RICE UNIVERSITY

Directed Evolution and Structure-Function Studies of Oxidosqualene Cyclases

by

Michael Joseph Robert Segura

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Seiichi P.T. Matsuda, Associate Professor
Department of Chemistry
Department of Biochemistry and Cell Biology

Ronald J. Parry, Professor
Department of Chemistry
Department of Biochemistry and Cell Biology

Yousif Shamee, Assistant Professor
Department of Biochemistry and Cell Biology

Houston, Texas
December 2002
December 2002

ABSTRACT

Directed Evolution and Structure-Function
Studies of Oxidosqualene Cyclases

by

Michael J. R. Segura

Although oxidosqualene cyclase reaction mechanisms have been established, how these enzymes promote specific transformations, generate product diversity, and achieve product specificity is only partially understood. Protein mutagenesis experiments described herein have identified catalytically important residues involved in deprotonation and cyclization steps of oxidosqualene cyclase catalysis.

Residues influencing specific deprotonation in lanosterol and cycloartenol synthase were identified. The Saccharomyces lanosterol synthase Thr384Tyr mutation compromised product specificity and caused the formation of the novel lanosta-24-ene-3β,9α-diol, a compound previously unidentified in nature. Furthermore, the Arabidopsis thaliana cycloartenol synthase His477 position was identified by random mutagenesis. Interestingly, subtle amino acid changes at this position show dramatic and different influences on product structure; the His477Asn mutant makes predominantly lanosterol, whereas the His477Gln mutant makes mostly parkeol. The AthCAS1 His477Asn and His477Gln mutants are currently the most accurate lanosterol and parkeol synthases, respectively, generated by protein mutagenesis. AthCAS1 His477 mutations were combined with other catalytically important mutations (AthCAS1 Tyr410Thr Ile481Val)
to see how the combined mutations would interact to influence product structure. Surprisingly, the His477 mutations did not influence catalysis when combined with the other mutations. The catalytic behavior of these *AtHCAS1* triple mutants is the first demonstration of dominant and recessive properties of catalytically important oxidosqualene cyclase mutations.

DNA shuffling of cycloartenol and lupeol synthase was undertaken to determine what residues or motifs control substrate folding, a reaction step that has important consequences for product structure. Hybrid enzymes isolated after one round of DNA shuffling revealed that the N- and C-terminal ends (115 and 140 a.a., respectively) of cycloartenol synthase do not contain residues that are required for protosteryl cation formation or cyclopropyl ring formation. Those hybrids possessing lupeol synthase sequence at the termini but retaining internal cycloartenol synthase sequence maintained cycloartenol biosynthetic ability, demonstrating that catalytic components required for cycloartenol biosynthesis are located in the internal portions of the enzyme sequence. Site-directed mutagenesis and chimeragenesis experiments further defined catalytically important regions indicated by DNA shuffling experiments.

In addition, this thesis describes the complete characterization of the *Arabidopsis* lupeol synthase and the construction of a novel yeast strain that possesses a plant sterol biosynthetic pathway.
ACKNOWLEDGEMENTS

I would like to thank my research adviser, Dr. Seiichi P.T. Matsuda, for providing me with a thoroughly fulfilling graduate experience. He was patient enough to let me chase many different ideas and was always more than willing to provide much needed instruction and guidance. I would like to also thank Dr. Ronald Parry and Dr. Yousif Shamoo for serving on my thesis committee and for their valuable comments regarding this thesis. In addition, I am indebted to Dr. William Wilson, Mr. Jihai Pang, and Dr. Bonnie Bartel for their expert advice.

Throughout my graduate studies, I have had the pleasure of working with many talented colleagues. I am grateful to both current and former lab members who have helped me, both scientifically and personally. I would like to extend a special word of thanks to Renee LeClair and Silvia Lodeiro for critical comments regarding this thesis.

I could not have done any of this without the love and support of my family. My family has shaped the person I am and they have always encouraged me. Thank you cannot adequately express my heartfelt gratitude to my wife, Casey. She is the reason for any success I may have had and is a terrific wife, mother, and friend. She has suffered through my highs and lows and has never failed to support and encourage me. My daughter, Sydney, has also added incredible joy to my life. I am sure Sydney and her soon-to-arrive little brother will continue to be my pride and joy throughout all of my life.

Finally, with much humility, I would like to thank God for His providence, without which, none of this would have been possible.
To my wife Casey
# Table of Contents

**Abstract**

ii

**Acknowledgements**

iv

**Dedication**

v

**List of Figures and Tables**

ix

## Chapter 1: Introduction and Background

1

- Overview
  1
- Cycloartenol Synthase
  6
- Lanosterol Synthase
  14
- Lupeol Synthase
  19
- Squalene-Hopene Cyclase
  24
- Mutagenesis Summary
  26

## Chapter 2: Materials and Methods

29

- Instrumentation
  29
- Equipment
  30
- Materials
  31
  - Enzymes
    31
  - Bacterial Media
    31
  - Yeast Media
    32
- Protocols
  33
  - DNA Plasmid Mini-Prep
    33
  - Preparative DNA Plasmid Prep
    33
PREPARATION OF PLASMID DNA FOR SEQUENCING 34
DNA SEQUENCING 35
DNA RESTRICTION DIGEST 35
DNA LIGATION 36
_E. coli_ TRANSFORMATION 37
SINGLE-STRANDED DNA PREPARATION FOR SITE-DIRECTED MUTAGENESIS 37

OLIGONUCLEOTIDE SYNTHESIS 38
SITE-DIRECTED MUTAGENESIS 39
YEAST TRANSFORMATION 40
PREPARATION OF 20 × (±)-OXIDOSQUALENE SOLUTIONS 40
SMALL SCALE OXIDOSQUALENE CYCLASE ASSAY 41
LARGE SCALE OXIDOSQUALENE CYCLASE INCUBATION 42
PURIFICATION OF TRITERPENE ALCOHOLS BY SILICA GEL CHROMATOGRAPHY 43

DERIVATIZATION OF TRITERPENE ALCOHOLS WITH TRIMETHYLSILANE (TMS) 43

CHAPTER 3: EXPERIMENTAL PROCEDURES 44

STRUCTURE DETERMINATION OF LANOST-24-ENE-3β,9α-DIOL 44
_Ath_CAS1 His477 SITE-DIRECTED MUTAGENESIS 45
_Ath_CAS1 His477 MUTANT ASSAYS 45
PRODUCT CHARACTERIZATION OF _Ath_CAS1 His477_ MUTANTS 46
CONSTRUCTION AND VALIDATION OF YEAST STRAIN MSY4 47
PURIFICATION OF _AthLUP1_ TRITERPENE ALCOHOLS BY ARGENTATION-HPLC 48

STRUCTURE DETERMINATION OF _AthLUP1_ PRODUCTS 48
_AthLUP1_ PRODUCT RATIO DETERMINATION 50
SYNTHESIS OF THE _Olea europaea_ LUPEOL SYNTHASE 51
DNASE I DIGEST FOR DNA SHUFFLING 53
PURIFICATION OF DNASE I DIGESTED FRAGMENTS 54
DNA FRAGMENT REASSEMBLY FOR DNA SHUFFLING 54
DNA SHUFFLE PCR AMPLIFICATION
DNA SHUFFLE LIBRARY SUB-CLONING
DNA SHUFFLE LIBRARY CONSTRUCTION AND PROPAGATION
DNA SHUFFLED LIBRARY SCREENING
CHARACTERIZATION OF CAS:LUP SHUFFLED OXIDOSQUALENE CYCLASES
CONSTRUCTION OF HYBRID OXIDOSQUALENE CYCLASES
ATHCAS1 ASN727 AND ILE732 MUTANTS

CHAPTER 4: RESULTS AND DISCUSSION

PART 1. DIFFERENTIAL DEPROTONATION

CHARACTERIZATION OF SCEERG7 THR384TYR AND RELATED MUTANTS
IDENTIFICATION OF THE ATHCAS1 HIS477ASN MUTANT
CHARACTERIZATION OF ATHCAS1 HIS477ASN AND HIS477GLN MUTANTS
CHARACTERIZATION OF ATHCAS1 HIS477 RELATED TRIPLE MUTANTS

PART 2. PROTOSTERYL VERSUS DAMMARENYL CATION FORMATION

CONSTRUCTION OF PLANT STEROL BIOSYNTHETIC PATHWAY IN YEAST
CHARACTERIZATION OF THE ARABIDOPSIS THALIANA LUPEOL SYNTHASE
DNA SHUFFLING OF CYCLOARTENOL SYNTHASE AND LUPEOL SYNTHASE
SEQUENCE SWAPPING OF ATHCAS1 AND OEU/LUP C-TERMINAL SEQUENCES
CHARACTERIZATION OF ATHCAS1 ASN727 AND ILE732 MUTANTS

CONCLUSIONS

REFERENCES

APPENDIX I: LIST OF ABBREVIATIONS
APPENDIX II: LIST OF STRAINS
LIST OF FIGURES AND TABLES

Figure 1.1. Substrate folding has stereochemical consequences 4
Figure 1.2. Cycloartenol and lanosterol formation 6
Figure 1.3. Conservation pattern of AthCAS1 Tyr410 9
Figure 1.4. Products formed by cycloartenol synthase mutants 10
Figure 1.5. Conservation pattern of AthCAS1 Ile481 11
Table 1.1. Product profiles of A. thaliana cycloartenol synthase Ile481 mutants 12
Figure 1.6. Reaction catalyzed by lanosterol synthase 14
Figure 1.7. Proposed model for epoxide ring opening 16
Figure 1.8. Cyclization of oxidosqualene analogs suggests the intermediacy of a 6-6-5 tricyclic intermediate en route to tetracyclic product formation 17
Figure 1.9. Reaction catalyzed by lupeol synthase 20
Figure 1.10. Products formed by lupeol and β-amyrin synthase mutants 22
Figure 1.11. Reaction catalyzed by the Alicyclobacillus acidocaldarius squalene-hopene cyclase 25
Figure 4.1. Product profiles of SeeERG7 Thr384 mutants 63
Figure 4.2. Conservation pattern of AthCAS1 His477 67
Figure 4.3. Product profiles of AthCAS1 His477 mutants 70
Figure 4.4. Cycloartenol is not a suitable substrate for AthCPI1 77
Figure 4.5. Sterol biosynthetic pathway in yeast strain MSY4 79
Figure 4.6. Arabidopsis thaliana lupeol synthase generates six different triterpene products 81
Figure 4.7. DNA shuffling procedure 85
Table 4.1. Comparison of *AthCAS1* and *OeuLUP* used for DNA shuffling 86

Figure 4.8. Hybrid constructs recovered from DNA shuffling and selection experiments 89

Figure 4.9. Protein sequence alignment of cycloartenol and lupeol synthase 91

Figure 4.10. The *AthCAS1* C-terminal “Island” is required for cyclase activity 93

Figure 4.11. Protein sequence alignment of oxidosqualene cyclase sequences corresponding to the *AthCAS1* C-terminal island 95
CHAPTER 1: INTRODUCTION & BACKGROUND

This thesis describes directed evolution and structure-function studies of several oxidosqualene cyclases to further understand how they control catalysis, determine product structure, and achieve product specificity. Site-directed and random mutagenesis along with DNA shuffling experiments have been used to determine what specific portions of these proteins control various aspects of the catalyzed reaction such as substrate folding, cyclization, cation rearrangement, and cation-quenching steps. This chapter provides a brief overview of oxidosqualene cyclase catalysis. Included are several sections highlighting important studies that have advanced our understanding of these enzymes. Specifically, work describing the reaction mechanism and mutagenic studies of cycloartenol synthase, lanosterol synthase, and lupeol synthase will be discussed. This chapter then closes with a discussion of important, but unanswered, questions that have been the focus of my research projects.

OVERVIEW

Oxidosqualene cyclases are a unique family of eukaryotic enzymes that catalyze the cation-initiated polyolefinic cyclization of oxidosqualene to a wide variety of structurally complex triterpene compounds. The structural diversity found among oxidosqualene cyclase products—over 100 identified in nature—is generated by the enzymes’ ability to promote different cyclization reactions, various rearrangements of cation intermediates, and alternative termination steps.
These enzymes catalyze what are possibly the most complex single-step chemical transformations found in nature. As many as sixteen bonds are broken and sixteen new bonds are formed in the course of some cyclizations.\(^2\) In addition, these enzymes display remarkable regio- and stereospecificity during the reaction. In some cases, ten different stereocenters are established in the product molecule.\(^3,^4\) Furthermore, the chemical transformations catalyzed by oxidosqualene cyclases proceed through highly reactive cationic intermediates. The short-lived cation intermediates are protected and stabilized within the active site cavity of the oxidosqualene cyclases. Despite being surrounded by numerous nucleophiles, both in the amino acid side-chains and in the protein backbone, these cations are not quenched prematurely.

It is important to note here that many biomimetic studies modeling these reactions have been performed and several excellent reviews detailing those studies are available.\(^5,^6\) These elegant, non-enzymatic transformations have established some of the chemical principles that underlie the polyene cyclizations catalyzed by the oxidosqualene cyclases and have demonstrated that some of the complexity of the cyclization reaction is determined by the chemical properties of the substrate itself (concertedness in the ring forming steps, stereospecificity at the ring fusions, etc.). The biomimetic studies have also shown that purely chemical methods can form some polycyclic products in good yield and with stereospecific control. However, successful chemical cyclizations rely on substrate modifications to control the reaction and do no specifically answer how the oxidosqualene cyclases direct the accurate cyclization of oxidosqualene to diverse triterpene compounds. Consequently, direct examination of the cyclases is required to determine how they mediate oxidosqualene cyclization and determine product structure.
The oxidosqualene cyclases and their products are not only interesting because of their elaborate chemistry, but they are also implicated in a wide variety of biological functions. The oxidosqualene cyclase product lanosterol is a required sterol precursor found in both mammals and fungi. In humans, lanosterol is ultimately converted to cholesterol, steroid hormones, and other sterol compounds. Cycloartenol, an isomer of lanosterol, serves a similar role in plants\textsuperscript{7} and some protists.\textsuperscript{8-12}

The biological function of most triterpenes remains unknown; this is particularly true of plant-derived oxidosqualene cyclase products. Higher plants are prolific triterpene producers and are responsible for most of the triterpene structural diversity found in nature. Increasing evidence indicates that many triterpenes may serve as defense compounds to fend off invaders such as insects and fungi.\textsuperscript{13-15} Whatever roles they may play in the native organisms, many triterpenes display valuable biological activities including (but not limited to) anti-fungal, antibiotic, insecticidal, anti-cancer, and other medicinal activities. Thus, the oxidosqualene cyclases and their derived products are important research targets from both chemical and biological perspectives.

The oxidosqualene cyclases can be grouped into two broad sub-families, the protosteryl-type cyclases and the dammarenyl-type cyclases. Protosteryl-type cyclases, such as lanosterol and cycloartenol synthase, share the protosteryl cation (Figure 1.1) as a common intermediate. This intermediate is formed via cyclization of oxidosqualene that is folded into a pre-chair-boat-chair conformation (sterol folding). All of the dammarenyl-type cyclases, such as lupeol and \( \beta \)-amyrin synthase, form the dammarenyl cation. The dammarenyl cation arises from oxidosqualene cyclization when the same substrate is folded into the energetically favored pre-chair-chair-chair conformation (non-
sterol folding). Protosteryl-type cyclases and their derived products are typically involved in primary metabolism of sterol biosynthesis (i.e. lanosterol and cycloartenol biosynthesis). In contrast, the dammarenyl-type cyclases and their derived products are most often involved only in secondary metabolism. Furthermore, the dammarenyl-type cyclases are usually found only in plants and some fungi, organisms with active secondary metabolic pathways. Mammals have only one oxidosqualene cyclase (lanosterol synthase) and it is part of primary metabolism.

![Chemical structures](image)

**Figure 1.1. Substrate folding has stereochemical consequences.** Cyclization of oxidosqualene folded in a pre-chair-boat-chair conformation (sterol folding) yields the protosteryl cation intermediate. When folded in a pre-chair-chair-chair conformation, oxidosqualene cyclization leads to the dammarenyl cation intermediate. These folding differences result in opposite stereochemistry at every chiral center from the B-C-ring junction outward.

A complete understanding of how oxidosqualene cyclases control these extended chemical transformations and achieve product specificity remains unclear. The specific structural features of these enzymes that impart the catalytic distinction between the protosteryl-type cyclases and the dammarenyl-type cyclases remain unknown.
Determining how some enzymes fold oxidosqualene into one conformation whereas others fold the same substrate into another conformation would provide significant insights into triterpene biosynthesis.

Furthermore, important distinctions exist between cyclases belonging to the same sub-family. Cycloartenol synthase and lanosterol synthase, both protosteryl-type cyclases, promote different deprotonation steps. It is still only partially understood which amino acid residues control those specific deprotonations. Similarly, lupeol and β-amyrin synthase, both dammarenyl-type cyclases, promote different rearrangement of cationic intermediates leading to products with different carbocyclic skeletons. Identifying the catalytic differences between enzymes within the same sub-family is also an important goal in the understanding of oxidosqualene cyclase catalysis.

My investigations have revolved around several protein mutagenesis approaches designed to advance our current knowledge of oxidosqualene cyclases catalysis. These efforts were aided significantly by important studies that laid much of the groundwork in the oxidosqualene cyclase field. The following sections will highlight key advances regarding the catalytic differences between the protosteryl-type cyclases, cycloartenol synthase and lanosterol synthase. Some critical mechanistic studies of the lanosterol synthase will also be included. Finally, studies directed towards the dammarenyl-type cyclases will be discussed. Although less attention has been given to the enzymatic reactions catalyzed by the dammarenyl-type cyclases, several important studies involving lupeol and β-amyrin synthase are available.
CYCLOARTENOL SYNTHASE

Cycloartenol synthase\(^{16}\) is a protosteryl-type oxidosqualene cyclase that initially cyclizes oxidosqualene to the protosteryl cation intermediate and then mediates a series of hydride and methyl shifts to form the C-9 cation intermediate. In the final step of the reaction, cycloartenol synthase induces cyclopropyl ring formation and abstracts a proton from C-19 to form cycloartenol (Figure 1.2). This reaction is very similar to that catalyzed by lanosterol synthase. All steps in cycloartenol and lanosterol biosynthesis are identical with the exception of the final deprotonation reaction. Cycloartenol synthases form the cyclopropane ring and abstract a proton from C-19, whereas lanosterol synthases remove a different proton and form lanosterol, the \(\Delta^8\) isomer of cycloartenol.

Figure 1.2. Cycloartenol and lanosterol formation. Cycloartenol and lanosterol arise by identical cyclization reactions up to the deprotonation step. Both compounds can be formed from the C-9 cation; proton abstraction from C-19 yields cycloartenol, but deprotonation at C-8 affords lanosterol.
Given the high degree of similarity in the reactions catalyzed by lanosterol and cycloartenol synthase, it seems plausible that a modest number of amino acid differences in the active site may be responsible for the difference between the two types of enzymes. Furthermore, because the cyclopropyl ring formation of cycloartenol biosynthesis is thermodynamically unfavorable relative to lanosterol formation, some of those amino acid differences are probably specifically required to induce cyclopropyl ring formation and to exclude the formation of more energetically favored products by the cycloartenol synthase. Substantial efforts have been devoted to determining which residues impart control over these deprotonation steps.

A crystal structure of an enzyme is typically the best guide for mutagenesis studies. The spatial arrangement of the amino acid residues in the active site frequently indicates which positions affect catalysis. Determination of the key components responsible for the catalytic differences between lanosterol and cycloartenol synthases has been hampered by the absence of a crystal structure. The oxidosqualene cyclases are membrane-bound proteins that have proven difficult to purify and crystallize; consequently, there are no three-dimensional structures available to guide mutagenesis studies. The crystal structure of the *Alicyclobacillus acidocaldarius* squalene-hopene cyclase\textsuperscript{17,18} is the only structural guide currently available. The squalene-hopene cyclases cyclize the oxidosqualene analog, squalene, using a mechanism similar to that of the oxidosqualene cyclases. However, there are significant differences between the squalene-hopene cyclases and the oxidosqualene cyclases (discussed below). Thus, while the squalene-hopene cyclase is the best available tool for studying oxidosqualene cyclases, it should not be used as an exact representation of oxidosqualene cyclases.
The advent of molecular biology has generated the tools to probe the enzymatic differences between the lanosterol and cycloartenol synthases, even in the absence of a crystal structure. Sequence alignments of characterized lanosterol and cycloartenol synthases have provided significant guidance in the identification of positions that may be responsible for the catalytic distinction between these two enzymes. When the protein sequences of known enzymes are aligned, conservation patterns (where a certain amino acid is specifically maintained at a specific position in a number of enzymes) become apparent. In some cases, certain residues are strictly conserved in all of the oxidosqualene cyclases, regardless of the enzymatic activity. These kinds of patterns suggest that those residues may play a role in catalysis but that their function is common to all cyclases and is not responsible for the different product formation. In other cases, strict but differential conservation patterns are revealed. In these instances, enzymes with one type of cyclase activity maintain a certain amino acid at a position whereas enzymes with a different cyclase activity specifically maintain a different amino acid at that same position. These types of strict, differential conservation patterns suggest those amino acids may play a role in the reaction and may also be partly responsible for the different enzymatic activities.

Using this sequence-analysis strategy, Herrera and co-workers identified four residues that were strictly but differentially conserved between known lanosterol and cycloartenol synthases. *Arabidopsis thaliana* cycloartenol synthase (*AthCAS1*) mutants Tyr410Thr, Gly488Ala, Phe717Trp, and Met731Ala (in which the *AthCAS1* position was changed to its lanosterol synthase counterpart) were generated by site-directed mutagenesis. The mutants were functionally characterized and only one mutant,
Tyr410Thr (Figure 1.3), converted oxidosqualene to compounds other than cycloartenol. The Tyr410Thr mutant made 65% lanosterol, 2% parkeol, and 33% 9β-lanosta-7,24-dien-3β-ol (Δ7-lanosterol). This single mutation abolished cycloartenol biosynthetic ability and also resulted in formation of the previously unidentified Δ7-lanosterol (Figure 1.4). Interestingly, the newly formed Δ7-lanosterol must arise via the C-8 cation, whereas parkeol and cycloartenol formation require a hydride shift from C-9 to C-8. Lanosterol formation can arise from either the C-8 or C-9 cation (Figure 1.4). The nature of the products formed by this mutant suggests that the Tyr410 residue may be involved in C-9 cation formation or facilitates the hydride shift from C-9 to C-8.

Figure 1.3. Conservation pattern of AthCAS1 Tyr410. Tyrosine is strictly conserved in the known cycloartenol synthases (CAS1) from *A. thaliana*¹⁶ (Ath), *Pisum sativum*²⁰ (Psa), *Panax ginseng*²¹ (Pgi), *Glycyrrhiza glabra*²² (Ggl), *Luffa cylindrica*²³ (Lcy), *Avena sativa*¹³ (Asa), and *Dictyostelium discoideum*²⁴ (Ddi). Threonine is conserved at the corresponding position in fungal and mammalian lanosterol synthases represented by *S. cerevisiae*²⁵,²⁶ (Sce), *Candida albicans*²⁷,²⁸ (Cal), *Cephalosporium caerulens*²⁹ (Cca), *Schizosaccharomyces pombe*³⁰ (Spo), *Rattus norvegicus*³¹,³² (Rno), and *Homo sapiens*³³,³⁴ (Hsa).
This work established that the Tyr410 residue is an essential component for cycloartenol synthase activity; mutating this position abolished all cycloartenol biosynthetic ability. The strong catalytic influence of the Tyr410Thr mutation and the conservation pattern of AthCAS1 position 410 also indicate that this position is specifically involved in cycloartenol formation. AthCAS1 Tyr410 is therefore probably one of several key positions that control the different deprotonation steps catalyzed by lanosterol and cycloartenol synthase.

Figure 1.4. Products formed by cycloartenol synthase mutants. Some AthCAS1 mutant products are derived exclusively from the C-8 cation (Δ7-lanosterol) and others arise from only the C-9 cation (cycloartenol and parkeol). Lanosterol can be formed from either the C-8 or the C-9 cation. Achilleol A and camilleol C are abortive cyclization products.
In addition to site-directed mutagenesis, random mutagenesis of cycloartenol synthase has also proven to be a valuable tool in the identification of residues that impart the catalytic distinction between cycloartenol and lanosterol synthase. Hart and co-workers screened libraries of randomly generated *AthCAS1* mutants for their ability to produce lanosterol. The *AthCAS1* Ile481Val mutant was identified in this manner. Cycloartenol synthases maintain an Ile at this position and a Val is conserved at the corresponding position in the lanosterol synthases (Figure 1.5). Functional characterization of the *AthCAS1* Ile481Val mutant demonstrated that it forms cycloartenol (55%), lanosterol (24%), and parkeol (21%). Because neither the isoleucine nor valine amino acid side chains have charged functional groups, they cannot be the deprotonating species. However, they may participate directly or indirectly by either positioning the base or the cationic intermediate(s).

![Conservation pattern of AthCAS1 Ile481](image)

**Figure 1.5. Conservation pattern of AthCAS1 Ile481.** Isoleucine is strictly conserved in the known cycloartenol synthases (CAS1) from *A. thaliana* (Ath), *Pisum sativum* (Psa), *Panax ginseng* (Pgi), *Glycyrrhiza glabra* (Ggl), *Luffa cylindrica* (Lcy), *Avena sativa* (Asa), and *Dictyostelium discoideum* (Ddi). Valine is conserved at the corresponding position in trypanosomal, fungal, and mammalian lanosterol synthases represented by *Trypanosoma brucei* (Tbr), *Trypanosoma cruzi* (Tcr), *S. cerevisiae* (Sce), *Candida albicans* (Cal), *Cephalosporium caerulens* (Cca), *Schizosaccharomyces pombe* (Spo), *Rattus norvegicus* (Rno), and *Homo sapiens* (Hsa).
To investigate this possibility, Matsuda and co-workers made a series of Ile481 mutations (Ile481Leu, Ile481Gly, Ile481Ala, Ile481Phe) to determine how steric bulk at that position affects catalysis. The product profiles of each of the mutants are shown in the table below (Table 1.1). The isoleucine and leucine amino acid side chains are the same size but have different shapes. Interestingly, introducing this small change by substituting the native isoleucine residue with leucine compromises product specificity. Along with cycloartenol (83%), the AthCAS1 Ile481Leu mutant makes parkeol (16%) and a trace of lanosterol (1%). Similarly, substitution of Ile481 with smaller amino acids (such as Ala and Gly) leads to formation of a mixture of cycloartenol, lanosterol, and parkeol. In addition, the Ile481Ala and Ile481Gly mutations also cause the formation of the monocyclic products achilleol A and camilleol C. The Ile481Phe mutant is inactive, possibly resulting from the phenylalanine amino acid side chain protruding into the active site cavity.

<table>
<thead>
<tr>
<th>AthCAS1 Mutant</th>
<th>lanosterol</th>
<th>cycloartenol</th>
<th>parkeol</th>
<th>achilleol A</th>
<th>camilleol C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile481Leu</td>
<td>1%</td>
<td>83%</td>
<td>16%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ile481Ala</td>
<td>54%</td>
<td>12%</td>
<td>15%</td>
<td>13%</td>
<td>6%</td>
</tr>
<tr>
<td>Ile481Gly</td>
<td>23%</td>
<td>17%</td>
<td>4%</td>
<td>44%</td>
<td>12%</td>
</tr>
<tr>
<td>Ile481Val</td>
<td>24%</td>
<td>55%</td>
<td>21%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1.1. Product profiles of *A. thaliana* cycloartenol synthase Ile481 mutants. A subtle shape change in the Ile481Leu mutant results in broadened product specificity. Substitution of Ile481 with the smaller alanine and glycine residues also results in miscyclization (achilleol A and camilleol C formation) in addition to broadened product specificity.
This series of Ile481 mutants demonstrated that steric changes in the active site can also have an important impact on catalysis and product specificity. Mutating \textit{Ath}CAS1 Ile481 leads to lanosterol formation, a result consistent with the conservation pattern of lanosterol and cycloartenol synthases at this position. In addition to a role in deprotonation, Ile481 may also be involved in proper substrate folding and cyclization; mutating Ile481 to smaller residues (Ala and Gly) led to monocyclization products.

The cycloartenol synthase mutagenesis studies described above identify enzyme residues that play key roles in the deprotonation step required for cycloartenol formation. Mutation of Tyr410 results in loss of product specificity and abolishes all cycloartenol synthase activity. Mutating that position also results in product formation from the C-8 cation in addition to the native C-9 cation intermediate. The series of Ile481 mutations demonstrate that steric changes at this position have important effects on product specificity as well; Ile481 mutations resulted in multiple deprotonation products. Furthermore, some Ile481 mutations (Ile481Ala and Ile481Gly) also caused the formation of abortive cyclization products, suggesting another role for Ile481 in the cyclization step of the reaction. These important mutagenesis studies revealed key features of the cycloartenol synthase and its distinctive deprotonation machinery. However, a complete understanding of the catalytic differences between the cycloartenol and lanosterol synthases cannot be achieved without similarly investigating the lanosterol synthase.
LANOSTEROL SYNTHASE

Lanosterol synthase, found in mammals, fungi, and trypanosomes, converts oxidosqualene to lanosterol. Proton-initiated epoxide ring opening and cyclization lead to the protosteryl cation, from which a series of hydride and methyl shifts occur to form the final cation intermediate. Lanosterol formation can arise by abstracting a proton from either a C-8 or C-9 cationic intermediate (Figure 1.6). Most of the pioneering mechanistic work detailing these chemical transformations was done using the lanosterol synthase system. Several key mechanistic insights were established by substrate analog and enzyme mutagenesis studies and those studies will be described briefly. Lanosterol synthase mutagenesis studies corresponding to the cycloartenol synthase work (described above) will then be discussed.

![Chemical diagram of lanosterol synthase reaction]

**Figure 1.6. Reaction catalyzed by lanosterol synthase.** Lanosterol synthase cyclizes oxidosqualene to the protosteryl cation and then promotes a series of hydride and methyl shifts to form either a C-8 or C-9 cationic intermediate. Deprotonation of either intermediate can result in lanosterol formation.

Although epoxides are somewhat acid labile, no natural amino acid is acidic enough to open the epoxide ring of oxidosqualene. Anchimeric assistance from the proximate double bond is required for epoxide ring opening and A-ring formation in the
oxidosqualene cyclases.\textsuperscript{39,40,41} This was established in part by a series of mechanistic investigations in which substrate analogs possessing various functional groups (Cl or H rather than CH\textsubscript{3}) at the C-6 carbon of oxidosqualene were incubated with the \textit{Saccharomyces cerevisiae} lanosterol synthase (SceERG7). The rate of cyclization decreased as the nucleophilicity of the proximate double bond decreased,\textsuperscript{41} demonstrating that the olefin does indeed participate in epoxide ring opening.

Enzyme mutagenesis also revealed important aspects of the epoxide-opening/initiation step. Sequence analysis of known oxidosqualene cyclases reveals a strongly conserved DCTA motif. Because this conservation does not correlate with a specific cyclization activity, the motif is likely involved in a catalytic step common to all oxidosqualene cyclase reactions. Several mutagenesis studies showed that altering this motif typically abolished activity,\textsuperscript{42,43} leading some groups to postulate that it was responsible for stabilization of cationic intermediates. Corey and co-workers proposed a different role for this oxidosqualene cyclase motif. From a series of mutagenesis experiments, they determined that mutation of the Asp456 and also His146 abolished activity. In addition, protein-labeling studies that employed suicide-substrate inhibitors implicated the catalytic involvement of Asp456 and His146 residues. Based on these results, Corey \textit{et al.} proposed a model in which the acidic aspartate residue (Asp456 using SceERG7 numbering) of the DCTA motif was the proton donor that initiated epoxide ring opening and subsequent cyclization.\textsuperscript{39} The conserved histidine (His146 using SceERG7 numbering) was proposed to enhance the acidity of Asp456. The model is shown below (Figure 1.7).
The two studies described above were fundamental advances in a mechanistic understanding of oxidosqualene cyclases. They established that both a "push" (the anchimeric assistance from the proximate double bond) and a "pull" (an acidic proton donor generated by the Asp456 and His146 pair) were required to initiate enzymatic oxidosqualene cyclization.

![Diagram of oxidosqualene (sterol folding)](image.png)

**Figure 1.7. Proposed model for epoxide ring opening.** Corey and co-workers proposed a model in which the acidic aspartate residue (Asp456 using SceERG7 numbering) is the proton donor that initiates epoxide ring opening and cyclization. The histidine residue (His146 using SceERG7 numbering) enhances the acidity of Asp456.

Additional substrate analog experiments suggest that formation of the 6-6-6-5 ring nucleus of the final tetracyclic product first proceeds through a 6-6-5 tricyclic intermediate. Oxidosqualene analogs possessing non-natural substituents at C-15 led to abortive 6-6-5 cyclization products when incubated with pig liver oxidosqualene cyclase\(^44\) (Figure 1.8). These results show that Markovnikov addition is favored, presumably because it forms the more stable tertiary cation intermediate. Other analogs
with chemical modifications of the C-18 double bond also suggest that the five-membered C-ring expansion is concerted with D-ring formation. Incubation of the 18,19-dihydrooxidosqualene with rat liver oxidosqualene cyclase\textsuperscript{45} again led to an abortive 6-6-5 tricyclic product (Figure 1.8). Without concurrent formation of the five-membered D-ring, there was no C-ring expansion to form the six-membered C-ring system.

![Figure 1.8. Cyclization of oxidosqualene analogs suggests the intermediacy of a 6-6-5 tricyclic intermediate en route to tetracyclic product formation.](image)

In addition to substrate analog experiments, protein mutagenesis studies of lanosterol synthase have also provided important discoveries regarding how the enzyme operates. Conservation patterns among known lanosterol and cycloartenol synthases and random mutagenesis studies have led to the identification of catalytic residues in the *Arabidopsis thaliana* cycloartenol synthase (discussed above) and those catalytic
positions in the *AthCAS1* background are probably also catalytically relevant in the lanosterol synthase background. Functional characterization of the complementary mutations in lanosterol synthase has provided further insight into how the lanosterol and cycloartenol synthases control their distinct deprotonation steps.

In the *AthCAS1*, the Ile481Val mutation was shown to influence catalysis. That position corresponds to Val454 in the yeast lanosterol synthase. Isoleucine is conserved at this position in all known cycloartenol synthases and a valine at the corresponding position in all known lanosterol synthases (Figure 1.5). Joubert and co-workers therefore investigated the catalytic effects of *SceERG7* Val454 mutations. They created and characterized the *SceERG7* Val454Leu, Val454Ile, Val454Ala, Val454Gly, and Val454Phe mutants. Interestingly, only modest catalytic perturbations were observed in these mutants. The Val454Leu and Val454Ile mutants did not influence product structure and made only lanosterol. Mutating to a smaller residue in the Val454Ala and Val454Gly mutants resulted in a small amount of achilleol A formation (5% and 17%, respectively) in addition to lanosterol biosynthesis. The Val454Phe mutant was inactive, perhaps by protrusion into the active site. In all of these mutants, the aliphatic valine was substituted with other non-polar residues. Thus, the mutations probably altered only the steric environment of the area surrounding the mutation. The formation of achilleol A in some of the mutants suggests that Val454 may be involved in proper substrate folding. The catalytic influence of *SceERG7* Val454 mutations is consistent with the behavior of corresponding mutations in the cycloartenol synthase; however, the catalytic changes in these lanosterol synthase mutants are much less pronounced than those observed in the cycloartenol synthase. Similar to some of the corresponding *AthCAS1* Ile481 mutations,
some of these SceERG7 Val454 mutations influence the cyclization step (mono- vs. tetracyclization). However, unlike the corresponding AthCAS1 mutations, the SceERG7 Val454 mutations do not appear to influence the deprotonation step once the cyclization has been completed.

**LUPEOL SYNTHASE**

Lupeol is an oxidosqualene cyclase product formed in many higher plants, and the lupeol synthase genes from several different plant organisms have been cloned. Although several basic features of lupeol biosynthesis resemble those in lanosterol and cycloartenol biosynthesis, lupeol and many other triterpenes result from a ring-forming mechanism distinct from that employed in lanosterol and cycloartenol formation. Identification of the catalytic differences between enzymes such as lupeol synthase and lanosterol and cycloartenol synthase is an important step in understanding ring formation in oxidosqualene cyclase catalysis.

Lupeol synthase is a member of the dammarenyl-type oxidosqualene cyclases. These enzymes fold oxidosqualene into a pre-chair-chair-chair conformation prior to cyclization whereas the protosteryl-type cyclases (i.e., lanosterol and cycloartenol synthase) fold the same substrate into a pre-chair-boat-chair conformation. This may appear to be a subtle difference but it results in the opposite stereochemistry at every chiral center from the B-C ring junction and beyond. Moreover, the stereochemical differences between the products of protosteryl- and dammarenyl-type cyclases are also useful phylogenetic markers. Triterpenes derived from the protosteryl cation (sterol folding) are found in all types of eukaryotic species and are usually involved in primary
sterol metabolism. Those triterpenes derived from the dammarenyl cation (non-sterol folding) are found almost exclusively in plants and are typically secondary metabolites.

In addition to the substrate folding differences and resultant stereochemical changes caused by the dammarenyl-type cyclases, the oxidosqualene cyclases in this sub-family typically promote additional cyclization steps prior to any hydride and methyl shifts. Thus, the triterpene products of the dammarenyl-type cyclases are often pentacyclic compounds, whereas those products derived from the protosteryl-type cyclases are commonly tetracyclic. Lupeol synthase catalyzes D-ring expansion of the dammarenyl cation and concurrent five membered E-ring formation, followed by proton elimination. This results in the 6-6-6-6-5 pentacyclic product (Figure 1.9). D-ring expansion and E-ring formation probably take place in the same manner as C-ring expansion and D-ring formation as discussed above in relation to lanosterol synthase and substrate analog experiments.

![Figure 1.9](image)

**Figure 1.9. Reaction catalyzed by lupeol synthase.** Lupeol synthase first cyclizes oxidosqualene to the dammarenyl cation. The D-ring expansion, E-ring formation, and deprotonation then yield lupeol.
Clearly, some elements of the oxidosqualene cyclization reaction sequence are shared in common between protosteryl- and dammarenyl-type oxidosqualene cyclases. Both mediate epoxide ring opening followed by a cyclization cascade to form a 6-6-6-5 cation intermediate. In addition, the reaction terminating steps (deprotonation or addition of water) are similar in both. Nevertheless, it is equally clear from the skeletal and stereochemical differences in the products that there must be significant active site differences between these two families of enzymes. Dammarenyl-type cyclases fold the substrate into a pre-chair B-ring conformation, whereas the protosteryl cyclases fold the same substrate into the energetically unfavorable pre-boat B-ring conformation. Also, from the tetracyclic protosteryl cation intermediate, the protosteryl cyclases typically promote hydride and methyl shifts to place the cation towards the A-ring in the ring nucleus. From a similar intermediate (dammarenyl cation), the dammarenyl-type cyclases, such as lupeol synthase, instead promote further cyclization and extend the carbocyclic skeleton to an E-ring. The active-site differences between the protosteryl- and dammarenyl-type cyclases responsible for controlling these different reaction modes remain unknown.

To date, few mutagenesis studies have been done to investigate lupeol synthase or any other dammarenyl-type cyclases. The existing studies have investigated the catalytic differences between lupeol synthase and β-amyrin synthase. β-Amirin synthase, identified in several different plants, is another dammarenyl-type cyclase that essentially catalyzes an extended version of the lupeol synthase reaction. From the lupyl cation (a 6-6-6-6-5 intermediate common to both lupeol and β-amyrin synthesis), the β-amyrin synthase promotes another ring expansion step resulting in a 6-6-6-6-6 pentacyclic cation
intermediate that then undergoes a series of hydride shifts and deprotonation step to form β-amyrrin (Figure 1.10).

![Chemical structures and reactions]

**Figure 1.10. Products formed by lupeol and β-amyrrin synthase mutants.** A variety of tetracyclic and pentacyclic triterpenes, all derived from the dammarenyl cation intermediate, are formed by lupeol and β-amyrrin synthase mutants.

Kushiro *et al.* made a series of chimeras of lupeol and β-amyrrin synthases in an effort to determine what portions of the genes control product specificity. They found that only relatively small portions of the respective genes were responsible for controlling product specificity. For instance, they described a lupeol synthase/β-amyrrin synthase chimera with only 25% β-amyrrin sequence that made four times as much β-amyrrin as it did lupeol. In that same work, they used a mixed PCR method to generate chimeras and
independently confirmed the importance of that region and further narrowed it down to an 80 amino acid span. It is interesting to note that the product control elements of these two enzymes appear to cluster together in a rather narrow region rather than being spread across the entire length of the protein.

In further work, Kushiro and co-workers went on to more closely investigate the small sections of sequence identified in their chimera experiments. Sequence analysis along with the chimera experiments suggested that Trp259 of the *Panax ginseng* β-amyrin synthase (PNY) and Leu256 of the *Olea europaea* lupeol synthase (OEW) might play a part in the catalytic difference between the two enzymes. They therefore made and characterized the PNY (β-amyrin synthase) Trp259Leu mutation and showed that it converted oxidosqualene to lupeol as its major product instead of β-amyrin. The corresponding mutation was made in the lupeol synthase background. The OEW (lupeol synthase) Leu256Trp mutant also showed inverted product specificity; β-amyrin (75%) was identified as the major product (Figure 1.10) along with some lupeol (7%), germanicol (8%), and butyrosperrmol (10%). These two mutants demonstrate that this position plays a role in the E-ring expansion step. The authors suggest that the aromatic tryptophan residue stabilizes the oleanyl cation formed in β-amyrin biosynthesis and that if absent (as is the case where leucine is at that position in the lupeol synthase), the reaction is terminated with lupeol formation.

The same authors also investigated a position that was strictly conserved in all dammarenyl-type cyclases (Tyr261 and Tyr258 in PNY and OEW, respectively). All known protosteryl-type cyclases possess a His residue at the corresponding position. Product analysis of the PNY (β-amyrin synthase) Tyr261His mutant revealed several
dammarradienol isomers as the only products (Figure 1.10). Relative to β-amyrin formation, these are abortive cyclization products. The mutation eliminated D-ring expansion and subsequent E-ring formation.

SQUALENE-HOPENE CYCLASE

The squalene-hopene cyclases (SHC) catalyze a mechanistically similar reaction to those found among the oxidosqualene cyclases. These enzymes are found primarily among prokaryotes and accept the oxidosqualene analog, squalene, as their native substrate. Like the oxidosqualene cyclases, the squalene-hopene cyclases mediate a cation initiated polyolefinic cyclization to form an intermediate cation that is typically quenched either by deprotonation or addition of water. The squalene-hopene cyclase from Alicyclobacillus acidocaldarius cyclizes squalene to the 6-6-6-6-5 hopyl cation, which is then either deprotonated to form hopene or quenched with water to form hopanol (Figure 1.11). Although the reactions catalyzed by the squalene-hopene cyclase and oxidosqualene cyclases are similar, the squalene-hopene cyclase catalyzed reactions are less sophisticated as they typically do not promote hydride and methyl shifts to generate multiple cationic intermediates.

Like the oxidosqualene cyclases, the squalene-hopene cyclase has been well studied and many experiments parallel those found in the oxidosqualene cyclase literature. Important substrate analog experiments and protein mutagenesis studies have described squalene-hopene cyclase activity and, by analogy, have also advanced the understanding of oxidosqualene cyclase catalysis in many cases. Those experiments detailing the squalene-hopene cyclase reaction mechanism and protein structure-function
relationships are beyond the scope of this work and will not be discussed here, but the reader is encouraged to read several excellent reviews\textsuperscript{50-52} concerning the catalytic mechanism of the squalene-hopene cyclases.

![Diagram of squalene, hopanyl cation, hopene, and hopanol]

**Figure 1.11. Reaction catalyzed by the *Alicyclobacillus acidocaldarius* squalene-hopene cyclase.** The bacterial squalene-hopene cyclase cyclizes the oxidosqualene analog, squalene, to generate the pentacyclic hopene and hopanol compounds. Unlike many oxidosqualene cyclases, the squalene-hopene cyclase does not promote any hydride or methyl shifts.

The *Alicyclobacillus acidocaldarius* squalene-hopene cyclase crystal structure\textsuperscript{17,18} has also been a useful guide for oxidosqualene cyclase investigations. The three-dimensional structure of the squalene-hopene cyclase reveals the placement of specific amino acids in relation to one another and, more importantly, shows the positions of active site residues. That information in turn indicates what amino acid residues may be directly participating in the reaction and what roles they may play. Because no oxidosqualene cyclase structure has yet been resolved, the squalene-hopene cyclase
crystal structure is currently the only suitable three-dimensional structure to guide oxidosqualene cyclase mutagenesis experiments.

As was mentioned earlier, the squalene-hopene cyclase crystal structure must be used with some degree of caution. Even though the reactions catalyzed by squalene-hopene cyclase and oxidosqualene cyclases are similar, there are probably also inevitable structural differences between the two types of enzymes. The bacterial squalene-hopene cyclases and eukaryotic oxidosqualene cyclases are only 19-25% identical at the amino acid level and the squalene-hopene cyclases are about 100 amino acids shorter than oxidosqualene cyclases. As a result, the squalene-hopene cyclase crystal structure should not be used as an exact representation of the oxidosqualene cyclases and models derived from that structure are only sound when corroborated with experimental evidence.

**MUTAGENESIS SUMMARY**

Significant advances have been made in understanding oxidosqualene cyclase activity, as highlighted in some of the studies described above. Site-directed and random mutagenesis approaches have led to the identification of key catalytic residues. Furthermore, structure-function relationships have helped establish what roles certain positions might play in catalysis. Novel cyclase activities and products have also resulted from mutagenesis of oxidosqualene cyclases.

Despite these advances, significant questions still remain. Many catalytically important single mutations in several different oxidosqualene cyclase backgrounds are now known. In every case reported, the product specificity of the mutant enzyme has been broadened. No mutant enzyme described to date has been altered to accurately
generate a product distinct from the native product. With this in mind, it remains unknown how the oxidosqualene cyclases not only promote certain reactions, but also how they exclude competitive reaction pathways. Similarly, it is unknown exactly how different active-site environments must be to generate accurate and distinct chemical transformations in the oxidosqualene cyclases. Are several amino acid changes sufficient to accurately convert a lanosterol synthase into a cycloartenol synthase, or are global changes required? It should be noted that this is probably the simplest case to investigate. Lanosterol and cycloartenol synthase perform identical cyclization and rearrangement reactions and differ only in the deprotonation step. Despite this apparent simplicity, mutant lanosterol or cycloartenol synthases have not yet been developed that accurately control product specificity. Accurate product inversion would likely prove even more difficult between two oxidosqualene cyclases with widely different activities (such as cycloartenol synthase and β-amyrin synthase).

My research efforts have been directed toward answering these kinds of questions and advancing our understanding of the oxidosqualene cyclases. I have employed a variety of protein mutagenesis studies to establish which portions of various cyclases are catalytically important and responsible for product specificity. The identification and characterization of those product control elements then creates the opportunity for rational protein design. If an adequate understanding of oxidosqualene cyclase catalysis exists, new cyclase activities can be engineering in a predictive manner. This type of protein engineering would then allow the creation of novel cyclases for which a natural counterpart is either unknown or unavailable. Furthermore, those designed cyclases
could perhaps be useful biocatalysts providing access to novel, valuable, or inaccessible compounds.

Research efforts investigating oxidosqualene cyclase catalysis are described herein. Chapter 2 contains a list of materials and methods used in this research. Chapter 3 describes specific experimental details. Finally, chapter 4 includes experimental results and a discussion of their significance.
CHAPTER 2: MATERIALS & METHODS

INSTRUMENTATION

Nuclear magnetic resonance (NMR) spectra were collected on a Bruker Avance 500 NMR spectrometer. Proton \(^1\)H NMR spectra were collected at 500 MHz and all shifts are reported in ppm referenced to internal trimethylsilane (TMS). Carbon \(^{13}\)C NMR spectra were collected at 125 MHz and all chemical shifts are reported in ppm referenced to CDCl\(_3\) at 77.0 ppm. Unless otherwise noted, all spectra were collected at 25\(^o\)C. Deuterated chloroform (CDCl\(_3\)) was filtered through K\(_2\)CO\(_3\) prior to use and all samples were prepared in 5 mm glass tubes (Wilmad Glass Co. or Shigemi Inc.)

Gas chromatography (GC) was performed using a Hewlett-Packard 6890 instrument (equipped with a 30 m, 0.25 mm ID, 0.10 \(\mu\)m film thickness, Rtx-5 column from Restek). Unless otherwise noted, the TRITERP2 separation method was used. This method employed the following conditions: inlet and FID-detector, 290\(^o\)C; oven, 280\(^o\)C isothermal for 30 min; split ratio 40:1, column flow (helium) 0.6 mL/min at constant pressure; helium makeup flow.

GC-MS data was collected on an instrument composed of a Hewlett-Packard 5890 GC (equipped with a 30 m, 0.25 mm ID, 0.10 \(\mu\)m film thickness DB-5MS column from J&W Scientific) and a VG-ZAB-HF mass spectrometer (50-500 m/z range, 70 eV electron impact (EI) ionization energy, accelerating voltage of 8 kV). The nature of the sample determined what GC conditions were used.
High-performance liquid chromatography (HPLC) separations were performed using a Hewlett-Packard G1312A binary pump, a Hewlett-Packard G1322A solvent degasser, and a Hewlett-Packard G1314A UV detector. Normal phase separations were carried out using mobile phases composed of hexane and methyl t-butyl ether (MTBE) solvents (EM Scientific). Either a silver-ion column (Alltech Nucleosil SA 5U, 250 mm × 10 mm I.D., washed repeatedly with silver nitrate) or a semi-prep normal phase column (YMC Co., Ltd., YMC-Pac-Sil, 250 mm × 10 mm I.D., S-10P μm, 120 Å) was used for triterpene alcohol separations.

**EQUIPMENT**

**PCR THERMOCYCLER**

Polymerase chain reactions (PCR) were performed using an Ericomp (San Diego, CA) Easy Cycler Series PowerBlock II System.

**CENTRIFUGATION**

Centrifugations of 1.5 mL tubes were performed in variable speed Eppendorf Centrifuge Models 5415C, 5415D, or 5810R. Centrifugations (3-250 mL) were performed using a Beckman Model TJ-6 centrifuge, an Eppendorf Model 5810R centrifuge (variable temperature), or a Sorvall Model RC-5B centrifuge (variable temperature).
INCUBATORS

*E. coli* and *S. cerevisiae* plate cultures were grown in Fisher Scientific Isotemp incubators. Liquid cultures were grown in either a New Brunswick Series 25 Incubator Shaker or a New Brunswick G24 Environmental Shaker.

CELL LYSIS

Yeast cells were lysed using either a French pressure cell press (SLM Instruments) or an Emulsiflex-C5 homogenizer (Avestin).

MATERIALS

ENZYMES

All restriction enzymes, T4 DNA ligase, T4 DNA polymerase, T4 DNA kinase, and M13KO7 helper phage were purchased from New England BioLabs (Beverly, MA). The Fast-Link DNA Ligase kit was purchased from Epicentre Technologies (Madison, WI). ExTaq DNA polymerase was purchased from Panvera Company (Madison, WI) and PfuTurbo DNA polymerase was purchased from Stratagene (La Jolla, CA).

BACTERIAL MEDIA

*E. coli* cultures were grown in Luria broth (LB) which was prepared by dissolving 10 g of LB powder (5 g tryptone, 2.5 g yeast extract, and 2.5 g NaCl) in 500 mL dH₂O. The LB media was sterilized by autoclaving at 121° C for 45 min. Solid LB media was prepared as described above except with the addition of 7.5 g agar prior to sterilization. Media supplementation with the antibiotic ampicillin (100 µg/mL) ensured the bacteria maintained the plasmid construct.
YEAST MEDIA

Yeast growth requires both a nitrogen and carbon source and yeast media contained equal volumes of each. Nitrogen sources were either YP (2 × YP: 5 g yeast extract and 10 g peptone in 500 mL dH₂O) or synthetic complete (2 × SC: 1.7 g yeast nitrogen base, 5.0 g ammonium sulfate, 2.0 g amino acid mix, and 2.0 mL of 1M NaOH to adjust final pH to ~5.5). The amino acid mix used for synthetic complete media was composed of 10 g leucine and 2.0 g of each of the following: alanine, arginine, asparagine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine, adenine, and uracil. Selective SC media, such as SC-Ura, SC-Leu, SC-His, and SC-Trp, was made using an amino acid mix (described above) lacking the specified amino acid or nucleotide. Carbon sources were either dextrose (2 × Dex: 20 g dextrose in 500 mL dH₂O) or galactose (2 × Gal: 20 g galactose in 500 mL dH₂O). Galactose was used when gene expression under the GAL promoter was required.

Agar (7.5 g per 500 mL) was added prior to sterilization to make the corresponding solid yeast media. All media was sterilized prior to use. Dextrose, galactose, and YP solutions were autoclaved at 121° C for 20-45 min. Synthetic complete media was autoclaved at 121° C for 20 min.

When required (as determined by the yeast strain), media was supplemented with 1.3 mg/mL hemin (100 × Heme: 65 mg hemin hydrochloride, 25 mL absolute EtOH, 25 mL dH₂O, and 0.75 mL of 1 M NaOH) and 2.0 mg/mL Tween 80-ergosterol solution (100 × Erg: 20 mg ergosterol, 5 mL absolute EtOH, and 5 mL Tween 80).
PROTOCOLS

DNA PLASMID MINI-PREP

Bacterial cultures (1-3 mL) were grown overnight in selective LB media. Cells were pelleted by centrifugation and the supernatants discarded. The cells were then resuspended in 200 μL P1 buffer plus RNase A (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 20 mg/mL RNase A). To the suspensions were added 200 μL P2 lysis buffer (200 mM NaOH, 1% SDS (w/v)) and the samples were then inverted about 5 times to mix. These solutions were then allowed to sit at room temperature for 5 min. The lysis was terminated by addition of 200 μL of chilled P3 buffer (3.1 M potassium acetate, pH 5.5, ~ 4° C). Samples were mixed by inversion and chilled on ice for about 15 min. The cellular debris was pelleted by centrifugation and the DNA containing supernatants were transferred to new 1.5 mL tubes. To these samples was added 2 volumes of absolute ethanol (EtOH). They were then mixed by vortexing and stored in the freezer for at least 15 min. DNA was pelleted by centrifugation and the supernatants discarded. To remove residual salt, the DNA pellets were washed with ~500 μL of 70% EtOH and then allowed to air dry. The pellets were then redissolved in 50-100 μL of T8 buffer (10 mM Tris-HCl, pH 8.0) and stored indefinitely at ~20° C.

PREPARATIVE DNA PLASMID PREP

A bacterial culture (50-100 mL) was grown in selective LB media overnight. The cells were pelleted by centrifugation in a 50 mL tube and the supernatant discarded. The resultant pellet was resuspended in 2 mL P1 buffer with RNase A. P2 buffer (3 mL) was
then added, mixed by inversion, and the solution allowed to sit at room temperature for 5 min. The cell lysis step was neutralized by addition of P3 buffer (3 mL) followed by mixing and chilling on ice. The cellular debris was removed by filtration through a 10 cc disposable syringe plugged with Miracloths (Calbiochem). The filtered solution was collected in a new tube. Isopropanol (0.7 volumes) was added and the solution mixed with vortexing. The sample was then placed in the freezer for a minimum of 15 min to enhance the yield in the precipitation step. Aliquots (1.5 mL) were transferred to 1.7 mL centrifuge tubes and the DNA pelleted for 18 min at maximum speed in a microcentrifuge. The supernatants were discarded and the pellets washed with 1 mL 70% EtOH. The DNA pellets were air dried and then redissolved in 50-200 µL of T8 buffer. Identical samples were combined and stored at −20°C.

**Preparation of Plasmid DNA for Sequencing**

Plasmid DNA from either the mini- or preparative DNA preps is suitable for sequencing. To a DNA sample was added 5 M NaCl (to a final concentration of 100 mM) and two volumes of absolute EtOH. The sample was mixed by vortexing and placed at −20°C for 15-30 min. DNA was pelleted by centrifugation and the supernatant discarded. The pellet was washed with 1 mL of 70% EtOH and then allowed to air dry. The DNA pellet was then redissolved in an appropriate amount of T8 buffer and the concentration determined by gel electrophoresis.
DNA Sequencing

DNA was sequenced using an Applied Biosystems International sequencer and was performed by Lone Star Labs, Inc. (Houston, TX)

DNA Restriction Digest

To 100-500 ng plasmid DNA was added 1.5 μL of 10 × restriction digest buffer, 0.3 μL (~3 U) of restriction enzyme, and water to bring the total volume to 15 μL. The digest was incubated (at the temperature specified by the manufacturer for the specific enzyme) for a minimum of 1 h. To the digest was added 1.5 μL of 10 × gel loading buffer (20% Ficoll 400, 0.1 mM EDTA, pH 8.0, 0.25% bromphenol blue, and 0.25% xylene cyanol). The sample was then analyzed by agarose gel electrophoresis. 1% agarose gels were made by microwaving 5 g agarose in 500 mL 1 × TAE buffer (40 mM tris base, 20 mM acetic acid, 1 mM EDTA) followed by addition of ethidium bromide (5 μL/100 mL gel). This solution was poured into a mold and allowed to set up at room temperature. Gels were run in 1 × TAE buffer and samples analyzed against an appropriate DNA molecular weight marker (typically λ DNA digested with BstE II). The applied voltage varied from 20 to 110 V, depending on the size of the gel and the desired electrophoresis speed.

Preparative DNA digests were carried out on 5-20 μg of plasmid DNA. To the DNA was added 5 μL of 10 × restriction enzyme buffer, 2-5 μL restriction enzyme (~20-50 U), and Milli-Q H₂O to a final volume of 50 μL. The digest was incubated at the
optimal temperature designated for the restriction enzyme (typically 37° C) for a minimum of two h. When complete, 5 μL of 10 × gel loading buffer was added to terminate the reaction and prepare the sample for gel electrophoresis. The desired DNA fragment was purified using preparative agarose gel electrophoresis. Gels were made and run as described above except GTAE buffer (40 mM tris base, 20 mM acetic acid, 1 mM EDTA, 1 mM guanosine) was substituted for TAE in both the gel and the running buffer. GTAE buffer is supplemented with guanosine and helps protect the DNA from UV damage that could hinder later manipulations. The preparative gel was run until the desired DNA fragment was adequately resolved for purification. The target band was cut out of the gel with a razor blade and the DNA recovered using gel extraction procedures.

**DNA LIGATION**

DNA fragments with cohesive ends were ligated using the Fast-Link DNA Ligation Kit (Epicentre Technologies). To a 1.5 mL tube were added 1.5 μL of 10 × Fast-Link ligation buffer and 1.5 μL of 10 mM ATP. Vector DNA and insert DNA were then added in amounts to achieve a 2:1 molar ratio of insert to vector. Milli-Q H₂O was added to bring the volume to 14 μL. Fast-Link DNA ligase (1.0 μL) was added and the ligation allowed to incubate at room temperature for a minimum of 15 min (could also go for many hours). The ligation was heat killed at 70° C for 15 min to increase transformation efficiency. The ligation was either used immediately or stored in the freezer.
**E. coli Transformation**

Chemically competent DH5α (stored at -80° C) were thawed on ice. DNA was added to the cells, which were then incubated on ice for 20 min. The cells were then heat shocked at 37° C for 5 min. Aliquots of the transformation (typically 10 µL and 90 µL) were plated on selective LB agar plates.

**Single-stranded DNA Preparation for Site-Directed Mutagenesis**

The plasmid construct to be used for mutagenesis was transformed into the bacterial strain RZ1032 which allows uracil incorporation into DNA. A single colony was picked and used to inoculate 1 mL of 2XYT media supplemented with ampicillin (100 µg/mL). The culture was grown at 37° C for 3 h. To this was added helper phage M13K07 to a final concentration of ~ 5x10⁸ pfu/mL. It was then grown for an additional 1.5 h at 37° C. The 1 mL culture was then diluted into 20 mL of 2XYT supplemented with ampicillin (100 µg/mL) and kanamycin (50 µg/mL). This culture was then grown with shaking at 37° C for 12 to 15 h.

The cells were pelleted at 7000 rpm, 4° C for 10 min in a SS34 rotor and Sorvall centrifuge. The supernatant was transferred to a new tube and 20% PEG-8000, 2.5 M NaCl solution (5 mL) was added. The solution was mixed and incubated on ice for a minimum of 1 h. The phage particles were then pelleted as described above and supernatant discarded. The phage pellet was resuspended in 1.2 mL TE8 buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) and transferred to a clean 1.5 mL tube. This suspension was incubated on ice for a minimum of 5 min. It was then pelleted in a microcentrifuge
for 10 min to pellet any residual debris. The resultant supernatant was carefully transferred to a new 1.5 mL tube and 20% PEG-8000, 2.5 M NaCl solution (300 µL) was added and the samples were mixed. This was incubated on ice for a minimum of 15 min and the phage then harvested by centrifugation at maximum speed in a microcentrifuge for 10 min. The supernatant was discarded and the phage pellet resuspended in TE8 buffer (200 µL).

Phenol (200 µL) and 5 M NaCl (4 µL) were added to the phage suspension and mixed by vortexing. The aqueous (upper) phase was collected and transferred to new 1.5 mL tube and was then extracted again with CHCl₃ (200 µL). The aqueous (upper) layer was collected and transferred to a new 1.5 mL tube. Two volumes of absolute EtOH were added, vortexed, and stored at -20° C for a minimum of 15 min. The DNA was pelleted at maximum speed in a microcentrifuge for 15 min. The supernatant was discarded and the pellet allowed to briefly air dry at room temperature. The ssDNA pellet was then dissolved in T8 buffer (50 µL) and stored indefinitely at -20° C.

**OLIGONUCLEOTIDE SYNTHESIS**

Oligonucleotides were synthesized by Integrated DNA Technologies, Inc. or the Great American Gene Co. Oligonucleotide stock solutions were typically made by dilution with Milli-Q H₂O to a final concentration of 100 pmol/µL. These stock solutions were stored indefinitely at -20° C.
SITE-DIRECTED MUTAGENESIS

Oligonucleotide Phosphorylation: To a 1.5 mL tube was added the mutagenesis oligonucleotide (360 pmol), 10 × T4 DNA kinase buffer (2 μL), 10 mM ATP (1 μL), and T4 DNA kinase (0.5 μL), and dI H2O to a final volume of 20 μL. This was incubated at 37⁰ C for 1 h followed by addition of 0.5 M EDTA (0.5 μL). The kinase was then heat killed by incubation at 70⁰ C for 10 min. The solution was diluted with dI H2O (15 μL) and stored at -20⁰ C.

Annealing Oligo and Template: To a 1.5 μL tube was added ssDNA (1 μg), phosphorylated oligonucleotide (25 pmol), 20 × SSC (1 μL), and dI H2O to a final volume of 20 μL. The contents were mixed with vortexing and the tube was then placed in a 70⁰ C heat block. The block was then placed on the bench and allowed to cool gradually to < 35⁰ C. This primed DNA sample was then placed on ice.

Second Strand Synthesis: To the 20 μL primed DNA sample (above) was added dI H2O (49 μL), 10 × T4 DNA ligase buffer (10 μL), 2.5 mM dNTPs (20 μL), single-stranded DNA binding protein (1 μL), T4 DNA polymerase (1 μL), and T4 DNA ligase (0.5 μL). This solution was mixed, incubated on ice for 5 min, incubated at room temperature for 5 min, and then incubated at 37⁰ C for 1.5-2 h.

DNA Precipitation: 5 M NaCl (2 μL) and absolute EtOH (200 μL) were mixed with the second strand synthesis reaction. It was incubated at -20⁰ C for at least 15 min before the DNA was pelleted at maximum speed in a microcentrifuge. The supernatant was aspirated off and the pellet allowed to air dry at room temperature. The DNA was redissolved in dI H2O (10 μL) and stored at -20⁰ C.
**YEAST TRANSFORMATION**

A 5-10 mL yeast culture was grown to saturation at 30° C. The cells were then harvested by centrifugation, the supernatant discarded, and the pellet washed twice with sterile H₂O. The pellet was then resuspended in 100 μL H₂O with vortexing. Hot salmon sperm ssDNA solution (1 mg/mL) and plasmid DNA (~10 μg) were added to the suspension which was then vortexed to mix. To this was added sterile yeast transformation buffer (2 mL). The transformation buffer was a sterile aqueous solution of 40% polyethylene glycol (PEG) 3350, 0.1 M lithium acetate, 10 mM Tris buffer (pH 7.5), 1 mM EDTA, and 100 mM dithiothreitol (DTT). The cell suspension was vortexed again and left to incubate at room temperature for 6-15 h.

The cells were then pelleted and the supernatant discarded. The pellet was washed twice with sterile H₂O (1 mL each time) and resuspended in a final volume of 500 μL sterile H₂O. Aliquots (20 μL and 200 μL) were then plated on selective plates and spread with sterile glass beads.

**PREPARATION OF 20 × (±)-OXIDOSQUALENE SOLUTIONS**

Two oxidosqualene substrate solutions that include different detergents were typically used. To a round bottomed flask was added racemic oxidosqualene (200 mg) and Triton X-100 detergent (2 g). Methylene chloride (1 mL) was added to aid mixing and then removed under vacuum. This was repeated a second time to ensure that the oxidosqualene and detergent were thoroughly mixed. After the complete removal of methylene chloride, dl H₂O (8 mL) was added and the solution was stirred at room
temperature for about 1 h to ensure complete mixing. The solution (which when used as a 20 × stock, yields a reaction with 1% Triton X-100 and 1 mg/mL oxidosqualene) was stored at 4°C.

Another solution using a different detergent was prepared in a similar manner. This solution was prepared as described above except Tween 80 (200 mg) replaced the Triton X-100 detergent and more dI H₂O (9.8 mL) was added. When used as a 20 × stock, this solution yields a reaction with 0.1% Tween 80 and 1 mg/mL oxidosqualene.

**SMALL SCALE OXIDOSQUALENE CYCLASE ASSAY**

Selected yeast constructs were grown in appropriate expression media on 5-20 mL scale. The yeast cells were pelleted by centrifugation and the supernatant discarded. The pellet was resuspended in 100 mM sodium phosphate buffer (pH 6.2) to provide a 20-50% cell suspension. To this suspension was added 20 × oxidosqualene to a final concentration of 1 mg/mL. Negative control reactions were set up where no substrate was added to the cell suspension. Approximately 300 μL of acid-washed glass beads were added to the cell suspensions for cell lysis. The samples (with beads) were vortexed on high for 2.5 min and then placed on ice for 2.5 min. The cell suspensions were then vortexed again on high for 2.5 min. The assays were then left to incubate at room temperature and followed by TLC analysis.
Large Scale Oxidosqualene Cyclase Incubation

Yeast were cultured in inducing media (1 L) in sterile Erlenmeyer flasks (2 L). When cultures reached saturation, the cells were pelleted by centrifugation and the supernatant discarded. The cells were then resuspended in two volumes of 100 mM sodium phosphate buffer (pH 6.2). Cells were lysed by passing the suspension through a French press at a minimum of 10,000 psi. The suspension was passed through the apparatus several times to ensure complete lysis. A small aliquot was removed to serve as a negative control for the assay. To the rest of the suspension was added $20 \times$ oxidosqualene to a final concentration of 1 mg/mL. The solution was swirled to mix and left to incubate at room temperature. The reaction was monitored by TLC and quenched with two volumes of absolute ethanol when completed.

The cellular debris was pelleted by centrifugation and the ethanolic extract recovered. A portion of the crude extract was removed, concentrated to dryness, and stored indefinitely at $-20^\circ$ C. Most of the ethanol was removed from the remaining material by rotary evaporation. The aqueous sample was extracted three times with ethyl acetate (50 mL) and the organic extract was collected and dried down on to neutral alumina (typically 5 g). This material was loaded on to a silica gel plug (10 g) and washed with diethyl ether (100 mL). The material collected from the silica plug was then analyzed using GC, GC-MS, and NMR.
Purification of Triterpene Alcohols by Silica Gel Chromatography

Triterpene alcohol products were purified from crude reaction mixtures using 50 to 100 times as much silica gel (240-400 mesh) as sample material (e.g., 100 mg of material used 5 g of silica gel). Columns were wet-packed using the same solvent mixture to be used for the separation. Compounds were typically separated using 10-20% diethyl ether in hexanes. Fractions were collected in disposable test tubes and the separations monitored by TLC.

Derivatization of Triterpene Alcohols with Trimethylsilane (TMS)

For GC and GC-MS analysis, triterpene alcohol products were converted to their corresponding trimethylsilyl (TMS) ethers. To the sample was added a 1:1 mixture of pyridine and (bis-trimethylsilyl)-trifluoroacetamide (BSTFA). The sample was then sealed and left to sit at either room temperature or 37°C for several hours. Samples derivatized in this manner were stable for several months if properly sealed and stored at -20°C.
CHAPTER 3: EXPERIMENTAL PROCEDURES

STRUCTURE DETERMINATION OF LANOST-24-ENE-3β,9α-DIOL

The SceERG7 Thr384Tyr mutant converted oxidosqualene to a product mixture that contained two fractions that were separable by silica gel chromatography: one fraction that comigrated with lanosterol and a more polar component. The polar fraction was determined to be that of lanosta-24-ene-3β,9α-diol (9α-hydroxylanosterol) by $^1$H, $^{13}$C, DEPT, COSYDEC, HSQC, and HMBC NMR experiments. Structure determination and chemical shift assignments of the ring system were aided by comparisons with the characterized 24,25-dihydro analog (lanostane-3β,9α-diol). All samples were run in 5 mm 528-PP tubes (Wilmad Glass Co.). $^1$H chemical shifts were referenced to internal TMS standard and $^{13}$C shifts were referenced to CDCl$_3$ at 77.0 ppm. $^1$H NMR (500 MHz, CDCl$_3$, 25 °C): δ 0.786 (3H, s, H-18), 0.808 (3H, s, H-29), 0.910 (1H, m, H-21), 0.924 (3H, s, H-30), 0.996 (3H, s, H-28), 1.035 (3H, s, H-19), 1.602 (3H, s, H-27), 1.681 (3H, s, H-26), 3.201 (1H, m, H-3α). $^{13}$C NMR (125 MHz, CDCl$_3$, 25 °C): δ 14.59 (C-18), 15.40 (C-29), 16.73 (C-19), 17.62 (C-27), 18.27 (C-30), 18.58 (C-21), 21.42 (C-6), 23.74 (C-7), 24.85 (C-23), 25.71 (C-26), 27.36 (C-2), 27.95 (C-16), 27.95 (C-11), 28.28 (C-28), 29.13 (C-12), 29.52 (C-1), 33.82 (C-15), 35.80 (C-20), 36.39 (C-22), 38.80 (C-4), 40.54 (C-8), 42.64 (C-10), 45.18 (C-5), 45.69 (C-13), 47.45 (C-14), 50.36 (C-17), 77.17 (C-9), 78.64 (C-3), 125.16 (C-24), 130.95 (C-25). Product ratios of the SceERG7 Thr384Tyr,

* The mutation and enzymatic incubations were performed by Michelle M. Meyer, and isolation, structure determination, and quantitation were performed by M. J. R. S.
Thr384Tyr Val454Leu, and Thr384Tyr Val454Ile mutants were determined by $^1$H NMR quantitation using well-resolved signals.

**ATHCAS1 HIS477 SITE-DIRECTED MUTAGENESIS**

Site-directed mutagenesis, with the primer 5'-GCAGTCAGAGATGGGCAACCTTGATCGGAGAAAAGGCAAG-3' (Pst I restriction site removed), was employed to create the His477Gln mutation in the AthCAS1 gene encoded by the pSM60.21 plasmid (AthCAS1 in pRS305GAL). AthCAS1 triple mutants incorporating His477Asn and His477Gln mutations were created in a similar manner. The pSM60.21 Tyr410Thr His477Asn Ile481Val triple mutant was generated using the primer 5'-GCAGTCAGAGACGGGCAACCGTTATCGGAGAAAAGGCAACCAAG-3' (Pst I restriction site removed) to incorporate the His477Asn mutation in the AthCAS1 gene encoded by the pSM60.21 Tyr410Thr Ile481Val plasmid. The corresponding pSM60.21 Tyr410Thr His477Gln Ile481Val triple mutant was generated by using the 5'-GCAGTCAGAGACGGGCAACCGTTATCGGAGAAAAGGCAACCAAG-3' (Pst I restriction site removed) primer to insert the His477Gln mutation in the pSM60.21 Tyr410Thr Ile481Val plasmid. All mutants were sequenced to ensure that only the desired mutations had been incorporated.

**ATHCAS1 HIS477 MUTANT ASSAYS**

Four AthCAS1 His477 mutants (pSM60.21 His477Asn, pSM60.21 His477Gln, pSM60.21 Tyr410Thr His477Asn, Ile481Val, and pSM60.21 Tyr410Thr His477Gln Ile481Val) were linearized with BstE II and transformed into LHY4. Transformants
were selected on SCD-Leu,H,E and positive transformants were used to inoculate SCD-Leu,H,E liquid cultures (10 mL). All yeast cultures were grown at 30 °C and liquid cultures were shaken at 250 rpm. Upon saturation (typically 2 days), the 10 mL cultures were used to inoculate YPD,H,E media (100 mL). Large-scale induction cultures of YPG,H,E (1 L) were then inoculated with the 100 mL cultures after 1-2 days. Upon saturation of the large scale cultures (2-3 days), cells were harvested by centrifugation, resuspended in two volumes of 0.1 M sodium phosphate buffer (pH 6.2), and lysed at ~10,000 psi using a French Pressure Cell Press or Emulsiflex-C5 homogenizer. After the addition of 20 × racemic oxidosqualene solution (Tween 80 detergent), the reactions were incubated at room temperature for one to two days and monitored by TLC (silica gel, 1:1 hexane/ether).

**PRODUCT CHARACTERIZATION OFATHCAS1 His477 MUTANTS**

The large scale in vitro incubations were quenched with two volumes of absolute EtOH and the cellular debris was removed by centrifugation. Most of the EtOH was removed from the extract by rotary evaporation and the resultant aqueous suspension was extracted 3 times with EtOAc (50 mL each time). The organic extract was concentrated onto neutral alumina (5 g) and was then eluted over a silica plug (5 g) with diethyl ether (100 mL). The resultant partially purified fraction contains starting material, triterpene alcohols and triterpene diols, but lacks more polar compounds. Polycyclic triterpene alcohols are nearly impossible to resolve by silica gel chromatography and mixtures can therefore be purified from crude preparations without distorting the product ratios. Most oxidosqualene cyclase products can be identified by GC-MS and quantitated by
integrating well-resolved product signals (GC-FID) in the partially purified material. $^1$H NMR of partially-purified material is required to characterize 9β-lanosta-7,24-dien-3β-ol ($\Delta^7$-lanosterol), which GC does not readily resolve from lanosterol. Integration of $^1$H NMR peaks in the column purified material provided the lanosterol / $\Delta^7$-lanosterol ratios, which were then incorporated with product ratios derived from GC integration to provide the complete product profile of the mutants.

CONSTRUCTION AND VALIDATION OF YEAST STRAIN MSY4

The Arabidopsis thaliana cyclopropyl-sterol isomerase (AthCPI1) in the integrative yeast expression vector pRS305GAL (pML1.5) was linearized with BstE II and transformed into the yeast lanosterol synthase mutant SMY8. Transformants were selected on SCD-Leu,H,E and the new yeast construct was named MSY4.

Arabidopsis thaliana cycloartenol synthase (AthCAS1) in the high-copy yeast expression vector pRS426GAL (pSM60.22) was transformed into MSY4. The pSM60.22 plasmid was maintained on selective SCD-Ura,H,E media. MSY4[pSM60.22] was then plated on SCG-Ura,H and YPG,H media (no ergosterol supplementation) to induce expression and to test the ability of the AthCPI1 / AthCAS1 pair to genetically complement the lanosterol synthase mutation in MSY4. Viable MSY4[pSM60.22] colonies were visible in about 1 week. MSY4[pSM60.22] yeast cultures grown without exogenously supplied sterol were also extracted and ergosterol was identified (TLC, GC) in the extracts. This confirmed that MSY4 was able to ultimately convert cycloartenol to ergosterol in vivo.
PURIFICATION OF *Ath*LUP1 TRITERPENE ALCOHOLS BY ARGENTATION-HPLC

Michelle M. Meyer provided a mixture of acetylated *Ath*LUP1 triterpene alcohol products collected from silica gel chromatography. This mixture was subjected to argentation-HPLC using a silver-ion column (Alltech Nucleosil SA 5U, 250 mm × 10 mm I.D., washed repeatedly with silver nitrate). The following conditions were used: flow rate, 1.5 mL/min; detection, 210 nm; solvent, 20% (w/w) methyl tert-butyl ether (MTBE) in hexanes. Fractions (1.5 mL) were collected using an automated fraction collector. Two runs (5 mg material per run) were required to collect enough material for NMR characterization. Four distinct product fractions from each run were collected and combined.

STRUCTURE DETERMINATION OF *Ath*LUP1 PRODUCTS

The structures of five acetylated *Ath*LUP1 triterpene alcohols collected from argentation-HPLC were determined by \(^1\)H, \(^13\)C, DEPT, COSYDEC, HSQC, and HMBC NMR. A purified sample (free alcohol) of the polar component of the *Ath*LUP1 product mixture was provided by Michelle M. Meyer and its structure was determined in a similar manner. All samples were run in 5 mm CDCl\(_3\) Shigemi tubes. \(^1\)H chemical shifts were referenced to internal TMS standard and \(^13\)C shifts were referenced to CDCl\(_3\) at 77.0 ppm. In addition to previously identified lupeol and β-amyrin, *Ath*LUP1 also makes germanicol, taraxasterol, \(\psi\)-taraxasterol, and 3β,20-dihydroxy-lupane.

* The LUP1 product mixture was generated by Michelle M. Meyer, and purification, structure determination, and quantitation were performed by M. J. R. S.
**Germanicol:** $^1$H NMR (500 MHz, CDCl$_3$, 25° C): $\delta$ 0.733 (3H, s, H-27), 0.844 (3H, s, H-24), 0.849 (3H, s, H-23), 0.904 (3H, s, H-25), 0.937 (3H, s, H-29), 0.944 (3H, s, H-30), 1.017 (3H, s, H-28), 1.078 (3H, s, H-26), 2.046 (3H, s, acetate), 4.485 (1H, m, H-3α), 4.862 (1H, m, H-19). $^{13}$C NMR (125 MHz, CDCl$_3$, 25° C): $\delta$ 14.55 (C-27), 16.08 (C-26), 16.52 (C-24), 16.76 (C-25), 18.14 (C-6), 21.12 (C-11), 21.32 (acetate) 23.69 (C-2), 25.25 (C-28), 26.18 (C-12), 27.51 (C-15), 27.91 (C-23), 29.18 (C-30), 31.36 (C-29), 32.35 (C-20), 33.33 (C-21), 34.34 (C-17), 34.52 (C-7), 37.14 (C-10), 37.36 (C-22), 37.69 (C-16), 37.82 (C-4), 38.39 (C-13), 38.62 (C-1), 40.76 (C-8), 43.32 (C-14), 51.13 (C-9), 55.58 (C-5), 80.96 (C-3), 129.77 (C-19), 142.67 (C-18), 171.02 (acetate).

**Taraxasterol:** $^1$H NMR (500 MHz, CDCl$_3$, 25° C): $\delta$ 0.843 (3H, s, H-24), 0.850 (3H, s, H-23), 0.855 (3H, s, H-28), 0.877 (3H, s, H-25), 0.927 (3H, s, H-27), 1.021 (3H, s, H-26), 1.021 (3H, d, H-29), 2.046 (3H, s, acetate), 4.485 (1H, m, H-3α), 4.609 (2H, m, H-30). $^{13}$C NMR (125 MHz, CDCl$_3$, 25° C): $\delta$ 14.72 (C-27), 15.88 (C-26), 16.33 (C-25), 16.49 (C-24), 18.18 (C-6), 19.47 (C-28), 21.32 (acetate), 21.46 (C-11), 23.69 (C-2), 25.49 (C-29), 25.62 (C-21), 26.15 (C-12), 26.64 (C-15), 27.94 (C-23), 33.99 (C-7), 34.53 (C-17), 37.05 (C-10), 37.80 (C-4), 38.29 (C-16), 38.44 (C-1), 38.86 (C-22), 39.16 (C-13), 39.38 (C-19), 40.92 (C-8), 42.04 (C-14), 48.65 (C-18), 50.40 (C-9), 55.45 (C-5), 80.98 (C-3), 107.11 (C-30), 154.67 (C-20), 171.02 (acetate).

$\psi$-**Taraxasterol:** $^1$H NMR (500 MHz, CDCl$_3$, 25° C): $\delta$ 0.733 (3H, s, H-28), 0.844 (3H, s, H-24), 0.854 (3H, s, H-23), 0.879 (3H, s, H-25), 0.948 (3H, s, H-27), 0.989 (3H, d, H-29), 1.045 (3H, s, H-26), 1.635 (3H, s, H-30), 2.046 (3H, s, acetate) 4.485 (1H, m, H-3α),
5.261 (1H, d, H-21). $^{13}$C NMR (125 MHz, CDCl$_3$, 25$^\circ$ C): $\delta$ 14.69 (C-27), 16.04 (C-26), 16.35 (C-25), 16.52 (C-24), 17.70 (C-28), 18.18 (C-6), 21.32 (acetate), 21.61 (C-11), 21.62 (C-30), 22.53 (C-29), 23.69 (C-2), 27.02 (C-15), 27.59 (C-12), 27.94 (C-23), 34.16 (C-7), 34.38 (C-17), 36.32 (C-19), 36.69 (C-16), 37.01 (C-10), 37.79 (C-4), 38.44 (C-1), 39.20 (C-13), 41.08 (C-8), 42.17 (C-22), 42.33 (C-14), 48.68 (C-18), 50.33 (C-9), 55.38 (C-5), 80.99 (C-3), 118.88 (C-21), 139.86 (C-20), 171.02 (acetate).

3β,20-Dihydroxylupane: $^1$H NMR (500 MHz, CDCl$_3$, 25$^\circ$ C): $\delta$ 0.764 (3H, s, H-24), 0.810 (3H, s, H-28), 0.841 (3H, s, H-25), 0.957 (3H, s, H-27), 0.973 (3H, s, H-23), 1.059 (3H, s, H-26), 1.122 (3H, s, H-29/30), 1.225 (3H, s, H-29/30), 3.198 (1H, m, H-3α). $^{13}$C NMR (125 MHz, CDCl$_3$, 25$^\circ$ C): $\delta$ 14.83 (C-27), 15.38 (C-24), 16.15 (C-25), 16.15 (C-26), 18.33 (C-6), 19.21 (C-28), 21.38 (C-11), 24.76 (C-29/30), 27.38 (C-2), 27.56 (C-15), 27.98 (C-23), 28.74 (C-21), 29.05 (C-12), 31.54 (C-29/30), 34.54 (C-7), 35.55 (C-16), 37.07 (C-10), 37.44 (C-13), 38.68 (C-1), 38.84 (C-4), 40.20 (C-22), 41.34 (C-8), 43.51 (C-14), 44.64 (C-17), 48.29 (C-18), 49.93 (C-19), 50.27 (C-9), 55.17 (C-5), 73.50 (C-20), 79.00 (C-3).

**ATHLUP1 PRODUCT RATIO DETERMINATION**

Product ratios of the *AthLUP1* products were determined using a combination of GC-FID and NMR analysis. The ratios of triterpene acetates (mono-hydroxy) were determined by integration of GC-FID peak areas. The following GC conditions were used: inlet, 260 ºC split injection (40:1 split ratio); detector, 260 ºC; column, 250 ºC, isothermal, flow rate ~ 0.5 mL/min. $^1$H NMR integration of well resolved peaks
distinctive of lupeol and 3β,20-dihydroxylupane was used to determine the proportion of the diol relative to the other components. *AthLUP1* was shown to convert oxidosqualene to lupeol (39%), 3β,20-dihydroxylupane (39%), β-amyrin (8%), germanicol (7%), taraxasterol (4%), and ψ-taraxasterol (3%).

**SYNTHESIS OF THE *OLEA EUROPAEA* LUPEOL SYNTHASE**

Seventy oligonucleotides (35 forward, 35 reverse) corresponding to a customized *Olea europaea* lupeol synthase were ordered and received from the Great American Gene Co. and were diluted with Milli-Q H₂O to 125 pmol/µL. All (1 µL each) were combined to provide an oligonucleotide mixture of 125 pmol/µL (in total oligonucleotide content). The oligonucleotide mixture (1 µL) was assembled in a primerless PCR reaction under the following conditions: 1 µL (125 pmol) oligo mixture, 5 µL 10 × PC2 Buffer, 4.0 µL dNTPs (2.5 mM each), and Milli-Q H₂O to a final volume of 50 µL. Following a 2 min hot start, 0.5 µL of Taq polymerase (Fisher) and 0.2 µL Pfu polymerase (Stratagene) were added to the PCR reaction. The "ASSEM" PCR program consisted of a 2 min 94 °C hot start followed by 60 cycles of denaturation at 94°C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 30 s plus 3 s/cycle. Gel electrophoresis of the PCR product revealed a low molecular weight (MW) smear (< 700 bp).

An attempt to PCR amplify the full length *OeuLUP* coding sequence from the assembly PCR reaction (described above) was performed as follows. PCR reactions composed of 1 µL of assembly PCR sample (described above), 5 µL 10 × Pfu polymerase buffer, 4 µL dNTPs (2.5 mM each), 1 µL T7 primer (20 pmol/µL), 1 µL T3 primer (20 pmol/µL), 1.0 µL Pfu polymerase, and Milli-Q H₂O to 50 µL final volume were
amplified using the “PFU” program (40 cycles of denaturation at 95 °C for 45 s, annealing at 56 °C for 45 s, and extension at 72 °C for 5 min and ending with a 10 min extension at 72 °C). Gel analysis did not reveal a distinct band but instead showed a high MW smear (extending up to ~8 kb).

The presence of higher MW species in the second PCR (described immediately above) increased the likelihood of amplifying the full-length synthetic lupeol synthase gene. Therefore, another amplification PCR reaction was performed essentially as described above. Gel analysis of this PCR product revealed a strong product band of the expected size (~2.3 kb).

The PCR product was EtOH precipitated, digested with Not I and Xho I, and sub-cloned into Sal I, Not I cut pBluescript II KS(+) vector. The resultant construct was named pMS7.0 and sequenced. Surprisingly, frameshift deletions (~1 every 150 bases) were found throughout the length of the gene. Similar results were found in additional clones. In any given clone, only a maximum of 250-300 bases of continuous sequence could be found.

PCR primers were designed to amplify out twelve “clean” stretches of sequence that could be used to construct the correct full-length sequence. Portions of pMS7.0 sequence were amplified using the PFU program (above) and the primer sets LUP2a-F, LUP2a-R; LUP2b-F2, LUP2b-R; LUP2c-F, LUP2c-R; LUP2d-F, LUP2d-R; LUP2e-F, LUP2e-R; LUP2f-F, LUP2f-R; LUP2g-F, LUP2g-R2; LUP2h-F, LUP2h-R; LUP2i-F, LUP2i-R; LUP2j-F, LUP2j-R; LUP2k-F, LUP2k-R; and LUP2l-F, LUP2l-R as follows: 1.0 μL pMS7.0 construct, 5 μL 10 × Pfu polymerase buffer, 4 μL dNTPs (2.5 mM each), 1.0 μL each primer (20 pmol/μL each), 1.0 μL cloned Pfu polymerase, and Milli-Q H₂O
to 50 µL total volume. The PCR products were gel purified and pooled together (5 µL each).

A portion of the clean fragment PCR mixture was then assembled and amplified as described above with the following exceptions. Only cloned Pfu polymerase (1.0 µL) was used in the assembly PCR reaction (rather than a Taq/Pfu mixture) and only one assembly reaction was required to assemble the fragments prior to amplification. The PCR product was sub-cloned into pBluescript II KS(+) as described above and sequenced. Two clones (pMS9.0-2 and pMS9.0-3) were relatively error free and a ~900 bp BsrG I fragment of pMS9.0-2 was combined with a ~4400 bp BsrG I fragment of pMS9.0-3 to construct the synthetic *Olea europaea* lupeol synthase gene. The construct (in pBluescript II KS(+) was named pMS13.0 and encoded a gene that encoded a peptide identical to the published *OeuLUP* sequence with the exception of one nucleotide change that resulted in the Asn731Ile mutation. This plasmid was used for subsequent DNA shuffling experiments. The synthetic lupeol synthase Asn731Ile mutation was later corrected by site-directed mutagenesis and the corrected constructs were named pMS30.3 (pRS426GAL) and pMS30.0 (pBluescript II KS(+)).

**DNase I Digest for DNA Shuffling**

pSM60.40 (*AthCAS1* in pBluescript II KS(+) and pMS13.0 (*OeuLUP* in pBluescript II KS(+) were fragmented by DNase I. To 1.5 mL centrifuge tubes were added 10 × DNase I buffer (5 µL), pSM60.40 and pMS13.0 DNA (15 µg total), DNase I (0.2 U), and Milli-Q H₂O to a final volume of 50 µL. Two different ratios of the plasmid constructs were used (a 1:1 ratio of pSM60.40 : pMS13.0 and a 1:2 ratio of pSM60.40 :
pMS13.0). The digests were allowed to incubate at room temperature for 15 min and were then quenched with 25 mM EDTA (4 µL) and heat killed at 70 °C for 10 min.

**Purification of DNase I Digested Fragments**

The DNase I digests were run out on a 2% agarose prep gel (GTAE) against a 100 bp ladder. To keep the DNA fragments concentrated enough for gel purification, the samples were electrophoresed only until the 100 and 200 bp bands of the DNA ladder were sufficiently resolved. Gel slices containing fragments ≤~200 bp were excised with a razor blade and the DNA was recovered using a Qiaquick gel purification kit (Qiagen).

**DNA Fragment Reassembly for DNA Shuffling**

Fragments of pSM60.40 and pMS13.0 (8 µL of each of the 1:1 and 1:2 mixtures generated by DNase I digestion) were assembled in a primerless PCR reaction. To PCR tubes were added 5 µL 10 × cloned Pfu polymerase buffer, 4 µL dNTPs (2.5 mM each), 8 µL of gel purified DNA fragments, and Milli-Q H₂O to a final volume of 50 µL. PfuTurbo polymerase (1.0 µL) was added after the 2 min hot start. The fragments were assembled using the "ASSEMBL" PCR program that consisted of a 2 min hot start at 95 °C followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, extension at 72 °C for 1 min plus 4 s/cycle, and a final 5 min extension step at 72 °C.

The PCR products from the first round of fragment assembly were carried through a second round of assembly. The first round PCR product (6 µL) was mixed with 5 µL 10 × cloned Pfu polymerase buffer, 4 µL dNTPs (2.5 mM each), and 34 µL Milli-Q H₂O.
PfuTurbo polymerase (1.0 μL) was added after the 2 min hot start of the “ASSEM2" PCR program. This program consisted of a 2 min hot start at 95 °C; 30 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, extension at 72 °C for 3 min plus 4 s/cycle; and a final 8 min extension step at 72 °C.

**DNA SHUFFLE PCR AMPLIFICATION**

Full-length genes were amplified from the assembled DNA fragments (two rounds of assembly) using the T7 and T3 primers. Reactions were composed of 6 μL 10 × ExTaq buffer, 5 μL dNTPs (2.5 mM each), 1.6 μL of each primer (both at 20 pmol/μL), 4 μL of assembled DNA fragments, and Milli-Q H₂O to a final volume of 60 μL. ExTaq polymerase (0.3 μL) was added after the 1 min hot start. PCR amplification was done using the “AMP” PCR program that was designed as follows: 1 min hot start at 95 °C; 30 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 3 min; followed by a 10 min extension step at 72 °C.

**DNA SHUFFLE LIBRARY SUB-CLONING**

PCR products from DNA shuffle amplification PCR reactions (100 to 300 μL) were brought to 492.5 μL with T8 buffer. To these were added 1μg/μL SDS (2.5 μL) and 500 mM EDTA pH 8 (5.0 μL) and the samples were then mixed. Proteinase K (25 μg) was added, the samples incubated at 37 °C for 1 h, and the digests were heat killed at 70 °C for 15 min. Digests were extracted once with phenol:chloroform:iso-amyl alcohol 25:24:1 (500 μL) and once with chloroform (500 μL). Absolute EtOH (2.5 volumes) and 3 M NaOAc (0.1 volume) were added to the aqueous extract and the samples were mixed
and stored at −20 °C for ~ 30 min. The DNA was then pelleted at maximum speed in a microcentrifuge for 15 min and the supernatants were discarded. DNA pellets were washed with 70% EtOH and then allowed to air dry. T8 buffer (50 μL) was added to the DNA pellets and the samples dissolved by vortex mixing. This material was then additionally purified prior to restriction digest using the Qiaquick PCR purification kit (Qiagen) according to manufacturer’s instructions.

**DNA Shuffle Library Construction and Propagation**

Purified DNA shuffle PCR products were digested with Not I and Xho I, gel purified with the Qiaquick gel purification kit (Qiagen), and ligated into Not I, Sal I cut pRS426GAL yeast expression vector. The ligations were heat killed and transformed into UltraMax chemically competent DH5α (Gibco BRL) according to manufacturer’s instructions. Small aliquots (1 and 10 μL) were plated on LB/Amp plates to count transformation efficiency and the rest of the bacterial transformation was plated on three 15 × 150 mm LB/Amp plates. Liquid LB/Amp media was added to the transformation plates and the colonies were scraped using a sterile plate spreader. The cell suspension was collected, glycerols were made, and a plasmid prep culture inoculated.

**DNA Shuffled Library Screening**

Shuffled libraries (in pRS426GAL) were transformed into MSY4 using the standard yeast transformation procedure described in Chapter 2. After sitting at room temperature for 7 h., the transformations (0.5 mL) were plated on selective media. About 20 μL was plated on SCD-Ura,H,E (to count transformation efficiency) while the
remaining transformation material was plated on two 15 × 150 mm YPG,H plates. The plates were placed in the 30 °C incubator and checked periodically for growth of sterol-independent colonies.

**Characterization of CAS:LUP Shuffled Oxidosqualene Cyclases**

After 7-10 d, several of the fastest grown colonies were picked and used to inoculate SCD-Ura,H,E (10 mL) cultures for plasmid recovery. Upon saturation, the yeast cells were collected by centrifugation, resuspended in Milli-Q H₂O (1 mL), and transferred to 1.5 mL centrifuge tubes. The cells were pelleted again, resuspended in 150 μL SCE/Lyticase/β-mercaptoethanol solution (5 mL SCE buffer, 3.0 mg lyticase, and 40 μL β-mercaptoethanol), and incubated at 37 °C for 40 min. At that time, 200 μL SDS solution (100 mM Tris-HCl pH 8, 10 mM EDTA, 2% SDS) was added and the sample incubated at 65 °C for 5 min. Cellular debris was precipitated by adding 200 μL potassium acetate (5 M) and chilling on ice followed by centrifugation. The supernatant was transferred to a new tube and 5 M ammonium acetate (200 μL) and isopropanol (850 μL) were added. The sample was mixed by vortexing and stored overnight at -20 °C. DNA was pelleted and the supernatant was discarded. The pellet was washed with 70% EtOH (1 mL), allowed to air dry, and dissolved in 0.5 mL T8 buffer. The samples remained cloudy with insoluble material, presumably genomic DNA.

A Wizard DNA purification kit (Promega) was used to purify the DNA recovered from yeast. *E. coli* (DH5α) was transformed with the yeast DNA and the plasmids were recovered and sequenced. Several different complementing genes were identified (pMS14.3, pMS15.3, pMS16.3, pMS17.3, pMS18.3, pMS19.3, and pMS20.3) but
pMS16.3, pMS17.3, and pMS18.3 were false positives that encoded an intact *Ath*CAS1. The pMS14.3 and pMS19.3 genes (discussed in chapter 4) contained the most amount of DNA exchange. The pMS14.3 construct was an *Ath*CAS1 protein in which 140 amino acids of the extreme C-terminus were encoded by *Oeu*LUP sequence. The pMS19.3 construct encoded an *Ath*CAS1 protein in which 115 amino acids of the extreme N-terminus were encoded by *Oeu*LUP sequence.

MSY4 was retransformed with these plasmids and the complementation experiments were repeated to ensure the results were reproducible. In addition, small scale (10 mL) assays were performed as described in chapter 2 and the assays were monitored by TLC to ensure the hybrid proteins had cyclase activity.

**Construction of Hybrid Oxidosqualene Cyclases**

Several oxidosqualene cyclase hybrids were generated by piecing together different portions of *Oeu*LUP, *Ath*CAS1, *Ath*CAS1 mutants, and shuffled genes. The .0 suffix (ex. pMSXX.0) designates the pBluescript II KS(+) vector, the .3 suffix (ex. pMSXX.3) designates the yeast expression pRS426GAL, and the .9 suffix (ex. pMSXX.9) designates the pRS246GPD yeast expression vector.

**pMS23**

The pMS23.3 construct was created by sub-cloning a pMS14.3 *Bsr*G I, *Not* I fragment (~600 bp) into a pMS19.3 *Bsr*G I, *Not* I fragment (~ 8100 bp). The ~2200 bp *Xho* I, *Not* I fragment (gene insert) of pMS23.3 was sub-cloned into *Sal* I, *Not* I cut pBluescript II KS(+) to generate pMS23.0.
pMS47

The pMS47.3 CAS:LUP chimera was constructed by sub-cloning a pMS30.3 Hind III, Not I fragment (~600 bp) into a similarly cut pMS23.3 fragment (~8100 bp).

pMS55.9 and pMS56.9

The pMS55.9 and pMS56.9 constructs were derived from the pMS50.9 and pMS51.9 plasmids. A BstAP I, Hind III fragment (~900 bp) of pSM60.21 Tyr410Thr His477Gln Ile481Val that contained three point mutations was sub-cloned into similarly cut pMS47.0 and pMS23.0 to yield the pMS50.0 and pMS51.0 constructs, respectively. The genes encoded by pMS50.0 and pMS51.0 (~2300 bp) were excised with Xho I and Not I and then sub-cloned into Sal I, Not I cut pRS426GPD to provide pMS50.9 and pMS51.9, respectively.

The pMS55.9 and pMS56.9 genes (shown in chapter 4), which have different LUP synthase compositions at the C-terminal end of the enzyme, were constructed by swapping Xma I fragments of pMS50.9 and pMS51.9. The ~2200 bp Xma I fragment of pMS50.9 and the ~6800 bp Xma I fragment of pMS51.9 were ligated to create pMS55.9. Conversely, the ~6800 bp fragment of pMS50.9 and the ~2200 bp fragment of pMS51.9 were combined to construct pMS56.9.

**AttnCAS1 ASN727 AND ILE732 MUTANTS**

Site directed mutagenesis, with primers 5'-GCATATGTTATCATGCAATTTGC GTGCGAAGACTCCCATTATT-3' (Mfe I restriction site added) and 5'-GCATAT GTTATCATGCAATTTGCGATGAAGACTCCCATTATT-3' (Mfe I restriction site added), was employed to make the AttnCAS1 Asn727Ala and Asn727Met mutations,
respectively, in the AthCAS1 gene encoded by pSM60.21. Similarly, the Ile732Ala and Ile732Leu mutations were made in the pSM60.21 encoded AthCAS1 gene using the 5'-CGATACGCGCATATGTGGCCATGCAATTGCGATTGAAGACTCCCATTA-3' (Mfe I restriction site added) and 5'-CGATACGCGCATATGTGAACCATGCAATTGCGATTGAAGACTCCCATTA-3' (Mfe I and Hinc II restriction sites added) primers, respectively.

Each of the mutant constructs were linearized with BstE II and transformed into SMY8. Positive transformants were selected on SCD-Leu,H,E. Small scale assays (10 mL YPG,H,E media) were performed to determine if the mutations had a catalytic role in oxidosqualene cyclization. TLC analysis revealed strong conversion of oxidosqualene to triterpene alcohol product. GC and GC-MS analysis of the small scale assays revealed that cycloartenol was the only product formed in each of the mutants, indicating that the single mutations at these positions did not significantly influence the reaction catalyzed by cycloartenol synthase.
CHAPTER 4: RESULTS AND DISCUSSION

The oxidosqualene cyclases generate a diverse array of structurally complex carbocyclic products through complex chemical transformations. How the oxidosqualene cyclases promote specific reactions and achieve product specificity by excluding competing reaction modes is still not completely understood. I have used protein mutagenesis to identify structural elements of oxidosqualene cyclases that participate in the deprotonation and substrate-folding aspects of catalysis—two reaction steps that are partly responsible for product diversity among the triterpenes. The results of these research efforts will be discussed in two sections. Part 1 describes the identification of amino acid residues that influence the different deprotonation steps catalyzed by lanosterol and cycloartenol synthase. Part 2 describes DNA shuffling of cycloartenol synthase (a protosteryl-type cyclase) and lupeol synthase (a dammarenyl-type cyclase) in an effort to determine what enzymatic elements control substrate folding and subsequent protosteryl or dammarenyl cation formation.

PART 1. DIFFERENTIAL DEPROTONATION

CHARACTERIZATION OF SceERG7 Thr384Tyr AND RELATED MUTANTS

In a continued effort to determine the catalytic differences between cycloartenol and lanosterol synthase, we typically construct complementary mutations of catalytically relevant positions in both backgrounds. Previous work established that mutating the Arabidopsis thaliana cycloartenol synthase (AthCAS1) Tyr410 residue dramatically
altered catalysis.\textsuperscript{19} At this position, cycloartenol synthases maintain a Tyr but the lanosterol synthases conserve a Thr. In addition to the catalytic behavior of the \textit{AthCAS1} Tyr410Thr mutant and this conservation pattern, the squalene-hopene cyclase residue corresponding to \textit{AthCAS1} Tyr410 is in the active site. This evidence suggests that the corresponding position in the lanosterol synthase may also be catalytically relevant.

We therefore made the corresponding \textit{Saccharomyces cerevisiae} lanosterol synthase (\textit{SceERG7}) mutation. The cycloartenol synthase residue was incorporated at the corresponding position in the \textit{Saccharomyces cerevisiae} lanosterol synthase (\textit{SceERG7}) to create the Thr384Tyr mutant. This lanosterol synthase mutant was expressed in the yeast strain LHY\textsuperscript{35} (which has no oxidosqualene cyclase activity nor the ability to produce the oxidosqualene substrate), and a homogenate was incubated with racemic oxidosqualene. Formation of triterpene alcohol products and another more polar product was evident by TLC. The two product fractions were isolated by silica gel chromatography. Comparison of GC-FID and GC-MS data and NMR spectra of the triterpene alcohol fraction with authentic standards confirmed the presence of both lanosterol and parkeol. The polar component of the product mixture was identified as lanosta-24-ene-3β,9α-diol\textsuperscript{55} (9α-hydroxylanosterol) by NMR spectroscopy (\textsuperscript{1}H, \textsuperscript{13}C, DEPT, COSY, DECO, HMQC, and HMB). This diol is a known synthetic compound but had not been previously identified in nature. The Thr384Tyr mutant converts oxidosqualene to lanosterol, parkeol, and 9α-hydroxylanosterol in a 79:11:10 ratio\textsuperscript{56,57} (Figure 4.1).

\textit{SceERG7} Val454 mutations had also previously been shown to influence catalysis and alter product structure.\textsuperscript{46} In an effort to determine if and how the Thr384Tyr and
Val454 mutations interact, we created the double mutant where both mutated positions were changed to the corresponding cycloartenol synthase residue. The *SceERG7 Thr384Tyr Val454Ile* double mutant converts oxidosqualene to lanosterol (13%), parkeol (64%), and 9α-hydroxylanosterol (23%).\textsuperscript{56,57} This double mutant has catalytic properties distinct from either of the parents. Addition of the Val454Ile mutation seems to strongly favor parkeol formation and also leads to a modest increase in 9α-hydroxylanosterol formation (Figure 4.1).

![Chemical structures](image)

<table>
<thead>
<tr>
<th>Mutant Configuration</th>
<th>Lanosterol</th>
<th>Parkeol</th>
<th>9α-Hydroxylanosterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>SceERG7 Thr384Tyr</td>
<td>79%</td>
<td>11%</td>
<td>10%</td>
</tr>
<tr>
<td>SceERG7 Thr384Tyr Val454Ile</td>
<td>13%</td>
<td>64%</td>
<td>23%</td>
</tr>
<tr>
<td>SceERG7 Thr384Tyr Val454Leu</td>
<td>20%</td>
<td>60%</td>
<td>20%</td>
</tr>
</tbody>
</table>

**Figure 4.1. Product profiles of *SceERG7 Thr384* mutants.** The Thr384Tyr mutation results in a loss of product specificity and causes formation of the novel 9α-hydroxylanosterol. Combining the Thr384Tyr mutation with either the Val454Ile or Val454Leu mutations caused a strong preference for parkeol formation and a modest increase in diol formation.
The synergism of these two mutations led us to investigate the influence of minor steric changes at position 454 by constructing the Thr384Tyr Val454Leu double mutant. Thus, the only difference between the double mutants is exchange of the sec-butyl group of isoleucine with the isobutyl moiety of leucine. Similar to the other double mutant, the SceERG7 Thr384Tyr Val454Leu mutant favors parkeol formation and converts oxidosqualene to lanosterol, parkeol, and 9α-hydroxylanosterol in a 20:60:20 ratio.

The formation of 9α-hydroxylanosterol provides several important insights about these mutants and perhaps the oxidosqualene cyclases in general. First, the water molecule was added to the central part of the cyclization product, indicating that coordinated water may be present in this and other oxidosqualene cyclase active sites. Furthermore, the stereochemistry of the triterpene diol product suggests that the enzyme controlled water addition. The water molecule was added stereospecifically to the α face of the cationic intermediate, consistent with addition concerted with C-9 to C-8 hydride shift. This is the first example of lanosta-24-ene-3β,9α-diol arising from an enzyme, and its formation confirms that triterpene diols can be direct cyclization products.

Consistent with sequence alignment data, structural data, and the behavior of the corresponding cycloartenol synthase mutant, the SceERG7 Thr384Tyr mutation significantly influences catalysis. The synergy observed upon combining the Thr384Tyr and Val454Ile or Val454Leu mutations is quite dramatic. All of the single Thr384 and Val454 mutants biosynthesize lanosterol as their major or sole product. However, when combined, the double mutants become fairly accurate parkeol synthases in which lanosterol is the minor component. The catalytic behavior of the double mutants is
especially interesting when compared to the Val454Ile and Val454Leu single mutants, which do not influence product structure and form lanosterol as their sole product.

**IDENTIFICATION OF THE *AthCAS1** His477Asn MUTANT**

A substantial number of *AthCAS1* mutations that alter product specificity have been reported,\(^{19,35,38,58,59}\) but broadened product specificity is also reported in each case. Some attempts to combine some of these known mutations have narrowed the product specificity,\(^{19}\) but they have not yet generated an accurate enzyme that generates a product distinct from the native cyclase. Specifically, efforts in our group have found that the known mutations are not sufficient to convert a cycloartenol synthase into an accurate lanosterol synthase. We therefore pursued random mutagenesis and selection experiments with the *Arabidopsis thaliana* cycloartenol synthase\(^{16}\) (*AthCAS1*) to identify additional positions that may be involved in the catalytic distinction between the lanosterol and cycloartenol synthases.

Amplification of a plasmid encoding *AthCAS1* in the mutagenic *E. coli* strain XL1-Red (Stratagene) created a library of randomly generated *AthCAS1* mutants. Several DNA repair mechanisms are inactivated in this bacterial strain, which consequently allows the incorporation of random nucleotide changes that arise during DNA replication. The mutant library was then transformed into the yeast lanosterol synthase mutant SMY8\(^{30}\) for genetic selection of mutant enzymes that allow lanosterol biosynthesis. In yeast, lanosterol is the triterpene precursor of the required sterol, ergosterol. Because the SMY8 yeast strain is defective in lanosterol biosynthesis, ergosterol must either be supplemented in the culture media, or a heterologously
expressed gene that promotes lanosterol biosynthesis must be added for cell growth. However, it is important to note that only traces of ergosterol are sufficient to support yeast growth (oxidosqualene cyclase mutants that make as little as 13% lanosterol allow yeast to grow at a rate similar to wild-type).\textsuperscript{56,57}

Yeast transformants were first selected on synthetic complete media lacking uracil and supplemented with heme and ergosterol. Functional selection of the mutant library was then initiated by replica-plating the yeast transformants onto galactose media containing heme but not ergosterol. Several colonies were able to grow in the absence of exogenous ergosterol, and they were subsequently re-cultured several times to ensure this sterol-independent behavior was reproducible. As a secondary screen, all of the positive yeast constructs were analyzed for ergosterol content. The presence of this lanosterol metabolite in culture extracts indicates that the \textit{Ath}CAS1 mutant biosynthesizes some lanosterol. The plasmids were recovered and sequenced and each mutant gene contained only one mutation. The mutant \textit{Ath}CAS1 plasmid constructs were then re-transformed into SMY8 and re-screened to ensure that the sterol-independent phenotype was plasmid linked.

Two of the best complementing mutants contained the His477Asn mutation and the Tyr532His mutation. The Tyr532His mutation had already been identified and described in a previous study.\textsuperscript{58} His477 is a newly identified amino acid residue that influences cycloartenol formation.\textsuperscript{59} All known cycloartenol synthases maintain a His at this position. Lanosterol synthases lack strict conservation at the corresponding position maintaining either a glutamine (fungi) or a cysteine (mammals) (Figure 4.2). As was discussed in Chapter 1, sequence analysis for positions that are strictly and differentially
conserved in the cycloartenol and lanosterol synthases has proven to be a useful way to identify residues that may impart the catalytic distinction between the two enzymes. However, *AthCAS1* His477 would not have been identified using those criteria. This may indicate that strict conservation in one type of enzyme (rather than both) is an appropriate indicator of catalytic relevance.

![Alignment of amino acid sequences](alignment.png)

**Figure 4.2. Conservation pattern of *AthCAS1* His477.** Histidine is strictly conserved in the known cycloartenol synthases (CAS1) from *A. thaliana*16 (*Ath*), *Pisum sativum*20 (*Psa*), *Panax ginseng*21 (*Pga*), *Glycyrrhiza glabra*22 (*Ggl*), *Luffia cylindrica*23 (*Lcy*), *Avena sativa*13 (*Asa*), and *Dictyostelium discoideum*24 (*Ddi*). Glutamine is conserved at the corresponding position in trypansomal (*Trypanosoma brucei*26 (*Tbr*)), *Trypanosoma cruzi*27 (*Tcr*) and fungal (*S. cerevisiae*25,26 (*Sce*), *Candida albicans*27,28 (*Cal*), *Cephalosporium caerulescens*29 (*Cca*), *Schizosaccharomyces pombe*30 (*Spo*) lanosterol synthases, whereas mammalian (*Rattus norvegicus*31,32 (*Rno*), and *Homo sapiens*33,34 (*Hsa*)) lanosterol synthases maintain a cysteine.

**CHARACTERIZATION OF *AthCAS1* HIS477ASN AND HIS477GLN MUTANTS**

The *AthCAS1* His477Asn mutant, described above, was expressed and characterized in the yeast strain LHY4.35 LHY4 is similar to SMY8 but also has a
squalene synthase (ERG9) deletion that precludes *in vivo* oxidosqualene biosynthesis. Consequently, all oxidosqualene cyclization products are derived from the *in vitro* assay. This ensures unbiased and accurate product ratio determination by eliminating further metabolism of products formed *in vivo*. Yeast grown on liter scale were homogenized and incubated with racemic oxidosqualene. Partially purified material was isolated and analyzed by GC, GC-MS and NMR. GC and GC-MS showed two peaks with chromatographic and mass spectral properties identical to those of authentic lanosterol and parkeol samples. No other peaks with masses consistent with being oxidosqualene cyclase products were evident in the GC-MS. The structural assignments of the products were confirmed by $^1$H NMR. GC-FID quantitation of the product mixture showed that the *A. thaliana* CAS1 Hs477Asn mutant makes lanosterol and parkeol in a 88:12 ratio$^{60}$ (Figure 4.3).

The asparagine residue in the His477Asn mutant is not found at that position in either lanosterol or cycloartenol synthase. Because most lanosterol synthases possess a glutamine at that position, we anticipated that mutating *A. thaliana* CAS1 His477 to the native yeast lanosterol synthase residue (glutamine) would further enhance the lanosterol biosynthetic ability of the mutant. Using site-directed mutagenesis, we therefore created the *A. thaliana* CAS1 Hs477Gln mutant and characterized it as described above. Interestingly, the His477Gln mutant is a rather accurate parkeol synthase. GC, GC-MS, and NMR demonstrated that this mutant makes parkeol (73%), lanosterol (22%), and 9β-lanosta-7,24-dien-3β-ol (5%).$^{60}$ It is important to note here that several orthogonal analysis methods were required for accurate characterization of this mutant and, in general, are required for accurate characterization of all oxidosqualene cyclase mutants. Many
triterpene alcohols are chromatographically and spectroscopically similar and are consequently difficult to purify. For example, lanosterol and its Δ7 isomer, 9β-lanosta-7,24-dien-3β-ol (formed by the His477Gln mutant described above) are not resolved under a wide variety of GC conditions. Consequently, analysis of this mutant using only GC would have caused incorrect reporting of the product composition. Unfortunately, inaccurate mutant characterizations have been reported in studies where appropriate analytical methodology was not used.\textsuperscript{59}

The \textit{AthCAS1} His477Asn and His477Gln mutants are currently the best lanosterol and parkeol synthases, respectively, generated by mutagenesis. Quite surprisingly, the His477Asn and His477Gln mutations result in dramatically different catalytic properties in the \textit{AthCAS1} background. The asparagine and glutamine residues differ by only one methylene unit, yet they result in almost the opposite product specificity—lanosterol is the major product (88\%) of the His477Asn mutant but parkeol is the major product (73\%) in the His477Gln mutant. Furthermore, the native lanosterol synthase residue (glutamine) actually promotes less lanosterol production than does the asparagine residue, which is not found at that position in either the lanosterol or cycloartenol synthases.

Despite the strong catalytic influence of His477 mutations, His477 is not in the active site based on comparison with the squalene-hopene cyclase crystal structure. Other models have also suggested a non-active site position for \textit{AthCAS1} His477.\textsuperscript{59} However, the conservation of histidine at this position in the cycloartenol synthases, the catalytic behavior of His477 mutants, and the manner in which a very subtle structural change between the His477Asn and His477Gln mutations results in dramatically different
catalytic behavior are all consistent with an active site role for His477. The physical location of *AthCAS1* position 477 will not likely be resolved without a cycloartenol synthase crystal structure. If, however, a crystal structure does demonstrate that the 477 position does indeed lie outside the active site, this position would have remarkable catalytic influence (as demonstrated by the His477 mutations) for a residue outside of the active site.

![Chemical structures and reaction pathway](image)

**Figure 4.3. Product profiles of *AthCAS1* His477 mutants.** In the native *AthCAS1* background, a histidine, asparagine, or glutamine residue at position 477 have very different effects on catalysis. However, when His477 mutations are combined with the Tyr410Thr Ile481Val mutations, the identity of that position is catalytically irrelevant.
CHARACTERIZATION OF *Ath*CAS1 His477 RELATED TRIPLE MUTANTS

As discussed above, an increasing number of catalytically relevant cycloartenol synthase mutants are now known and in all cases described, the mutations broaden product specificity. However, little success has been found in narrowing the product specificity of a mutant enzyme toward some compound distinct from that produced by the native enzyme. Due to the complexity of the oxidosqualene cyclase reaction and the precedent of broadened product specificity among single mutants, multiple mutations are probably required to specifically promote the formation of one product and to exclude the formation of alternative products.

The *Ath*CAS1 Tyr410Thr and Ile481Val mutations have been shown to influence catalysis (they produce 65% and 24% lanosterol, respectively) and when combined, the resultant double mutant has catalytic characteristics distinct from either of the single mutants (75% lanosterol production). In a continued effort to convert a cycloartenol synthase into a specific lanosterol synthase, the His477Asn mutation (described above) was incorporated and the triple mutant was made by site-directed mutagenesis.

The *Ath*CAS1 Tyr410Thr His477Asn Ile481Val triple mutant was expressed and assayed in the yeast strain LHY4 as described above. Product determination was carried out as described previously (GC, GC-MS, NMR) and interestingly, the triple mutant was experimentally indistinguishable from the Tyr410Thr Ile481Val double mutant (Figure 4.3). The His477Asn mutation did not contribute to lanosterol production nor did it influence byproduct composition (parkeol is the side-product in the His477Asn mutant but it is not formed by the triple mutant).
We suspected that this behavior was perhaps due to the fact that the asparagine residue is not found at the position corresponding to AthCAS1 His477 in either the lanosterol or cycloartenol synthases. We therefore made the AthCAS1 Tyr410Thr His477Gln Ile481Val triple mutant in which all three mutated positions were converted to their lanosterol synthase counterparts. It is important to note here that although the His477Gln single mutation favors parkeol formation (73%) in the AthCAS1 background, its presence in lanosterol synthases show that a glutamine at this position is consistent with lanosterol biosynthesis. The AthCAS1 Tyr410Thr His477Gln Ile481Val triple mutant was characterized and its product profile was also indistinguishable from that of the Tyr410Thr Ile481Val double mutant (Figure 4.3). The lack of catalytic influence by the His477Gln mutation in this triple mutant is even more dramatic because the His477Gln single mutant caused predominant parkeol formation. Yet, no parkeol is identified in the triple mutant. Once again, the catalytic influences of the Tyr410Thr Ile481Val mutations dominate those of the His477 mutation.

This is the first demonstration of a hierarchy among catalytically important oxidosqualene cyclase mutations.\textsuperscript{61} Previous reports have shown that combinations of oxidosqualene cyclase mutations create mutant enzymes with catalytic behavior distinct from either of the parents, indicating that mutations typically have additive effects. However, these His477 single and triple mutants are an important illustration of the dominant influence some mutations can have over other mutations. In the native AthCAS1 background where the Tyr410 and Ile481 residues are unchanged, a histidine, asparagine, or glutamine at position 477 have dramatic and different effects on the catalyzed reaction. Yet, when combined with the Tyr410Thr and Ile481Val mutations,
the identity of the amino acid at position 477 is irrelevant for catalysis. This behavior suggests that the catalytic influence of the Tyr410 and His477 mutations are mutually exclusive; only the influence of one or the other may be displayed, but not both.

Because no cycloartenol synthase crystal structure exists, it is difficult to accurately define how the dominant nature of some mutations arise and how in other cases mutations show synergistic influences. However, some plausible explanations can be proposed. The dominant behavior of certain mutations could arise from steric or electronic changes within the active site. If one mutation alters the positioning of the substrate or reaction intermediate(s), the manner in which another mutation interacts with the substrate or reaction intermediate(s) might be altered. Differential energetics of various reaction pathways could also render some mutations dominant. A given mutation (or mutations) might allow access to an energetically favorable pathway (such as lanosterol formation relative to cycloartenol formation); although two pathways might be accessible, the energetically favored one would be dominant. Finally, the physical arrangement of the mutations might be another way dominant behaviors of certain mutations arise. A mutation that influences a late step in the reaction sequence could be precluded by another mutation affecting an earlier catalytic step. For example, one mutation may cause deprotonation of a specific intermediate but its influence would be preempted by another mutation that blocked the formation of that specific intermediate.

The latter model best fits the experimental data in hand. Parkeol and cycloartenol formation require a hydride shift from C-9 to C-8. However, the lanosterol and 9β-lanosta-7,24-dien-3β-ol products identified in the double and triple mutants do not require that shift and can be formed via the C-8 cation. Tyr410 has previously been
proposed to stabilize the C-9 cation or facilitate the hydride shift from C-9 to C-8.\textsuperscript{56,57} Appetiable amounts of parkeol have been identified in lanosterol and cycloartenol synthase mutants that have a Tyr at the position corresponding to \textit{AthCAS1} position 410, but mutants with Thr at that position do not generate significant amounts of C-9 cation derived products. Tyr\textsuperscript{410}Thr Ile481Val mutations appear to block the C-9 to C-8 hydride shift and thereby preclude the influence of the His477 position in determining C-9 cation derived product formation.

\textbf{PART 2. PROTOSTERYL VERSUS DAMMARENYL CATION FORMATION}

This section describes experiments designed to identify oxidosqualene cyclase positions that direct either protosteryl or dammarenyl cation formation. The substrate folding conformations that yield these intermediate cations ultimately determine a wide variety of structural and stereochemical diversity found among the triterpenes. DNA shuffling of cycloartenol synthase (protosteryl-type cyclase) and lupeol synthase (dammarenyl-type cyclase) coupled with a genetic selection for cycloartenol synthase activity was used to elucidate key enzyme positions. Construction of the novel MSY4 yeast strain used for the genetic selection and complete product characterization of the \textit{Arabidopsis thaliana} lupeol synthase will be described as a prelude to the DNA shuffling work.
CONSTRUCTION OF PLANT STEROL BIOSYNTHETIC PATHWAY IN YEAST

Cycloartenol is the triterpene sterol precursor in higher plants and some protists. However, plant membrane sterols typically lack the cyclopropyl ring found in cycloartenol, suggesting that some enzymatic activity was required for cyclopropyl ring opening. Several reports described both direct\textsuperscript{62,63} and indirect\textsuperscript{64-67} measurements of such an enzymatic activity in plants that isomerized cycloecalenol to obtusifoliol. We therefore screened an Arabidopsis thaliana cDNA library for a gene that complemented the sterol auxotrophy of yeast lanosterol synthase mutant expressing cycloartenol synthase. A positive strain was identified, the plasmid recovered, and the cDNA sequenced.

The 842 bp cDNA encodes a predicted 36 kDa protein. A bacterial expression host was used to confirm the activity of the putative enzyme in a series of in vitro assays. The gene product was first assayed for its ability to convert cycloecalenol to obtusifoliol. Although the activity was low (perhaps due to poor bacterial expression), it clearly established that the newly identified gene and encoded protein caused the ring-opening/isomerization of cycloecalenol to obtusifoliol. The harsh conditions (10% H\textsubscript{2}SO\textsubscript{4} in refluxing isopropanol for 24 h) required for non-enzymatic isomerization of cyclopropyl sterols precluded that the detected isomerization was non-enzymatic. Appropriate controls ensured that the cyclopropane ring opening was linked to the gene expression. We also investigated if cycloartenol was an acceptable substrate for this enzyme, but no isomerase activity was detected. The newly identified cycloecalenol cycloisomerase gene was named cyclopropyl sterol isomerase (CPII).\textsuperscript{68}
We anticipated that this newly identified gene could be exploited to create a metabolically engineered yeast strain with a plant-like sterol biosynthetic pathway. Such a strain would cyclize oxidosqualene to cycloartenol rather than lanosterol, and then use cyclopropyl sterol isomerase to convert cycloartenol-derived triterpenes to tetracyclic sterol precursors that yeast could convert to ergosterol. A derivative lanosterol synthase mutant expressing the cyclopropyl sterol isomerase but lacking the cycloartenol synthase could be very useful for cloning cycloartenol synthases. Although the sterol auxotrophy of a simple lanosterol synthase mutant cannot be genetically complemented by a cycloartenol synthase, the cyclopropyl sterol isomerase might allow the conversion of cycloartenol to ergosterol precursors.

A potential problem with this approach was suggested by previous experiments that established that cycloecalenol was an acceptable cyclopropyl-sterol isomerase substrate whereas cycloartenol was not. The 4-β methyl group of cycloartenol apparently blocks enzymatic access to the β-oriented cyclopropyl group. Only the 4-α methyl group is present in cycloecalenol, which probably allows better access for the cycloecalenol cycloisomerase to the cyclopropyl moiety (Figure 4.4). However, many of the enzymes along the yeast sterol biosynthetic pathway have broad substrate specificity, and yeast sterol biosynthesis is better viewed as a network rather than a strictly linear pathway. It therefore seemed possible that even though cycloartenol is not a natural yeast compound, the native yeast lanosterol 4-demethylase might oxidize cycloartenol to 4-desmethyl cycloartenol, which cyclopropyl sterol isomerase could isomerize to a tetracyclic sterol precursor. Cycloartenol synthase should therefore satisfy the cell’s sterol requirement,
and would genetically complement the lanosterol synthase deletion if cycloartenol demethylation and isomerase activities were sufficient.

![Chemical structures](image)

**Figure 4.4. Cycloartenol is not a suitable substrate for AthCPII.** *In vitro* assays demonstrated that the *AthCPII* was able to convert cycloeucalenol to obtusifoliol. Isomerization of cycloartenol was not observed. Removal of the 4-β methyl group of cyclopropyl sterols is probably required to provide the CPII access to the cyclopropane moiety.

To create this strain, the *Arabidopsis thaliana* cyclopropyl sterol isomerase (*AthCPII*) was sub-cloned into the integrative yeast expression vector pRS305GAL.\(^{69,70}\) This construct was then transformed into the yeast lanosterol synthase deletion mutant SMY8. Transformants were isolated on selective media and re-grown on selective media to make sure they were stable. The resultant yeast strain was named MSY4.
MSY4 was transformed with the pSM60.22 plasmid containing the *Arabidopsis thaliana* cycloartenol synthase (AthCAS1) in the high-copy, inducible yeast expression vector pRS426GAL. Transformants were first isolated on ergosterol-supplemented selective dextrose medium, which does not induce plasmid gene expression and consequently selects for plasmid incorporation rather than plasmid-born gene activity. Positive transformants were then grown on galactose media (to induce expression of both AthCPI1 and AthCAS1) without ergosterol supplementation (to ensure that all sterols are cycloartenol metabolites). MSY4 reproducibly grew under these conditions, establishing the genetic complementation. In addition, an extract of yeast grown under these conditions was shown by chromatographic and spectroscopic analysis to contain ergosterol, confirming that the intended pathway had been constructed (Figure 4.5).

MSY4 is a novel yeast strain that incorporates portions of the plant sterol biosynthetic pathway. The unique characteristics of this metabolically engineered strain make it a useful tool for selection/cloning of cycloartenol synthases from a wide variety of organisms. Furthermore, it is also suitable for any experiments that require detection of cycloartenol synthase activity (such as mutagenesis and DNA shuffling).

**Characterization of the *Arabidopsis thaliana* Lupeol Synthase**

The *Arabidopsis thaliana* lupeol synthase (AthLUP1) was originally isolated by Herrera and co-workers, who partially characterized its activity by identifying lupeol as the major enzymatic product in addition to a small amount of β-amyrin. However, they also suspected several other minor components were formed. I, therefore, undertook experiments to rigorously characterize the product profile of this enzyme.
Figure 4.5. Sterol biosynthetic pathway in yeast strain MSY4. The novel yeast strain MSY4 is a yeast lanosterol synthase mutant that expresses the Arabidopsis cyclopropyl-sterol isomerase. This strain converts the plant sterol precursor cycloartenol to the essential yeast sterol ergosterol.

The lupeol synthase was expressed in the yeast lanosterol synthase mutant SMY8\textsuperscript{30} and grown on liter scale. The cells were then harvested and lysed and the resultant homogenate was incubated with racemic oxidosqualene (400 mg). The reaction was monitored by TLC and two chromatographically distinct product bands (one with a chromatographic mobility typical of triterpene alcohols and another more polar band) were purified by column chromatography. A portion of the triterpene alcohol fraction was acetylated and was then purified further by argentation-HPLC.\textsuperscript{54} Triterpene alcohols can be very difficult to purify one from another using silica gel chromatography because these are isomeric compounds that frequently differ only in the location of a double bond.
Argentation-chromatography is a very powerful and useful method that can separate compounds such as those based on double bond character. Electron rich double bonds coordinate to the silver ions, and compounds are bound to the column more or less tightly based on the chemical environment of the double bond (substituents, conjugation, steric, etc.). Three pure triterpene alcohols were isolated from argentation-HPLC purification and two other compounds were isolated as a mixture.

These compounds were analyzed by GC and GC-MS and had parent masses consistent with acetylated derivatives of $C_{30}H_{50}O$ triterpene alcohols. The compound structures were determined by a combination of NMR experiments ($^1H$, $^{13}C$, DEPT, COSYDEC, HSQC, and HMBC) and comparison with chemical shift data found in the literature. The acetates of germanicol, taraxasterol, and $\psi$-taraxasterol were identified, establishing the parent free alcohols as enzymatic products in addition to the previously established lupeol and $\beta$-amyrin.

The polar component of the enzymatic product was purified by silica gel chromatography and GC-MS analysis established a parent mass consistent with $C_{30}H_{52}O_2$. A similar battery of NMR experiments ($^1H$, $^{13}C$, DEPT, COSYDEC, HSQC, and HMBC) established the structure as $3\beta,20$-dihydroxylupane.

The total composition of the product mixture was determined by a combination of GC-FID and NMR integration. The relative amounts of the triterpene alcohols were determined by integration of GC-FID signals from the acetylated triterpene alcohol fraction. The proportion of triterpene diol was determined by integration of characteristic and well-resolved NMR signals for both lupeol and $3\beta,20$-dihydroxylupane in a crude extract of total enzymatic product. The *Arabidopsis thaliana* lupeol synthase converts
oxidosqualene to lupeol (39%), 3β,20-dihydroxylupane (39%), β-amyрин (7%),
germanicol (8%), taraxasterol (4%), and ψ-taraxasterol (3%). These products are shown
in Figure 4.6.

Figure 4.6. *Arabidopsis thaliana* lupeol synthase generates six different triterpene
products. The native product diversity of this enzyme demonstrates that dedicated
enzymatic activities are not required to account for each discreet oxidosqualene
cyclization product identified in nature. Furthermore, 3β,20-dihydroxylupane formation
demonstrates that triterpene diols can be direct oxidosqualene cyclase products.
This diverse product profile was unknown in oxidosqualene cyclases, but does have precedent in the lower terpene synthases where several enzymes have been shown to make multiple products.\textsuperscript{72,73} It is worth noting that it is unlikely that the enzyme is catalyzing interconversion between products as a diversity of products is observed. If interconversion were occurring, the product composition would converge to one or two thermodynamically favored structures. Although it is currently unknown what selective advantage a multi-functional enzyme might provide for \textit{Arabidopsis}, several interesting possibilities exist. Most triterpene structural diversity is generated by plants, which are believed to use these products to defend against infection, infestation, and herbivory. Generating the most diverse possible array of defense compounds from a single enzyme could confer a significant competitive advantage when defending against a broad array of potential threats. Mixtures of compounds may also have useful physical characteristics. Plants have a waxy layer over their leaves and stems that facilitates water retention, and a mixture of triterpene structures would be less prone to crystal formation, maintaining the physical integrity of this water barrier.

In addition to the unprecedented product diversity generated by this single oxidosqualene cyclase, this is also the first demonstration that triterpene diols can be direct oxidosqualene cyclase products. Several natural triterpene diols are known.\textsuperscript{74-76} Although hydration of a triterpene olefin remains a plausible origin for some compounds, the formation of a triterpene diol by the \textit{AthLUP1} clearly demonstrates that their formation directly from the cyclization reaction is plausible.

There are over 80 oxidosqualene cyclase products identified in nature,\textsuperscript{1} yet only very few of the enzymes responsible for their biosynthesis are known. The \textit{Arabidopsis}
lupeol synthase provides a demonstration that each structure does not imply a single enzyme dedicated to its formation. This enzyme certainly does not preclude the existence of those dedicated enzymes, but simply establishes that fewer than 80 enzymes could generate the over 80 known oxidosqualene cyclase structures.

**DNA SHUFFLING OF CYCLOARTENOL SYNTHASE AND LUPEOL SYNTHASE**

Cycloartenol synthase folds oxidosqualene into a pre-chair-boat-chair conformation, cyclizes it to the protosteryl cation intermediate, and then promotes another series of steps leading to cycloartenol formation. In contrast, lupeol synthase folds oxidosqualene into a pre-chair-chair-chair conformation and consequently forms the dammarenyl cation intermediate *en route* to lupeol. How the oxidosqualene cyclases control this substrate-folding step of the reaction sequence is not known. I undertook experiments to determine what motifs in these enzymes control the two different substrate folding conformations and the resultant cyclization steps.

Because of the global changes that are likely required to accommodate the two different oxidosqualene conformations, it is very probable that multiple changes within the enzyme active site are necessary to control how folding takes place. The numerous mutations to convert one kind of enzyme into another precluded the use of both site-directed and random mutagenesis approaches. Making a large number of site-directed mutants and combinations thereof is cumbersome and there is insufficient structural information to guide site-directed mutagenesis experiments. Random mutagenesis was similarly unattractive. Although random mutagenesis can readily generate multiple
mutations, the extreme mutagenic load required to achieve a suitable number of beneficial mutations would probably cause too many deleterious changes.

Therefore, a DNA shuffling strategy was adopted because it allows the generation of large, randomly mutated libraries without the heavy mutagenic load required of traditional random mutagenesis. DNA shuffling is a homology-based PCR strategy that requires the fragmentation of different genes, fragment mixing, and primerless PCR assembly of those fragments to create a random population of full-length, hybrid genes.\textsuperscript{77,78} In the primerless PCR reaction, gene fragments with sufficient identity anneal to one another and then serve as "primed" DNA species for the primerless PCR reaction. Extension of the primed DNA species creates a crossover point where the DNA sequence of the amplicon changes from one gene to the other; when performed iteratively in the PCR reaction, a semi-random distribution of these crossover points arise over the length of the coding sequence (Figure 4.7).

The shuffled library can then be put through iterative rounds of selection for those genes that complement the sterol auxotrophy of a suitable expression host such as MSY4. We sought to replace the maximum amount of cycloartenol synthase sequence with lupeol synthase sequence while maintaining cycloartenol synthase activity. Hybrid, functional genes isolated from these shuffling and selection experiments would reveal what portions of the cycloartenol synthase sequence are required for cycloartenol synthase activity (and protosterol cation formation). Regions of the cycloartenol synthase that can be substituted with lupeol synthase without compromising activity are not required for cycloartenol formation. Conversely, those portions of the cycloartenol
synthase that cannot be functionally replaced by lupeol synthase sequence may have a catalytic role specifically required for cycloartenol biosynthesis.

![Diagram of DNA shuffling procedure](image)

**Figure 4.7. DNA shuffling procedure.** Different genes are fragmented by DNase I digestion and small fragments are recovered and mixed. The fragments anneal to one another in a primerless PCR reaction and are extended to create randomly generated, full-length hybrid genes. The hybrid genes are then amplified and subcloned for screening or selection experiments.

Because DNA shuffling is a homology-based strategy, we elected to shuffle the *Arabidopsis thaliana* cycloartenol synthase\(^\text{16}\) (*AthCAS1*) and the *Olea europaea* lupeol synthase\(^\text{47}\) (*OeuLUP*). Of the cloned and characterized oxidosqualene cyclases, this pair of a protosteryl-type cyclase and dammarenyl-type cyclase share the highest degree of nucleotide sequence identity (59%). Although DNA shuffling of genes with 59% homology is possible, higher homology facilitates heterologous assembly. We therefore synthesized the lupeol synthase where every silent nucleotide was changed to match the
AthCAS1 sequence. The synthetic lupeol synthase amino acid sequence remained unchanged, but the nucleotide homology between the Arabidopsis cycloartenol and synthetic lupeol synthase increased to 79% (Table 4.1). Seventy oligos (35 forward and 35 reverse) corresponding to the newly designed lupeol synthase sequence were assembled in a primerless PCR reaction similar to that described for DNA shuffling experiments. Amplicons were sub-cloned into an appropriate vector, recovered, and sequenced. A full-length clone with the correct sequence was identified and carried on for further experiments. We found that gel purification of the oligos was critical for timely gene synthesis. Many literature reports state that purified DNA oligomers are not essential\textsuperscript{77,78} but that is true only in cases where there is an accessible and convenient functional assay for the integrity of the gene.

<table>
<thead>
<tr>
<th>Nucleotide homology of native AthCAS1 and native OeuLUP</th>
<th>59%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide homology of native AthCAS1 and synthetic OeuLUP</td>
<td>79%</td>
</tr>
<tr>
<td>Amino acid homology of native AthCAS1 and native OeuLUP</td>
<td>63%</td>
</tr>
<tr>
<td>Amino acid homology of native AthCAS1 and synthetic OeuLUP</td>
<td>63%</td>
</tr>
<tr>
<td>Number of amino acids in the AthCAS1 and OeuLUP proteins</td>
<td>758</td>
</tr>
<tr>
<td>Number of amino acid differences between AthCAS1 and OeuLUP</td>
<td>281</td>
</tr>
</tbody>
</table>

Table 4.1. Comparison of AthCAS1 and OeuLUP used for DNA shuffling. A customized lupeol synthase was designed and synthesized to achieve maximum nucleotide identity with the AthCAS1 without altering the protein sequence. Some of the 281 residue differences between the two enzymes determine the substrate folding conformation.

The AthCAS1 and synthetic lupeol synthase genes (both in the pBluescript vector) were fragmented by DNase I digestion and appropriate size fragments (~30-200 bp) were isolated by gel purification. Those fragments were then combined (in varying
proportions) and assembled in a primerless PCR reaction. PfuTurbo (Stratagene) was empirically determined to be the best polymerase for the assembly reactions and three rounds of assembly were typically required to generate full-length hybrid genes. The assembled PCR product then served as the template for a traditional primer-based PCR reaction (using vector based primers) to amplify the shuffled library for sub-cloning and subsequent selection. The PCR product was then purified and digested using a proteinase K clean-up procedure\(^7\) and ligated into the inducible yeast expression vector pRS426GAL.\(^7\) The library was transformed and propagated in high efficiency, chemically competent DH5α.

The cycloartenol synthase / lupeol synthase shuffled library was transformed into the yeast strain MSY4. Most of the transformations were plated directly on galactose media and supplemented only with heme, conditions which immediately select for cycloartenol biosynthesis (in many cases, transformants are isolated on media selective for plasmid incorporation only and are then replica plated onto media appropriate for gene expression studies).

Sterol-independent transformants were identified and re-streaked again on galactose media supplemented with heme (no ergosterol) to ensure they maintained their sterol independent phenotype. Those transformants were also extracted and checked for ergosterol content (ergosterol is indicative of an active cyclase with some degree of cycloartenol or lanosterol biosynthetic ability). The plasmids containing the hybrid cyclases were recovered and sequenced.

Two clones with the greatest amount of sequence exchange were named pMS14.3 and pMS19.3. Several other shuffled clones were identified but their crossover regions
were smaller and were already included in the pMS14.3 and pMS19.3 constructs. Both pMS14.3 and pMS19.3 retained mostly cycloartenol synthase sequence (~80% in each). In pMS14.3, 140 amino acid residues at the C-terminal end of the enzyme were lupeol synthase sequence. The pMS19.3 construct had 115 amino acids of the N-terminal end of the protein encoded by lupeol synthase sequence. A diagram of their crossover points and sequence composition are shown below (Figure 4.8). This round of DNA shuffling actually produced chimeras (at least among those hybrid genes that complemented) although that was not the intention. The formation of chimeras probably resulted from poor control of fragment size. If the shuffle fragments were rather large relative to the size of the entire gene, only a few crossover points would arise.

GC and GC-MS established that both pMS14.3 and pMS19.3 derived hybrid enzymes produce only cycloartenol, showing that the termini of cycloartenol synthase lacks motifs and structural elements that are unique to cycloartenol biosynthesis. Catalytically essential positions may reside in the regions substituted with lupeol synthase sequence, but if so those positions must participate in steps common to both lupeol and cycloartenol biosynthesis (i.e., proton-initiated epoxide opening). We then wanted to determine whether a combination of pMS14.3 and pMS19.3 would still retain cycloartenol synthase activity when both termini were substituted with lupeol synthase sequence.

A gene incorporating the relevant features of both pMS14.3 and pMS19.3 was constructed and named pMS23.3. This construct was transformed into MSY4 and the cycloartenol synthase / lupeol synthase hybrid enzyme encoded by the pMS23.3 plasmid was shown to genetically complement the sterol auxotrophy as described above. The
ability of this enzyme to convert oxidosqualene to cycloartenol was also confirmed by *in vitro* assays and product characterization (GC and GC-MS). A diagram of the pMS23.3 encoded protein is shown below (Figure 4.8).

![Diagram](image)

**Figure 4.8. Hybrid constructs recovered from DNA shuffling and selection experiments.** Hybrid cycloartenol synthases with lupeol sequence at the N- and C-termini of the proteins were isolated by genetic complementation of MSY4. pMS23.3 was constructed and shown to have cycloartenol synthase activity. Diagrams are not drawn to scale.

The first round of cycloartenol synthase / lupeol synthase DNA shuffling has provided important information about the cycloartenol synthase and it has also narrowed the number and location of residues that may control protosteryl versus dammareneryl cation formation. Of the original 281 candidate *AthCAS1* residues (the number of amino acid differences between the native cycloartenol and lupeol synthases) that may control the different substrate folding steps in cycloartenol synthase, 108 (38%) of those have
been eliminated (Figure 4.9). Cycloartenol synthase activity is retained in the pMS23.3 construct although 108 AthCAS1 positions have been replaced with lupeol synthase residues. Furthermore, the *Olea europaea* lupeol synthase and native *Arabidopsis thaliana* cycloartenol synthase share only 63% amino acid identity, but the cycloartenol synthase encoded by the pMS23.3 hybrid gene and the native *Olea europaea* lupeol synthase now share 77% amino acid identity. In addition, the pMS23.3 hybrid cycloartenol synthase will be a valuable component in subsequent rounds of shuffling because it still retains cycloartenol biosynthetic ability but now shares even higher homology with the lupeol synthase, thus aiding crossover generation and elimination of non-essential cycloartenol synthase sequence.

**Sequence Swapping of AthCAS1 and OeuLUP C-terminal Sequences**

The DNA shuffling experiments described above revealed that rather large portions of the cycloartenol synthase could be replaced with lupeol synthase sequence and still retain cycloartenol biosynthetic ability. In the pMS14.3 and pMS23.3 constructs, 140 C-terminal amino acids were replaced with lupeol synthase residues except for a small 17 amino acid portion (*AthCAS1* positions 720-736).

Anticipating that the small "island" of cycloartenol synthase sequence was not critical for catalysis, I decided to exchange that *AthCAS1* sequence with lupeol synthase sequence using a shared *Hind* III site several hundred base pairs upstream of the C-terminal crossover point (Figure 4.10). To remove the CAS island sequence and to also insert additional lupeol synthase sequence up to the *Hind* III site, I used this site to subclone the 3'-end of the lupeol synthase onto the pMS23.3 construct. The resultant clone, pMS47.3, was tested for its ability to genetically complement the yeast strain MSY4.
Figure 4.9. Protein sequence alignment of cycloartenol and lupeol synthase. Positions where the cycloartenol and lupeol synthase sequences differ (281 positions) are boxed in black. The underlined sequence denotes sections of AthCAS1 that can be functionally replaced with OeuLUP sequence.
Surprisingly, the newly generated construct was not able to complement the sterol auxotrophy of MSY4, indicating that it had lost cyclase activity.

Two different portions of the pMS23.3 sequence were changed in construction of the pMS47.3, and one of them (or both) resulted in a loss of activity. Therefore, constructs were prepared in which only one of the two possible changes were present in a single enzyme. pMS56.9 had only the island sequence substituted with lupeol synthase sequence. The pMS55.9 was constructed so that it retained cycloartenol synthase sequence in that island but otherwise included lupeol synthase sequence up to the Hind III site (Figure 4.10).

A technical detail is important to note here. During these experiments, I grew concerned that the hybrid enzymes would have reduced activity and would consequently not complement MSY4. Without suitable sensitivity in the genetic complementation, these experiments could not successfully identify functional hybrid enzymes. I therefore re-designed some of the hybrid constructs for genetic selection in SMY8 (lanosterol biosynthesis) rather than MSY4 (cycloartenol biosynthesis). Genetic complementation of SMY8 by lanosterol biosynthesis is more robust and less complex than complementation of MSY4 with cycloartenol formation (which requires adequate activity in both the cycloartenol synthase and cyclopropyl-sterol isomerase enzymes). Lanosterol biosynthetic ability was “added” to the hybrid constructs by making three point mutations of catalytically important AthCAS1 residues. The AthCAS1 Tyr410Thr, His477Gln, and Ile481Val mutations, discussed throughout this thesis, were incorporated in the pMS55.9 and pMS56.9 constructs described in the preceding paragraph. Thus, domain swapping
of the cycloartenol and lupeol synthase still helps identify catalytically important regions as the experiment was originally designed. However, the three point mutations cause lanosterol formation rather than cycloartenol formation and allow SMY8 to be used as the selection host. These modified constructs (Figure 4.10) were therefore tested for lanosterol biosynthetic ability in SMY8 complementation experiments.

![Diagram](image)

**Figure 4.10.** The *AthCAS1* C-terminal "Island" is required for cyclase activity. Several constructs were generated to determine if the island of C-terminal *AthCAS1* sequence in pMS23.3 is required for activity. Only constructs that possess cycloartenol synthase sequence in that region retain activity. The asterisks denote the three *AthCAS1* Tyr410Thr, His477Gln, and Ile481Val mutations that were incorporated to allow for genetic selection of lanosterol (rather than cycloartenol) biosynthetic ability. Diagrams are not drawn to scale.
Interestingly, only the pMS55.9 construct, which leaves the C-terminal *Ath*CAS1 island intact, retains activity and is able to complement the sterol auxotrophy of SMY8. These findings demonstrate that that 17 amino acid stretch is required for cyclase activity in these hybrid constructs. Additional experiments were performed to determine if specific amino acid changes with in that region were responsible for the apparent significance of that 17 amino acid span.

**CHARACTERIZATION OF *Ath*CAS1 ASN727 AND ILE732 MUTANTS**

Functional analysis of *Ath*CAS1 and *Oeu*LUP hybrid cyclases (described above) suggested that 17 amino acids (*Ath*CAS1 720-736) of C-terminal sequence were important for cyclase activity. Protein alignments of cloned and characterized oxidosqualene cyclases revealed two specific positions in this region that are strongly and differentially conserved in the protosteryl-type and dammarenyl-type cyclases.

Most protosteryl-type cyclases conserve an asparagine at the position corresponding to *Ath*CAS1 Asn427 whereas the dammarenyl-type cyclases maintain a methionine. Similarly, the protosteryl-type cyclases strictly conserve an isoleucine and the dammarenyl-type cyclases possess a leucine at the position corresponding to *Ath*CAS1 Ile732 (Figure 4.11). The conservation pattern at these two positions indicates that they are good targets for site-directed mutagenesis.

The *Ath*CAS1 Asn727Ala, Asn727Met, Ile732Ala, and Ile732Leu mutants were made to determine if these two positions have any catalytic significance in the *Ath*CAS1 enzyme. They were constructed in pRS305GAL, sequenced, and transformed into the yeast lanosterol synthase mutant SMY8. Transformants were selected on synthetic
complete media lacking leucine and supplemented with heme and ergosterol. Product profiles of each of the mutants were determined by in vitro assay and product characterization (GC and GC-MS). Unfortunately, none of the mutants showed altered product profiles as they all made cycloartenol as their sole product (GC and GC-MS).

Figure 4.11. Protein sequence alignment of oxidosqualene cyclase sequences corresponding to the AthCAS1 C-terminal island. Two positions (AthCAS1 Asn727 and Ile732) are conserved among the oxidosqualene cyclases in the region corresponding to the AthCAS1 island identified in hybrid enzyme experiments. Protosteryl-type cyclases are represented by lanosterol synthases (ERG7) from S. cerevisiae\textsuperscript{25,26} (Sce), Candida albicans\textsuperscript{27,28} (Cal), Schizosaccharomyces pombe\textsuperscript{30} (Spo), Rattus norvegicus\textsuperscript{31,32} (Rno), and Homo sapiens\textsuperscript{33,34} (Hsa) and cycloartenol synthases (CAS1) from D. discoideum\textsuperscript{24} (Ddi), Pismum sativum\textsuperscript{20} (Psa), Panax ginseng\textsuperscript{21} (Pgi), Glycyrrhiza glabra\textsuperscript{22} (Ggl), Luffa cylindrica\textsuperscript{23} (Lcy), and A. thaliana\textsuperscript{15} (Ath). Dammarenyl-type cyclases are represented by amyrin and lupeol synthases from Olea europaea\textsuperscript{47} (Oeu), Taraxacum officinale\textsuperscript{47} (Taf), Pismum sativum\textsuperscript{80} (Psa), Panax ginseng\textsuperscript{21,81} (Pgi), and A. thaliana\textsuperscript{3} (Ath).

The accurate cycloartenol formation in the four AthCAS1 mutants described above demonstrates that in isolation, the AthCAS1 Asn727 and Ile732 residues are not catalytically relevant. The catalytic behavior of these mutants provide indirect evidence
that the 17 amino acid span identified in the hybrid protein experiments does not specifically influence the distinctive lupeol or cycloartenol synthase reactions, but rather influences the cyclase in a non-specific manner. Based on the sequence alignment shown in Figure 4.11, the *AthCAS1* Asn727 and Ile732 are the most likely catalytically relevant residues; however, mutating these positions does not alter catalysis in the single mutants. The apparent importance of this region in the hybrid protein work therefore probably results from protein folding or protein stability problems. The C-terminal *AthCAS1* island may specifically interact with another portion of the protein, and when it is removed, the global structure of the protein is changed and catalytic activity abolished.
CONCLUSIONS

Oxidosqualene cyclases convert oxidosqualene to a wide variety of structurally diverse compounds with remarkable accuracy. The manner in which these enzymes are able to promote specific cyclization-rearrangement-deprotonation sequences is not fully understood. Experiments described in this thesis have identified residues that influence the distinct deprotonation steps carried out by lanosterol and cycloartenol synthase. Mutations of SceERG7 Thr384 and AthCAS1 His477 compromised the product specificity of the enzymes and resulted in an array of aberrant deprotonation products. These results are generally consistent with other findings that point mutations typically broaden product specificity. Interestingly, Asn and Gln mutations at AthCAS1 H477 yielded dramatically different product profiles from one another and from wild-type, and suggest that His477 may be an active site residue.

No success has been reported in generating oxidosqualene cyclases that are specific for a product distinct from that formed by the parent enzyme. Combinations of relevant oxidosqualene cyclase mutations have in some cases biased formation towards one new product but it has always been accompanied by significant amounts of by-product. In an effort to achieve product specificity by mutagenesis and to observe how different catalytically important mutant enzymes interact, we combined AthCAS1 His477Asn and His477Gln mutations with the known Tyr410Thr and Ile481Val mutations. Surprisingly, the His477 mutations (which have dramatic catalytic influence as the single mutants) showed no influence on catalysis in the Tyr410Thr Ile481Val background. These findings are the first demonstration of the mutually exclusive catalytic influence of specific oxidosqualene cyclase mutations and they provide some
insight into the roles specific amino acids may serve within the enzyme. Furthermore, understanding which mutations show synergistic influence and which ones show hierarchical influence is an important advance towards rational, predictive protein design of cyclases with non-natural activities.

Oxidosqualene folding is another important step in oxidosqualene cyclase biocatalysis that has important consequences for product structure. DNA shuffling experiments were initiated to determine how some cyclases fold oxidosqualene into one conformation while other cyclase fold it in a different conformation. Initial results from this study demonstrated that the termini of \( \text{AthCAS1} \) are not required for cycloartenol biosynthesis and structural elements specific to cycloartenol formation (and characteristic substrate folding) are located within the internal portions of the sequence. Additional iterations of DNA shuffling will be required to identify where those elements are located.

Many questions still surround the complex chemical transformations catalyzed by oxidosqualene cyclases, but rapidly evolving mutagenesis and directed evolution techniques are powerful tools that can be used to answer those questions. As a comprehensive understanding of these enzymatic transformations becomes established, exciting opportunities will arise in which the oxidosqualene cyclases can be exploited to generate inaccessible, novel, and valuable biocatalysts and triterpene compounds.
REFERENCES


(62) Hartmann-Bouillon, M. A.; Benveniste, P. *Phytochemistry* 1978, 17, 1037-1042.


(64) Schmitt, P.; Benveniste, P.; Leroux, P. *Phytochemistry* 1981, 20, 2153-2159.


(81) Kushiro, T.; Shibuya, M.; Ebizuka, Y. In *Towards Natural Medicine Research in the 21st Century, Excerpta Medica International Congress Series 1157*; Ageta,
H., Aimi, N., Ebizuka, Y., Fujita, T., Honda, G., Eds.; Elsevier Science:
### APPENDIX I: LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>alanine</td>
</tr>
<tr>
<td>Amp</td>
<td>ampicillin</td>
</tr>
<tr>
<td>Asn</td>
<td>asparagine</td>
</tr>
<tr>
<td>AthCAS1</td>
<td><em>Arabidopsis thaliana</em> cycloartenol synthase</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>COSYDEC</td>
<td>decoupled $^1\text{H}-^1\text{H}$ correlation NMR spectroscopy</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleic acid triphosphates</td>
</tr>
<tr>
<td>DEPT</td>
<td>distortionless enhancement by polarization transfer</td>
</tr>
<tr>
<td>dI H$_2$O</td>
<td>deionized water</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>EtOAc</td>
<td>ethyl acetate</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>GC</td>
<td>gas chromatography</td>
</tr>
<tr>
<td>GC-MS</td>
<td>gas chromatography-mass spectrometry</td>
</tr>
<tr>
<td>Gln</td>
<td>glutamine</td>
</tr>
<tr>
<td>Gly</td>
<td>glycine</td>
</tr>
<tr>
<td>GTAE</td>
<td>40 mM tris base, 20 mM acetic acid, 1 mM EDTA, 1 mM guanosine</td>
</tr>
<tr>
<td>His</td>
<td>histidine</td>
</tr>
<tr>
<td>HMBC</td>
<td>heteronuclear multiple bond correlation</td>
</tr>
<tr>
<td>HSQC</td>
<td>heteronuclear single-quantum coherence</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Ile</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>Leu</td>
<td>Leucine</td>
</tr>
<tr>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>OeuLUP</td>
<td><em>Olea europaea</em> lupeol synthase</td>
</tr>
<tr>
<td>OS</td>
<td>(S)-2,3-oxidosqualene</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>SCE</td>
<td>1 M sorbitol, 0.1 M sodium citrate (pH 7.0), 60 mM EDTA</td>
</tr>
<tr>
<td>SceERG7</td>
<td><em>Saccharomyces cerevisiae</em></td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>T8</td>
<td>10 mM tris-HCl, pH 8</td>
</tr>
<tr>
<td>TAE</td>
<td>40 mM tris base, 20 mM acetic acid, 1 mM EDTA</td>
</tr>
<tr>
<td>TE8</td>
<td>10 mM tris-HCl (pH 8), 1 mM EDTA</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>TMS</td>
<td>Trimethylsilane</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>Tris(hydroxymethyl)aminomethane hydrochloride</td>
</tr>
<tr>
<td>Trp</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>Ura</td>
<td>Uracil</td>
</tr>
<tr>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>Media</td>
<td>Composition</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>LB</td>
<td>1% tryptone, 0.5% yeast extract, 0.5% NaCl</td>
</tr>
<tr>
<td>YPD</td>
<td>1% yeast extract, 2% peptone, 2% dextrose</td>
</tr>
<tr>
<td>YPG</td>
<td>1% yeast extract, 2% peptone, 2% galactose</td>
</tr>
<tr>
<td>SCD-Ura</td>
<td>1% ammonium sulfate, 0.34% yeast nitrogen base, 0.4% amino acid mix lacking uracil, 2% dextrose</td>
</tr>
<tr>
<td>SCD-Leu</td>
<td>1% ammonium sulfate, 0.34% yeast nitrogen base, 0.4% amino acid mix lacking leucine, 2% dextrose</td>
</tr>
<tr>
<td>SCG-Ura</td>
<td>1% ammonium sulfate, 0.34% yeast nitrogen base, 0.4% amino acid mix lacking uracil, 2% galactose</td>
</tr>
<tr>
<td>H</td>
<td>13 μg/mL heme</td>
</tr>
<tr>
<td>E</td>
<td>20 μg/mL ergosterol, 5 μg/mL Tween 80</td>
</tr>
</tbody>
</table>
APPENDIX II: LIST OF STRAINS

**Escherichia coli Strain**

DH5α: $F^{+}$80dlacZDM15 $\Delta$(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(rκ, mκ+)
phoA supE44 λ thi-1 gyrA96 relA1

**Saccharomyces cerevisiae Strains**

JBY575: MATa ura3-52 trp1-Δ63 leu2-3,112 his3-Δ200 ade2 GAL⁺

SMY8: MATa erg7::HIS3 hem1::TRP1 ura3-52-trpl-Δ63 leu2-3,112 his3-Δ200 ade2

GAL⁺

LHY3: MATa erg7::leu2::hisG-URA3-hisG erg9::HIS3 hem1::TRP1 ura3-52-trpl-Δ63

leu2-3,112 his3-Δ200 ade2 GAL⁺

LHY4: MATa erg7::leu2::hisG erg9::HIS3 hem1::TRP1 ura3-52-trpl-Δ63