RICE UNIVERSITY

Mutagenesis of Cycloartenol Synthase and Lanosterol Synthase: Broadening and Narrowing Product Profile

by

Silvia Lodeiro

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

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

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

Jonathan Silberg, Assistant Professor
Department of Biochemistry and Cell Biology

HOUSTON, TEXAS
MAY 2006
Copyright

Silvia Lodeiro

2006
ABSTRACT

Mutagenesis of Cycloartenol Synthase and Lanosterol Synthase:

Broadening and Narrowing Product Profile

By

Silvia Lodeiro

This thesis describes mutagenesis experiments in cycloartenol synthase and lanosterol synthase that allowed the identification of important catalytic residues necessary to broaden and narrow product profiles. These results provided insights into factors critical for redesigning enzyme function.

Mutagenesis in Arabidopsis thaliana cycloartenol synthase revealed that His477 is an essential component of the catalytic distinction between cycloartenol synthase and lanosterol synthase. Mutations at position 477 abolish cycloartenol biosynthesis, and subtle structural changes at this position dramatically alter product profile. The His477Asn mutant produces lanosterol as its major product (88%) whereas His477Gln produces primarily parkeol (73%). His477 influences deprotonation more strongly than any of the previously studied catalytic residues, but changes at this position are catalytically irrelevant in the presence of Tyr410Thr and Ile481Val mutations. The His477Asn Tyr410Thr Ile481Val and His477Gln Tyr410Thr Ile481Val triple mutants have the same product profile as the Tyr410Thr Ile481Val double mutant. Homology modeling studies established that His477 is a second-sphere residue that affects catalysis indirectly through its interactions with the active-site residue Tyr410. Changes at His477
strongly affect the location, orientation, and electronics of the Tyr410 side chain.

Efforts to modify the catalytic specificity of enzymes consistently show that it is easier to broaden the substrate or product specificity of an accurate enzyme than to restrict the selectivity of one that is promiscuous. Careful examination of the homology model allowed the identification of a combination of mutations necessary to redesign cycloartenol synthase into a highly accurate lanosterol synthase. A double mutant was constructed and characterized and was shown to cyclize oxidosqualene accurately to lanosterol (99%). This catalytic change entailed both relocating polarity with a His477Asn mutation and modifying steric constraints with an Ile481Val mutation, and is among the best examples of redesigning an enzyme to accurately generate a new product.

Known lanosterol synthase mutants make monocyclic or tetracyclic byproducts from oxidosqualene. Mutation of *Saccharomyces cerevisiae* lanosterol synthase at Tyr510 caused partial substrate misfolding and generated a tricyclic byproduct. This novel triterpene, (13αH)-isomalabarica-14(27),17E,21-trien-3β-ol, is the putative biosynthetic precursor of the isomalabaricane triterpenoids in sponges. The results suggest the facile evolution of tricyclic terpenoids as secondary metabolites in sponges.
Acknowledgments

I would like to thank my research advisor, Prof. Seiichi P. T. Matsuda for his guidance and mentoring, and for believing that an old school organic chemist could learn molecular biology and get a kick out of tinkering with enzymes and DNA. I would also like to thank him for encouraging me to write and for helping me with my English. Thanks to Prof. Ron Parry for being in my Thesis Committee and for all the thoughts and advice he gave me over these years and for his patience to sit through all my seminars and stay awake. Thanks to Prof. Joff Silberg for his valuable advice on my manuscript on Enzyme Redesign and for accepting to integrate my Thesis Committee.

I would like to thank Dr. Tanja Schulz-Gasch, our collaborator at Hoffmann-La Roche Ltd., for developing the homology model which made possible very nice results described in this thesis.

I am grateful to Prof. Bonnie Bartel who gave me valuable advice on molecular biology techniques whenever I had trouble, and to Prof. James Mc New for letting me use the cell disruptor and for being very kind and understanding every time it was not working.

I would like to thank Dr. William K. Wilson, for being a great thought partner, and friend. Bill’s professional and personal advice over the years was invaluable and his help with NMR analysis was key to the results obtained.

I thank Dr. Hui Shan for her support, encouragement, friendship, and for her valuable professional help with HPLC and LC-MS.

I want to thank other members in the Matsuda Lab: Gia Fazio for her friendship
and for offering a shoulder to lean on and being a good listener; Dr. Quanbo Xiong for sharing his knowledge and giving valuable advice; Caroline Pardue for helping with the proof reading of my manuscripts and for always having good supplies of chocolate every time I needed some; Dr. Uttam Dasgupta whose presence in the lab gave me the opportunity to feel helpful and be part of a different interesting project; Yulia Ivanova and Maria Kolesnikova for teaching me my first words in Russian.

I would like to thank former members of the lab: Dr. Renee Leclair with whom I shared not only the bench lab but also laughs, tears, stories and hopes dear to my heart. Renee also helped me learn all the DNA techniques and kept the lab well provided with good energizing music. Dr. Mike Segura with whom I shared my first projects and who taught me some of the procedures I used herein, Dr. Ran Xu for sharing her knowledge and friendship, and Michelle Meyer and Akash Patel for starting some of the projects.

Thanks to Prof. Patrick Moyna, Prof. Gustavo Seoane, Prof. Hugo Cerecetto and Prof. Karen Ofejevi at the chemistry department at Universidad de la Republica in Uruguay, for giving me the opportunity to join their faculty and make my first inroads into the world of organic chemistry.

I thank my husband Martin and my children Natalia and Santiago for their love and support through the ups and downs of my research, and for their understanding when I could not participate in family projects because of work. My parents Ramon and Marcelina who always encouraged me and reminded me I would succeed, and specially my mom who came all the way from Montevideo to help me take care of the kids during the thesis writing; and for those mouth watering Spanish recipes she cooked through that period. I thank my sister Isabel for supporting me from the distance.
There would not be enough pages to thank everybody and I am sure my memory may have missed to be fair with all. If so please forgive me.

Thank you Lord for making all these things possible.
To my family
# Table of Contents

Abstract

Acknowledgments

Dedication

List of Figures and Tables

Chapter 1: Introduction

Previous Mutagenesis Studies in Cycloartenol Synthase

Previous Mutagenesis Studies in Lanosterol Synthase

Squalene-Hopene Cyclase

Lanosterol Synthase Crystal Structure

Overview

Chapter 2: Results and Discussion

Part I: Cycloartenol Synthase Mutants

*At*hCAS1 His477Asn and *At*hCAS1 His477Gln Single Mutants

*At*hCAS1 Tyr410Thr His477Asn Ile481Val and *At*hCAS1 Tyr410Thr His477Gln Ile481Val Triple Mutants

*At*hCAS1 His477Asn Ile481Val and *At*hCAS1 His477Gln Ile481Val Double Mutants

Part II: Lanosterol Synthase Mutants

*Sce*Erg7 Tyr510His and *Sce*Erg7 Tyr510Phe Mutants

Chapter 3: Conclusions
Chapter 4: Experimental Procedures

Materials 58

Nuclear Magnetic Resonance (NMR) 58

Gas Chromatography (GC) 59

Gas Chromatography-Mass Spectrometry (GC-MS) 59

High-Performance Liquid Chromatography (HPLC) 60

UV Spectroscopy 60

Centrifugation 60

Incubators 60

Cell Lysis 61

Strains 61

Bacterial Media 61

Yeast Media 62

Oligonucleotides 63

Site-Specific Mutagenesis 63

Bacterial Transformation 65

Miniprep DNA Purification 65

Preparative DNA Purification 66

DNA Mapping 67

DNA Gel Electrophoresis Analysis 67

Preparative DNA Digest 68

DNA Ligation 68

DNA Sequencing 68
DNA Linearization 69
Yeast Transformation 69
Complementation Assay 71
Synthesis of (±)-2,3-Oxidosqualene 71
20 × (±)-2,3-Oxidosqualene Solution 72
20 × Tween 80 Solution 72
Small Scale in Vitro Assay 72
Large Scale in Vitro Assay 74
Large Scale in Vivo Assay 75
Column Chromatography 75
Derivatization of Triterpene Alcohols 76
Characterization of AthCAS1 His477Asn and AthCAS1
His477Gln Single Mutants 77
Characterization of AthCAS1 Tyr410Thr His477Asn Ile481Val
and AthCAS1 Tyr410Thr His477Gln Ile481Val Triple Mutants 79
Construction and Characterization of AthCAS1 His477Asn
Ile481Val and AthCAS1 His477Gln Ile481Val Double Mutants 82
Construction and Characterization of SceErg7 Tyr510His and
SceErg7 Tyr510Phe 86
References 97
Appendix A: List of Abbreviations 106
Appendix B: Sequencing primers 109
List of Figures and Tables

Figure 1.1 Conversion of oxidosqualene to membrane sterols in eukaryotes.

Figure 1.2 Cyclization of oxidosqualene by lanosterol synthase and cycloartenol synthase.

Figure 1.3 Conservation pattern of \textit{AthCAS1} Ile481.

Figure 1.4 Oxidosqualene cyclization products of \textit{AthCAS1} Ile481 mutants.

Table 1.1 Percentage product composition of \textit{AthCAS1} Ile481 mutants.

Figure 1.5 Conservation pattern of \textit{AthCAS1} Tyr410.

Table 1.2 Percent product composition of \textit{AthCAS1} and \textit{AthCAS1} mutants.

Figure 1.6 Oxidosqualene cyclization products of \textit{AthCAS1} Tyr410Thr mutants.

Table 1.3 Percentage product composition of \textit{SceErg7 Val454} mutants.

Figure 1.7 Oxidosqualene cyclization products of lanosterol synthase mutants.

Table 1.4 Percentage product composition of lanosterol synthase mutants.

Figure 1.8 \textit{Alicyclobacillus acidocaldarius} squalene-hopene cyclase catalyzes the cyclization of squalene to hopene and hopanol.

Figure 1.9 Ribbon representation of homodimeric squalene-hopene cyclase.

Figure 1.10 Active site of squalene-hopene cyclase with the inhibitor 2-azasqualene.

Figure 1.11 Two different views of a ribbon representation of human lanosterol synthase.

Figure 1.12 Active site of human lanosterol synthase with the reaction product lanosterol.

Figure 1.13 Stereo view of the active site of human lanosterol synthase with the
reaction product lanosterol.

Figure 2.1  Product profiles of wild-type *AthCAS1* and *AthCAS1* His477 single mutants.

Figure 2.2  Conservation pattern of His477.

Figure 2.3  Conservation pattern of Tyr410, His477 and Ile481.

Figure 2.4  Percentage product composition of *AthCAS1* and *AthCAS1* mutants.

Figure 2.5  *AthCAS1* homology model showing mutated residues and residues crucial for the deprotonation step.

Figure 2.6  Stereo representation of the *AthCAS1* homology model showing mutated residues and residues crucial for the deprotonation step.

Figure 2.7  Effects of mutations in position 477 on the orientation of Tyr410 side chain.

Figure 2.8  Cyclization of oxidosqualene to lanosterol and cycloartenol.

Figure 2.9  Percentage yields of oxidosqualene cyclization products in *AthCAS1* and *AthCAS1* mutants.

Figure 2.10  Superposition of wild-type *AthCAS1* with A) *AthCAS1* H477Q I481V and B) *AthCAS1* H477N I481V.

Figure 2.11  Tyr510 and other selected residues in the active site of lanosterol synthase.

Table 2.1  Percentage product composition of oxidosqualene cyclases.

Table 2.2  Comparison of in vivo and in vitro product ratios.

Figure 2.12  Cyclization of oxidosqualene by *SceErg7* and *SceErg7* Tyr510 mutants.

Figure 2.13  Proposed biosynthesis of isomalabaricane triterpenoids via isomalabaricatrienol.
Figure 4.1  TMS-derivatization of lanosterol.

Figure 4.2  Synthesis of (±)-2,3-Oxidosqualene.

Figure 4.3  $^1$H NMR spectrum of HPLC-purified isomalabaricatrienol.

Figure 4.4  $^{13}$C NMR and DEPT spectra of HPLC-purified isomalabaricatrienol.

Figure 4.5  $^1$H and $^{13}$C NMR signal assignments for isomalabaricatrienol.

Figure 4.6  $^1$H NMR spectrum of late fractions ($Rt$ 27 min) from the HPLC purification.

Figure 4.7  Electron-impact mass spectrum of underivatized isomalabaricatrienol.

Figure 4.8  Electron-impact mass spectrum of the TMS-derivative of isomalabaricatrienol

Figure 4.9  Total ion chromatogram from GC-MS analysis of underivatized isomalabaricatrienol, purified by HPLC.

Figure 4.10  Total ion chromatogram from GC-MS analysis of TMS-derivative of HPLC-purified isomalabaricatrienol
Chapter 1: Introduction

The cyclization of oxidosqualene (OS) and squalene to polycyclic triterpenes has fascinated and captivated scientists for more than fifty years.¹⁻¹³ These complex reactions occur with extraordinary regiospecificity and stereoselectivity. Oxidosqualene cyclases (OSCs) are a family of eukaryotic enzymes that convert acyclic (S)-2,3-oxidosqualene to over 100 cyclic triterpene alcohols (C₃₀H₅₀O) and triterpene diols (C₃₀H₅₂O₂).¹⁰ The reaction is initiated by epoxide protonation, followed by cation-olefin cyclization and cation rearrangement, and terminated either by deprotonation or water addition. Squalene cyclases (SCs) are a related enzyme family that convert acyclic squalene by similar mechanisms to over 60 cyclic triterpenes (C₃₀H₅₀) and triterpene alcohols (C₃₀H₅₂O).¹⁰ These enzymes control the reactivity of carbocations with precision unrivalled by nonenzymatic catalysts, but how they utilize steric bulk and polar groups to guide carbocation reactivity remains poorly understood. Structural diversity is achieved by different active-site geometries and by alternative rearrangement, deprotonation and water quenching steps.

The best studied member of the oxidosqualene cyclases family is lanosterol synthase (Erg7). Lanosterol synthase cyclizes (S)-2,3-oxidosqualene to lanosterol,¹⁴ the initial tetracyclic sterol precursor in fungi¹⁵ and animals¹⁶. Lanosterol is metabolized to ergosterol in fungi which serves as a cell membrane sterol and is essential for cell viability (Figure 1.1). Animals convert lanosterol to cholesterol which also serves as cell membrane component, and to steroid hormones. Cycloartenol synthase (CAS1) is a paralogous enzyme that cyclizes (S)-2,3-oxidosqualene to cycloartenol, an isomeric
pentacyclic sterol precursor in plants$^{17-19}$ and some protists$^{20-24}$. Plants transform cycloartenol to membrane sterols such as sitosterol and stigmasterol. These enzymes are only moderately related (~40% identical) but promote mechanistically similar reactions. Both enzymes cyclize (S)-2,3-oxidosqualene, pre-folded in a chair-boat-chair conformation, to the protosteryl cation and promote hydride and methyl shifts to form the

Figure 1.1. Conversion of oxidosqualene to membrane sterols in eukaryotes.
C-8 cation and then diverge by abstracting a different proton (Figure 1.2). Cycloartenol synthase promotes an additional hydride shift and terminates the reaction with cyclopropyl ring formation and deprotonation from C-19, whereas lanosterol synthase deprotonates from C-9. This pair of enzymes provides an ideal venue in which to investigate structure-function relationships in terpene cyclization because they catalyze closely related but distinct reactions, and numerous examples of each have been cloned and characterized from a broad variety of organisms.

![Diagram of cyclization process](image)

**Figure 1.2.** Cyclization of oxidosqualene by lanosterol synthase and cycloartenol synthase.

This thesis describes the use of mutagenesis experiments and homology modeling to identify catalytic residues that control cyclization and deprotonation in cycloartenol synthase and lanosterol synthase. The next two sections describe previous mutagenesis
studies of these enzymes that constituted the foundation for this thesis work. Then, an overview is given of the crystal structure of squalene-hopene cyclase (SHC) which is a related enzyme that was used as template for the homology model. The last section is dedicated to the crystal structure of lanosterol synthase which was reported later in the course of this work.

**Previous mutagenesis studies in cycloartenol synthase**

A series of mutagenesis experiments illuminated the roles of specific cycloartenol synthase residues by altering the catalytic capabilities of the enzyme. Hart et al.\textsuperscript{25} used directed evolution experiments in which *Arabidopsis thaliana* cycloartenol synthase\textsuperscript{26} (*AthCAS1*) was expressed in the yeast lanosterol synthase deletion mutant SMY8\textsuperscript{27} under ergosterol deprived conditions. Under these conditions, growth of the parent strain is limited by sterol deficiency. Although the cell can generate oxidosqualene and cyclize it to cycloartenol using the foreign enzyme, it cannot open the cyclopropyl ring to produce conventional tetracyclic sterols. Under these selective conditions a mutant enzyme arose that genetically complemented the mutation; sequencing of the recovered plasmid uncovered two mutations His158Gln and Ile481Val. They prepared the single mutants by site directed mutagenesis, and complementation experiments revealed that only the Ile481Val mutation conferred sterol independent growth. In vitro experiments and product characterization established that *AthCAS1* Ile481Val converts oxidosqualene to cycloartenol, lanosterol and parkeol (54:25:21). Ile481 is strictly conserved in all known cycloartenol synthases, and lanosterol synthases strictly conserve Val at the corresponding position (Figure 1.3). *AthCAS1* Ile481 corresponds to the active site
residue Asp374 in *Alicyclobacillus acidocaldarius* squalene-hopene cyclase\(^{28-30}\) (AacSHC) and therefore Hart et al. suggested that Ile481 is in the active-site of *AthCAS1*. The recently described X-ray structure of lanosterol synthase\(^3\) and homology modeling within this thesis confirmed the Ile481 location. Removing a methylene group in *AthCAS1* Ile481Val slightly enlarges the active site and compromises accurate cycloartenol formation allowing alternative deprotonation to yield lanosterol and parkeol.

\[\begin{array}{cccccccccccc}
\text{A} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{A} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{A} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{A} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{A} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{V} & \text{D} & \text{H} & \text{G} & \text{W} & \text{P} & \text{|} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{R} & \text{P} & \text{Q} & \text{A} & \text{W} & \text{Q} & \text{|} & \text{V} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{A} & \text{S} & \text{Q} & \text{S} & \text{W} & \text{Q} & \text{|} & \text{V} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{K} & \text{T} & \text{Q} & \text{G} & \text{Y} & \text{T} & \text{|} & \text{V} & \text{A} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{K} & \text{E} & \text{Q} & \text{G} & \text{Y} & \text{T} & \text{|} & \text{V} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{K} & \text{D} & \text{Q} & \text{G} & \text{Y} & \text{A} & \text{|} & \text{V} & \text{S} & \text{D} & \text{C} & \text{T} & \text{S} & \text{E} \\
\text{I} & \text{T} & \text{Q} & \text{G} & \text{Y} & \text{T} & \text{|} & \text{V} & \text{S} & \text{D} & \text{T} & \text{T} & \text{T} & \text{S} & \text{E} \\
\text{L} & \text{O} & \text{C} & \text{G} & \text{W} & \text{I} & \text{|} & \text{V} & \text{A} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\text{L} & \text{O} & \text{C} & \text{G} & \text{W} & \text{I} & \text{|} & \text{V} & \text{S} & \text{D} & \text{C} & \text{T} & \text{A} & \text{E} \\
\end{array}\]

Figure 1.3. Conservation pattern of *AthCAS1* Ile481. Isoleucine is strictly conserved in the known cycloartenol synthases (CAS1) from *Arabidopsis thaliana*\(^{26}\) (*Ath*), *Pisum sativum*\(^{32}\) (*Psa*), *Panax ginseng*\(^{33}\) (*Pgi*), *Glycyrrhiza glabra*\(^{34}\) (*Ggl*), *Luffa cylindrica*\(^{35}\) (*Lcy*), *Avena sativa*\(^{36}\) (*Asa*), and *Dictyostelium discoideum*\(^{37}\) (*Ddi*). Valine is conserved at the corresponding position in lanosterol synthases represented by *Trypanosoma brucei*\(^{38}\) (*Tbr*), *Trypanosoma cruzi*\(^{39}\) (*Tcr*), *Saccharomyces cerevisiae*\(^{40,41}\) (*Sce*), *Candida albicans*\(^{42,43}\) (*Cal*), *Cephalosporium caerulens*\(^{44}\) (*Cca*), *Schizosaccharomyces pombe*\(^{27}\) (*Spo*), *Rattus norvegicus*\(^{45,46}\) (*Rno*), and *Homo sapiens*\(^{47,48}\) (*Hsa*).

To investigate the role of steric bulk at this position, Matsuda et al.\(^{49}\) generated and characterized the *AthCAS1* mutants where Ile481 was changed to Leu, Ala, Gly and Phe. The *AthCAS1* Ile481Phe was inactive whereas all the other mutants had a broadened product profile (Figure 1.4, Table 1.1). The *AthCAS1* Ile481Leu mutant
produced mainly cycloartenol with parkeol and a trace of lanosterol as by-products. The \textit{AthCAS1 Ile}481\textit{Ala} and \textit{AthCAS1 Ile}481\textit{Gly} mutants produced cycloartenol, parkeol, lanosterol and the monocycles achilleol A\textsuperscript{50} and camelliol C\textsuperscript{51}. The subtle change induced by replacing the \textit{sec}-butyl side chain in Ile with the isobutyl side chain in Leu shifts steric bulk and compromises cyclopropyl ring formation but does not allow much deprotonation to form lanosterol. Further reduction in steric hindrance with an Ile481Ala and Ile481Gly mutations compromised not only specific deprotonation, but also cyclization resulting in early truncation of cyclization and formation of monocyclic compounds. Steric bulk at position 481 seems necessary for correct substrate folding to avoid miscyclization and early deprotonation.

Figure 1.4. Oxidosqualene cyclization products of \textit{AthCAS1 Ile}481 mutants.
<table>
<thead>
<tr>
<th>481 residue</th>
<th>cycloartenol</th>
<th>parkeol</th>
<th>lanosterol</th>
<th>achilleol A</th>
<th>camelliol C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile</td>
<td><img src="image" alt="" /></td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Leu</td>
<td><img src="image" alt="" /></td>
<td>83</td>
<td>16</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Val</td>
<td><img src="image" alt="" /></td>
<td>54</td>
<td>21</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Ala</td>
<td><img src="image" alt="" /></td>
<td>12</td>
<td>15</td>
<td>54</td>
<td>13</td>
</tr>
<tr>
<td>Gly</td>
<td><img src="image" alt="" /></td>
<td>17</td>
<td>4</td>
<td>23</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 1.1. Percentage product composition of AthCAS1 Ile481 mutants.

By looking at conservation patterns between known cycloartenol and lanosterol synthases, Herrera et al.\textsuperscript{52} identified in addition to Ile481 four other residues that are strictly and differentially conserved: Tyr410, Gly488, Phe717, and Met731. They constructed and characterized AthCAS1 mutants with each of these residues mutated to their lanosterol synthase counterpart: Tyr410Thr, Gly488Ala, Phe717Trp, and Met731Ala. Only AthCAS1 Tyr410Thr had an altered product profile, but the other three mutants remained accurate cycloartenol synthases. AthCAS1 Tyr410Thr abolished cycloartenol formation and produced lanosterol along with the novel 9\(\beta\)-lanosta-7,24-dien-3\(\beta\)-ol (9\(\beta\)-A7-lanosterol) (deprotonation from C-7) and parkeol in a 65:33:2 ratio (Figure 1.6, Table 1.2). Tyr410 is essential for cyclopropyl ring formation and is therefore a key residue to the catalytic difference between CAS1 and Erg7. Herrera et al
suggested that Tyr410 is in the active site cavity because Ser307, the corresponding residue in SHC, is in the active site.\textsuperscript{29,30} As for Ile481, recent structural data of lanosterol synthase\textsuperscript{31} and homology modeling of \textit{AthCAS1} described within this thesis work support the proposed Tyr410 location. Because the Tyr410Thr mutation yields products from the C-8 cation, Herrera et al. suggested that Tyr410 could be involved in stabilizing the C-9 cation or in facilitating the hydride shift from C-9 to C-8 that are necessary for cycloartenol formation.

\textbf{Figure 1.5.} Conservation pattern of \textit{AthCAS1} Tyr410. Tyrosine is strictly conserved in the known cycloartenol synthases (CAS1) from \textit{Arabidopsis thaliana}\textsuperscript{26} (Ath), \textit{Pisum sativum}\textsuperscript{32} (Psa), \textit{Panax ginseng}\textsuperscript{33} (Pgi), \textit{Glycyrrhiza glabra}\textsuperscript{34} (Ggl), \textit{Luffa cylindrica}\textsuperscript{35} (Lcy), \textit{Avena sativa}\textsuperscript{36} (Asa), and \textit{Dictyostelium discoideum}\textsuperscript{37} (Ddi). Threonine is conserved at the corresponding position in fungal lanosterol synthases represented by \textit{Saccharomyces cerevisiae}\textsuperscript{40,41} (Sce), \textit{Candida albicans}\textsuperscript{42,43} (Cal), \textit{Cephalosporium caerulens}\textsuperscript{44} (Cca), \textit{Schizosaccharomyces pombe}\textsuperscript{27} (Spo), and mammalian lanosterol synthases represented by \textit{Rattus norvegicus}\textsuperscript{45,46} (Rno), and \textit{Homo sapiens}\textsuperscript{47,48} (Hsa).

To investigate possible synergistic effects between the known catalytically relevant residues, they generated the \textit{AthCAS1} Tyr410Thr Ile481Val double mutant where two lanosterol synthase residues were introduced. The double mutant produced
75% lanosterol, 24% 9β-Δ7-lanosterol and < 1% parkeol. The Tyr410Thr and the Ile481Val mutations have synergistic effects, *AthCAS1*Tyr410Thr Ile481Val biosynthesizes lanosterol more efficiently than either parent mutant (Table 1.2).

<table>
<thead>
<tr>
<th>mutant</th>
<th>cycloartenol</th>
<th>parkeol</th>
<th>lanosterol</th>
<th>9β-Δ7-lanosterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>native</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ile481Val</td>
<td>54</td>
<td>21</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Tyr410Thr</td>
<td>0</td>
<td>2</td>
<td>65</td>
<td>33</td>
</tr>
<tr>
<td>Tyr410Thr Ile481Val</td>
<td>0</td>
<td>&lt; 1</td>
<td>75</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 1.2. Percent product composition of *AthCAS1* and *AthCAS1* mutants.

![Diagram of cycloartenol cyclization products](image)

Figure 1.6. Oxidosqualene cyclization products of *AthCAS1* Tyr410Thr mutants.

Meyer et al.\textsuperscript{53} used directed evolution in *Dictyostelium discoideum* cycloartenol
synthase\textsuperscript{37} (\textit{DdiCAS1}) to identify new catalytically relevant positions. Four mutants \textit{DdiCAS1} Tyr363Cys, \textit{DdiCAS1} Phe424Ser, \textit{DdiCAS1} Ser434Pro, and \textit{DdiCAS1} Tyr481His were selected for their ability to genetically complement the lanosterol synthase deletion in SMY8. To study the effects of these mutations in the \textit{AthCAS1} background, they generated the corresponding \textit{AthCAS1} derivatives by site-specific mutagenesis: \textit{AthCAS1} Tyr410Cys, \textit{AthCAS1} Phe472Ser, \textit{AthCAS1} Ser482Pro, and \textit{AthCAS1} Tyr532His. The \textit{AthCAS1} Tyr410Cys and \textit{AthCAS1} Tyr532His strongly complemented SMY8 as their \textit{DdiCAS1} analogs did in the original selection. In contrast, the \textit{AthCAS1} Phe472Ser complemented poorly and the \textit{AthCAS1} Ser482Pro did not complement whereas their \textit{DdiCAS1} analogs were selected for their strong complementation ability. In vitro experiments established that \textit{AthCAS1} Tyr410Cys cyclized oxidosqualene to lanosterol, 9\(\beta\)-\(\Delta\)7-lanosterol and achilleol A in a 75:24:1 ratio. The \textit{AthCAS1} Tyr410Cys mutant introduces Cys which is structurally similar to the Thr residue found at the corresponding position in wild-type \textit{Erg7}. Both the \textit{AthCAS1} Tyr410Cys and the \textit{AthCAS1} Tyr410Thr mutants produce lanosterol as the major compound. \textit{AthCAS1} Tyr532His generated lanosterol, parkeol and achilleol (45:31:24). Tyr532 is strictly conserved in cycloartenol synthases, and the Tyr532His mutation abolished cycloartenol biosynthesis indicating that Tyr532 is essential for the formation of the strained cyclopropyl ring. However, lanosterol synthases also show strict conservation of Tyr at the corresponding position which suggests that this residue should play a role common to both mechanisms. \textit{AthCAS1} Tyr532His generates a monocyclic compound, so it might play a role in promoting cyclization. Tyr532 corresponds to the active-site residue Tyr420 in SHC which has also been shown to be involved in
promoting cyclization.$^{54}$

\textit{AthCAS1} Phe472Ser produced lanosterol as deduced from the complementation experiments, and in vitro assays indicated very low activity and did not allow further product characterization. Phe472 is also strictly conserved in all cycloartenol and lanosterol synthases and therefore is not expected to be part of the catalytic difference between these enzymes. Phe472 corresponds to the strictly conserved active-site residue Phe365 in SHC that was shown to be critical for catalysis as a SHC Phe365Ala generated only abortive cyclization products.$^{55}$

In vitro experiments with \textit{AthCAS1} Ser482Pro did not afford any oxidosqualene cyclization products. This result combined with the lack of complementation of SMY8 was surprising given that this mutant was initially selected in \textit{DdiCAS1} for its complementation ability.

\textbf{Previous mutagenesis studies in lanosterol synthase}

Several mutagenesis experiments in lanosterol synthase have also illuminated the roles of specific residues by altering the catalytic capabilities of the enzyme. Corey et al.$^{27}$ looked at the conservation pattern of the five lanosterol synthases known at the time$^{40,42,45,47,48}$ and observed that there were fourteen conserved tryptophan residues. Buntel and Griffin$^{42}$ had suggested that because Trp residues are electron-rich they could stabilize the intermediate carbocations during cyclization. Alternatively, Trp residues could have a structural function though hydrophobic interactions or \(\pi\)-stacking. If this were the case a phenylalanine substitution would not alter the enzyme activity. To investigate the role of these Trp residues Corey et al.$^{27}$ mutated eight Trp in \textit{SceErg7} to
phenylalanine: Trp194Phe, Trp198Phe, Trp207Phe, Trp218Phe, Trp232Phe, Trp583Phe, Trp587Phe, and Trp632Phe. Complementation experiments in SMY8 revealed that all the mutants maintained their ability to biosynthesize lanosterol, demonstrating that these Trp residues are not essential for lanosterol biosynthesis.

Corey et al.\textsuperscript{56} carried out mutagenesis experiments in SceErg7 to investigate the possibility that an aspartic or glutamic acid within lanosterol synthase could be acting as the protonating agent that initiates cyclization. All the Asp (residues: 140, 286, 370, 456, 580 and 629) and Glu (residues: 216, 264, 460, 483, 487, 511, 520, 526, and 634) residues conserved in the known lanosterol synthases were mutated to Asn and Gln respectively. The mutants were expressed in SMY8 and grown in absence of ergosterol to check for lanosterol production. Of all the mutants constructed, only the Asp456Asn was inactive demonstrating that Asp456 is essential for lanosterol biosynthesis. Considering these results and given that Asp456 is in a highly conserved region (DCTA) Corey et al. proposed that Asp456 could be the catalytic acid that initiates cyclization by protonating the epoxide in oxidosqualene. Because experiments with OS showed that OS was stable for several hours in glacial acetic acid and a stronger acid was required to protonate the epoxide,\textsuperscript{56,57} Corey et al. suggested that the acidity of Asp456 might be increased by a positively charged nearby residue. To investigate further, Corey et al.\textsuperscript{58} made 76 SceErg7 mutants in which highly conserved residues were mutated to different amino acids. The mutants were checked for viability by expressing them in SMY8 and growing them without ergosterol supplementation. In addition to Asp456Asn, five inactive mutant enzymes were identified: His146Ala, His234Ala, Met532Ala, His234Lys and His234Arg. A yeast lanosterol synthase mutant expressing the His234Phe mutant
derivative grew without ergosterol whereas the His234Ala, His234Lys and His234Arg mutants did not. In view of these results, Corey et al. suggested that His234 could be involved in cation stabilization and substrate folding. In contrast, the His146Ala mutant was inactive whereas the Lys and Arg mutants were active; these results led to the hypothesis that protonated His146 could be hydrogen-bonded to Asp456 to increase its acidity. Corey et al. also suggested that unprotonated His146 could be the base that abstracts the proton to form lanosterol regenerating protonated His146 for the next catalytic cycle. The crystal structure of AacSHC reported afterward indicated that Asp376, the corresponding residue of Asp456, was appropriately positioned to be the protonating acid. In contrast, the residue corresponding to His146 was not located in the active site or in the vicinity of the protonating acid.²⁹ No specific role was proposed for Met532.

To investigate the role of the N- and C- termini, Corey et al.⁵⁸ generated truncated enzymes missing 19, 31, 52, 65, or 75 amino acids at the N-terminus or 21 amino acids at the C-terminus. Complementation experiments in SMY8 showed that only the mutant missing the first 19 amino acids in the N-terminus was able to grow in absence of ergosterol. These results indicated that these 19 amino acids are not essential for lanosterol biosynthesis.

Joubert et al.⁵⁹ studied the role of sterics in SceErg7 at position 454, which corresponds to AthCAS1 Ile481 that had been found to be a key residue for product specificity and tetracyclization in AthCAS1. They mutated the strictly conserved Val454 residue to Ile, Leu, Ala, Gly and Phe. SceErg7 Val454Phe was found to be inactive as was the corresponding AthCAS1 mutant Ile481Phe. The aromatic ring might either
disrupt protein folding, or project into the active site cavity and interfere with substrate folding. The \textit{SceErg7 Val454Ile} and \textit{SceErg7 Val454Leu} mutants remained accurate lanosterol synthases suggesting that the extra steric bulk imposed by an extra methylene does not interfere with the catalytic mechanism. However, reducing steric hindrance in the \textit{SceErg7 Val454Ala} and \textit{SceErg7 Val454Gly} mutants led to miscyclization and formation of achilleol A in addition to the native product lanosterol (Table 1.3). Reducing steric bulk in the \textit{SceErg7 Val454 residue} influences cyclization but not deprotonation. Once the cyclic intermediate is formed deprotonation is specific. In contrast, in \textit{AthCAS1} steric bulk changes at Ile481 influences both cyclization and deprotonation.

<table>
<thead>
<tr>
<th>454 residue</th>
<th>lanosterol</th>
<th>achilleol A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
<td>&lt;chem&gt;CH&lt;sub&gt;3&lt;/sub&gt;-CH&lt;sub&gt;3&lt;/sub&gt;-CH&lt;sub&gt;3&lt;/sub&gt;&lt;/chem&gt;</td>
<td>100</td>
</tr>
<tr>
<td>Ile</td>
<td>&lt;chem&gt;CH(CH&lt;sub&gt;2&lt;/sub&gt;)&lt;sub&gt;2&lt;/sub&gt;-CH&lt;sub&gt;3&lt;/sub&gt;&lt;/chem&gt;</td>
<td>100</td>
</tr>
<tr>
<td>Leu</td>
<td>&lt;chem&gt;CH&lt;sub&gt;3&lt;/sub&gt;-CH&lt;sub&gt;2&lt;/sub&gt;-CH&lt;sub&gt;3&lt;/sub&gt;&lt;/chem&gt;</td>
<td>100</td>
</tr>
<tr>
<td>Ala</td>
<td>&lt;chem&gt;CH&lt;sub&gt;3&lt;/sub&gt;&lt;/chem&gt;</td>
<td>95</td>
</tr>
<tr>
<td>Gly</td>
<td>H&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>83</td>
</tr>
</tbody>
</table>

\textbf{Table 1.3.} Percentage product composition of \textit{SceErg7 Val454} mutants.
Figure 1.7. Oxidosqualene cyclization products of lanosterol synthase mutants.

Meyer et al.\textsuperscript{60} studied \textit{SceErg7} residue Thr384, which corresponds to Tyr410 in \textit{AthCAS1} that had been found to be essential for cycloartenol biosynthesis and therefore part of the catalytic difference between both enzymes. Tyr410 is strictly conserved in cycloartenol synthases and corresponds to an active-site residue in the \textit{AacSHC}\textsuperscript{28} crystal structure\textsuperscript{29,30}. They generated and characterized the \textit{SceErg7} Thr384Tyr mutant, which bore the corresponding CAS1 residue. \textit{SceErg7} Thr384Tyr had a broadened product profile with lanosterol as major product and parkeol and lanost-24-ene-3β,9α-diol (9α-hydroxylanosterol) as by-products (79:11:10). Lanost-24-ene-3β,9α-diol had not previously been found in nature or as an enzymatic product, but had been obtained synthetically.\textsuperscript{61} Lanost-24-ene-3β,9α-diol derives from water quenching of the C-9 intermediate carbocation possibly concerted with the hydride shift from C-9 to C-8 (Figure 1.7). This was the first example of a water-quenched triterpene in a lanosterol synthase mutant, but several examples of water-quenched cyclization products from other
enzymes are known.\textsuperscript{62,63}

To investigate how this residue interacts with the previously known catalytically relevant \textit{SceErg7} V454 residue, Meyer et al. constructed the \textit{SceErg7} Thr384Tyr Val454Ile mutant with both residues mutated to the corresponding CAS1 residues. The double mutant had catalytic properties different from either parent. The Val454Ile mutation alone in the parent mutant did not alter product profile. However, in the double mutant, Val454Ile amplified the effect of Thr384Tyr increasing parkeol formation while decreasing lanosterol and increasing moderately lanost-24-ene-3β,9α-diol (64:13:23). They also generated the \textit{SceErg7} Thr384Tyr Val454Leu double mutant to observe the effects of having an isobutyl side-chain at position 454 instead of a sec-butyl side-chain. \textit{SceErg7} Thr384Tyr Val454Leu also formed parkeol, lanosterol and lanost-24-ene-3β,9α-diol but with a slightly different ratio (Table 1.4). Again, even though Val454Leu does not change product specificity in the parent single mutant, when combined with Thr384Tyr it favors parkeol formation. The Thr384Tyr mutation broadens product profile acquiring parkeol biosynthetic ability, and an increase in steric bulk at position 454 through a Val454Ile or Val454Leu mutation amplifies the effect making parkeol the major product.
<table>
<thead>
<tr>
<th>mutant</th>
<th>lanosterol</th>
<th>parkeol</th>
<th>lanost-24-ene-3β,9α-diol</th>
</tr>
</thead>
<tbody>
<tr>
<td>native</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Val454Ile</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Val454Leu</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thr384Tyr</td>
<td>79</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Thr384Tyr Val454Ile</td>
<td>13</td>
<td>64</td>
<td>23</td>
</tr>
<tr>
<td>Thr384Tyr Val454Leu</td>
<td>20</td>
<td>60</td>
<td>20</td>
</tr>
</tbody>
</table>

**Table 1.4.** Percentage product composition of lanosterol synthase mutants.

**Squalene-hopene cyclase**

Squalene-hopene cyclases\(^4,\text{5}\) convert squalene to hopene and hopanol (hopan-22-ol, diplopterol) which are the biosynthetic precursors of a wide variety of hopanoids that serve as membrane constituents\(^6\) in bacteria. The mechanism of cyclization is similar to that of OSCs although is less complex because it does not involve cation rearrangement. The reaction is initiated by protonation of the C2-C3 double bond of squalene in an all pre-chair conformation (non-sterol folding) to give the hopanyl cation which can then undergo deprotonation from C29 to yield hopene or nucleophilic attack by water to give hopanol.\(^7,\text{9}\) The *Alicyclobacillus acidocaldarius* squalene-hopene cyclase\(^28\) (*AacSHC*) was initially crystallized with medium resolution (2.9 Å)\(^29\) and later with high resolution\(^30\) (2.0 Å) in a new crystal form. This was the only structural information available at the outset of this thesis work. Squalene-hopene cyclases and oxidosqualene cyclases are only ~20 % identical, and SHCs are ~100 amino acids shorter than OSCs. Therefore,
SHC structural information can be used as a guide to study OSCs but with extreme caution.\(^6\)

\begin{center}
\begin{tikzpicture}
\node (squalene) at (0,0) {squalene};
\node (hopanyl_cation) at (2,0) {hopanyl cation};
\node (hopene) at (4,2) {hopene};
\node (hopanol) at (4,-2) {hopanol};
\draw[->] (squalene) -- (hopanyl_cation);
\end{tikzpicture}
\end{center}

**Figure 1.8.** *Alicyclobacillus acidocaldarius* squalene-hopene cyclase catalyzes the cyclization of squalene to hopene and hopanol.

Squalene-hopene cyclase is a homodimeric monotopic\(^6\) membrane protein. Each subunit has 631 amino acids and is organized in two domains. Domain 1 consists of an \(\alpha_6\)-\(\alpha_6\) barrel of two concentric rings of \(\alpha\) helices and domain 2 contains an \(\alpha\)-\(\alpha\) barrel and is inserted into domain 1. The center contains long loops from both domains and small \(\beta\)-strands surrounding a large cavity. The dimer interface is more polar than usual oligomer interfaces.
Figure 1.9. Ribbon representation of homodimeric squalene-hopene cyclase.

OSC\textsubscript{S}s and SHC\textsubscript{S}s contain a distinctive motif that consists of a QW-sequence repeat\textsuperscript{69} that is found 5-8 times within the sequence and that is not present in other enzymes. The crystal structure of \textit{AacSHC} reveals eight QW-motifs that interconnect the external helices by hydrogen bonds and give stability to the enzyme. Wendt et al.\textsuperscript{29} proposed that these motifs could absorb the energy released during the highly exothermic ($\sim$200 kJ/mol) cyclization reaction and therefore avoid structural damage. The protein stabilization energy is typically $\sim$50 kJ/mol.

The active site is located in the large central cavity and squalene can access it through a non polar channel between the helices of domain 2. The non polar channel has a constriction that seems to be mobile enough to allow the access of the substrate and the
exit of the products, hopene and hopanol. Surrounding the channel entrance on the enzyme surface there is a large non-polar plateau with a ring of positively charged residues. In the dimer the non-polar plateaus are parallel. This arrangement seems to facilitate the fusion of the enzyme with the membrane.

The active site cavity is lined with several conserved aromatic residues and has a polar end at the top where the conserved sequence motif DXDD (residues 374 to 377, X = any amino acid) is located. Wendt et al.\textsuperscript{29} proposed that Asp376 is the catalytic residue that initiates the cyclization by protonating the double bond in squalene. Previous mutagenesis experiments had shown that Asp376 was essential for catalysis.\textsuperscript{70} Usually a carboxylic acid is not acidic enough to protonate a carbon-carbon double bond. However, other residues within the enzyme contribute to decrease the pKa of Asp376. His451 is probably positively charged and is hydrogen-bonded to Asp376, so it might increase the acidity of Asp376 by stabilizing the negative charge after squalene’s double bond protonation. Asp374:Asp377 are a hydrogen-bonded pair that might be negatively charged to stabilize the positively charged Asp376: His451 pair. There is a water molecule that connects Asp376 to Tyr495, this system could increase the acidity of Asp376 and could also facilitate reprotonation of Asp376 after each catalytic cycle. The proton could originally come from bulk water and could be shuttled through disorder water in the upper cavity to Asp476 through Tyr495 and the bridged water molecule.

Wendt et al.\textsuperscript{29,30} suggested that the aromatic residues that lined the active site could stabilize intermediate carbocations by cation-π interactions.\textsuperscript{42,71,72} The conserved aromatic residues Trp312, Trp489 and Phe365 could stabilize the intermediate cations at C4, C10 and C8 (hopene numbering), respectively. Trp169 could stabilize the anti-
Markovnikov intermediate cation at C13 during C-ring formation. Phe601 is at a proper position to stabilize the Markovnikov intermediate cation at C18 during the proposed five-membered D-ring closure.\textsuperscript{73} Phe605 is only conserved in squalene cyclases and is in an appropriate position to stabilize the secondary cation at C17 facilitating D-ring expansion to get the six-membered ring observed in hopene. Phe605 could also stabilize the hopanyl cation during E-ring closure. Wendt et al.\textsuperscript{30} suggested that additional stabilization of the first intermediate carbocations could be provided by the negatively charged hydrogen-bonded Asp374:Asp377 pair and by the dipoles of the conserved residues Tyr609 and Tyr612. Very recent work in our laboratory suggests that stabilization of intermediate carbocations is not a major factor in controlling selectivity of cyclization and rearrangement.

The crystal structure does not show any residue that could act as the base near the E-ring where deprotonation occurs to give hopene. However, at the bottom of the cavity there is a network of water molecules in contact with a hydrogen-bonded network that includes Gln262, Glu45, Glu93 and Arg127. Wendt et al.\textsuperscript{29} proposed that a water molecule polarized by this polar network could serve as the catalytic base to yield hopene, or could quench the intermediate cation to give hopanol. The proton could then be shuttled via the hydrogen-bonded water molecules to the bulk solvent.

Reinert et al.\textsuperscript{74} reported later a crystal structure of \textit{AacSHC} bound to the inhibitor 2-azasqualene\textsuperscript{75} which is very similar to squalene. The new structure provided insights on the cyclization mechanism and the conformation of squalene prior to cyclization. Squalene prefolds in a meandering conformation enforced by the shape of the active site and contracts from both ends during cyclization. Reinert et al. suggested that the
geometry of the squalene molecule is set in a way that cyclization of rings A through D (6-6-6-5 chair-chair-chair tetracycle) can occur with minimal conformational changes and therefore in a concerted manner. After the D-ring is formed there is a long-lived intermediate cation (C19) before D-ring expansion and E-ring formation can occur.\textsuperscript{74} The last isoprene unit has to move to get the 22-23 double bond into proximity with the cation at C19 and this movement comprises an energy barrier. This mechanism could explain the observed byproducts in SHC and several products formed by mutant enzymes.\textsuperscript{9,73} After the 22-23 double bond gets close to C19, a D-ring expansion can occur concomitantly with E-ring formation to form the hopenyl cation, which then gets quenched by deprotonation or by water addition.

\textbf{Figure 1.10.} Active site of squalene-hopene cyclase with the inhibitor 2-azasqualene.
Lanosterol synthase crystal structure*

Lanosterol synthase is a membrane-bound protein associated with the endoplasmic reticulum and therefore has proven to be very difficult to purify and crystallize. Membrane proteins require the use of detergents to disrupt the membrane and solubilize the protein. Choosing the right conditions to solubilize and purify the enzyme without protein denaturation is not an easy assignment. Ruf et al.\textsuperscript{76} overexpressed human lanosterol synthase\textsuperscript{47,48} (HsaErg7) in Pichia pastoris, solubilized it with Triton X-100, and after purification they got milligram amounts of pure enzyme (95\% pure). Using analytical ultracentrifugation they determined that the protein is monomeric. Activity assays determined that HsaErg7 is most active as a monomer confirming that this is the native state. The highest enzyme activity was observed with Triton X-100. Crystallization attempts under different conditions identified octyl-\(\beta\)-D-glucopyranoside (\(\beta\)-OG) as the appropriate detergent, probably because it keeps the enzyme in a monodisperse state\textsuperscript{77}, and led to crystals suitable for structure determination.

Late in the course of this thesis work, Thoma et al.\textsuperscript{31} reported the crystal structure of human lanosterol synthase in complex with the inhibitor Ro48-8071 and in complex with lanosterol. Human lanosterol synthase is a monomeric monotopic membrane protein with two (\(\alpha/\alpha\)) barrel domains connected by loops and three small \(\beta\)-stuctures. Five QW-sequence repeats\textsuperscript{69} are located in domain one in the external barrel helices. As suggested for SHC,\textsuperscript{29,30} Thoma et al. proposed that these motifs might provide additional stability to the enzyme to avoid structural damage due to the high energy released during

* Residue numbering within this section corresponds to human lanosterol synthase, whereas elsewhere numbering corresponds to yeast lanosterol synthase (SceErg7). For comparisons, see protein sequence alignments in Figures 1.3 and 1.5.
the cyclization reaction. The amino terminal region is located in between both domains and seems to be required to determine their relative orientation. SHCs lack this N-terminal region.

The active-site is located in a large cavity in the center of the enzyme between domains 1 and 2. The membrane binding region is a hydrophobic surface located in domain 2. A hydrophobic channel connects the active-site with the membrane insertion site to allow the uptake of substrate from the membrane interior and the release of lanosterol. This channel has a constriction site that could make traffic of molecules difficult. However, Thoma et al. suggested that changes in the side chain orientations of three residues (Tyr237, Cys233 and Ile524) around this site or reorganization of the loops 516-524 and 697-699, would allow access of substrate and exit of product. The active-site cavity is mainly hydrophobic but has a polar end at the top. The crystal structure shows Asp455 at the top of the cavity hydrogen-bonded to the 3β-OH group of lanosterol suggesting a role for this residue in agreement with the previous belief that Asp455 was the catalytic acid.\textsuperscript{56,58} Corey et al.\textsuperscript{56,57} had shown that oxidosqualene is stable for \~1 day in glacial acetic acid, so the sole carboxylic group of Asp455 is not acidic enough to protonate the epoxide group. Cys456 and Cys533 are hydrogen-bonded to Asp455 and therefore activate Asp455 by decreasing its pKa. The acidity of Asp455 is lower than the acidity of the corresponding residue Asp376 in SHC; Wendt\textsuperscript{78} speculated that this could be so that the enzyme could discriminate between the epoxide in one end of the oxidosqualene molecule and the olefin in the other end. Lanosterol synthase does not cyclize squalene whereas squalene-hopene cyclase can accept and cyclize oxidosqualene.\textsuperscript{79,80}
**Figure 1.11.** Two different views of a ribbon representation of human lanosterol synthase. Bottom view shows the space between domains with the amino terminal sequence region, the loops and the β structures.
The crystal structure suggests that Tyr98 imposes the unfavorable B-ring boat folding by forcing the methyl group at C8 (lanosterol numbering) below the molecular plane. Schulz-Gasch and Stahl\textsuperscript{81} had observed that OSCs have a one residue insertion above the molecular plane and a one residue deletion below the plane compared to SHCs that could help enforce the B-ring boat conformation.

The active site cavity is lined with conserved aromatic residues. Thoma et al. suggested that these residues could stabilize the intermediate carbocations through cation-π interactions. Phe444, Tyr503 and Trp581 could stabilize the intermediate tertiary cations at C10 and C8 (lanosterol numbering) after A- and B-ring formation. His232 and Phe696 could stabilize the anti-Markovnikov intermediate secondary cation at C13 during C-ring formation. Thoma et al. suggested that after the protosteryl cation (C20) is formed, rearrangement to the lanosteryl C8 cation occurs because there is a higher π-electron density around C8. However, this rearrangement can occur spontaneously without much involvement of the enzyme because hydride and methyl shifts have been observed in non enzymatic Lewis acid catalyzed cyclization of OS, and in the reaction of protosterol analog compounds with Lewis acids.\textsuperscript{82-84} His232 is the only basic residue available in the proximity of the C8 lanosteryl cation to abstract a proton from C9 to form lanosterol. The X-ray structure shows His232 hydrogen-bonded to Tyr503 which is in a better position to abstract the proton, so Tyr503 might initially abstract the proton which can then be shifted to His232. This proton can be shuttled back to Asp455 after the catalytic cycle. Alternatively, the proton for Asp455 reprotonation could come from bulk solvent through a chain of water molecules and Glu459.
Figure 1.12. Active site of human lanosterol synthase with the reaction product lanosterol (blue). Asp455 (yellow) is the catalytic acid and His232 (magenta) is the catalytic base. Residues mentioned within the text are labeled.
Figure 1.13. Stereo view of the active site of human lanosterol synthase with the reaction product lanosterol. Lanosterol is shown in blue, Asp455 in yellow and His232 in magenta.

Overview

Several catalytic residues in lanosterol synthase and cycloartenol synthase have been identified by means of random and site-directed mutagenesis. Typically, mutations compromise product specificity and mutant enzymes have broadened product profiles. However, there are not any examples of mutant enzymes that had been modified to generate a new product accurately. Reported mutants disrupt substrate folding or modify cation quenching to generate monocyclic or tetracyclic byproducts. No characterized mutants generate bicyclic or tricyclic compounds.

This thesis describes mutagenesis experiments and homology modeling that identified a second sphere AthCAS1 residue that is a critical component of the catalytic
distinction between cycloartenol synthase and lanosterol synthase. Additional experiments are described in which homology modeling was combined with site-specific mutagenesis to redesign cycloartenol synthase into and accurate lanosterol synthase. Finally, lanosterol synthase mutants are described that generate a novel tricyclic compound.
Chapter 2: Results and Discussion

Part I: Cycloartenol Synthase Mutants

\textit{AthCAS1 His477Asn} and \textit{AthCAS1 His477Gln} single mutants

Random mutagenesis and selection experiments uncovered the \textit{AthCAS1 His477Asn} mutant for its ability to genetically complement the lanosterol synthase deletion in the yeast strain SMY8.\textsuperscript{85} The \textit{AthCAS1 His477Asn} mutant enzyme was expressed and characterized in the yeast strains LHY4 (squalene synthase/lanosterol synthase deletion mutant) and RXY6 (squalene epoxidase/lanosterol synthase deletion mutant). These strains are useful for expressing oxidosqualene cyclases because they cannot biosynthesize oxidosqualene and consequently do not accumulate triterpene alcohols in vivo, thus facilitating accurate quantitation of products. In vitro assays with racemic oxidosqualene, followed by extraction and partial purification yielded a triterpene alcohol fraction. GC, GC-MS and \textsuperscript{1}H NMR analyses established that the \textit{AthCAS1 His477Asn} mutant forms 88\% lanosterol and 12\% parkeol (Figure 2.1). Among the previously known cycloartenol synthase mutants, the \textit{AthCAS1 Tyr410Thr Ile481Val} double mutant was the best lanosterol synthase producing 75\% lanosterol along with 25\% 9\(\beta\)-lanosta-7,24-dien-3\(\beta\)-ol (9\(\beta\)-\(\Delta\)7-lanosterol).\textsuperscript{52} \textit{AthCAS1 His477Asn} biosynthesizes lanosterol quite cleanly. His477 is strictly conserved in all known cycloartenol synthases (Figure 2.2). However, native lanosterol synthases in fungi and trypanosomes, encode glutamine at the corresponding position and mammals encode cysteine. Hypothesizing that an enzyme encoding the native glutamine residue could
increase lanosterol production further, we generated and assayed the *AthCAS1* His477Gln mutant. Surprisingly, this mutant's major product is parkeol. A combination of GC, GC-MS, and $^1$H NMR techniques established that this mutant makes parkeol (73%), lanosterol (22%), and 9β-Δ7-lanosterol (5%).

These mutagenesis experiments illuminated the role of His477, which is strictly conserved in the known cycloartenol synthases (Figure 2.2). A histidine residue at position 477 seems to be essential for cyclopropyl ring formation because both the His477Asn and His477Gln mutations abolished cycloartenol biosynthesis. Although Gln differs from Asn only by having an additional methylene that slightly increases steric bulk and shifts the polar functionality, the *AthCAS1* His477Gln mutant has radically different catalytic properties from the *AthCAS1* His477Asn mutant.

The *AthCAS1* His477Gln mutant was the most accurate parkeol synthase known at that time, whereas the *AthCAS1* His477Asn mutant produced lanosterol more accurately than any previously described cycloartenol synthase mutant. Asn induces lanosterol formation more effectively than Gln in the *AthCAS1* background, but Gln is preferred in native lanosterol synthases. This finding is a vivid illustration that a residue with optimal features for catalysis in one background may interact differently with the reaction intermediate or neighboring amino acid residues when its environment changes. A model derived from the squalene-hopene cyclase crystal structure$^{29,30}$ positioned *AthCAS1* His477 outside of the cycloartenol synthase active site.$^{86}$ However, subtle changes in the amino acid structure at position 477 dramatically alter catalysis. His477 influences deprotonation more strongly than any of the previously studied catalytic residues.
Figure 2.1. Product profiles of wild-type AthCAS1 and AthCAS1 His477 single mutants.

<table>
<thead>
<tr>
<th></th>
<th>native</th>
<th>His477Asn</th>
<th>His477Gln</th>
</tr>
</thead>
<tbody>
<tr>
<td>cycloartenol</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>lanosterol</td>
<td>0</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>parkeol</td>
<td>1</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>9β-Δ7-lanosterol</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 2.2. Conservation pattern of His477. Histidine (●) is strictly conserved in the known cycloartenol synthases (CAS1) from Arabidopsis thaliana26 (Ath), Pism sativum33 (Psa), Panax ginseng33 (Pgi), Glycyrrhiza glabra34 (Ggl), Luffa cylindrica35 (Lcy), Avena sativa36 (Asa), and Dictyostelium discoideum37 (Di). Glutamine is conserved at the corresponding position in both trypanosomal and fungal lanosterol synthases represented by Trypanosoma brucei38 (Tbr), Trypanosoma cruzi39 (Tcr), Saccharomyces cerevisiae40 (Sce), Candida albicans42,43 (Cal), Cephalosporium caerules44 (Cca) and Schizosaccharomyces pombe27 (Spo). The known mammalian lanosterol synthases from Rattus norvegicus45,46 (Rno) and Homo sapiens47,48 (Hsa) have cysteine at the corresponding position.
*Ath*CAS1 Tyr410Thr His477Asn Ile481Val and AthCAS1 Tyr410Thr His477Gln Ile481Val triple mutants

Tyr410, His477, and Ile481 are catalytically important *Ath*CAS1 residues. These are strictly conserved cycloartenol synthase residues, but animal and fungal lanosterol synthases maintain Thr, Cys or Gln, and Val at the corresponding positions (Figure 2.3). Mutating these residues in *Ath*CAS1 to the corresponding lanosterol synthase residues allows some degree of lanosterol biosynthesis.25,52,53,85,86 The *Ath*CAS1 Ile481Val mutant produces 25% lanosterol, along with cycloartenol and parkeol (Figure 2.4),25 and the *Ath*CAS1 Tyr410Thr mutant forms 65% lanosterol with 9β-Δ7-lanosterol and parkeol as byproducts (Figure 2.4).52 Of all *Ath*CAS1 residues studied to date, His477 has the strongest effect on product profile. The *Ath*CAS1 His477Asn mutant biosynthesizes lanosterol more accurately (88%) than any other described *Ath*CAS1 mutant,85 whereas the His477Gln mutation strongly favors parkeol formation (73%).85

![Conservation pattern of Tyr410, His477 and Ile481. Representative cycloartenol synthases are from Dictyostelium discoideum (DdiCAS1) and Arabidopsis thaliana (AthCAS1). Representative lanosterol synthases are from the fungi Saccharomyces cerevisiae (SceERG7) and Schizosaccharomyces pombe (SpoERG7) and the animals Homo sapiens (HsaERG7) and Rattus norvegicus (RnoERG7). Tyr410 (•), His477 (*) and Ile481 (▼) are labeled above the sequences.](image-url)

*Figure 2.3.* Conservation pattern of Tyr410, His477 and Ile481. Representative cycloartenol synthases are from *Dictyostelium discoideum* (DdiCAS1) and *Arabidopsis thaliana* (AthCAS1). Representative lanosterol synthases are from the fungi *Saccharomyces cerevisiae* (SceERG7) and *Schizosaccharomyces pombe* (SpoERG7) and the animals *Homo sapiens* (HsaERG7) and *Rattus norvegicus* (RnoERG7). Tyr410 (•), His477 (*) and Ile481 (▼) are labeled above the sequences.
Mutations at positions 410 and 481 have synergistic effects. The *AthCAS1* Tyr410Thr Ile481Val double mutant biosynthesizes lanosterol more accurately than either single mutant (Figure 2.4).\textsuperscript{52,67} We anticipated that the three mutations that promote lanosterol biosynthesis (His477Asn, Tyr410Thr and Ile481Val) would act synergistically to form lanosterol even more accurately than the Tyr410Thr Ile481Val double mutant. In contrast, we expected the parkeol-forming properties of the His477Gln mutant to diminish lanosterol biosynthesis in the Tyr410Thr Ile481Val double mutant background. We generated the triple mutants and expressed them in the yeast strains LHY4 and RXY6. In vitro assays coupled with GC-FID, GC-MS, and NMR analyses established that both mutants generate essentially the same product profile as the Tyr410Thr Ile481Val double mutant (Figure 2.4). The His477Asn mutation does not increase lanosterol biosynthesis in the Tyr410Thr Ile481Val background; the triple mutant actually produces less lanosterol than does the His477Asn single mutant. Similarly, the Tyr410Thr and Ile481Val mutations completely abolish the influence of His477Gln so that parkeol is not even produced as a measurable byproduct. Although His477Asn and His477Gln mutations impart radically different catalytic properties on native *AthCAS1*, this influence is preempted by the Tyr410Thr and Ile481Val mutations. This catalytic hierarchy is unusual. Previous studies show synergistic effects between catalytically relevant oxidosqualene cyclase mutations; combining catalytically relevant mutations generates a multiple mutant with catalytic properties distinct from either parent.\textsuperscript{52,60}
Figure 2.4. Percentage product composition of AthCAS1 and AthCAS1 mutants.

We applied homology modeling studies to investigate how changing structure through mutations affects product specificity. The AthCAS1 structure was modeled as described previously for the closely-related (44% identical) human lanosterol synthase. A key feature of the active site is a hydrogen-bonding network of residues that includes Tyr118, His257, Tyr410, Asp483, Tyr532, Tyr616, Tyr734 and Tyr737 (Figure 2.5 and 2.6). These residues are strictly conserved in the known cycloartenol synthases, and all but Tyr410 are also present in the known lanosterol synthases. Tyr410 and Tyr532 are essential for cycloartenol biosynthesis; Asp483 and His257 correspond to the catalytically essential yeast lanosterol synthase residues that were proposed to be the active-site acid and base, respectively. This hydrogen-bonding network would order the active-site structure and may shuttle a proton from C-19 to the active-site acid (Figure 2.5).
Figure 2.5. *AthCAS1* homology model showing mutated residues and residues crucial for the deprotonation step.

The Ile481 side chain is at the top of the active site, and by interacting with the A ring it orients the substrate in the binding site to avoid early reaction termination.²⁵,⁸¹ Lanosterol synthase has smaller residues at positions corresponding to *AthCAS1* Ile481 and Tyr410. The active-site enlarges when these residues are mutated to their lanosterol synthase counterparts, which may compromise the mutants’ ability to fix the position of
Figure 2.6. Stereo representation of the AthCAS1 homology model showing mutated residues and residues crucial for the deprotonation step.

the lanosteryl cation to promote deprotonation from C-19.\textsuperscript{25,52}

His257 and Tyr410 are a H-bonded pair positioned near the C-19 angular methyl group where deprotonation yields cycloartenol. Distances after Molecular Dynamics (MD) simulations with constraint backbone and flexible active-site side-chains were 2.5 Å for the Tyr410 oxygen and 2.7 Å for the His257 Ne2 to either hydrogen atom of the C-19 angular methyl group. Although Tyr410 appears to be slightly closer to accept the proton, His257 is not dramatically more distant. The difference in distances is relatively minor, and considerable error could arise from assumptions made in homology modeling and constraints during MD simulations. Either residue could be the base or could influence deprotonation by participating in the H-bonding network. The experimentally established change in deprotonation position in Tyr410 mutants is consistent with either possibility. If Tyr410 is the base, the Tyr410Thr mutation would abolish cycloartenol
biosynthesis by removing the phenolic hydroxyl that would be the proton acceptor. Alternatively, the Tyr410Thr mutation could alter the hydrogen-bonding network and induce the proton acceptor His257 to change positions. The Tyr410Thr mutant remains a competent catalyst, suggesting that the change in H-bonding does not abolish reprotonation of Asp483.

His477 is not in the active site, but is a 2nd sphere residue that is hydrogen bonded to Tyr410 (Figure 2.5 and 2.6) and strongly affects its location and side chain orientation. His477 is hydrogen bonded to a backbone carbonyl via N81 (donor) and to Tyr410 via Ne2 (acceptor). Changes at His477 strongly affect the location, orientation, and electronics of the Tyr410 side chain. Mutating His477 to Asn keeps a nitrogen atom in the position of N81 to form the hydrogen backbone interaction. However, the Asn carbonyl oxygen is farther away, approximately in the position of C82 of the native His. As a result, the hydrogen bond with Asn477 pulls Tyr410 away from the active site (Figure 2.7). Mutating His477 to Gln induces a less dramatic change. The carbonyl oxygen of Gln is located between Ne2 of His477 in wild-type AthCAS1 and the carbonyl oxygen of Asn in the His477Asn mutant. Changing the orientation of Tyr410 disturbs the hydrogen bonding pattern with His257 and Tyr532, shifting the potential proton acceptors in these three residues (Figure 2.7). In addition, the steric changes induced by moving Tyr410 may affect product formation. Pulling Tyr410 only slightly out of the active site with Gln477 may allow sufficient mobility for the base to access the C-11 proton to form parkeol. Further reduction of steric hindrance through Tyr410 by Asn477 would further enlarge the active site, which might allow sufficient rotation of the intermediate cation that the C-9 proton would be accessible. The His477Asn and
His477Gln mutations are catalytically irrelevant in the Tyr410Thr Ile481Val background because they are outside of the active site and influence catalysis only through interaction with Tyr410. Mutating Tyr410 to Thr relocates the polar group away from position 477, and interrupts H-bonding so that these residues no longer influence each other's positions.

![Diagram](image)

**Figure 2.7.** Effects of mutations in position 477 on the orientation of Tyr410 side chain (wild-type: green, His477Gln: yellow, His477Asn: blue). The lanosteryl cation is shown with a transparent Connolly surface. His477 mutants pull Tyr410 out of the active site and allow re-orientation of the intermediate cation to form other products.

A previous effort to model *AthCAS1* threaded the sequence onto the *Alicyclobacillus acidocaldarius* squalene-hopene cyclase (*AacSHC*) crystal structure.
This model oriented the His477 side chain away from Tyr410 to mimic the AacSHC positioning, and it consequently could not explain the observed effects of His477 mutations.\textsuperscript{86} Although the primary sequence alignment lacks inserts or gaps near His477, steric and electrostatic differences between residues require some reorientation of backbone atoms (this rearrangement does not further affect secondary structure). For this reason, the AthCAS1 model generated with a fully automated procedure could not predict atoms in this region.

The influence of 2\textsuperscript{nd} sphere residues on oxidosqualene cyclase product profile may be a necessary adaptation to the unusual reactivity of the carbocationic intermediates that these enzymes handle. Their active sites must be constructed from a limited set of amino acids to preclude nucleophilic attack on an intermediate carbocation. In particular, these active sites are dominated by electron-rich but relatively unreactive\textsuperscript{29,30,74} aromatic residues, which may stabilize carbocations through $\pi$-interactions.\textsuperscript{71} However, relatively few aromatic amino acids are available for natural protein construction, and this set may lack sufficient structural diversity to generate the subtle alterations in active-site topology needed to form the nearly 200 different cyclic products of the oxidosqualene cyclase/squalene cyclase family.\textsuperscript{10} The experiments and modeling describe herein show that 2\textsuperscript{nd} sphere residues can generate diversity in active-site topology by interactions with active-site residues. The preponderance of Tyr in the active site may reflect that its hydrogen-bonding potential makes it more responsive than Phe or Trp to neighboring residues, which can alter the position or electronic properties of the tyrosine by interaction with the phenolic hydroxyl group.
AthCAS1 His477Asn Ile481Val and AthCAS1 His477Gln Ile481Val double mutants

Nature has used random mutagenesis and selection to generate an enormous diversity of enzymes. Protein scientists have recently begun to make progress in engineering enzymes with broadened substrate, reaction, and product specificity.\textsuperscript{87-90} It has proven dramatically more difficult to redesign enzymes to have stringent specificity. Herein, homology modeling was combined with site-directed mutagenesis to redesign cycloartenol synthase to make lanosterol with 99% accuracy.

Cycloartenol synthase and lanosterol synthase are paralogous enzymes. Despite being only moderately related (\textsim 40% identical), they promote mechanistically similar reactions. Both enzymes cyclize oxidosqualene to the protosteryl cation and promote hydride and methyl shifts to form the C-8 cation, and then diverge by abstracting a different proton (Figure 2.8). Cycloartenol synthase promotes an additional hydride shift and terminates the reaction with cyclopropyl ring formation and deprotonation from C-19, whereas lanosterol synthase deprotonates from C-9. We elucidated how the known catalytically important residues impact the difference in deprotonation using a structural model derived from the Alicyclobacillus acidocaldarius squalene-hopene cyclase (AacSHC) crystal structure.\textsuperscript{29,30,67,74} The orientation of active-site side chains was refined using the recently reported crystal structure of the human lanosterol synthase.\textsuperscript{31} As stated above, Tyr410, His477, and Ile481 are strictly conserved, catalytically important residues in AthCAS1. These residues synergistically promote cycloartenol biosynthesis, and mutations at these positions allow lanosterol formation. Ile481 is conserved in all cycloartenol synthases, whereas Val is present in lanosterol synthases. The Ile481 γ-
methyl promotes accurate cycloartenol formation by preventing rotation of the intermediate through steric interactions with C-2 and the two axial methyl groups of the

![Chemical structures](image)

oxidosqualene  →  protosteryl cation  →  C-8 cation  →  C-9 cation

↓

lanosterol  →  cycloartenol

**Figure 2.8.** Cyclization of oxidosqualene to lanosterol and cycloartenol.

A-ring (Figure 2.5 and 2.6). Removing the γ-methyl group with an Ile481Val mutation results in 25% lanosterol in addition to cycloartenol and parkeol (Figure 2.9). Tyr410 hydrogen-bonds with His257 to form part of the ceiling of the active site. Both Tyr410 and His257 are close to C-19 and one of these could be the base that deprotonates to form cycloartenol. Alternatively, their influence on deprotonation could result from participation in the H-bonding network. Tyr410 is present in all cycloartenol synthases, but animal and fungal lanosterol synthases maintain Thr at the corresponding position. The *AthCAS1* Tyr410Thr mutant forms 65% lanosterol along with 9β-Δ7-lanosterol and
parkeol. Removing the aromatic ring of Tyr410 decreases steric bulk above the intermediate. Because the hydroxyl in Thr is closer to the α-carbon than in Tyr, the polar groups of Tyr410Thr, Tyr532, and His257 are repositioned in the Tyr410Thr mutant. This combination of steric and electronic changes abolishes cycloartenol biosynthesis and allows deprotonation of the C-8/C-9 lanosteryl cation to form lanosterol, parkeol and 9β-Δ7-lanosterol.

![Cycloartenol, Lanosterol, Parkeol, 9β-Δ7-Lanosterol](image)

<table>
<thead>
<tr>
<th></th>
<th>Cycloartenol</th>
<th>Lanosterol</th>
<th>Parkeol</th>
<th>9β-Δ7-Lanosterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>99</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>His477Asn</td>
<td>0</td>
<td>88</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>His477Gln</td>
<td>0</td>
<td>22</td>
<td>73</td>
<td>5</td>
</tr>
<tr>
<td>Ile481Val</td>
<td>54</td>
<td>25</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Tyr410Thr</td>
<td>0</td>
<td>65</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>Tyr410Thr Ile481Val</td>
<td>0</td>
<td>78</td>
<td>&lt;1</td>
<td>22</td>
</tr>
<tr>
<td>Tyr410Thr His477Asn Ile481Val</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Tyr410Thr His477Gln Ile481Val</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>His477Asn Ile481Val</td>
<td>0</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>His477Gln Ile481Val</td>
<td>0</td>
<td>94</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 2.9.** Percentage yields of oxidosqualene cyclization products in *AthCAS1* and *AthCAS1* mutants.
His477 is a second-sphere residue that affects the product profile through interactions with the side chain of Tyr410. His477 is strictly conserved in the known cycloartenol synthases, whereas lanosterol synthases maintain either Gln or Cys. The *AthCAS1* His477Gln mutant is a parkeol synthase (73%) that also forms lanosterol (22%). Although known lanosterol synthases lack Asn at this position, an *AthCAS1* His477Asn mutant biosynthesizes 88% lanosterol and 12% parkeol. The structural model suggests that these His477 mutations alter steric and electronic surfaces of the active site by offering a smaller interaction partner to Tyr410, which is consequently withdrawn from the active site. The Tyr410Thr mutation facilitates lanosterol production by decreasing active-site bulk and relocating potential proton acceptors, and the His477 mutants achieve a similar result by moving Tyr410 away from the intermediate cation.

Careful examination of the homology model suggested that the His477Gln and His477Asn mutations should act synergistically with the Ile481Val mutation to more accurately biosynthesize lanosterol. The *AthCAS1* His477Gln mutant has the polar functionality moved towards C-11 and consequently biosynthesizes more parkeol than lanosterol. *AthCAS1* His477Asn forms lanosterol by positioning the base near C-9/C-8, but close enough to C-11 to form some parkeol. *AthCAS1* Ile481Val allows some lanosterol biosynthesis by introducing a smaller side-chain which enlarges the active site cavity permitting rotation of the intermediate cation. Computer models show that both the *AthCAS1* His477Gln Ile481Val and the *AthCAS1* His477Asn Ile481Val double mutants relocate polarity to a position more favorable for lanosterol formation (Figure 2.10). The decreased sterics allow the intermediate to rotate, moving C-9/C-8 towards the base. Because *AthCAS1* His477Gln and *AthCAS1* His477Asn form 73% and 12%
parkeol respectively, the reduction in parkeol biosynthesis in *AthCAS1* His477Gln Ile481Val was expected to be less than in *AthCAS1* His477Asn Ile481Val.

![Figure 2.10](image)

**Figure 2.10.** Superposition of wild-type *AthCAS1* (green) with A) *AthCAS1* H477Q I481V (yellow) and B) *AthCAS1* H477N I481V (blue). The lanosteryl cation is shown with a transparent Connolly surface. The hydrogen bonding pattern is indicated by dotted lines.

The *AthCAS1* His477Asn Ile481Val and the *AthCAS1* His477Gln Ile481Val double mutants genetically complemented the yeast lanosterol synthase deletion mutant SMY8.\(^{27}\) Expression in the yeast strain RXY6\(^{91}\) provided cyclase free of in vivo products, and in vitro assay with racemic oxidosqualene generated a triterpene alcohol fraction. After purification by silica gel chromatography, GC-FID, GC-MS and \(^1\)H NMR analyses established that the *AthCAS1* His477Asn Ile481Val mutant produces 99% lanosterol and 1% parkeol. Similar analyses on the *AthCAS1* His477Gln Ile481Val mutant products revealed the same compounds in a ratio of 94 : 6. These rationally designed cycloartenol synthase derivatives are the best known examples of terpene synthases modified to make different products accurately, and they compare favorably
with remodeled proteins in other systems. The experimental results for both double mutants confirmed the working hypothesis and further validate the homology model. Activity assays indicated that the \textit{Ath}CAS1 His477Asn Ile481Val mutant has about half the activity of wild type \textit{Ath}CAS1. This mutant has quite good efficiency considering that many triterpene cyclase mutants have very low efficiency compared to wild-type enzyme.\textsuperscript{9}

The synergistic behavior of the His477Asn/Gln and Ile481Val mutations resembles that seen with the Tyr410Thr and Ile481Val mutations; the \textit{Ath}CAS1 Tyr410Thr Ile481Val double mutant produces lanosterol more accurately (78\%) than the parent mutants.\textsuperscript{52,67} Neither mutation depends on direct interaction with the other or preempts the effects of the other. In contrast, the His477Asn and His477Gln mutations lack influence when both Ile481Val and Tyr410Thr are present (Figure 2.9).\textsuperscript{67} These three mutations are not synergistic because when Tyr410 is mutated to Thr, His477Asn/Gln cannot interact with the smaller and more distant Thr side chain and therefore cannot influence catalysis.

The \textit{Ath}CAS1 His477Asn Ile481Val double mutant (with native Tyr410) is the most accurate example of an enzyme mutated to biosynthesize lanosterol, and the \textit{Ath}CAS1 His477Gln Ile481Val double mutant is nearly as accurate. Neither enzyme has the Tyr410Thr Ile481Val maintained by animal and fungal lanosterol synthases. The \textit{Ath}CAS1 His477Gln Ile481Val double mutant has the residues of the phylogenetically distinct trypanosomal lanosterol synthases.\textsuperscript{38,39} No native lanosterol synthases were known at the time of these experiments that used the motifs of the more accurate
At/CAS1 His477Asn Ile481Val double mutant. However, a lanosterol synthase gene from Arabidopsis thaliana was recently cloned and characterized and has these motifs.\(^9\)

These experiments show how mutagenesis coupled with a high-quality model allows the design of an efficient and highly accurate lanosterol synthase. An improved understanding of structure and catalytic mechanisms should facilitate future engineering of enzymes with tailored catalytic activities.\(^9\)

**Part II: Lanosterol Synthase Mutants**

*SceErg7 Tyr510His and SceErg7 Tyr510Phe mutants*

The catalytic importance of Tyr510 (residue numbering corresponds to SceErg7) was initially established in directed evolution experiments with cycloartenol synthase (CAS1) from Dictyostelium discoideum\(^3\) (DdiCAS1) and Arabidopsis thaliana\(^2\) (AthCAS1). Mutating the corresponding Tyr in AthCAS1 to His abolishes cycloartenol biosynthesis and results in a mixture of the monocycle achilleol A (24%) and two tetracycles, lanosterol (45%) and parkeol (31%).\(^5\) This catalytic outcome is consistent with the homology model that predicts this Tyr residue to be part of a hydrogen-bond network that facilitates cation deprotonation.\(^6\) The human lanosterol synthase crystal structure indicates a different role for Tyr510; its phenol is hydrogen-bonded directly to the proposed active-site base His234 (Figure 2.11). The proximity of the phenolic oxygen to the C9 proton suggests that Tyr510 is the initial proton acceptor and transfers its phenolic proton to His234.\(^3\) To investigate experimentally the roles of Tyr510, I generated the Tyr510His and Tyr510Phe mutants of lanosterol synthase from
Saccharomyces cerevisiae (SceErg7).\textsuperscript{40} SceErg7 Tyr510Phe had previously been constructed and shown to complement SMY8, but no product characterization was done.\textsuperscript{58}

\textbf{Figure 2.11.} Tyr 510 and other selected residues in the active site of lanosterol synthase.

The SceErg7 Tyr510 mutants were generated by site-specific mutagenesis and were expressed in the yeast strain RXY6. In vitro assays were conducted by incubating racemic oxidosqualene with homogenates obtained from 1-L cultures of yeast expressing
SceErg7 Tyr510 mutants. After extraction and partial purification, GC, GC-MS, and NMR analyses established that the SceErg7 Tyr510His triterpene fraction contained achilleol A (2), lanosterol (4), parkeol (5), and an unidentified C\textsubscript{30}H\textsubscript{50}O triterpene alcohol (3) with a mass spectral fragmentation suggesting incomplete cyclization (Table 2.1, Figure 2.12). The SceErg7 Tyr510Phe reaction gave lanosterol and the same unknown 3 in a 95:5 ratio.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>achilleol</th>
<th>isomalabaricatrienol</th>
<th>lanosterol</th>
<th>parkeol</th>
</tr>
</thead>
<tbody>
<tr>
<td>AthCAS1 Tyr532His</td>
<td>24</td>
<td>0</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>SceErg7</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SceErg7 Tyr510Phe</td>
<td>0</td>
<td>5</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>SceErg7 Tyr510His</td>
<td>45</td>
<td>4</td>
<td>42</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2.1. Percentage product composition of oxidosqualene cyclases.

SceErg7 Tyr510His and SceErg7 Tyr510Phe were also expressed in SMY8. Both mutants genetically complemented SMY8 and showed growth comparable to wild-type SceErg7 expressed under identical conditions. In SMY8 cultures expressing SceErg7 Tyr510Phe, the lanosterol was largely converted to ergosterol, whereas 3 was evidently not metabolized. Consequently, the triterpene alcohols that accumulated in vivo were dramatically enriched in the unknown. These SMY8 cultures showed a distorted triterpene product ratio relative to the RXY6 ratios (Table 2.2). In vivo SMY8 product ratios varied from 80:20 to 90:10. This variability is unsurprising because time of harvest, amount of ergosterol in the medium, degree of aeration, and other culture conditions
could affect the extent of lanosterol metabolism. By contrast, in vitro RXY6 product ratios are not affected by such conditions and are reproducible. The results illustrate the pitfalls of using in vivo accumulation to estimate the product profile of a triterpene synthase.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>isomalabaricatrienol</th>
<th>lanosterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>RXY6 (in vitro)</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>SMY8 (in vivo)</td>
<td>90</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 2.2. Comparison of in vivo and in vitro product ratios.

I exploited the distorted product profile of SMY8 [SceErg7 Tyr510Phe] to obtain enough unknown 3 for structure elucidation. A 1-L culture gave after saponification and partial purification a triterpene alcohol fraction containing predominantly 3, with some lanosterol, 4,4-dimethylcholesta-8,24-dien-3β-ol (T-MAS), and traces of sterol intermediates and possibly unidentified triterpenes. Further purification by HPLC yielded an analytical sample of 3 which was used for structure elucidation. Compound 3 was identified by NMR spectroscopy (1H, 13C, DEPT, COSYDEC, HSQC, HMBC, NOE) as (13αH)-isomalabarica-14(27),17E,21-trien-3β-ol.
Figure 2.12. Cyclization of oxidosqualene by SceErg7 and SceErg7 Tyr510 mutants.
The novel tricycle 3 is the parent skeleton of isomalabaricanes, found in nature only in certain Asian sponges.\textsuperscript{94-97} These secondary metabolites are tricyclic triterpenoids that display the distinctive \textit{trans-syn-trans} ring fusion found in lanosterol biosynthesis. Sponges synthesize sterols from lanosterol,\textsuperscript{98,99} and the enzyme that constructs the isomalabaricatrienol skeleton probably evolved from a lanosterol synthase. Sponges could cyclize oxidosqualene to isomalabaricatrienol and produce the isomalabaricane triterpenoids by additional desaturation and specific oxidation (Figure 2.13). The $\Delta 13(14)$ isomer of 3 could alternatively be a precursor of isomalabaricanes.\textsuperscript{10} However, 3 should be easier to generate enzymatically since a C14 methyl proton is more readily abstracted than the more hindered C13 methine proton. Also, abstraction of the C13 proton would probably require rotation of the side chain to align the C-H bond with the cationic 2p orbital.\textsuperscript{100}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.13.png}
\caption{Proposed biosynthesis of isomalabaricane triterpenoids via isomalabaricatrienol.}
\end{figure}
These sessile marine animals produce numerous secondary metabolites for a variety of purposes.\textsuperscript{101,102} Among the many 6-6-5 tricycles isolated from nature or from experiments with substrate analogs and cyclase mutants, nearly all have a \textit{trans-anti-trans} stereochemistry. The sole characterized oxidosqualene cyclase that forms tricycles is thalianol synthase,\textsuperscript{91} a plant enzyme that generates a 6-6-5 malabaricatrienol from a \textit{trans-anti-trans} all-chair cation. Interestingly, thalianol synthase shares close phylogenetic affinity to plant β-amyrin and lupeol synthases which generate tetracyclic all-chair intermediates. These observations are consistent with the hypothesis that cyclases rarely evolve from a B-ring chair to B-ring boat mechanism or vice versa.\textsuperscript{103}

Some reported cyclase mutants generate 6-6-5 \textit{trans-anti-trans} tricycles.\textsuperscript{9,104} Malabarica-14(27),17\textit{E},21-trien-3β-ol, a \textit{trans-anti-trans} isomer of 3, has been produced from 2,3 oxidosqualene by a squalene-hopene cyclase mutant\textsuperscript{104} and non-enzymatically under mildly acidic\textsuperscript{82} or free-radical\textsuperscript{105} conditions. Other biomimetic reactions also produce 6-6-5 tricycles, all with the \textit{trans-anti-trans} stereochemistry. The only reported cyclizations to isomalabaricane skeletons involve reaction of oxidosqualene substrate analogs with lanosterol synthase.\textsuperscript{106-109} However, none of these reactions produce the novel (13α\textit{H})-isomalabarica-14(27),17\textit{E},21-trien-3β-ol.

The product profiles of the \textit{SceErg7} Tyr510 mutants provide some insight into the cyclization mechanism, although interpretation of the results is necessarily speculative in the absence of crystallographic studies of mutants. Considering the critical role of the Tyr510 hydroxyl in deprotonating H-9 of the C-8 cation (D) to form lanosterol,\textsuperscript{31} it is interesting that the Tyr510Phe mutant makes predominantly lanosterol despite loss of the deprotonating hydroxyl. The mutant may have a similar active-site geometry except that
the native phenolic OH is replaced by an ordered water. This water could accept the H-9 proton and transfer a proton to its hydrogen bonding partner His234.

Alternatively, the new hydrogen-bond network may locate His234 close enough to H-9 for direct deprotonation. As shown in Figure 2.11, the aryl ring of Tyr510 does not face the active-site cavity and thus affects folding only through steric interactions. The occasional misfolding that leads to isomalabaricatrienol in Tyr510Phe may result from mobility of the aryl ring, which is no longer anchored to the hydrogen-bond network. Alternatively, steric changes caused by a modified hydrogen-bond network may slightly retard folding for D-ring formation, and the delay could allow some deprotonation of intermediate B at the C-14 methyl. This methyl is on the β face of the intermediate near H-9β, and deprotonation by His234 could be mediated by ordered waters in the active site. Molecular modeling and the absence of tricycles with alternate deprotonation positions (0.1% detection limit) suggest that steric constraints block all Wagner-Meerwein rearrangements other than the ring expansion en route to the tetracyclic cation.

The Tyr510His mutation has substantial effects on cyclization, reducing the proportion of molecules that undergo tetracyclization to 50%. The imidazole ring is only slightly smaller than a phenol, but the dramatically different electronic properties could result in extensive rearrangement of the hydrogen-bond network and induce the significant misfolding that is observed. Minor change in location of His234 could alter deprotonation to form parkeol. Likewise, the His510 may be suitably located for facile deprotonation at C-11 to form parkeol or at the C-10 methyl to form achilleol A.
Similar considerations could account for the extensive misfolding reported recently for Tyr510Ala, Tyr510Lys, and Tyr510Trp mutants. However, these reports should be interpreted cautiously; the mutants were characterized by analysis of the products accumulated in vivo, and my results in Table 2.2 indicate that such product ratios can severely underestimate lanosterol content.

In summary, *SceErg7* Tyr510 mutants produce a tricyclic alcohol that is the putative biosynthetic precursor of cytotoxic isomalabaricane triterpenoids in sponges. Isomalabaricatrienol synthase may have evolved from lanosterol synthase. The ability of simple lanosterol synthase mutants to generate isomalabaricatrienol supports this hypothesis. Further evidence will await the cloning and characterization of sponge cyclases. Animals produce only three different triterpene skeletons: lanosterol, parkeol, and isomalabaricatrienol. Notably, the Tyr510His mutant makes all three of these.
Chapter 3: Conclusions

The present work describes mutational studies of cycloartenol synthase and lanosterol synthase. The experimental results give new insights on how oxidosqualene cyclases achieve product specificity by controlling cyclization and promoting specific deprotonation of the cationic intermediates.

Mutagenesis experiments and computer modeling established that second sphere oxidosqualene cyclase residues are a critical component of the catalytic distinction between cycloartenol synthase and lanosterol synthase. *AthCAS1* His477 is a second sphere residue essential for cycloartenol biosynthesis, and mutations at this position strongly affect product profile. An *AthCAS1*His477Asn mutant was described that forms 88% lanosterol whereas an *AthCAS1*His477Gln mutant forms mainly parkeol (73%). Further analysis of the homology model suggested that combining these mutations with an Ile481Val mutation would allow the redesigning of cycloartenol synthase into lanosterol synthase. The double mutants were constructed and characterized. The *AthCAS1* His477Gln Ile481Val makes 94% lanosterol and 6% parkeol and the *AthCAS1*His477Asn Ile481Val makes 99% lanosterol and 1% parkeol. These rationally designed enzymes are the best known examples of terpene synthases modified to make different products accurately. Only two mutations were necessary to convert cycloartenol synthase to lanosterol synthase.

The crystal structure of human lanosterol synthase suggests that Tyr510 is appropriately located to abstract the proton from the C8 intermediate cation to form lanosterol. However, this work revealed that a Tyr510Phe lanosterol synthase mutant
makes primarily lanosterol (95%) and that Tyr510 is not essential for lanosterol biosynthesis. This mutation causes partial misfolding and compromises D-ring formation yielding a tricyclic compound. This novel tricycle was identified as isomalabarica-14(27),17E,21-trien-3β-ol and is the putative biosynthetic precursor of isomalabaricane triterpenoids found in nature in certain sponges. Isomalabaricatrienol has never been isolated from sponges probably because it gets metabolized quickly to oxidized isomalabaricanes. The enzyme that constructs the isomalabaricatrienol skeleton in sponges probably evolved from a lanosterol synthase. The ability of lanosterol synthase mutants to generate isomalabaricatrienol supports this hypothesis.

Valuable data has been gathered over the past fifty years on the cyclization of oxidosqualene by these enzymes. The recently reported crystal structure of lanosterol synthase did not resolve all the unanswered questions but comprises a valuable resource. All these bits of data are like pieces of a big puzzle that is still not complete. The experiments described in this thesis add new important pieces to this puzzle and provide improved understanding of the catalytic mechanisms, which will facilitate future engineering of enzymes with tailored catalytic activities.
Chapter 4: Experimental procedures

**Materials.** All restriction enzymes, T4 DNA polymerase, T4 DNA ligase, T4 DNA kinase, single-stranded binding protein, 100 bp DNA ladder, *Bst*II-digested lambda DNA molecular weight marker and Quick ligation kit were purchased from New England BioLabs (Beverly, MA). The Qiaquick gel extraction kit was purchased from Qiagen (Valencia, CA). Luria-Bertani (LB) media, dextrose, galactose, peptone, yeast extract, yeast nitrogen base and agar were obtained from Fisher Scientific (Pittsburgh, PA). Heme, ergosterol and other chemical reagents were obtained from Sigma/Aldrich Chemical Company (St. Louis, MO). Organic solvents, Silica Gel 60 and Silica Gel 60 plates were purchased from EM Science (Gibbstown, NJ). Double deionized water was obtained by filtering deionized water through a Super-Q water system (Millipore, Bedford, MA).

**Nuclear Magnetic Resonance (NMR).** Proton ($^1$H) NMR spectra were collected at 500 MHz on a Bruker Avance 500 NMR spectrometer equipped with a 5 mm inverse-geometry probe. Tetramethylsilane (TMS) was used as internal standard. All spectra were collected at 25 °C. All chemical shifts are reported in parts per million (ppm) referenced to TMS at 0 ppm. Carbon ($^{13}$C) NMR spectra were collected at 125 MHz and all shifts are reported in ppm referenced to CDCl$_3$ at 77.0 ppm. Deuterated chloroform (CDCl$_3$) (Cambridge Isotope Laboratories; Andover, MA) was filtered through basic alumina (ICN Alumina B, Activity I) prior to use and all samples were prepared in 5 mm glass tubes (Wilmad Glass Co., Inc. or Shigemi Inc.).
Gas Chromatography (GC). At the beginning of this thesis work, GC spectra were obtained on an Agilent 6890 Series GC System equipped with an Rtx-5MS column (30 m, 0.25 mm id, 0.10 μm df; Restek, Bellefonte, PA). GC conditions: inlet and FID-detector temperatures were 290 °C; the oven was kept isothermal at 280 °C; split injection was used with a ratio of 40:1; helium was the carrier gas at a constant flow of 0.6 mL/min. Under these conditions some of the triterpene alcohols described within this work were comigrating, making the product characterization very difficult. Several alternative conditions (temperature and flow) were tested with the same column with no results in separation. Later separations used the same instrument equipped with an Rtx-35MS column (30 m, 0.25 mm id, 0.10 μm df; Restek, Bellefonte, PA). GC conditions: inlet and FID-detector temperatures were 280 °C; the oven was kept isothermal at 260 °C; split injection was used with a ratio of 40:1; helium was the carrier gas at a constant flow of 1 mL/min.

Triterpene alcohol products were analyzed either without derivatization or as trimethylsilyl (TMS) ethers

Gas Chromatography-Mass Spectrometry (GC-MS). GC-MS analysis was performed on an Agilent 6890 gas chromatograph interfaced to a 5973 MSD using electron impact at 70 eV. The GC column and conditions were identical to those for GC analysis.
High-Performance Liquid Chromatography (HPLC). Reversed-phase separations were carried out with a Phenomenex 5-μm ODS3 column (250 × 4.6 mm i.d) using an Agilent 1100 HPLC system with a Rheodyne 7125 injector (500 μL loop) and ultraviolet (UV) detection at 210 nm. Separations were performed at ambient temperature (22 ± 1 °C) with gradient elution by low-pressure mixing. Elution was done at a flow-rate of 1.0 ml/min with a linear gradient from 9:1 MeOH-water to 100% MeOH during the first 15 min, followed by isocratic elution with MeOH for another 25 min. Samples were dissolved in 200 μL MeOH for injection.

UV Spectroscopy. UV measurements were done on a UV-Visible spectrophotometer Shimadzu UV-1601.

Centrifugation. Microcentrifugations were performed in variable speed Eppendorf Centrifuge 5415D. Centrifugations were done using an Eppendorf centrifuge 5810R (variable speed and temperature), a Beckman Model TJ-6 centrifuge (variable speed), or a Sorvall Model RC-5B centrifuge (variable speed and temperature).

Incubators. Bacterial and yeast plate cultures were grown in Fisher Scientific Isotemp incubators. Bacterial liquid cultures were grown in a New Brunswick Scientific G24 Environmental Incubator Shaker. Yeast liquid cultures were grown in either a New Brunswick Scientific Series 25 Incubator Shaker or a New Brunswick Scientific C25 Incubator Shaker. Cultures were usually shaken at 250 rpm unless otherwise indicated.
**Cell Lysis.** Large scale yeast cells were lysed using an Emulsiflex-C5 homogenizer (Avestin) or a French Pressure Cell Press (SLM Instruments, Inc.). Small scale yeast cells were lysed in bead beater.

**Strains.** *E. coli* strain DH5α$^{112}$ (F$^-$ φ80dLacZΔM15 Δ(lacZYA-argF) U169 deoR recA1 endA1 hsdR17(rk$, m_k^+$) phoA supE44 λ- thi-1 gyrA96 relA1) was the host for plasmid construction. Three different yeast strains were used to express oxidosqualene cyclases. SMY8$^{27}$ (*MATa erg7::HIS3 hem1::TRP1 ura3-52 trp1- 63 leu2-3,112 his3- Δ200 ade2Gal$^+$) is a lanosterol synthase deletion mutant that also bears a *hem1* deletion to facilitate sterol import. LHY4$^{25}$ (*MATa erg7::leu2::hisG erg9::HIS3 hem1::TRP1 ura3-52 trp1- Δ63 leu2-3,112 his3- Δ200 ade2Gal$^+$) is a related strain that also has a squalene synthase deletion. This strain is useful for expressing oxidosqualene cyclases because it cannot biosynthesize the precursor oxidosqualene and it consequently does not accumulate triterpenes in vivo. RXY6$^{91}$ (*MATa erg1::KanMX4 erg7::HIS3 hem1::TRP1 ura3-52 trp1-Δ63 leu2-3,112 his3–Δ200 ade2Gal$^+$) is a strain that is similar to LHY4, but instead of a squalene synthase deletion, it lacks squalene epoxidase. Whereas LHY4 accumulates farnesyl pyrophosphate, RXY6 accumulates squalene. RXY6 grows much faster than LHY4, apparently because squalene accumulation is less problematic than farnesyl pyrophosphate accumulation.

**Bacterial Media.** LB-amp liquid media$^{112}$ contained Luria broth (LB; 5 g tryptone, 2.5 g yeast extract and 2.5 g NaCl dissolved in 500 mL deionized water, autoclaved for 35-45 min at 121°C) and ampicillin (250 × stock: 25 mg/mL, filter
sterilized) to a final concentration of 0.1 mg/mL. Solid LB-amp media contained the same ingredients as liquid media with the addition of 7.5 g agar/500 mL prior to sterilization. LB-carb liquid and solid media were prepared in the same manner, using carbenicillin (250 × stock: 25 mg/mL, filter sterilized) instead of ampicillin.

**Yeast Media.** YPDHE\textsuperscript{26} media contained nutrient rich yeast peptone (2 × YP: 10 g yeast extract, 20 g peptone dissolved in 500 mL deionized water, autoclaved 20-35 min at 121 °C) as nitrogen source, and dextrose (2 × D: 20 g dextrose dissolved in 500 mL deionized water, autoclaved 20-35 min at 121 °C) as carbon source. It also contained ergosterol (100 × erg: 20 mg ergosterol dissolved by sonication in 5 mL ethanol and 5 mL Tween 80) and hemin (100 × hem: 13 mg hemin dissolved in 10 mL 50% ethanol, 10 mM NaOH) as supplements.

ScD-LeuHE\textsuperscript{27} media contained synthetic complete media (2 × Sc.: 1.7 g yeast nitrogen base, 5 g ammonium sulfate, and 2 g amino acid dropout mix without leucine dissolved in 500 mL deionized water; NaOH pellets to pH ~ 5.5, autoclaved 20 min at 121 °C) as nitrogen source. The amino acid drop out mix contained 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. The carbon source and supplements are the same as in YPDHE. ScD-LeuHE media was used to select for recombinant yeast strains containing the plasmid with the leucine biosynthetic gene marker.

YPGHE\textsuperscript{26} and ScG-LeuHE\textsuperscript{27} have galactose (2 × G: 20 g galactose dissolved in
500 mL deionized water, autoclaved 20-35 min at 121 °C) as carbon source instead of dextrose. Galactose was used to induce heterologous expression of recombinant genes under the Gal promoter, and is referred to as inducing media.

Liquid media were prepared with equal volumes of the nitrogen source and the carbon source and the appropriate amount of supplements. For solid media, 7.5 g agar/500 mL was added to the carbon source prior to sterilization. Plates were prepared by mixing 15 mL of a nitrogen source, 15 mL of a carbon source containing agar, and appropriate nutritional supplements, if needed. All yeast strains were grown at 30 °C.

**Oligonucleotides.** Custom designed oligonucleotides were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA). Oligonucleotides were dissolved in mqH₂O to a final concentration of 100 pmol/µL. These stock solutions were stored at -20 °C.

**Site-Specific Mutagenesis.** Site-specific mutagenesis was performed according to Kunkel¹¹³ with some modifications. Mutagenic oligonucleotides were designed to incorporate the desired change, and introduce a restriction site that did not otherwise alter the encoded protein sequence. Oligos were phosphorylated by combining 360 pmol of the oligo (3.6 µL of 100 pmol/µL stock solution), 2 µL 10 × kinase buffer, 1 µL 10 mM ATP, 0.5 µL T₄ DNA kinase and deionized water to make a total volume of 20.5 µL. The solution was mixed by flicking the tube, and the reaction was spun to the bottom and incubated at 37 °C for 1 h. Then, 0.5 µL 0.5 M EDTA was added and the mixture was incubated at 70 °C in a heating block for 10 min. The phosphorylated oligo was then
diluted by adding 15 μL of deionized water to get a concentration of 10 pmol/μL and stored at -20 °C.

To anneal the oligo and the template, 1 μg of ssDNA was combined with 25 pmol phosphorylated oligo, 1 μL 20 × SSC (20 × SSC: 3 M NaCl, 0.3 M sodium citrate, pH 7.0) and deionized water to get a total volume of 20 μL. The solution was mixed by flicking the tube, and the reaction was spun to the bottom and placed at 70 °C in a heating block. The block was removed from the heater and allowed to cool to < 35 °C at room temperature. The tube containing the primed DNA was spun to the bottom and placed on ice.

The second strand was synthesized by adding to the primed DNA on ice 49 μL deionized water, 10 μL 10 × ligase buffer, 20 μL 2.5 mM dNTPs, 1 μL single-stranded binding protein, 1 μL T4 DNA polymerase, and 0.5 μL T4 DNA ligase. The solution was mixed by flicking the tube. Then the reaction was spun to the bottom and incubated on ice for 5 min, at room temperature for 5 min, and then at 37 °C for 2 h. The DNA was precipitated by adding 2 μL 5 M NaCl and 200 μL ethanol. The reaction was mixed by inverting the tube several times and incubated at -20 °C for 1 h. Double-stranded DNA (dsDNA) was pelleted by centrifugation (13000 × g, 15 min), and the supernatant was decanted. This sequence was done twice. The DNA pellet was air-dried for 15 min and dissolved in 10 μL deionized water with brief incubation at 37 °C to facilitate dissolution. The dsDNA (4 μL) containing the desired mutations was used to transform DH5α. Two control reactions were performed along with the mutagenesis, one without oligo and one without T4 DNA polymerase. In both cases, deionized water was used instead.

ssDNA of pSM61.21, *Saccharomyces cerevisiae* lanosterol synthase\(^40\) (SceErg7p)
expressed in the integrative galactose-inducible yeast expression vector pRS305GAL (amp<sup>R</sup>, LEU2, GAL1 promoter) was used as template for site-specific mutagenesis in lanosterol synthase. This ssDNA was a gift from Kristin Krukenberg. ssDNA of pSM60.21, *Arabidopsis thaliana* cycloartenol synthase (AthCAS1) expressed in the integrative galactose-inducible yeast expression vector pRS305GAL, was used as template for site-specific mutagenesis in cycloartenol synthase. This ssDNA was a gift from Dr. Michael J. R. Segura.

**Bacterial Transformation.** Chemically competent DH5α *E. coli* cells (100 µL) were thawed on ice. The cells were then added to a chilled microcentrifuge tube containing recombinant plasmid DNA. The cells were gently mixed with the DNA by pipetting up and down, and the mix was incubated on ice for 30 min. The solution was heat-shocked at 37 °C for 5 min and then chilled on ice for 2 min. Aliquots of the cells (10 µL and 90 µL) were spread onto pre-warmed LB-amp or LB-carb agar plates and incubated overnight at 37 °C. In some cases when transformation efficiency could be low, cells were recovered after heat-shock by incubating with 900 µL LB or SOC medium at 37 °C for 1 h with horizontally shaking. Aliquots of 100-300 µL were used in these cases to spread onto selective plates.

**Miniprep DNA Purification.** Plasmid DNA was propagated in DH5α cells and isolated using alkaline lysis methodology. An isolated colony from a bacterial transformation plate was inoculated into 1.5 mL of selective LB media and the resultant culture was grown overnight (12-16 h) at 37 °C. Cells were harvested by centrifugation
and resuspended in 100 μL P1 buffer plus RNase A (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 20 mg/mL RNase A, stored at 4 °C). This suspension was added 100 μL P2 lysis solution (200 mM NaOH, 1% SDS (w/v)), was mixed by inverting the tube five times and was incubated at room temperature for 5 min. The lysis was terminated by neutralization with 100 μL of chilled P3 buffer (3.1 M potassium acetate, pH 5.5). The sample was mixed immediately by inversion and chilled on ice for 30 minutes. The mixture was centrifuged (12000 × g, 10 min), and the supernatant containing plasmid DNA was transferred to a clean 1.5 mL tube and 750 μL (2.5 volumes) of absolute EtOH was added. The solution was mixed by vortexing and incubated at -20 °C for 15 min or longer to increase DNA yields. The DNA was obtained by centrifugation (12000 × g, 20 min, 4 °C), and the supernatant was discarded. The DNA pellet was washed with 300 μL of 70% EtOH and centrifuged (12000 × g, 10 min, 4 °C). The ethanol wash was pipetted off and the pellet was allowed to air dry. The DNA pellet was dissolved in 30 μL of T8 buffer (10 mM Tris-HCl, pH 8.0) with brief incubation at 37 °C to facilitate dissolution. DNA solution was stored at −20 °C.

**Preparative DNA Purification.** A 20-mL bacterial culture was grown in selective LB media overnight. The cells were harvested by centrifugation (3000 × g, 10 min), and the DNA was purified from the pellet by anion exchange resin using CONCERT High Purity Plasmid Purification System (Life Technologies). The kit was used essentially following the manufacturer’s procedures, except for the use of T8 buffer (10 mM Tris-HCl, pH 8.0) instead of TE (10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA) to dissolve the purified DNA pellet. Plasmid DNA obtained in this manner was suitable for
sequencing without further purification.

**DNA Mapping.** Enzymatic digestion reactions were performed by mixing in a digestion well plate, 1-3 μL plasmid DNA, 0.3 μL restriction enzyme, 1.5 μL recommended buffer and deionized water to get the total reaction volume to 15 μL. The reactions were sealed to prevent evaporation and incubated at the temperature specified by the manufacturer for 1-2 h. Digested DNA samples were analyzed by agarose gel electrophoresis.

**DNA Gel Electrophoresis Analysis.** DNA samples were analyzed by gel electrophoresis on 1% agarose gels in TAE buffer (50 × TAE stock solution: 242 g Tris base, 57.1 mL glacial acetic acid, 37.2 g Na₂EDTA.2H₂O, pH adjusted to 8.5, di H₂O to 1L total volume). Agarose gels were prepared by mixing and melting in microwave 5 g agarose and 500 mL 1 × TAE. Once mixing was complete, ethidium bromide (5 μL/100 mL gel from a 10mg/mL stock) was added and the solution was poured into a cast electrophoresis block. After solidification, the gel was covered with 1 × TAE. DNA samples were loaded on the gel after addition of gel loading buffer (10 × stock: 0.25% Bromophenol Blue, 0.25% Xylene Cyanol, 25% Ficoll 400 in 100 mM EDTA) to help visualization. DNA molecular weight markers, 100 bp DNA ladder or BstE II-digested lambda DNA, were loaded and ran alongside to facilitate fragment size determination. The applied voltage varied between 48 and 110 volts according to the speed required for each particular experiment.
Preparative DNA Digest. Enzymatic digestion reactions for subcloning were performed by mixing in an Eppendorf tube, 10 µg plasmid DNA, 5 µL restriction enzyme, 10 µL recommended NEB buffer and deionized water to get the total reaction volume to 100 µL. The reactions were sealed to prevent evaporation and incubated at the temperature specified by the manufacturer for 2+ h. The DNA fragments were then separated by gel electrophoresis on 1% agarose gel in GTAE buffer (TAE buffer supplemented with guanosine, 0.283g/L, to avoid UV damage of the DNA) with ethidium bromide. Desired DNA fragments were excised from the agarose gel with a sharp scalpel, trying to get the minimum amount of agarose. This procedure was done under UV light, so it had to be done quickly to prevent DNA damage. DNA fragments were purified from agarose gels using Qiaquick Gel Extraction kits according to manufacturer’s directions.

DNA Ligation. Ligation of purified DNA fragments with cohesive ends was accomplished using Quick Ligation Kit (New England BioLabs) according to manufacturer’s procedures. Vector DNA was usually combined with a 3- or 5- fold molar excess of insert. The ligation mixture was either used immediately for bacterial transformation or stored at -20 °C.

DNA Sequencing. All DNA constructs described within this work were completely sequenced to guarantee that only the intended mutations had been incorporated. Sequencing was performed by either Lone Star Labs, Inc. (Houston, TX) or SeqWright, Inc. (Houston, TX) using an Applied Biosystems International sequencer.
**DNA Linearization.** All the recombinant DNA plasmids studied within this work were expressed in the integrative galactose-inducible yeast expression vector pRS305GAL.\textsuperscript{27} Since this is an integrative vector it has to be linearized before it can be used to transform yeast. Linear plasmids integrate by homologous recombination into the yeast genome at the *LEU2* locus.

Plasmid DNA (5-10 μg) was combined in an Eppendorf tube with 1.5 μL *Bst*E II, 5 μL buffer NEB 3, and deionized water to bring the total reaction volume to 50 μL. The tube was sealed and incubated at 60 °C in heating block for 1.5-2 h.

**Yeast Transformation.** A fresh yeast colony (RXY6, SMY8, or LHY4) was inoculated into 10 mL YPDHE medium and incubated overnight at 30 °C with shaking. Cells were collected by centrifugation (3000 × g, 5 min), and the supernatant was discarded. The yeast pellet was washed with 1 mL of sterile deionized water to remove remaining media, and the pellet was collected again by centrifugation (3000 × g, 30 s). This washing step was performed twice. The yeast pellet was then resuspended in 100 μL of sterile deionized water by vortexing. Pre-boiled single-stranded carrier DNA (50 μL)\textsuperscript{114} was added along with 5-10 μg of linear plasmid DNA to the yeast pellet, this solution was mixed by gentle vortexing. Then, 2 mL of yeast transformation buffer (40% polyethylene glycol (PEG) 3350, 0.1 M lithium acetate, 10 mM Tris pH 7.5, 1 mM EDTA, 0.1 M dithiothreitol; filter sterilized) were added and mixed thoroughly by vortexing. The reaction was incubated at room temperature overnight (10-16 h). The yeast cells were collected by centrifugation (3000 × g, 2 min) and the supernatant was
decanted. The cell pellet was washed twice with 1 mL of sterile deionized water to remove transformation buffer. Then, the cells were resuspended in 0.5 mL of sterile deionized water with the help of a vortex. The cells were then inoculated onto two plates of selective media (SCD-LeuHE), 200 µL in one and 20 µL in the other, spread with glass beads and incubated at 30 °C until colonies were observed in 3-4 days.

An alternative procedure was used when high efficiency transformations were needed. This method was based on the Gietz protocol\textsuperscript{114} with some modifications. A fresh yeast colony was inoculated into 10 mL YPDHE medium and grown to saturation at 30 °C with shaking. A 0.5 mL aliquot of the saturated culture was used to inoculate 50 mL YPDHE and was grown until the culture reached a density of 2 x 10\textsuperscript{7} cells/mL (1:10 dilution, OD\textsubscript{600} = 0.2). A 20 mL aliquot of this culture was transferred to a 50 mL sterile Falcon tube and the cells were harvested by centrifugation (3000 x g, 5 min). The supernatant was discarded and the cells were resuspended in 1 mL sterile deionized water and transferred to a 2 mL sterile Eppendorf tube. The cell suspension was centrifuged one minute and the supernatant was discarded. The cell pellet was resuspended in 500 µL 0.1 M LiAc (450 µL sterile di H\textsubscript{2}O + 50 µL 1 M LiAc filter sterilized) and incubated 10 min at 30 °C. The cell suspension was centrifuged for 1 min and the supernatant discarded. The transformation reagents were added to the cell pellet in the following order: 1 mL 50% PEG (25 g PEG 3350, di H\textsubscript{2}O to 50 mL, filter sterilized) (mixing with cells by pipetting up and down), 150 µL 1 M LiAc, 100 µL carrier ssDNA (2 mg/mL, pre-boiled for 5 min and cooled on ice), 5 µg linear plasmid DNA, sterile di H\textsubscript{2}O to a total of 1.5 mL. The mix was vortexed to get a homogeneous suspension and incubated at 30 °C for 30min. The suspension was then heat-shocked in 42 °C water bath for 20 min,
then it was centrifuged 1 min and the supernatant was thoroughly removed. The cell
pellet was added 1 mL sterile di H₂O, incubated at room temperature 5 min and then
resuspended by pipetting up and down and gentle vortexing. The suspension was
centrifuged for 1 min and the supernatant was discarded. The cells were resuspended in
500 μL sterile di H₂O and plated on selective plates (250 μL/ plate). Plates were
incubated at 30°C until colonies were observed (3-4 days).

Complementation Assay. The DNA constructs were linearized with BstE II and
used to transform the yeast lanosterol synthase deletion mutant SMY8. Transformants
were selected at 30 °C on SCD-LeuHE plates. Positive transformants were streaked on
YPGH plates without ergosterol supplementation, to induce expression and check the
ability of the mutants to genetically complement the lanosterol synthase deletion. SMY8
expressing Saccharomyces cerevisiae lanosterol synthase⁴⁰ (SceErg7p) was used as
positive control and SMY8 expressing Arabidopsis thaliana cycloartenol synthase⁵
(AthCAS1) was used as negative control. These plates were incubated at 30 °C for as
long as necessary to get the complementation answer (up to ten days). SceErg7p, the
positive control, complements the deletion with colonies visible in two days whereas
AthCAS1 does not complement.

Synthesis of (±)-2,3-Oxidosqualene. (±)-2,3-Oxidosqualene was synthesized
from squalene following the procedure described by Nadeau and Hanzlik.¹¹⁵
**Figure 4.1.** Synthesis of (±)-2,3-Oxidosqualene.

20 × (±)-2,3-Oxidosqualene Solution. (±)-2,3-Oxidosqualene (200 mg) and Tween 80 detergent (200 mg) were placed in a round bottom flask. Methylene chloride was added to facilitate mixing and was then completely removed by rotary evaporation. dH₂O (9.8 mL) was added and the solution was stirred at room temperature until mixing was complete and a homogeneous solution was observed. The solution was stored at 4 °C. This solution was stirred at room temperature for about 15 min before being used.

20 × Tween 80 Solution. Tween 80 (200 mg) and dH₂O (9.8 mL) were placed in a round-bottomed flask and were stirred at room temperature until mixing was complete and a homogeneous solution was observed. The solution was stored at 4 °C.

Small Scale in Vitro Assay. Recombinant yeast strain (RXY6 or LHY4) was grown at 30 °C with shaking in 10 mL SCD-LeuHE media up to saturation (2-3 days). A 1-mL aliquot of this culture was reinoculated into 10 mL YPGHE media in a pre-weighed Falcon tube and grown to saturation at 30 °C with shaking (1-2 days). The cells were collected by centrifugation (3000 × g, 5 min), the supernatant was discarded and the
pellet weighed. A 50% slurry of cell pellet (w/v) was prepared in 0.1 M sodium phosphate buffer (pH 6.2). After mixing, the slurry was divided into two Eppendorf tubes, one for the oxidosqualene reaction and one for the control reaction. 20 × (±)-2,3-Oxidosqualene solution was added to the oxidosqualene reaction tube, and 0.1% Tween 80 was added into the control reaction tube. Acid-washed glass beads (200 μL) were added into each tube. The cells were lysed in a Bead-Beater by vortexing for 3 min at maximum velocity, chilling them on ice 3 min and vortexing them again for 3 min. The tubes were incubated at room temperature (16-24 h).

The reactions were monitored by TLC. Aliquots of the enzymatic reaction mixtures (5 μL) were spotted on a TLC plate as thin vertical bands. A mixture of authentic lanosterol and ergosterol standards was spotted alongside. The TLC plate was air dried and was then partially developed twice with ethyl ether to about one fourth the length of the plate to focus each spot to a single point. The plate was then developed with either CH₂Cl₂ or 1:1 hexane/ethyl ether up to a few mm below the top of the plate. Products were visualized by staining with p-anisaldehyde (3.7 mL p-anisaldehyde, 15 mL acetic acid, 50 mL concentrated H₂SO₄, 1350 mL ethanol) and moderate heating.

Each reaction was quenched with two volumes of ethanol, typically after 24 h. The cellular debris was removed by centrifugation and the supernatant was transferred to a glass tube and concentrated to dryness under a nitrogen stream. The residue was resuspended in ethyl ether and loaded onto a short silica gel plug in a Pasteur pipet. The sample was then eluted with ethyl ether into a clean glass tube and dried under stream of nitrogen. The sample was then derivatized and analyzed by GC and GC-MS analysis.
Large Scale in Vitro Assay. A single colony of recombinant yeast strain (RXY6 or LHY4) was used to inoculate SCD-LeuHE liquid culture (10 mL) and was grown at 30 °C with shaking (250 rpm). Upon saturation, the 10 mL culture was used to inoculate YPDHE media (100 mL) and was grown in the same conditions. When saturated, the 100-mL culture was used to inoculate 1-L scale induction medium (YPGHE). Upon saturation, cells were harvested by centrifugation (4000 × g, 15 min, 4 °C) in pre-weighed centrifuge bottles and the supernatant was discarded. The cells pellet was weighed and resuspended in two volumes of 100 mM sodium phosphate buffer (pH 6.2). The cell suspension was chilled on ice, and the cells were lysed by passing the suspension twice through an Emulsiflex-C5 homogenizer or a French Pressure Cell Press. 20 × (±)-2,3-Oxidosqualene solution was added to the cell homogenate to a final concentration of 1 mg/mL. The homogenate was mixed by swirling and was incubated at room temperature for about 24 h. An aliquot of the cell homogenate (usually 1/10) with no substrate added was removed to serve as negative control for the assay. The reactions were monitored by silica gel TLC (as described above) in two solvent systems (methylene chloride and 1:1 hexane/ethyl ether) and visualized by staining with p-anisaldehyde.

After 24 h of incubation, the in vitro assays were quenched with two volumes of ethanol and the cellular debris was removed by centrifugation (3000 × g, 5 min). The ethanol was eliminated by rotary evaporation and the remaining aqueous suspension was extracted with methyl tert-butyl ether (MTBE). The combined MTBE extracts were washed with brine and then concentrated to dryness by rotary evaporation. Aliquots of these crude extracts were analyzed by GC, GC-MS, and 500 MHz ¹H NMR.
An alternative work up used ethyl acetate instead of MTBE to extract the aqueous suspension. The ethyl acetate layer was then concentrated onto 5 g of neutral alumina by rotary evaporation. The alumina bound crude was loaded onto a silica gel plug (5 g) and eluted with ethyl ether (100 mL). The ethyl ether was removed by rotary evaporation. Aliquots of this crude were analyzed by GC, GC-MS, and 500 MHz $^1$H NMR.

Large Scale in Vivo Assay. SMY8 harboring the plasmid of interest was grown in inducing medium without ergosterol supplementation (YPGH, 1L) until the culture reached high cell density. Cells were harvested by centrifugation and the cell pellet was saponified with 10% KOH in 80% EtOH for 2.5 hours at 70 °C. The reaction mixture was diluted with one volume of water and was allowed to cool to room temperature. The resultant suspension was extracted with hexanes and the combined hexanes layers were washed with brine. The hexanes layer was concentrated to dryness by rotary evaporation to yield the nonsaponifiable lipid (NSL) fraction. An aliquot of the crude NSL was analyzed by GC and GC-MS after TMS derivatization.

Column Chromatography. Triterpene alcohol products were purified from crude extracts by column chromatography using silica gel as adsorbent. Usually 6-7 g of silica gel were used to separate 100-150 mg of crude material. Separations were carried out using either methylene chloride or gradients of ethyl ether:hexanes as eluents. Fractions were collected in disposable glass tubes. Separations were monitored by TLC.
**Derivatization of Triterpene Alcohols.** Triterpene alcohol products in crude or purified samples were analyzed by GC and GC-MS as their corresponding trimethylsilyl (TMS) ethers (Figure 4.2). The sample to be analyzed was resuspended in ether and the desired aliquot was transferred to a GC vial with insert. After evaporation of solvent under a nitrogen stream, the sample residue was derivatized by treatment with 100 μL of dry pyridine and 100 μL of bis(trimethylsilyl)trifluoroacetamide (BSTFA). The sample was then sealed, mixed by gentle vortexing and left at 37 °C for >2 h. The resulting solution was used directly for GC or GC-MS.

Authentic standards were derivatized in the same manner and analyzed under the same conditions for comparison.

Product ratios were determined by GC-FID quantitation in partially purified fractions and in crude extracts.

![Diagram of derivatization process]

**Figure 4.2.** TMS-derivatization of lanosterol.
Characterization of *AthCAS1* His477Asn and *AthCAS1* His477Gln Single Mutants. The *AthCAS1* His477Asn and *AthCAS1* His477Gln constructs\(^\text{85}\) were obtained from the plasmid glycerol library. Plasmid DNA was propagated in DH5\(\alpha\), isolated, linearized with *Bst*E II and used to transform the yeast strains SMY8, LHY4 and RXY6. Complementation experiments in SMY8 revealed that both mutant enzymes complemented the lanosterol synthase deletion. In vitro assays were performed by incubating cell homogenates of LHY4 and RXY6 expressing the mutant enzymes with racemic oxidosqualene. TLC analysis showed that both mutant enzymes had a product spot that comigrated with a lanosterol standard (\(R_f\) 0.5 in 1:1 hexane/ethyl ether). GC-FID analysis of the ethyl ether crude after TMS derivatization revealed that the *AthCAS1* His477Asn mutant was producing two compounds that comigrated with TMS-derivatized standards of lanosterol (88%) and parkeol (12%). GC-MS analysis of this sample confirmed the structural assignments and did not reveal any other signals with masses consistent with being TMS derivatives of oxidosqualene cyclization products (m/z 498). Silica gel chromatography with gradients of ethyl ether:hexanes (5-20%) separated the triterpene alcohol fraction from other components (squalene, oxidosqualene, ergosterol and more polar yeast components). \(^1\)H NMR analyses of this partially purified triterpene alcohol fraction showed distinctive chemical shifts\(^\text{116}\) for parkeol (\(\delta\) 0.647, 0.737, and 5.224 ppm) and lanosterol (\(\delta\) 0.687 and 0.810 ppm) angular methyl groups and vinyl protons confirming the structural assignments. Integration of the peak areas of both angular methyl groups and vinyl protons confirmed the quantitation.

Similar GC-MS analysis of the ethyl ether crude after TMS derivatization of the *AthCAS1* His477Gln mutant revealed a major peak with retention time and mass spectral
characteristics of TMS-parkeol and one other peak with retention time and mass spectral consistent with TMS-lanosterol. GC-FID provided a ratio of 73:27 parkeol:lanosterol. $^1$H NMR of the partially purified triterpene alcohol fraction confirmed the presence of parkeol ($\delta$ 0.647, 0.737, and 5.224 ppm) and lanosterol ($\delta$ 0.687 and 0.810 ppm), but integration of the 0.647 and 0.687 signals provided a 77:23 ratio. Meticulous analysis of the $^1$H NMR spectra exposed signals at $\delta$ 5.556, 1.020, and 1.017 ppm which indicated the presence of 9$\beta$-lanosta-7,24-dien-3$\beta$-ol (9$\beta$-$\Delta^7$-lanosterol). Integration of the H-7 proton of 9$\beta$-$\Delta^7$-lanosterol relative to the H-11 proton of parkeol showed that 9$\beta$-$\Delta^7$-lanosterol made up 5% of the total product mixture. This compound was not detected in the GC and GC-MS analyses because its TMS ether comigrates with TMS-lanosterol on a Restek Rtx-5MS column under diverse GC conditions available at the time of the experiment (Later in the course of my thesis work, I found new conditions that separated these two compounds. See below.). To get the correct product ratios we combined the GC quantitation of 73% parkeol and 27% $\Delta^8$-lanosterol + $\Delta^7$-lanosterol isomers with the numbers derived from $^1$H NMR and got a product ratio of 73% parkeol, 22% lanosterol, and 5% 9$\beta$-$\Delta^7$-lanosterol. Integrating the $^1$H NMR signals of H-7 (9$\beta$-$\Delta^7$-lanosterol), H-11 (parkeol), and H-24 (side-chain representative of total product) provided a product ratio of 71% parkeol, 24% lanosterol, and 5% 9$\beta$-$\Delta^7$-lanosterol. All these numbers were in close agreement.

GC-FID and GC-MS analyses of the partially purified fractions for both mutants after TMS- derivatization yielded cleaner traces confirming structural assignments and quantitation.
Characterization of \textit{AthCAS1} Tyr410Thr His477Asn Ile481Val and \textit{AthCAS1} Tyr410Thr His477Gln Ile481Val Triple Mutants. The \textit{AthCAS1} Tyr410Thr His477Asn Ile481Val and \textit{AthCAS1} Tyr410Thr His477Gln Ile481Val triple mutants constructs\textsuperscript{67} were obtained from the plasmid glycerol library. Plasmid DNA was propagated in DH5\textalpha, isolated, linearized with \textit{Bst}E II and used to transform the yeast strain SMY8. Complementation experiments in SMY8 revealed that both mutant enzymes complemented the lanosterol synthase deletion, allowing growth comparable to the lanosterol synthase control. Both mutant enzymes were also expressed in the yeast strains LHY4 and RXY6 to perform in vitro assays. Recombinant yeast strains were grown on 1-L scale and in vitro enzymatic reactions were carried out with racemic oxidosqualene. TLC analyses of both \textit{AthCAS1} Tyr410Thr His477Asn Ile481Val and \textit{AthCAS1} Tyr410Thr His477Gln Ile481Val triple mutants assays revealed a product spot that comigrated with an authentic lanosterol standard (\textit{R}_f 0.5 in 1:1 hexane/ethyl ether). No oxidosqualene cyclization products were visualized in the negative control reactions.

GC-FID and GC-MS analyses were performed on the ethereal crude extracts of both mutant enzymes after TMS derivatization. Spectra obtained on a Restek Rtx-5MS column under diverse GC conditions showed only one product peak (\textit{m/z} 498) that comigrated with a TMS-lanosterol standard. Triterpene alcohol products were purified from ethyl ether crude by silica gel chromatography using gradients of ethyl ether:hexanes (5-20\%). 500 MHz \textsuperscript{1}H NMR analyses of the purified triterpene alcohol fractions of both \textit{AthCAS1} Tyr410Thr His477Asn Ile481Val and \textit{AthCAS1} Tyr410Thr His477Gln Ile481Val triple mutants revealed the presence of lanosterol and 9\textbeta-\Delta7-lanosterol in a roughly 3:1 ratio. This ratio was considered to be identical within
experimental errors to the one previously obtained for the *AthCAS1* Tyr410Thr Ile481Val double mutant. The presence of 9β-Δ7-lanosterol in all these mutants and the necessity to know whether these ratios were identical or not, urged me to search for the right GC conditions to separate lanosterol and 9β-Δ7-lanosterol. Finally, lanosterol and 9β-Δ7-lanosterol were successfully resolved by using an Rtx-35MS column (30 m, 0.25 mm id, 0.10 μm df) and GC conditions: inlet and FID-detector, 280 °C; oven, 260 °C isothermal; split injection, ratio 40:1; column, helium 1 mL/min constant flow. Retention times: TMS-lanosterol, 15.9 min; TMS-9β-Δ7-lanosterol, 16.2 min.

Ratios were determined by GC-FID quantitation in the partially purified fractions and confirmed by GC-FID quantitation in the crude extracts. Both *AthCAS1* Tyr410Thr His477Asn Ile481Val and *AthCAS1* Tyr410Thr His477Gln Ile481Val triple mutants have an identical product profile with 78% lanosterol and 22% 9β-Δ7-lanosterol.

The *AthCAS1* Tyr410Thr Ile481Val double mutant was re-characterized under the new conditions to confirm whether or not the product profile was identical to the triple mutants. The *AthCAS1* Tyr410Thr Ile481Val double mutant construct was obtained from the plasmid library, propagated in DH5α, isolated, linearized with Bst EII and transformed into the yeast strain RXY6. In vitro assays with racemic oxidosqualene, followed by extraction and partial purification yielded a triterpene alcohol fraction. GC-FID and GC-MS analyses of the crude extract and the partially purified fraction after TMS derivatization showed that the *AthCAS1* Tyr410Thr Ile481Val double mutant produced 78% lanosterol, 22% 9β-Δ7-lanosterol and < 1% parkeol.

The *AthCAS1* His477Asn and *AthCAS1* His477Gln single mutants were re-analyzed by GC-FID and GC-MS with the newly developed conditions. GC-FID
provided a ratio of 88% lanosterol and 12% parkeol for the *AthCAS1* His477Asn mutant and 73% parkeol, 22% lanosterol, and 5% 9β-Δ7-lanosterol for the *AthCAS1* His477Gln mutant. These product ratios agreed with the ones obtained previously by a combination of NMR and GC-FID quantitation confirming the legitimacy of those ratios.

Molecular modeling studies were performed by Dr. Tanja Schulz-Gasch on a Silicon Graphics Fuel R14000 Workstation. MOE 2003.04 (Chemical Computing Group Inc., Montreal, Quebec, Canada) was used for sequence alignment and homology modeling. Sequence information on *AthCAS1*\(^{26}\) and *AacSHC* was obtained from the Swiss-Prot database (P38605 and P33247 respectively). The *AacSHC* crystal structure\(^{29,30,74}\) (PDB entry 2SQC, resolution 2.00 Å) obtained from the Protein Data Bank (PDB) was used as template for homology modeling. Mutations D376C and C435S were reverted to wild-type residues. All water and inhibitor molecules, as well as the B-chain were removed from the structure. Sequence alignment was performed using the MOE sequence alignment module applying the BLOSUM 40 substitution matrix.\(^{117}\) Conserved motifs and catalytic residues were found to align appropriately. The stereochemical quality of the final model structure was checked with the MOE protein report and manually refined. The mutations were manually introduced to the wild-type *AthCAS1* homology model. To relax the mutated structures, short MD simulations (100 ps, gas phase) followed by a minimization were performed with Moloc (Gerber Molecular Design). The cationic intermediate was manually docked to the active sites of the model structures using Moloc.
Construction and characterization of *AthCAS1* His477Asn Ile481Val and *AthCAS1* His477Gln Ile481Val Double Mutants. The *AthCAS1* His477Asn Ile481Val double mutant construct was created by sub-cloning a *Bgl* II fragment (~650 bp) containing the His477Asn/Ile481Val mutations from the *AthCAS1* Tyr410Thr His477Asn Ile481Val triple mutant\(^67\) into the wild type *AthCAS1* Bgl II fragment (~8110 bp). Both original constructs were in the integrative galactose-inducible yeast expression vector pRS305GAL. DNA plasmids were excised with *Bgl* II; fragments were gel-purified, ligated and used to transform DH5α competent cells. Plasmid DNA was isolated for mapping, sequencing and yeast transformation.

The *AthCAS1* His477Gln Ile481Val double mutant was generated similarly by subcloning the corresponding *Bgl* II fragment from the *AthCAS1* Tyr410Thr His477Gln Ile481Val triple mutant\(^67\) into the wild-type *AthCAS1* Bgl II fragment.

The *AthCAS1* His477Asn Ile481Val plasmid was named pSLT 11.1 and the *AthCAS1* His477Gln Ile481Val was named pSLT 12.1.

The *AthCAS1* His477Asn Ile481Val and *AthCAS1* His477Gln Ile481Val constructs were linearized with *Bst*E II and used to transform the yeast strain SMY8. Complementation experiments in SMY8 revealed that both mutants supported growth without exogenous ergosterol at a rate comparable to wild type *SceErg7p* (colonies visible in two days) whereas *AthCAS1* did not.

The *AthCAS1* His477Asn Ile481Val and *AthCAS1* His477Gln Ile481Val constructs were linearized with *Bst*E II and used to transform the yeast strain RXY6. In vitro assays with cell homogenates from 1-L scale and racemic oxidosqualene (1 mg/mL) were incubated at room temperature for 24 h. TLC analysis of both assays showed a
triterpene alcohol spot that comigrated with a lanosterol standard (R_f 0.5 in 1:1 hexane/ethyl ether) and that was well separated from the oxidosqualene substrate (R_f 0.8) and yeast sterols such as ergosterol (R_f 0.3). No oxidosqualene cyclization products were visualized in the negative control assays.

The in vitro assays were extracted with methyl-tert-butyl ether (MTBE) and aliquots of the MTBE crude extracts were analyzed by GC, GC-MS, and 500 MHz \(^1\)H NMR. Triterpene alcohol fractions were purified from crude extracts by silica gel column chromatography using gradients of ethyl ether/hexane. Column separation was monitored by TLC. Products were identified by 500 MHz \(^1\)H NMR analysis and by GC and GC-MS analysis of the TMS-derivatives. Fraction A (elution with 2% ether in hexanes) gave squalene (m/z 410). Fraction B (elution with 5% ether in hexanes) gave 2,3-oxidosqualene (m/z 426). Fraction C (elution with 5-10% ether in hexanes) gave triterpene alcohols (lanosterol and parkeol as TMS ethers, m/z 498). Fraction D (elution with 20% ether in hexanes) gave ergosterol-TMS (m/z 468) and minor amounts of ergosterol-TMS derivatives (m/z 466, 468, and 470). A polar fraction was eluted with 50-100% ether in hexanes. Only fraction C contained peaks corresponding to m/z 498.

Product ratios were determined by GC-FID quantitation of the partially purified fractions and confirmed by GC-FID quantitation of the MTBE crude extracts. \textit{AthCASM1 His477Asn Ile481Val} produced lanosterol and parkeol in a 99:1 ratio and \textit{AthCASM1 His477Gln Ile481Val} formed the same compounds in a 94:6 ratio. If any additional byproducts were present, they were at levels <1%. Integration of well-resolved \(^1\)H NMR peaks in the column purified material provided similar product ratios. The triterpene alcohol fraction of \textit{AthCASM1 His477Asn Ile481Val} had the following resolved \(^1\)H NMR
signals characteristic of lanosterol.\textsuperscript{116} $\delta$ 0.690 (3H, s, H-18), 0.811 (3H, s, H-29), 0.982 (3H, s, H-19), 1.001 (3H, s, H-28), and 3.236 (1H, dd, H-3\alpha); and the following resolved signals for parkeol.\textsuperscript{85,116} $\delta$ 0.647 (3H, s, H-18) and 0.737 (3H, s, H-30). The triterpene alcohol fraction of AthCAS1 His477Gln Ile481Val had the following $^1$H NMR resolved signals characteristic of lanosterol:\textsuperscript{116} $\delta$ 0.690 (3H, s, H-18), 0.811 (3H, s, H-29), 0.982 (3H, s, H-19), 1.001 (3H, s, H-28), and 3.235 (1H, dd, H-3\alpha); and the following resolved signals for parkeol:\textsuperscript{85,116} $\delta$ 0.647 (3H, s, H-18) and 0.737 (3H, s, H-30).

It is notoriously difficult to measure $K_m$ values in the presence of detergents because the concentrations of substrate and enzyme are distorted by the biphasic aqueous and micellar system. Although most of the enzyme and substrate are probably constrained to the restricted volume of the micelles, the soluble proportion is not readily determined. I consequently compared the catalytic competence of the native and mutant enzymes using homogenate assays with substrate at a concentration (1.2 mM) well above the literature $K_m$ of 25-125 $\mu$M described for several plant cyclases.\textsuperscript{118-120}

Duplicate samples of AthCAS1 His477Asn Ile481Val and wild type AthCAS1 expressed in RXY6 were grown under identical conditions as described above and collected by centrifugation. Each yeast strain was suspended in 2 volumes of 100 mM sodium phosphate buffer pH 7 and lysed using an Emulsiflex-C5 homogenizer. After lysis, 100 mM sodium phosphate buffer pH 7 was added to make a 20% slurry. A solution of racemic 2,3-oxidosqualene and Triton X-100 was added to the homogenate aliquots (350 $\mu$L) to give a final concentration of 1 mg/mL substrate and 0.1% Triton X-100. After 0.5, 1, 3, 5, and 10 hours, the reactions were terminated by adding two volumes of ethanol. Cholesterol (8 $\mu$g) was added to each reaction as an internal
standard. The denatured protein was removed by centrifugation and the supernatant was concentrated to dryness under a nitrogen stream. The residue was resuspended in ethyl ether, loaded onto a short silica gel plug and eluted with ethyl ether. Extracts were evaporated to dryness and residues were derivatized with BSTFA-pyridine (1:1) and analyzed by GC-FID. The native enzyme generated 20 µg cycloartenol/(h × mg yeast) (n=2, mean error 2 µg / (h × mg yeast)), and the AthCAS1 His477Asn Ile481Val mutant produced 10 µg lanosterol/(h × mg yeast) (n=2, mean error 2 µg / (h × mg yeast)).

These experiments indicate that the AthCAS1 His477Asn Ile481Val mutant has about half the activity of the wild-type AthCAS1 enzyme. Since the substrate concentration should be well above the K_m and since the micellar distortions of the functional volume could provide even higher local substrate concentrations, the mutant’s modest decrease in catalytic efficiency is probably primarily a k_cat effect.

A homology model of AthCAS1 was built by Dr. Tanja Schultz-Gasch as described previously using the latest software release of MOE (Chemical Computing Group Inc, Montréal, Québec, Canada), version 2004.03. The AacSHC crystal structure (PDB entry 2SQC, resolution 2.00 Å) obtained from the Protein Data Bank (PDB) was used as template. Late in the course of our work the crystal structure of human lanosterol synthase (PDB entry 1W6K, resolution 2.1Å) was reported. AthCAS1 and human lanosterol synthase are 44% identical while AthCAS1 is only 22% identical to AacSHC, the original template for the homology model. Therefore, the model was further refined by using the new structural information. During optimization of the model no major structural changes were observed for active site residues, and refinement consisted only on fine-tuning the coordinates of active-site side chain atoms. Modeling of mutants and
manual substrate docking was carried out with the modeling package Moloc (Gerber Molecular Design, Basel, Switzerland).

**Construction and Characterization of SceErg7 Tyr510 His and SceErg7 Tyr510Phe.** The SceERG7 Tyr510His and SceERG7 Tyr510Phe were prepared by site-specific mutagenesis. The sequence of the mutagenic oligonucleotide for SceERG7 Tyr510 Phe was 5' TAGTGGGCGCTTGATTTTTTTTCCgaAGGTTGCAAAGGAACCATATTC-3'. It contained a new BstI restriction site as well as one amino acid mutation. The sequence of the mutagenic oligonucleotide for SceERG7 Tyr510His was 5' TAGTGGGCGCTTGATTTTTTTTTCATgaGTTGCAAAGGAACCATATTTCAAA-3'. It contained a new BspH I restriction site and one amino acid mutation. The plasmids were used to transform DH5α competent cells and DNA plasmid was then isolated for mapping, sequencing and yeast transformation. The SceERG7 Tyr510 Phe plasmid was named pSLT 16.1 and the SceERG7 Tyr510His was named pSLT 17.1.

pSLT 16.1 and pSLT 17.1 were linearized with BstE II and used to transform yeast strains SMY8 and RXY6. Complementation experiments in SMY8 demonstrated that both mutant enzymes were capable of supporting growth in absence of ergosterol at a rate comparable to wild type SceErg7p.

In vitro assays were conducted by incubating racemic oxidosqualene (1 mg/mL) with homogenates obtained from 1-L YPG,H,E cultures of RXY6 [pSLT 16.1] and RXY6 [pSLT 17.1], each of which yielded about 25 g of yeast each. The reactions were monitored by TLC. RXY6 [pSLT16.1] showed one new spot when compared to the
control reaction lacking oxidosqualene, while in RXY6 [pSLT17.1] two new spots were identified. Both strains presented one spot that comigrated with a lanosterol standard (Rf 0.5 in 1:1 hexane/ethyl ether) and RXY6 [pSLT17.1] had an additional spot, which comigrated with an achilleol A standard (Rf 0.6). The homogenates were partitioned with methyl-tert-butyl ether and silica gel column chromatography with gradients of ethyl ether/hexanes was used to isolate the oxidosqualene cyclization products from the crude extracts. GC and GC-MS analyses of the MTBE crude sample and of the triterpene alcohol fraction of RXY6 [pSLT16.1] showed two peaks with mass appropriate to be TMS derivatives of triterpene alcohols (m/z 498). One of them had a retention time and mass fragmentation pattern identical to an authentic TMS-lanosterol standard. The second peak corresponded to an unknown compound with a mass spectrum suggesting an incompletely cyclized compound and with a retention time of 8.8 min. Parallel GC and GC-MS analyses on the RXY6 [pSLT17.1] crude and triterpene alcohol fraction revealed three peaks that comigrated with authentic standards of TMS: achilleol A, TMS-lanosterol and TMS-parkeol; a fourth peak corresponded to the same unknown compound found in RXY6 [pSLT16.1]. Retention times: TMS-achilleol, 6.5 min; TMS-lanosterol, 12.9 min; TMS-parkeol, 14.4 min. GC-FID quantitation of products in the crude and the partially purified fraction of RXY6 [pSLT16.1] showed that this mutant produces 95% lanosterol and 5% of the unknown compound. Similar analysis showed that RXY6 [pSLT17.1] produces 42% lanosterol, 45% achilleol A, 9% parkeol and 4% unknown. \(^1\)H NMR analysis of the triterpene alcohol fractions confirmed the presence and ratios of these compounds.

SMY8 [pSLT 16.1] was grown in inducing medium without ergosterol
supplementation (YPGH, 1L) until the culture reached high cell density. Saponification of the cell pellet, followed by hexanes extraction yielded the nonsaponifiable lipid (NSL) fraction. An aliquot of the crude NSL was analyzed by GC and GC-MS after TMS derivatization. These analyses showed that SMY8 [pSLT16.1] in vivo accumulated the C_{30}H_{50}O unknown triterpene, 4,4-dimethylcholesta-8,24-dien-3β-ol (T-MAS), ergosterol and ergosterol derivatives, and a small amount of lanosterol. GC-FID quantitation revealed a ratio of lanosterol:unknown of 10:90, which was dramatically different from the ratio of 95:5 obtained in vitro. This assay showed a sharp increase in the amount of unknown relative to the amount of lanosterol because lanosterol was largely converted to ergosterol. This distorted ratio greatly facilitated the purification and characterization of the unknown. Silica gel column chromatography with gradients of ethyl ether/hexanes gave a triterpene alcohol fraction that comigrated with a lanosterol standard. GC-MS analysis after TMS derivatization revealed that this fraction consisted mainly of the unknown triterpene, with some lanosterol and T-MAS. Since lanosterol and the unknown triterpene comigrated on silica gel in both TLC and column chromatography, further separation by reversed-phase HPLC was necessary to get an analytical sample. HPLC separation was performed with the valuable help and advice of Dr. Hui Shan. The injected sample from the SMY8 [pSLT 16.1] culture contained unknown, lanosterol, and T-MAS in a 12:2:1 ratio, together with traces of other sterol intermediates en route to ergosterol, as judged by \textsuperscript{1}H NMR analysis. The unknown triterpene was eluted and collected at 21.0 min. Lanosterol and T-MAS coeluted at 27 min.

The HPLC fractions were analyzed by NMR. The unknown triterpene alcohol was identified by NMR spectroscopy by Dr. William K. Wilson (\textsuperscript{1}H, \textsuperscript{13}C, DEPT,
COSYDEC, HSQC, HMBC, NOE) as (13αH)-isomalabarica-14(27),17E,21-trien-3β-ol.\textsuperscript{10} The \textsuperscript{1}H NMR spectrum is shown in Figure 1 and indicated a purity of about 99\%. The \textsuperscript{13}C NMR and DEPT spectra shown in Figure 2 also show high purity. \textsuperscript{1}H and \textsuperscript{13}C NMR signal assignments for (13αH)-isomalabarica-14(27),17E,21-trien-3β-ol (isomalabaricatrienol) are shown in Figure 3.

Other fractions of the HPLC separation of in vivo products were also analyzed by \textsuperscript{1}H NMR. Fractions preceding elution of isomalabaricatrienol at $R_t$ 21 min contained no triterpenes other than traces of isomalabaricatrienol. The fraction at $R_t$ 27 min contained a 2:1 mixture of lanosterol and T-MAS, with traces of other ergosterol intermediates (Figure 4). Fractions eluted between $R_t$ 21 and 27 min contained small amounts of isomalabaricatrienol and several unidentified compounds that appeared to be sterols, 4,4-dimethylsterols, and/or triterpenes. These unidentified compounds represent 1\% of the amount of isomalabaricatrienol obtained from in vivo incubation. As judged by in vitro experiments, isomalabaricatrienol comprises only 5\% of the SceERG7 Tyr510Phe product; assuming that no tricyclic triterpenes are metabolized by yeast, any unidentified triterpenes would represent 0.05\% of the enzymatic products. Thus, I estimate, at a detection limit of <0.1\%, that no tricyclic products other than isomalabaricatrienol are generated by the SceERG7 Tyr510Phe mutant.

An aliquot of the HPLC purified fraction of isomalabaricatrienol was analyzed by GC-MS, both as the free alcohol and as the TMS-derivative. Electron-impact mass spectra are shown for the isomalabaricatrienol (Figure 5) and its TMS derivative (Figure 6), and the corresponding total ion chromatograms are shown in Figure 7 and Figure 8.
Figure 4.3. $^1$H NMR spectrum of HPLC-purified isomalabaricatrienol. Insets show expansions of methyl and olefinic regions.
Figure 4.4. $^{13}$C NMR and DEPT spectra of HPLC-purified isomalabaricatrienol.
Figure 4.5. $^1$H and $^{13}$C NMR signal assignments for isomalabaricatrienol.
Figure 4.6. 1H NMR spectrum of late fractions ($R_t$ 27 min) from the HPLC purification. Insets show expansions of methyl and downfield regions.
Figure 4.7. Electron-impact mass spectrum of underivatized isomalabaricatrienol.

Figure 4.8. Electron-impact mass spectrum of the TMS-derivative of isomalabaricatrienol.
**Figure 4.9.** Total ion chromatogram from GC-MS analysis of underivatized isomalabaricatrienol, purified by HPLC.

**Figure 4.10.** Total ion chromatogram from GC-MS analysis of TMS-derivative of HPLC-purified isomalabaricatrienol.
To observe the effects on lanosterol metabolism and accumulation of isomalabaricatrienol, SMY8 [pSLT 16.1] was also cultured in inducing medium with ergosterol supplementation (YPGHE). Under these conditions a ratio of isomalabaricatrienol:lanosterol of 87:13 was observed. The effects of harvesting time in the in vivo cultures grown without ergosterol supplementation was studied by harvesting cells at saturation (2 days) and one or two days past saturation. The ratios obtained varied from 80:20 to 90:10.
References


# Appendix A: List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>AacSHC</em></td>
<td><em>Alyciclobacillus acidocaldarius</em> squalene hopene cyclase</td>
</tr>
<tr>
<td>Ala, A</td>
<td>alanine</td>
</tr>
<tr>
<td>Amp</td>
<td>ampicillin</td>
</tr>
<tr>
<td>Asn, N</td>
<td>asparagine</td>
</tr>
<tr>
<td><em>AthCAS1</em></td>
<td><em>Arabidopsis thaliana</em> cycloartenol synthase</td>
</tr>
<tr>
<td>Carb</td>
<td>carbenicillin</td>
</tr>
<tr>
<td>CAS1</td>
<td>cycloartenol synthase</td>
</tr>
<tr>
<td>COSYDEC</td>
<td>decoupled $^1$H-$^1$H correlation NMR spectroscopy</td>
</tr>
<tr>
<td>D</td>
<td>dextrose</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleic acid triphosphate</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>Erg7</td>
<td>lanosterol synthase</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>G</td>
<td>galactose</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, Q</td>
<td>glutamine</td>
</tr>
<tr>
<td>Gly, G</td>
<td>glycine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GTAE</td>
<td>40mM tris base, 20 mM acetic acid, 1mM EDTA, 1mM guanosine</td>
</tr>
<tr>
<td>His, H</td>
<td>histidine</td>
</tr>
<tr>
<td>HMBC</td>
<td>heteronuclear multiple bond correlation</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-performance liquid chromatography</td>
</tr>
<tr>
<td>HSQC</td>
<td>heteronuclear single-quantum coherence</td>
</tr>
<tr>
<td>Ile, I</td>
<td>isoleucine</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>LiAc</td>
<td>lithium acetate</td>
</tr>
<tr>
<td>Lys, K</td>
<td>lysine</td>
</tr>
<tr>
<td>Leu, L</td>
<td>leucine</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular Dynamics</td>
</tr>
<tr>
<td>MeOH</td>
<td>methanol</td>
</tr>
<tr>
<td>MTBE</td>
<td>methyl tert-butyl ether</td>
</tr>
<tr>
<td>NBS</td>
<td>N-bromosuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>OS</td>
<td>(S)-2,3-oxidosqualene</td>
</tr>
<tr>
<td>OSC(s)</td>
<td>oxidosqualene cyclase(s)</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>Phe, F</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>SC</td>
<td>synthetic complete</td>
</tr>
<tr>
<td>SceErg7</td>
<td><em>Saccharomyces cerevisiae</em> lanosterol synthase</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>Scr</td>
<td>serine</td>
</tr>
<tr>
<td>SHC(s)</td>
<td>squalene hopene cyclase (s)</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, 1mM EDTA</td>
</tr>
<tr>
<td>TE8</td>
<td>10 mM tris-HCl (pH 8), 1mM EDTA</td>
</tr>
<tr>
<td>Thr, T</td>
<td>threonine</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>TMS</td>
<td>trimethylsilane / trimethylsilyl</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>tris(hydroximethyl)aminomethane hydrochloride</td>
</tr>
<tr>
<td>Trp, W</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Tyr, Y</td>
<td>tyrosine</td>
</tr>
<tr>
<td>Ura</td>
<td>uracil</td>
</tr>
<tr>
<td>Val, V</td>
<td>valine</td>
</tr>
<tr>
<td>YP</td>
<td>yeast extract/peptone</td>
</tr>
</tbody>
</table>
Appendix B: Sequencing Primers

CAS1 P1: 5’-CATATTGAGGGGCCCTAGCACCAT-3’
CAS1 P2: 5’-GTCCCGTGAACAAAGGTATTAATAT-3’
CAS1 P3: 5’-CTATCCAAGCTTTGATATCATTTCCG-3’
SM60A: 5’-TAAACAAAGAAATGCGCCGTT-3’
SM60B: 5’-CAACTGGAATGAAGCAGCA-3’
SM60C: 5’-GTGTTAGAAGACTGCCCTGG-3’
SM60D: 5’-GCAGATGCTCATGGTGATGG-3’
SM60E: 5’-ATAATGTGGAAACTGAAGAT-3’
SM60F: 5’-TCCCCGCGGTCATTCTCCTTGTTGCAA-3’
426GALF1: 5’-AACTAATACCTTTCAACATTTTCCG-3’
SM61H: 5’-TACCGGATCCCGCTCGAGG-3’
SM61I: 5’-TCCGTTGATGGTGCTGG-3’
SM61J: 5’-TACCGTATGTGGAGTAGACC-3’
SM61K: 5’-TGCACCGAGTGCGTTCCAGG-3’
T7: 5’-CTGAATACGACTCAGTATAG-3’