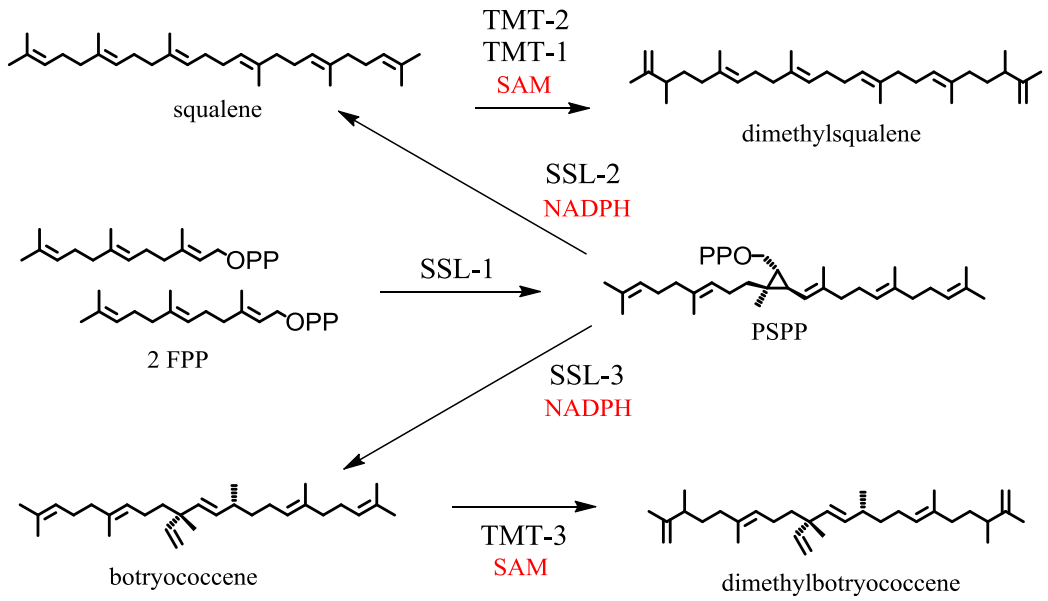


Figure 3.7 The methyl-triterpene biosynthetic pathways in *B. braunii*.

SSL-1 converts two FPP molecules to PSPP, which is converted in an NADPH-dependent manner to either squalene or botryococcene by SSL-2 or SSL-3, respectively. TMT-1 and TMT-2 can transfer a methyl group from SAM to squalene to form mono- or di-methyl squalene, while TMT-3 acts on botryococcene to form mono- or dimethyl botryococcene.

The substrate specificity of the TMTs is unexpected when one considers the symmetry and asymmetry of squalene, botryococcene and the mono-methylated intermediates, and the successive nature of these catalytic events. Most small molecule MTs catalyze mono-methylation reactions, with some notable exceptions such as the tri-methylation of phosphoethanolamine in the biosynthesis of the choline head group in phospholipid biosynthesis (71). In contrast, the successive methylation of the sterol side-chain at C-24 requires distinct enzymes, sterol methyltransferases 1 and 2 (72, 73.). The successive nature of the *B. braunii* TMTs appears to represent yet another permutation in the activities of this diverse family of enzymes. For TMT-1 and -2, the symmetry of squalene affords equal probability of methylation at either end of the molecule, but these enzymes also introduce a second methylation at the equivalent position on the other side of the molecule. While TMT-1 appears to perform this second methylation with great facility, this is not the case for TMT-2. The accumulation of di-methylated squalene exceeds that for mono-methylated squalene greater than 2-fold in yeast expressing the TMT-1 gene (Figure 3.5A), but di-methylated squalene only accumulates to approximately 20% of that for mono-methylated squalene in yeast expressing TMT-2 (Figure 3.5B). TMT-3 functionally resembles TMT-1 with regards to the ease with which it introduces the second methylation into the botryococcene backbone, that is, the accumulation of dimethyl-botryococcene slightly exceded that of monomethyl-botryococcene (Figure 3.5F). Based on NMR analysis of the mono-methylated botryococcene produced in yeast, showacene accumulated to higher levels than isoshowacene in an approximate ratio of 1.7 to 1.0. It is unclear whether this ratio arises from a preference for methylating botryococcene at the C-20 rather than the C-3 position with both monomethylated-botryococcenes serving as equal substrates in the second methylation reaction, or if both the C-3 and C-20 positions of botryococcene are methylated with equal efficiency but isoshowacene (methylated at C-3) is the preferred substrate for the second methylation reaction, or a combination of both possibilities. Regardless, the same 1.7 to 1.0 ratio of showacene to isoshowacene is seen in mono-methylated botryococcene isolated from *B. braunii* (73), suggesting that yeast expressing TMT-3 and SSL-1-3m recapitulate the same biochemical bias as observed in *B. braunii*.

The unique specificities of TMT-1, -2 and -3 offer opportunities to gain insights into functional aspects of these enzymes. Given the large collection of the class 1 MT crystal structures and their utility for molecular modeling and mapping residues important for catalysis in the wider family of MT enzymes, a similar strategy might facilitate identifying those regions of these *Botryococcus* MTs specifying substrate selectivity and target site selection for methylation.

It is important to note that the data provided here is not sufficient to conclude that we have identified all the mechanisms responsible for *in vivo* methylation of botryococcene and squalene in *Botryococcus braunii* race B. We have identified several genes encoding for methyltransferases capable of introducing terminal methyl substituents at

C-3 and C-22/C-20 of squalene and botryococcne. It is also important to point out that botryococcene and squalene accumulate in *B. braunii* largely in their tetramethylated forms. Hence, additional MTs or other mechanisms for the complete methylation pattern of these triterpenes remain to be discovered.

Methods

Cloning the SMT-like genes

The triterpene methyltransferase-3 (TMT-3) was identified through a random sequencing effort of ESTs using a *B. braunii* phage cDNA library as previously described (58). Briefly, phages were converted to their plasmid form using the mass excision protocol as described by the manufacturer (Stratagene), and ~500 individual colonies were randomly selected for automated DNA sequencing using sequencing primers flanking the cDNA insertion sites. Manually assembled cDNA sequences were then screened against the NCBI tBlastn search function across all available databases and TMT-3 was identified as exhibiting similarity to 24-C-sterol methyltransferase (SMT) genes. All other SMT-like genes were identified in a *B. braunii* 454 transcriptomic dataset as previously described (58). This dataset was screened computationally using a NCBI BLAST search window with the *C. reinhardtii* SMT-1 protein sequence (EDP05221) and the *Arabidopsis thaliana* SMT-1 sequence (AAG28462) as the queries, which led to the identification of six full-length ORFs that were at least 42% identical and 59% similar to *C. reinhardtii* SMT. Full sequence data is available from Genbank.

Primers flanked by the BamHI and NotI or EcoRI and NotI restriction enzyme sites were designed to amplify each of the six SMT-like genes from *Botryococcus braunii* mRNA, the amplification products digested with the corresponding restriction enzymes, then ligated into the standard yeast expression vectors YEp352-Ura or pESC-Leu (58). All constructs were verified by DNA sequencing.

Yeast expression

Yeast lines previously developed for high level accumulation of squalene and botryococcene were used for evaluating the putative triterpene methyltransferase genes (22, 41, 58). These lines consist of the TN7 parental strain harboring an insertional mutation in the native yeast squalene synthase gene (ERG 9) transformed with expression vectors containing either the full-length *Botryococcus* squalene synthase (BSS) gene (37) or a fusion of the Botryococcus SSL-1 and SSL-3 genes (functional equivalent of botryococcene synthase) including a sequence encoding for the carboxy-terminal membrane targeting domain of the *Botryococcus* squalene synthase protein (SSL-1-3m) (58). The various methyltransferase expression vectors were introduced into these two yeast line using the lithium acetate transformation protocol, followed by selection for complementation of the uracil and leucine auoxtrophic growth markers (41). Transformants were confirmed to possess the various expression vectors using colony PCR with primers selective for the methyltransferase genes. Individual colonies were subsequently grown in 30 mL of the appropriate Yeast Synthetic Drop-out medium (selection) containing 5 mg/l ergosterol for the indicated time at 30°C before analyzing the cultures for production of novel triterpene components. In brief, 1 mL aliquots of the

culture were combined with 1 mL of acetone, mixed vigorously, and incubated at room temperature for 10 min. One mL of hexane was added and mixed vigorously for 60 sec. The mixture was then centrifuged briefly at 500 *g* to separate the phases, and an aliquot of the hexane (1-3 µL) analyzed by GC-MS with a Varian CP-3800 GC coupled to a Varian Saturn 2200 MS/MS (Varian Medical Systems) using a Supelco SLB-5ms fused silica capillary column (30 m x 0.25 mm x 0.25 µM film thickness, Supelco). The initial oven temperature was set at 220°C for 1 min., ramped to 280°C at 20°C/min., then ramped to 298°C at 3°C/min.

Purification of mono- and di-methylated triterpenes

Yeast lines containing the respective triterpene synthase and TMT expression cassettes were grown in 1L Yeast Synthetic Drop-out medium media containing 5 mg/l ergosterol at 28°C for 8 days, after which hexane extracts were prepared. The crude extracts were then subject to HPLC separation on a Waters 2695 HPLC with a Waters 2996 Photodiode Array detector (Waters Corporation) and a Develosil 60-3, 250 mm x 20 mm column (Nomura Chemical), run in isocratic mode (100% n-hexane) at 8 ml/min. Under these conditions, C32 botrycooccene, C31 botryococcene, C32 squalene, and C31 squalene eluted at ~18, 22, 32, and 34 minutes, respectively. Repetitive chromatographic runs afforded further purification of the various compounds.

NMR of methylated triterpenes

^{1}H and ^{13}C NMR spectra were recorded on a JEOL alpha 600 NMR spectrometer at 300K. Chemical shifts were referenced relative to solvent peaks, namely δ_H 7.24 and δ_C 77.1 for $CDCl_3$. Each product was identified as shown in Figure X by referring ^{13}C chemical shifts for botryococcenes and methylsqualenes previously reported (16, 20, 31, 34, 35).

In vitro assays for the methyltransferase activities

The various *B. braunii* SMT-like genes or empty vector control were expressed in TN-7 yeast and grown in 100 mL selection media for 3 days, after which microsomes were prepared according to the methods of Pompon *et al.* (*77*). Enzyme assays contained 50 mM HEPES, pH 7.5, either 0.01% (TMT-3) or 0.1% (all other samples) DHPC, 2 mM acceptor substrate (botryococcene, squalene, C32-botryococcene, C32-squalene, cycloartenol, zymosterol, or lanosterol), 50 µM ^{3}H-SAM (~150 dpm/pmole), 2 µl microsomes, in 100 µl total volume. Assays were set up by first combining everything except ^{3}H-SAM and microsomes and treating with a sonicating water bath (Branson 2510) for ~1 min. until the solution became cloudy due to micelle formation, after which ^{3}H-SAM and microsomes were added and the reaction incubated at 37°C for 5 min. Reactions were stopped by adding an equal volume of 10% KOH in methanol, followed by extraction of hydrocarbon products with 400 µl n-hexane. An aliquot of the hexane

extract was spotted on silica TLC plates and developed with n-hexane:MTBE (25:1). Triterpenes were visualized with iodine vapor and the corresponding zones were scraped and subject to scintillation analysis.

Chapter 4 Enzymology of the SSL Enzymes

Introduction

An interesting feature of SSL-2 is that in addition to catalyzing the conversion of PSPP to squalene, it can also catalyze the formation of bisfarnesyl ether from FPP or FPP and farnesol. Bisfarnesyl ether catalysis is dependent on NADPH although no reducing equivalents are formally required. Catalytic constants were determined for bisfarnesyl ether production by SSL-2 (K_m = 57.1µM, k_{cat} = 1.7x10-3/s), which showed that this activity was nearly 100-fold less efficient than other squalene synthase-like enzymes. NADPH3 was able to stimulate bisfarnesyl ether biosynthesis consistent with the NADPH requirement of this reaction being mediated by a structural change in the enzyme, rather than serving in a catalytic role.

Background

Based on the observation that *B. braunii* race B accumulates very little bisfarnesyl ether or its derivatives, we have assumed that the primary role of the *B. braunii* SSL-2 enzyme is to catalyze the conversion of PSPP produced by SSL-1 to squalene (58). Nonetheless, an unusual and interesting aspect of SSL-2 is its ability to catalyze the biosynthesis of bisfarnesyl ether from either FPP or FPP and farnesol as substrates (58). Based on the proposed scheme for bisfarnesyl ether biosynthesis (see Figure 2.8), this compound may arise from an alternative of the typical squalene synthase-like catalytic cascade. In squalene synthesis, two molecules of FPP associate with the enzyme. One of the FPP molecules generates a reactive carbocation species that attacks the second FPP molecule to form PSPP. If reducing equivalents are present, PSPP can then undergo a reductive rearrangement to form squalene in an NAD(P)H-dependent manner (52). In the case of bisfarnesyl ether biosynthesis, the initial ionized FPP is quenched by a water capture mechanism and the resulting farnesol (FOH) is positioned to react with the other bound FPP molecule in a Williamson ether synthesis-like reaction to form bisfarnesyl ether (38).

There have been no prior reports of bisfarnesyl ether accumulation in any organism, including *B. braunii*, raising questions about whether this molecule is a true *in vivo* metabolite in *B. braunii* and whether it plays any physiological roles. Because molecular genetic tools for knocking out gene function or introducing gene silencing constructs are not available for *B. braunii*, one must relie on indirect means and correlations to infer the physiological contribution of an enzyme activity like that for bisfarnesyl ether biosynthesis by SSL-2. As a first approximation for assessing the *in vivo* contribution of SSL-2 towards bisfarnesyl ether biosynthesis, we have sought the catalytic constants for this enzyme in comparison to the activities of the SSL enzymes found in *B. braunii*.

Another interesting aspect of bisfarnesyl ether biosynthesis is that this reaction is NADPH dependent, even though there is no chemical requirement for reducing equivalents in its proposed biosynthetic scheme. Hence, it is not clear what role(s) this cofactor might play in the catalytic process. NADPH3, a substrate analogue of NADPH that is non-hydrolyzable, has previously been shown to stimulate the rate of PSPP production by squalene synthase. It binds to squalene synthase and induces a conformational change creating a more catalytic efficient enzyme (15). NADPH3 stimulation of bisfarnesyl ether production by SSL-2 would suggest that nicotinamide cofactors play a structural rather than catalytic role.

The work presented here describes our efforts to define some of the catalytic constants of SSL-2 and investigate the role of nicotinamide cofactors in this reaction.

Results

Bacterial expressed purified SSL-2 was used in enzyme assays to help characterize its *in vitro* activity. Initial experiments showed that 30°C was the optimal temperature for bisfarnesyl ether enzyme activity, with activity dropping off at higher temperatures and significantly declining at 42°C. All assays described in this chapter were performed at 30°C. Assays were set up with varying pH (Figure 4.1) that showed that activity is significantly inhibited at pH above 8.0.

Assays set up with varying concentrations of FPP (Figure 4.2) showed that as the concentration of FPP increases, the ratio of squalene to bisfarnesyl ether formation increases slightly from about 10% of the total product at 2 µM FPP to about 20% at 80 µM FPP.

The data from Figure 4.2 was used to determine catalytic constants of SSL-2 (Figure 4.3). The K_m of FPP for bisfarnesyl ether synthesis was found to be 57.1 µM with a v_{max} of 124.8 pmoles/µg/h. The k_{cat} was calculated to be $1.7x10^{-3}$/s.

The ability of the nicotinamide cofactors, NADP+, NADPH, and NADPH3 to stimulate bisfarnesyl ether production were assessed (Figure 4.4). While NADP+ did not stimulate bisfarnesyl ether production to detectable levels, NADPH3 stimulated ether synthesis to about 20% the level of NADPH.

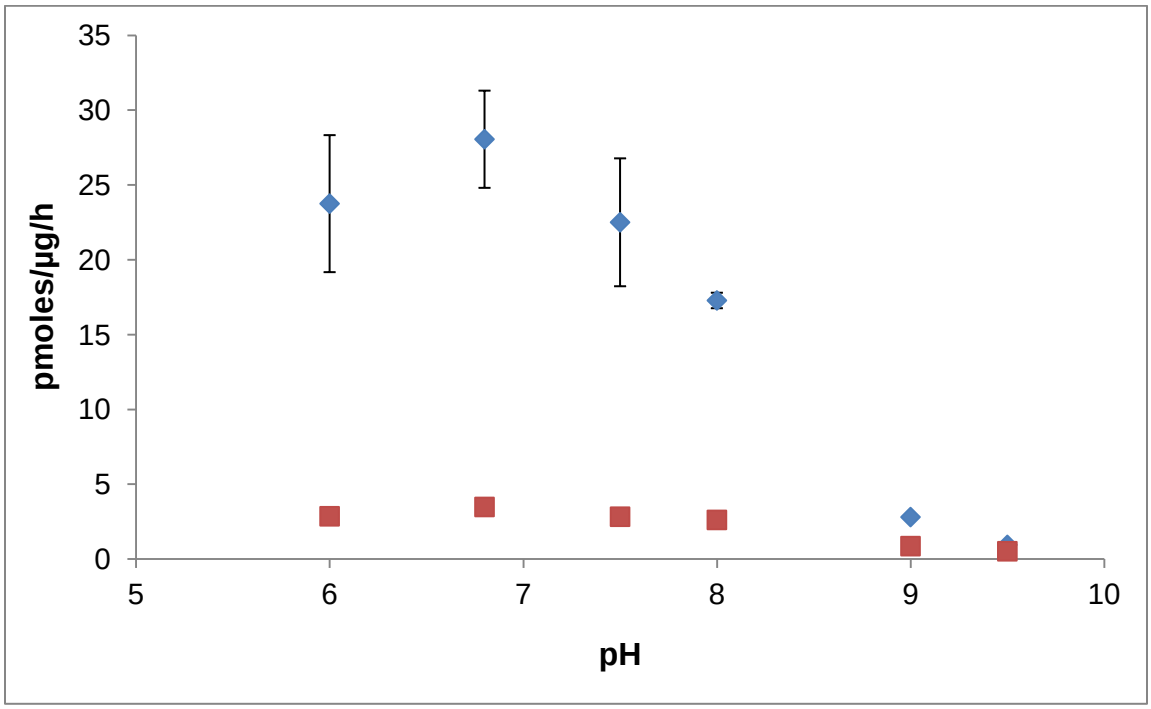


Figure 4.1 Enzyme activity of SSL-2 at varying pH.

Bisfarnesyl ether (blue diamonds) and squalene (red squares) synthase enzyme activity of SSL-2 with varying pH. Enzyme assays were conducted with 50 mM buffer (acetate pH 6.0 or Tris pH 6.8, 7.5, 8.0, 9.0, and 9.5), 20 mM $MgCl_2$, 2.5 mM β-mercaptoethanol, 2.5 mM NADPH, 10 µM ^{3}H-FPP ($4.8x10^5$ dpm total), and 2 µg bacterial expressed purified SSL-2 in 50 µl total volume. Reactions were incubated 15 min at 30°C and stopped by adding 50 µl of 0.5 M EDTA. Reaction products were extracted with 100 µl MTBE and an aliquot spotted on silica TLC plate with standards of squalene and bisfarnesyl ether. TLC was developed in n-hexane:MTBE (25:1), standards visualized with iodine vapor, and the corresponding zones were scraped and subject to scintillation counting. Data represents mean ± S.D., n=2.

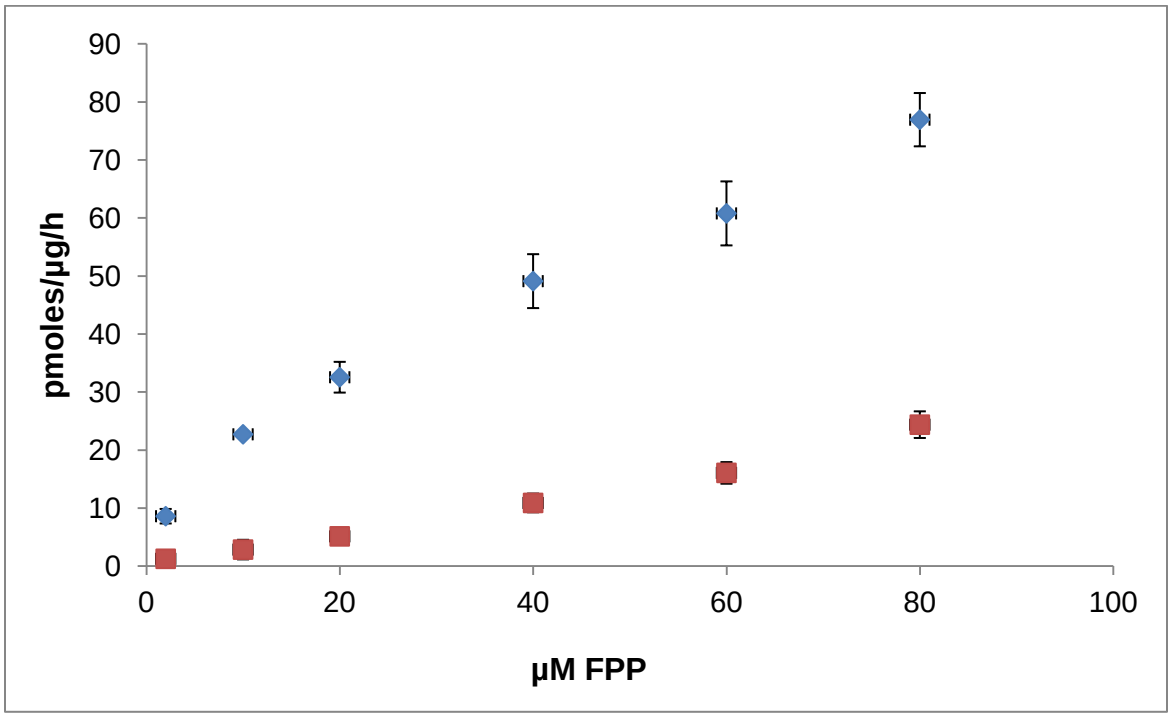


Figure 4.2 Enzyme activity of SSL-2 at varying FPP concentrations.

Bisfarnesyl ether (blue diamonds) and squalene synthase (red squares) enzyme activity of SSL-2 with varying concentrations of FPP. Enzyme assays were conducted with 50 mM MOPS, pH 7.3, 20 mM $MgCl_2$, 2.5 mM β-mercaptoethanol, 2.5 mM NADPH, the indicated concentration of ^{3}H-FPP ($4.8x10^5$ dpm total), and 2 µg bacterial expressed purified SSL-2 in 50 µl total volume. Reactions were incubated 15 min. at 30°C and stopped by adding 50 µl of 0.5 M EDTA. Reaction products were extracted with 100 µl MTBE and an aliquot spotted on silica TLC plate with standards of squalene and bisfarnesyl ether. TLC was developed in n-hexane:MTBE (25:1), standards visualized with iodine vapor, and the corresponding zones were scraped and subject to scintillation counting. Data represents mean ± S.E.M., n=3.

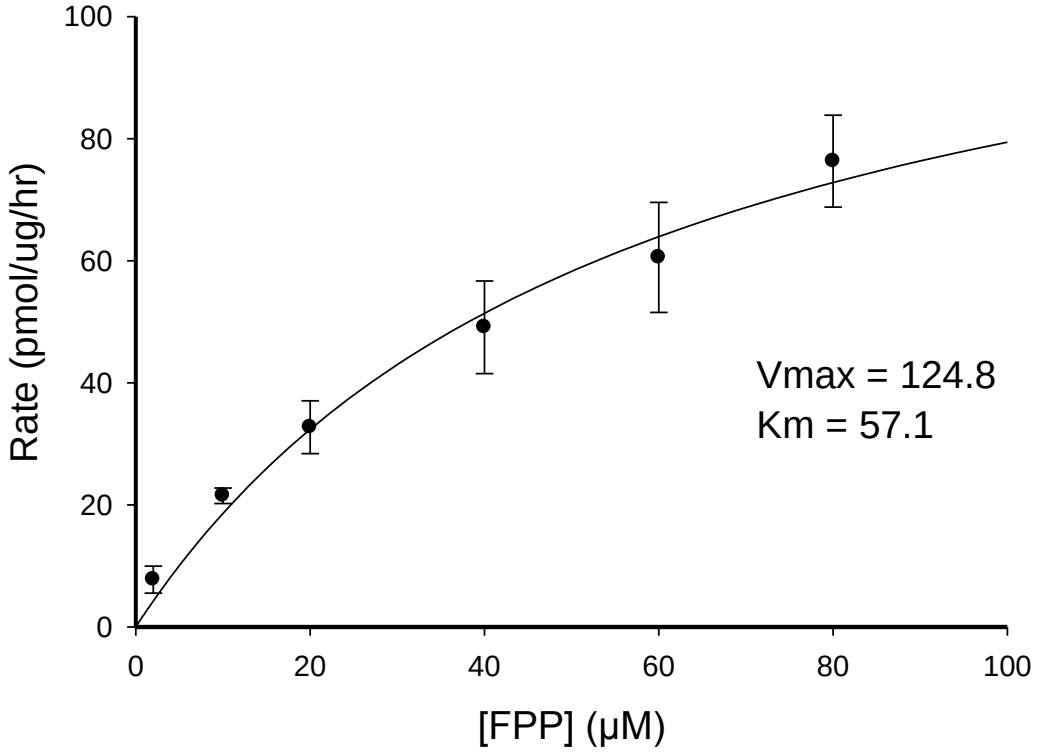


Figure 4.3 Enzyme kinetics of SSL-2.

Data from Figure 4.2 was analyzed using the SigmaPlot Enzyme Kinetics 1.3 module. Data represents bisfarnesyl ether synthase enzyme activity of SSL-2.

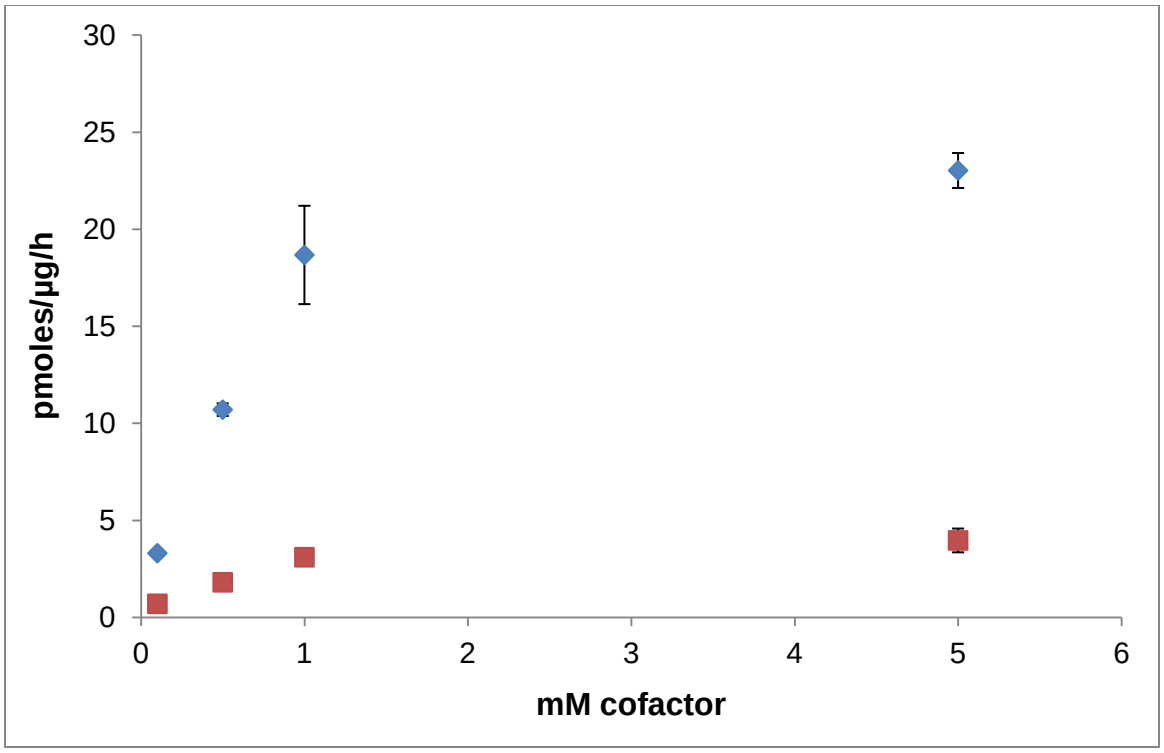


Figure 4.4 Effect of nicotinamide cofactors on SSL-2 enzyme activity.

Enzyme activity of SSL-2 with varying concentrations of NADPH (blue diamonds) or NADPH3 (red squares). Enzyme assays were set up with 50 mM MOPS, pH 7.3, 20 mM $MgCl_2$, 2.5 mM β-mercaptoethanol, the indicated concentration of NADPH or NADPH3, 10 µM ^{3}H-FPP ($4.8x10^5$ dpm total), and 2 µg bacterial expressed purified SSL-2 in 50 µl total volume. Reactions were incubated 15 min at 30°C and stopped by adding 50 µl of 0.5 M EDTA. Reaction products were extracted with 100 µL MTBE and an aliquot spotted on silica TLC plate with standards of squalene and bisfarnesyl ether. TLC was developed in n-hexane:MTBE (25:1), standards visualized with iodine vapor, and the corresponding zones were scraped and subject to scintillation counting. Data represents mean ± S.E.M., n=3.

Discussion

The optimal pH for bisfarnesyl ether biosynthesis was observed at pH 6.8, with enzyme activity decreasing sharply at higher pHs and almost completely lost at pH 9.5 (Figure 4.1). Activity was slightly less at pH 6.0 than 6.8, however it is unclear whether this is due to a lowering of enzyme activity or a consequence of spontaneous chemical oxidation of NADPH at lower pH values. For comparison, *B. braunii* squalene synthase exhibited only a slight loss of activity at pH 9.5 (data not shown).

The K_m of FPP for bisfarnesyl ether biosynthesis by SSL-2 was found to be 57.1 µM with a k_{cat} of $1.7x10^{-3}$/s. For comparison, SSL-1 was shown to have a K_m of 12.8 µM and k_{cat} of $2.7x10^{-2}$/s (Figure 2.4). If one compares the catalytic efficiency (k_{cat}/K_m) of SSL-1 and SSL-2, $2.1x10^{-3}$/s/M and $3.0x10^{-5}$/s/M respectively, SSL-1 is nearly 100-fold more efficient at converting FPP to PSPP than SSL-2 is at converting FPP to bisfarnesyl ether. This explains why enzyme assays containing both SSL-1 and SSL-2 accumulate only squalene and very little if any bisfarnesyl ether (see Figure 2.5H). With SSL-2 having such low efficiency with FPP as a substrate, it is unlikely that bisfarnesyl ether is significantly produced *in vivo* in *B. braunii* as long as there is sufficient SSL-1 enzyme present.

Bisfarnesyl ether biosynthesis is dependent on NADPH, but there is no chemical necessity for why NADPH was needed for this reaction to proceed (see table 2.2). NADPH may play a structural role when bound to SSL-2, resulting in a conformation change in the SSL-2 enzyme that makes the SSL-2 enzyme more catalytically competent for ether formation. However, NADP+ did not stimulate bisfarnesyl ether production even at concentrations as high as 50 mM. We tested an NADPH synthetic analogue, NADPH3, for its effect on SSL-2 and found that it stimulated ether biosynthesis to levels only about 20% of the NADPH stimulation. Because NADPH3 is a non-hydrolyzable co-factor, we conclude that the nicotinamide cofactors play a structural role in bisfarnesyl ether production by SSL-2.

Methods

Enzyme assays

SSL-2 enzyme was expressed in BL21 *E. coli* and purified as described in the methods section of chapter 2. Typical enzyme assays were set up with 50 mM MOPS, pH 7.3, 20 mM $MgCl_2$, 2.5 mM β-mercaptoethanol, 2.5 mM NADPH, 10 µM ^{3}H-FPP ($4.8x10^5$ dpm total), and 2 µg bacterial expressed purified SSL-2 in 50 µl total volume, unless otherwise indicated. Reactions were incubated 15 min. at 30°C and stopped by adding 50 µl of 0.5 M EDTA. Reaction products were extracted with 100 µl MTBE and an aliquot spotted on silica TLC plate with standards of squalene and bisfarnesyl ether. TLC was developed in n-hexane:MTBE (25:1), standards visualized with iodine vapor, and the corresponding zones were scraped and subject to scintillation counting.

NADPH3 synthesis

Synthesis of 12% Pd on $BaCO_3$ catalyst. To 1 g (4.10 mmol, 1.05 equivalents) of $BaCl_2$ in 10 mL of distilled water at 70^{o}C in a 50 mL centrifuge tube was added 8.2 mL of 5% Na_2CO_3 solution (3.86 mmol) dropwise over a 15 min period. The mixture was stirred at 70^{o}C for 1 h and centrifuged for 5 min at low speed. The supernatant was decanted. To the remaining white precipitate was added slowly 0.5 mL of 0.5M $PdCl_2$ in 3.5M HCl (1.75 mmol of HCl). When gas evolution ceased, 0.5 mL of 37% formalin (Fisher) was added. The mixture was stirred for 20 h at 25^{o}C to afford a black precipitate. The mixture was centrifuged, and the aqueous supernatant was discarded. The remaining catalyst was washed successively with 20 mL portions of water, ethanol and diethyl ether and was dried *in vacuo* to afford *ca*. 500 mg of 12% Pd on $BaCO_3$ catalyst. This calculated percentage takes into account the $BaCO_3$ neutralized when adding the $PdCl_2$/HCl solution. The catalyst was used within a week of its preparation in the following hydrogenation. Commercial samples of Pd on $BaCO_3$ proved ineffectual.

Synthesis of $NADPH_3$. A Lab-Crest pressure reaction vessel (100 mL volume) was charged with 97 mg of $NADP^+$, 17 mg of ammonium sulfate (130 mg/mmol of $NADP^+$), 188 mg of 12% Pd on $BaCO_3$ catalyst, and 10 mL of cold, distilled water. The tube was maintained at 0^{o}C until the hydrogenation commenced. The mixture was hydrogenated in a fume hood at 20 psi H_2 at 25^{o}C for 6 h. The mixture was filtered by gravity through filter paper to afford a clear, pale yellow solution.

NADPH3 purification

The resulting pale yellow solution from NADPH3 synthesis was filtered at 0.2 µM, divided into 2 equal aliquots, and lyophilized. The freeze-dried powder from one aliquot was resuspended in 1 mL of 50 mM NH_4HCO_3, and loaded on a 10 mL DEAE-Sephadex A-25 (Sigma) column equilibrated with 50 mM NH_4HCO_3 and chased with 10 mL 50 mM NH_4HCO_3. After completion of the 10 mL chase, a 20 mL gradient from 50 mM to 300 mM NH_4HCO_3 was run and 1 mL fractions were collected. Afterwards a 20 mL gradient form 300 mM to 1 M NH_4HCO_3 was run. The 1 mL fractions were analyzed for absorbance at 263 and 291 nm, and the fractions corresponding to the peak with the highest 263/291 nm ratio were pooled and lyophilized. An aliquot of the resuspended lyophilized product was subject to GC/MS analysis. The purified product had a parent ion of 748, while NADP+ had a parent ion of 744, confirming the product was indeed NADPH3.

Chapter 5 Identification of a carboxy-terminal, non-catalytic domain of fungal squalene synthases that is essential for directing squalene flux into the sterol biosynthetic pathway

Introduction

Expression studies of various heterologous squalene synthase genes in *S. cerevisiae* corroborated an earlier but surprising observation reported in the literature. While the squalene synthase gene of *S. cerevisiae* was able to complement an erg9 (squalene synthase) knockout in yeast, squalene synthase genes from plants and animals were not. Chemical profiles revealed that squalene accumulated to significant levels in yeast expressing the squalene synthase of plant, animal, or *S. cerevisiae.* Further, squalene synthase enzyme activity could be detected in lysates from yeast expressing the various synthases. This suggested that it was not the ability of these heterologous enzymes to produce squalene, but their inability to feed squalene into the native sterol biosynthetic pathway that prevented them from restoring normal ergosterol biosynthesis in *S. cerevisiae.* The ability of chimera squalene synthase enzymes to complement the erg9 mutation was examined. This led to the identification of a descrete sequence of amino acids near the C-terminus of the enzyme that is necessary and sufficient for allowing any squalene synthase to restore normal sterol metabolism.

Background

The first squalene synthase (SS) gene to be functionally characterized was isolated from *Saccharomyces cerevisiae* and cloned concurrently by the Karst and Robinson groups (78, 79). Both groups utilized the strategy of screening *S. cerevisiae* genomic library clones for their ability to functionally complement a squalene synthase (erg9)-deficient yeast line. Interestingly, Jennings, *et al.* (79) found that a genomic clone containing only a partial SS gene fragment was able to restore ergosterol prototrophy even though it only restored 5% of the normal level of SS enzyme activity. This finding suggested that low levels of SS enzyme activity were sufficient to complement the erg9 deficiency in yeast. Soon afterwards, Robinson, *et al.* (36) attempted to clone the SS gene from *Homo sapiens* and *S. pombe* using the same strategy, but isolated only the *S. pombe* gene by screening for complementation of the er9-deficient line (36). Having two SS genes from two species of fungi, these investigators were able to identify conserved regions within the deduced protein sequences to which they designed degenerate primers and cloned the human SS homolog using PCR (36). Robinson, *et al*. (36) confirmed that the human squalene synthase gene was unable to restore ergosterol prototrophy to the erg9-deficient yeast line, but a chimera SS gene constructed by combining a 5' region of the human gene containing the putative catalytic domain with a 3' region of the *S. cerevisiae* gene containing a membrane-anchoring domain was able to complement the erg9 deficiency. Robinson, *et al.* (36) suggested that the inability of human SS to functionally complement the erg9-deficient yeast line was due to problems with expression or stability of the human protein in *S. cerevisiae*. A few years later, Soltis, *et al.* (80) isolated a similar allele of the human squalene synthase gene by screening a human cDNA library with a rat squalene synthase gene probe. These investigators also determined that the human squalene synthase gene was not able to complement an erg9 deficiency in yeast. They were, however, able to document expression of the human squalene synthase gene in yeast by recording the corresponding protein by immuno-blotting methodology, as well as measuring inducible enhancement of SS enzyme activity. This result conflicted with the notion that a heterologously expressed SS was not able to complement the erg9 deficiency in yeast because of problems with transgene expression or protein stability in yeast, and Soltis, *et al.* (80) hypothesized that structural differences in the carboxy-termini of the yeast and human SS may affect localization or folding of the proteins in association with intracellular membranes.

The first plant SS was cloned from *Arabidopsis* (81) and soon after from *Nicotiana benthamiana* (82). Nakashima, *et al.* (81) failed to isolate an *Arabidopsis* SS gene by screening for complementation of an erg9 deficient yeast line, and instead screened plaques of an *Arabidopsis* cDNA library with a mouse squalene synthase cDNA probe. Hanley, *et al.* (82) used a degenerate primer/PCR approach to isolate a *N. benthamiana* SS, and likewise noted that the tobacco SS gene was unable to restore growth when expressed in an erg9 deficient yeast strain. Later, Kribii *et al.* (83) reported that the *Arabidopsis* genome contained two highly homologous SS genes organized in a tandem

array. This group confirmed that the *Arabidopsis* SS could not complement the erg9 (SS gene) disruption in yeast, but they measured significant SS enzyme activity in the microsomal fraction of these yeast. These investigators went on to show that a chimeric *Arabidopsis* SS gene containing a substitution corresponding to the 66 carboxy-terminal amino acids of *Arabidopsis* SS with 111 carboxy-terminal amino acids of the S. pombe SS were sufficient to restore prototrophic growth of the erg9 knockout in yeast without exogenous sterol. Radiolabeling studies were also performed with [^{3}H]-FPP fed to microsomes isolated from yeast expressing either the full length *Arabidopsis* SS or the *Arabidopsis-S. pombe* chimera SS genes, or from wild type yeast. Radiolabel was incorporated by either the wild type yeast microsomes or microsomes from the erg9-deficient yeast over-expressing the *Arabidopsis-S. pombe* chimera SS into squalene, squalene-2,3-epoxide, and lanosterol. However, when [^{3}H]-FPP was incubated with microsomes from erg9 deficient yeast expressing the full length Arabidopsis SS, only radiolabeled squalene was detected. No SS enzyme activity was detectable in the cytosolic (soluble) fractions of these yeast lines. These results strongly suggested that active SS was being expressed and targeted to membrane in all the constructs tested; however, the carboxy-terminal 111 amino acids of S. pombe were necessary for channeling of squalene into the ergosterol biosynthetic pathway (83).

In 2000, another fungal squalene synthase was isolated from *Yarrowia lipolytica* using a degenerate primer approach (84). The *Y. lipolytica* SS was found to complement an erg9 deficient yeast line, albeit the complemented yeast grew slower than the yeast complemented with the *S. cerevisiae* SS gene. Altogether, this result and those of the other investigators demonstrated that at least three different fungal SS could complement the erg9 knockout in *S. cerevisiae*, but no other SS isolated from animal or plant could accomplish this task.

In 2008, Busquets, *et al.* reported that of the two annotated SS genomic sequences in *Arabidopsis*, only one coded for a functional SS enzyme (85). Busquets, *et al.* also performed some fluorescence microscopy experiments to determine the intracellular location of Arabidopsis SS. GFP was tagged to the N-terminus of a full length SS, a SS lacking the equivalent of the carboxy-terminal 67 amino acids, or the GFP was fused directly to a gene fragment corresponding to that encoding for the carboxy-terminal 67 amino acids of the SS. All three constructs were transiently co-expressed in onion epidermal cells with an ER-targeted version of DsRed. Both the GFP linked to the full length SS and the carboxy-terminal 67 amino acids of SS co-localized with DsRed, which indicated that these two SS enzymes were localized to the ER membrane. The GFP-SS fusion lacking the carboxy-terminal 67 amino acids appeared localized to only the cytosol. These authors concluded that the membrane-spanning region at the carboxy-terminus of SS was critical for correct targeting of SS to the ER membrane (85).

These results and our observations that the algal *Botryococcus braunii* SS also could not complement the erg9 mutant in yeast suggested that it was not simply targeting of squalene synthase enzyme activity to the ER membrane of yeast that was important.

Some additional protein domain within the carboxy-terminal region of the yeast squalene synthase was necessary to facilitate the complementation phenotype. The work reported here represents our efforts to use the erg9 complementation test in yeast to map a specific peptide domain within the carboxy-terminal region of the yeast squalene synthase protein necessary for the complementation phenotype.

Results

Complementing the erg9 knockout mutation in yeast requires more than active squalene synthase enzyme activity and squalene accumulation

Various squalene synthases were cloned into the yeast expression vector, Yep352-ADH1, including those from *S. cerevisiae* (YSS), *B. braunii* (BSS), *N. benthamiana* (TSS), A*. thaliana* (ASS), and *R. norvegicus* (RSS), as well as a C-terminal truncated form of YSS in which 24 amino acids comprising the ER membrane-spanning region were eliminated (YSStr). These constructs were transformed into the yeast line CALI7-1, which has been selected for high level FPP biosynthesis (41) and has an erg9 knockout mutation. It is unable to synthesize sterols *de novo* and is dependent on exogenous ergosterol for growth. To test if introducing the various squalene synthases could restore ergosterol prototrophy to CALI7-1 yeast line, colonies testing positive for the respective SS transgene by PCR screens were grown in selection media containing ergosterol and serial dilutions were spotted on plates not containing exogenous ergosterol (Figure 5.1). To evaluate the encoded enzymes for catalytic activity, extracts prepared from the CALI7-1 yeast lines were assayed for squalene synthase enzyme activity. Constructs were also expressed in TN-7 yeast line and cultures were grown in liquid media with ergosterol and subsequently analyzed for squalene accumulation. Squalene levels were determined in TN-7 because in this yeast line, there is no possibility of squalene feeding into the ergosterol biosynthetic pathway. This allows for accurate comparisons of squalene production by the various SS enzymes.

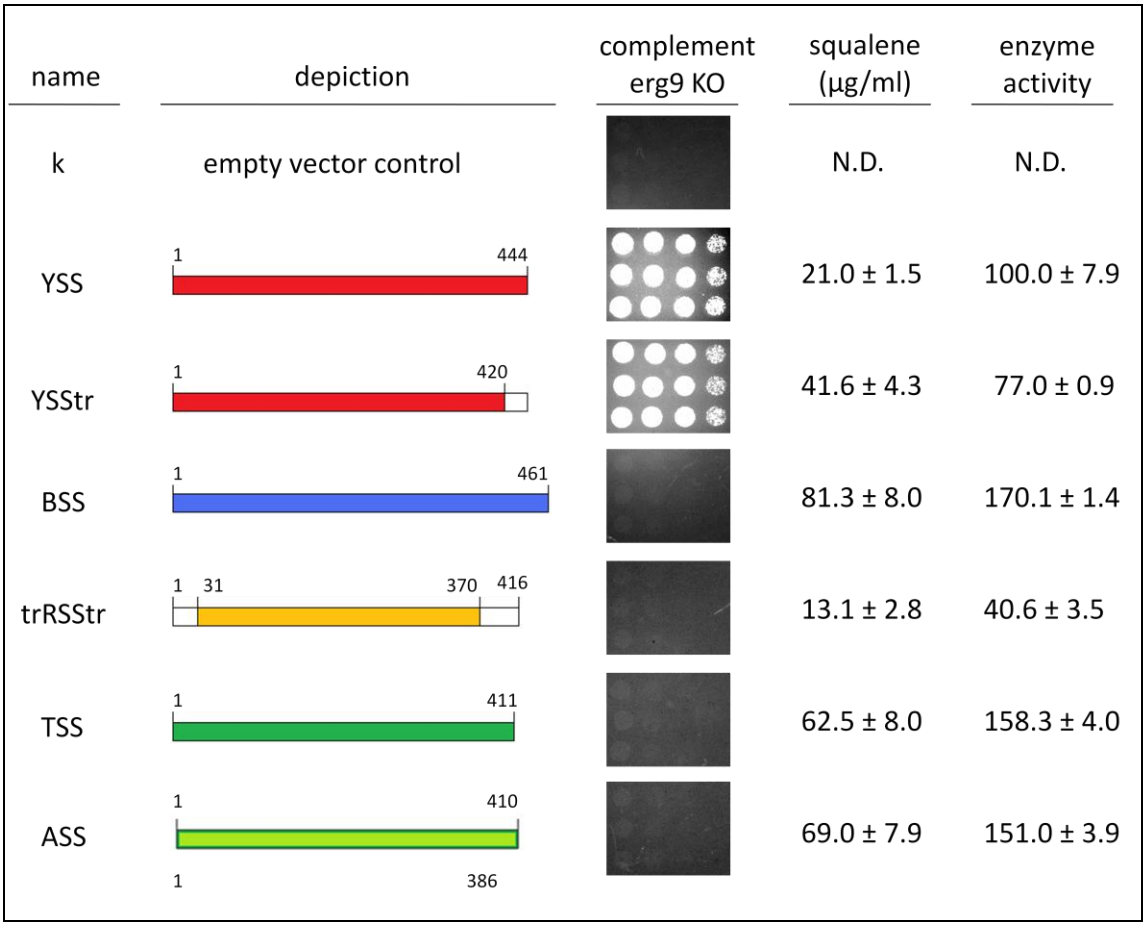


Figure 5.1 Various full length and carboxy-terminally truncated squalene synthases genes were expressed in yeast line CALI7-1 to test for their ability to complement the erg9 deletion.

Terminal truncations are indicated by a lack of color in the linear boxed enzyme model with the specific amino acid coordinates labeled above each. The squalene synthases tested are those of *S. cerevisiae* (YSS), *B. braunii* (BSS), *R. norvegicus* (RSS), *N. benthamiana* (TSS), and *A. thaliana* (ASS). For the erg9 complementation tests, three independent CALI7-1 transformants of each construct were randomly selected and grown in 2 mL Yeast Synthetic Drop-out media (Sigma) containing 5 mg/L ergosterol at 28^{o}C for three days, after which the culture was serially diluted with water to optical densities (600 nm) equal to 1, 0.2, 0.04, and 0.008, and 5 µL of each dilution spotted onto Yeast Synthetic Drop-out media plates without any exogenous ergosterol. Plates were incubated at 28^{o}C for 72 hours. Liquid cultures of each transformant in TN-7 line were grown in 10 mL of Yeast Synthetic Drop-out media containing 5 mg/L ergosterol at room temperature for seven days. Organic extracts were prepared and analyzed by GC-MS for their squalene content. To validate each gene construct, the squalene synthase enzyme activity encoded by each gene was assessed when the gene was expressed Cali-7 yeast. Briefly, 2000 x g supernatants were prepared and used in enzyme assays containing 3H-FPP and radiolabeled products separated by TLC and analyzed. Enzyme activity (pmoles/h/µg total protein) is recorded as a percent of YSS. n=3

Interestingly, only the full-length YSS gene and a carboxy-terminal truncated form (deletion of the terminal 24 amino acids) could restore ergosterol prototrophy to the Cali7-1 yeast line. Further, significant squalene accumulation and SS enzyme activity was observed in all squalene synthase constructs tested, suggesting that the enzyme was properly expressed in an active form in all cases (Figure 5.1). This suggested that all SS enzymes tested properly expressed in vivo, but only YSS could complement the erg9 deficient CALI7-1 yeast line.

The terminal membrane-spanning domain of squalene synthase is not necessary for functional enzyme nor complementation of the erg9 mutant

To corroborate and extend the earlier observations of Kirbii et al (83) that a non-catalytic, carboxy-terminal domain was important for complementation, we created reciprocal molecular swaps of the terminal 100 or so amino acids of various plant and algae SS enzymes with that corresponding to the 91 carboxy-terminal amino acids from *S. cerevisiae* SS and tested each for its ability to complement the erg9 mutation in yeast line CALI7-1 (Figure 5.2). As expected, all the constructs that contained the carboxy-terminus of *S. cerevisiae* SS were able to complement the erg9 knockout mutation in Cali7-1, while none of the *S. cerevisiae* SS enzymes that had their carboxy-terminus replaced with that of a plant or algae SS displayed ergosterol prototrophy.

Of equal interest is the observation that constructs containing a deletion of the terminal 24 amino acids of the yeast SS (YSStr, Figure 5.1) or appending this modified terminal domain to the algal SS (BSS-YSStr, Figure 5.2) did not alter the ability of these gene constructs to complement the erg9 deletion. On the basis of hydropathy plots of this region of the SS enzymes (36), this terminal region has been referred to as a membrane-spanning domain and by inference the domain mediating tethering of the SS enyzme activity to the ER membrane system in eukaryotic cells (36). Regardless of these inferences, the YSStr and BSS-YSStr constructs shown in Figs. 1 and 2 encode for functionally soluble SS enzyme activity (enzyme activity found in 20,000 g supernatants of *E. coli* expressing this genes) and these constructs complemented the erg9 mutation in yeast equally well as the full-length gene constructs. Hence, the carboxy-terminal 24 amino acids of the yeast SS are not necessary for the complementation phenotype, but some other element(s) within the proximal 67 amino acids of the carboxy-terminus appears to be both necessary and sufficient for complementation.

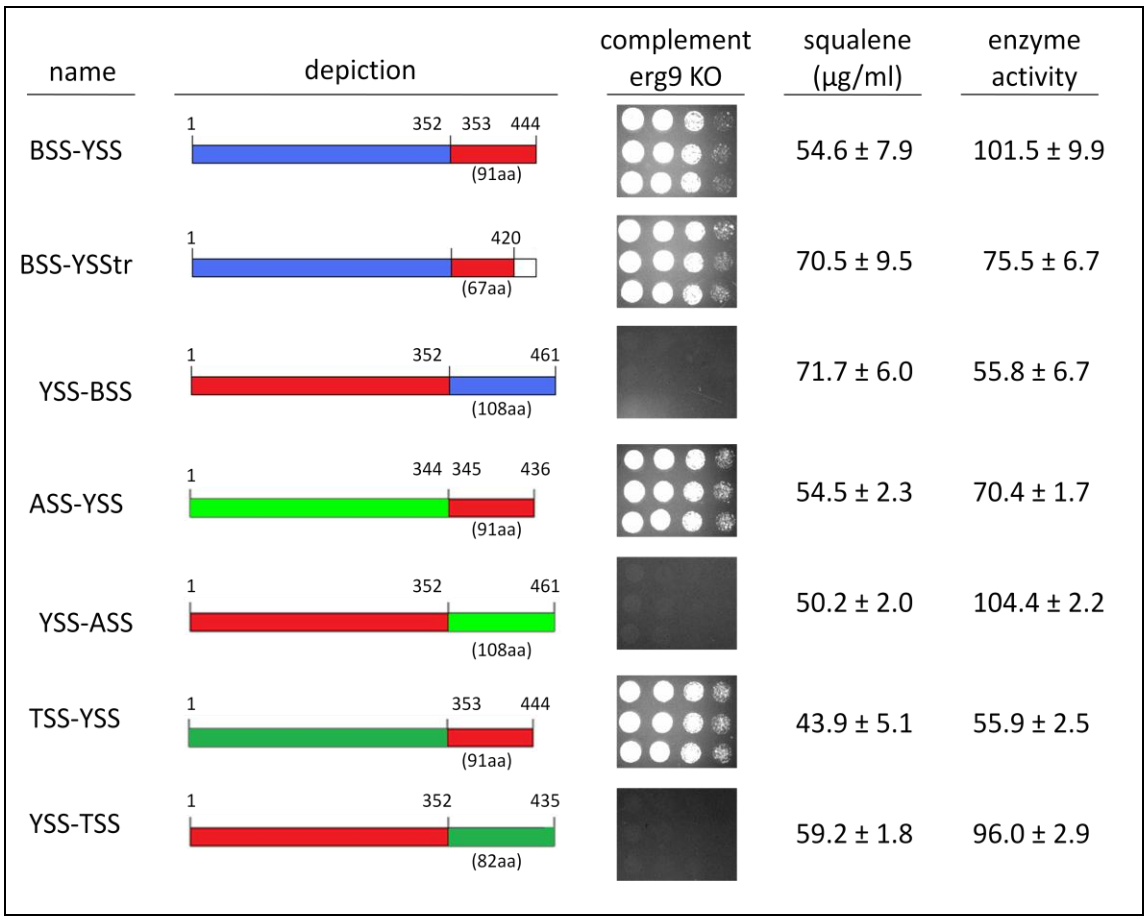


Figure 5.2 The carboxy-terminal amino acids of YSS are necessary and sufficient when appended to heterologous squalene synthase genes to confer ergosterol prototrophic growth to an erg9 knockout yeast line.

Constructs were created by reciprocal swapping of the DNA sequences coding for the 91 carboxy-terminal amino acids of YSS with the corresponding DNA segments of the algal (BSS, *Botryococcus* squalene synthase) and plant (ASS and TSS, *Arabidopsis* and tobacco squalene synthase). An additional 24 amino acid truncation of the YSS carboxy domain is indicated by a lack of color in the linear enzyme model. Each construct is annotated with the specific amino acids labeled above each depiction. Constructs were independently transformed into a yearst erg9 knockout line, and three independent transformants for each construct were randomly selected for growth in Yeast Synthetic Drop-out media containing 5 mg/l ergosterol. After three days, each culture was serially diluted with water to OD600=1, 0.2, 0.04, and 0.008, and 5 µl of each dilution spotted on Yeast Synthetic Drop-out media plates lacking ergosterol. Plates were incubated at 28°C for 72 h. Liquid cultures of each transformant in TN-7 line were grown in 10 mL of Yeast Synthetic Drop-out media containing 5 mg/L ergosterol at room temperature for seven days. Organic extracts were prepared and analyzed by GC-MS for their squalene content. To validate each gene construct, the squalene synthase enzyme activity encoded by each gene was assessed when the gene was expressed

Cali-7 yeast. Briefly, 2000 x g supernatants were prepared and used in enzyme assays containing 3H-FPP and radiolabeled products separated by TLC and analyzed. Enzyme activity (pmoles/h/µg total protein) is recorded as a percent of YSS. n=3

Computational screens for possible carboxy terminal domains responsible for the complementation phenotype

Because squalene synthase genes from *S. cerevisiae* (this work), *S. pombe* (36), and *Y. lipolytica* (84) have been demonstrated to complement the erg9 knockout in yeast, but squalene synthases from plant, algae and animals cannot, and because the results in Figs. 1 and 2 pointed to a proximal carboxy terminal region of 67 amino acids being responsible for the specificity of this complementation, amino acid sequence comparisons of this region between relevant squalene synthases were performed. No over-arching sequence similarities were observed when comparing the sequences across this region from algae, plants, animals and fungi, although there were greater similarities within the first 26 amino acids of this region (Figure 5.3A). This degree of similarity became much more apparent when the alignments of only the fungal squalene synthase were compared (Figure 5.3B). Within this short segment of amino acids, 8 residues are absolutely conserved and the degree of amino acid similarity across the entire 26 amino acids reaches upwards of 45%.

To functionally evaluate the contribution of this domain to the complementation phenotype, reciprocal constructs where these 26 amino acids of YSS were exchanged with the corresponding amino acids regions within the BSS and ASS were generated. These constructs were transformed into the erg9 knockout yeast line, and 3 independent colonies from each transformation were screened for their ability to grow in the absence of ergosterol (Figure 5.4). When the carboxy-terminal 26 amino acid sequence from S. cerevisae was substituted into the algal (BSS-YSS-BSS), complementation was readily apparent. When the carboxy-terminal 26 amino acid sequence from S. cerevisae was substituted into the Arabidopsis backbone (ASS-YSS-ASS), complementation was restored but the growth rate was noticeably affected. Further analysis revealed that squalene accumulation in this yeast line was only about 5% that of the wild type YSS. Further, squalene synthase enzyme activity in this yeast line was only 7.6% the level as the YSS-expressing yeast line. This result suggested that the ASS-YSS-ASS was compromised in its squalene synthase enzyme activity, but was still able to complement the erg9 knockout. The reciprocal substitutions of inserting the algal or plant amino acid sequences into the corresponding site of the S. cerevisiae SS resulted in a loss of its ability to complement the erg9 knockout mutations. The loss of this complementation capability was not due to a loss in the catalytic activity of the yeast SS. When these lines were grown in the presence of ergosterol, greater levels of squalene accumulated in these cultures relative to those lines transformed with the wild type YSS construct or other erg9 complementing constructs and SS enzyme activity could be readily measured in yeast lysates.

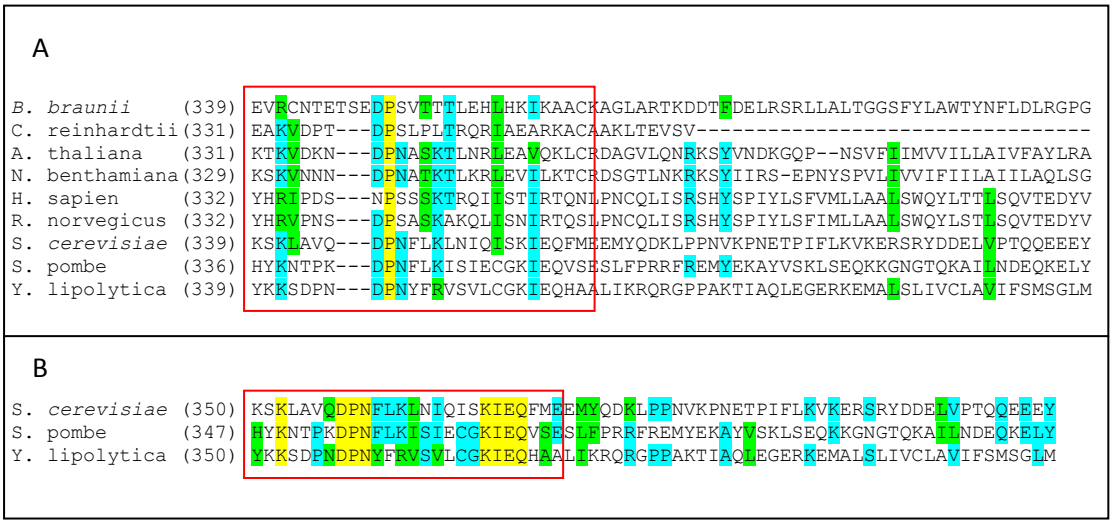


Figure 5.3 Alignment of the 26 amino acids.

Amino acid alignment of the *B. braunii (*AF205791*)*, *C. reinhardtii (*XM_001703395*)*, *A. thaliana (*NM_119630*)*, *N. benthamiana (*U46000.1*)*, *H. sapien (*NM_004462*)*, and *R. norvegicus (*NM_019238*)* squalene synthases relative to those for *S. cerevisiae (*X59959*)*, *S. pombe (*NM_001021271*)*, and *Y. lipolytica (*AF092497*)* (A). The alignment is limited to the sequences corresponding to the 67 amino acid domain of the *S. cerevisiae* squalene synthase that are necessary and sufficient to restore ergosterol prototrophy to erg9 deficient yeast. The region boxed in red corresponds to a stretch of 26 amino acids that appear more conserved than other regions, and particularly well conserved amongst the fungal squalene synthase's (B).

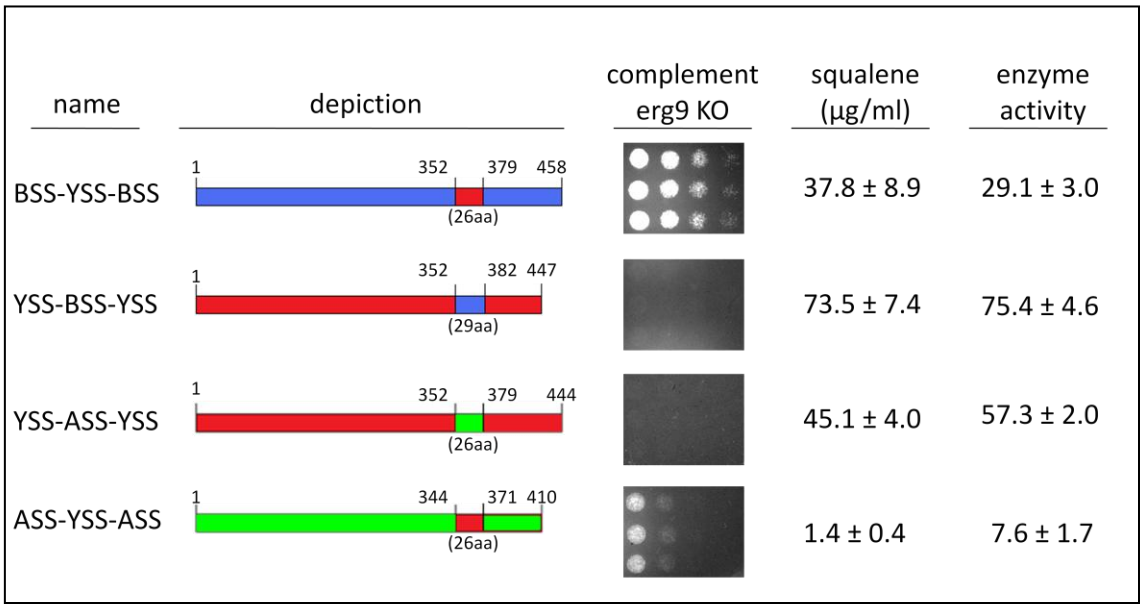


Figure 5.4 Functional assessment of the role a 26 amino acid peptide sequence within fungal squalene synthases plays in facilitating the complementation of the erg9 mutant.

Reciprocal constructs were created by swapping the indicated 26 amino acids of the yeast squalene synthase (YSS) for the corresponding amino acids of an algal squalene synthase (BSS) and a higher plant squalene synthase (ASS). Three independent Cali-7 transformants for each construct were randomly selected, grown in Yeast Synthetic Drop-out media (Sigma) containing 5 mg/l ergosterol for three days, after which serial dilutions were prepared corresponding to optical densities at 600 nm equivalent to 1, 0.2, 0.04, and 0.008, and 5 µl of each dilution spotted onto Yeast Synthetic Drop-out media plates lacking ergosterol. Plates were incubated at 28˚C for 72 h. Liquid cultures of each transformant in TN-7 line were grown in 10 mL of Yeast Synthetic Drop-out media containing 5 mg/L ergosterol at room temperature for seven days. Organic extracts were prepared and analyzed by GC-MS for their squalene content. To validate each gene construct, the squalene synthase enzyme activity encoded by each gene was assessed when the gene was expressed Cali-7 yeast. Briefly, 2000 x g supernatants were prepared and used in enzyme assays containing 3H-FPP and radiolabeled products separated by TLC and analyzed. Enzyme activity (pmoles/h/µg total protein) is recorded as a percent of YSS. n=3

Mapping the specific amino acids contributing to the complementation phenotype

To further assess the contribution of individual amino acids within this key stretch of 26 amino acids, fine mapping substitution series constructs were generated (Figure 5.5). First, mutants were created in which either the first half (13 residues) or the second half of the 26 amino acid domain was swapped from the *A. thaliana* SS into the *S. cerevisiae* backbone (Figure 5.5, mutants a and b). Swapping out the first 13 amino acids of the *S. cerevisiae* SS (Figure 5.5, mutant a) had no effect on the ability of the subsequent construct to complement the erg9 mutation, but swapping the second half resulted in a complete loss of the complementation phenotype (Figure 5.5, mutant b). Swapping out the first 13 amino acids of *A. thaliana* SS with those in *S. cerevisiae* also had no effect on the ability of this construct to complement the erg9 mutant (Figure 5.5, mutant f). However, exchanging the second half of these residues with those of the yeast squalene synthase enabled the construct to restore partial growth to the erg9 mutant (Figure 5.5, mutant g). Thus, it appeared that the residues in the second half of the 26-amino acid domain were largely responsible for complementation phenotype and this stretch of amino acids was evaluated further.

Because the KIEQ and FLKLNIQ stretches of amino acids seem particularly conserved amongst the fungal squalene synthases, various combinations of these peptide domains were exchanged between the ASS and YSS constructs (Figure 5.5, mutants c-e and h-j). When the “FLKLNIQ” stretch of YSS was replaced with the corresponding domain of ASS (Figure 5.5, mutant d), complementation growth was not affected, and likewise the reciprocal swap in ASS (Figure 5.5, mutant i) did not restore complementation. However, when the “KIEQ” stretch of YSS was replaced with the corresponding domain of ASS (Figure 5.5, mutant c), growth of the yeast was significantly impaired, suggesting that complementation had been affected. The reciprocal swap in ASS (Figure 5.5, mutant h) did not restore growth of the erg9 mutant. Two additional mutants in which the entire domains spanning from “FLKLNIQ” to “KIEQ” were exchanged (Figure 5.5, mutant e and j) were also evaluated for their ability to complement the erg9 mutation. Consistent with our expectations from the other mutants, substituting the yeast amino acids of this domain with those from the Arabidopsis squalene synthase resulted in a complete lose in the ability of this construct to complete the erg9 mutation. However, unexpectedly, substitution of this domain in the Arabidopsis squalene synthase gene with that of the yeast squalene synthase only restored a very modest level of growth to mutant yeast in the complementation tests. All constructs tested in Figure 5.5, when expressed in yeast, accumulated higher levels of squalene and had SS enzyme activity in excess of yeast expressing YSS.

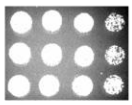
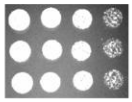
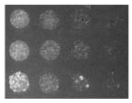
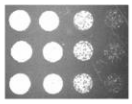
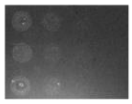
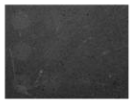
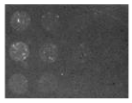
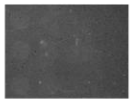
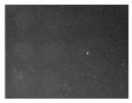
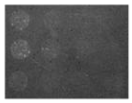
mutant	construct	complement erg9 KO	squalene (µg/ml)	enzyme activity
YSS	KSKLAVQDPNFLKLNIQISKIEQFME		21.0 ± 1.5	100.0 ± 7.9
M(a)	KTKVDKNDPNASKLNIQISKIEQFME		65.4 ± 3.1	160.5 ± 3.8
M(b)	KSKLAVQDPNFLKTLNRLEAVQKLCR		46.6 ± 9.7	152.7 ± 1.6
M(c)	KSKLAVQDPNFLKLNIQISAVQKFME		28.0 ± 1.3	156.1 ± 5.8
M(d)	KSKLAVQDPNASKTLNRISKIEQFME		50.7 ± 13.5	161.8 ± 5.3
M(e)	KSKLAVQDPNASKTLNRLEAVQKFME		69.1 ± 7.1	123.1 ± 6.0
ASS	KTKVDKNDPNASKTLNRLEAVQKLCR		69.0 ± 7.9	151.0 ± 3.9
M(f)	KSKLAVQDPNFLKTLNRLEAVQKLCR		82.8 ± 10.3	137.4 ± 8.6
M(g)	KTKVDKNDPNASKLNIQISKIEQFME		47.9± 6.0	151.4 ± 4.0
M(h)	KTKVDKNDPNASKTLNRLEKIEQLCR		64.9 ± 8.9	166.1 ± 1.3
M(i)	KTKVDKNDPNFLKLNIQLEAVQKLCR		85.2 ± 5.6	170.8 ± 1.2
M(j)	KTKVDKNDPNFLKLNIQISKIEQLCR		41.7 ± 1.6	159.7 ± 5.8

Figure 5.5 Evaluating the contribution of a carboxy-terminal sequence of 26 amino acids conserved amongst fungi to the complementation and restoration of ergosterol pleiotrophy to an erg9 knockout yeast line.

Full-length S. *cerevisiae* and Arabidopsis squalene synthase genes were constructed in which the indicated amino acids corresponding to residues 353 to 378 of YSS and

residues 345 to 370 of ASS were exchanged with one another in either the YSS gene (mutants a-e) or the ASS gene (mutants f-j) (in green, ASS amino acids substituted into the YSS gene; in red, YSS residues substituted into the ASS gene). Each construct was independently transformed into the Cali7-1 erg9 mutant line, 3 independent transformants were randomly selected and grown in Yeast Synthetic Drop-out media (Sigma) containing 5 mg/l ergosterol for 3 days. Aliquots of each culture were then diluted with water to optical densities (600 nm) corresponding to 1, 0.2, 0.04, and 0.008, and 5 µl of each dilution spotted on Yeast Synthetic Drop-out media plates lacking ergosterol. Plates were incubated at 28˚C for 72 h. Liquid cultures of each transformant in TN-7 line were grown in 10 mL of Yeast Synthetic Drop-out media containing 5 mg/L ergosterol at room temperature for seven days. Organic extracts were prepared and analyzed by GC-MS for their squalene content. To validate each gene construct, the squalene synthase enzyme activity encoded by each gene was assessed when the gene was expressed Cali-7 yeast. Briefly, 2000 x g supernatants were prepared and used in enzyme assays containing 3H-FPP and radiolabeled products separated by TLC and analyzed. Enzyme activity (pmoles/h/µg total protein) is recorded as a percent of YSS. n=3

Discussion

While the observation that a non-catalytic, carboxy-terminal domain of the fungal squalene synthases was important for allowing heterologous squalene synthase genes to complement a genetic deficiency of squalene synthase (erg9 mutants) in yeast was noted by several groups of investigators almost 20 years ago (36, 83, 84), the actual mechanism responsible for this phenomenon has remained unresolved. Kirbii, *et al*. (83) suspected that because this carboxy-terminal region of all squalene synthase enzymes served to anchor the enzymes to membrane systems, the membrane-spanning domain encompassed by this region of the protein might be important for correctly targeting the various SS proteins to the ER membrane in yeast, and thus complementing the erg9 deficiency. Unfortunately, how important the 24 carboxy-terminal amino acids comprising the membrane-spanning region were for the complementation phenotype was never tested. Instead, Kirbii, *et al.* (83) demonstrated that a heterologous squalene synthase unable to complement the erg9 mutant in yeast could be converted to a complementing form if its native 67 carboxy-terminal amino acids were replaced with 111 amino acids of the S. pombe SS carboxy-terminus. In the work presented here, we have demonstrated that the 24 carboxy-terminal amino acids of the native yeast squalene synthase are not necessary for this enzyme to complement the erg9 knockout mutation (Figure 5.1), suggesting that the membrane-spanning region may not be as important for complementation as suggested by Kirbii, *et al*. (83). In contrast, we have identified a stretch of 26 amino acids located outside of the membrane- spanning domain that appear highly conserved to most fungal squalene synthases but not squalene synthases belonging to animals, plants or other eukaryotic organisms, and appears sufficient, when appropriately attached to any squalene synthase enzyme, for the so engineered squalene synthase to complement the erg9 knockout in yeast. Further mutagenesis and substitution assessment of this 26 amino acid domain revealed that several of the small stretches of amino acids, especially the KIEQ domain, were essential for the complementation activity of this domain. Other residues and small domains distributed across this region could also contribute to the complementation activity, but these might be more influenced by the context of the surrounding residues. Altogether, our results suggest that, at least in yeast, it is not the membrane-spanning region of squalene synthases that is important for establishing ergosterol prototrophy, but a stretch of amino acids located near the carboxy-terminus that is essential. Our results also demonstrate that this domain does not influence the catalytic performance of the squalene synthase enzyme directly, yet must be mediating the complementation phenotype by some other means.

Although plant and animal squalene synthases were not able to complement the erg9 knockout in yeast, squalene accumulation was readily apparent and was often in vast excess to the levels observed for the wild type yeast or any of the complemented erg9 mutant lines. Hence, the erg9 mutation cannot be complemented by simple over-production of squalene. Perhaps it is equally important as to where the squalene might be produced such that it could feed more directly into the ergosterol biosynthetic

pathway. Support for this notion comes from the feeding studies performed by Kribii, *et al.*, in which ^{3}H-FPP was fed to microsomes from yeast expressing either ASS (non-complementing the erg9 mutant) or an ASS-YSS fusion (complementing). With the ASS microsomes, ^{3}H-labeled squalene was the only detectable metabolite. However, with the ASS-YSS microsomes, radiolabel was also found in squalene epoxide and lanosterol, two reaction intermediates downstream of squalene in the ergosterol pathway. Equally interestingly, nearly all the squalene synthase enzyme activity reported for these independently transformed lines was associated with the microsomes, indicative that the squalene synthases seem to be correctly targeted to the ER membrane.

How then can SS enzymes targeted to the correct membrane system (i.e. the ER) still be active and generating squalene, yet the squalene produced doesn't seem sufficient to complement the erg9 mutation? Perhaps the simplest explanation is that the squalene biosynthesized under these conditions is not in the right subcellular environment to be properly channeled into sterol metabolism. Such a notion would be consistent with the carboxy-terminal located 26 amino acid domain essential for complementation playing some role in protein:protein interactions and, thus possibly serving to insert the squalene synthase enzyme activity into the right environment and context for the channeling of squalene onto the next enzymes in the sterol biosynthetic pathway. The identification of ERG28 as a protein scaffold that directly influences the efficiency of sterol biosynthesis by coordinating the interaction between enzymes in the sterol biosynthetic pathway (86) provides a compelling example for our suggestion here.

The stretch of 26 amino acids is well conserved among fungal squalene synthases, and particularly the KIEQ sequence is strongly conserved. For example, the fungal species *Kluyveromyces lactis*, *Candida glabrata*, *Pichia jadinii*, and *Botryotinia fuckeliana* all share greater than 50% identity to the *S. cerevisiae* 26 amino acid domain and the KIEQ sequence is completely conserved. Other fungal species such as *Zygosaccharomyces rouxii*, *Clavispora lusitaniae*, and *Aspergillus clavatus* share greater than 50% identity within the 26 amino acid domain but contain a similar RIEQ instead of KIEQ sequence. If this stretch of amino acids is vital for proper sterol biosynthesis in other fungi, as it is in *S. cerevisiae*, then peptides or small molecules that block or mimic these residues could disrupt sterol metabolism. Such molecules could potentially be used as fungal specific antibiotics.

Another important question arising from the current observations is if analogous domains to the 26 amino acid stretch identified here in the yeast squalene synthase also exists in other squalene synthases, like that for man? If so, then blocking the protein:protein interactions that such a squalene synthase might participate in could disrupt the flow of squalene to the sterol biosynthetic pathway and thus represent in a novel mechanism to regulate cholesterol in man. There are now good examples of drugs targeting protein-protein interactions such as anti-tumor and anti-cancer drugs. Brooks *et al.* (87) showed that MMP2 binding to the αVβ3 integrin played a major role in the process of

angiogenesis, but that the interaction of these proteins could be blocked with a small peptide fragment of MMP2. Using a combinatorial chemistry approach, they also identified a small molecule that could block these protein-protein interactions in a dose-dependent manner, and demonstrated that these same small molecule blocked tumor growth in *in vivo* model studies.

In conclusion, we have shown that a stretch of 26 amino acids near the carboxy-terminus of fungal squalene synthase is both necessary and sufficient for any squalene synthase to restore ergosterol prototrophy to erg9 deficient yeast, and that the function of this stretch of amino acids is not likely involved in correctly targeting squalene synthase to the ER membrane, but instead may play a role in metabolon assembly of sterol biosynthetic proteins.

Methods

Cloning the various squalene synthases

Squalene synthases from S. *cerevisiae* (YSS), *B. braunii* (BSS), R. norvegicus (RSS), N. benthamiana (TSS), and A. thaliana (ASS) were cloned from original cell or tissue sources. First, total RNA was isolated from the respective species using the RNeasy Plant mini kit (Qiagen) or Trizol (Invitrogen, for R. norvegicus and S. *cerevisiae*) according to the manufacturers recommendations, and first strand cDNA was synthesized using the SMARTer PCR cDNA synthesis kit (Clontech). The first strand cDNA was used as a template to amplify the various squalene synthase genes using the primer sets listed in Table 5.1 (restriction site in bold). YSS was cloned into the Yep352 vector with the AscI and XbaI sites, and a 3'-truncated form cloned into Yep352 and pET28a using the AscI and XbaI or BamHI and XhoI sites, respectively. BSS and its 3'-truncated form were cloned into Yep352 with the EcoRI and HindIII sites. A 5'- and 3'-truncated RSS was cloned into Yep352 and pET28a using the EcoRI and HindIII sites or BamHI and XhoI sites, respectively. TSS was cloned into Yep352 with the EcoRI and NotI sites and its 3'-truncated form was cloned into Yep352 and pET28a with the EcoRI and NotI sites. ASS was cloned into Yep352 with the EcoRI and NotI sites. All constructs were verified by automated DNA sequencing.

Table 5.1 Squalene Synthase Primer Sequences.

gene	primer	Sequence
YSS	AscI For	A**GGCGCGCC**AAAACAATGGGAAAGCTATTACAATGGC
	BamHI For	CGC**GGATCC**AAAACAATGGGAAAGCTATTACAATGGC
	XbaI Rev	GC**TCTAGA**TCACGCTCTGTGTAAAGTGTATAT
	NotI Rev	ATAAGAAT**GCGGCCGC**TCACGCTCTGTGTAAAGTG
	XbaI Rev trunc	GC**TCTAGA**TCACTTGTACTCTTCTTC
	XhoI Rev trunc	GGG**CTCGAG**TCACTTGTACTCTTCTTC
	NotI Rev trunc	ATAAGAAT**GCGGCCGC**TCACTTGTACTCTTCTTCTTG
BSS	EcoRI For	CCG**GAATTC**AAAACAATGGGGATGCTTCGCTGGGGAGTGG
	HindIII Rev	ATCCC**AAGCTT**TTAGGCGCTGAGTGTGGGTCTAGG
	NotI Rev	ATAAGAAT**GCGGCCGC**TTAGGCGCTGAGTGTGGGTCTAGG
	HindIII Rev trunc	CCC**AAGCTT**TCACCACCAATGTTGGGTTACAGA
RSS	EcoRI For trunc	CCG**GAATTC**AAAACAATGGACCGGAACTCGCTCAGC
	BamHI For trunc	CGC**GGATCC**ATGGACCGGAACTCGCTCAGC
	HindIII Rev trunc	CCC**AAGCTT**TCAGCTCTGCGTCCTGATGTTGGAG
	XhoI Rev trunc	CCG**CTCGAG**TCAGCTCTGCGTCCTGATGTTGGAG
TSS	EcoRI For	G**GAATTC**ATGGGGAGTTTGAGGGCTATTC
	XbaI Rev	GC**TCTAGA**CTAAGATCGGTTTCCGGATAGC
	NotI Rev	ATAAGAAT**GCGGCCGC**CTAAGATCGGTTTCCGGATAGC
	NotI Rev trunc	ATAAGAAT**GCGGCCGC**TCAGTAATTAGGCTCGCTCCTG
ASS	EcoRI For	CCG**GAATTC**AAAACAATGGGGAGCTTGGGGACGATGCTG
	XbaI Rev	GC**TCTAGA**TCAGTTTGCTCTGAGATATGC
	NotI Rev	ATAAGAAT**GCGGCCGC**TCAGTTTGCTCTGAGATATGCAAAG

Creating the BSS-YSS fusion

A reverse primer was designed to pair with the BSS EcoRI For (see Table 5.1), to amplify the first 352 codons of BSS except that a single nucleotide mutation was introduced into the 352nd codon to introduce a HindIII restriction site without changing the encoded amino acid (ATCCC**A**__**AGCTT**__CTCTGCTAATTTGAGG, HindIII site in bold, mutation underlined). This was cloned into the pET28a vector with the corresponding restriction sites, giving BSS_{352}-pET28a. Another primer was designed to pair with either primer, YSS NotI Rev or YSS NotI Rev trunc, to amplify YSS starting from codon 353 (ATCC**AAGCTT**AAATCTAAATTGGCTGTGC, HindIII site in bold), and these fragments cloned into BSS_{352}-pET28a cut with HindIII and NotI to give the BSS-YSS and BSS-

YSStr constructs. These were cut from the pET28a vector using EcoRI and NotI and ligated into the corresponding sites of Yep352. The construct was verified by automated DNA sequencing.

Creating the BSS-YSS-BSS expression cassette

A primer was designed to pair with primer, BSS EcoRI For, to amplify a fragment of the BSS-YSS construct with NgoMIV and NotI restriction sites (ATAAAGAAT**GCGGCCGC**GAAT**GCCGGC**TTCCATAAACTGTTCGATCTTGG, NgoMIV and NotI sites in bold). This was cloned into the EcoRI and NotI sites of YEp352, which was later cut with NgoMIV and NotI. Meanwhile a primer was designed to pair with primer, BSS NotI Rev, to amplify a 3' region of BSS except that two nucleotide mutations were introduced to add an NgoMIV restriction site without changing the encoded amino acids (GCAAAGAAT**GC<u>C</u>GG<u>C</u>**CTGGCACGCACAAAAGATGACACC, NgoMIV site in bold, mutations underlined). This fragment was cloned into the NgoMIV and NotI sites of the cut Yep352 vector to give BSS-YSS-BSS. The construct was verified by automated DNA sequencing.

Creating other fusion constructs

All other fusion constructs were created by employing an assembly PCR strategy as described by Niehaus *et al.* (58), using the primers listed in Tables 1 and 2. For example, YSS-BSS was created by using YSS as a template with the primer set, YSS-BSS 1R and YSS AscI For, to amplify a fragment of YSS with a 3' overhang, and using BSS as the template with the primer set, YSS-BSS 1R and BSS HindIII Rev, to amplify a fragment of BSS with a 5' overhang. These two fragments were both used as templates in a PCR reaction with the primer set, YSS AscI For and BSS HindIII Rev, to give the YSS-BSS construct, which was cloned into the YEp352 vector with the corresponding restriction sites. YSS-BSS-YSS was created by using YSS-BSS and YSS as templates in the initial PCR reaction, and cloning the finished construct into YEp352 with the AscI and XbaI sites. All other constructs were created in a similar manner. TSS-YSS and YSS-TSS were cloned into YEp352 with the EcoRI and NotI, or AscI and XbaI restriction sites, respectively. YSS-ASS and ASS-YSS were cloned into YEp352 with AscI and XbaI, or EcoRI and XbaI restriction sites, respectiviely. All YSS M(a)-M(e) constructs were cloned into YEp352 with the AscI and XbaI restriction sites, and all ASS M(f)-M(j) constructs were cloned into YEp352 with EcoRI and XbaI restriction sites. All constructs were verified by automated DNA sequencing.

Table 5.2 Chimeric Squalene Synthase Primer Sequences.

construct	direction	sequence
YSS-BSS	1F	CTTACGTGATATCGAAGTCAGATGCAACACCGAGACCAGCGAGGATCCC
	1R	GCATCTGACTTCGATATCACGTAAGTAATAGTCAAAAATCTCGACACAGCC
YSS-ASS	1F	CTTACGTGATATCAAGACAAAGGTTGACAAGAACGATCCAAATGCCAG
	1R	CAACCTTTGTCTTGATATCACGTAAGTAATAGTCAAAAATCTCGACAC
ASS-YSS	1F	CCTGCATGCTGAAATCTAAATTGGCTGTGCAAGATCCAAATTTCTTA
	1R	GCCAATTTAGATTTCAGCATGCAGGAAAAATCATAGAAAGCACCATAG
YSS-TSS	1F	CTTACGTGATATCAAATCCAAGGTTAATAATAATGATCCAAATGCAAC
	1R	TTAACCTTGGATTTGATATCACGTAAGTAATAGTCAAAAATCTCGACAC
TSS-YSS	1F	GACTTTTCTTGTATGCTGAAATCTAAATTGGCTGTGCAAGATCCAAATTTCTT
	1R	GCCAATTTAGATTTCAGCATACAAGAAAAGTCAAAAAAAGCACCATATACATC
YSS-BSS-YSS	1F	CAAAGCTGCCTGCAAGGAAATGTACCAGGATAAATTACCTCCTAACGTGAAGCC
	1R	CCTGGTACATTTCCTTGCAGGCAGCTTTGATCTTATGCAGGTGTTCCAGAG
YSS-ASS-YSS	1F	AAGACAAAGGTTGACAAGAACGATCCAAATGCCAGTAAGACACTAAACCGACTTGAAGCC
	1R	TCTGCAGAGTTTCTGAACGGCTTCAAGTCGGTTTAGTGTCTTACTGGCATTTGGATCGTT
ASS-YSS-ASS	1F	AAATCTAAATTGGCTGTGCAAGATCCAAATTTCTTAAAATTGAACATTCAAATCTCCAAG
	1R	TTCCATAAACTGTTCGATCTTGGAGATTTGAATGTTCAATTTTAAGAAATTTGGATCTTG
YSS M(c)	1F	CTCCGCCGTTCAAAAGTTTATGGAAGAAATGTACCAGGATAAATTACC
	1R	CCATAAACTTTTGAACGGCGGAGATTTGAATGTTCAATTTTAAG
YSS M(d)	1F	AATGCCAGTAAGACACTAAACCGTATCTCCAAGATCGAACAGTTTATGG
	1R	GATACGGTTTAGTGTCTTACTGGCATTTGGATCTTGCACAGCCAATTTAG
YSS M(e)	1F	GCCAGTAAGACACTAAACCGTCTTGAAGCCGTTCAGAAGTTTATGGAAGAAATGTACCAG
	1R	CTTCTGAACGGCTTCAAGACGGTTTAGTGTCTTACTGGCATTTGGATCTTGCACAGCC
YSS M(a)	1F	AAGACAAAGGTTGACAAGAACGATCCAAATGCCAGTAAGTTGAACATTCAAATCTCCAAG
	1R	CTTACTGGCATTTGGATCGTTCTTGTCAACCTTTGTCTTGATATCACGTAAGTAATAGTC
YSS M(b)	1F	ACACTAAACCGACTTGAAGCCGTTCAGAAACTCTGCAGAGAAATGTACCAGGATAAATTA
	1R	TCTGCAGAGTTTCTGAACGGCTTCAAGTCGGTTTAGTGTTTTTAAGAAATTTGGATCTTG
ASS M(h)	1F	
	1R	
ASS M(i)	1F	AATTTCTTAAAATTGAACATTCAACTTGAAGCCGTTCAGAAACTCTGCAG
	1R	AAGTTGAATGTTCAATTTTAAGAAATTTGGATCGTTCTTGTCAACCTTTG
ASS M(j)	1F	TTCTTAAAATTGAACATTCAAATCTCCAAGATCGAACAGCTCTGCAGAGACGCTGGAG
	1R	CTGTTCGATCTTGGAGATTTGAATGTTCAATTTTAAGAAATTTGGATCGTTCTTGTCAAC
ASS M(f)	1F	AAATCTAAATTGGCTGTGCAAGATCCAAATTTCTTAAAAACACTAAACCGACTTGAAGCC
	1R	TTTTAAGAAATTTGGATCTTGCACAGCCAATTTAGATTTCAGCATGCAGGAAAAATCATA
ASS M(g)	1F	TTGAACATTCAAATCTCCAAGATCGAACAGTTTATGGAAGACGCTGGAGTTCTTCAAAAC
	1R	TTCCATAAACTGTTCGATCTTGGAGATTTGAATGTTCAACTTACTGGCATTTGGATCGTT

The erg9 complementation assay

The Cali7-1 yeast line, which has an erg9 deletion so that it cannot synthesize sterols de novo and requires exogenous ergosterol for growth, was used for these purposes (41). The various squalene synthase constructs were transformed into Cali7-1 yeast using the lithium acetate method and plated on Yeast Synthetic Drop-out medium (selection media) lacking uracil and containing 5 mg/l ergosterol. Three independent CALI7-1 transformants of each construct were randomly selected and grown in 2 mL Yeast Synthetic Drop-out media (Sigma) containing 5 mg/L ergosterol at 28^{o}C for three days, after which the culture was serially diluted with water to optical densities (600 nm) equal to 1, 0.2, 0.04, and 0.008, and 5 µL of each dilution spotted onto Yeast Synthetic Drop-out media plates without any exogenous ergosterol. Plates were incubated at 28^{o}C for 72 hours.

Liquid cultures of each transformant in TN-7 line were grown in 10 mL of Yeast Synthetic Drop-out media containing 5 mg/L ergosterol at room temperature for seven days. Organic extracts were prepared and analyzed by GC-MS for their squalene content. In brief, 1 mL aliquots of the culture were combined with 1 mL of acetone, vigorously mixed, and incubated at room temperature for 10 min. One mL of hexane was added and mixed vigorously for 60 sec. The mixture was then centrifuged briefly at 500*xg* to separate the phases, and an aliquot of the organic phase removed and 1-2 µL aliquots analyzed by GC-MS with a Varian CP-3800 GC coupled to a Varian Saturn 2200 MS/MS (Varian Medical Systems) using a Supelco SLB-5ms fused silica capillary column (30 m × 0.25 mm × 0.25 µM film thickness, Supelco). Initial oven temperature was set at 220 °C for 1 min., ramped to 280 °C at 20 °C/min., then ramped to 300 °C at 3 °C/min.

The various SS constructs were expressed in Cali-7 yeast and grown for 3 days before 1.5 ml of culture was collected by centrifugation and stored at -80C until further analysis. Yeast pellets were resuspended in 0.5 ml buffer (50 mM NaH_2PO_4, pH 7.8, 300 mM NaCl, 1 mM $MgCl_2$, 1 mM PMSF, 1% glycerol (v/v)) then sonicated 3 x for 5 sec with a microprobe sonicator at 60% maximum power. The samples were cooled on ice for 2 min between sonication treatments. The sonicate was centrifuged at 2,000 *g* for 10 min at 4°C and the supernatant used in enzyme assays. Assays contained 50 mM Mops, pH 7.3, 20 mM $MgCl_2$, 2.5 mM 2-mercaptoethanol, 10 µM [1-^{3}H]-FPP (~1x10^{5} dpm total), and 2 mM NADPH in total reaction volume of 50 µL. Reactions were incubated at 37ºC for 1 h and then extracted with 100 µl n-hexane and an aliqout spotted on TLC with a squalene standard. TLC was developed with n-hexane and the squalene zone was scraped and analyzed by scintillation spectroscopy. The amount of total protein in the yeast supernatants was determined by Bradford Dye assays. Enzyme activity (pmole/h/µg total protein) is expressed as a percent of S. *cerevisiae* squalene synthase (YSS) enzyme activity. n=3.

Chapter 6 Concluding Remarks

Chapter 2 describes the pathway for botryococcene biosynthesis in *B. braunii* which is interesting because new mechanisms for botryococcene and squalene catalysis are suggested. Squalene biosynthesis in other organisms has been well documented, consisting of in initial condensation of 2 molecules of FPP to form PSPP, then an NAD(P)H dependent reductive rearrangement of PSPP to form squalene. Both of these steps occur within a single enzyme unit, and this mechanism is highly conserved amongst all eukaryotes. In contrast, *B. braunii* seems to have evolved completely different mechanisms for botryococcene and squalene formation. First, the enzyme SSL-1 catalyzes the condensation of 2 molecules of FPP to form PSPP as its sole reaction product. Two additional enzymes, SSL-2 and SSL-3, cannot use FPP as a substrate but instead are able to reductively rearrange PSPP generated by SSL-1 to form squalene or botryococcene, respectively. With this new suite of squalene synthase-like enzymes available which are specialized in one particular half-reaction, one can predict specific residues that may be important for particular steps in the reaction cascade and test these predictions with mutagenesis experiments. Thus these enzymes will greatly facilitate the understanding of the squalene synthase enzymology; which has remained vague despite decades of study.

Through biochemical studies of the SSL enzymes described in chapter 2 it was discovered that SSL-2 catalyzed the formation of bisfarnesyl ether from either FPP or FPP and farnesol as substrates. Bisfarnesyl ether is an unusual molecule that has never been reported to accumulate in any organism, including *B. braunii*, and likely results from a perturbation of the normal catalytic cascade of the SSL-2 enzyme. Chapter 4 describes some of the kinetic parameters and cofactor requirements for bisfarnesyl ether catalysis by SSL-2. This information could aid in studies aimed at understanding the enzymology of squalene synthase-like enzymes. Additionally, the NADPH3 that was synthesized for this work may also be used in crystallography studies of the SSL enzymes to determine how nicotinamide cofactors interact with such proteins.

Chapter 2 also describes the production of botryococcene in a yeast heterologous host. Expression of SSL-1 and SSL-3 on separate plasmids caused a modest accumulation of botryococcene in yeast, however, fusing these two peptides with a short amino acid linker and adding a peptide sequence at the carboxy terminus that targeted to yeast ER membrane caused accumulation to increase 5-fold, demonstrating that high levels of botryococcene can be produced in a heterologous host. However, for this high-value oil to be produced in a clean, renewable, and sustainable means, this pathway will have to be engineered in a photosynthetic organism capable of converting atmospheric CO2 into botryococcene, such as fast growing plants or algae. Efforts to deploy high level botryococcene biosynthesis in a photosynthetic platform will benefit from the examples shown in yeast.

Triterpene methyltransferases (TMTs) that can introduce methyl groups to either squalene or botryococcene are described in Chapter 3. These TMTs are capable of converting the C30 triterpenes into more highly branched C32 forms, thus increasing there efficacy as petrochemical feedstocks. It was demonstrated that high levels of methylated triterpenes accumulated in yeast expressing TMT's along with triterpene synthases. Attempts to produce triterpenes in photosynthetic organisms should also include the TMTs described in chapter 3 to produce an optimal feedstock for fuel production. These TMT genes also offer an opportunity to further dissect substrate specificity and methylation target site selectivity.

Chapter 5 describes a specific stretch of amino acids that are necessary and sufficient to allow any squalene synthase to functionally complement an erg9 (squalene synthase) deficient yeast line by allowing squalene to feed into the sterol biosynthetic pathway. This represents a nice example of metabolon assembly and increases our understanding of sterol metabolism. Additionally, these studies could lead to future drug design. For example, small molecules could be designed that mimic the specific stretch of amino acids and block yeast squalene synthase from forming a metabolon, thus disrupting sterol biosynthesis in yeast. Such drugs could be used to specifically target yeast or other fungal pathogens. Also interesting, does this phenomenon of a specific stretch of amino acids of squalene synthase being necessary to feed squalene into the sterol biosynthetic pathway occur in other organisms besides yeast? The work in chapter 5 warrants studies to see if similar mechanisms occur is plants and animals. If this mechanism is conserved across Kingdoms, then it could lead to a new class of cholesterol lowering drugs in human.

Bibliography

1. Glikson M, Lindsay K, Saxby J (1989) *Botryococcus* - a planktonic green-alga, the source of petroleum through the ages - transmission electron microscopical studies of oil shales and petroleum source rocks. *Org Geochem* 14:595-608.
2. Derenne S*, et al.* (1997) Chemical structure of the organic matter in a pliocene maar-type shale: Implicated *Botryococcus* race strains and formation pathways. *Geochim Cosmochim Acta* 61:1879-1889.
3. Mastalerz M, Hower JC (1996) Elemental composition and molecular structure of *Botryococcus* alginite in westphalian cannel coals from Kentucky. *Org Geochem* 24:301-308.
4. McKirdy DM, *et al*. (1986). Botryococcane in a New Class of Australian Non-marine crude oils. *Nature* 320: 57-59
5. Gelpi E, Oro J, Schneide HJ, Bennett EO (1968) Olefins of high molecular weight in 2 microscopic algae. *Science* 161:700-701.
6. Okada S, Murakami M, Yamaguchi K (1995) Hydrocarbon composition of newly isolated strains of the green microalga *Botryococcus braunii*. *J Appl Phycol* 7:555-559.
7. Metzger P, Allard B, Casadevall E, Berkaloff C, Coute A (1990) Structure and chemistry of a new chemical race of *Botryococcus braunii* (Chlorophyceae) that produces lycopadiene, a tetraterpenoid hydrocarbon. *J Phycol* 26:258-266.
8. Weiss TL*, et al.* (2010) Raman spectroscopy analysis of botryococcene hydrocarbons from the green microalga *Botryococcus braunii*. *J Biol Chem* 285:32458-32466.
9. Zhaohui Z, Metzger P, Sachs JP (2007) Biomarker evidence for the co-occurrence of three races (A, B and L) of *Botryococcus braunii* in El Junco Lake, Gala´pagos *Org Geochem* 38:1459-1478.
10. Brown AC, Knights BA, Conway E (1969) Hydrocarbon content and its relationship to physiological state in green alga *Botryococcus braunii*. *Phytochemistry* 8:543-547.
11. Lee DY, Fiehn O (2008) High quality metabolomic data for *Chlamydomonas reinhardtii. Plant Methods* 4:7
12. Hillen LW, Pollard G, Wake LV, White N (1982) Hydrocracking of the oils of *Botryococcus braunii* to transport fuels. *Biotech Bioeng* 24:193-205.
13. Poulter CD (1990) Biosynthesis of non-head-to-tail terpenes - formation of 1'-1 and 1'-3 linkages. *Acc Chem Res* 23:70-77.
14. Rilling HC (1966) A new intermediate in biosynthesis of squalene. *J Biol Chem* 241:3233-3236.
15. Blagg BSJ, Jarstfer MB, Rogers DH, Poulter CD (2002) Recombinant squalene synthase. A mechanism for the rearrangement of presqualene diphosphate to squalene. *J Amer Chem Soc* 124:8846-8853.
16. Okada*, et al.* (2004) Characterization of botryococcene synthase enzyme activity, a squalene synthase-like activity from the green microalga *Botryococcus braunii*, Race B. *Arch Biochem Biophys* 422:110-118.

17. Jarstfer MB, Zhang DL, Poulter CD (2002) Recombinant squalene synthase. Synthesis of non-head-to-tail isoprenoids in the absence of NADPH. *J Amer chem Soc* 124:8834-8845.
18. Pan JJ, Bugni TS, Poulter CD (2009) Recombinant squalene synthase. Synthesis of cyclopentyl non-head-to-tail triterpenes. *J Org Chem* 74:7562-7565.
19. Banerjee A, Sharma R, Chisti Y, Banerjee UC (2002) *Botryococcus braunii*: A renewable source of hydrocarbons and other chemicals. *Crit Rev Biotech* 22:245-279.
20. Okada S, Tonegawa I, Matsuda H, Murakami M, Yamaguchi K (1997) Braunixanthins 1 and 2, new carotenoids from the green microalga *Botryococcus braunii*. *Tetrahedron* 53:11307-11316.
21. Metzger P, Rager MN, Largeau C (2007) Polyacetals based on polymethylsqualene diols, precursors of algaenan in *Botryococcus braunii* race b. *Org Geochem* 38:566-581.
22. Takahashi S*, et al.* (2007) Metabolic engineering of sesquiterpene metabolism in yeast. *Biotech Bioeng* 97:170-181.
23. Morrison RT, Boyd RN (1973) Organic Chemistry 3rd Edition, Allyn and Bacon, Boston MA, 109-110 p.
24. Metzger P, Largeau C (2005) *Botryococcus braunii*: A rich source for hydrocarbons and related ether lipids. *Appl Microbiol Biotech* 66:486-496.
25. Huang Z, Poulter CD (1989) Tetramethylsqualene, a triterpene from *Botryococcus braunii* var showa. *Phytochemistry* 28:1467-1470.
26. Metzger P, Berkaloff C, Casadevall E, Coute A (1985) Alkadiene-producing and botryococcene-producing races of wild strains of *Botryococcus braunii*. *Phytochemistry* 24:2305-2312.
27. Metzger P (1999) Two terpenoid diepoxides from the green microalga *Botryococcus braunii*: Their biomimetic conversion to tetrahydrofurans and tetrahydropyrans. *Tetrahedron* 55:167-176.
28. Metzger P, Rager MN, Largeau C (2002) Botryolins A and B, two tetramethylsqualene triethers from the green microalga *Botryococcus braunii*. *Phytochemistry* 59:839-843.
29. Moldowan JM, Seifert WK (1980) 1st discovery of botryococcane in petroleum. *J Chem Soc-Chem Comm* 912-914.
30. Zhang DL, Poulter CD (1995) Biosynthesis of non-head-to-tail isoprenoids - synthesis of 1'-1-structures and 1'-3-structures by recombinant yeast squalene synthase. *J Amer Chem Soc* 117:1641-1642.
31. Bergstrom JD*, et al.* (1993) Zaragozic acids - a family of fungal metabolites that are picomolar competitive inhibitors of squalene synthase. *Proc Natl Acad Sci USA* 90:80-84.
32. Sasiak K, Rilling HC (1988) Purification to homogeneity and some properties of squalene synthetase. *Arch Biochem Biophys* 260:622-627.
33. Jarstfer MB, Blagg BSJ, Rogers DH, Poulter CD (1996) Biosynthesis of squalene. Evidence for a tertiary cyclopropylcarbinyl cationic intermediate in the rearrangement of presqualene diphosphate to squalene. *J Amer Chem Soc* 118:13089-13090.

34. Gu PD, Ishii Y, Spencer TA, Shechter I (1998) Function-structure studies and identification of three enzyme domains involved in the catalytic activity in rat hepatic squalene synthase. *J Biol Chem* 273:12515-12525.
35. Pandit J*, et al.* (2000) Crystal structure of human squalene synthase - a key enzyme in cholesterol biosynthesis. *J Biol Chem* 275:30610-30617.
36. Robinson GW, Tsay YH, Kienzle BK, Smithmonroy CA, Bishop RW (1993) Conservation between human and fungal squalene synthetases - similarities in structure, function, and regulation. *Mol Cell Biol* 13:2706-2717.
37. Okada S, Devarenne TP, Chappell J (2000) Molecular characterization of squalene synthase from the green microalga *Botryococcus braunii*, race B. *Arch Biochem Biophys* 373:307-317.
38. Houben J, Weyl T, Müller E (1965) *Methoden der organischen chemie* (G. Thieme, Stutgart, Germany) 4th Ed p 832.
39. Agnew WS, Popjak G (1978) Squalene synthetase - stoichiometry and kinetics of presqualene pyrophosphate and squalene synthesis by yeast microsomes. *J Biol Chem* 253:4566-4573.
40. Radisky ES, Poulter CD (2000) Squalene synthase: Steady-state, pre-steady-state, and isotope-trapping studies. *Biochemistry* 39:1748-1760.
41. Song LS (2003) Detection of farnesyl diphosphate accumulation in yeast erg9 mutants. *Anal Biochem* 317:180-185.
42. Nara T, Hshimoto T, Aoki T (2000) Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene* 257:209-222.
43. Peters RJ (2010) Two rings in them all: The labdane-related diterpenoids. *Nat Prod Rep* 27:1521-1530.
44. Field B,Osbourn AE (2008) Metabolic diversification - independent assembly of operon-like gene clusters in different plants. *Science* 320:543-547.
45. Toyomasu T*, et al.* (2000) Cloning of a full-length cDNA encoding ent-kaurene synthase from *Gibberella fujikuroi*: Functional analysis of a bifunctional diterpene cyclase. *Biosci Biotechnol Biochem* 64:660-664.
46. Prisic S, Xu MM, Wilderman PR, Peters RJ (2004) Rice contains two disparate ent-copalyl diphosphate synthases with distinct metabolic functions. *Plant Physiol* 136:4228-4236.
47. Xu MM, Hillwig ML, Prisic S, Coates RM, Peters RJ (2004) Functional identification of rice syn-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. *Plant J* 39:309-318.
48. Xu MM, Wilderman PR, Peters RJ (2007) Following evolution's lead to a single residue switch for diterpene synthase product outcome. *Proc Natl Acad Sci USA* 104:7397-7401.
49. Xu MM*, et al.* (2007) Functional characterization of the rice kaurene synthase-like gene family. *Phytochemistry* 68:312-326.
50. Cao R*, et al.* (2010) Diterpene cyclases and the nature of the isoprene fold. *Proteins* 78:2417-2432.

51. Haralampidis K*, et al.* (2001) A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proc Natl Acad Sci USA* 98:13431-13436.
52. Lin F-Y*, et al.* (2010) Mechanism of action and inhibition of dehydrosqualene sythase. *Proc Natl Acad Sci USA* 107:21337-21342.
53. Liu CI*, et al.* (2008) A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* 319:1391-1394.
54. Traverse A (1955) Occurrence of the oil-forming alga *Botryococcus* in lignites and other tertiary sediments. *Micropaleontology* 1:343-350.
55. Wang, L. L., Kao, R., Ivey, F. D.Hoffman, C. S. (2004) Strategies for gene disruptions and plasmid constructions in fission yeast *Methods* **33:** 199-205.
56. Sun, A. H., Mao, Y. F., Hu, Y., Sun, Q.Yan, J. (2009) Sensitive and specific ELISA coated by TpN15-TpN17-TpN47 fusion protein for detection of antibodies to Treponema pallidum *Clinical Chemistry and Laboratory Medicine* **47:** 321-326.
57. Metzger P, Casadevall E, Coute A (1988) Botryococcene distribution in strains of the green alga *Botryococcus braunii*. *Phytochemistry* 27(5): 1383-8
58. Niehaus TD*, et al.* (2011) Identification of Unique Mechanisms for Triterpene Biosynthesis in *B. braunii*. *Proc Natl Acad Sci USA* 108(30), 12260-12265.
59. Casadevall E, *et al.* (1983) Hydrocarbon production by the unicellular alga *Botryococcus braunii*. *Bio-Sciences* 8: 129-33
60. Achitouv E, Metzger P, Rager MN, Largeau C (2004) C31-C34 methylated squalenes from a Boli*vi*an strain of *Botryococcus braunii. Phytochemistry* 65(23), 3159-3165
61. Ginger ML, Portman N, McKean PG. (2008) Swimming with protists: perception, motility and flagellum assembly. *Nature Reviews Microbiology* 6(11), 838-850
62. Ahmad I, Rao DN (1996) Chemistry and biology of DNA methyltransferases. *Crit Rev Biochem Mol Biol* 31(5&6), 361-380.
63. Schubert, H. L., Blumenthal, R. M., and Cheng, X. D. (2003) *Trends in Biochemical Sciences* **28**, 329-335.
64. Kagan, R., and Clarke, S. (1995) *Archives Of Biochemistry And Biophysics* **316**, 657-657
65. Noel, J. P., Dixon, R. A., Pichersky, E., Zubieta, C., and Ferrer, J.-L. (2003) *Recent Advances in Phytochemistry* **37**, 37-58
66. Liscombe, D. K., Usera, A. R., and O'Connor, S. E. (2010) *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18793-18798
67. Carland F, Fujioka S, Nelson T (2010) The sterol methyltransferases SMT1, SMT2, and SMT3 influence Arabidopsis development through nonbrassinosteroid products. *Plant Physiology* 153(2), 741-756.
68. Merchant *et al*. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318(5848), 245-251.
69. Diener AC *et al.* (2000) Sterol methyltransferase 1 controls the level of cholesterol in plants. *The Plant Cell* 12:853-870
70. Bohlmann J, Keeling CI (2008) Terpenoid Biomaterials. *Plant Journal* 54(4): 659-669

71. Wu S *et al*. (2007) Cloning, characterization, and transformation of the phosphoethanolamine N-methyltransferase gene (ZmPEAMT1) in maize (Zea mays L.). *Molecular Biotechnology* 36(2), 102-112
72. Malhotra HC, Nes WR (1971) Mechanism of introduction of alkyl groups at C-24 of sterols. IV. Inhibition by triparanol. *J Biol Chem* 246(16), 4934-7
73. Bouvier-Nave P *et al*. (1997) Identification of cDNAs encoding sterol methyl-transferases involved in the second methylation step of plant sterol biosynthesis. *Eur J Biochem* 246(2), 518-529
74. Huang Z, Poulter DC (1989) Isoshowacene, a C31 hydrocarbon from *Botryococcus braunii* var. showa. *Phytochemistry*, 28, 3043-3046
75. Metzger P, Casadevall E, Pouet MJ, Pouet Y (1985) Structures of some botryococcenes: branched hydrocarbons from the B-race of the green alga *Botryococcus braunii*. *Phytochemistry* 24(12), 2995-3002
76. Huang Z, Poulter DC (1989) Tetramethylsqualene, a triterpene from *Botryococcus braunii* var. showa. *Phytochemistry* 28(5), 1467-70
77. Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods in Enzymology* 272: 51-64
78. Fegueur M, Richard L, Charles AD, Karst F (1991) Isolation and primary structure of the ERG9 gene of Saccharomyces cerevisiae encoding squalene synthetase. *Current Genetics* 20(5), 365-72
79. Jennings SM, Tsay YH, Fisch TM, Robinson GW (1991) Molecular cloning and characterization of the yeast gene for squalene synthetase. *Proceedings of the National Academy of Sciences of the United States of America* 88(14), 6038-42
80. Soltis DA, *et al*. Expression, purification, and characterization of the human squalene synthase: use of yeast and baculoviral systems. *Archives of Biochemistry and Biophysics* 316(2), 713-23
81. Nakashima T, Inoue T, Oka A, Nishino T, Osumi T, Hata S, (1995) Cloning, expression, and characterization of cDNAs encoding Arabidopsis thaliana squalene synthase. *Proceedings of the National Academy of Sciences of the United States of America* 92(6), 2328-32
82. Hanley K *et al*. (1996) Molecular cloning, *in vitro* expression and characterization of a plant squalene synthetase cDNA. *Plant Molecular Biology* 30(6), 1139-1151
83. Kribii RA *et al*. (1997) Cloning and characterization of the Arabidopsis thaliana SQS1 gene encoding squalene synthase. Involvement of the C-terminal region of the enzyme in the channeling of squalene through the sterol pathway. *European Journal of Biochemistry* 249(1), 61-69
84. Merkulov S *et al*. (2000) Cloning and characterization of the Yarrowia lipolytica squalene synthase (SQS1) gene and functional complementation of the Saccharomyces cerevisiae erg9 mutation. *Yeast* 16(3), 197-206
85. Busquets A *et al*. (2008) Arabidopsis thaliana contains a single gene encoding squalene synthase. *Plant Molecular Biology* 67(1-2), 25-36
86. Mo C, Bard M (2005) Erg28p is a key protein in the yeast sterol biosynthetic enzyme complex. *Journal of Lipid Research* 46(9), 1991-1998

87. Brooks PC *et al.* (1996) Localization of matrix metalloproteinase MMP-2 to the surface of invasive cells by interaction with integrin αvβ3. *Cell* 85(5), 683-693

Vita

1) Born at Central Baptist Hospital, Lexington, KY on October 25, 1982

2) Bachelor of Science in Agricultural Biotechnology at the University of Kentucky, Fall 2006.

3) No professional positions held.

4) Graduated *cum laude* from the University of Kentucky. Won best student presentation at Terpnet 2009.

5) Publications:

   Niehaus TD, Okada S, Devarenne TP, Watt DS, Sviripa V, Chappell J (2011) Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii. Proc Natl Acad Sci USA* **108:** 12260-12265

   Niehaus TD, Kinison S, Okada S, Devarenne TP, Chappell J (2011) Functional Identification of triterpene methyltransferases from *Botryococcus braunii Race B.* (In Review)

   Niehaus TD, Chappell J (2011) Identification of a carboxy-terminal, non-catalytic domain of fungal squalene synthases that is essential for directing squalene flux into the sterol biosynthetic pathway. (In preparation)

6) Thomas D Niehaus