DIRECTED EVOLUTION OF TERPENE SYNTHASES FOR NON-NATURAL SUBSTRATES

Thesis by

Karolina Zbigniew Kalbarczyk

In Partial Fulfillment of the Requirements for the degree of

Chemical Engineering

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

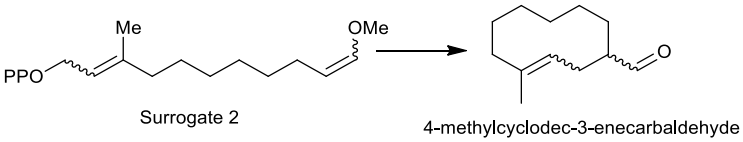
2013

ACKNOWLEDGEMENTS

The author would like to acknowledge the support of the SURF office and the California Institute of Technology. Special thanks to the Arnold Laboratory, in particular Professor Frances H. Arnold and Dr. Ryan Lauchli.

ABSTRACT

Terpenes represent about half of known natural products, with terpene synthases catalyzing reactions to increase the complexity of substrates and generate cyclizations of the linear diphosphate substrates, therefore forming rings and stereocenters. With their diverse functionality, terpene synthases may be highly evolvable, with the ability to accept a wide range of non-natural compounds and with high product selectivity. Our hypothesis is that directed evolution of terpene synthases can be used to increase selectivity of the synthase on a specific substrate. In the first part of the work presented herein, three natural terpene synthases, Cop2, BcBOT2, and SSCG_02150, were tested for activity against the natural substrate and a non-natural substrate, called Surrogate 1, and the relative activities on both the natural and non-natural substrates were compared. In the second part of this work, a terpene synthase variant of BcBOT2 that has been evolved for thermostability, was used for directed evolution for increased activity and selectivity on the non-natural substrate referred to as Surrogate 2, shown below. Mutations for this evolution were introduced using random mutagenesis, with error prone polymerase chain reactions, and using site-specific saturation mutagenesis, in which an NNK library is designed with a specific active site amino acid targeted for mutation. The mutant enzymes were then screened and selected for enhancement of the desired functionality. Two neutral mutants, 19B7 W367F and 19B7 W118Q, were found to maintain activity on Surrogate 2, as measured by the screen.



Surrogate 2 and Product.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION/BACKGROUND

a. Introduction to Terpene Synthases

Terpenes represent about half of known natural products, with terpene synthases having a variety of uses ranging from functions in primary metabolism, including as retinals and steroids, to functions in secondary metabolism, such as chemical attractants and repellents [19]. Amongst the most valuable terpenes are the pharmaceutically used products paclitaxel and artemisinin; the former being used for chemotherapy and the latter to treat malaria (Fig. 1) [20]. Natural terpenes can often have highly desirable functionality, stemming from their remarkable selectivity in binding to receptors. This selectivity is due to their complexity in terms of containing chiral centers, dense functionality, and multi-cyclic aliphatic structures.

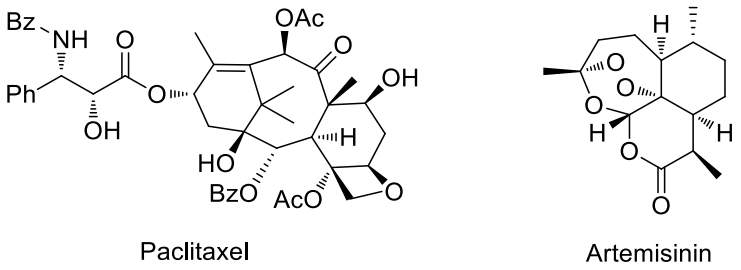


Figure 1. Valuable potential pharmaceutical products Paclitaxel and Artemisinin.

To make terpenes, nature starts with the five-carbon building blocks isopentyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) [4]. These molecules are enzymatically joined to form longer chain diphosphates, which behave as substrates for the terpene synthases. The terpene synthases catalyze reactions that increase the complexity of the linear molecules, effecting cyclizations thereby forming rings and stereocenters (Fig. 2) [15]. From these cyclized structures, a large variety of compounds can be generated through further enzymatic conversions [11].

Furthermore, terpenes generally function as secondary metabolites, and terpene synthases produce a range of products [19]. In primary metabolic enzymes, natural selection has produced highly specific enzymes, with a catalytic landscape as seen in Figure 3. With specialized metabolic enzymes, such as terpene synthases, there exists a wider landscape of products that are produced with less specificity. With this diverse functionality, we believe that terpene synthases have the potential to be highly evolvable, with the ability to accept a wide range of non-natural compounds, and to be evolved to have high product selectivity [6], [17].

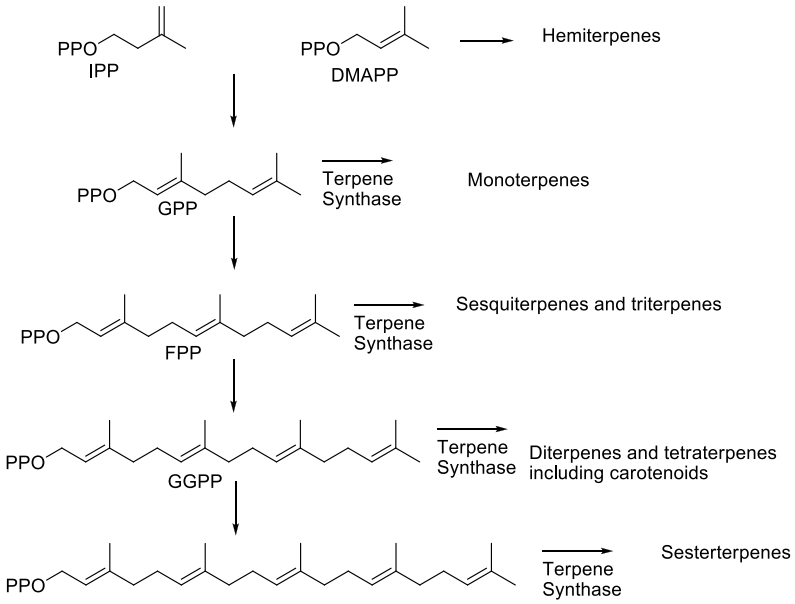


Figure 2. Terpene Biosynthetic Pathway.

The C15 isoprenoid diphosphate, farnesyl diphosphate (FPP), serves as the precursor for about 300 distinct terpene synthase products, called sesquiterpenes [11]. Further diversity of products can be created through biosynthetic derivatization in the form of P450-mediated oxidations, enzymatic halogenations, and esterifications [18].

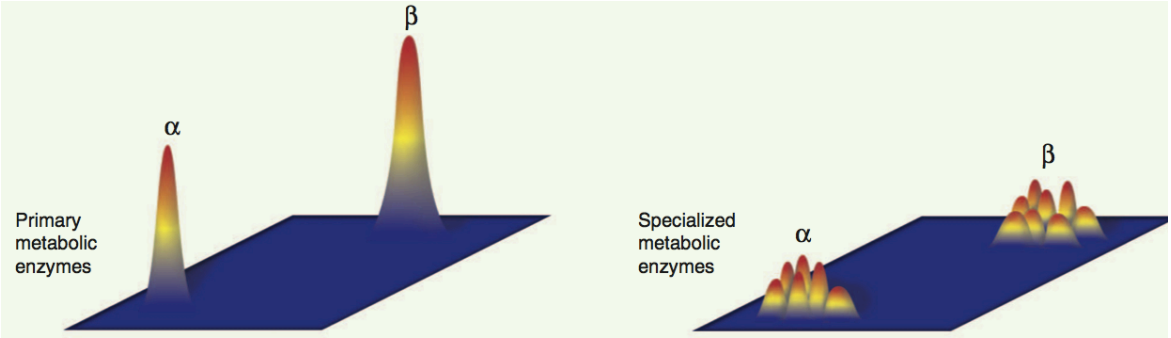


Figure 3. Specificity of catalytic landscapes of a primary metabolic enzyme vs. a specialized metabolic enzyme [19].

b. Previous Engineering Studies of Sesquiterpene Synthases

Current literature has shown that terpene synthases can be engineered to use natural substrates to make new products, where the initial FPP substrate is capable of making a new product [21]. In work by Yoshikuni, rational design and random mutagenesis were combined to evolve a sesquiterpene synthase, (+)-δ-cadinene synthase, to function as a germacrene D-4-ol synthase. The focus was to create a variant that could shift from producing primarily (+)-δ-cadinene, which is the function of the natural terpene synthase, to produce germacrene D-4-ol, in order to demonstrate terpene synthase engineering (Fig. 4). Twenty-one active variants were found, which were able to take the natural FPP substrate and generate products different from the parent enzymes [21].

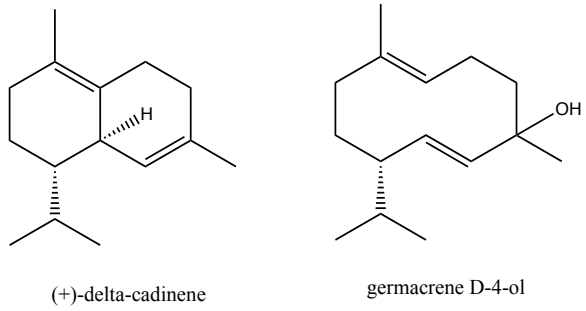


Figure 4. Comparison of the natural product, (+)-δ-cadinene, to the selected germacrene D-4-ol product.

In order to identify active mutants, the group analyzed potential variants with GC-MS. This method was low throughput, and could only screen about 100 samples/day, so the group developed a selection for properly folded proteins to narrow down potentially improved mutants for screening. In this selection, variants were fused to chloramphenicol acetyltransferase (CAT), and proteins that exhibited proper folding were therefore soluble. Thus, recombinant *E. coli* containing this fusion construct grew in the presence of chloramphenicol and cells with improperly folded protein did not. This assay served to determine which variants should be analyzed by GC-MS, but was limited to a determination of enzyme solubility. The test was also not a good indicator of activity, which could have been valuable in predicting improved variants [21].

The results attained by Yoshikuni's screening determined that the variants with the most altered functionality were N403P and L405H. The parent terpene synthase produced (+)-δ-cadinene with 98% selectivity, while in the final N403P and L405H variants germacrene D-4-ol was produced with 52%-53% selectivity. Identification of the products without an overlayer revealed that the L405H variant has a selectivity of 93% for the germacrene D-4-ol product, a result that may be affected by evaporation of other products [21].

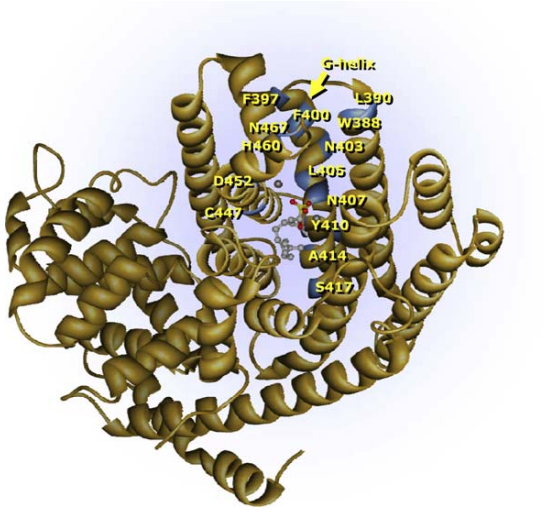


Figure 5. Homology structure of (+)-δ-cadinene (modeled on the crystal structure of 5-*epi*-aristolochene) and structural annotation of mutations obtained from random mutagenesis [21].

Identifying the ten most successful variants capable of producing germacrene D-4-ol, these corresponding mutations were mapped to a homology structure of (+)-δ-cadinene synthase, in order to determine the location of the active site. The mutations identified were centered primarily on the G-helix, shown in Figure 5 [21]. Overall, Yoshikuni's work indicated the potential to evolve terpene synthases for product selectivity by creating mutations in the active site.

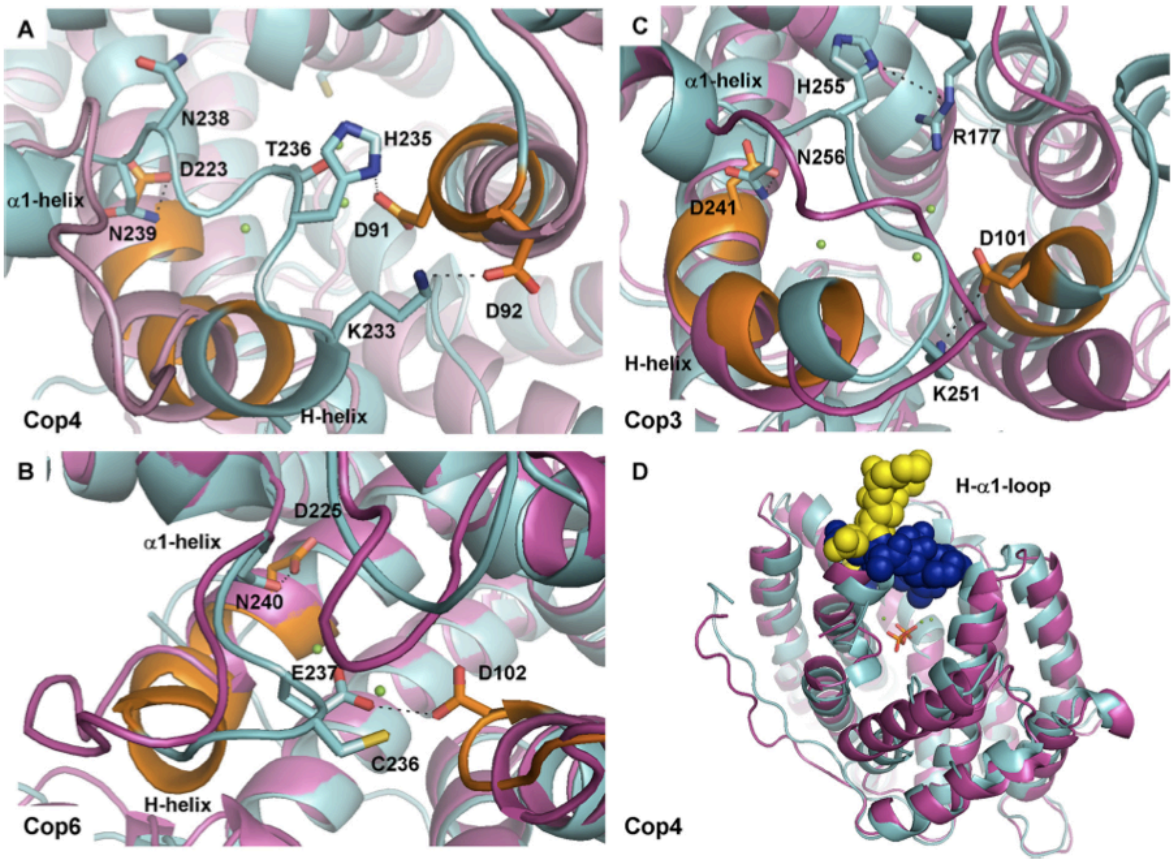


Figure 6. Homology models of (A) Cop4, (B) Cop6, (C) Cop3, and (D) superimposed Cop4 open structure (yellow H-α1 loop) and Cop4 closed structure (blue H-α1 loop). Magenta indicates the open conformation; cyan indicates the closed conformation [12].

López-Gallego was interested in modifying natural terpene catalysts using site-directed mutagenesis, to vary the product profiles of the wild-type synthases. The research focused

on comparing conserved active sites of three terpene synthases: Cop3, Cop4, and Cop6. To identify the important qualities of the active sites, the models were compared at both the open (unbound) and closed (ligand –bound) conformations. Homology models of Cop3 and Cop4 were made using the crystal structure of aristolochene synthase, while Cop6 was modeled using the structure of trichodiene synthase (Fig. 6) [12].

When superimposed, the three models (Fig. 6) indicated the H-α1 loop region was part of the active site, which in the closed conformation allowed the magnesium ion cluster, where the diphosphate binds, to move inwards. In this closed conformation, the three enzymes had active site cavities of the same volume. With Cop4 and Cop3 in the open conformation, the region was larger and more promiscuous, while the Cop6 in the open conformation had an active site that was smaller and more selective, at about half the volume. Looking at the product profiles for the three enzymes, the Cop6 was already very specific, while the Cop3 and Cop4 had more range in the product profile. In performing site-directed mutagenesis, mutation sites within the H-α1 loop for the Cop6 sesquiterpene synthase had little effect on the product profile (Fig 7A). For the Cop4 and Cop3 enzymes, specific mutations enabled a greater diversity of products in some instances, and created more specialized variants of the sesquiterpene synthases in other cases (Fig. 7) [12].

In addition to the site-specific mutations, the authors also attempted to switch the more promiscuous active site of Cop4 with the extremely selective active site of Cop6. The results of this switch indicated that the Cop6 and Cop6L4 (with the Cop4 acitve site) maintained essentially the same activity, while the Cop4 and Cop4L6 (with the Cop6 active site) exhibited completely different product selectivity (Fig. 8). Overall, this work suggested that identifying the H-α1 loop active site and targeting it for mutation was capable of generating product diversity of sesquiterpene synthases.

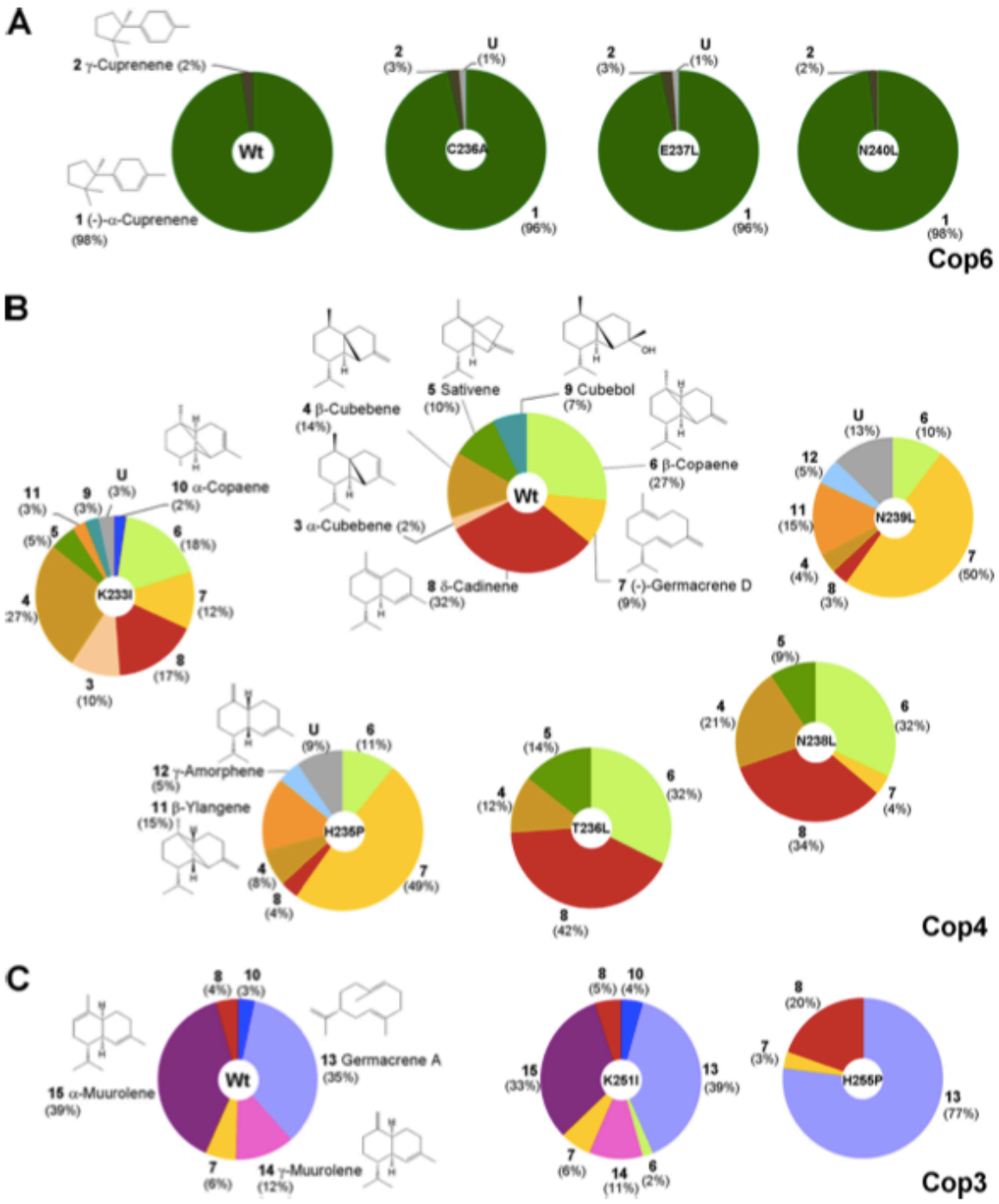


Figure 7. Influence of site-specific H-α1 loop mutations on product profiles of (A) Cop6, (B) Cop4, and (C) Cop3. Relative ratios of product can be seen in the charts, and compared to the wild-type charts given as Wt [12].

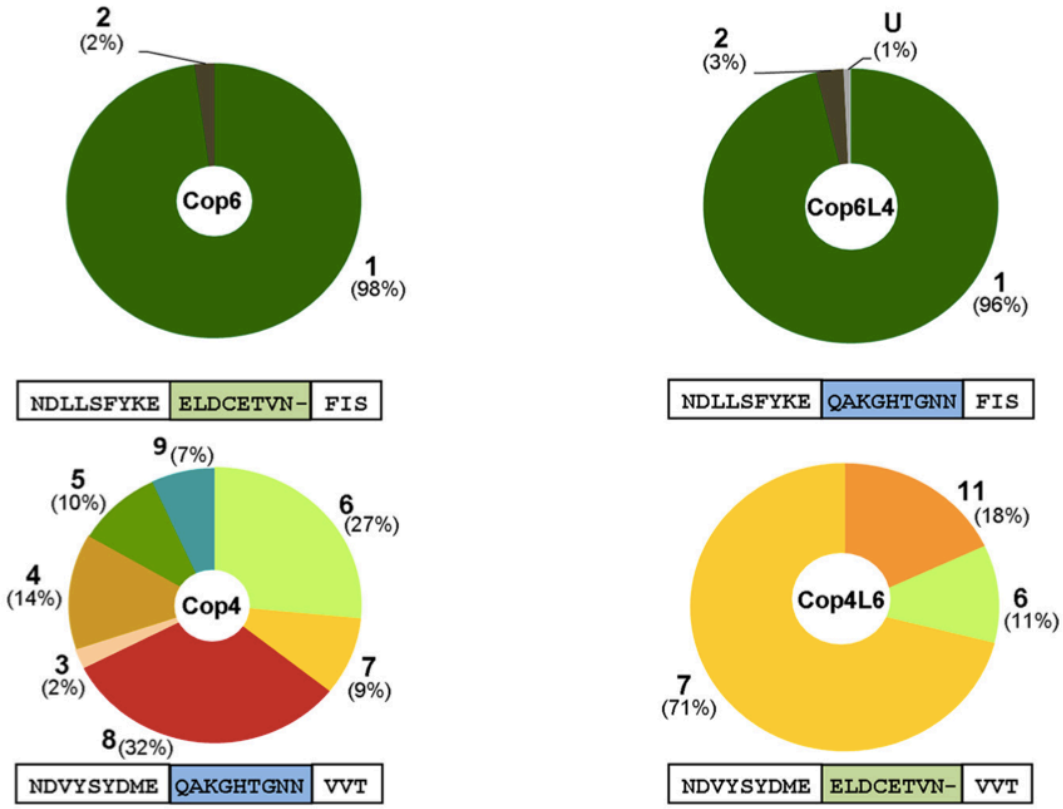


Figure 8. Comparison of the influence of loop grafted Cop4 and Cop6 on the product profiles. Relative ratios of product can be seen in the charts, and with the original wild-type enzyme product profiles on the left and the loop grafted variants on the right [12].

Yoshikuni's work was limited by the lack of a convenient screen to determine sesquiterpene synthase activity, which restricted the choice of terpene synthase possible to screen as well as the number of variants that could be easily tested and manipulated. López-Gallego was restricted by site-directed mutagenesis, because they used GC-MS for characterization of product. With an improved screen to determine activity of sesquiterpene synthases, efficient directed evolution using random mutagenesis of many different terpene synthases could be performed, which could produce selective enzymes for a wide range of new terpene products.

c. Reactions of Analogs of FPP with Sesquiterpene Synthases

Above it has been shown that terpene synthases can be engineered to follow new pathways and make new products. Work has also been done to look at the effects of changing the natural substrate, and thereby generate novel terpene compounds.

In work by Cascón, analogous FPP substrates and two plant terpene synthases were used to make a library of new germacrene products [3]. The two plant sesquiterpene synthases used were germacrene A synthase (GAS), which converts FPP into germacrene A, and germacrene D synthase (GDS), which converts FPP into germacrene D. The original FPP was modified into FPP analogues, shown in Figure 9 [3].

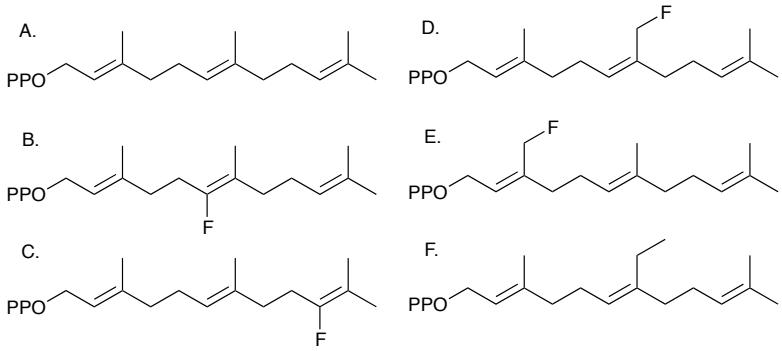


Figure 9. FPP (A) and five modified analog substrates (B-F).

Reactions of the two synthases with the FPP analogues generated a library of germacrene analogues. The work has shown that the terpene synthase enzymes were capable of processing farnesyl diphosphate analogues into cyclized products [3].

Observed success with the FPP analogues supports the conjecture that directed evolution of terpene synthases could be used to create enzymes capable of more effectively processing non-natural substrates. There are many opportunities to construct terpene synthases with altered functionality, and which could therefore access a whole new range of diverse and useful products (Fig. 10) [7], [14].

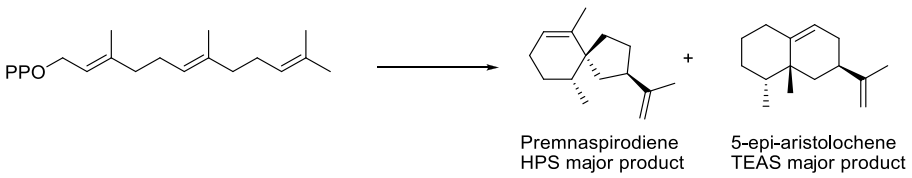


Figure 10. Cyclization of FPP to give two representative products.

d. Project Hypothesis

The hypothesis is that directed evolution can be used to increase selectivity of a terpene synthase on a specific substrate. The specific aims of this research were to:

1. Test three natural terpene synthases, Cop2, BcBOT2, and SSCG_02150, for activity against the natural substrate and a non-natural substrate, called Surrogate 1, and then to compare the relative activities on both the natural and non-natural substrates.
2. Use directed evolution of a terpene synthase variant of BcBOT2 that has been evolved for thermostability to increase activity and selectivity on non-natural substrate Surrogate 2.

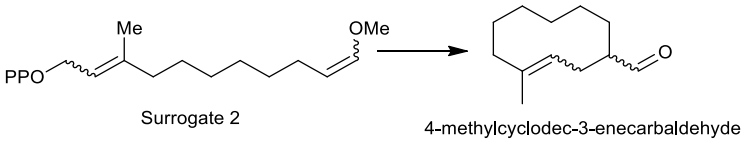


Figure 11. Surrogate 2 and product.

The reaction of Surrogate 2 with the terpene synthase BcBOT2 creates the product shown in Figure 11. Creating compounds with *trans* double bonds within a ring is difficult with synthetic chemistry, especially in small rings with high strain. This project could allow a faster and more efficient way to create new terpene synthases for the purpose of synthesizing new products. For the first time, terpene synthases could be evolved for non-natural terpenes.

e. Approach

i. AOX-Purpald Assay Screening

In order to screen for activity, an assay was developed to test for the specific product formation. To apply this assay, surrogate FPP analog substrates with a methoxy group were created (Fig. 12).

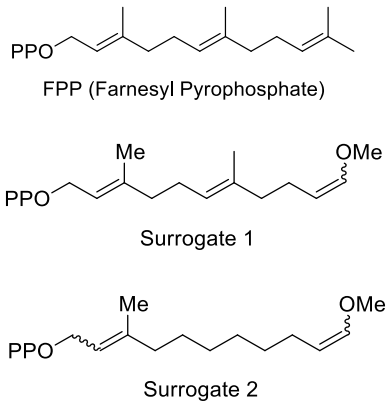


Figure 12. Structures of FPP (the natural substrate), Surrogate 1, and Surrogate 2.

During reaction of the substrate and the terpene synthase, the cyclization would cause the release of methanol as a side product alongside the desired complex ring structure product. In previous studies, reaction of Surrogate 1 with the terpene synthase produced the products germacrene A and methanol, shown in Figure 13.

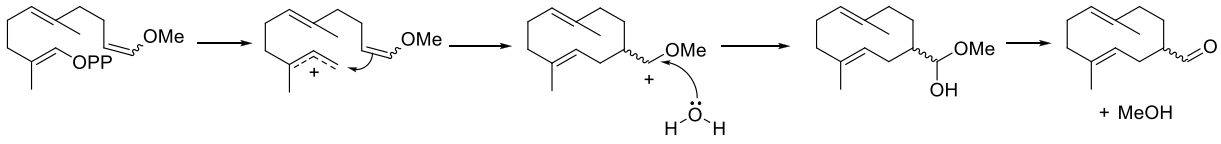


Figure 13. Cyclization reaction of Surrogate 1.

In the presence of alcohol oxygenase (AOX), the methanol was converted to formaldehyde. Ethylenediaminetetraacetic acid, EDTA, was used to kill the reaction between the lysate and substrate, and finally, purpald was added to cause a colorimetric response. In the AOX-

Purpald assay, the formaldehyde reacted with purpald to produce a purple color corresponding to the concentration of product (Fig. 14) [8].

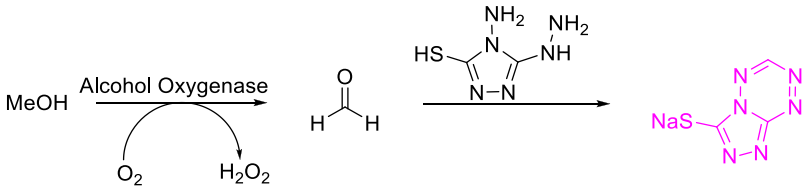


Figure 14. Reaction of methanol with AOX-Purpald assay to produce colorimetric response.

After the reaction with the purpald took place, some wells appeared purple while others appeared clear, and there was a gradient of different intensities of purple. These color changes were measured using a plate reader, recording the absorption at 550 nm. This color indicated which mutants were capable of producing methanol. Following the preliminary assays, mutants producing the purple color were re-screened using gas chromatography to screen for specific products.

ii. Directed Evolution through Mutagenesis

Directed evolution was chosen as a strategy to create variants for screening with the AOX-Purpald assay. To develop terpene synthases with desired functions, the thermostable terpene synthase referred to as 19B7 was used as a starting point. This thermostable variant was a mutant of the original fungal terpene synthase BcBOT2, shown modeled on pentalenene synthase in Figure 15.

This particular terpene synthase was used as a starting point for directed evolution because the terpene synthase BcBOT2 is one of the fastest natural terpene synthases available. In the model, the black region of the model indicates the magnesium ion binding region, which allows the diphosphate complexes to bind. The red region represents the areas that were mutated in order to produce the 19B7 thermostable structure. The increased

thermostability of the terpene synthase allows for more mutations to be introduced while maintaining activity [2], [5].

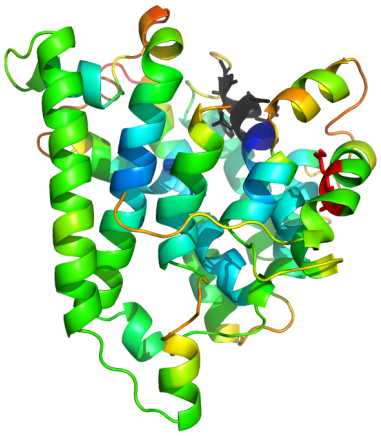


Figure 15. Proposed BcBOT2 structure modeled on pentalenene synthase; SWISS-MODEL, Protein Structure Bioinformatic group at the Swiss Institute of Bioinformatics and the Biozentrum University of Basel.

The primary goal of the research presented in this thesis was to use directed evolution to increase activity of the terpene synthase on a specific substrate. Mutations for this evolution were introduced using random mutagenesis, with error prone polymerase chain reactions, and using site-specific saturation mutagenesis, where an NNK library is designed with a specific active site amino acid targeted for mutation. The mutant enzymes were then screened and selected for enhancement of the desired functionality.

Chapter 2

MATERIALS AND METHODS

a. Double Digestion

In preparing a vector or insert from a given cryostock, cultures were grown in Luria Broth (LB) with Kanamycin overnight and then purified via the Qiagen Miniprep kit. Double digest reactions were set up for the vector and inserts in 50 µL samples. With 1 µL of XhoI New England Biolabs enzyme, 1 µL of NdeI New England Biolabs enzyme, 5 µL of New England Bio Labs Buffer 4, 0.5 µL of New England Biolabs BSA, and 1-2 µg of DNA, the remainder of the volume was filled up to 50 µL with autoclaved distilled water. The sample was then left at 37 °C for 1 hour, and purified through gel extraction, to be later used in ligation.

b. Agarose Gel and Extraction

A 1% agarose gel was poured (adding gel red) and left for about 20 minutes to solidify. The samples were run against a 1 KB ladder, with a ratio of 5:1 for dye to sample. The gel was run at about 100 V for about 20 minutes, and the resulting gel was analyzed under UV light. From preparatory gels, the gel was extracted using the procedure given by the Promega PCR clean-up system.

c. Error Prone PCR

The PCR mix was made of 1 µL of BcBOT2 forward primer (100 ng/µL for all primers), 1 µL of BcBOT2 reverse primer, 1 µL of 2 mM dNTP's, 5 µL of 2 mM $MnCl_2$, 5 µL 10x Roche Taq Polymerase Buffer, 0.5 µL of 19B7 DNA template, and 0.5 µL of Roche DNA Taq polymerase, with the remainder of the volume filled up to 50 µL with autoclaved distilled water. The program began with 98 °C for 3 minutes, followed by 30 cycles of 98 °C for 0:30 / 62 °C for 0:30 --0.5 °C where the temperature was gradually lowered by 0.5

°C per cycle/ 72 °C for 1:45 ++0:05 where the extension time was gradually extended by 5 seconds each cycle. With a final extension at 72 °C for 10 minutes, the program ended at 10 °C for the remaining time.

d. NNK library generation using overlap extension PCR

The PCR program consisted of three rounds in order to combine the NNK library primers with the primers for the gene of interest. First there was the fragment PCR, where the PCR mix was made of 50 µL samples with 0.5 µL of template DNA (19B7), 1 µL of Phusion Polymerase, 10 µL of the 5x High Fidelity Phusion Buffer, 1 µL of dNTP's, 1 µL of forward primer, 1 µL of reverse primer, and the rest with autoclave distilled water. [Note the enzymes used were generally acquired from New England Biolabs.] Two mixes were set up for each of the sets of primers, one with the specific forward primer and the BcBOT2 reverse primer, the other with the specific reverse primer and the BcBOT2 forward primer. The fragment PCR program began with 98 °C for 2 minutes, followed by 35 cycles of 98 °C for 0:30 / 55 °C to 75 °C (gradient PCR annealing temperatures) for 0:30 / 72 °C for 1:30, and then a final extension at 72 °C for 10 minutes. The program ended at 10 °C for the remaining time, followed by preparation of the PCR assembly mix to combine the two products for each primer.

Before the assembly step, a preliminary gel was run to see that there was product (as well as a gel extraction to purify the product). To set up the assembly PCR, the following 50 µL PCR mix was prepared, with 1 µL of template DNA from each of the fragment samples for the specific NNK, 1 µL of Phusion Polymerase, 10 µL of the 5x High Fidelity Phusion Buffer, 1 µL of dNTP's, and the rest with autoclave distilled water. The mix was run in an assembly PCR program which again began with 98 °C for 2 minutes, but this time followed by 20 cycles of 98 °C for 0:30 / 55 °C to 75 °C (gradient PCR annealing temperatures) for 0:30 / 72 °C for 1:30, and a final extension at 72 °C for 10 minutes. The program ended at 10 °C for the remaining time. The final assembly PCR program was run after 1 µL of BcBOT2 forward primer and 1 µL of BcBOT2 reverse primer had been

added. The program was the same as the previous assembly PCR program, but with 25 cycles instead of 20.

Note that for the final libraries, an improved set of PCR conditions was used. In this procedure, the fragment assembly PCRmix was1µL of forward primer, 1µL of reverse primer, 1µL of dNTP's, 10µL of New England Biolabs 5x High Fidelity Phusion Buffer, 1µL of New England Biolabs Phusion Polymerase, 0.5µL of DNA (19B7) and finally 35.5µL of autoclaved distilled water. The fragment PCR program again began with 98 °C but this time only for 0:30, followed by 30 cycles of 98 °C for 0:10 / 55 °C to 75 °C (gradient PCR annealing temperatures) for 0:20 / 72 °C for 0:20, and a final extension at 72 °C for 10 minutes. The program ended at 10 °C for the remaining time.

To set up the assembly PCR, the same 50 µL PCR mix was prepared as before. The mix was run in an assembly PCR program which began with 98 °C for 0:30, followed by 20 cycles of 98 °C for 0:10 / 55 °C to 75 °C (gradient PCR annealing temperatures) for 0:20 / 72 °C for 0:20, and a final extension at 72 °C for 10 minutes. The program ended at 10 °C for the remaining time. The final assembly PCR program was run after 1 µL of BcBOT2 forward primer and 1 µL of BcBOT2 reverse primer had been added. The program was at 98 °C for 2 minutes, followed by 25 cycles of 98 °C for 0:30 / 55 °C to 75 °C (gradient PCR annealing temperatures) for 0:30 / 72 °C for 1:45, and a final extension at 72 °C for 10 minutes. The program ended at 10 °C for the remaining time.

e. Ligation

Ligation of an insert into a vector was typically done with 50 ng of the epPCR insert, 50 ng of vector, 2 µL of ligase buffer, 1 µL of New England Biolabs T4 DNA ligase enzyme, and the remainder of the volume filled up to 20 µL with autoclaved distilled water. The ligation was left overnight at 16 °C, and then transformed following the standard transformation procedure.

Note that for the final libraries, the ligation mix was modified for the NNK library. Due to inaccuracy of DNA concentration measurements, the mix was modified to have 10 µL of insert, 2.5 µL of vector, 1.5 µL of ligase buffer, 2 µL of New England Biolabs T4 DNA ligase enzyme, and no added water, to make a total volume of 15 µL. The ligation was still left overnight at 16 °C, and then transformed following the standard transformation procedure.

f. Transformation

From an existing ligation, transformation into LB agar plates with Kanamycin (50 mg/L) was typically done with 1 µL of the ligation added to 25 µL of thawed BL21 *E. coli* electrocompetent cells. The cells and DNA were transferred to an electroporation cuvette and pulsed at 2.5 kV. Immediately after, 950 µL of recovery media were added to the cuvette, and it was left to recover in a 37 °C shaker for 1 hr. From this mix, 200 µL were plated onto a large plate using bead shaking to spread the cells around the plate. The colonies were left in a 37 °C incubator overnight to grow.

g. Making Libraries

To make the library, colonies were picked from a large plate where error-prone PCR had been transformed into a large Luria Broth (LB) agar plate with Kanamycin. With sterilized toothpicks, individual colonies were picked into 96-well shallow plates with 300 µL of LB with Kanamycin in each well. These plates were left in a 37 °C shaker at 250 rpm and 80% humidity overnight. The next day 50 µL of these cultures were transferred to deep well plates with 800 µL of Terrific Broth (TB) with Kanamycin and left in a 37 °C shaker at 250 rpm and 80% humidity for 4 hours, at which point the temperature was lowered to 25 °C for 30 minutes. The samples were then induced with 50 µL of 9 mM IPTG in TB and left in a 25 °C shaker at 250 rpm and no humidity overnight. The resulting cultures were spun down at 3,000 g at 4 °C for 10 minutes, after which the supernatant was poured out and the resulting pellets were frozen at -20 °C, ready for screening after freezing overnight.

h. Glycerol Stocks

To make glycerol stocks of plates from the libraries, 120 µL of the material in the shallow wells was combined with 80 µL of 50% glycerol in microtiter plates, for a final concentration of 20% glycerol. Note that the plates were sterilized before use by spraying with ethanol and drying under the hood. These were sealed with USA Scientific Cold Storage Foil, and stored at -80 °C.

i. Streaking Plates

From an existing glycerol stock, a streaked out plate was made to isolate individual colonies, to allow these to grow into cultures to eventually attain purified DNA for sequencing or use as DNA template in PCR reactions. Streaking was done onto LB agar plates with Kanamycin, using sterilized toothpicks or disposable inoculating loops. The first toothpick or loop was used to transfer the glycerol stock to the plate using three strokes. Then a second clean toothpick or loop was used to streak three more lines that start from the strokes of the first toothpick or loop, to further separate the colonies out. A third clean toothpick or loop was used to stroke out three more lines, starting from the strokes of the second toothpick or loop, to help further distinguish the colonies. The plate was then left overnight in a 37 °C incubator overnight.

j. Miniprep DNA

Minipreping DNA started by picking a colony from a transformed or streaked out plate into 4 mL of LB with Kanamycin, in a 15 mL round-bottomed culture tube. The culture was grown overnight in a 37 °C shaker, at a rate of 225 rpm or 250 rpm. The resulting cells were spun down for 10 minutes at 4,000g, and DNA was extracted using the protocols and materials of the QIAprep Spin Miniprep Kit by Qiagen. The resulting purified DNA was sent for sequencing, or used as DNA template for various PCR reactions.

Note that sequencing samples were prepared by first streaking a cryostock, then growing cultures from the plate, and finally performing minipreps.

k. Stamping Plates

Plates were stamped using a metal 96 well stamp, which wes sterilized with ethanol under a flame. The stamp was then placed in the desired 96-well cryostock plate, finally to be transferred to shallow 96-well plates with 300 µL LB with Kanamycin in each well. The remainder of the procedure followed exactly as with making the original libraries.

l. Screening Substrates (with AOX Purpald Assay): Small Scale

Screening for all libraries and substrates started by thawing deep well plates with cell pellets at room temperature for 30 minutes. Then, to lyse the cells, 300 µL of lysis buffer were added to each well, where the lysis mix for one plate was made with 35 mL of standard buffer (pH: 7.6, 50 mM 1,4-piperazinediethanesulfonic acid (PIPES), 10 mM magnesium chloride, 100 mM sodium chloride, 2 mM dithiothreitol) with 0.5 mg/mL lysozyme and 0.02 mg/mL DNase I. The plate was vortexed to ensure resuspension of the pellets, and then left in a 37 °C incubator for 1 hr. After incubation, the plate was spun down at 5,000 g at 4 °C for 15 minutes, separating the liquid lysate from the solid pellet at the bottom of the wells. Using a liquid-handling robot, 100 µL of lysate from each well was transferred to 96-well microtiter plates, where 100 µL of substrate in standard buffer was added. (Note that for the final, most recent library, this procedure was done by hand underneath the hood in order to prevent contamination.) The substrate was made so as to allow for a 0.5 mM final concentration. Typically, the original stock of substrate was at 10 mM, so a 20x dilution into standard buffer was done immediately preceding addition of the substrate to the lysate. (This was done because the dithiothreitol in the standard buffer has the potential to degrade.)

After the lysate and substrate were given 1-2 hours to react at 37 °C, 10 µL of alcohol oxygenase (50 µL of AOX stock solution dissolved in 950 µL of 0.1 M KPi, pH 8.0) was added to each well for 10 minutes at room temperature and 600 rpm, converting the side product methanol into formaldehyde. Next 16 µL of 0.5 M Ethylenediaminetetraacetic acid (EDTA) of pH 8.0 was added to each well for 1 minute at room temperature and 600rpm,

to stop the reaction. Finally, 50 µL of 0.16 M purpald in 2 M NaOH was added to each well. The plate was shaken at 600 rpm for 25 minutes at room temperature, after which a microtiter plate reader at 550nm and 10 flashes was used to record the absorbance of the wells. With the reaction of formaldehyde and purpald, a purple color appeared, typically indicative of high activity enzymes to be screened further.

m. Screening Substrates (with Gas Chromatography): Medium Scale

Screening done at the medium scale was done with 15 µL of 10mM substrate, 480 µL of standard buffer (with 10% betaine for certain reactions), and 5 µL of 2 µM enzyme. After these three components were mixed, 500 µL of hexane was added on top, carefully so as to not disturb the reaction. The samples were left to react at 30 °C, after which they were vortexed to stop the reaction, and centrifuged for 2 minutes at 10,000 g to separate the material into three layers – two liquid layers separated by a gel layer. The top layer of hexanes contained the product, which was extracted with a pipette and then run through a magnesium sulfate filter to remove excess water before running through a gas chromatography (GC) machine. Larger scale reactions were then performed, and analyzed using the gas chromatograph, which detected and distinguished the products.

n. Screening Substrates (with Protein Purification): Large Scale

Screening done at the large scale started with a 4 mL LB and Ampicillin pre-culture with a colony of the successful initial sample left overnight in a 37 °C incubator. The next day the sample was transferred to 500 mL of TB with ampicillin and metals mix, left shaking at 37 °C and 225 rpm for about 4 hours until the OD_{600} reaches 1.8. Then the sample was put in a 25 °C shaker for 30 min, and induced with 9 mM IPTG in TB for a final concentration of 0.1 mM. The large culture flask was left shaking overnight.

After the large culture grows overnight, it was poured into bottles and centrifuged at 5000 rpm for 25 min at 4 °C. The supernatant was discarded and the remaining pellet is left at -20 °C overnight. The next day the pellet was resuspended in 15 mL of Buffer A (pH: 7.6,

50 mM 2-amino-2-hydroxymethyl-propane-1,3-diol (TRIS), 5 mM magnesium chloride, 300 mM sodium chloride, 5 mM dithiothreitol, and 10 mM imidazole) and sonicated by pulsing for 0.5 seconds then resting for 0.5 seconds for the course of 1 minute. This sonication program was repeated three times, inverting the tube in between each round. The sample was then centrifuged at 15,000rpm for 20 minutes at 4°C, after which it was run through a purification column. The purification column was first saturated with Buffer A, and then the protein (in Buffer A) was poured in and forced through the column as well. Then the Buffer A was run through the column again, followed by Buffer B (pH: 7.6, 50 mM TRIS, 5 mM magnesium chloride, 300 mM sodium chloride, 5 mM dithiothreitol, and 500 mM imidazole) which eluted the purified protein in a clear but deep brown/amber color. This material was concentrated using 50mL Millipore concentration tubes, spinning four times at 3000g for 20min at 4°C, adding standard buffer to resuspend and concentrate. The final sample was frozen at -80°C.

o. Compilation Plate

A compilation plate was prepared using streaked plates or glycerol stocks, where individual cell samples or colonies were picked, using sterilized toothpicks or disposable inoculating loops, into a 96-well plate with 300 μL LB with kanamycin in each well. The resulting plate was grown with the remainder of the procedure following exactly as with making the original libraries.

p. DpnI Treatment

DpnI treatment was performed on the error prone PCR insert, or the fragment PCR insert. The addition of DpnI enzyme removed the excess template DNA. The initial minipreped template from *E. coli* was methylated. After PCR, only the original template retained methylation, so when the DpnI bound to the methylated DNA and degradeed it, the excess template decreases. In order to do this treatment, 1 μL of enzyme was added to the PCR sample and left at 37 °C for 1 hour.

q. Alkaline Phosphatase Treatment

Alkaline phosphatase treatment was performed on the digested vector or plasmid DNA portion. When the vector was digested, then there was phosphate remaining on the ends of the DNA, which allowed the vector to close back upon itself. The addition of CIP (calf intestine alkaline phosphatase) removed the phosphate ends, and lessened the likelihood of the DNA binding to itself. This increased the likelihood of successful gene insertions during the ligation stage. In order to do this treatment, 1 μL of enzyme was added to the PCR sample and left at 37 °C for 1 hour.

r. Random Mutagenesis

A gene of interest was evolved through mutagenesis, where random mutations are introduced into the DNA to make a variety of different sequences and corresponding enzymes. When a gene of interest was selected then the gene was digested with two restriction endonucleases, as was a plasmid of DNA. The gene was then inserted into the vector (portion of the plasmid) through a ligation, and the resulting DNA was transformed into cells using electroporation. After plating, the cells grew small colonies with the gene in their DNA plasmid, and these colonies were grown into cultures from which this final DNA was extracted. With the gene of interest now in a plasmid, the DNA was used as a template to evolve and make a new library of enzymes.

In directed evolution, the gene of interest in the vector was mutated through error prone PCR, making use of the natural error rate of Taq polymerase. The epPCR returned a mix of DNA with different mutations, and this DNA was transformed into electrocompetent cells to be plated to grow colonies with a variety of DNA sequences. These colonies were then grown as individual cultures, which were members of a library, where larger libraries had the benefit of expressing a greater range of mutations. Protein expression was induced, and the cells were frozen and kept at -20 °C.

Once the library was expressed, the cells were lysed and the enzyme was extracted to be screened for activity. Extracted lysate was combined with a substrate in standard buffer (see materials and methods), and left to react at 37°C. After the reaction has been given time to occur, a variety of preliminary assays were applied to determine which enzymes exhibit increased product formation. Larger scale reactions were repeated with the lysate, substrate, and buffer, which were then run through the gas chromatography machine to determine exactly which products were present.

Note that the 19B7 library was generated by picking an overnight culture, which was then extracted using a miniprep procedure. Next, epPCR samples were prepared from the 19B7 gene and subsequently treated with DpnI. Following gel extraction, the samples underwent double digestion using two endonucleases, XhoI and NdeI.

The vector 03688 was also picked and grown in an overnight culture, from which DNA was extracted using the miniprep procedure. After miniprepping, the samples were digested with XhoI and NdeI, and treated with alkaline phosphatase or CIP (calf intestine phosphatase). Finally, a ligation was set up with the 19B7 epPCR into the digested vector 03688. The ligation was left overnight at 16 °C, and then cleaned and concentrated using the Zymo Clean and Concentrate kit. This mix was then transformed into electrocompetent BL21 *E. coli* cells. These colonies were picked to make a library, which could then be screened.

s. Site-Directed Mutagenesis

With site-directed mutagenesis the process began with a fragment assembly PCR. The set up was for 50 µL samples with 0.5 µL of template DNA (19B7), 1 µL of Phusion Polymerase, 10 µL of the 5x High Fidelity Phusion Buffer, 1 µL of dNTP's, 1 µL of forward primer, 1 µL of reverse primer, and the rest with autoclave distilled water. Two mixes were set up for each of the sets of primers, one with the specific forward primer and the BcBOT2 reverse primer, the other with the specific reverse primer and the BcBOT2 forward primer. Then followed the fragment PCR program which was run as described in

the methods section. Before the assembly step, a preliminary gel was run to see that there was product, and the results were good, so this was followed with DpnI treatment to remove excess template.

Following the same gel extraction procedure as before, the samples underwent double digestion with XhoI and NdeI. Ligation was performed overnight, followed by transformation of the cells onto LB Kanamycin plates. The cells were left to grow overnight, proving to have been successful. Therefore a library was picked.

Using the library making procedure from before, a library of two plates was picked from each of the sets of primers, and screened for hits, also using the procedure for small scale screening described in the methods section. Note, the primers used for this procedure are given in Figure 16.

NNK Site	Forward Primer (5'->3')	Reverse Primer (5'->3')
W118NNK	CTGGCGAGCATGNNKGCTCCG	CTGGTGCATCCGGAGCMNNCATGC TCGCCA
A119NNK	GGCGAGCATGTGGNNKCCGGAT GCAACAG	CTGGTGCATCCGGMNNCCACATGC TCGCC
W133NNK	GGTTATGATGCTGGATNNKAAC CACTGGG	CCCAGTGGTTMNNATCCAGCATCA TAACC
V137NNK	GGAACCACTGGNNKTTTCTGTTT GATGACC	GGTCATCAAACAGAAAMNNCCAG TGGTTCC
F138NNK	CCACTGGGTANNKCTGTTTGATG ACCAGTTC	GAACTGGTCATCAAACAGMNNTAC CCAGTGG
W367NNK	GGCAACCTGTATNNKAGCTTCC AGACCG	CGGTCTGGAAGCTMNNATACAGGT TGCC

Figure 16. Primers for the NNK sites for 19B7.

Chapter 3

RESULTS AND DISCUSSION

a. Comparison of FPP and Surrogate 1 Activity

Data was gathered to compare the reactions that occur between a variety of natural terpene synthases with the natural substrate (FPP, farnesyl pyrophosphate), as well as Surrogate 1. The goal was to determine how the activity of various natural terpene synthases would be affected by the use of this non-natural Surrogate 1 substrate. Three wild-type terpene synthases were tested: Cop2, BcBOT2, and SSCG_02150. The primary products of these sesquiterpene synthases are shown in Figure 17 [10], [13], [16].

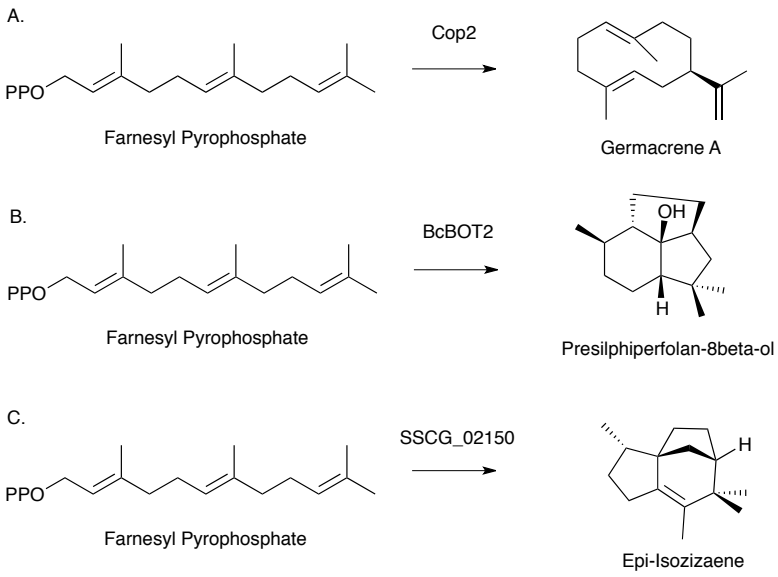


Figure 17. FPP cyclization products with (A) Cop2, (B) BcBOT2, and (C) SSCG_02150.

Cop2 catalyzes the cyclization of FPP to germacrene A. The mechanism for the reaction can be seen in Figure 18. Cop2 performs a 1,10 cyclization to create germacrene A, which could serve as a valuable precursor for biofuels [1].

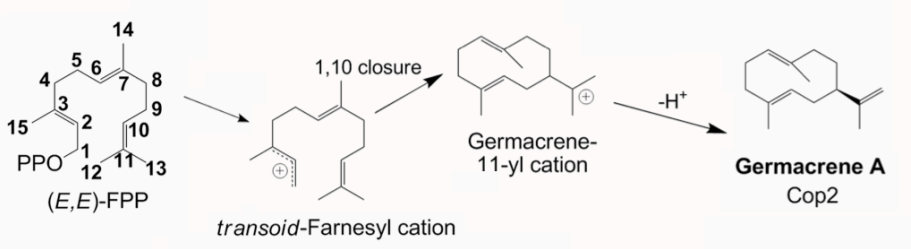


Figure 18. FPP cyclization mechanism in reaction with Cop2 [1].

BcBOT2 catalyzes the cyclization of FPP to presilphiperfolan-8β-ol [16]. The mechanism for the reaction can be seen in Figure 19. Presilphiperfolan-8β-ol is created by a 1,11 cyclization [13].

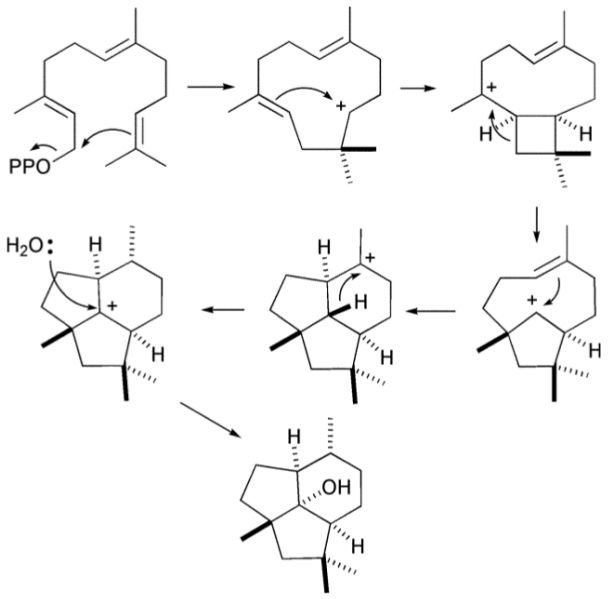


Figure 19. FPP cyclization mechanism in reaction with BcBOT2 [13].

SSCG_02150 catalyzes the cyclization of FPP to epi-isozaene [9]. The mechanism for reaction can be seen in Figure 20. In order to form epi-isozizaene, SSCG_02150 performs a 1,6 cyclization [13].

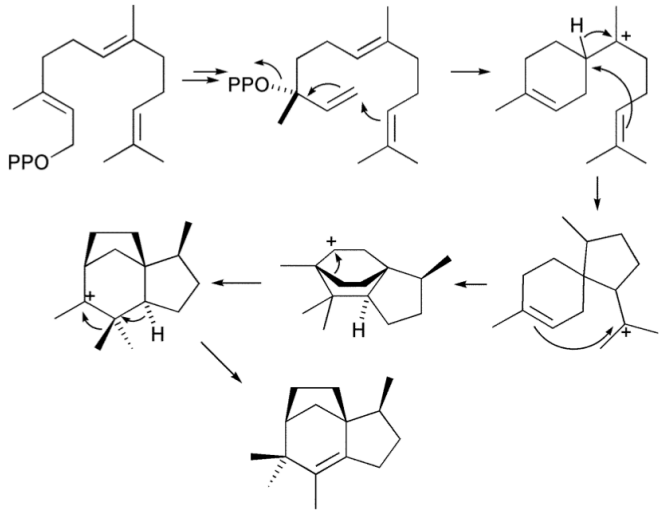


Figure 20. FPP cyclization mechanism in reaction with SSCG_02150 [13].

Taking purified protein of three terpene synthases, GC reactions were prepared with both the natural and the surrogate substrates. Data was gathered to compare behavior of FPP to Surrogate 1 with Cop2, BcBOT2, and SSCG_02150 as purified enzymes, for reaction times of 1 hour and over 24 hours (Fig. 21). The data represents the relative area-under-the-curve product formation, gathered from GC analysis.

For all of the synthases, at 24 hours, there is about double the amount of product formed from FPP compared to product from Surrogate 1. In fact, FPP always shows greater product formation than S1. BcBOT2 appears to be the fastest catalyst, creating the most product material with both FPP and S1 after 1 hour, as was expected because BcBOT2 is one of the fastest terpene synthases according to literature values. With Cop2, we see slower product formation, but after 24 hours this terpene catalyst can produce as much product as BcBOT2, for both the FPP and S1 substrates. Finally, with SSCG_02150, we see that in 1 hour, very little product is formed.

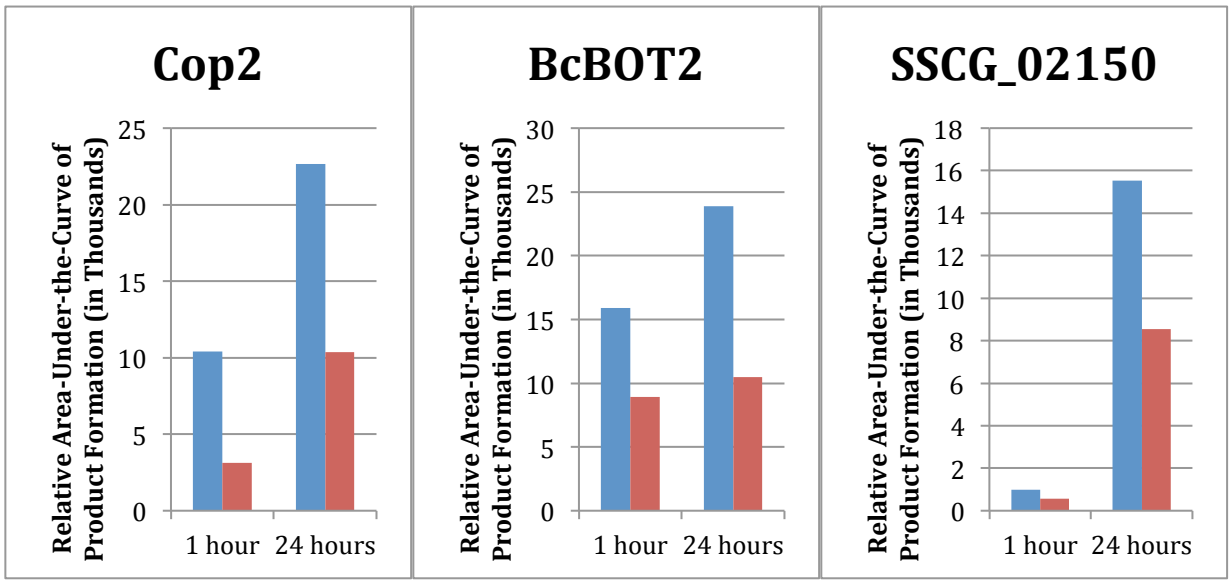


Figure 21. Product formation from farnesyl pyrophosphate (FPP) BLUE and Surrogate 1 (S1) RED with Cop2, BcBOT2, and SSCG_02150.

This data shows that BcBOT2 is the fastest functioning terpene synthase, showing that it is possible for a terpene synthase to perform a 1,11 cyclization efficiently. Cop2 is slower than BcBOT2, though both produce about the same amount of product after 24 hours. The mechanism for Cop2 involves a simpler cyclization than the mechanism for SSCG_02150, which was the slowest of the three sesquiterpene synthases and appeared poorly active at 1 hour. In order for the reactions of the substrates with SSCG_02150 to undergo 1,6 cyclization, the synthase must first undergo a rearrangement, which may slow down the overall product formation. SSCG_02150's continued productivity over the 24-hour period showed that the terpene synthase stays active for a relatively long time, suggesting that it is stable at the conditions tested. This stability suggests that SSCG_02150 could be a good starting point for evolution, as stable enzymes are less likely to lose function during mutagenesis.

For each of the terpene synthases, there has been analogous behavior between the natural substrate and the surrogate compound, with conversion of S1 at about half the rate of FPP. This consistent relationship between the two substrates was surprising, especially

considering the mechanistic differences between the enzymes. With BcBOT2 the cyclization includes a water-quenching termination step, and with SSCG_02150 and Cop2 there is a carbocation elimination, both of which require different chemistry. Despite these differences, all of the enzymes were capable of S1 turnover, which suggest that many terpene synthases could be useful for converting non-natural substrates into products. Furthermore, it may be possible to incorporate methyl vinyl ether groups into non-natural substrates which terpene synthases can be active upon, such as with Surrogate 2.

b. Round 1: Original Random Mutagenesis Library of 19B7

The reaction of Surrogate 2 with 19B7 was performed. This reaction was analyzed by gas chromatography, which indicated a wide range of products (Fig. 22). Therefore, by pursuing random mutagenesis, there was the potential to be able to evolve selectivity for one of these specific products.

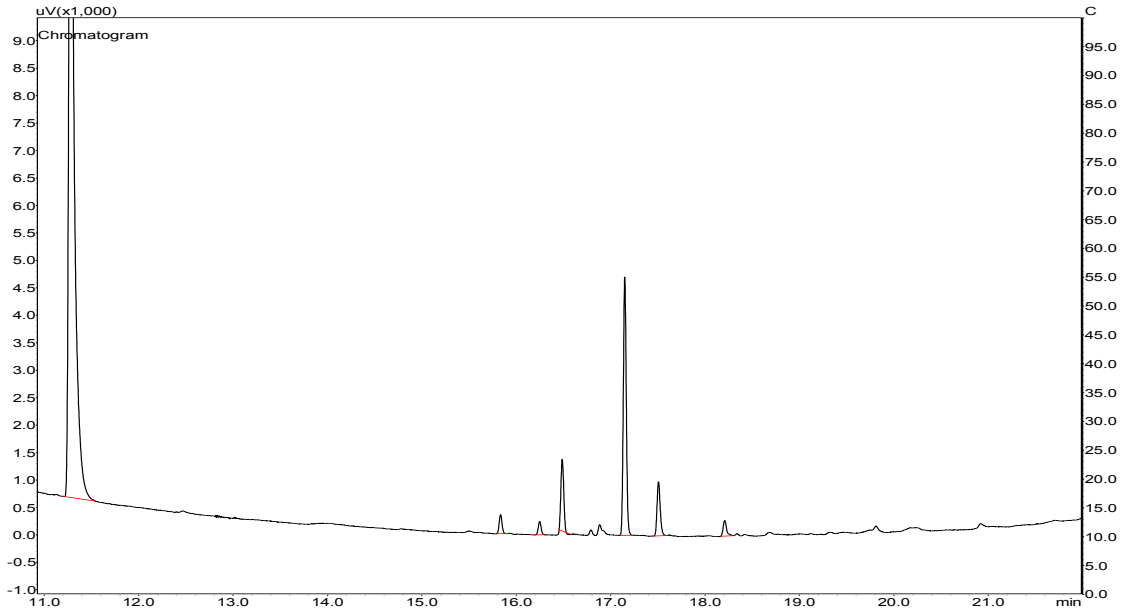


Figure 22. Gas chromatography results of 19B7 with Surrogate 2.

Within the Arnold laboratory, a library was generated through random mutagenesis of the thermostable parent terpene synthase variant 19B7. This mutant has a K85R mutation, and

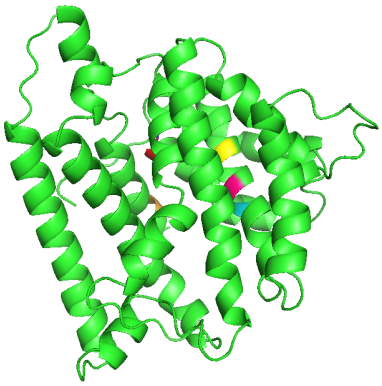
a silent mutation at base 93. Re-stamping was done to grow a library of enzymes that were screened using Surrogate 2, and with the AOX-Purpald assay. No improved mutants were found.

c. Round 2: New Random Mutagenesis Library of 19B7

A new random mutagenesis library was created, using 19B7 as the parent enzyme. The library consisted of seventeen 96-well plates. From this new library, twelve of the plates were chosen at random for screening, with 24 potential hits combined to make a compilation plate, and four as potential hits. However, when re-screened, the compilation plate indicated the results were only parent. With an error rate of 2.1 base pair mutations per gene, and a coefficient of variation of 17%, this lack of improved mutations is as expected. The error rate is low for a mutagenesis library, indicating that more of our colonies than desired are likely to be parent. With a high coefficient of variation, we see that there is a large range in which the colorimetric assay will indicate variability, while still being parent. Therefore, false hits could appear to have higher absorbency in the screen, but in fact have no improvements. The lack of improved mutants may have stemmed from the fact that in random mutagenesis, large libraries of several thousand mutants should be screened in order to find improved mutants.

d. Round 3: Directed Mutagenesis Library of 19B7

The approach was shifted from random mutagenesis to site directed mutagenesis. Rather than using random mutations with libraries of twenty to thirty plates, specific amino acid residues were mutated (usually in the active site region) to create libraries of 1-2 plates. This allowed me to screen the library while conserving more of the Surrogate 2 substrate. In literature, active site saturation mutagenesis has been a successful method of approach to evolve enzymes for non-natural substrates.



Mutation at:
W118: Red
A119: Blue
W133: Yellow
V137: Pink
F138: Light Blue
W367: Orange

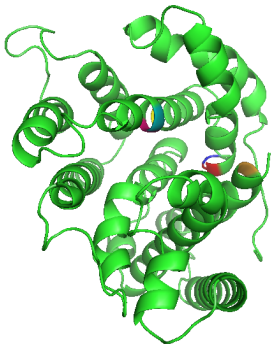


Figure 23. Homology model indicating NNK sites, model based on pentalenene synthase.

For the new libraries, six sets of primers with selected mutations in the active site were designed, which are referred to by the number of the amino acid to be mutated. The first letter indicates the amino acid to be changed, and the number is the amino acid location. These primers were designed to create an NNK library, where the amino acid in question was mutated to all other possible amino acids, and it was from these various primers that my new 19B7 libraries were prepared. The residues chosen for saturation mutagenesis were W118, A119, W133, V137, F138, and W367. These locations can be seen in the model in Figure 23.

Work initially started with the three W libraries, the W118NNK, W133NNK, and W367NNK, with a fragment assembly PCR. The potential hits were combined into a 96 well plate referred to as Comp 1 (Fig. 23), with each column corresponding to one variant. The entire procedure for the first three sets of primers was exactly repeated for the second three sets, being the A119NNK, the V137NNK, and the F138NNK.

Comp 1 AOX-Purpald Screen

	Blank	Parent	W367F	W118V	W118V	Blank	W118Q	W133A	W133A	W367F	Blank	Blank
A	0.1794	0.1455	0.1592	0.1395	0.1681	0.1309	0.1964	0.2163	0.1453	0.1535	0.1162	0.1134
B	0.1186	0.1847	0.2143	0.2049	0.1229	0.1332	0.2042	0.1609	0.1508	0.2093	0.1202	0.1229
C	0.1309	0.1724	0.1832	0.1719	0.1451	0.1291	0.2237	0.1709	0.1448	0.1886	0.1012	0.1166
D	0.1704	0.261	0.2403	0.1449	0.1237	0.1146	0.1961	0.1638	0.1485	0.2213	0.1047	0.1202
E	0.2788	0.2988	0.256	0.1946	0.2096	0.1567	0.2642	0.2007	0.147	0.2506	0.1405	0.1404
F	0.1869	0.2856	0.2491	0.2477	0.2058	0.1797	0.2982	0.1708	0.1837	0.2144	0.148	0.1514
G	0.1677	0.2719	0.2492	0.2443	0.1714	0.163	0.2875	0.1842	0.177	0.2116	0.1545	0.1751
H	0.1782	0.2984	0.3166	0.1872	0.2273	0.1822	0.2993	0.2587	0.2268	0.2663	0.1728	0.1643

Figure 24. Compilation plate from W118NNK, W133NNK, and W367NNK. Top three rows with DMSO, bottom five with normal screening. Boxes indicate potential hits from the NNK libraries from the AOX-Purpald screen.

Note that the top three rows (A, B, and C) were screened with dimethyl sulfoxide, DMSO (Fig. 24). In the reaction between Surrogate 2 and the terpene synthase, the rate-limiting step is believed to be the release of product from the active site. The hypothesis was that treatment with DMSO could surround the product with DMSO in order to help the product release from the enzyme complex. However, the DMSO did not prove effective with Surrogate 2.

There were three clones which displayed activity comparable to the parent. These clones proved to be two single mutants, 19B7 W367F and 19B7 W118Q. Mutating tryptophan (W) to phenylalanine (F) in 19B7 W367F is not expected to destroy the functionality of the enzyme, because it is switching from one aromatic amino acid to another aromatic amino acid. This mutation could in fact be helpful in increasing the size of the active site, because

phenylalanine is smaller than tryptophan. The other mutant, 19b7 W118Q, involves a change from an aromatic amino acid to a polar amino acid. This could be advantageous because having a more polar amino acid in the active site could help with releasing the product from the active site by allowing water into the active site to dislodge the product. In order to further characterize these mutants, both mutants were grown and expressed for protein purification. 19B7 W367F was successfully purified, and GC-scale reactions were performed with purified enzyme, to test for improved selectivity on the products of Surrogate 2.

Having only neutral mutations with respect to activity on Surrogate 2, the NNK libraries did not deliver the hoped-for increase in product generation. This fact may be due to insufficient accuracy in modeling the amino acids that comprised the active site. The nearest homolog to BcBOT2 is only about 23% identical at the amino acid level, meaning that the active site model may be inaccurate, and the proposed NNK sites could be imperfectly selected. Using X-ray crystallography to obtain an accurate model of the enzyme would allow for better determination of the active site residues, and improved selection of NNK mutation sites.

With directed evolution, it is important to consider the path of evolution that may be required to change the functionality of BcBOT2. In order to switch the activity of BcBOT2 from FPP to S2, several mutations may be required. If the number of mutations required is high, then it may be necessary to implement an incremental approach to the evolution, where BcBOT2 is first evolved for high activity on Surrogate 1. When the enzyme is evolved to have a better K_m for the vinyl methyl ether Surrogate 1, this enzyme could be easier to evolve for Surrogate 2. These intermediate variants may require neutral mutations in order to reach a significantly improved enzyme.

Chapter 4

CONCLUSIONS

Three natural terpene synthases, Cop2, BcBOT2, and SSCG_02150, were tested for activity against the natural substrate and a non-natural substrate, where the relative activities on both the natural and non-natural substrates were compared. BcBOT2 proved to be the fastest, with Cop2 following next, and SSCG_02150 producing compatible product amounts after 24 hours. All of the natural terpene synthase enzymes were more active on FPP than the Surrogate 1 compound, but the activity between the two was highly correlated.

With directed evolution for increased activity and selectivity on the unnatural substrate Surrogate 2, mutagenesis libraries have not been able to significantly improve enzyme activity. Initial random mutagenesis and site-specific saturation mutagenesis libraries produced two neutral mutants, 19B7 W367F and 19B7 W118Q, which were found by the screen to maintain activity on Surrogate 2. It is possible that the non-polar product of the Surrogate 2 reaction is difficult to separate from the enzyme, and therefore it may be difficult to evolve terpene synthases for this application. Furthermore, the site saturation mutagenesis towards the non-natural compound Surrogate 2 may have suffered from a lack of an accurate homology model, and the understanding of the active site configuration. Developing a better homology model by obtaining a successful crystal structure could improve the choice of amino acids for NNK libraries, and therefore produce mutations for all non-natural substrates of interest.

REFERENCES

1. Agger, S., Lopez-Gallego, F., & Schmidt-Dannert, C. (2009). Diversity of sesquiterpene synthases in the basidiomycete Coprinus cinereus. *Molecular Microbiology*, *72*(5), 1181-1195.

2. Arnold, F. H. (1998). Design by Directed Evolution. *Accounts of Chemical Research*, *31*(3), 125-131.

3. Cascón, O., Touchet, S., Miller, D. J., Gonzalez, V., Faraldos, J. A., & Allemann, R. K. (2012). Chemoenzymatic preparation of germacrene analogues. *Chemical Communications*, *48*, 9702-9704.

4. Chou, H. H., & Keasling, J. D. (2012). Synthetic Pathway for Production of Five-Carbon Alcohols from Isopentenyl Diphosphate. *Applied and Environmental Microbiology*, *78*(22), 7849-7855.

5. Diaz, J. E., *et al* (2011). Computational design and selections for an engineered thermostable terpene synthase. *Protein Science*, *20*, 1597-1606.

6. Gonzales-Vigil, E., Hufnagel, D. E., Kim, J., Last, R. L., & Barry, C. S. (2012). Evolution of TPS20-related terpene synthases influences chemical diversity in the glandular trichomes of the wild tomato relative Solanum habrochaites. *The Plant Journal*, *71*, 921-935.

7. Greenhagen, B. T., O'Maille, P. E., Noel, J. P., & Chappell, J. (2006). Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *PNAS*, *103*(26), 9826-9831.

8. Hopps, H. B. (2000). Purpald®: A Reagent That Turns Aldehydes Purple! *Aldrichimica ACTA*, *33*(1), 28-30.

9. Hu, Y., Chou, W. K.W., Hopson, R., & Cane, D. E. (2011). Genome Mining in Streptomyces clavuligerus: Expression and Biochemical Characterization of Two New Cryptic Sesquiterpene Synthases. *The Cell: Chemistry & Biology*, *18*, 32-37.

10. Lin, X., Hopson, R., & Cane, D. E. (2006). Genome Mining in Streptomyces coelicolor: Molecular Cloning and Characterization of a New Sesquiterpene Synthase. *Journal of the American Chemical Society*, *128*(18), 6022-6023.

11. Lopez-Gallego, F., Agger, S. A., Abate-Pella, D., Distefano, M. D., & Schmidt-Dannert, C. (2010). Sesquiterpene Synthases Cop4 and Cop6 from Coprinus

cinereus: Catalytic Promiscuity and Cyclization of Farnesyl Pyrophosphate Geometric Isomers. *ChemBioChem*, *11*, 1093-1106.

12. Lopez-Gallego, F., Wawrzyn, G. T., & Schmidt-Dannert, C. (2010). Selectivity of Fungal Sesquiterpene Synthases: Role of the Active Site's H-1alpha Loop in Catalysis. *Applied and Environmental Microbiology*, *76*(23), 7723-7733.

13. Miller, D. J., & Allemann, R. K. (2011). Sesquiterpene synthases: Passive catalysts or active players? *Natural Product Reports*, *29*(60), 60-71.

14. O'Maille, P. E., Chappell, J., & Noel, J. P. (2006). Biosynthetic potential of sesquiterpene synthases: Alternative products of tobacco 5-epi-aristolochene synthase. *Archives of Biochemistry and Biophysics*, *448*(1-2), 73-82.

15. Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B., & Keasling, J. D. (2012). Microbial engineering for the production of advanced biofuels. *Nature*, *486*, 320-328.

16. Pinedo, C., *et al* (2008). Sesquiterpene Synthase from the Botrydial Biosynthetic Gene Cluster of the Phytopathogen Botrytis cinerea. *ACS Chemical Biology*, *3*(12), 791-801.

17. Trapp, S. C., & Croteau, R. B. (2001). Genomic Organization of Plant Terpene Synthases and Moelcular Evolutionary Implications. *Genetics*, *158*, 811-832.

18. Wawrzyn, G. T., Bloch, S. E., & Schmidt-Dannert, C. (2012). Discovery and Characterization of Terpenoid Biosynthetic Pathways of Fungi. *Methods in Enzymology*, *515*, 83-105.

19. Weng, J.-K., Philippe, R. N., & Noel, J. P. (2012). The Rise of Chemodiversity in Plants. *Science*, *336*, 1667-1670.

20. Withers, S. T., & Keasling, J. D. (2007). Biosynthesis and engineering of isoprenoid small molecules. *Applied Microbiology and Biotechnnology*, *73*, 980-990.

21. Yoshikuni, Y., Martin, V. J.J., Ferrin, T. E., & Keasling, J. D. (2006). Engineering Cotton (+)-δ-Cadinene Synthase to an Altered Function: Germacrene D-4-ol Synthase. *Chemistry & Biology*, *13*, 91-98.