國 立 交 通 大 學

生物科技研究所

博士論文

利用突變策略針對氧化鯊烯環化酵素進行其結構與功能及其產物專一性/多樣性之研究

Mutagenesis Approach to Investigate the Structure-Function Relationships and Product Specificity/Diversity of Oxidosqualene Cyclase

研究生：劉媛婷

Student: Yuan-Ting Liu

指導教授：吳東昆 博士

Advisor: Prof. Tung-Kung Wu Ph.D

中華民國九十九年七月

利用突變策略針對氧化鯊烯環化酵素進行其結構與功能及其產物專一性/多樣性之研究

Mutagenesis Approach to Investigate the Structure-Function Relationships and Product Specificity/Diversity of Oxidosqualene Cyclase

研究生：劉媛婷　　Student: Yuan-Ting Liu

指導教授：吳東昆 博士　　Advisor: Prof. Tung-Kung Wu Ph.D

國 立 交 通 大 學

生物科技研究所

博士論文

A Dissertation

Submitted to Department of Biological Science and Technology

College of Biological Science and Technology

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

in Biological Science and Technology

Hsinchu, Taiwan, Republic of China

July, 2010

中華民國九十九年七月

摘要

突變策略被視為是一種強而有力的工具，其廣泛地應用於針對蛋白質的結構、功能、或反應機制的研究上。因此，在本論文中我們透過某些功能性已知的環化酵素，對於其胺基酸序列和蛋白質結構間的關係進行比較，並利用丙胺酸掃描式突變（alanine-scanning）或定點／飽和突變的策略（site-directed/saturated mutagenesis）並配合基因互補和產物收集/鑑定的方法，針對啤酒酵母菌中的氧化鯊烯環化酵素(*Saccharomyces cerevisiae* ERG7) 其活性區域內多個重要的胺基酸位置，它們在環化酵素所負責的複雜環化與重組反應機制中所扮演的角色加以確定。一開始的工作是針對活性區域內的組胺酸-234 號位置。當此位置被置換成其它不同性質的胺基酸時，我們透過所獲得的多樣性產物，包括單環、三環、和不同去質子化的四環產物進一步地闡明此胺基酸參予在環化酵素所催化之反應中的角色。此外，利用相似的突變策略對於其它重要的胺基酸位置，我們可以透過來自各突變株所收集到的不完全環化產物或是不同重組階段的替代產物加以解釋各個胺基酸所賦予在各環化重組過程中的重要性。同時，為了要更進一步的細看最後的去質子步驟，對於決定環化酵素其產物特異性的關係，我們建立多個連續定點突變株。其中一個突變酵素上第 384 號位置的蘇胺酸我們將之換成酪胺酸，而谷氨醯胺 450 位置將其換成組胺酸，最後我們將纈胺酸 454 位置換成異亮胺酸（也就是我們同時突變這三個胺基酸位置）。由實驗的結果發現，此突變株劇烈地改變它的特異性，從原本的羊毛硬脂醇合成酵素變成一個極度精密的帕克醇合成酵素，另外我們亦發現其產物中伴隨著兩個骨架已被下游連續酵素所修飾的產物。這個現象不僅反映當有不尋常量的三萜類化合物累積在酵母菌體時，其代謝流向的改變情況，而且也說明了下游酵素對於受質的選擇性效應。除了上述的突變實驗外，我們也同步針對植物中的一個三萜類合成酵素，即 β-香桂素合成酵素進行系列的研究。包括從碗豆 (*Pisum sativum*) 種子中進行分子選殖以獲

得 β-香桂素合成酵素的基因，並利用異質性的表現系統將其成功地在酵母菌中表現，其後我們利用電腦模擬的方式建立其同源性結構，並透過上述的定點突變實驗對於其重要的 19 個胺基酸位置，進行研究並分析突變效應後所產生的環化產物的樣式。雖然我們並沒有發現在各個突變株中有差異的非皂化脂質粗萃取物，但這一系列的實驗說明當 β-香桂素合成酵素中的重要胺基酸被置換時，會顯著地影響此酵素之催化活性。因此，透過上述的實驗結果我們可以對於不同的氧化鯊烯環化酵素其在環化/重組反應機制上提供更多的了解。

謝誌 (Acknowledgement)

論文即將付梓之際，心中滿是感恩。感謝主的安排，匯集如此大的恩典，讓我得以享受充實與精彩的過程，體會流淚灑種必歡笑收割的甘甜。

而論文能順利完成，最要感謝的是我的指導教授-吳東昆博士，八年前給了一個不是出自名門學府、毫無化學背景、懵懂無知的我加入這個實驗室的機會。其間老師對於整體研究方向、觀念的啟迪，實驗架構與邏輯的匡正，與對於結果的辯證和思維，著實讓我受益良多。對於學術研究的熱情與堅持，以及正面積極、不畏挫折的勇氣，更是學生學習與仿效的榜樣。另外對於學生求學態度的斧正，以及給予生活細節的叮嚀與照顧，更是帶動學生持續往前走的動力，於此獻上最深的敬意與謝意。口試委員李耀坤老師、鄭建中老師、林敬堯老師以及許鉦宗老師陪我度過修業期間的幾個重要考試，謝謝您們不吝提供寶貴意見以及對於疏漏處的指正，使得論文更臻完備。清大貴儀中心彭菊蘭小姐，因著有您對於 NMR 光譜上的幫忙，才得以將那微乎其為的化合物驗明正身，謝謝您。師母賴美伶小姐常以著開朗與樂觀的態度分享著實驗或生活上的建議，每每與您談話，總讓我像是重新充電般，再次獲得能量，得以再出發。

漫長的旅途中，雖然總是挫折積累，幸運的是，有一群小天使，一路陪伴、鼓勵著我，給予最溫暖的支持，使我得以堅強地走到現在，如今有個完美的結果，您們功不可沒。程翔學長是我的良師益友，秉著熱情與耐心伴著我成長。每當徬徨無措，總是理性的分析，像是盞明燈指引著我該走的方向。您那化腐朽為神奇的魔力，著實令人讚嘆，從您的身上我真的獲得許多。晉豪是我的好夥伴。頭腦聰穎的你，總是給予我不管是實驗還是生活上很大的幫忙。真的，還好有你 Cover 我一切的不足。晉源是開心果。總是能讓煩悶的過程增添了許多的歡笑與樂趣，有顆赤子之心的你，是很有魅力的喔！和小紅妳為室友的日子，夜裡不論是對於實驗的討論，或是小女人間的話家常、說八卦，還是追星、瘋偶像劇，都是我難忘的回憶。好麻吉怡親，總是給予我最真誠的關心。說心事、聊是非，一起逛街、買拍賣，妳為我的研究生涯灑下了不一樣的精采。我的好鄰居文祥，超級行動派的你總是實現每一個不可能。那段南征北討，通霄”學習”的歲月，現在想起來還是覺得很酷！裕仁總是適時的給予我加油、打氣與肯定，重要時刻也沒缺席，你就是一個這麼樣體貼的小男生。斯文的小宇學長在最初給了剛進實驗室懷著忐忑不安心情的我那抹善意的微笑，至今我還記得。大景傻得可愛的氣息，讓我的生活充滿著爆笑噴飯的事件。熱心的宏城和裕國是我的同窗伙伴，和你們之間的切磋討論，讓我獲益匪淺。還有貼心的文鴻，不管是實驗還是生活的細節，總是幫忙許多。告訴你喔，我真的覺得你很厲害，所以要繼續堅持喔！另外同屬 Cyclase Project 的學弟妹美婷、文暄、皓宇、采婷、亦諄與宗璟，更是我學習生涯缺一不

可的好幫手。猶記得那段熬夜通銀染的日子，絞盡腦汁地玩著連連看，到最終產物鑑定，並得以發表。這一切成就感的獲得，全因著你們的認真與執著，所收割的豐碩果實，真的謝謝你們了。天昶是最挺我的學弟，總是努力認真的幫忙我完成每一個實驗的假想。時常與奕齊你激辯、交鋒，也著實讓我成長。可愛的卡通人靜婷，則是個認真有規劃的好學妹。怡臻的貼心，常為我枯竭的靈感注入新的想法。新新人類欣怡，總是帶給我許多年輕人的新想法。和世穎一天一音樂的遊戲，振奮夜裡萎靡的精神，也給了漫長的論文撰寫過程中一個期待。因著你們的年輕有活力，使得實驗室充滿著生氣。也謝謝你們在庶務上面的幫忙，分擔了許多的麻煩，真的謝謝了。而分屬其它計劃的衣鵑、Mili、妍希、佳宜、令宗、宏明、庭翊、禕婷、書涵、宜芳與欣芳，謝謝你們豐富了我的生活。

另外還要感謝我的好朋友們- 建勝一路走來對我的關心與照顧，如果當初沒有你，我想也不會有現在的我，謝謝。和智凱你那段慢跑的時光，以及之後書本小說和音樂小品的分享，是我放鬆休憩的最佳良藥。千惇、佩陵與怡萱，常常給予我加油、打氣，妳們給的溫暖，讓我明白這並不是個人單打獨鬥的旅程。晉嘉品容雙寶，常常帶給我意外的捧腹大笑，使得原本抑鬱的心情展為笑顏。

最後，特將本論文獻給我最親愛的家人- 爸爸、媽媽以幾近縱容的態度，讓我自由探索想走的方向。每每失去勇氣，您們的溫柔撫慰，總是再一次讓我獲得站立的力量。因著您們的全然支持，是使得資質平庸的我，能夠獲得博士學位，最重要的基石。妹妹和弟弟則是我最大的驕傲，更是促使我不斷向上的原動力。還有已在天堂的爺爺和奶奶，您們溫良謙讓的處世態度，尤其令我銘記在心，也是我日後為人處事的最大的後盾。還有要謝謝親朋好友給予的關心與代禱。

感謝所有關心我，陪伴著我的人。小小的版面，實在不足以聊表我最深忱的謝意。

『參透為何，才能迎接任何』，願以此期勉自己的未來，能夠發揮所學，並開創或將其應用至每一個可能性。加油!

媛婷 謹誌於
交通大學生物科技研究所
中華民國九十九年七月二十六日

Abstract

The mutagenesis is regarded as a powerful tool for investigating the structure, function, and reaction-mechanism relationships of proteins. Sequence alignment with other known cyclases and various forms of mutagenesis were used to identify and study catalytically important residues in *Saccharomyces cerevisiae* ERG7. Using mutagenic techniques coupled with genetic complementation and product characterization, it became possible to further characterize the cyclization and rearrangement mechanism of 2,3-oxidosqulene. Several mutations of the His-234 residue in ERG7 generated diverse product profiles with various monocyclic, tricyclic, and tetracyclic products while similar mutagenesis strategies on other catalytically important residues resulted in derailed cyclization and alternative deprotonation in other critical steps of the cyclization cascade. A series of site-directed mutations were also made to probe the crucial residues involved in the final deprotonation stage of the reaction for product specificity. For example, the ERG7$^{T384Y/Q450H/V454I}$ mutant changed its product specificity from the original lanosterol synthase into a parkeol synthase accompanied by two scaffold-modifying products that were generated speculatively by additional tailoring enzymes. The trend of metabolite flux changed when unusual levels of triterpene accumulated in the yeast cell and impacted the substrate selectivity of some downstream enzymes. A parallel experiment was conducted with a triterpene synthase, β-amyrin synthase from *Pisum sativum* using functional expression by yeast and site-directed mutagenesis on 19 residues. Examination of its product profile revealed no divergence of the non-saponifiable lipid patterns among any of the mutants, suggesting that the exchange of the important functional residue of β-amyrin synthase might influence its enzymatic

activity dramatically. The aforementioned results, when combined, provides for a better understanding for the cyclization and rearrangement mechanism of various oxidisqualene cyclases.

Table of Contents

Table of Figures

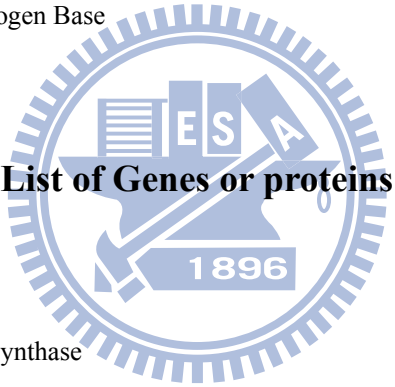
Table of Schemes

Table of Tables

List of Abbreviations

A, Ade	Adenine
A. acidocaldarius	*Alicyclobacillus acidocaldarius*
A. thaliana	*Arabidopsis thaliana*
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
bp	Base Pair
cDNA	Complementary DNA
Cys	Cystine
ddH_2O	Double Distilled Water
DEPC	Diethylpyrocarbonate
dNTP	Deoxynucleoside triphosphate
GC-MS	Gas Chromatography-Mass Spectromerty
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
H, His	Histidine
H. sapiens	*Homo sapiens*
HPLC	High Performance Liquid Chromatography
Ile	Isolucine
kb(s)	kilobase(s)
L, Lys	Lysine
LA	Lanosterol
LB	Luria-Bertani
Leu	Leucine
M, Met	Methonine
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NSL	Non-Saponifiable Lipid
O. europa	*Olea europa*
OS	Oxidosqualene
P. ginseng	*Panax ginseng*
P. sativum	*Pisum sativum*
PCR	Polymerase Chain Reaction

Phe	Phenylanine
Pro	Prolne
rpm	revolutions per minute
RT	Room Temperature
S. cerevisiae	*Saccharomyces cerevisiae*
Ser	Serine
ssDNA	Single-stranded DNA
T, Trp	Tryptophan
Thr	Threnine
TLC	Thin Layer Chromatography
Tyr	Tyrosine
U, Ura	Uracil
Val	Valnine
WT	Wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YNB	Yeast Nitrogen Base



List of Genes or proteins

βAS	β-amyrin synthase
CAS	Cycloartenol synthase
ERG1	Squalene epoxidase
ERG6	Sterol C-24 methyltransferase
ERG7	Oxidosqualene Cyclase ; Lanosterol Synthase
ERG11	Lanosterol 14α-demethylase
ERG25	Sterol C-4 demethylase
ERG26	Sterol C-3 dehydrogenase
LUS	Lupeol synthase
OSC	Oxidosqualene cyclase
SHC	Squalene-hopene cyclase
SMT	Steroll methyl-transferase
URA3	Orotidine-5'-Phosphate (OMP) Decarboxylase gene

CHAPTER 1

General introduction

1.1 Overview of oxidosqualene cyclase

The 2,3-oxidosqualene cyclization is the most remarkable example of a single enzyme-catalyzed reaction that triggers a common polyolefin for the production of a vast diversity of tetracyclic and pentacyclic products with exquisite control of stereoselective cyclization and skeletal rearrangements. This impressive enzymatic reaction catalyzes an acyclic squalene or 2,3-oxidosqualene with one or no stereogenic center into polycyclic sterols and triterpenes that contain multiple rings and several chiral centers, along with at least 10 formal covalent-bond cleavages and formations. The cyclization cascades include an initial epoxide protonation, four or five cation-olefin mediated ring annulations, multiple hydride/methyl group rearrangements, and the final termination either by deprotonation or water addition.[1] The lanosterol is further converted to the end product, cholesterol, via additional 18 enzymatic reactions in the human steroid framework, which serves as the membranous components, precursors of steroid hormones, and other secondary metabolites. Certainly, the oxidosqualene cyclase (OSC) is responsible for the most amazing step among the natural chemical reactions (Figure 1.1).

Over a half century ago, this powerful and efficient chemical transformation catalyzed by OSC intrigued and fascinated scientists. The first rudimentary model was proposed in 1934, and, in addition, other different radioisotopic feeding experiments were used to demonstrate the final carbon and hydrogen source of cholesterol.[2,3] In 1953, Woodward and Bloch suggested a hypothesis regarding lanosterol formation

derived from squalene.[4] However, further experiments demonstrated that 2,3-oxidosqualene, rather than squalene, is the direct intermediate in the biosynthesis of lanosterol in which the 3*S* isomer of oxidosqualene was used as the exclusive substrate.[5,6] Moreover, the more detailed information about lanosterol production, such as the [1,2]-hydride shifts and the migration of methyl groups, was examined from the trapping of truncated intermediates in the substrate-analogue mediated cyclization.[5-8] These early pioneering chemical studies of OSC provided fundamental background information concerning its biological importance and the stereochemical course of the cyclization.

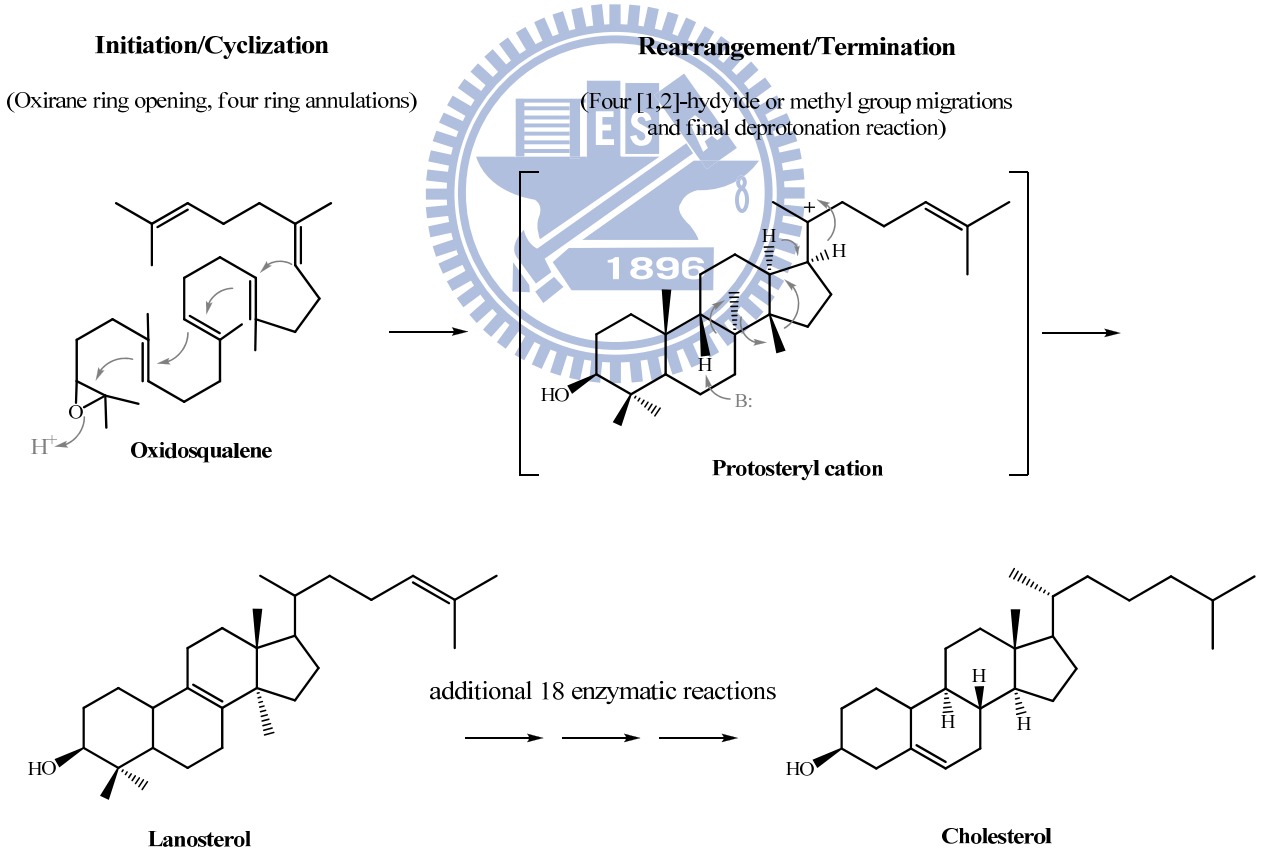


Figure 1.1. The complexity of the cyclization cascade from 2,3-oxidosqualene to lanosterol and the end product, cholesterol, in the biosynthesis pathway for cholesterol in human.

A wide array of tetracyclic or pentacyclic backbones is generated by the species-dependent cyclization manners. It is widely accepted that the distinct stereochemical configuration of the second cyclohexyl B-ring of 2,3-oxidosqualene in the active-site cavity of the enzyme resulted in two major cationic intermediates for its corresponding cyclization pathways. In the protosteryl cationic pathway, the six-membered B-ring is used as the "boat" form, whereas the "chair" form of the cyclohexyl B-ring takes place in the dammarenyl cationic pathway. Thus, either for lanosterol production in animals or for cycloartenol production in plants, the substrate, 2,3-oxidosqualene, proceeds via an energically-unfavorable pre-folded conformation, i.e., the "chair-boat-chair conformation" to produce a protosteryl C-20 cation, followed by a series of hydride and methyl group shifts and the elimination of an individual final proton. In parallel, after generation of the tetracyclic dammarenyl C-20 cation from the "all-chair" conformation, the substrate is triggered to form dammaradienol, lupeol, α-amyrin, or β-amyrin in the higher plants, algae through similar rearrangements, ring expansions, and final elimination steps (Figure 1.2).[1]

The cyclization reaction catalyzed by oxidosqualene cyclases

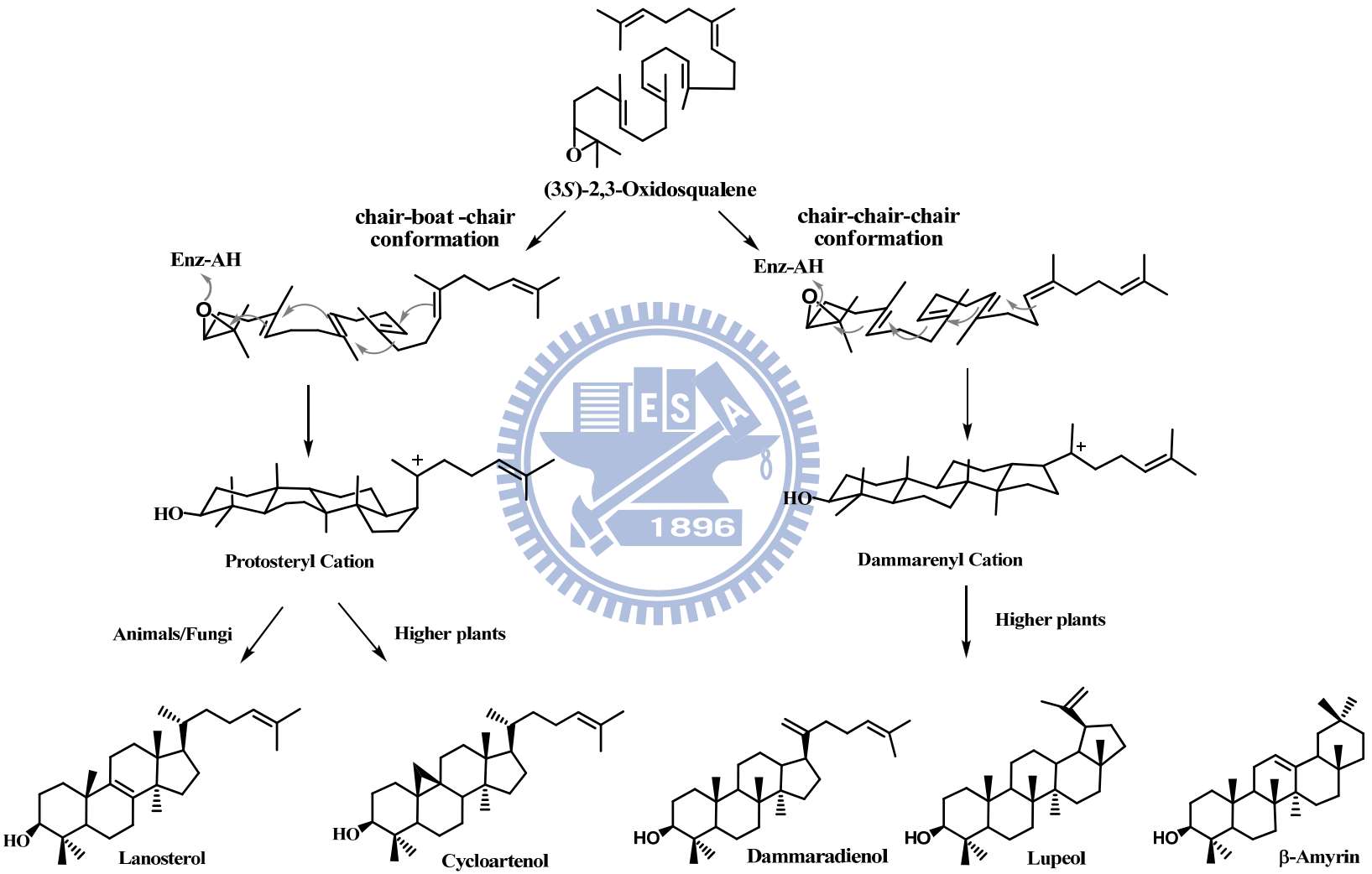


Figure 1.2. Product diversity of oxidosqualene cyclase is acquired by the species-dependent cyclizations and stereochemical configuration of carboncationic intermediates.

1.2 Historical hypothesis of the cyclization mechanism

1.2-1 Mechanistic and stereochemical insights relative to the 2,3-oxidosqualene cyclization cascade

Converting a mobile and linear polyolefinic substrate into tetracyclic and pentacyclic triterpene products is a highly exothermic chemical reaction with a low activation energy requirement.[9] Thus, it is hard to control tendency for complex reactions in the non-enzymatic cyclization of oxidosqualene.[9] In the biological system, OSC can trigger an efficient cyclization reaction and lead to a variety of product specificities by overcoming the energy barrier and diminishing the activation energy. Therefore, many experiments have been performed, mainly the synthesis of substrate analogues or inhibitors, to understand the relationship between the cyclization mechanism and the enzyme itself.

According to the "biogenic isoprene rule,"[10] cyclizing 2,3-oxidosqualene to lanosterol should proceed via a direct, concerted, and "non-stop" processes. However, van Tamelen and his-coworker proposed that the cyclization proceeded through a series of discrete, conformationally-rigorous, and partially-cyclized carbocationic intermediates.[11,12] These debates have been clarified by direct experimental evidence.

The reaction rate of the cyclization of various polyene monoepoxides with a Lewis acid suggested that the cleavage of the oxirane ring, accompanied by A-ring cyclization, occurs via a high degree of anchimeric assistance from the neighboring double bond (C-6/C-7).[12] The enzymatic cyclization of two 2,3-oxidosqualene analogs with the C-6 methyl group replacement by either hydrogen or chloride, i.e, 6-desmethyl-2,3-oxidosqualene and 6-chloro-6-desmethyl-2,3-oxidosqualene, demonstrated that electrophilic oxirane-ring cleavage and cyclization of the A-ring are concerted and initiated by a particularly acidic and highly conserved aspartic acid

residue, Asp-456.[13] Moreover, the V_{max}/K_M values of the enzymatic conversion of 2,3-oxidosqualene, 6-desmethyl-2,3-oxidosqualene, and 6-chloro-6-desmethyl-2,3-oxidosqualene are 138, 9.4, and 21.9, respectively. This implied that the π-nucleophilicity of the proximate double bond might influence the rate of cleavage of the oxirane ring.[13]

The boat conformation of the B-ring might occur temporally once a positive charge has built at the C-5 of the A-ring. It is still unknown exactly how the cyclase enzyme mediates this process. One possibility is that cyclase may lower the activation energy of the boat form by delivering a negative point charge to the α–face at C-8.[14] In addition, the isolation of "trapped" products from natural sources has provided direct experimental evidence that cyclization might proceed via different, stepwise, cyclized stages (Fig 1.3).[15,16]

A secondary 6,6,6-tricyclic cation with a unfavorably thermodynamic tendency was proposed to exist via an anti-Markovnikov addition during the closure of the six-membered C-ring.[14] However, a 6,6,5-tricyclic product was found via a thermodynamically-favorable, tertiary, cationic intermediate in the cyclization of 18,19-dihydro-2,3-oxidosqualene (Fig 1.4).[17] In addition, the enzymatic cyclization of 20-oxa-2,3-oxidosqualene produced a 6,6,5,4-tetracyclic compound that further exhibited the precursor scaffold with a 6,6,5-ring (Fig 1.5). These observations strongly indicated that the cyclization of 2,3-oxidosqualene must proceed via a discrete tricyclic cation, from a Markovnikov closure of the cyclopentylcarbinyl cation, followed by a subsequent ring expansion to form a tricyclic anti-Markovnikov cyclohexyl C-ring intermediate cation.[18,19]

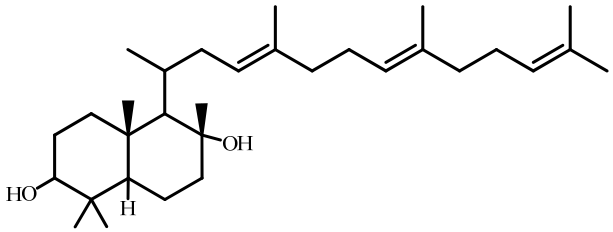
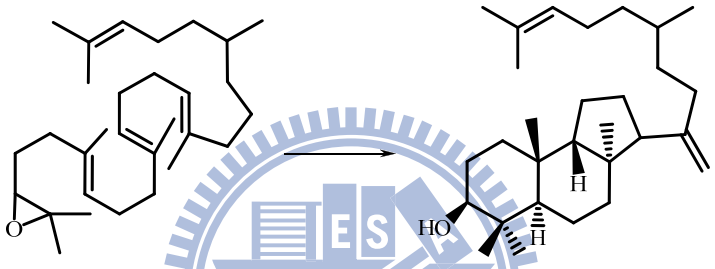
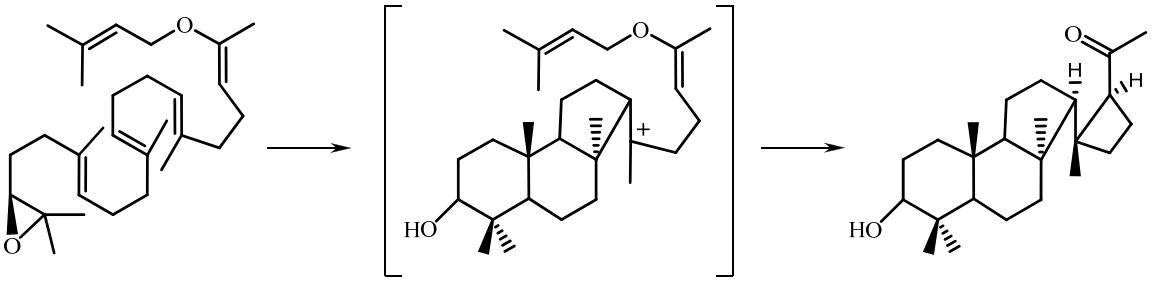


Figure 1.3. Isolation of a novel bicyclic triterpenoid from gum mastic of *Pistacia lentiscus*[15,16]



18,19-Dihydrooxidosqualene 6,6,5-Fused ring

Figure 1.4. Conversion of 18,19-dihydrooxidosqualene into a 6,6,5-ring fused product by rat liver OSC[17]



20-Oxa-2,3-oxidosqualene *p*-Bromobenzoate derivative

Figure 1.5. Enzyme converted a substrate analogue, 20-oxa-2,3-oxidosqualene, into a product with a *p*-bromobenzoate derivative.[18]

The stereochemistry at C-20 during the oxidosqualene cyclization has also been illustrated.[1,20] Isolation of the trapped protosteryl derivatives from the enzyme-catalyzed cyclization of 20-oxa-2,3-oxidosqualene and (20*E*)-20,21-dehydro-2,3-oxidosqualene, respectively, demonstrated that the stereochemical β–side chain conformation, but not a α–side chain conformation, is formed at C-17 of the protosteryl derivative (Figure 1.6). These findings further explained that only a smaller rotation angle (60^{o}) of the C-17 side chain is required prior to the leading of the natural 20*R* stereo configuration during the D-ring closure of the C-20 protosteryl cation.[21,22]

On the other hand, lanosterol-like products that contained an unnatural 6,6,5-tricyclic scaffold accompanied by various methyl group migrations, were also found in the cyclase-triggered catalysis.[23] In parallel, the cyclase enzyme could convert the 2,3-oxidosqaulene analogues, which lack the methyl substituent at C-10 and C-15, into a cyclized product without rearrangement of carbon and hydrogen skeleton (Figure 1.7).[24] This phenomenon indicated that the methyl group at C-10 of 2,3-oxidosqualene is very critical for the cyclization-rearrangement step.[25] After the skeletal rearrangement goes through a series of [1,2]-hydride/methyl group shifts, the lanosteryl C-8 or C-9 cation is formed. The cyclization is terminated by a proton removement step. The deprotonation of the intermediate cation was performed by either water addition or proton abstraction from a well-positioned enzymatic base, which is elegantly controlled by the cyclase enzymes to avoid the disruption of the early product prior to the final reaction step. Therefore, efforts were made to understand the more-detailed, and diagrammatic elucidation of the oxidosqualene-catalyzed cyclization reaction. The extensive mechanistic hypothesis of the formation of lanosterol is summarized in Figure 1.8.

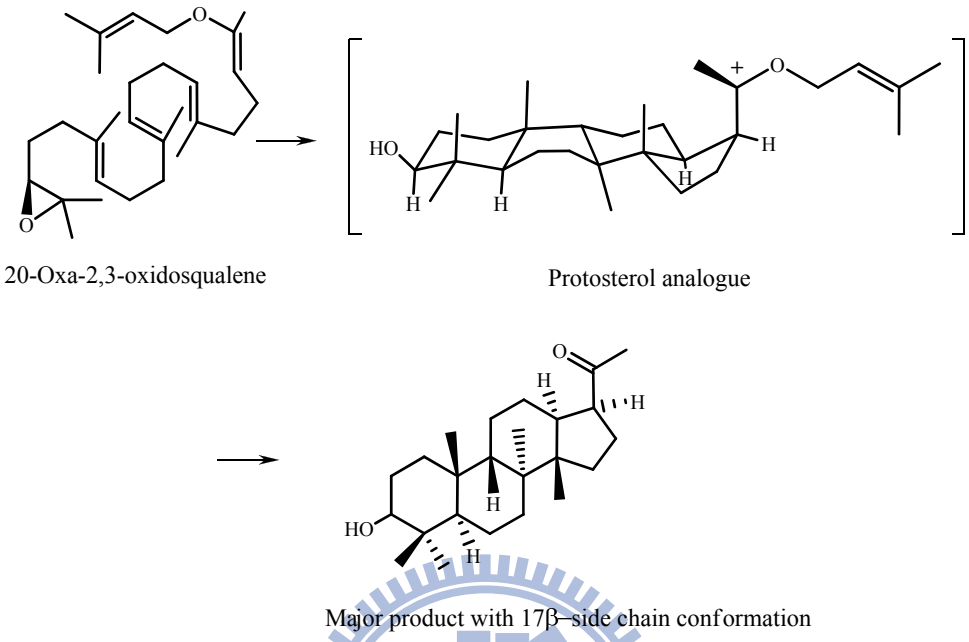


Figure 1.6. Enzymatic transformation of an 2,3-oxidosqualene analogue, 20-oxa-2,3-oxidosqualene, into a protosterol derivative with a 17β–side chain conformation.[26]

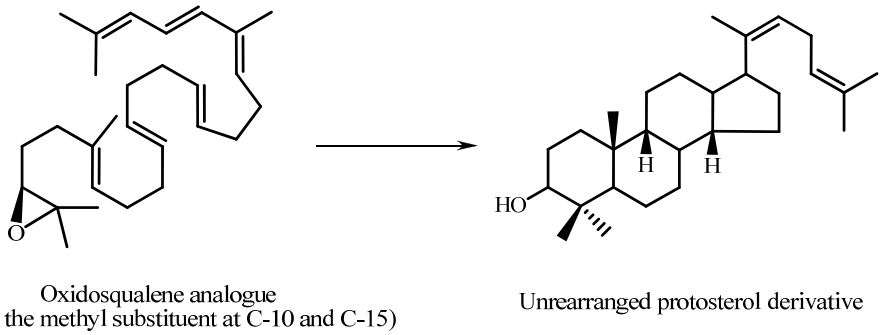


Figure 1.7. Enzyme conversion of the 2,3-oxidosqualene analogue, which lack the methyl substituent at C-10 and C-15, into an unrearranged protosterol derivative.[24]

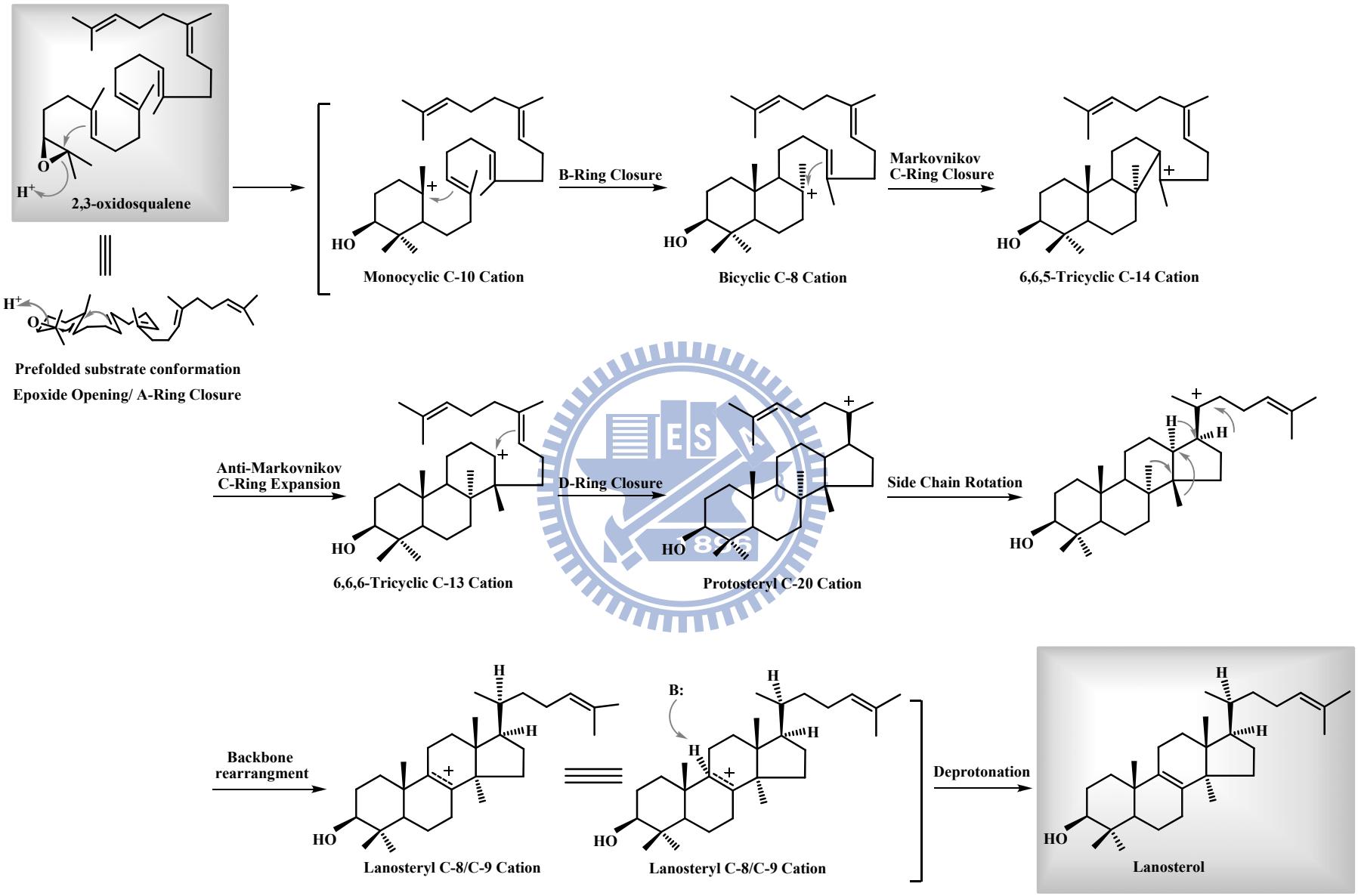


Figure 1.8. Hypothesized cyclization cascade of 2,3-oxidosqualene

1.2-2 The theoretical models of cyclase enzymes

Due to the intrinsic difficulty of purifying active cyclase enzymes, the critical issue concerning the enzymes themselves is illusive and poorly understood. Many theoretical models of enzymatic processes have been developed. Ourisson established the phylogenetic relationship among all of the known cyclase families.[27] It was suggested that, in the beginning, a primitive ancestral cyclase that contained several essential elements was responsible for the enzyme-catalyzed cyclization/rearrangement reaction. Through the divergent molecular evolution of these cyclase enzymes in different species, the shape of the cavities of active sites or some other critically important characteristics or functionalities were changed by mutations.[27] Extensive research focused on substrate specificity and the stereochemistry of the cyclization reaction promoted Johnson to propose that cyclization is initiated by proton donation from a Lewis acid residue on an epoxide ring and that three anionic sites in the enzymatic active-site cavity are involved in axial delivery to form ion pairs, thereby stabilizing the highly-energetic and developing cationic center of the cyclized substrate (Figure 1.9).[28,29]

Comparison of the amino acid sequence among all of the known cyclases in bacterial sources or in the eukaryotic cell, one unique motif was observed for five or six repetitions, i.e., the "Q-W motif" with the highly-conserved sequence, [(K/R)(G/A)X_{2-3}(F//Y/W)(L/I/V)$_3X_3QX_{2-5}$GXW] (Figure 1.10).[29,30] One of the "Q-W motifs" was further labeled with a mechanistic inhibitor, suggesting that these "Q-W motifs" might play a role in stabilizing the enzyme structure through surface-connecting α-helices or by mediating the carbocationic intermediate through cation-π interactions.[31,32,33] According to the "*Aromatic hypothesis,*" the essential feature of the enzymatic active site is the presence of tryptophan and tyrosine, which possess the electron-rich indole and phenol groups. In addition, a positively-charged

transition state and highly-energetic intermediates occur during the cyclization/rearrangement processes.[34] Thus, the cation-π interaction from these highly aromatic functional groups is responsible for stabilizing the electron-deficient cationic intermediates (Figure 1.11).[34]

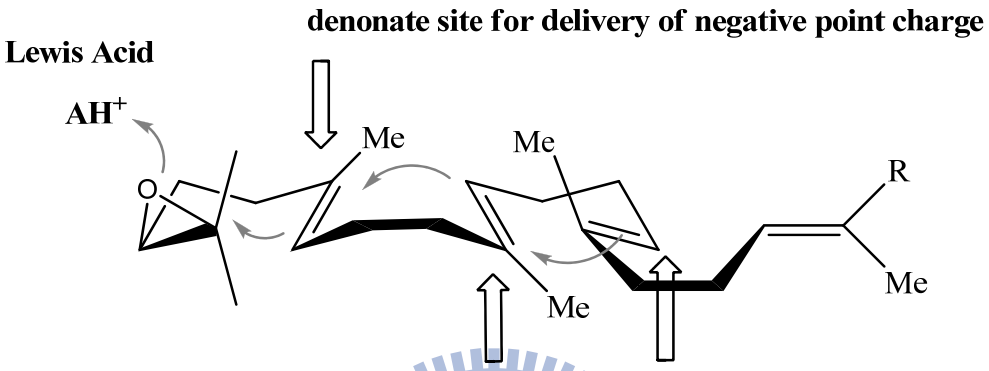


Figure 1.9. Three cation-stabilizing auxiliaries (vertical arrows) that were proposed by Johnson in the cyclase enzyme[28,29]

QW-1			**QW-4**		
Homo	672	RGVRCLLEKQLPNGDW	Homo	485	DSVNLLLSLQSKKGGF
Pisum	704	RAAKLLINSQLEQGDW	Pisum	516	DSVNLLLSLQSKKGGL
Arabidopsis	702	RAARYLINAQMENGDF	Arabidopsis	514	EAVNVIISLQNADGGL
Saccharomyces	675	RGIDLLKNRQEESGEW	Saccharomyces	488	EGIDVLLNLQNI-GSF
QW-2			**QW-5**		
Homo	614	RACDFLLSRQMADGGW	Homo	126	EIVRYLRSVQL-PDGGW
Pisum	642	KGVKFLLTTQREDGGW	Pisum	150	EILRYIYCHQN-EDGGW
Arabidopsis	640	KACEFLLSKQQPSGGW	Arabidopsis	149	EMRRYLYNHQN-EDGGW
Saccharomyces	617	KGCDFLVSKQMKDGGW	Saccharomyces	127	ELIRYIVNTAHPVDGGW
QW-3			**QW-6**		
Homo	562	QGLEFCRRQQRADGSW	Homo	79	N-GMTFYVGLQAEDGHW
Pisum	593	NAVRFLEDTQTEDGSW	Pisum	99	RRGTHHLATLQTSDGHW
Arabidopsis	591	KAVKFIESIQAADGSW	Arabidopsis	98	KRGLDFYSTIQAHDGHW
Saccharomyces	568	IAIEFIKKSQLPDGSW	Saccharomyces	79	NGASFFKLL-QEPDSGI

Figure 1.10. Summary of the highly-conserved QW motifs [(K/R)(G/A)XX(F/Y/W)(L/I/V)XXXQXXXGXW] in OSC for four species: Homo, *H. sapiens* OSC; Pisum, *P. sativum* BAS; Arabidopsis, *Arabidopsis thaliana* OSC (cycloartenol synthase); Saccharomyces, *S. cerevisiae* OSC; The residues that occur frequently are shown in respective colors, and the hyphens indicate the gap introduced to maximize the alignment.

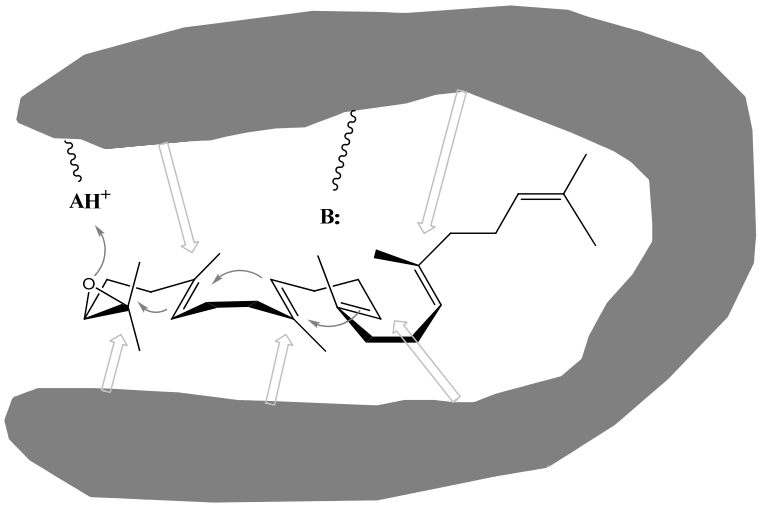


Figure 1.11. Aromatic hypothetical model for the cyclization mechanism (Vertical arrows show the proposed stabilizing delivery force for the carbonium ions)

Also, the region of the sequence of amino acids from Val-454 to Glu-461, i.e., the VXDCATE motif in the eukaryotic cyclases, which correspond to the DXDDTAE motif in the bacterial cyclases, has also been characterized.[31,35] According to the results obtained from affinity-labeling experiments, this peptide was specifically labeled with the first mechanism-based, irreversible inactivator, 29-methylidene-2,3-oxidosqualene (29-MOS), indicating the critical functional importance of this region in the oxidosqualene cyclization process.[35] Successes in similar experiments have provided extensive and valuable insight into the mechanism of 2,3-oxidosqualene cyclization.[36] The suicide substrates, which mimic monocyclic or bicyclic reactive intermediates, were specifically labeled on the fragments from Lys-489 to Lys-512 in *Sacchomyces cerevisiae* (*S. cerevisiae*) ERG7.[21,36] In addition, the substrate analogues, which imitate the C-20 cation intermediates, revealed a capability for covalent modification at His-234 and Trp-232 residues in the cyclase enzyme (Figure 1.12).[21,36]

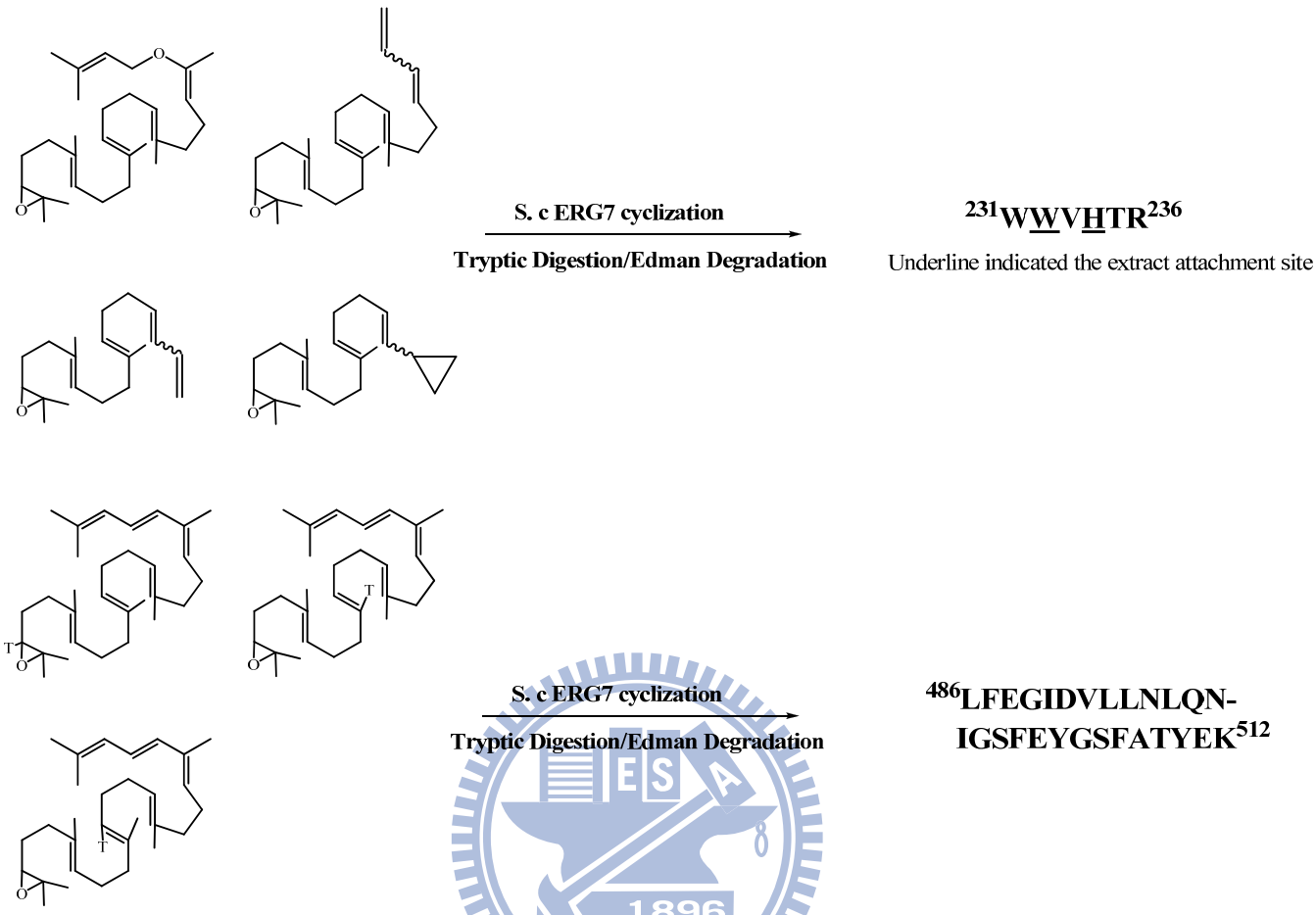


Figure 1.12. Affinity labeling experiments for cyclase enzymes with mechanistic suicide inhibitors

In addition, selective side-chain modifying agents have also been used to provide clues concerning the potential involvement of amino acid residues in the important aspects of the enzyme's function. The compounds, *p*-(chloromercuri)-benzenesulfonic acid and *N*-ethylmaleimide, the SH-selective modifying reagents, presented strong inhibitory activity toward cyclases.[37,38] It indicated that the presence of cysteinyl residue in the activity-site cavity of the enzyme is crucial for catalysis. Another chemical inactivation experiment was carried out to clarify the essentiality of histidine residue in cyclases.[38] The discordant results from the chemical inactivation experiment on various source cyclases further imply that there are subtle differences

between the enzyme-active sites among these homologous cyclases.[38,39]

These early studies of cyclases focused mainly on the use of a variety of substrate analogues, which mimic various cyclic intermediates, to estimate the mechanism of structural requirement, the substrate's specificity, and the stereochemistry of various stages in this chemical reaction. In addition, the conscientious analysis of the primary amino acid sequence among all known cyclases, in combination with the affinity labeling strategy, further elucidated the essential components of cyclase catalysis. However, the strict definition of which functional residues must be manipulated and how they are involved in the complex cyclization stages is still unknown at this time.

1.3 Crystallization and structural characterization of cyclases

Cyclase enzymes are considered to be an integral membrane protein based on the biophysical characterization of OSC for either cellular localization or protein solubility.[40] In order to access, steer, and release the lipidic substrate and the synthetic product, these enzymes permanently occupy the lipid bilayer. The membrane-binding region employs a large hydrophobic channel that connects the active-site cavity of the enzyme and the non-polar portion located on one leaflet of the bilayer. Thus, the protein structure is regarded as a path for interacting with the target substrate and delivering the reactants.[9]

In 1997, the X-ray structure of squalene-hopene cyclase (SHC) from *Alicyclobacillus acidocaldarius* (*A. acidocaldarius*) identified the cyclase as a homodimeric monotopic membrane protein.[33] The polypeptide chain of each subunit is organized into two sets of α-α barrel domains, which further jointly form a dumbbell-shaped molecule (Figure 1.13). The unique sequence fingerprint, Q-W repeats, is distributed throughout the surface of the enzyme structure, where the

functional residues of these amino acid repeats form an intricate hydrogen-bonding network and connect all outer helices of the α-α barrel domains by hydrophobic interactions.[33,41,42] It proposed that this arrangement might stabilize the whole protein structure against the reaction enthalpy that is released during the cyclization reaction.

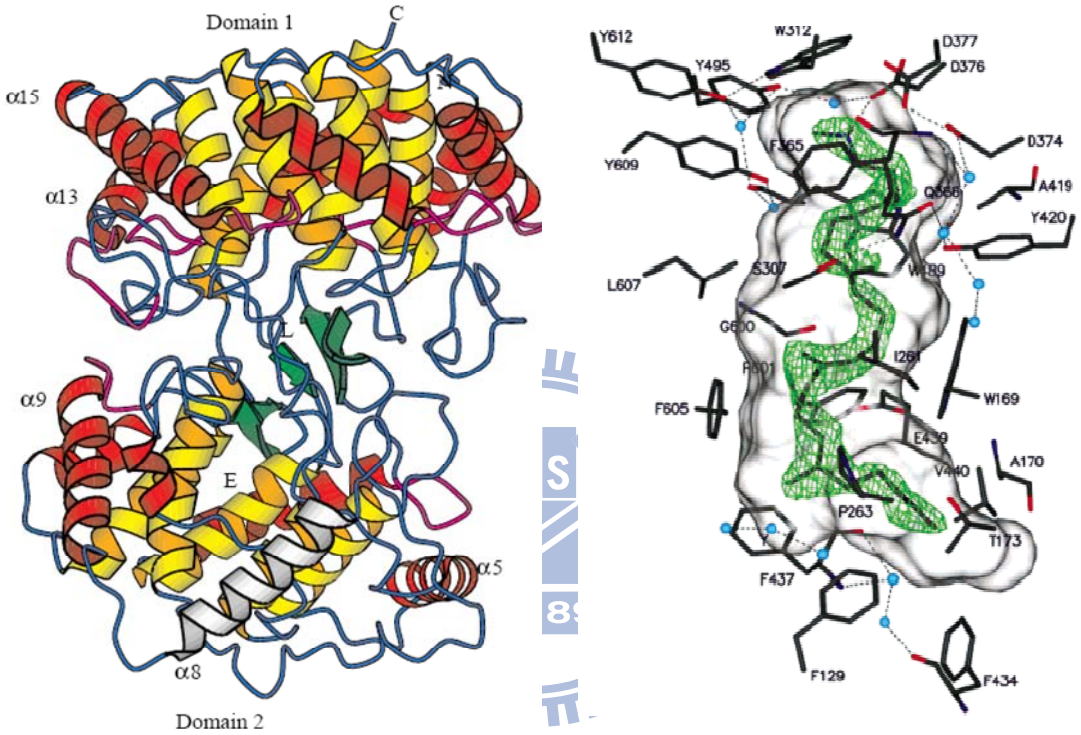


Figure 1.13. (Left) The overall structure of *A. acidocaldarius* SHC[33] (Right) The putative active-site cavity in the SHC[41]

Moreover, a large hydrophobic cavity located deep in the enzymatic active site of SHC is enclosed by long loops and several small β sheets that link barrel domains.[33] Several conserved aromatic amino-acid residues that are aligned in suitable positions in this cavity were assumed to access/constrict the mode of the substrate, thereby determining the manners in which the substrate is folded and stabilizing the individual carbocationic intermediates during the cyclization reaction.[9] Two smaller polar patches formed by networks of hydrogen-bonding around Asp-376 (at the top of

cavity) and Glu-45 (at bottom of cavity) were believed to participate in the initiation of the cyclization cascade via olefin protonation and in the termination of the reaction by effectively abstracting the proton or hydroxylating the cationic position with water, respectively.[9,33]

The structure-based mutagenesis of conserved amino acid residues in the active site cavity was subsequently examined based on the structure information of *A. acidocaldarius* SHC.[43-48] Functional residues that participated in the SHC-catalyzed cyclization from squalene to hopene/hopanol were determined by isolating the resulting early-truncated products from different protein mutants. [43-48]

The SHC in prokaryote is the counterpart of the OSC in eukaryotic cells, regardless of the variant amino acid sequences and the distinct enzyme-catalyzed mechanisms. Thus, the early three-dimensional structure and the mutagenesis studies of SHC provided a shallow insight for the first time that helped researchers to understand the relationship between functional residues in the active-site cavity of cyclase enzyme and the enzymatic reaction mechanism of the OSC in the eukaryotic cells (Figure 1.13).

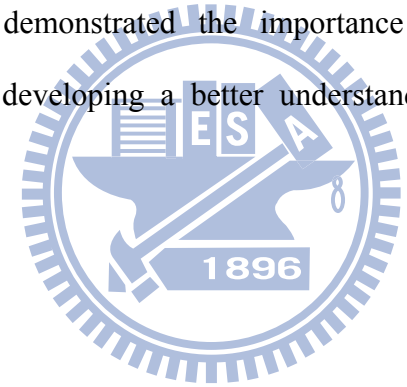
Recently, the long-awaited X-ray crystal structure of human OSC co-crystallized with the reaction product, lanosterol, was successfully determined (Figure 1.14).[49] This was a significant milestone for understanding the molecular interactions between lanosterol and the enzymatic active site, and it also provided an important additional snapshot of the triterpene polycyclization cascades. Overall, the structure of human OSC has a topology that is almost identical to that of *A. acidocaldarius* SHC. OSC is a monotopic membrane protein, and its membrane-inserted plateau with a diameter of 25 Å serves as hydrophobic substrate entrance channel that connects the active-site cavity and non-polar interior of the membrane. On the basis of crystallographic analysis of human OSC structure, this channel is constricted by Tyr-237, Cys-233,

Ile-524, and the two stained loops from 516 to 524, and from 679 to 699. Once the immigrated substrate moves into the appropriate position in the enzymatic active-site cavity, the polycyclization reaction can be triggered. Asp-455, which is located at the top of the cavity and form a hydrogen-bonding network with the 3-hydroxyl group of the lanosterol, is thought to be responsible for the initiation of the reaction. The surrounding Cys-456 and Cys-533 residues also act as hydrogen-bonding partners with the Asp-455 to increase its acidity. Thus, Asp-455 performs the role of a general acid to donate a proton to the epoxide oxygen of the substrate and promotes the initiation of the cyclization cascade.[33,41]

As proposed in a previous theoretical model, the abundant, highly-electron-riched aromatic residues, which are located in the hydrophobic enzymatic cavity, are responsible for stabilizing the carbocation intermediate in different stages (Figure 1.15). Phe-444, Tyr-503, and Trp-581, which are oriented around the A/B ring of the lanosterol molecule, might stabilize the monocyclic C-10 and bicyclic C-8 cation (lanosterol numbering), respectively. Tyr-98 is located spatially at a better position to push the methyl group at C-8 below the molecular plane and enforces the formation of the energetically-unfavorable boat conformation of the B-ring. During formation of the C-ring, His-232 and Phe-696 are in suitable orientations to stabilize the unusual secondary anti-Markovnikov cyclohexyl carbocation at C-14. It has been reported that Trp-169 and Phe-605 in SHC stabilize the long-lived 6,6,6,5-fused tetracyclic secondary cation as required for the formation of the E-ring of hopene. The lack of aromatic residues in the corresponding position of the OSC may lead to the termination of the cyclization cascade at the tertiary D-ring protosteryl cation.[49] Then the protosteryl C-20 cation undergoes a series of skeletal rearrangements to form the lanosteryl C-8/C-9 cation. Finally, His-232 and Tyr-503 are possible candidates for the role of a general base that extracts the proton from the lanosteryl C-8/C-9 cation to

terminate the cyclization/rearrangement reaction cascades.[49-51]

The insight on the biological and structural characteristics of the triterpene cyclase family provides a better understanding of the relationships between important catalytic residues in the active-site cavity and the mechanisms by which cyclization occurs. It also suggests that the strategies of rational, structure-based engineering of the cyclase enzymes and its experimental results might further illustrate the detailed interactions that occur between the enzyme itself and the reactants. Thus, in the following chapters, various mutagenesis approaches on OSC were carried out. The isolation/characterization of several novel truncated cyclizations, alternative deprotonations, as well as the distinct stereochemical conformation products from these mutations further demonstrated the importance and convenience of the mutagenesis strategy for developing a better understanding of the OSC-mediated reaction.



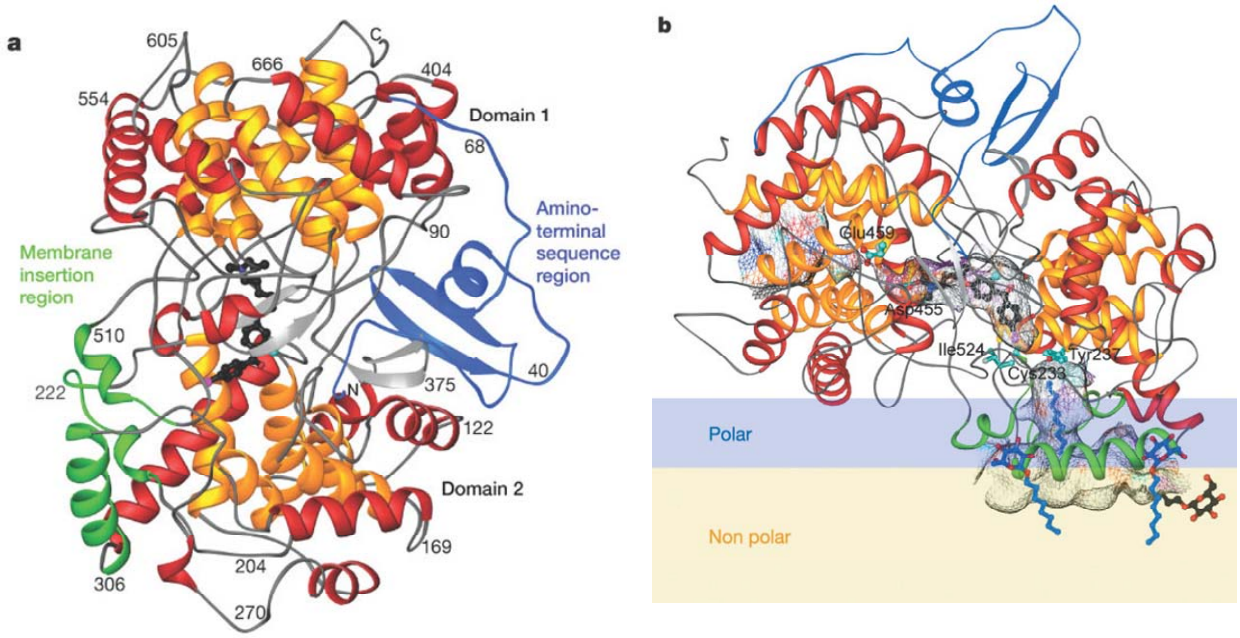


Figure 1.14. Crystal structure of human OSC; (a) Ribbon diagram of human OSC; (b) The orientation of OSC is relative to one leaflet of the membrane, and Ro48-8071 is bound in the central active-site cavity.[49]

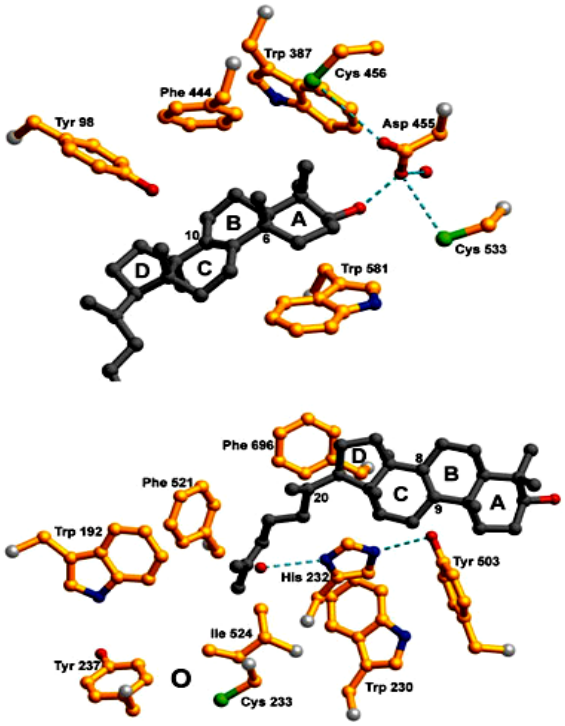


Figure 1.15. Local view of crystal structure of human OSC[49]

.

CHAPTER 2

Thesis organization

Advances in crystallographic illustration, structure-based mutagenesis, chemical inhibition and modification of substrate analogues for SHC have provided the insight to understand the relationship between the structural and functional details of the catalytic mechanism of enzyme-templated consequent ring-formation reactions. SHC appear to be primitive ancestors of OSCs, and this link such allow for the collection of extensive, plausible information about comparable OSCs. Recently, a three-dimensional structure of a related OSC from *Homo sapiens* (*H. sapiens*) was elucidated, presenting structural clues for the molecular interactions between the substrate and crucial residues in the active-site. Despite this information, various catalytic mechanisms are still being proposed.

The identification of the important residues in OSC that participate in the course of the complex cyclization and rearrangement cascade is critical for understanding the relationships in the structure-function-reaction mechanism. This information can be used to identify the residues involved in specific rearrangement and deprotonation steps, as well as those that determine substrate folding and product diversity. Herein, computer-assisted homology modeling, site-directed mutagenesis and site-saturated mutagenesis, coupled with product characterization strategies, were used to understand and solve these fascinating puzzles.

Chapter 1 summarizes a historical review as well as the biomimetic approach toward OSC, whether in chemistry, molecular biology, or enzymology, to give a fundamental framework for exploring the squalene and OSC-mediated reaction.

Chapter 3 presents the alignments of multiple sequences among various cyclases, each with a distinct substrate pre-folding conformation, to determine which residues

are critical for product control. Out of nine selected residues, His-234 was found to play the greatest role in catalysis as evidenced by many truncated, incomplete products, including monocyclic, tricyclic, and alternatively rearranged products. Many of these products were identified for the first time, representing the strong catalytic role of His-234 in the intermediate steps of oxidosqualene cyclization.

Chapter 4 describes the use of a computer-based model of *S. cerevisiae* ERG7 to determine the potential performance of the active-site cavity. Also, site-saturated mutagenesis experiments coupled with product isolation and identification were used to illustrate the specific amino-acid residues of OSC that contribute to the stabilization of carbocationic intermediates, hydride and methyl group rearrangements, and the deprotonation reaction.

Chapter 5 investigates the catalytic motif specific to the rearrangement cascade and the final deprotonation reaction in sterol-producing cyclases. Three residues, Thr-384, Glu-450, and Val-454, whose conserved patterns are different between OSC and cycloartenol synthas (CAS), were chosen for the mutagenesis studies. Mutation of these residues in OSC into the corresponding substitutions in CAS produced the most accurate parkeol synthase in the study.

Chapter 6 details the efforts involved in the molecular cloning and sequencing of a plant 2,3-oxidosqualene β-amyrin synthase from *P. sativum* (*P. sativum*). In addition, the structural comparison between *P. sativum* β-amyrin synthase and *S. cerevisiae* OSC and their site-directed mutagenesis experiments were used to identify and confirm the possible role of the properties in specific amino acids in the mutations.

Chapter 7 is a discussion on the results from the previous chapters and describes concepts that could be used in future projects.

Chapter 8 presents the detailed experimental aspects of the study of 2,3-oxidosqualene cyclase enzymes. Fundamental molecular biology techniques,

general experimental materials, and the construction of specific recombinant plasmid are also described.

The Appendix provides information on the ^{1}H-NMR, DEPT, HSQC, and HMQC NMR spectra that characterize the protosta-12,24-dien-3β-ol, protosta-20,24-dien-3β-ol, (13αH)-isomalabrica-14(26),17E,32-trien-3β-ol, 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol, and 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol.

CHAPTER 3

Site-directed mutagenesis of oxidosqualene cyclase to characterize the plasticity of the protein and the diversity of the product

3.1 Research background and aim

Nearly 200 distinct triterpenoid skeletons have been isolated from several sources.[52-54] The diverse product profiles with regio- and stereo-selectivity of oxidosqualene-catalyzed cyclization reaction might be controlled precisely by the pre-organized substrate conformation and the plastic enzyme active site. Recently, mutagenesis studies have shown that a slight modification of the enzymatic active sites is sufficient to lead to different cyclization/rearrangement reaction products.[55,56] For example, our group conducted random mutagenesis and *in vivo* selection experiments on *Arabidopsis thaliana* (*A. thaliana*) CAS to identify several critical residues that are responsible for the specific deprotonation at the C-8/C-9 position or for the formation of a strained cyclopropyl ring.[57] Five product specificity-determining point mutations of *A. thaliana* CAS, including Tyr410Cys, Ala469Val, His477Tyr, Ile481Thr, and Tyr532His, produced lanosterol to complement the *erg7*-deficient strain (Figure 3.1).[51] Subsequently, various other functional residues have been isolated independently by using similar strategies coupled with computer modeling illustration.[58,59]

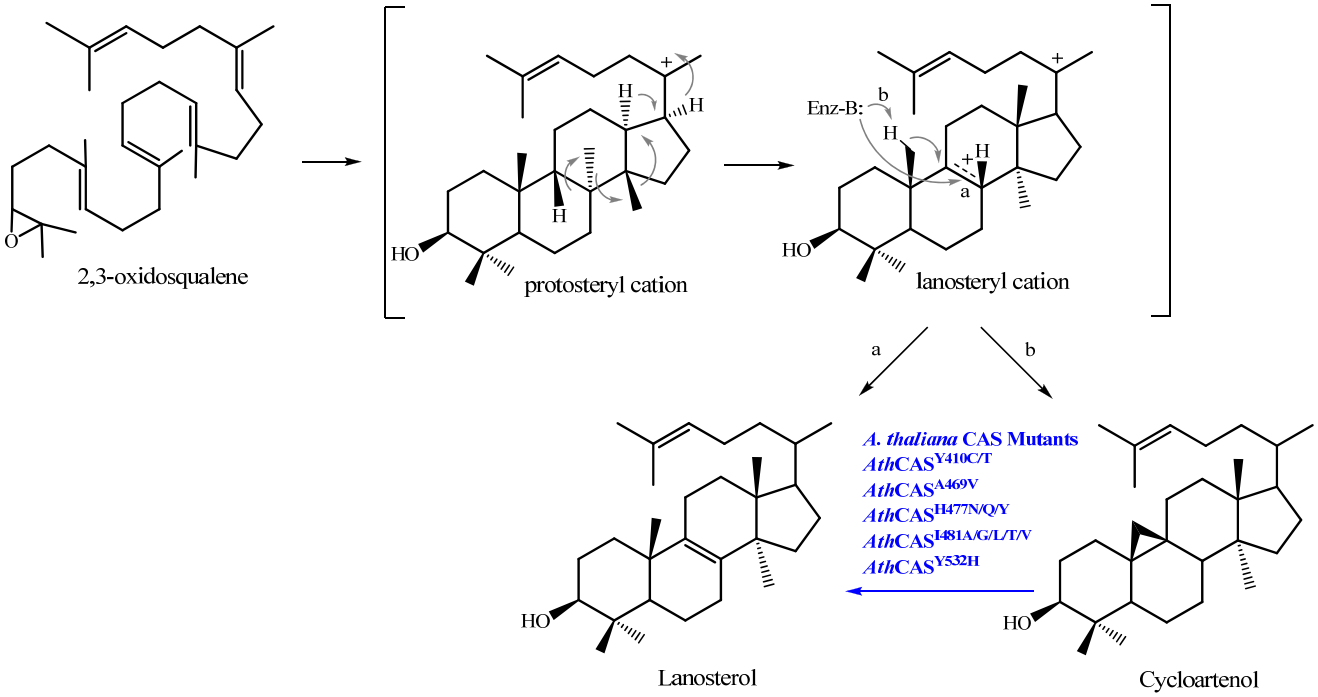


Figure 3.1. The *A. thaliana* CAS mutants changed their product specificity from cycloartenol to lanosterol

Recently, CAS was redesigned to be a lanosterol synthase with 99% accuracy by a $CAS^{H477N/I481V}$ double mutant.[60] In parallel, site-directed mutagenesis was also used to define specific residues responsible for the product specificity of the β-amyrin synthase and lupeol synthase.[61,62] Through a domain-swapping experiment, the chimeric enzyme between these two cyclases revealed a responsible region with about 80 amino acids for the products' specificity.[61] Furthermore, Trp-259 of *Panax ginseng* (*P. ginseng*) β-amyrin synthase was mutated into the corresponding residue, Leu-256, of lupeol synthase. This mutation altered the enzymatic specificity dramatically and produced lupeol as its dominant product along with other minor compounds (less than 10%). Interestingly, the corresponding mutation of Leu256Trp in *Olea europa* (*O. europa*) lupeol synthase became a functional β–amyrin synthase to further elucidate the important role of this residue in determining product specificity (Figure 3.2).[62]

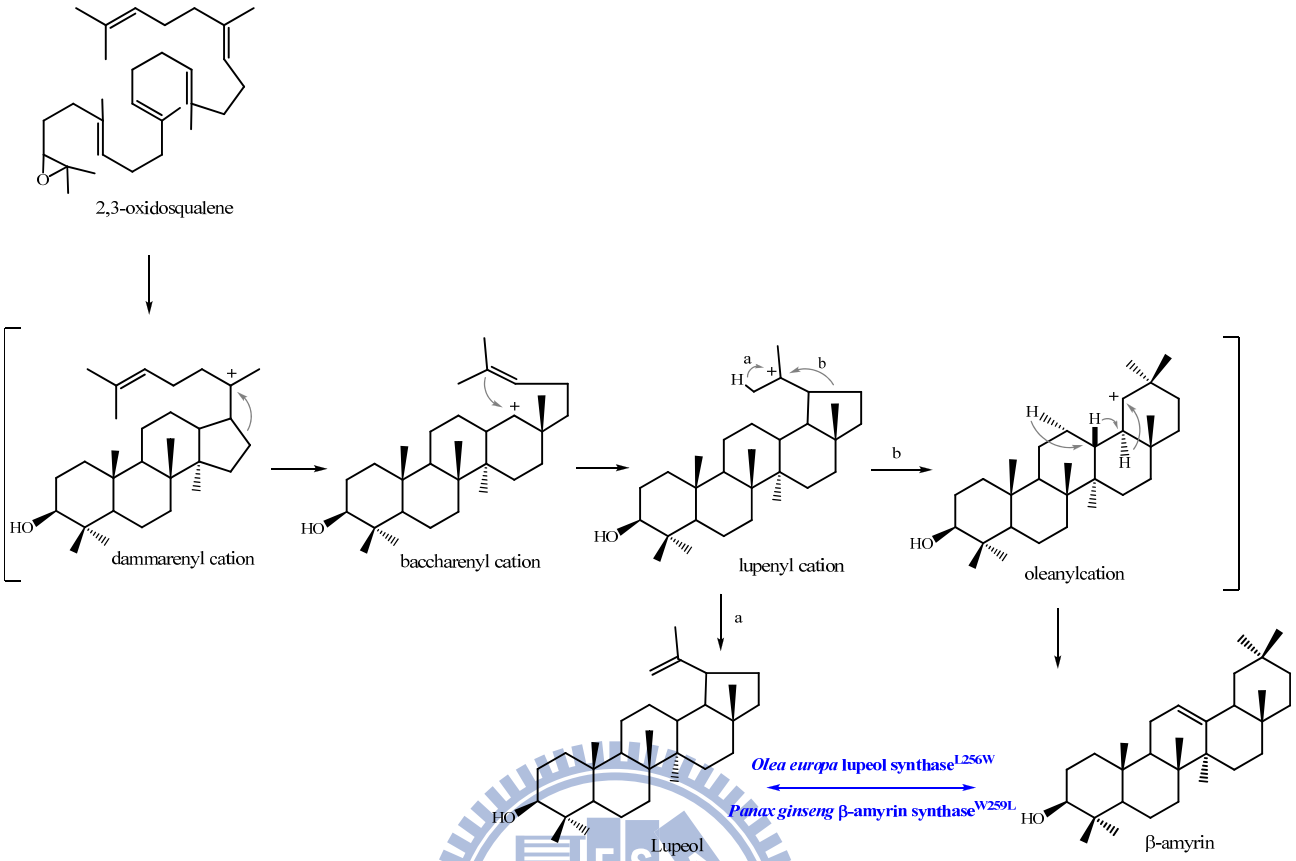


Figure 3.2. Single amino acid substitution converted the product specificity between lupeol synthase and β-amyrin synthase

The geometry of the enzymatic active site might adapt the substrate into a product-like, pre-folded conformation prior to the initiation of the reaction. However, there has been no study that focused on the relationship between substrate conformation and product specificity. How some ingenious performance of the enzyme forces the substrate into an energetically-unfavored boat form or the preferable chair form is very critical for cyclase to determine the following product formation in either the steroid or the triterpenoid biosynthesis pathway. Therefore, the alignment of the amino acid sequences among these cyclases with an opposite substrate conformation might provide valuable information for explaining the catalytic discrepancy among oxidosqualene-catalyzed reactions.[51] Site-directed mutagenesis, genetic selection, and the product-characterization approach, were

conducted to examine those artificial cyclases with novel catalytic activity or novel product specificity/diversity (Figure 3.3).

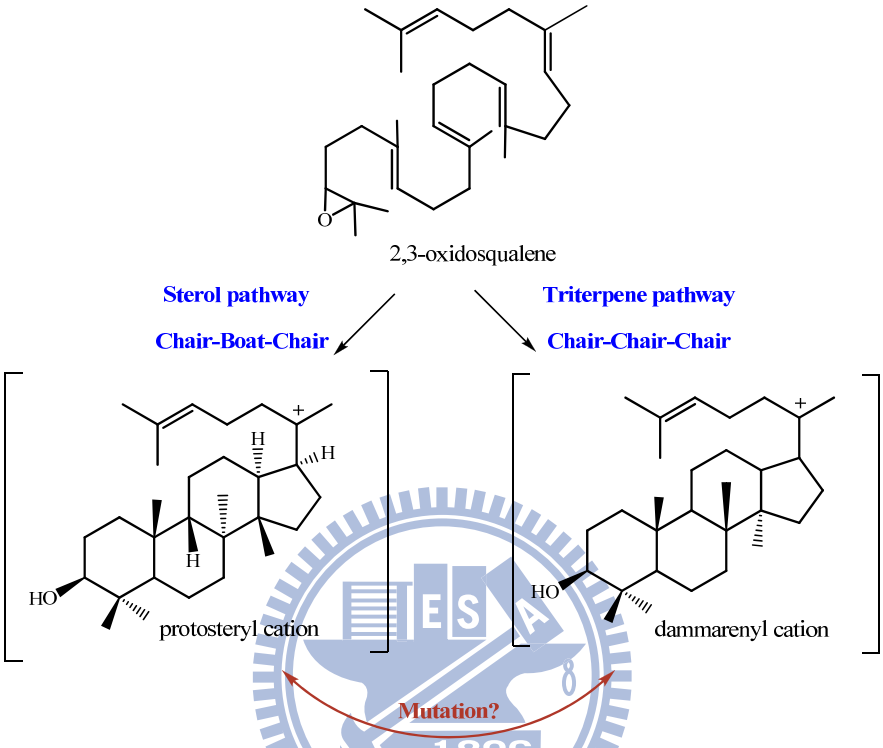


Figure 3.3. The mutagenesis approach might provide the opportunity to understand how the cyclase enzyme forces the substrate into two different conformations, which is critical for determining the following product formation in either the sterol or the triterpene biosynthesis pathway.

3.2 Results and Discussion

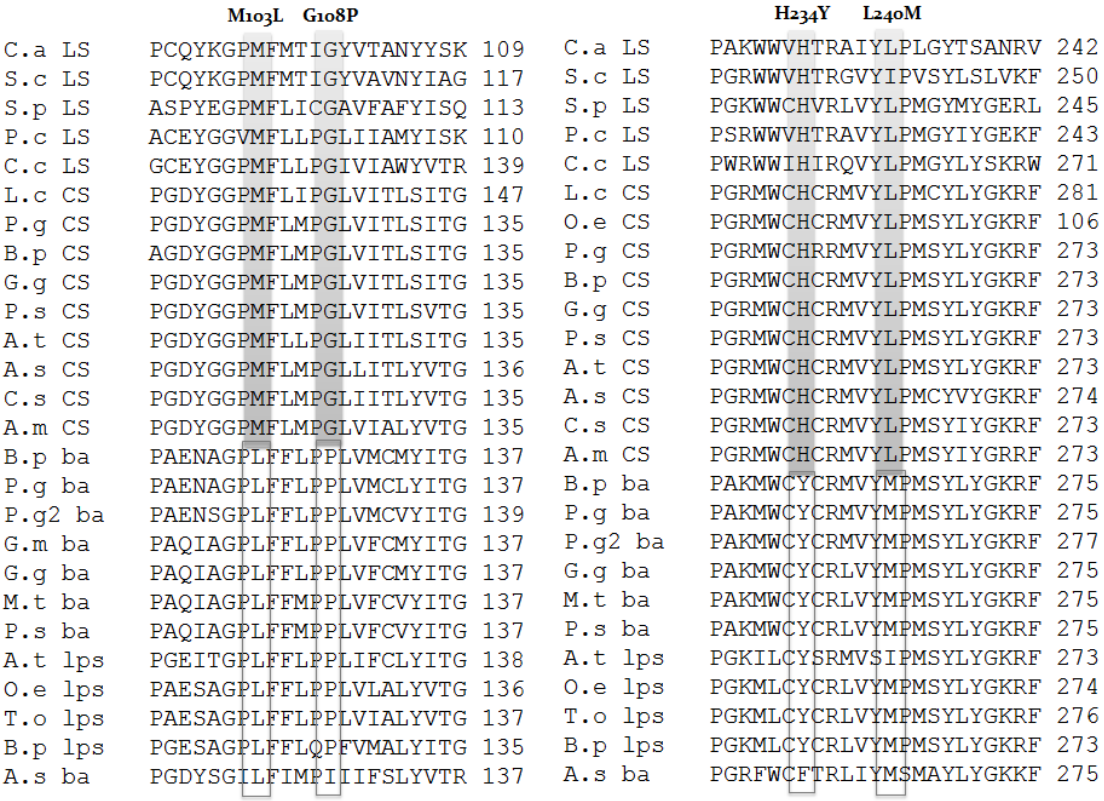
3.2-1 Alignments of multiple amino acid sequences for sterol cyclases and triterpene cyclases

Despite the low sequence identity (from 16% to 21%) in these two types of cyclases, which have the opposite conformations of the substrate, the triterpene cyclase families in prokaryotic cell and sterol cyclase in eukaryotic cell show similar structural topology.[49,63] It is considered that the conserved amino acid residues may be

responsible for the catalytic correlation and also for maintaining the secondary structure. The exclusively conserved amino-acid residues in either sterol OSCs or in triterpene OSCs might be required for the formation of the “boat” or “chair” substrate conformation, which further leads to a protosteryl cation or a dammarenyl cation, respectively. Therefore, the comparison of multiple amino acid sequences among all known cyclases was first utilized to elucidate the putative functional amino-acid residues that are critically involved in influencing the B-ring conformation (Figure 3.4).

3.1-2 Nine site-directed mutants of ERG7 gene from *S. cerevisiae*

The site-directed mutagenesis approach was first used to substitute the selected residues on ERG7 with various suitable amino acids. Nine conserved residues, i.e., Met-103, Gly-108, His-234, Ile-240, Thr-333, Pro-340, Gly-444, Ala-525, and Ile-705, from the ERG7 gene of *S. cerevisiae* were mutated into the corresponding residues that are conserved in the triterpene OSCs with a substrate that has an opposite, pre-folded chair-chair-chair conformation. These mutants, including Met103Leu, Gly108Pro, His234Tyr, Ile240Met, Thr333Ser, Pro340Cys, Gly444Thr, Ala525Thr, and Ile705Leu, were constructed by using the QuikChange™ Site-Directed Mutagenesis System or the overlapping extension PCR methods with the respective mutagenic primer pairs. A silent mutation was introduced simultaneously for screening the desired mutants according to a restrictive endonucleases analysis. The presence of the mutations was verified further by using the sequencing determination (Table 3.1).


M103L G108P
C.a LS PCQYKGPMFMTIGYVTANYYSK 109
S.c LS PCQYKGPMFMTIGYVAVNYIAG 117
S.p LS ASPYEGPMFLICGAVFAFYISQ 113
P.c LS ACEYGGVMFLLPGLIIAMYISK 110
C.c LS GCEYGGPMFLLPGIVIAWYVTR 139
L.c CS PGDYGGPMFLIPGLVITLSITG 147
P.g CS PGDYGGPMFLMPGLVITLSITG 135
B.p CS AGDYGGPMFLMPGLVITLSITG 135
G.g CS PGDYGGPMFLMPGLVITLSITG 135
P.s CS PGDYGGPMFLMPGLVITLSVTG 135
A.t CS PGDYGGPMFLLPGLIITLSITG 135
A.s CS PGDYGGPMFLMPGLLITLYVTG 136
C.s CS PGDYGGPMFLMPGLIITLYVTG 135
A.m CS PGDYGGPMFLMPGLVIALYVTG 135
B.p ba PAENAGPLFFLPPLVMCMYITG 137
P.g ba PAENAGPLFFLPPLVMCLYITG 137
P.g2 ba PAENSGPLFFLPPLVMCVYITG 139
G.m ba PAQIAGPLFFLPPLVFCMYITG 137
G.g ba PAQIAGPLFFLPPLVFCMYITG 137
M.t ba PAQIAGPLFFMPPLVFCVYITG 137
P.s ba PAQIAGPLFFMPPLVFCVYITG 137
A.t lps PGEITGPLFFLPPLIFCLYITG 138
O.e lps PAESAGPLFFLPPLVLALYVTG 136
T.o lps PAESAGPLFFLPPLVIALYVTG 137
B.p lps PGESAGPLFFLQPFVMALYITG 135
A.s ba PGDYSGILFIMPIIIFSLYVTR 137
H234Y L240M
C.a LS PAKWWVHTRAIYLPLGYTSANRV 242
S.c LS PGRWWVHTRGVYIPVSYLSLVKF 250
S.p LS PGKWWCHVRLVYLPMGYMYGERL 245
P.c LS PSRWWVHTRAVYLPMGYIYGEKF 243
C.c LS PWRWWIHIRQVYLPMGYLYSKRW 271
L.c CS PGRMWCHCRMVYLPMCYLYGKRF 281
O.e CS PGRMWCHCRMVYLPMSYLYGKRF 106
P.g CS PGRMWCHRRMVYLPMSYLYGKRF 273
B.p CS PGRMWCHCRMVYLPMSYLYGKRF 273
G.g CS PGRMWCHCRMVYLPMSYLYGKRF 273
P.s CS PGRMWCHCRMVYLPMSYLYGKRF 273
A.t CS PGRMWCHCRMVYLPMSYLYGKRF 273
A.s CS PGRMWCHCRMVYLPMCYVYGKRF 274
C.s CS PGRMWCHCRMVYLPMSYIYGKRF 273
A.m CS PGRMWCHCRMVYLPMSYIYGRRF 273
B.p ba PAKMWCYCRMVYMPMSYLYGKRF 275
P.g ba PAKMWCYCRMVYMPMSYLYGKRF 275
P.g2 ba PAKMWCYCRMVYMPMSYLYGKRF 277
G.g ba PAKMWCYCRLVYMPMSYLYGKRF 275
M.t ba PAKMWCYCRLVYMPMSYLYGKRF 275
P.s ba PAKMWCYCRLVYMPMSYLYGKRF 275
A.t lps PGKILCYSRMVSIPMSYLYGKRF 273
O.e lps PGKMLCYCRLVYMPMSYLYGKRF 274
T.o lps PGKMLCYCRLVYMPMSYLYGKRF 276
B.p lps PGKMLCYCRLVYMPMSYLYGKRF 273
A.s ba PGRFWCFTRLIYMSMAYLYGKKF 275
T333S P340C
C.a LS IVKEYQNTEYLCIAPVSFAFNMVV 343
S.c LS IKTELQNTDSLCIAPVNQAFCALV 349
S.p LS IKIEDQNTDYSCIGPVNAAMNTVC 344
P.c LS IEYENKNTDFLCIGPVNFSIHILA 342
C.c LS IDMEDKNTDYSDLAPVNAPMNTVA 371
L.c CS IHYEDENTRYICIGPVNKVLNMLC 384
O.e CS IHYEDENTRYICIGPVNKVLNMLC 209
P.g CS IHYEDENTRYICIGPVNKVLNMLC 376
B.p CS IHYEDENTRYICIGPVNKVLNMLC 376
G.g CS IHYEDDNTRYLCIGPVNKVLNMLC 376
P.s CS IHYEDENTRYICIGPVNKVLNMLC 376
A.t CS IHYEDENTRYICIGPVNKVLNMLC 376
A.s CS VHYEDENTRYICIGPVNKVLNMLT 377
C.s CS VHYEDENTRYICIGPVNKVLNMLC 376
A.m CS IHYEDENTRYICIGPVNKVLNMLC 376
B.p ba IHYEDENSRYITIGGVEKVLCMLA 379
P.g ba IHYEDENSRYITIGCVEKVLCMLA 378
P.g2 ba IHYEDENSRYITIGCVEKVLCMLV 380
G.m ba IHYEDETSRYITIGCVEKVLCMLA 356
G.g ba IHYEDETSRYITIGCVEKVLCMLA 381
M.t ba IHYEDENSRYLTIGCVEKVLCMLA 381
P.s ba IHYEDENSRYLTIGCVEKVLCMLA 381
A.t lps IHYEDENSHYITIGCVEKVLCMLA 379
O.e lps VHYEDMNSRYLCIGCVEKVLCLIA 379
T.o lps VHYEDKSSRYLCIGCVEKVLCLIA 381
B.p lps VHYEDENSKYLCIGSVEKVLCLIA 378
A.s ba IHYDDESTKYVGICPINKALNMIC 380
G444T
C.a LS FRDRRKGAWPFSTKEQGYTVSDCTAE 444
S.c LS YRDKRKGAWGFSTKTQGYTVADCTAE 460
S.p LS YRYNSLGAWPFSNITQGYTVSDTTSE 455
P.c LS YRHRRKGAWPFSTRQQGYTVSDCTAE 455
C.c LS YRQTRKGGWPFSNKDQGYAVSDCTSE 484
L.c CS YRHISKGAWPFSTADHGWPISDCTAE 495
O.e CS YRHISKGAWPFSTADHGWPISDCTAE 320
P.g CS YRHVSKGAWPFSTADHGWPISDCTAE 487
B.p CS YRHISKGAWPFSTADHGWPISDCTAE 487
G.g CS YRHISKGAWPFSTADHGWPISDCTAE 487
P.s CS YRHISKGAWPFSTADHGWPISDCTAE 487
A.t CS YRHISKGAWPFSTADHGWPISDCTAE 487
A.s CS YRHISKGAWPFSTADHGWPISDCTAE 488
C.s CS YRHISKGAWPFSTADHGWPISDCTSE 487
A.m CS YRHISNGAWPFSTRDHGWPISDCSSE 487
B.p ba HRHISKGSWTFSDQDHGWQVSDCTAE 489
P.g ba HRHISKGSWTFSDQDHGWQVSDCTAE 488
P.g2 ba YRHISKGSWTFSDQDHGWQVSDCTAE 490
G.m ba HRHISKGSWTFSDQDHGWQVSDCTAE 466
G.g ba YRHISKGSWTFSDQDHGWQVSDCTAE 489
M.t ba HRHISKGSWTFSDQDHGWQVSDCTAE 489
P.s ba HRHISKGSWTFSDQDHGWQVSDCTAE 489
A.t lps YRHISKGAWTFSDRDHGWQVSDCTAE 487
O.e lps YRHTSKGAWTFSMQDHGWQVSDCTAE 487
T.o lps YRHTSKGAWTFSIQDHGWQVSDCTAE 489
B.p lps YRHINKGAWTFSMQDHGWQVSDCTAE 486
A.s ba YRHRSKGSWTLSSVDNGWSVSDCTAE 488

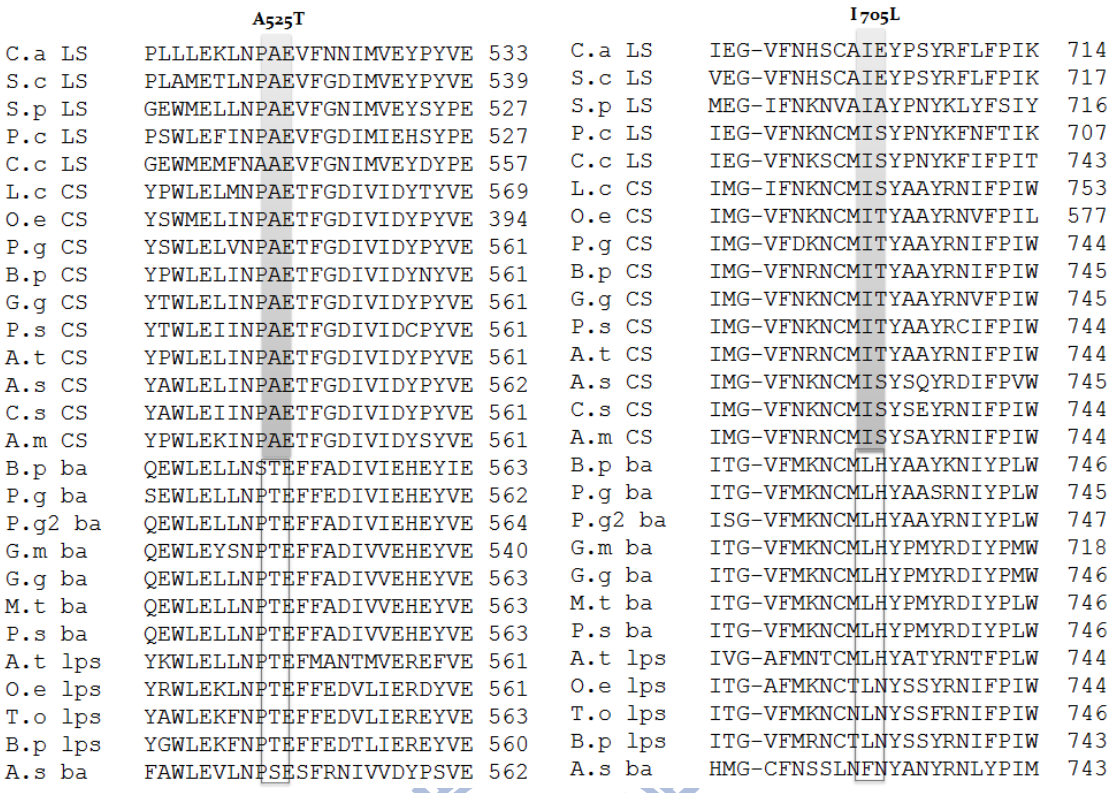


Figure 3.4. Alignment of multiple amino acid sequences for screening the putative residues involved in the product specificity/diversity. The sequences shown are *Candida albicans* OSC (C.a LS), *S. cerevisiae* OSC (S.c LS), *Schizosaccharomyces pombe* OSC (S.p LS), *Pneumocystis carinii* OSC (P.c LS), *Cephalosporium caerulens* OSC (C.c LS), *Luffa cylindrica* CAS (L.c CS), *Olea europaea* CAS (O.e CS), *P. ginseng* CAS (P.g CS), *Betula platyphylla* CAS (B.p CS), *Glycyrrhiza glabra* CAS (G.g CS), *P. sativum* CAS (P.s CS), *A. thaliana* CAS (A.t CS), *Avena strig*osa CAS (A.s CS), *Costus speciosus* (C.s CS), *Abies magnifica* CAS (A.m CS), *Betula platyphylla* βAS (B.p ba), *Panax ginseng* βAS (P.g ba), *Panax ginseng-2* βAS (P.g2 ba), *Glycine max* βAS (G.m ba), *Glycyrrhiza glabra* βAS (G.g ba), *Medicago truncatula* βAS (M.t ba), *P. sativum* βAS(P.s ba), *A. thaliana* LUS (A.t lps), *Olea europaea* LUS (O.e lps), *Taraxacum officinale* LUS (T.o lps), *Betula platyphylla* LUS (B.p lps), and *Avena strigosa* βAS (A.s ba), respectively.

Table 3.1. Construction strategies and individual silent mutation sites of the nine conserved residues

Constructs	Amino acid location	Mutation	Translation	Mutation strategies	Mapping site For silent mutation
pRSYTLOSCM103L : T1	103	ATG → CTG	M → L	QuikChange PCR	*Pvu l*
pRSYTLOSCG108P : T2	108	GGT → CCT	G → P	QuikChange PCR	*Apa l*
pRSYTLOSCH234Y : T3	234	CAT → TAT	H → Y	Overlapping extension PCR	*Xho l*
pRSYTLOSCI240M : T4	240	ATT → ATG	I → M	QuikChange PCR	*Xho l*
pRSYTLOSCT333S : T5	333	ACT → AGT	T → S	QuikChange PCR	*Pst l*
pRSYTLOSCP340C : T6	340	CCT → TGT	P → C	QuikChange PCR	*Hind lll*
pRSYTLOSCG444T : T7	444	GGC → ACC	G → T	QuikChange PCR	*Ban l*
pRSYTLOSCA525T : T9	525	GCT → ACT	A → T	QuikChange PCR	*Nco l*
pRSYTLOSCI705L : T11	705	ATT → CTT	I → L	QuikChange PCR	*Pvu ll*

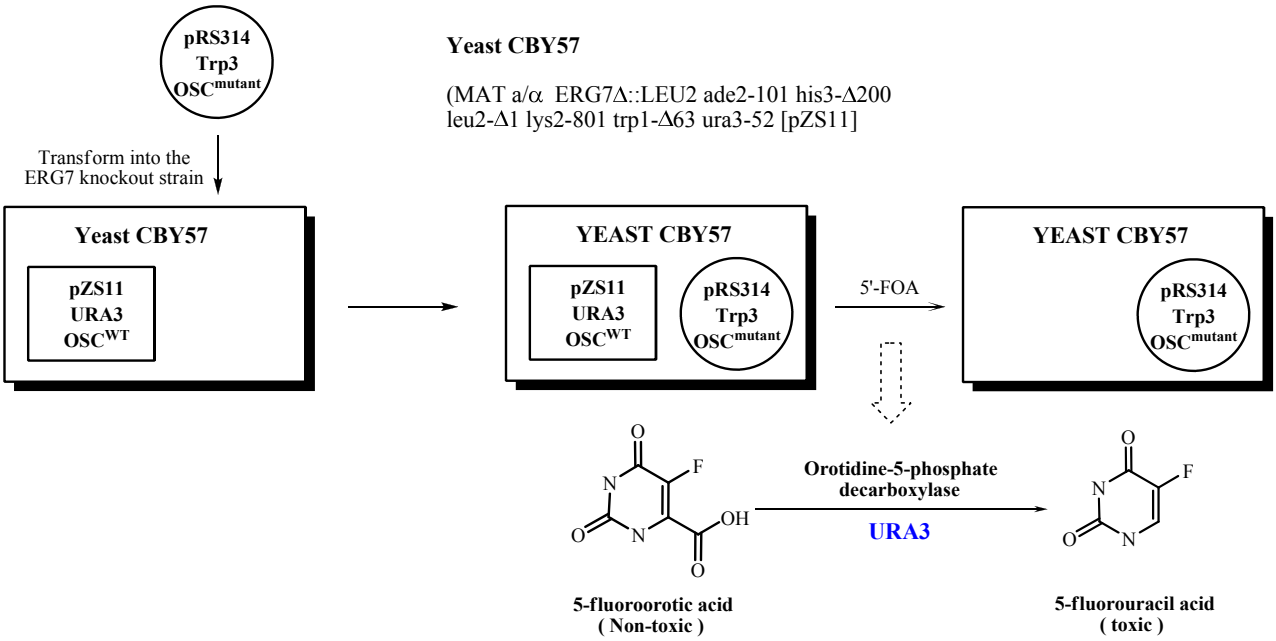
3-2.3 Principle of the plasmid shuffle method

The plasmid shuffle method, which involves the exchange of a plasmid bearing the wild-type gene sequence with a plasmid bearing the mutation of interest, is an effective method for analyzing the effect of mutants by their complementation ability to the *erg7*-deficient strain.[64]

Biosynthetic ergosterol, which is required for yeast viability, is generated through conversion of 2,3-oxidosqualene into lanosterol by OSC and another 18 additional enzymatic steps. Thus, *erg7* is regarded as an essential gene. In order to demonstrate the functional role of ERG7 mutants, the plasmid shuffle method is used. First, the *erg7* gene of a haploid yeast cell was disrupted by means of one-step gene disruption with a selectable genetic marker. Viability was maintained by transforming a plasmid bearing the wild-type gene, pZS11. Any random loss of the wild-type gene plasmid from the yeast cell would be lethal to the viability of the cell. The wild-type gene

plasmid also carries a second selectable/counter-selectable marker, URA3. This commonly used marker, the *ura3* gene, encodes orotidine-5'-phosphate decarboxylase, which converts the 5'-fluoroorotic acid (5'-FOA) into a toxic compound, 5'-fluorouracil. Thus, yeast that lacks URA3 activity is resistant to 5'-FOA. Another pRS314 plasmid bearing a gene with a mutated sequence of interest and a third selectable genetic marker is transformed into a yeast cell containing a disrupted *erg7* gene (*erg7*$^-$) and the wild-type gene plasmid (pZS11). Viability is the result of the presence of the first wild-type gene plasmid. Placing a supplement of the exogenous uracil into the medium, which is the end product of the URA3-involved biosynthetic pathway, could facilitate the loss of the first wild-type gene plasmid with frequencies up to 1% per generation.

Counter-selection against the wild-type gene plasmid with 5'-FOA allows the yeast cell to retain only the second transformed mutated gene plasmid. Therefore, the survival or the death of the yeast transformant depends on the functional activity of the mutated enzyme. If the mutant abolishes the enzymatic activity, the yeast transformant is expected to die upon loss of the first plasmids. Conversely, if the functional activity is not dramatically influenced by the mutation, the yeast transformants will still remain viable. The selectable/counter-selectable plasmid shuffle strategy provides a powerful tool for functional analysis of the mutated gene of interest (Scheme 3.1).[64]



Scheme 3.1. The plasmid shuffle strategy. Two step methods are used to exchange the wild type plasmid with the mutant plasmid by 5'-FOA, which is converted into 5'-flurouracil by the URA3 enzyme.

3.2-4 Screening inactive ERG7 mutants via the plasmid shuffle method

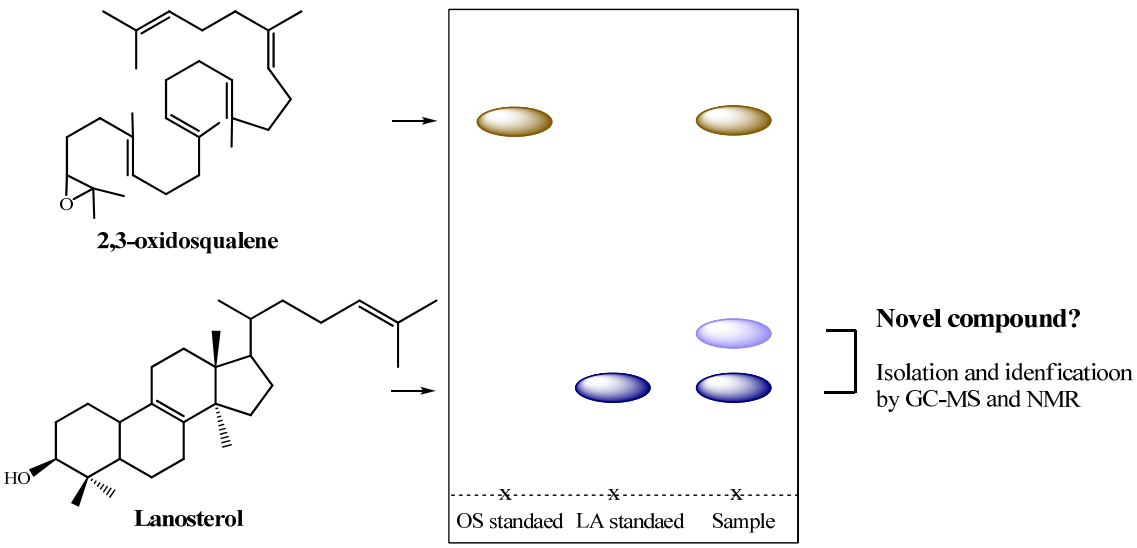
The genome type of a haploid yeast strain, CBY57, is MATa/α *ERG7Δ::LEU2 ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-52*. A CBY57[pZS11] with an *erg7*-disrupted genome and a pZS11 plasmid, which is a URA3 centromeric plasmid with a wild-type *erg7* gene, was used as the assay strain. The pRS314-derived ERG7-mutated plasmids with a selectable genetic marker of TRP1 were transformed into the yeast strain CBY57[pZS11] by using electroporation. The plasmid shuffle strategy, illustrated above, was used to analyze the mutation effect. Positive (pRS314ERG7WT) and negative (pRS314) controls were also carried out to confirm the accuracy of the plasmid shuffle. First, yeast transformants were selected on SD+Ade+Lys+His plates, and then, re-selected on SD+Ade+Lys+His+Ura+5'-fluoro-orotic acid media to elucidate the complementation effects. Among the results, all of the mutants grew colonies upon counter-selection for pZS11 with 5'-FOA, indicating

that Met103Leu, Gly108Pro, His234Tyr, Ile240Met, Thr333Ser, Pro340Cys, Gly444Thr, Ala525Thr, and Ile705Leu mutants are active and don't dramatically influence the enzymatic activity. This finding suggested that the nine selected residues in the *S. cerevisiae* ERG7 do not impair the activity below the basal level or affect the growth of yeast cells. Mutation might slightly influence the active site or the enzyme structure, but lanosterol was still produced to maintain the yeast viability. This finding implied that these residues might not be responsible for the determination of product specificity/diversity.

3.2-5 Lipid extraction, column chromatography, and product characterization of the yeast transformants

To deduce the possible catalytic or structural role of the selected residues, the non-saponifiable lipid (NSL) extract was prepared from 2.5 liters of cultured yeast transformants and applied to analysis of the product profiles. NSL extracts were dissolved in CH_2Cl_2, spotted on thin-layer chromatography (TLC), developed with 20% Ethyl Acetate (EA)/Hexane, and observed by using anisaldehyde staining. Furthermore, to examine the chemical structure of these triterpene products, the NSL extracts were separated roughly into several fractions by silica gel chromatography.

Different product patterns were collected independently (Scheme 3.2) by comparing to the control strain, pRS314ERG7WT. In addition, gas chromatography-mass spectrometry (GC-MS) was utilized to screen all fractions of the product pattern in the wild-type ERG7 and mutants.



Scheme 3.2. Preliminary analysis and separation of products on TLC plate

Based on the results of TLC and GC-MS, the product patterns of all mutants were similar to that of the wild-type ERG7 except for the His234Tyr mutant, demonstrating that Met-103, Gly-108, Ile-240, Thr-333, Pro-340, Gly-444, Ala-525, and Ile-705 of ERG7 are not the critical determinative amino acids that affect or determine the specificity/diversity of the enzyme.

A novel product, with a smaller polarity than lanosterol, was detected in the His234Tyr mutant by analyzing the TLC and GC-MS spectra. In addition, the GC-MS spectra revealed two additional novel compounds that had an identical molecular mass of m/z = 426; these compounds were indistinguishable from TLC chromatography in the Lanosterol fraction (LA fraction) (Figure 3.5). By characterizing the product structure, the possible role of His-234 of ERG7 involved in the cascades of cyclization/rearrangement will be elucidated in the next sections.

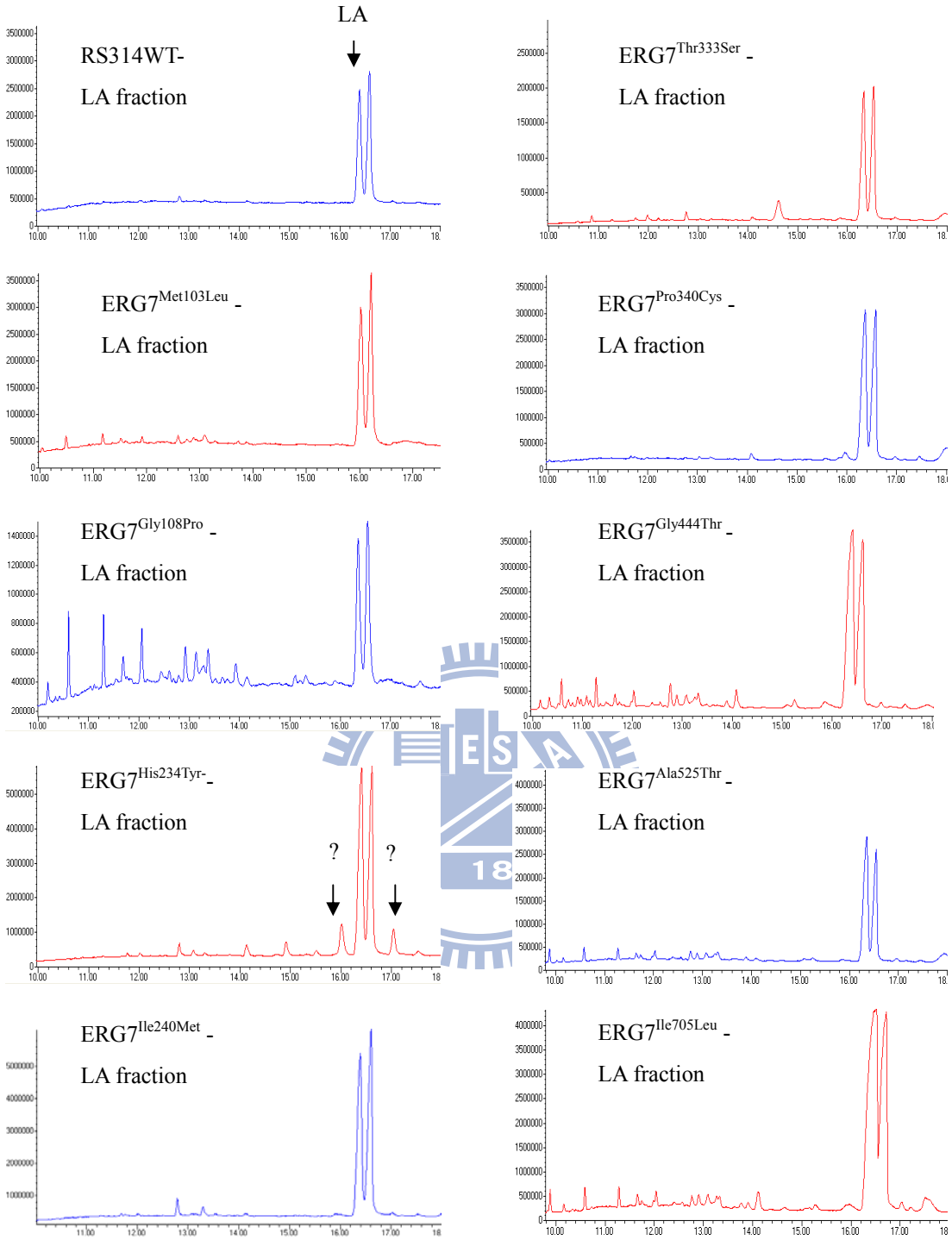


Figure 3.5. Comparison of the GC patterns of LA fraction from the NSL extracts among $ERG7^{WT}$ and nine mutants

The product of His234Tyr mutant with higher R_f value on the TLC plates was further isolated via silica gel chromatographic purification. After analyzing the GC-MS spectra and the ^{1}H-NMR spectroscopic data for this compound and

comparing the results with those of the authentic sample, this compound was identified as the monocyclic triterpene, achilleol A.[65]

3.2-6 Characterization of the mutant products in the novel gene disruption strain, TKW14c2 strain

In order to determine the functional role of inactive mutations on EER7 activity unambiguously, a new *erg7*-deficient strain, TKW14c2, derived from CBY57[pZS11], was generated.[65] Accordingly, the inherent lanosterol production from the wild-type ERG7 gene supported the survival of the yeast cells. However, it might lead to the calculation of erroneous product profiles and, thus, lead us to define an incorrect role for the residues of interest. Hence, the elimination of the pZS11 plasmid from the yeast cells is necessary, and the viability of the yeast cells can be mediated by the support of the supplementary ergosterol. It has been reported that mutation in the heme biosynthetic pathway could allow the uptake of sterol from the media under aerobic conditions, so efforts were made to incorporate a *hem1* gene disruption into a CBY57[pZS11] strain.[66,67] After successful use of gene homologous recombination technology as well as the removal of the pZS11 from the yeast cell by the counter-selection method, genetic knockout of the *hem1* with a Geneticin resistant gene ($G418^r$) was accomplished, and a novel yeast strain was generated and named TKW14c2 (MATa/α *ERG7*Δ*::LEU2 ade2*-101 *lys2*-801 *his3*-Δ200 *leu2*-Δ1 *trp1*-Δ63 *ura3*-52 *HEM1*Δ*::KanR*).[65]

The His234Tyr-mutated plasmid was transformed into the TKW14c2 strain, and the yeast transformant was viable in the absence of ergosterol media as a result of the counter selection of the CBY57 strain, indicating that the His234Tyr mutant retained the essential activity of ERG7. To analyze the product profile, a large-scale culture of the His234Tyr mutant strain was prepared. After harvesting the cells, the NSL extracts

were analyzed by TLC and GC-MS.

The GC-MS analysis of the lanosterol-positioned product revealed three compounds with identical molecular masses (*m/z* = 426). After co-injecting with the authentic samples, the major product was identified to be lanosterol, and the second compound was parkeol.[65] In order to examine the third product, the product mixture was acetylated with acetic anhydride/pyridine in order to modify the polar moiety of triterpene alcohol and applied to an $AgNO_3$-impregnated silica gel column with 15% diethyl ether/hexane as the mobile phase for separation. The isolated third product was deacetylated and characterized using GC-MS, ^{1}H-NMR, ^{13}C-NMR, different DEPT-NMR, HMBC, HSQC, and NOE spectra. The 150 MHz ^{13}C-NMR spectrum of this compound revealed the presence of two tertiary-quaternary substituted double bonds (δ 119.78, 148.76 and 125.12, 130.99 ppm). The 600 MHz ^{1}H-NMR spectrum also showed the presence of seven methyl singlets (δ 1.662, 1.580, 1.013, 0.954, 0.943, 0.878, 0.833, s, 3H) and one methyl doublet (δ 0.905, d, J 6.76 Hz, 3H), as well as two vinylic protons (δ 5.277 and 5.076, t, 1H). The HSQC spectrum showed that the methyl protons at δ 0.954 are attached to the carbon at δ 22.65 ppm and the methyl protons at δ 0.878 are attached to the carbon at δ 19.92 ppm. For the olefinic protons, the HSQC spectrum established that the vinylic proton at δ 5.277 is attached to the carbon at δ 119.78 ppm. Furthermore, nuclear over hauser effect (NOE) spectra identified the interaction between Me-19/Me-28, Me-30/H-5, and Me-18/H-9, as well as the absence of NOE spectra between Me-18/H-17 and Me-19/Me-30. Thus, the structure was characterized as protosta-12,24-dien-3β-ol, a product with $\Delta^{12,24}$ double bonds (Figure 3.6, Table 3.2, Appendix 3.1).[68]

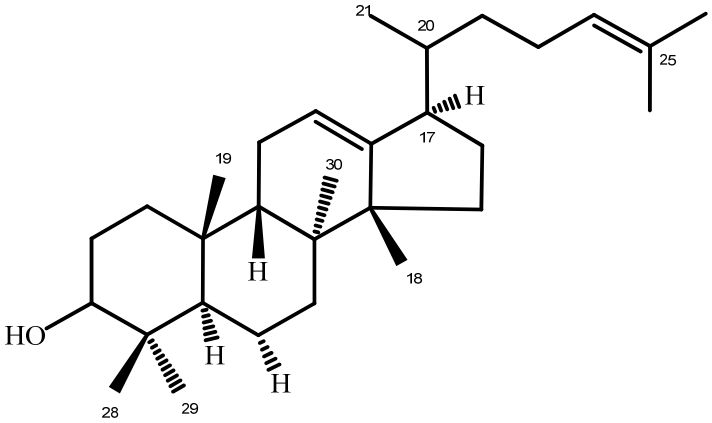


Figure 3.6. Structure characterization of protosta-12,24-dien-3β-ol

Table 3.2. ^{1}H- and ^{13}C-NMR spectra data of protosta-12,24-dien-3β-ol (150 MHz for ^{13}C-NMR and 600 MHz for ^{1}H-NMR in $CDCl_3$)

Carbon	^{13}C	Proton	^{1}H	Carbon	^{13}C	Proton	^{1}H
1-	35.04	1	1.43	16	26.99	16	1.7
2	28.44	2	1.64	17	49.43	17	2.170
3	79.33	3	3.196	18	22.65	18	0.954
4	39.61	4	0.83	19	26.29	19	1.013
5	45.19	5	1.51	20	37.56	20	1.508
6	18.17	6	1.42	21	18.27	21	0.905
7	30.21	7	1.85	22	34.11	22	1.02, 1.44
8	51.95			23	25.89	23	1.84, 2.029
9	45.42	9	1.74	24	125.15	24	5.075
10	36.08			25	130.98		
11	25.67	11	1.8-1.9	26	25.72	26	1.66
12	119.78	12	5.267	27	17.68	27	1.58
13	148.76			28	15.82	28	0.833
14	37.59			29	28.48	29	0.943
15	31.59	15	1.52, 1.18	30	19.92	30	0.878

GC-MS analysis of the NSL extract was performed, and the following products were identified in the His234Tyr mutant: Achilleol A (monocyclic), protosta-12,24-dien-3 β-ol (C-12 deprotonation), lanosterol (C-8 deprotonation), and parkeol (C-11 deprotonation). These products were found to have the percentage distribution of 14:26:51:9, respectively. The results we obtained for the His234Tyr mutant of ERG7 in *S. cerevisiae* demonstrated definitively that His-234 with an imidazole ring might be involved in stabilizing the cationic intermediate (probably during or after the formation of the protosteryl cation) and/or determining the deprotonation position. This is the first report of a truncated rearrangement product, protosta-12,24-dien-3β-ol, generated from a cyclase enzyme that adapt substrate in a pre-folded C-B-C conformation .[68]

The results of sequence alignment coupled with mutagenesis study on nine residues that were predicted to be responsible for promoting the different B-ring conformation are illustrated above. These results suggest that none of the nine residues is a critical, product-determining amino acid. Fortunately, the role of an important residue, His-234, was analyzed for the first time, and this analysis indicated that this residue is involved in the course of the complex cyclization/rearrangement cascade.

3.2-7 Site-saturated mutagenesis of His-234 to investigate its importance for ERG7 activity

In order to further elucidate the detailed role of His-234 residues, the site-saturated mutagenesis of His-234 of OSC in *S. cerevisiae* was performed.

The “site-saturated mutagenesis” approach replaces the target residue with each of other proteinogenic amino acids.[69] The different steric or electrostatic features of the other residues will facilitate the understanding of the functional role of His-234 in the catalytic activity and product profile of OSC. The His234X site-saturated mutants

were created by the QuikChange™ Site-Directed Mutagenesis System with the respective complementary mutagenic primer pairs containing degenerate base. Then, each mutant was transformed into TKW14c2 strain yeast to analyze individually the OSC activity. The results of genetic selection showed that His234Gly, His234Leu, His234Asn, His234Gln, His234Met, His234Phe, His234Asp, and His234Glu mutants were viable in the absence of ergosterol, indicating that these mutants were active and that they did not influence the function of ERG7. In contrast, His234Ala, His234Val, His234Ile, His234Pro, His234Cys, His234Ser, His234Thr, His234Trp, His234Lys, and His234Arg mutants failed to complement the ERG7 disruption. This finding suggested that the substitution of different functional groups may impair the catalytic activity of the enzyme, affect the distribution of electronic density in the active site, or influence the incomplete cyclization/rearrangement/deprotonation cascade (Table 4.3).

To elucidate the effect of different residues on His-234, the yeast transformants that contained mutated plasmids were grown in the 2.5 liters liquid media. The cells were harvested, NSL extracts were prepared, and products were analyzed by GC-MS. Among these mutations, H234Lys and H234Arg were detrimental to the essential activity of ERG7, and the product profiles were also similar to the results obtained for the negative control, TKW14c2[pRS314]. This suggests that Lys and Arg substitution of the His-234 residue may influence the correct folding of the substrate because the positively-charged side chain is unable to stabilize the high-energy cation intermediate.

Table 3.3. Complementary results of $ERG7^{H234X}$ mutants in an *erg7 hem1* double-knockout strain, TKW14c2. The symbol "+" indicates the mutant cyclase can complement the knockout strain. The symbol "-" means the inactive mutants.

Amino acid substitutions		Viability
Non polar, aliphate	Gly	+
	Ala	-
	Val	-
	Leu	+
	Ile	-
	Pro	-
Polar, uncharge	Cys	-
	Met	+
	Asn	+
	Gln	+
	Ser	-
	Thr	-
Aromatic	Tyr	-
	Phe	+
	Trp	+
Positive charge	Lys	-
	Arg	-
Negative charge	Asp	+
	Glu	+

Based on GC-MS analysis of the NSL extracts from His234X mutants with the authentic samples, various cyclized products were confirmed from these mutants, including achilleol A, lanosterol, parkeol, and protosta-12,24-dien-3β-ol. Additionally, some mutants also revealed a substantial amount of two additional products distinct from the known triterpenes. To further characterize and quantify these two compounds, the cells were cultured, and the NSL extracts were acetylated and placed in an $AgNO_3$-impregnated silica gel column for chromatographic purification and structure identification with NMR spectra (^{1}H-, ^{13}C-, different DEPT, HSQC, HMBC and 2-D NOE). The fifth compound isolated from His234Leu, His234Met, His234Asp, and His234Asn mutants was established as protosta-20,24-dien-3β-ol, a tetracyclic product with $\Delta^{20,24}$ double bonds (Figure 3.7, Table 3.4, and Appendix 3.2). The compound showed the distinct ^{1}H-NMR chemical shifts with one olefinic proton (δ 5.08) and two methylene protons (δ 4.88, 4.86), as well as seven methyl singlets (δ 1.66, 1.58, 1.07, 0.96, 0.87, 0.82, and 0.76). The 150 MHz ^{13}C-NMR spectrum revealed the presence of one secondary-quaternary substituted double bond and one tertiary-quaternary substituted double bond (δ 108.78, 152.04 and 124.36, 131.43 ppm, respectively). The latter is a characteristic of the double bond at the hydrocarbon side chain. The HSQC spectrum showed that the olefinic methylene protons at δ 4.88 and 4.86 are attached to the carbon at 108.78 ppm, while the methine proton at δ 2.60 is attached to the carbon at 43.88 ppm (C-17). In the ^{1}H-^{1}H COSY spectrum, the methine proton at δ 2.60 shows connectivity with the two olefinic methylene protons at δ 1.75 and the methane proton at δ 1.97, which are attached to carbons at 27.65 ppm (C-16) and 44.52 ppm (C-13), respectively. In the HMBC spectrum, the δ 2.60 methine proton is coupled by ^2J to carbons at 152.04 ppm (C-20), 27.65 ppm (C-16) and 44.52 ppm (C-13). The δ 2.60 methine proton is coupled by ^3J connectivity to a carbon at 108.78 ppm (C-21), as well as to 38.62 ppm (C-22) and 50.67 ppm (C-14).

The HMBC also established that the tertiary vinylic proton (δ 5.08) is coupled by ^{2}J to carbons at 27.30 ppm (C-23), and 131.43 ppm (C-25), as well as by ^{3}J connectivity to carbons at 38.62 ppm (C-22), 17.69 ppm (C-27), and 25.70 ppm (C-26). Additional HSQC and HMBC correlations showed ^{3}J connectivity between olefinic methylene protons (4.88 and 4.86) and carbons at 38.62 ppm (C-22) and 43.88 ppm (C-17). These correlations establish unambiguously the key structural features of the two double bonds located between C-20 and C-21 and between C-24 and C-25, respectively. Finally, the presence of NOE spectra between Me-19/Me-28, Me-18/H-21, Me-18/H-9, Me-19/H-9, Me-29/H-3, and H-3/H-5, as well as the absence of NOE spectra between Me-19/Me-30, Me-18/Me-30, and Me-18/H-17, confirmed the structure to be protosta-20,24-dien-3β-ol, a tetracyclic product with $\Delta^{20,24}$ double bonds and a C-17β hydrocarbon side-chain configuration. This is the first reported isolation of the product derived from the protosteryl cation without rearrangement of the C-17α hydride.[69]

On the other hand, the sixth distinct product isolated from the His234Leu, His234Met, His234Asp, and His234Asn mutants was determined to be (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol, a chair-boat (C-B) 6,6,5-tricyclic product with trans-syn-trans stereochemistry and $\Delta^{14(26),17,21}$ double bonds (Figure 3.8, Table 3.5, and Appendix 3.3). Confirmation of the structure, including the stereochemistry and the connectivity of C-8 and C-13, was obtained by analyzing the following NMR data. The ^{1}H-NMR showed signals of seven methyl singlets (δ 1.66, 1.582, 1.579, 1.09, 0.96, 0.93, 0.75), two singlets of double bond protons (δ 4.84, 4.60), and two multiplet, double-bond protons (δ 5.11-5.08). These NMR results indicated the presence of two hydrocarbon side-chain double bonds and one exocyclic double bond. The ^{13}C-NMR also showed the presence of two tertiary-quaternary substituted double bonds and one secondary-quaternary substituted double bond (δ 124.37, 131.29 and

124.15, 135.14, as well as 109.14, 155.08 ppm) that are similar to those of the tricyclic nucleus. Furthermore, the HSQC spectrum established that the double-bond protons at δ 4.84 and δ 4.60 are attached to the carbon at 109.14 ppm (C-26) and that the methine proton at δ 2.07 is attached to the carbon at 56.47 ppm (C-13). In the HMBC spectrum, the δ 2.07 methine proton is coupled by ^{2}J to carbons at 155.08 ppm (C-14), 44.79 ppm (C-8), and 28.36 ppm (C-12), as well as by ^{3}J to carbons at 109.14 ppm (C-26), 52.31 ppm (C-9), 39.22 ppm (C-15), 31.93 ppm (C-7), 29.64 ppm (C-27), and 20.74 ppm (C-11), thus establishing the connectivity of the germinal protons on the double bond adjacent to a tricyclic nucleus. Finally, the presence of NOE spectra between Me-28/Me-30, Me-27/H-13, and Me-28/H-9 and the absence of NOE spectra between Me-27/Me-28, Me-27/H-9, Me-28/H-5, and H-9/H-13 were uniquely consistent with the stereochemistry of the C-B 6,6,5-tricyclic nucleus. These findings established unambiguously the structure to be (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol, which is a tricyclic abortive cyclization product with $\Delta^{14(26),17,21}$ double bonds and trans-syn-trans stereochemistry on the 6,6,5-tricyclic nucleus. This work described the first tricyclic intermediate of the oxidosqualene cyclization/rearrangement cascade that has been derived from the direct trapping of the C-B 6,6,5-Markovnikov tricyclic cation.[69]

Product profiles of the His234X site-saturated mutations are summarized in Table 3.6. The H234X mutants generated a diverse profile of products. The His234Phe produced achilleol A, accompanied by lanosterol and parkeol; His234Asp produced (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol as part of the product profile. The His234Gly, His234Ala, and His234Cys mutants produced protosta-20,24-dien-3β-ol, in addition to other products. Interestingly, the His234Asn mutant simultaneously produced (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol, protosta-20,24-dien-3β-ol, protosta-12,24-dien-3β–ol, lanosterol, and parkeol, but it did not produce monocyclic

achilleol A. In contrast, the His234Ser mutant produced a minor amount of protosta-12,24-dien-3β-ol as its sole product. The His234Glu mutant produced lanosterol as the sole product. Finally, the His234Thr and His234Trp mutants produced minor amounts of parkeol as their sole product.

What is the altered enzyme activity from the various substitutions on His-234 of ERG7 for the yield of the different products? It appears that the enzyme first initiates the cyclization of OS from a pre-folded chair-boat-chair conformation to a monocyclic C-10 cation. Specific deprotonation at C-10 results in the production of achilleol A. Then, $ERG7^{H234X}$ mutants catalyze B-ring and C-ring cyclization continuously via a chair-boat conformation and a Markovnikov-favored 6,6,5-ring closure to form a C-14 cation. Direct abstraction of the C-26 proton generates the tricyclic (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol. C-ring expansion from the tertiary cyclopentylcarbinyl cation leads to a secondary anti-Markovnikov cyclohexyl carbocation, followed by D-ring annulation to generate the protosteryl C-20 cation, proceeded a finally rotation of about 60° to generate the natural C-20R configuration with a 17β-side chain. The hydrogen is abstracted directly at C-21 to yield the protosta-20,24-dien-3β-ol. Then, two backbone rearrangements of [1,2]-hydride shift from H17α→20α and H13α→17α to generate the protosteryl C-13 cation. The elimination of H-12 further forms protosta-12,24-dien-3β-ol with a Δ^{12} double bond. The generation of the lanosteryl C-8 cation from the C-13 cation involves two additional methyl-group shifts (Me-14β→Me-13β and Me-8α→Me-14α). Lanosterol formation is completed by a final deprotonation, either by the abstraction of the proton at C-9 or at C-8 via a hydride shift from C-9 to C-8. Parkeol was formed after another hydride shift from C-9 to C-8, coupled with the C-11 proton abstraction (Figure 3.9).

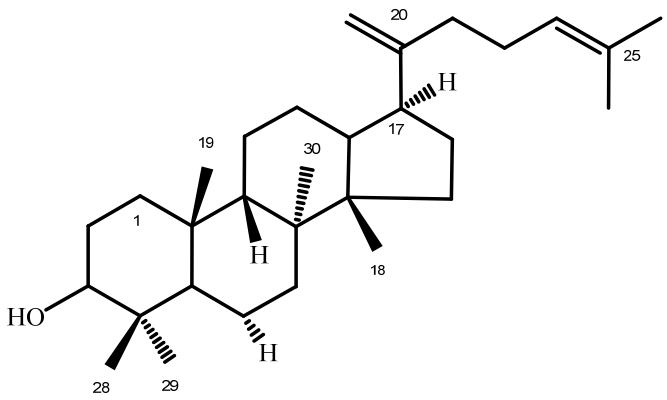


Figure 3.7. Structure characterization of protosta-20,24-dien-3β-ol

Table 3.4. ^{1}H- and ^{13}C-NMR data of protosta-20,24-dien-3β-ol (150 MHz for ^{13}C-NMR or 600 MHz for ^{1}H-NMR in $CDCl_3$)

Carbon	^{13}C	Proton	^{1}H	Carbon	^{13}C	Proton	^{1}H
1	32.88	1	1.38 (d), 1.37 (d)	16	27.65	16	1.75 (m, 2H)
2	29.11	2	1.68 (d), 1.52 (m)	17	43.88	17	2.60 (dt)
3	79.33	3	3.22 (dd)	18	17.48	18	0.82 (s,3H)
4	39.15			19	22.38	19	0.87 (s,3H)
5	47.66	5	1.43 (d)	20	152.04		
6	18.57	6	1.119-1.18	21	108.78	21	4.86 (s), 4.88 (s)
			(m), 1.49 (m)				
7	34.81	7	1.18-1.17 (m), 1.92 (m)	22	38.62	22	1.95 (m), 2.10 (m)
8	39.95			23	27.30	23	1.91 (m), 2.04 (m)
9	45.68	9	1.44 (t)	24	124.36	24	5.08 (m)
10	36.81			25	131.43		
11	23.09	11	1.20-1.19 (m), 1.46 (m)	26	25.70	26	1.66 (s, 3H)
12	25.14	12	1.16- 1.15 (m)	27	17.69	27	1.58 (s, 3H)
			1.20-1.19 (m)				
13	44.52	13	1.97 (d)	28	16.10	28	0.76 (s,3H)
14	50.67			29	29.08	29	0.96 (s,3H)
15	33.35	15	1.17-1.16 (m), 1.42 (d)	30	21.98	30	1.07 (s, 3H)

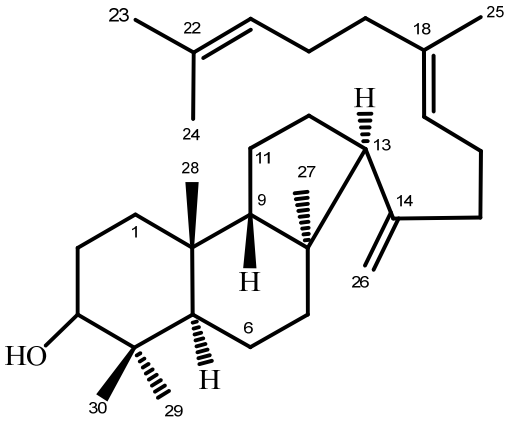


Figure 3.8. Characterization of the structure of (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol

Table 3.5. ^{1}H- and ^{13}C-NMR data of (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol (150 MHz for ^{13}C-NMR or 600 MHz for ^{1}H-NMR in $CDCl_3$)

Carbon	^{13}C	Proton	^{1}H	Carbon	^{13}C	Proton	^{1}H
1	34.27	1	1.39, 1.46 (m)	16	26.53	16	2.13 (m)
2	29.17	2	1.63-1.62 (m), 1.73-1.71 (m)	17	124.15	17	5.09 (m)
3	79.53	3	3.23 (dd)	18	135.14		
4	39.06			19	39.69	19	1.98, 1.94 (m, 2H)
5	46.70	5	1.50	20	26.72	20	2.06 (m), 2.04 (m)
6	18.55	6	1.52, 1.19 (m)	21	124.37	21	5.11 (m)
7	31.93	7	1.61 (m),1.24	22	131.29		
8	44.79			23	17.68	23	1.582 (s, 3H)
9	52.31	9	1.53	24	25.69	24	1.66 (s)
10	35.27			25	16.01	25	1.579 (s, 3H)
11	20.74	11	1.40(m), 1.53	26	109.14	26	4.60 (s), 4.84 (s)
12	28.36	12	1.55 (s), 1.96 (m)	27	29.64	27	1.09 (s,3H)
13	56.47	13	2.07 (d)	28	23.11	28	0.93 (s,3H)
14	155.08			29	29.08	29	0.96 (s,3H)
15	39.22	15	1.83-1.87 (m)	30	15.81	30	0.75 (s,3H)

Table 3.6. Product profile of *S. cerevisiae* TKW14c2 expressing the ERG7^{H234X} site-saturated mutants

Amino acid Substitutions		Product Profile						
		No product	Achilleol A	Isomalabarica-14 (26),17,21-trien-3β-ol	Protosta-12,24-dien-3β-ol	Lanosterol	Parkeol	Protosta-20,24-dien-3β-ol
	Gly				29	17	47	7
	Ala				17	30	40	13
Non polar,	Val					42	58	
aliphate	Leu			30		39	31	
	Ile					70	30	
	Pro					64	36	
	Cys				7	67	22	4
Polar,	Met		17	10		31	42	
Uncharge	Asn			26	23	27	10	14
	Gln					100		
	Ser				100			
	Thr						100	
	Tyr		14		26	51	9	
	Phe		66			14	20	
Aromatic	Trp						100	
	Lys	V						
Positive	Arg	V						
charge	His					100		
Negative	Asp			58		30	12	
charge	Glu					49	51	

The site-saturated mutation experiments clearly demonstrated that the His-234 of ERG7 is important and necessary for OSC catalysis. The intrinsic His-234:Tyr-510 hydrogen-bonding network and/or steric influence of this residue in the active-site cavity are optimal for the stabilization of the electron-deficient cationic intermediate during the native product formation. Different interactions with the reactive intermediated cations or the neighboring amino acid residues were expected and led to the production of diverse, unnatural product profiles among the various His-234 mutants. The substitution of different amino acids in the His-234 position of ERG7 resulted in the production of achilleol A, (13α*H*)-isomalabaricane-14(26),17*E*,21-trien-3β-ol, protosta-20,24-dien-3β-ol, protosta-12,24-dien-3β-ol, and parkeol. The product profiles illustrated how cyclase subtly controls the ERG7^{His234} residue in the π-electron-rich pocket for interaction with spatially adjacent amino acid residues and stabilization of the high-energy cationic intermediate at the cyclization/rearrangement/deprotonation reaction stage, respectively. On the other hand, isolation of (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol and protosta-20,24-dien-3β-ol provides long-sought support for the hypothesis that the formation of lanosterol from oxidosqualene in ERG7-catalyzed cyclization/rearrangement reactions must proceed via the Markovnikov C-B 6,6,5-tricyclic cation as well as the protosteryl C-20 cation.[69]

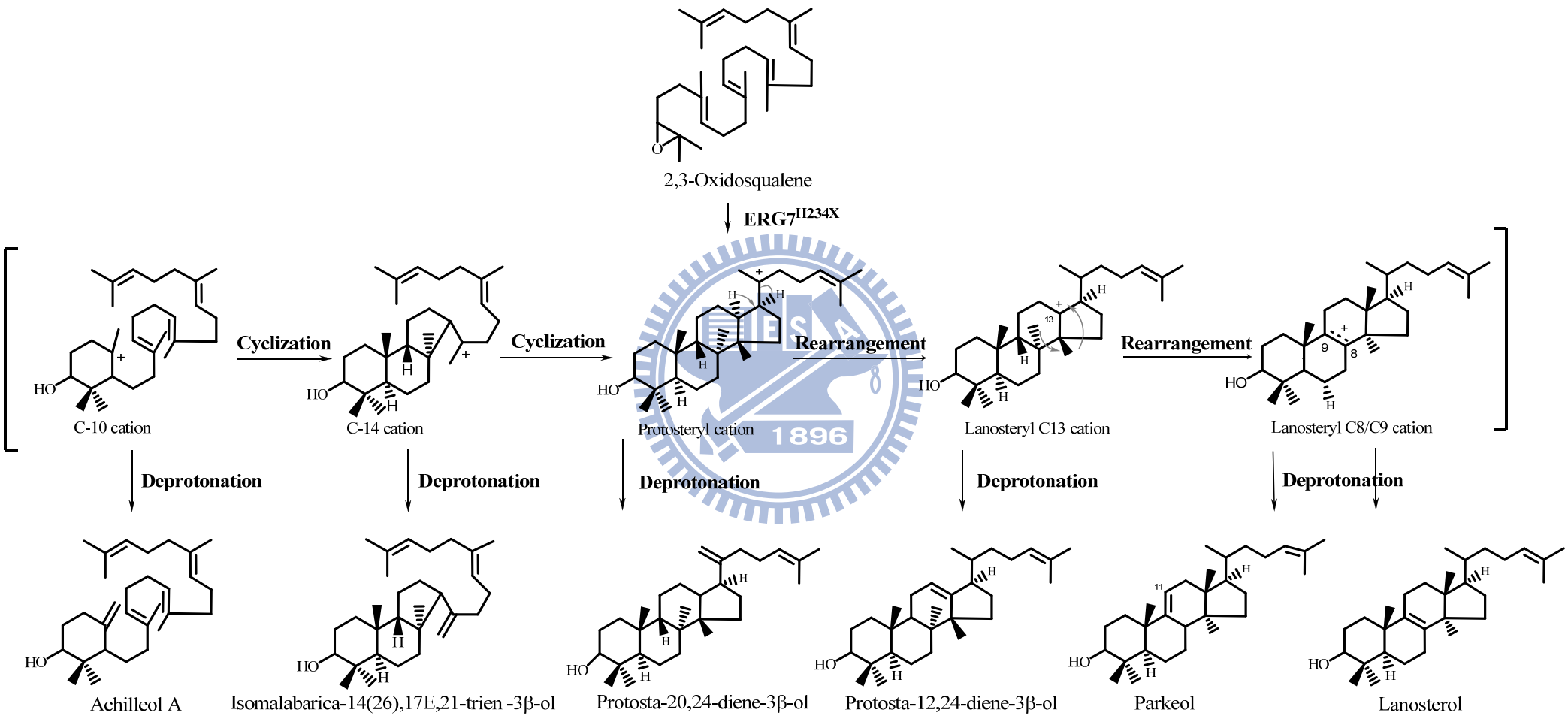


Figure 3.9. Product profile produced from TKW14c2 expressing the ERG7^{H234X} site-saturated mutations

CHAPTER 4

Site-saturated mutagenesis and product characterization to study the putative active-site residues from *S. cerevisiae* oxidosqualene cyclase

4.1 Research background and aim

The crystal structure of bacterial SHC in complex with a substrate mimic, 2-azasqualene, and human OSC in complex with the reaction product, lanosterol, provided the benefit of additional snapshots for understanding the cyclase enzyme-mediated cyclization cascade.[33,49] Several amino-acid residues, which are responsible for each of the steps, were clearly defined by the structural insight. Although these residues seem to result in useful models for explaining the cyclization steps, some of fundamental questions in the complex oxidosqualene cyclization still remained unsolved, including (i) how OSCs strictly trigger the substrate, oxidosqualene, to adopt the pre-folded conformation in the enzymatic active-site cavity for the different cyclization pathway; (ii) why the A/B ring formation is concerted or nonconcerted after the initiation reaction; (iii) which functional residues located in the proper position of the enzymatic active site are responsible for rigidly holding and stabilizing the cyclized carbocation to prevent early deprotonated truncation or nonspecifically nucleophilic water termination through cation-π interaction or the steric effect; (iv) validation or rejection of the argument regarding the formation of a tricyclic cyclopentyl or a carbinylcyclohexyl intermediate prior to the generation of tetracyclic protosteryl cation; (v) how the stereochemical course among various cationic intermediates, e.g. C-17α/β orientation or the C-20*R*/*S* configuration of the protosteryl cation is regulated precisely; and (vi) mutation on plastic residues for the product specificity/diversity in various species-dependent

cyclases.[51,65,70]

The availability of only the structural illustration of the OSC-mediated cyclization mechanism is not enough to clarify these anterior questions. Therefore, a series of alanine-scanning and site-saturated mutagenesis on putative site residues of *S. cerevisiae* ERG7 protein, coupled with product isolation and characterization, was further executed and described in this chapter.

4.2 Result and discussion

4.2-1 Generation of a homology model for *S. cerevisiae* OSC and functional analysis of alanine-scanning mutants of putative active-site residues

To verify the significant function of putative amino-acid residues from *S. cerevisiae* OSC, the homology structure of the *S. cerevisiae* OSC was determined by using either the crystal structure of *A. acidocaldarius* SHC (PDB entry 1UMP) or the crystal structure of *H. sapiens* OSC (PDB entry 1W6K) as the guide, respectively.[63,65] After the docking of the lanosterol molecule into the homology models by using GOLD 3.0 software, the ligand-protein complex structures were converged into the energy-minimized state by Sybyl 7.0. The superimposed *S. cerevisiae* OSC homology structure, either with the *A. acidocaldarius* SHC or *H. sapiens* OSC, showed a high similarity to the overall framework and the active-site cavity. Further, it also revealed the resemblant orientation among these putative amino-acid residues (Figure 4.1).

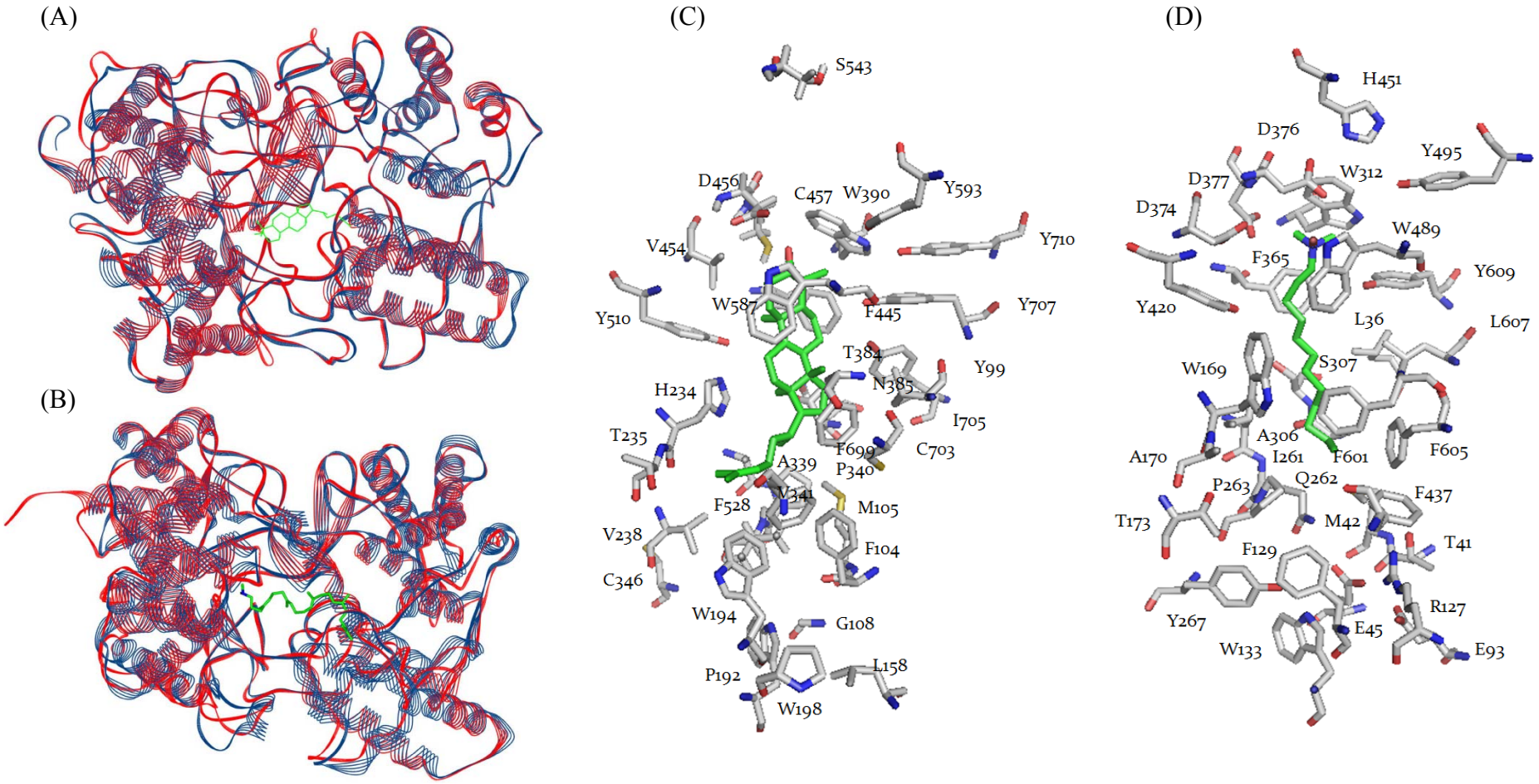


Figure 4.1. (A) Superimposition of the *S. cerevisiae* ERG7 homology modeling structure (red) with crystal structure of human OSC (blue), and the docked product, lanosterol (green), is included (B) Superimposition of the *S. cerevisiae* ERG7 homology modeling structure (red) with crystal structure of *A. acidocaldarius* SHC (blue), and the inhibitor, 2-azasqualene, is shown in green (C) Local views of the homology modeled *S. cerevisiae* ERG7 structure based on the crystal structure of human OSC. Lanosterol (green) is also included (D) Local views of the crystal structure of *A. acidocaldarius* SHC. 2-azasqualene (green) is also included.

Detailed examination of the modeling structure suggested that a π-electron-rich pocket was created to stabilize and fix the highly energetic carbocation intermediate through the previously-illustrated catalytically important aromatic amino-acid residues.[34,71] Therefore, these composing residues from the active-site appearance were selected for the mutagenesis to further explain how each residue contributed to a particular reaction mechanism.

Alanine-scanning mutagenesis is a widely used technique to determine the catalytic or functional role of protein residues.[72] The residue of interest was substituted to the alanine, which eliminated the side chain beyond the β-carbon but did not alter the main-chain conformation or impose excessive electrostatic or steric effects. Then, alanine-scanning mutagenesis on active-site residues showed that the yeast transformants that contained Tyr99Ala, Trp232Ala, His234Ala, Trp390Ala, Trp443Ala, Phe445Ala, Trp448Ala, Tyr510Ala, Phe699Ala, Tyr707Ala, and Ile705Ala mutants failed to grow on the ergosterol-free medium. These findings indicated that homology modeling of the structure of yeast ERG7 provides a reliable way of defining these putative active-site amino-acid residues.

4.2-2 Site-saturated mutagenesis and functional analysis of active-site residues from *S. cerevisiae* ERG7

After alanine-scanning mutagenesis on putative amino-acid residues, the catalytic residues that participated in the cyclase-mediated reaction were elucidated. To further explore the steric or electrostatic effects of the residues that were presumed to be involved in the complicated cyclization/rearrangement cascade, the site-saturated mutagenesis approach, coupled with the characterization of the bioorganic compounds, was systematically established on several functional residues, including Tyr-99, Trp-232, His-234, Trp-390, Trp-443, Phe-445, Lys-448, Tyr-510, Phe-699, Ile-705,

and Tyr-707 from *S. cerevisiae* OSC (Figure 4.2). The product profiles distributed from each mutant were isolated and identified. Different monocyclic, bicyclic, tricyclic, stereochemically distinct C-C-C tricyclic, tetracyclic, truncated rearranged tetracyclic, tetracyclic with a 17α stereochemical exocyclic side chain, and alternatively deprotonated tetracyclic scaffold products were characterized among the ERG7 mutants.[51,65,68,69,73-82] These acquired abortive products provided direct evidence that demonstrated the authenticity of long-deduced mechanistic intermediates during the process in which oxidosqualene is cyclized to form lanosterol. In addition, they also showed that the plasticity of the cyclase enzyme family could be obtained via only minor modification of the enzymatic active-site cavity to evolve numerous triterpene skeletons to allow various organisms to adapt to the environmental changes. The abortive products that were isolated from different mutations are listed in Figure 4. 3.

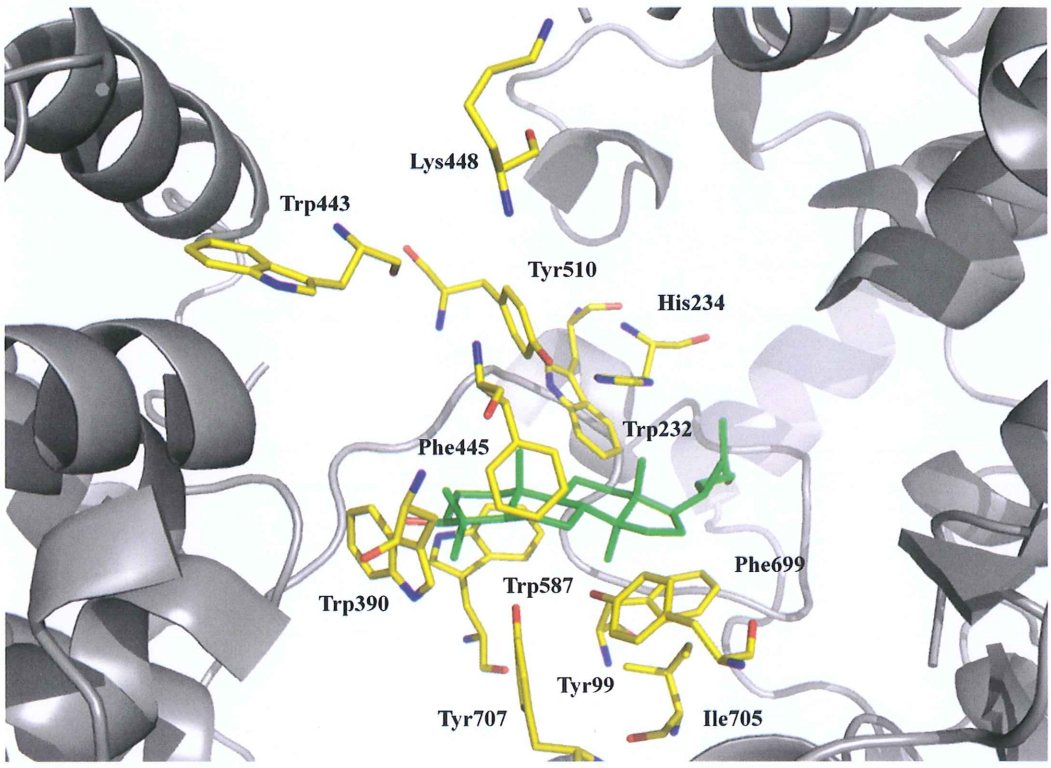
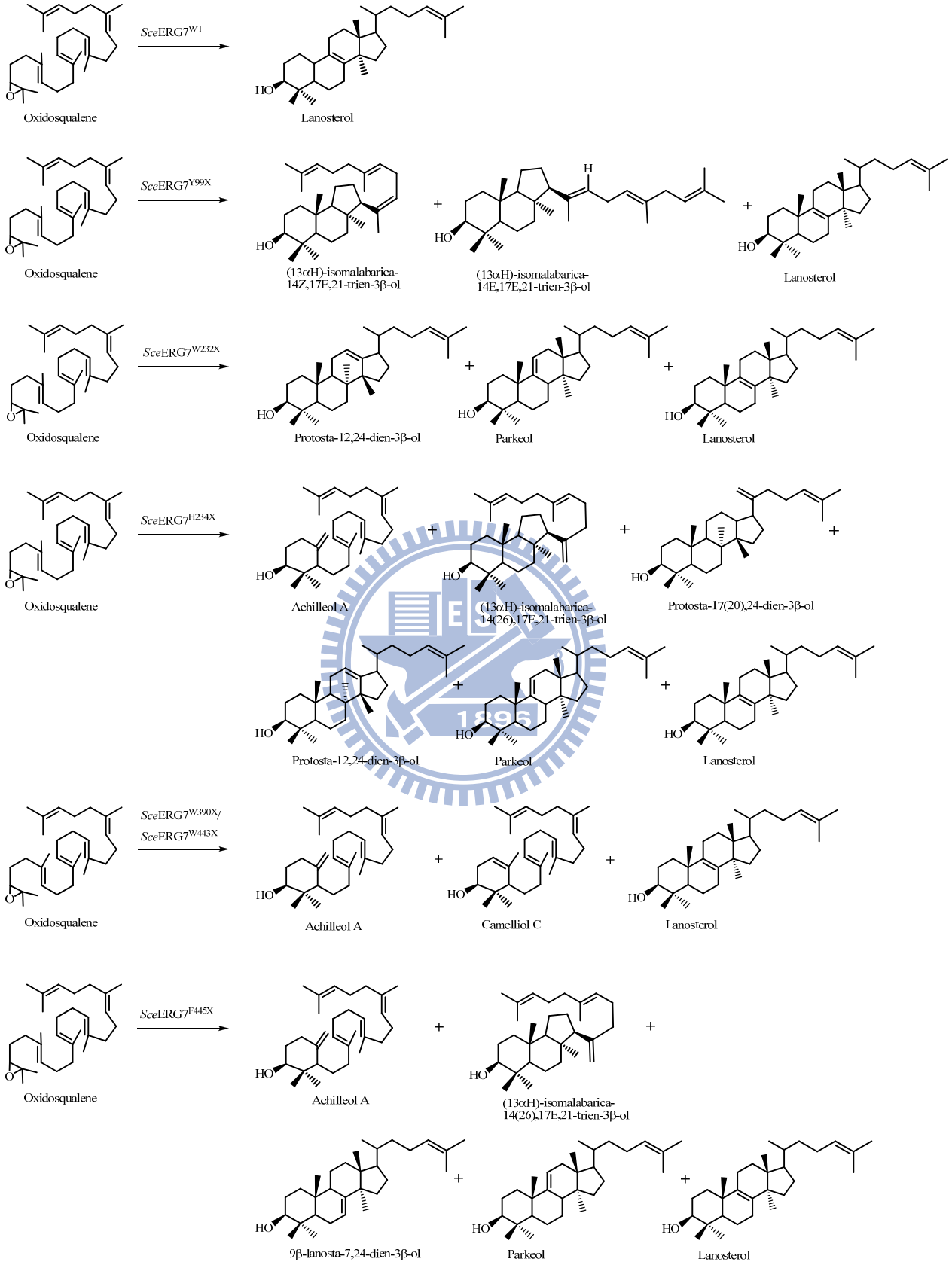


Figure 4.2. Local view of the homology modeled structure of *S. cerevisiae* ERG7 with reactant, lanosterol.


SceERG7WT
Oxidosqualene
Lanosterol
SceERG7Y99X
Oxidosqualene
(13αH)-isomalabarica-14Z,17E,21-trien-3β-ol
(13αH)-isomalabarica-14E,17E,21-trien-3β-ol
Lanosterol
SceERG7W232X
Oxidosqualene
Protosta-12,24-dien-3β-ol
Parkeol
Lanosterol
SceERG7H234X
Oxidosqualene
Achilleol A
(13αH)-isomalabarica-14(26),17E,21-trien-3β-ol
Protosta-17(20),24-dien-3β-ol
Protosta-12,24-dien-3β-ol
Parkeol
Lanosterol
SceERG7W390X/
SceERG7W443X
Oxidosqualene
Achilleol A
Camelliol C
Lanosterol
SceERG7F445X
Oxidosqualene
Achilleol A
(13αH)-isomalabarica-14(26),17E,21-trien-3β-ol
9β-lanosta-7,24-dien-3β-ol
Parkeol
Lanosterol

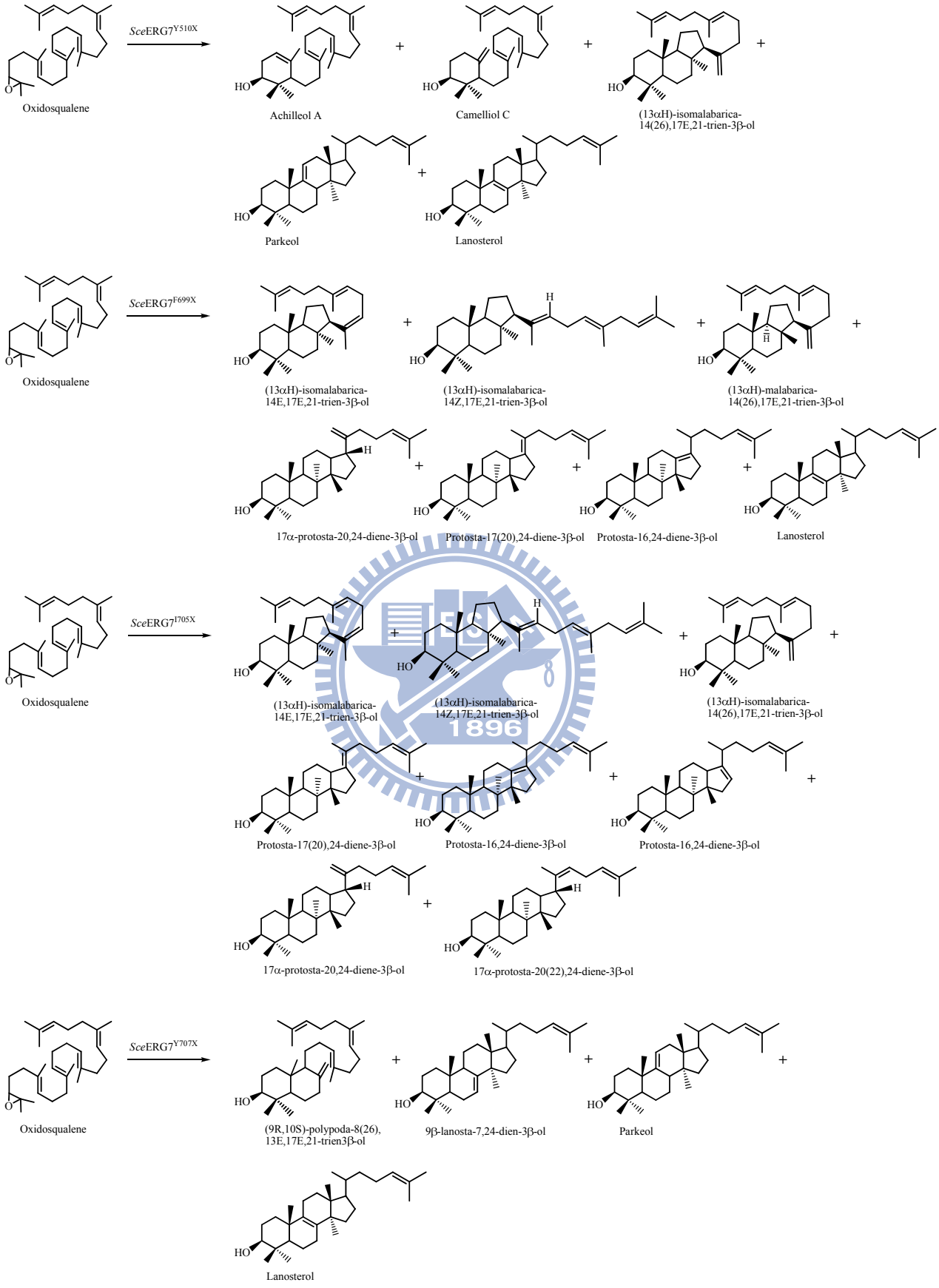


Figure 4.3. The isolated products structure from *S. cereviae* ERG7^{WT} and various ERG7^{mutants}

4.2-3 Correlation between theoretical model results and experimental evidence to investigate the cyclization/rearrangement mechanism of oxidosqualene cyclase

The long-standing question concerning whether the initial epoxide ring opens is concerted or not during A-ring formation and, possibly, B-ring formation has been answered via the experimental isolation of monocyclic achilleol A and camelliol C.[65,73] According to the *S. cerevisiae* ERG7 homology model, Tyr-510 is in an appropriate position and orientation to stabilize the intermediated tertiary cations at C-6, as well as at C-10, just like the role of the Tyr-503 in human OSC.[49] Thus, the mutagenic change of the position of Tyr-510 might disrupt the transition interaction between the hydroxyl group of residue and the carbocation intermediate, leading to premature cyclized products. Similarly, the replacement of Trp-443 with Ala or Lys and the replacement of Lys-448 with Ala resulted in the monocyclic achilleol A and camelliol C.[51,78,79] Although Trp-443 and Lys-448 are located upstream of the putative active site, the functional role of these residues for the cyclization cascade are different. Because of its relatively long distance from the substrate, Trp-443 might stabilize the highly energetic cationic intermediate through the indirect steric effect. Lys-448 is located in the flexible loop region between two helix motifs, and it interacts with Phe-426 and Asn-332 to maintain the correct conformation of the helixes.

Recently, the mutation on Tyr-707 produced a significant amount of the long-awaited C-B bicyclic triterpene, (9*R*,10*S*)-polypoda-8(26),13*E*,17*E*,21-tetraen-3β-ol.[51,77,80] The *S. cerevisiae* ERG7 homology model supported that Tyr-707 is spatially proximal to both the C-10 and C-8 position of lanosterol and is capable of stabilizing the bicyclic cationic intermediate via the hydroxyl group of the phenolic side chain pointing toward the B-ring. Interestingly, a controversial situation exists between the human crystal OSC structure and the *S. cerevisiae* OSC homology model.

X-ray crystallographic analysis of the human OSC structure suggested that Tyr-98 is situated in the spatial position to enforce the formation of the energetically-unfavorable boat B-ring conformation of lanosterol by pushing the methyl group at C-8 below the molecular plane.[49] Thus, the contradictive role between Tyr-707 of yeast OSC and Tyr-98 of human OSC should be further clarified in future research.

After the formation of a 6,6-bicylic A/B-ring, the additional, abnormal, anti-Markovnikov reaction on a six-membered C ring was discussed extensively also. The compound (13α*H*)-isomalabarica-14(26),17*E*,21-trien-3β-ol with a C-B 6,6,5-tricyclic skeleton was obtained by trapping a C-ring-closured, tertiary-Markovnikov-cation intermediate state, which might be stabilized by His-234, Phe-445, and Tyr-510 positions.[68,69,74,79] Homology modeling of the *S. cerevisiae* ERG7 indicated that Phe-445, located at the spatially proximal location relative to the B/C ring junction, exhibited the closest distance to the C-8/C-14 position of the lanosterol. In addition, experimental data further illustrated that Phe-445, through cation-π interaction, might have a definitive role, along with other π-electron-rich aromatic residues in the cyclase active site, in stabilizing the discrete transition-state intermediates.[74] His-234 of *S. cerevisiae* ERG7, positioned spatially near the C-13 and C-20 positions of lanosterol, might influence the stabilization of the C-13 and C-20 positions as well as the C-14 Markovnikov cation during the formation of the rings. In parallel, although Tyr-510 occupies the front portion above the molecular plane of the substrate in the cyclase active-site cavity, the replacement of this Tyr-510 position might disrupt the hypothetical hydrogen-bonding dyad between Tyr-510 and His-234, causing slight influence on the orientation of His-234 for stabilizing the C-14 cation and resulting in the isolation of a tricyclic byproduct.[51,79]

Moreover, C-B 6,6,5-tricyclic products with different stereochemical controls, i.e.,

(13α*H*)-isomalabarica-14*Z*,17*E*,21-trien-3β-ol and (13α*H*)-isomalabarica-14*E*,17*E*,21-trien-3β-ol, were isolated from the Tyr-99 mutants and the Phe-699 mutants.[75,76,78,79,83] Tyr-99 was located near the position of Phe-445 residue and close to the C/D ring junction of the substrate in the *S. cerevisiae* ERG7. The short distance allows Tyr-99 to stabilize the electron deficient C-14 cationic intermediate. The production of these two stereochemical skeletons from Tyr-99 mutants indicated that Tyr-99 might affect the enzymatic active-site cavity, resulting in a slight rotation of the long carbon side chain along the axis of the C-14/C-15 backbone. The role of Phe-699 was predicted to stabilize the C-17 cation. The truncated tetracyclic rearrangement products abstracted at the C-17 position further verify its role.[75,76,79] However, the two stereochemically distinct compounds illustrated above, along with a C-C-C tricyclic malabarica-14*E*,17*E*,21-trien-3β-ol, expanded the critical role of Phe-699 in the cyclization reaction. The corresponding residue of Phe-699 in the human OSC is Phe-696, which was suggested to act as a stabilizer for the anti-Markovnikov C-13 cation during C-ring formation. The correlation between human OSC and yeast ERG7 active-site cavities is still obscure, but elucidating the structure might provide insight so that the correlation could be determined. The deletion of one amino-acid residue between Gly-697 and Val-698 might result in a more closely-connected adjacent residue between Tyr-707 and Tyr-99 and form a restricted girdle around the B/C/D ring region and exocyclic side chain. Substitution on Phe-699 might cause an inappropriate position of the side chain and/or destroy the steric or electrostatic interaction within the enzymatic cavity, thus allowing the unconstrained folding of the substrate. Hence, the Phe-699 mutant changes the course of the oxidosqualene cyclization/rearrangement process via traversing the mechanistic barriers between the pre-folded C-B-C and C-C-C substrate conformations.

Many 6,6,6,5-truncated rearrangement tetracyclic products were isolated from

various *S. cerevisiae* ERG7 mutants, including protosta-20,24-dien-3β-ol,[69] protosta-17(20),24-dien-3β-ol,[76,79,81] protosta-16,24-dien-3β-ol,[81] protosta-13(17),24-dien-3β-ol,[79] protosta-12,24-dien-3β-ol,[68,84] parkeol,[68,74] 9β-lanosta-7,24-dien-3β-ol [74,77,80] with a 17β exocyclic hydrocarbon side chain, and two additional 17α-derivatives, i.e., 17α-protosta-20,24-dien-3β-ol[76,79] and 17α-protosta-20(22),24-dien-3β-ol[81]. When the D-ring was annulated, the protosteryl C-20 cation was generated. The Phe-699 located at a spatial position near the D-ring/exocyclic terminal hydrocarbon side chain junction that might utilize the π-rich phenyl group to act as an optimal assistance for stabilizing the C-17 and C-20 cationic intermediate. Incompletely rearranged products that eliminated the surrounding proton migration at C-13 or C-20, respectively, were isolated from Phe-699 mutants. These compounds provide evidence to explain the involvement and the role of Phe-699 in the H-17α→20α or H-13α→17α hydride shifts process. Otherwise, the argument regarding the stereochemical control of the exocyclic side chain with a 17β configuration, in the *S. cerevisiae* pathway, has also been answered. Phe-699 and a second-sphere residue, Ile-705, might contribute to the stereochemical selectivity by directly affecting the C-17 orientation and provide a correct spatial direction toward the C-17 position of lanosterol, respectively. As indicated in previously published results, the sterically bulky group of the amino acid residues might control the stereochemical density in the SHC.[9,85] Substitution of Phe-699 and Ile-705 positions gave 17α-protosta-20,24-dien-3β-ol and 17α-protosta-20(22),-

24-dien-3β-ol, which process an α-configuration of the exocyclic side chain, to elucidate the importance of these position for the stereochemical selectivity of oxidosqualene cyclization. In parallel, a direct deprotonation of the protosteryl cation generated protosta-20,24-dien-3β-ol, accompanied by subsequent skeletons of rearranged protosta-12,24-dien-3β-ol and parkeol, which was isolated from His-234

mutants with a small, non-polar, hydrophobic residues substitution. The resulting product distribution further confirmed that His-234 in the active-site cavity of *S. cerevisiae* ERG7 is the only basic residue in proximity to the termination site and, thus, could abstract the proton to stop the reactions, as mentioned previously. In addition, a hydrogen bond between His-234 and the hydroxyl group of Tyr-510 might locate in the more proper position to accept the proton from C-8/C-9 of the lanosteryl cation. The single substitution of Tyr-510 or simultaneous substitutions of Tyr-510 and His-234 for the spatially bulky groups or some steric residues, might disrupt the electrostatic interaction between Tyr-510 and His-234 and acquire abortive cyclized or alternatively deprotonated products that further support the relationship of the His-234:Tyr-510 hydrogen-bonding dyad. On the other hand, replacement of Trp-232 and Phe-445 produced many abnormal, alternatively-rearranged compounds that might imply indirect interactions with spatially-affected, predominant active-site residues. Homology simulation of the Trp-232 position with other residues revealed that change at this position might reduce the π-electron density, affect the His-234:Tyr-510 catalytic hydrogen-bonding network, and shift the relative distance between the His-234 and C-11/C-12 lanosterol positions. Similarly, the terminal phenyl group of Phe-445 might interact with the nearby aromatic residues, especially Tyr-510, during the rearrangement process to provide the high electron density gradient for the equilibrium shift toward the final deprotonation reaction at the lanosteryl C-8/C-9 cation. The partial disturbance of the transient cation-π interaction might affect the equilibrium tendency for the C-8/C-9 proton position and, consequently, result in the possible formation of alternatively-deprotonated products.

4.3 Conclusion

In summary, the above-illustrated results, from either the isolation and characterization of unnatural but mechanistically-postulated products in various site-saturated mutations or from the homology simulation of amino-acid residues of interest in the *S. cerevisiae* ERG7 structure, elucidated the functional role of these residues involved in the ERG7-templated oxidosqualene cyclization/rearrangement cascade. First, Asp-456, hydrogen-bonded to two cysteine residues (Cys-457 and Cys-540), is crucial for the initiation step in the ERG7-catalyzed reaction; Trp-443 might play a role in influencing substrate binding, stabilizing epoxide protonation, and inducing A-ring formation. Also, Val-454 employs the steric effect in assisting the orientation of the substrate during cyclization. After the epoxide ring opening and subsequent cation-π annulations, the C-B bicyclic C-8 cation and the C-B 6,6,5-Markovnikov tricyclic C-14 cation are stabilized by Tyr-510, Tyr-707, Tyr-99, and His-234, respectively. Then, His-234 and Phe-699 further stabilize the D-ring closured protosteryl C-20 cation as well as the first hydride rearranged lanosteryl C-17 cation by their electron-rich functional group. Trp-232 reveals an influence on the coordinative action between His-234 and Tyr-510, thereby becoming involved in the rearrangement process. His-234 and Tyr-510 further stabilize the incoming respective cationic intermediates during the rearrangement stages and guide the final deprotonation step. Moreover, Phe-445 influences both the C-ring formation and deprotonation step (Figure 4.4). These data demonstrate that: i) through the mutagenesis approach and concomitant product isolation/characterization, the significance of individual active-site residues and their roles in each step of the oxidosqualene cyclization/rearrangement cascade is made clear; ii) only a small modification of the enzymatic cavity may readily alter the product specificity to

further illustrate that the cyclase enzymes are evolutionarily plastic enzymes; iii) site-saturated mutagenesis with 19 proteogenic amino-acid residue substitutions provides insight concerning the impact of modifying the enzymatic cavity environment on the structure-function relationship.

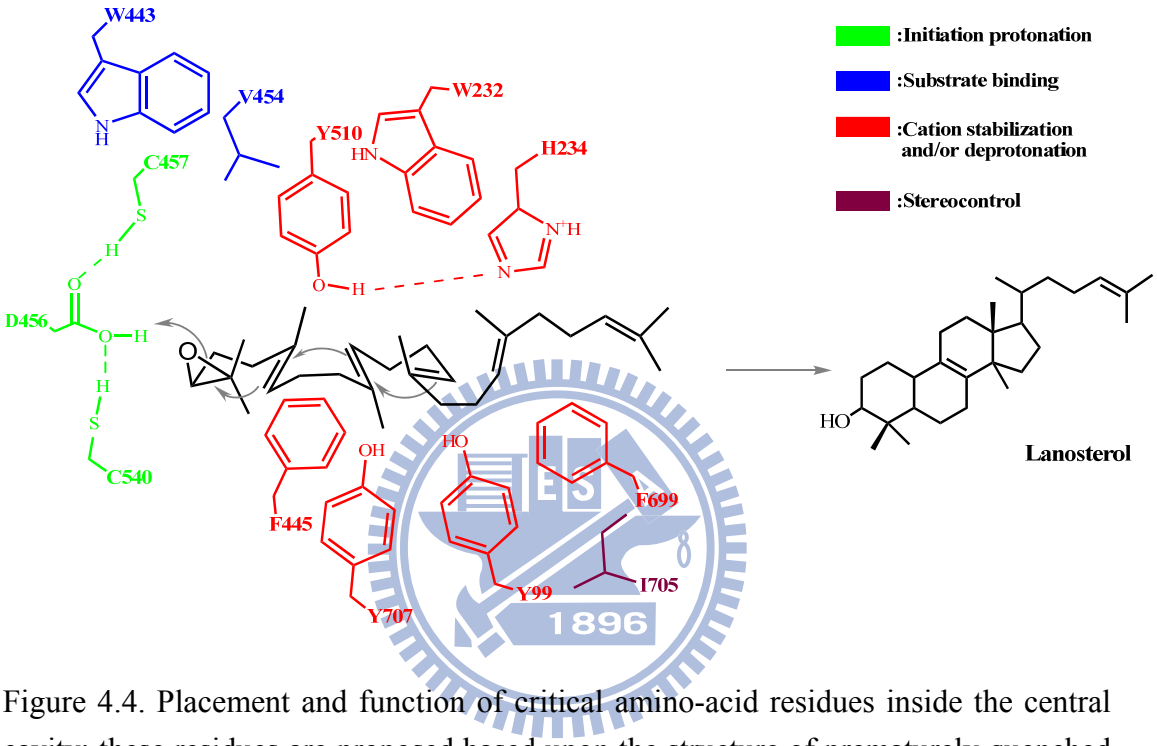


Figure 4.4. Placement and function of critical amino-acid residues inside the central cavity; these residues are proposed based upon the structure of prematurely quenched products in conjugation with the *S. cerevisiae* ERG7 homology model.

CHAPTER 5

Site-directed mutagenesis study on the deprotonation course of oxidosqualene cyclization

5.1 Research background and aim

The cyclase proteins rely on a small number of amino-acid modifications of the structure of the enzymes and, even though they share the same substrate, to generate diverse sterol and triterpene skeletons. It is generally believed that the result of divergent evolution is caused when the primitive enzyme adapts to the mutagenic pressure from its surrounding environment.[86-88] Accordingly, these cyclase enzymes are categorized as either accurate-cyclase or multifunctional-cyclase, defined by the patterns of their product distributions. Accurate-cyclases, such as sterol-producing enzymes, i.e., OSC and CAS, seem to accomplish the catalytic perfection to make a remarkable major product without any slight errors, whereas multifunctional cyclases generate a diverse array of products, and none of them is dominant. Most of triterpene alcohol-producing synthases are representative of this example. Therefore, experiments on these two classes of cyclases could provide some insight for better understanding the relationship between enzymatic structure-function and product specificity/diversity.[54,87] Domain-swapping strategies have been set up for these two classes of cyclases to elucidate in detail the functional residues that especially contribute to enzymatic specificity and result in the defined polycyclic product.[89,90]

A series of chimeric enzymes was created either for the construction of the chimeras between β-amyrin synthase from *Panex ginseng* and lupeol synthase from *A. thaliana* or for the construction of the chimeras between OSC form *S. cerevisiae* and CAS from *A. thaliana*.[65,89] According to a successful experiment on the triterpene

alcohol-producing synthases, a functional residue within the region-B of β-amyrin synthase was found to control and regulate product specificity.[89] However, the loss of enzymatic function was examined in individual chimera of sterol-producing enzymes. The difference between these two classes of cyclases might be the result of a discrepancy in the primary amino acid sequence. Due to the high-identity, large-scaled domains exchanging among these triterpene alcohol-producing synthases, only minor changes are generated on the primary amino acid sequence. Hence, the general protein skeleton was maintained and used to generate diverse product profiles. This explains why most of the triterpene alcohol-producing synthases are multi-functional enzymes. In contrast, most sterol-producing cyclases possess an extreme sequence distinction, implying that the determination of product specificity might be controlled by a few conserved residues or motifs. Thus, domain-swapping experiments might not be a suitable strategy for studying the product's accuracy of sterol-producing cyclase.

According to the results of the domain-swapping experiment, it is necessary to identify and characterize the catalytic residues or motifs that are responsible for the difference in the catalytic promotion of the specific deprotonation reaction to generate the distinct product in individual cyclase-mediated reactions. OSC and CAS catalyze a closely-related reaction via the same cyclization and rearrangement steps, but they abstract a distinct proton either at C-8/C-9 or at the C-19 position to form lanosterol or cycloartenol, respectively. The pairwise sequence alignment of OSC and CAS indicated the Thr-384, Gln-450, and Val-454 are highly-conserved patterns among all of the known OSCs. Also, the corresponding positions of these three amino-acid residues are associated with different conserved residues in CAS (Figure 5.1).

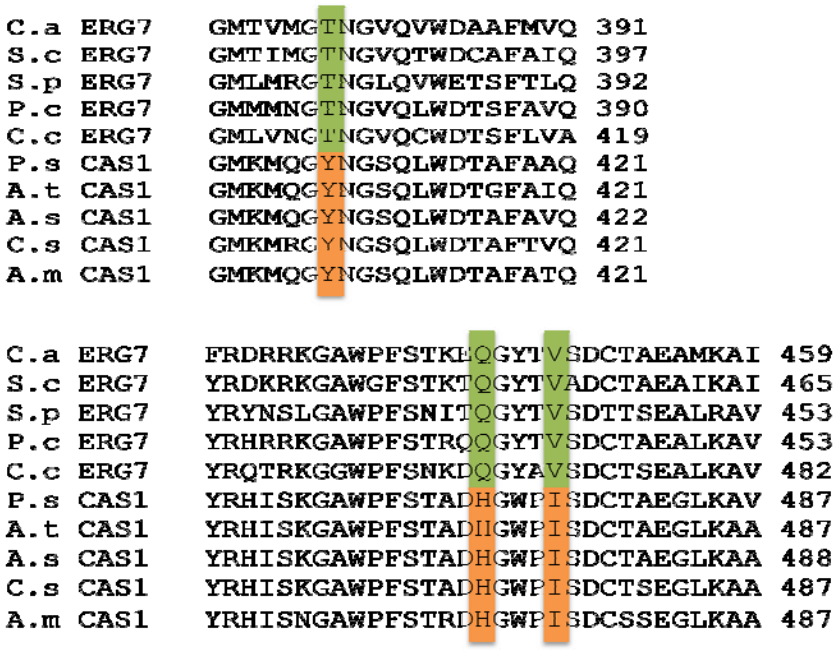


Figure 5.1. Alignment of multiple sequences of CAS and OSC (conserved residues labeled with color)

The observed alignments imply that these amino-acid residues might be involved in determining product profiles. According to the homology modeling for *S. cerevisiae* ERG7, these three amino-acid residues were located on the top of the molecular plane of the substrate (Figure 5.2). Previous studies have reported that the final deprotonation process was executed by specific amino-acid residues that were situated at the ceiling of the active-site cavity.[49,69] Hence, it is reasonable to assume that the crucial, product-altering residues or motifs might occupy the space at the top of the molecular plane of the substrate.

In this chapter, site-directed mutagenesis were used to examine several functional residues, especially on Thr-384, Gln-450, and Val-454, in order to discuss how sterol-producing cyclases precisely remove specific hydride positions to create the tendency of the product specificity under enzymatic catalysis.

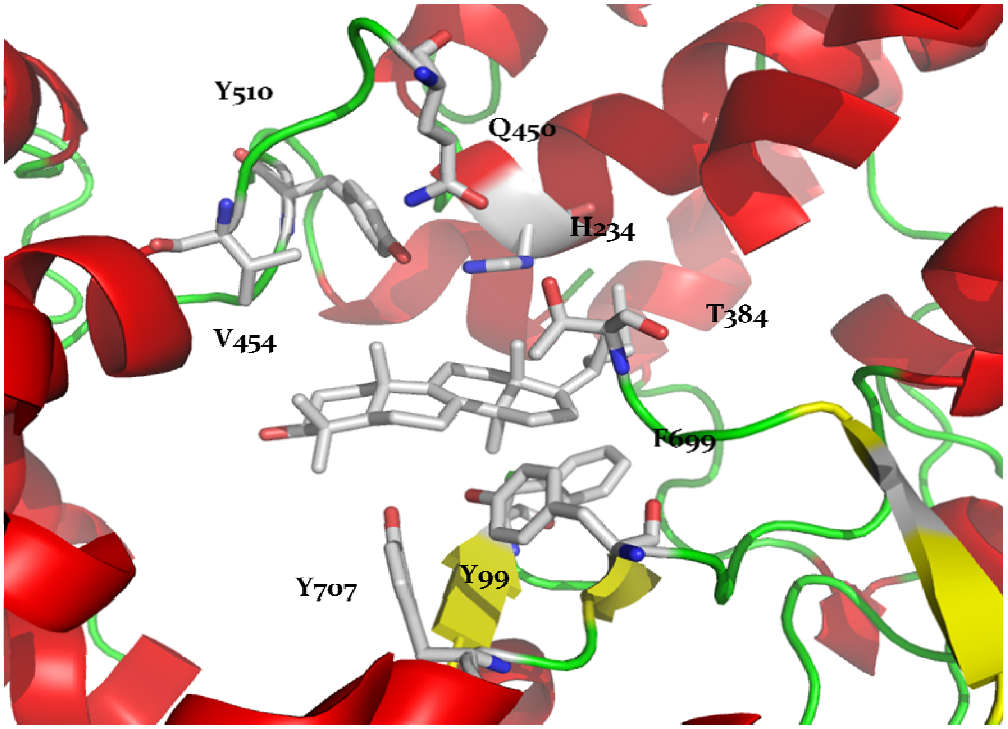


Figure 5.2. Local view of the homology modeled structure of *S. cerevisiae* ERG7; Thr-384, Gln-450, and Val-454 residues located on the top of the molecular plane of substrate.

5.2 Results and Discussion

5.2-1 Generation of the site-directed mutations on the *S. cerevisiae* ERG7 gene and functional analysis of site-directed mutations via the plasmid shuffle methodology and product characterization

Site-directed mutagenesis was conducted to exchange three amino acids in *S. cerevisiae* ERG7 gene into their corresponding amino-acid residues in CAS (i.e., Thr-384 into Tyr, Gln-450 into His, and Val-454 into Ile) by using the QuikChange™ Site-Directed Mutagenesis Kit. In addition to single-point mutation, different combinations of multiple-position mutants (double- or triple-point mutation) were also designed. Thus, the plasmid-carrying *S. cerevisiae* ERG7^{T384Y}, *S. cerevisiae* ERG7^{Q450H}, *S. cerevisiae* ERG7^{V454I}, *S. cerevisiae* ERG7$^{T384Y/Q450H}$, *S. cerevisiae* ERG7$^{Q450H/V454I}$, *S. cerevisiae* ERG7$^{T384Y/V454I}$, and *S. cerevisiae* ERG7$^{T384Y/Q450H/V454I}$

were constructed and then transformed into the *S. cerevisiae* TKW14c2 strain, as described in the previous chapter. The product profiles generated by various *S. cerevisiae* ERG7 mutant enzymes were further analyzed and characterized by GC-MS. Analysis of three single-point mutations (*S. cerevisiae* ERG7^{T384Y}, *S. cerevisiae* ERG7^{Q450H}, and *S. cerevisiae* ERG7^{V454I}) showed that the accurate enzymatic function of OSC was interrupted when residues with different characteristics were substituted into the respective active-site residues. The analysis further supported the conclusion that cyclases maintain their product specificity via a sophisticated control. The *S. cerevisiae* ERG7^{Q450H} mutant converted oxidosqualene into achilleol A. This implied that introducing the positively-charged residue to this position of the enzymatic active-site cavity might facilitate premature deprotonation, cause early closure of the A-ring, and lead to the formation of a monocyclic compound. The *S. cerevisiae* ERG7^{V454I} mutant formed lanosterol and (13α*H*)-isomalabrica-14(26),17*E*,21-trien-3β-ol. Surprisingly, Thr-384 caused the strongest effect on the product profile of all of the *S. cerevisiae* ERG7 mutants. The *S. cerevisiae* ERG7^{T384Y} mutant biosynthesized lanosterol and multiple alternative rearrangement products, including protosta-16,24-dien-3β-ol, protosta-13(17),24-dien-3β-ol, 9β-lanosta-7,24,dien-3β-ol, and parkeol. Simultaneously, double-position mutations at position 384 and either at position 450 or 454 have synergistic effects that favor the formation of parkeol. The *S. cerevisiae* ERG7$^{T384Y/Q450H}$ and *S. cerevisiae* ERG7$^{T384Y/V454I}$ mutants produced up to 80% parkeol as their major product, with minor amounts of some byproducts of lanosterol and 9β-lanosta-7,24,dien-3β-ol. Surprisingly, the simultaneous substitutions on these three amino-acid residues effectively presented a dominant tendency for the generation of parkeol. The *S. cerevisiae* ERG7$^{T384Y/Q450H/V454I}$ triple mutations completely eliminated the biosynthesis of lanosterol, but, alternatively, they generated the largest distribution of parkeol. These results indicated that these functional

residues have quite different catalytic properties than the wild-type *S. cerevisiae* ERG7. The cyclization and rearrangement cascades for these ERG7 mutants were identical to those of the wild-type *S. cerevisiae* ERG7; alternatively, the migration of the hydride/methyl groups of carboncationic intermediates was arrested at the C-9 position, with subsequent deprotonation occurring at C-10 to form parkeol as the sole product (Table 5.1, Figure 5.3).

Table 5.1. Product distribution and its ratio from *S. cerevisiae* ERG7 mutants

Mutants	Ergosterol complementary assay	Product distributions and ratio						
		AA	H234L-Tricyclic	13,17-LA	16,17-LA	PK	LA	9β-LA
T384Y	Live			1	1	31.5	63	3.5
Q450H	Die	100						
V454I	Live		10					90
T384Y+Q450H	Live (slow)					98	1	1
Q450H+V454I	Live		16		9		75	
T384Y+V454I	Live					84	10	6
T384Y+Q450H+V454I	Die					100		

Abbreviate: AA, Achilleol A; 9β-LA, 9β-lanosta-7,24,dien-3β-ol; H234L-tricyclic, (13αH)-isomalabrica-14(26),17*E*,21-trien-3β-ol; PK, parkeol; 13,17-LA, protosta-13(17),24-dien-3β-ol; 16,17-LA, protosta-16,24-dien-3β-ol.

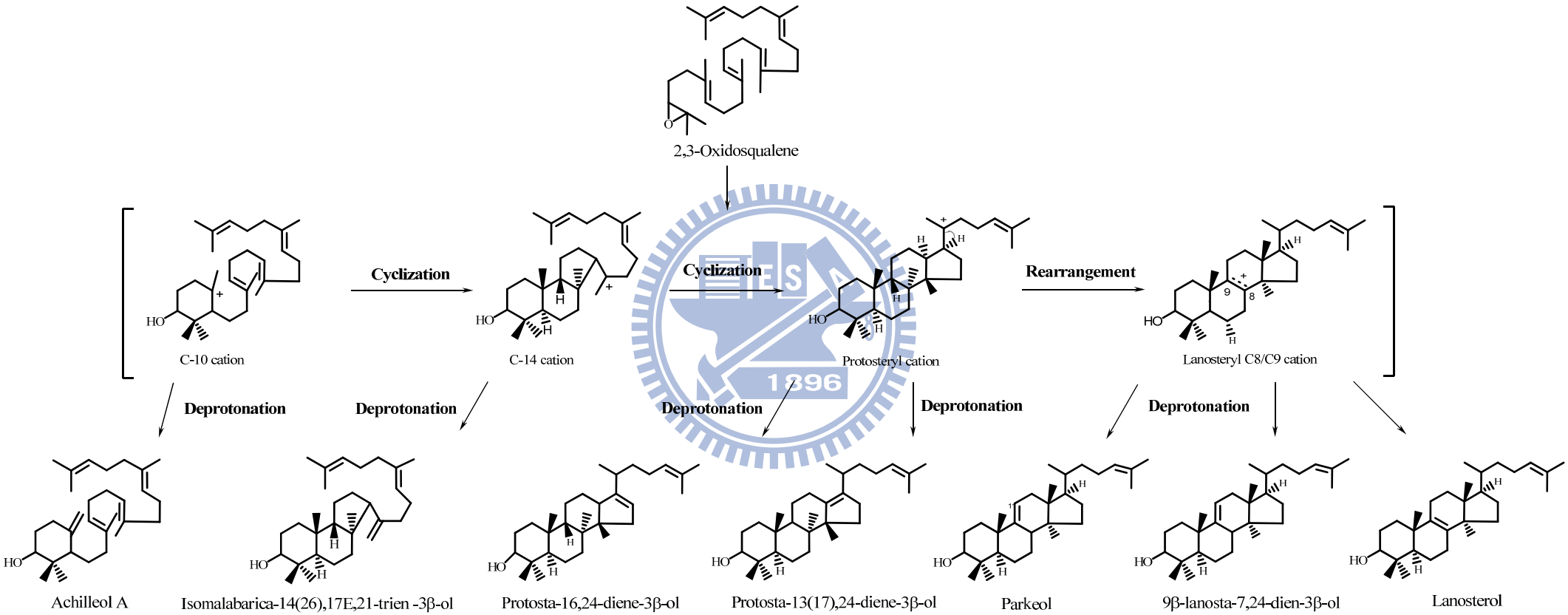


Figure 5.3. Oxidosqualene cyclization products formed by OSC mutants from *S. cerevisiae*

5.2-2 Homology modeling illustration of critical residues on enzyme function

From the observation of the homology model of *S. cerevisiae* ERG7, Gln-450 and Val-454 residues were found to be located outside of the active-site cavity in the *S. cerevisiae* ERG7, and they belonged to the second-tiered residues (Figure 5.2). The striking change in the product profile might be mediated through the molecular interaction between these two amino acids and the Thr-384 by side-chain orientation and steric effects. In the CAS, Tyr-410 (the corresponding residue of Thr-384) and His-257 (the corresponding residue of His-234) are hydrogen-bonded pairs and lie near the C-9/C-19 position where the deprotonation occurred to produce cycloartenol. Thus, the substitution of Thr-384 with Tyr might destroy the original Tyr-510:His-234 hydrogen-bonded dyad and provide the opportunity to make new hydrogen-bonded pairs. (It was observed that the original distance of 7 Å had obviously been reduced to 3.5 Å between Thr-384 and His-234.) It is also conceivable that the proton donors might be altered, causing a variety of rearranged products and product distributions. In addition, the simultaneous mutations at these three positions might cause steric effects between Gln-450 and Val-454, which could create a slight rotation of Tyr-384 and allow the acceptance of the C-11 proton, thereby acquiring parkeol.

Even so, the replacement of the active-site residues of *S. cerevisiae* ERG7 with the corresponding residues of CAS does not cause the enzyme to produce the envisioned product, i.e., cycloartenol. The accurate parkeol synthase from one *S. cerevisiae* ERG7 $^{T384Y/Q450H/V454I}$ mutant was first reported. These *S. cerevisiae* ERG7 mutants may assist the hydride shift from C-8 to C-9 and, therefore, promote the relatively stable formation of parkeol. Due to the energetically-unfavored formation of a strained cyclopropyl ring, it is resonable to explain why the nearly-exclusive production of parkeol and the absence of cycloartenol occurred in this *S. cerevisiae* ERG7 $^{T384Y/Q450H/V454I}$ mutant. It also indicated that some residues are still required to

facilitate the abstraction of the proton from C-19 for obtaining the cycloartenol. Hence, at least one of the additional residues that have yet to be identified in the *S. cerevisiae* ERG7 mutants, or perhaps other potential interactive amino acid pairs, must be sought for the next step to determine the entity that serves as the stabilizing force for the formation of the three-member ring.

Otherwise, parkeol, a product with $\Delta^{9(11)}$-lanosterol, and its chemical derivatives showed an important significance in echinoderms, i.e., that it probably can be used against environmental toxin action, related to the saponins.[91-94] Spheciosterol sulfates A-C and topsentiasterol sulfate E, which possess a parkeol-based skeleton, demonstrated the ability to act as PKCζ inhibitors, and they also inhibit the activation of the NF-κB pathway.[95] Therefore, parkeol might be used as a core framework, via the following tailoring enzymatic modification *in vivo*, or a chemical alternation *in vitro*, to create an opportunity for the development of compounds that are potentially bioactive. So far, the natural parkeol synthase gene has not been cloned, and sequence information is still unavailable; thus, the *S. cerevisiae* ERG7$^{T384Y/Q450H/V454I}$ mutant might describe and provide an approximate model for this enzyme.

5.2-3 Investigation of product modification by subsequent triterpene tailoring enzymes in *S. cerevisiae* by isolating and identifying unexpected downstream products

In parallel, there are two additional sterol products that were detected by GC-MS spectra of the NSL extract from the ERG7$^{T384Y/Q450H/V454I}$ mutant. The spectra show compounds with a tetracyclic-like skeleton, even though the molecular weight of one of the compounds is not *m/z* 426. After acquiring a large-scale culture of *S. cerevisiae* ERG7 $^{T384Y/Q450H/V454I}$, the NSL fractions were purified by HPLC to obtain the pure compounds, which were further characterized by extensive NMR spectra. The

structure of a novel compound with molecular ion at *m/z* 426 indicated that the compound was 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol, a tetracyclic product with a parkeol nucleus but only one methyl group remaining at the C-4 carbon; also, its C-24 carbon presents a germinal-substituted double bond (Figure 5.4, Table 5.2, Appendix 5.1). The distinct ^{1}H-NMR chemical shifts with one olefinic proton (δ 5.27) and two methylene protons (δ 4.61, 4.66), as well as seven methyl singlets (δ 1.07, 1.06, 1.03, 0.99, 0.95, 0.79, and 0.71), were observed. The 150 MHz ^{13}C-NMR spectrum revealed the presence of one secondary-quaternary-substituted and one tertiary-quaternary-substituted double bond (δ 106.47, 157.81 and 116.94, 147.38 ppm). The HSQC spectrum showed that the methylene protons at δ 4.61 and 4.66 are attached to the carbon at 106.47 ppm (C-24), while the olefinic proton at δ 5.27 is attached to the carbon at 116.94 ppm (C-11). In the HMBC spectrum, the δ 5.27 proton is coupled by ^{2}J to the carbon at 38.14 ppm (C-12), as well as by ^{3}J connectivity to carbons at 39.41 ppm (C-10), 42.17 ppm (C-8), and 45.03 (C-13). A methyl proton δ1.03 is coupled by ^{2}J to a carbon at 39.41 ppm (C-10), by ^{3}J to carbons at 36.23 ppm (C-1), 50.09 ppm (C-5), 40.24 ppm (C-4), 42.17 ppm (C-8), and 147.34 ppm (C-9). In addition, the carbon at 40.24 ppm (C-4) is connected to only one methyl singlet at δ3.06, indicating that it is a C-4 demethylated metabolite. The HMBC also established that a methyl proton δ 0.71 (Me-18) is coupled by ^{2}J to a carbon at 45.03 ppm (C-13), as well as by ^{3}J to carbons at 38.14 ppm (C-12), 47.85 ppm (C-14), and 51.68 ppm (C-17). Another methyl proton δ 0.95 (Me-21) is coupled by ^{2}J to a carbon at 36.92 ppm (C-20), as well as by ^{3}J to carbons at 28.73 ppm (C-16) and 35.83 ppm (C-22). A tertiary carbon at 36.92 ppm (C-20) is coupled by ^{2}J to protons at δ 1.65-1.65 (H-22) and δ 1.66-1.71 (H-17), and it is coupled by ^{3}J to a proton at δ 1.862-1.981 (H-22). In addition, a quaternary carbon at 157.80 ppm (C-24)

is coupled by ^{2}J to protons at δ 2.26–2.30 (H–25), δ 2.14-2.19 (H-23), and δ 1.86-1.98 (H-23), as well as by ^{3}J to protons at δ 1.05 (H–26) and δ 1.06 (H–27). Finally, the presence of NOE between Me-28/H-3, Me-18/H-7, and Me-29/H-17, as well as the absence of NOE between Me-19/H-3, confirmed the structure to be 4α,14α-dimethyl-24-methylene-5α-cholest-9(11),24-dien-3β-ol.

In parallel, the spectra information of another product with molecular ion at *m/z* 412 established its structure as 4α,14α-dimethyl-5α-cholest-9(11),24-dien-3β-ol, a tetracyclic compound presenting a 4-demethylated $\Delta^{9(11),24}$ metabolite of parkeol (Figure 5.5, Table 5.3, Appendix 5.2). As with the previously mentioned compound, the very closely matched values of NMR spectra were conserved between these two structures, except for the position of the methyl group substitution. The 150 MHz ^{13}C-NMR spectrum revealed the presence of two pairs of tertiary-quaternary substituted double bonds (δ 125.96, 131.55 and 116.96, 147.32 ppm). Connectivities from HMBC, COSY, and distinct NOE spectra further confirmed the structure as 4α,14α-dimethyl-5α-cholest-9(11),24-dien-3β-ol. These two compounds are being reported for the first time in this thesis, and this treasured information will facilitate the exploration of the effect of unusual triterpene products that are generated and that impact the action of the downstream tailoring enzymes in the *S. cerevisiae* sterol biosynthesis pathway.

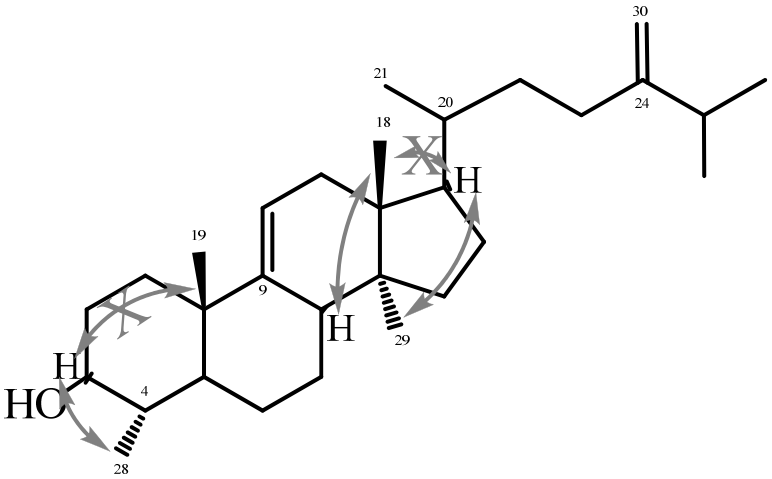


Figure 5.4. Characterization of the structure and the results of the NOE correlation of 4α,14α-dimethyl-24-methylene-5α-cholest-9(11),24-dien-3β-ol

Table 5.2. ^{1}H- and^{13}C-NMR data of 4α,14α-dimethyl-24methylene-5α-cholest-9(11), 24-dien-3β-ol. (150 MHz for ^{13}C-NMR or 600 MHz for ^{1}H-NMR in CD_2Cl_2)

Carbon	^{13}C	Proton	^{1}H	Carbon	^{13}C	Proton	^{1}H
1	36.23	1	1.420-1.457, 1.842-1.845	16	28.73	16	1.35-1.38, 1.94-1.97
2	32.08	2	1.50-1.58, 1.86-1.91	17	51.68	17	1.66-1.71
3	76.98	3	3.00	18	14.98	18	0.71
4	40.24		1.30-1.35	19	21.08	19	1.03
5	50.09	5	0.82-0.86	20	36.92	20	1.43-1.49
6	24.83	6	1.79-1.84, 1.23-1.56	21	18.95	21	0.95
7	28.25	7	1.63-1.67, 1.30-1.35	22	35.83	22	1.71-1.21, 1.62-1.65
8	42.17	8	22-226	23	32.01	23	1.86-1.98, 2.14-2.19
9	147.34			24	157.80		
10	39.41			25	34.57	25	2.26-2.30
11	116.94	11	5.27	26	22.54	26	1.06
12	38.14	12	1.94-1.99, 2.12-2.19	27	22.41	27	1.07
13	45.03			28	15.83	28	0.99
14	47.85			29	18.89	29	0.79
15	34.64	15	1.40-1.46	30	106.47	30	4.61, 4.66

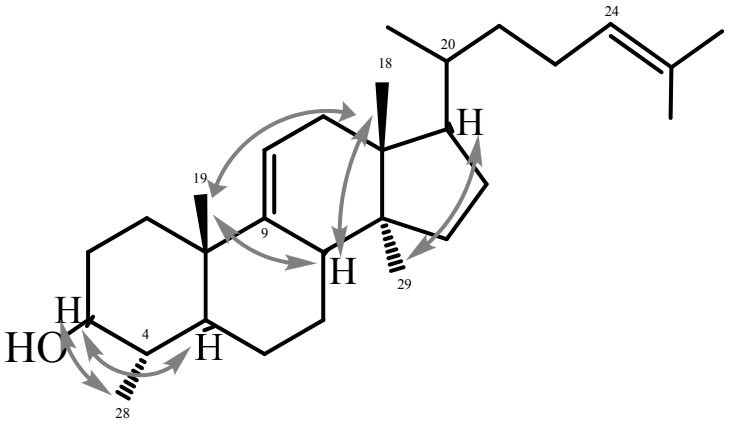


Figure 5.5. Characterization of the structure and the results of the NOE correlation of 4α,14α-dimethyl-5α-cholest-9(11),24-dien-3β-ol

Table 5.3. ^{1}H- and^{13}C-NMR data of 4α,14α-dimethyl-5α-cholest-9(11),24-dien-3β-ol. (150 MHz for ^{13}C-NMR or 600 MHz for ^{1}H-NMR in CD_2Cl_2)

Carbon	^{13}C	proton	^{1}H	Carbon	^{13}C	Proton	^{1}H
1	36.21	1	1.41, 1.43, 1.84-1.86	16	28.73	16	1.90-1.96
2	28.23	2	1.64-1.68	17	51.70	17	1.65-1.67
3	79.96	3	3.07	18	14.97	18	0.75
4	40.23		1.30-1.35	19	21.00	19	1.03
5	50.08	5	0.84	20	36.68	20	1.43-1.45
6	24.82	6	1.79-1.82, 1.25-1.36	21	18.87	21	0.94
7	32.06	7	1.45-1.60, 1.86-1.91	22	25.64	22	1.89-1.94, 2.06-2.09
8	42.14	8	2.23-2.26	23	37.14	23	1.10-1.06
9	147.32			24	125.96	24	5.14
10	39.40			25	131.55		
11	116.96	11	5.34	26	26.17	26	1.64
12	38.12	12	1.11-1.42, 1.95-1.97	27	18.06	27	1.71
13	45.00			28	15.82	28	1.00
14	47.83			29	18.07	29	0.78
15	34.63	15	1.36-1.45				

In general, in *S. cerevisiae*, the following key transformations from lanosterol to ergosterol take part in the order indicated: (1) C-14 demethylation/C-15 deprotonation by ERG11 and ERG24; (2) subsequent C-4 demethylation by ERG25, ERG26, and ERG27; (3) methylation/deprotonation at C-24 by ERG6; (4) double bond transformation in the nucleus from Δ^8 to Δ^7 and a proton abstraction at C-6 by ERG2 and ERG 3; and (5) double bond transformation in the side chain from C-24(28) to C-22(23) by ERG5 and ERG4 (Figure 5.6).[96,97] However, we propose that, when the parkeol accumulates in the yeast cell, one of C-4 methyl group is removed first to obtain 4α,14α-dimethyl-5α-cholest-9(11),24-dien-3β-ol, followed by subsequent methylation/deprotonation at C-24 to obtain 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol. Hence, acquiring these two compounds might be regarded as a subversively cognitive mechanism (Figure 5.7). Other alternative biosynthesis pathways probably exist for these non-established triterpene compounds or the downstream flux of individual enzymes, especially ERG11 or ERG24, that might possess a strict substrate specificity.

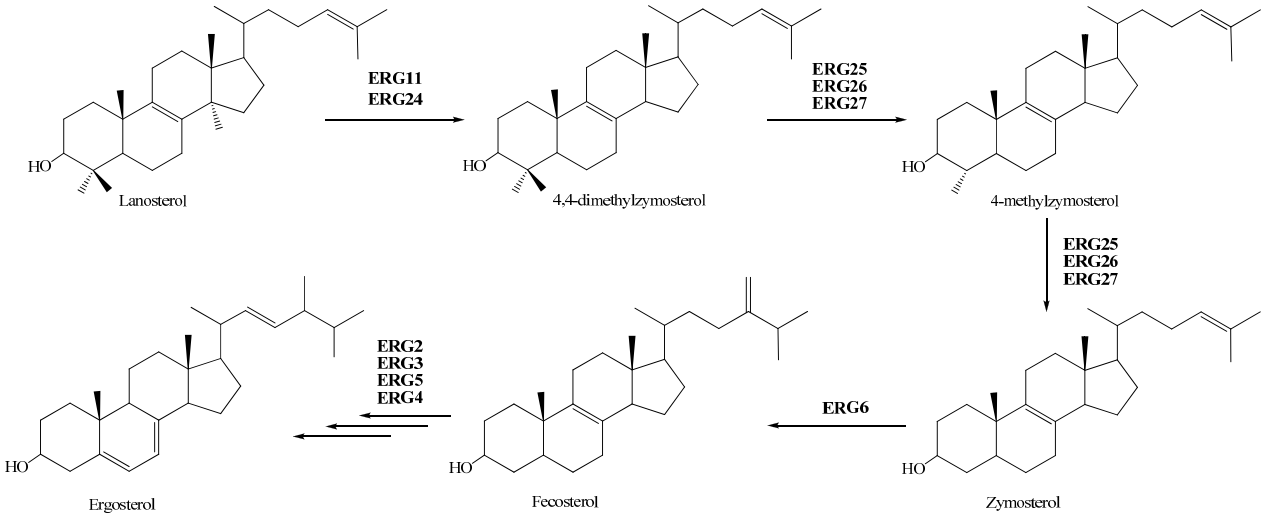


Figure 5.6. Lanosterol–ergosterol biosynthetic pathway in *S. cerevisiae*

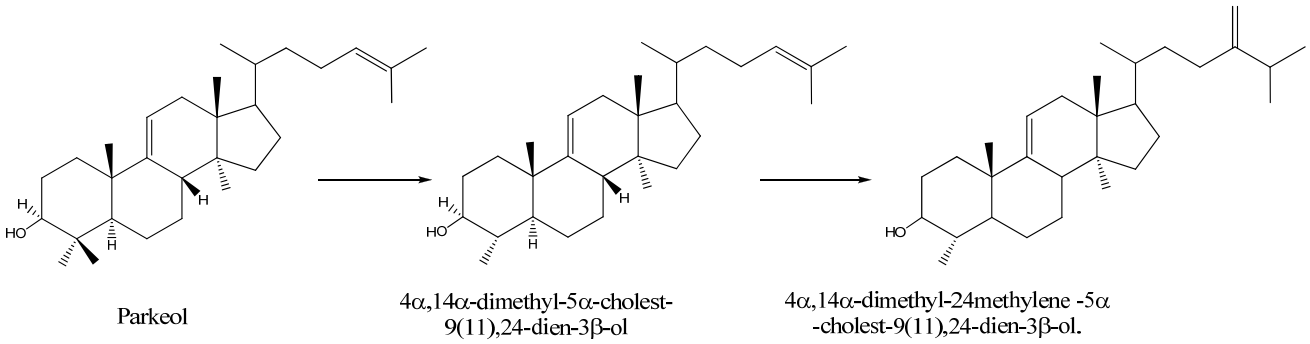


Figure 5.7. Proposed biosynthetic pathways in the *S. cerevisiae* ERG7$^{T384Y/Q450H/V454I}$

The comparison of the respective enzymatic characteristics with previous experimental results provided some clues. In the 14α-demethylase reaction, the C3-hydroxyl group, 14α-methyl group, and $\Delta^{8(9)}$-double bond substitution, as well as the flat substrate structure, are obligatory for the binding, recognition, and catalysis of the demethylase enzyme, which is caused mainly by the hydrophobic contact between the β-surface of the molecule and the residues of the protein.[98] Further, a series of experimental results with various sterol analogues suggested that the structure of the sterol nucleus and orientation of the side chain have important roles in this hydrophobic interaction.[99] It further demonstrated that the parkeol and cycloartenol are not suitable substrates for bacterial 14α-demethylase.[99] Recently, the gene cluster of helvoic acid biosynthesis has been established from *Aspergillus fumigatus*.[100] From the characterization of the structure of helvoic acid, it was apparent that the demethylation at C-4 was followed by the formation of protosta-17(20)*Z*,24,dien-3β-ol, and its C-14 methyl group was unchanged.[100,101] Logically, one would expect to be able to distinguish the proteins observed among different species, but, in this enzyme, the common features are shared with all species of 14α-demethylase.[99] Therefore, it is understandable that a compound could be acquired that still has its C-14 methyl group. In contrast, the C-24 methylase (SMT)

reaction reveals broad substrate specificity. The structural sterol isomers with a $\Delta^{8,(9)}$, Δ^{7}, Δ^{5} nucleus double bond (monoene or diene) and a 9β,19-cyclopropane ring are optimal substrates for the plant or fungal SMT enzyme.[96,102] Additionally, 4α-monomethyl-sterols, 14α-methyl-sterols, and 4α,14α-dimethyl-sterols also were converted to C-24-methylene products.[103] Hence, 4α,14α-dimethyl-5α-cholest-9(11), -24-dien-3β-ol seems to be regarded as the interpretable substrate for SMT to gain the 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol. Fortunately, through product isolation and structural elucidation, the ERG7-mediated reaction and the subsequent triterpene tailoring steps in the *S. cerevisiae* can be illustrated.

CHAPTER 6

Homology modeling coupled with site-directed mutagenesis study of plant oxidosqualene β-amyrin synthase to investigate the relationship between substrate folding geometry and the resulting diverse products

6.1 Research background and aim

Higher plants produce not only the indispensable phytosterol, which acts as a structural component in the cell membrane, but also a wealth of non-steroidal triterpenoids that are stored beneath the surface of the soil in form of glycosides or saponin derivatives and are classified as secondary metabolites. However, the understanding of the biological significance or the biosynthetic regulation of these compounds is quite limited.[104] β-amyrin or other triterpenoids are thought to serve as precursors of oleanolic acid or to associate with saponins, which have been found to have tremendous potential for anti-cancer scope development.[100]

Most of these compounds originated from triterpene synthase that initially cyclizes 2,3-oxidosqualene into the tetracyclic dammarenyl cation via a pre-chair-chair-chair folded substrate conformation. β-amyrin synthase further promotes ring expansions and annulations to form lupenyl cations or oleanyl cations; it also rearranges the skeletal hydride/methyl groups before the deprotonation reaction for producing β-amyrin (Figure 6.1). The conserved catalytic residues in the active site of each triterpene synthase might function as candidates for the responsibility of controlling the complicated reactions and the resulting diverse skeletons.

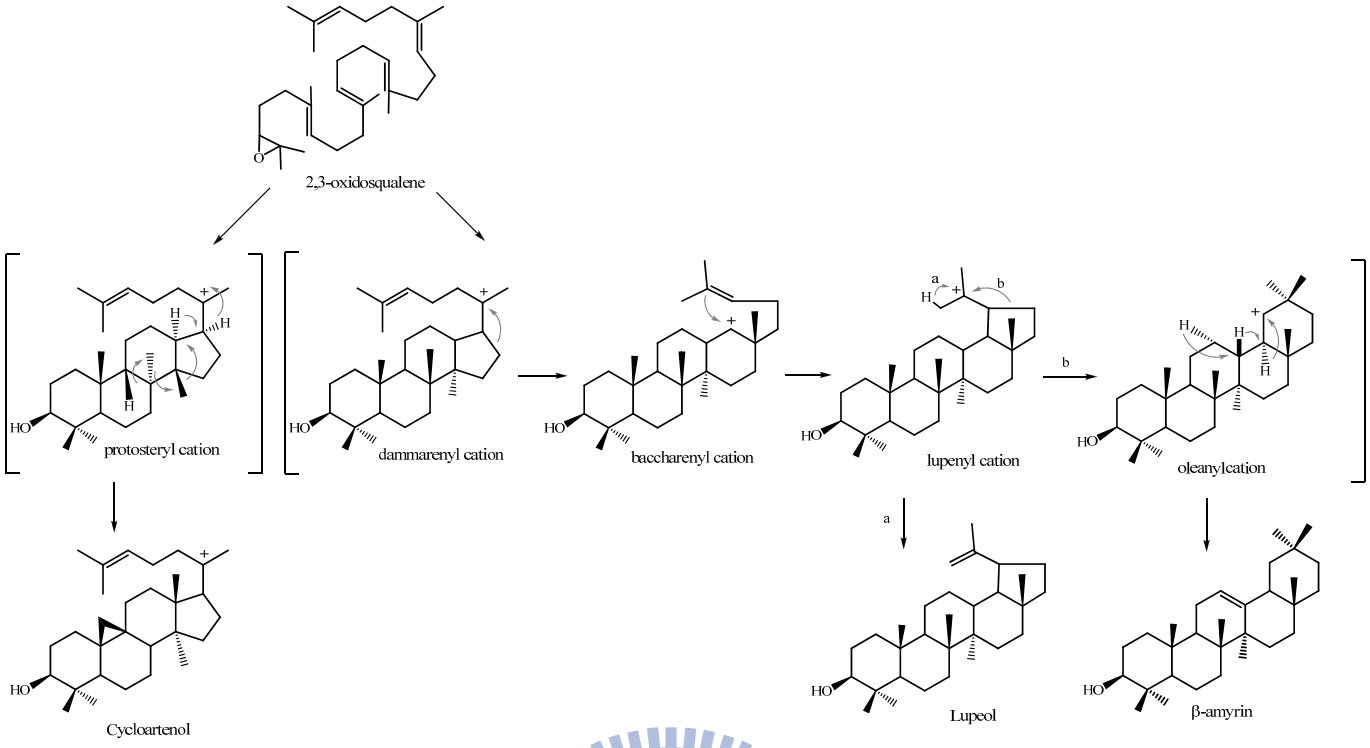


Figure 6.1. Formation of diverse triterpene skeletons from oxidosqualene cyclization in plants

As described previously in Chapter 3, sequence alignments of these two mainly divergent cyclases, which possess opposite substrate folding manners during their B-ring formation, were conducted to identify the residues that determine substrate-folding characteristics. Unfortunately, the selected residues from *S. cerevisiae* ERG7 did not influence substrate folding during B-ring closure.

Because of the different cyclization/rearrangement mechanisms and the sequence diversity among individual cyclase enzymes, insight from the structure is regarded as the critical reference for understanding enzyme-mediated cyclization. Recently, computational prediction has become a potential tool that could be used to study the functions of enzymes, especially in conjunction with various experimental approaches, such as site-directed mutagenesis.[105-107] Thus, various situations were simulated when the target amino acid positions were changed into a different residue with respective

functional groups. For example, an enlarged enzyme active-site cavity caused by introducing a smaller residue could allow the alternative substrate-folding pattern and might result in premature deprotonation.[56,59,93] Also, the decreasing steric bulk of the functional residues allowed irregular rotation of the intermediate, which might influence the accuracy of the expected product. After evaluation of reliable information from different models, it was concluded that the mutagenesis approach could be further applied to verify the expected results from bioinformatic studies. After a careful examination of the significant homology models, a representative example was attained from the proven synergistic effect of two mutations that converted a CAS into an accurate OSC.[108] The example provided the template for understanding the relationships between substrate folding pattern, enzymatic structure modification, and the expected product alternation in these triterpene cyclase families.

Obtaining the β-amyrin synthase gene is a pioneering task to deeply investigate the evolutionary diversity among these structurally-similar cyclase enzymes and to investigate the relationships between the folding geometries of the substrates and the resulting diverse products. Next, we performed the homology modeling construction related to the *S. cerevisiae* ERG7 and *P. sativum* β-amyrin synthase. After the generation of reliable models, the molecular ligands, lanosterol and β-amyrin, were docked into the enzymatic active-site cavity of wild-type *S. cerevisiae* ERG7 and *P. sativum* βAS, respectively, to provide a molecular basis for understanding the interactions between specific residues and ligands. Pairwise structure alignment of *S. cerevisiae* ERG7 and *P. sativum* βAS might also aid the understanding of the influence of structural distinction on substrate folding, cyclization discrepancy, and product variety among cyclase-catalyzed reactions. Finally, site-directed mutagenesis, together with an efficient genetic selection method and product pattern characterization, was conducted to define the role of critical amino acid residues of

the *P. sativum* βAS gene involved in substrate-folding geometry and the complex cyclization/rearrangement cascade. This information is very valuable for revealing the natural evolutionary relationship between sterol and triterpene biosynthesis.

6.2 Results and Discussion

6.2-1 Putative β-amyrin synthase cDNA amplified from *P. sativum*

To date, many triterpene cyclases have been cloned and sequenced [109-113]. From the observation of the sterol and triterpene biosynthetic discrepancy during the development of plant seeds, β-amyrin synthases exhibit a dominant activity just after the period of germination, whereas CAS increase their activity gradually and act as the major functional cyclases after several days of germination[104]. Therefore, the germinated seed of *P. sativum* is a rich source to be used as a template for acquiring the β-amyrin synthase gene. Total RNA of germinated pea seeds was extracted by using Poly(A)Pure™ mRNA Purification kit (Ambion, Inc.) and the reverse transcription was conducted for cDNA isolation and for several trials of the PCR reaction. Finally, the putative ORF with a full length of about 2300 bp was amplified successfully by using specific PCR oligonucleotide primers, three core primers and one pair of terminal primers, i.e., the N-terminus-based and C-terminus-based sequences. [109]

The overall nucleotide sequence of the *P. sativum psy* gene revealed the presence of an ORF with 2277 bp that encoded a 758 amino acid polypeptide with a molecular mass of 87.1 kDa. The theoretical pI of *P. sativum* β-amyrin synthase is 5.96 from the prediction of *protein Identification and analysis Tools* in the Expert Protein Analysis System server (ExPASy server). The deduced amino acid sequences of this putative ORF showed a very small difference from the historically-published β-amyrin

synthase PSY in the National Center for Biotechnology Information's (NCBI's) databank. The difference may result from a mistake in the *Taq* DNA polymerase's fidelity or the utilization of a different source of the pea species. The nucleotide sequence and deduced amino acid sequence of cDNA from *P. sativum* β-amyrin synthase are shown in Figure 6.2.

```
atgtggaggttgaagatagcagaaggtggtaacgatccatatttgttcagcacaaacaac
 M  W  R  L  K  I  A  E  G  G  N  D  P  Y  L  F  S  T  N  N      20
tttgttggaagacagacatgggagtatgatccagaagcaggtagtgaagaagaaagagct
 F  V  G  R  Q  T  W  E  Y  D  P  E  A  G  S  E  E  E  R  A      40
caagttgaagaagctcgtcgcaatttctacaacaaccgcttcgaggttaagccatgtggt
 Q  V  E  E  A  R  R  N  F  Y  N  N  R  F  E  V  K  P  C  G      60
gatctcctatggcgttttcaggttcttagagaaaataacttcaagcaaacaataggcggc
 D  L  L  W  R  F  Q  V  L  R  E  N  N  F  K  Q  T  I  G  G      80
gtaaagatagaagatgaagaagaaataacatatgaaaaaactacaacaacattgagaaga
 V  K  I  E  D  E  E  E  I  T  Y  E  K  T  T  T  T  L  R  R     100
ggcacacatcacttagcagcattgcaaaccagtgatggtcattggccagctcaaattgca
 G  T  H  H  L  A  A  L  Q  T  S  D  G  H  W  P  A  Q  I  A     120
ggtcctctatttttcatgcctcctttggttttctgtgtctacattactggacatcttgat
 G  P  L  F  F  M  P  P  L  V  F  C  V  Y  I  T  G  H  L  D     140
tcggtgttcccaccagagcatcgcaaagagattcttcgttatatatattgccaccagaac
 S  V  F  P  P  E  H  R  K  E  I  L  R  Y  I  Y  C  H  Q  N     160
gaagatggagggtgggggcttcacattgagggtcatagcaccatgttttgtactgcactc
 E  D  G  G  W  G  L  H  I  E  G  H  S  T  M  F  C  T  A  L     180
aactatatatgtatgcggattctcggagaagggcctgatggtggtgaagacaatgcttgc
 N  Y  I  C  M  R  I  L  G  E  G  P  D  G  G  E  D  N  A  C     200
ggtagagcaagaagttggattcgtcaacacggtggtgtcacacatataccttcgtggggg
 G  R  A  R  S  W  I  R  Q  H  G  G  V  T  H  I  P  S  W  G     220
aaaacttggctttcgatacttggtgtgtttgattggttgggaagtaacccaatgccccct
 K  T  W  L  S  I  L  G  V  F  D  W  L  G  S  N  P  M  P  P     240
gaattttggatccttccttcatttcttcctatgcatccagctaaaatgtggtgttattgt
 E  F  W  I  L  P  S  F  L  P  M  H  P  A  K  M  W  C  Y  C     260
cgattggtatacatgcctatgtcttatttgtatggaaaaagatttgtgggtcctatcaca
 R  L  V  Y  M  P  M  S  Y  L  Y  G  K  R  F  V  G  P  I  T     280
ccgctcattttacagttgcgagaagaactccatactgaaccttatgaaaagattaattgg
 P  L  I  L  Q  L  R  E  E  L  H  T  E  P  Y  E  K  I  N  W     300
acgaaaacacgtcatctatgtgcaaaggaagatatttactatcctcatcctttgatacaa
 T  K  T  R  H  L  C  A  K  E  D  I  Y  Y  P  H  P  L  I  Q     330
Gatcttatatgggatagcttgtatattttcaccgagccacttttgactcgttggcctttc
 D  L  I  W  D  S  L  Y  I  F  T  E  P  L  L  T  R  W  P  F     340
aacaaattggtcaggaaaagagctcttgaagtaacaatgaagcatattcactacgaggat
 N  K  L  V  R  K  R  A  L  E  V  T  M  K  H  I  H  Y  E  D     360
gagaatagtcgataccttactattggctgtgtggaaaaggttttatgtatgcttgcttgt
 E  N  S  R  Y  L  T  I  G  C  V  E  K  V  L  C  M  L  A  C     380
```

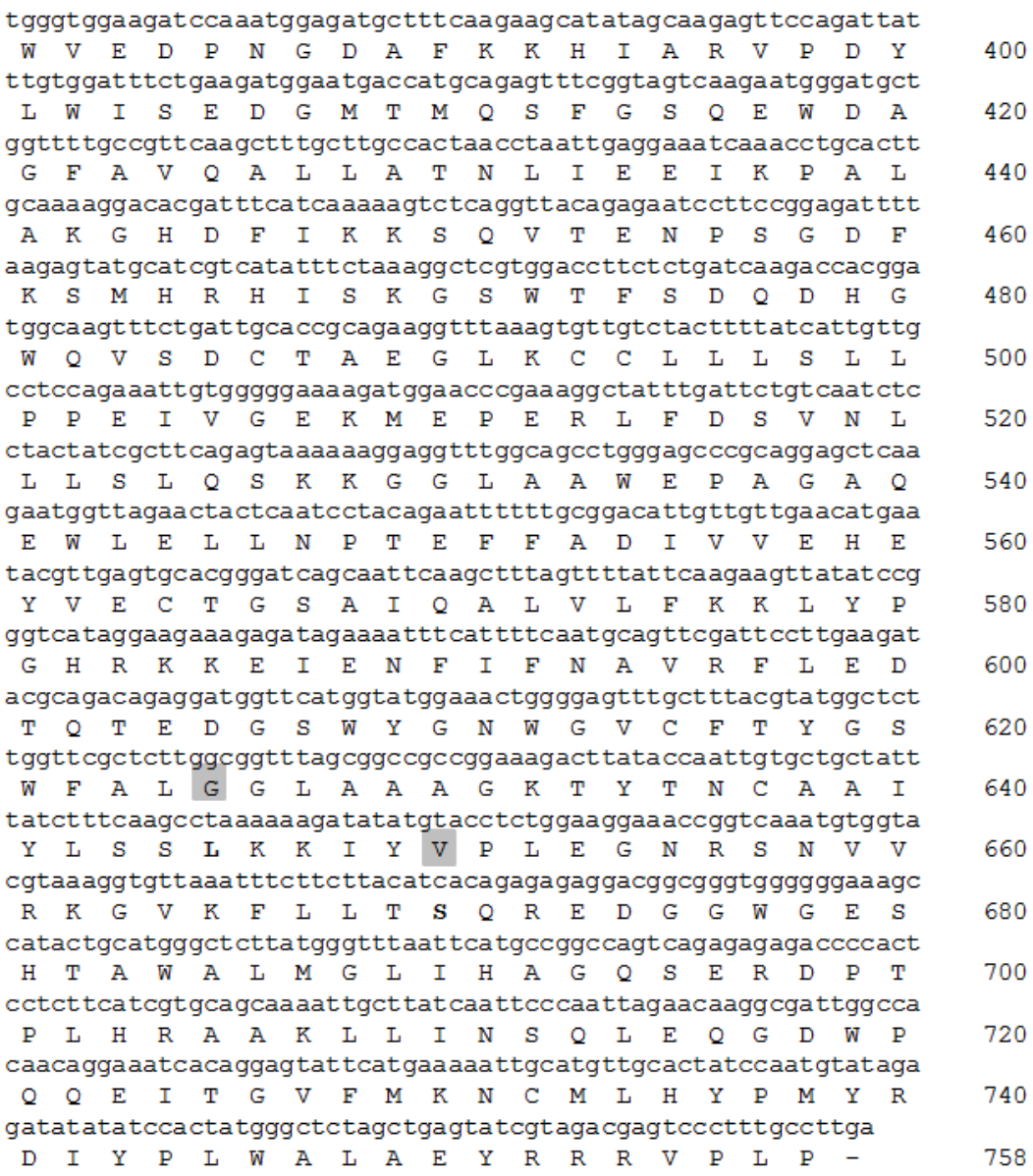


Figure 6.2. Nucleotide sequence and the deduced amino acid sequence of *P. sativum* β-amyrin synthase. The difference between the cloned *P. sativum* β-amyrin synthase gene and the historically published one[109] are labeled with gray block plan.

6.2-2 Heterologous expression of β-amyrin synthase in yeast

In order to elucidate the enzyme function of *P. sativum* β-amyrin synthase, the functional expression of this gene was conducted by using a yeast expression system. The PCR products were digested by *EcoR* I and *Xho* I and subcloned into the corresponding restriction sites of the pYES2 vector (Invitrogen), a yeast expression

vector with a *GAL1* promoter. The gene construction was named as pYES2-PSY. Subsequently, the pYES2-PSY was transformed into an *erg7*-deficient yeast strain, TKW14c2, for the protein expression. After the induction of galactose, the harvested cells were subjected to alkaline treatment. The NSL from the cells was extracted with petroleum ether, and the cyclized products were analyzed by GC-MS. As shown in Figure 5.3, pYES2-PSY produced β-amyrin as a sole product, and the mass fragments of this peak were identical to those of an authentic β-amyrin, conclusively proving that the *psy* gene encodes a *P. sativum* β-amyrin synthase.

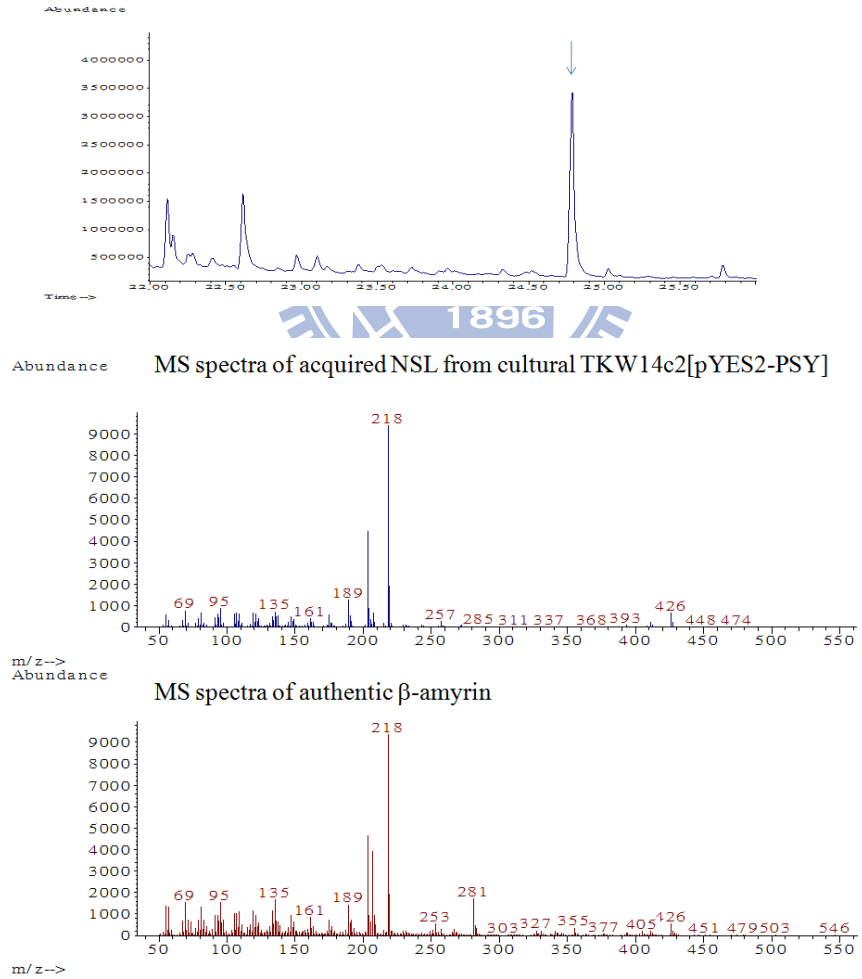


Figure 6.3. GC-MS analysis of the NSL isolated from the petroleum ether extraction of pYES2-PSY in a yeast strain, TKW14c2. The dominated product is identical to an authentic β-amyrin.

6.2-3 Study of enzyme activity using homology modeling coupled with the site-directed mutagenesis approach

Because of the parallel reaction mechanism and the high sequence similarity, the comparable three-dimensional structure would be expected among the triterpene synthases. Based on this hypothesis, the X-ray crystal structure of *H. sapiens* OSC (PDB entry 1W6K) was used as the primary template to generate the overall skeletons of *S. cerevisiae* ERG7 and *P. sativum* β-amyrin synthase, respectively. Then, to simulate the spatial relationship between the catalytically-critical amino acid residues and the reactants in the enzyme-action environment, the molecular-docking program was performed. Therefore, lanosterol and β-amyrin were used as ligands and docked into the homology models of *S. cerevisiae* ERG7 and *P. sativum* βAS, respectively. The preliminary superimposition of the *S. cerevisiae* ERG7 and *P. sativum* βAS homology model with either lanosterol or β-amyrin was performed by using Pymol software (Figure 6.3). Comparing these two virtual protein-modeling structures, it was apparent that the overall skeletons shared a moderate similarity, which indicated that the conformity of the general structure was conserved among these mechanistically-similar cyclases. Finally, the ligand-protein complexes were applied further into the energy minimization calculation by using Sybyl 7.0 software. By calculating geometrical distances, several residues were identified that were located at distances within 6 Å of the lanosterol molecule in the *S. cerevisiae* ERG7 and of the β-amyrin molecule in the *P. sativum* βAS (Figure 6.3, Table 6.1).

Via a careful elucidation of enzyme-substrate interactions and examination of the changes in the spatial environment of the protein upon residue substitution, nineteen residues in the β-amyrin synthase that were distinct from those of yeast ERG7 were replaced with the corresponding residue of yeast ERG7 by insight II software (Table 6.1). After computational execution, the superimposed “modified β-amyrin synthase”

was compared with native *S. cerevisiae* ERG7. The overall three-dimensional structure between these two enzymes showed a high similarity in their enzymatic cavities, which is an interesting result. Even though the parent β-amyrin synthase exhibited a total difference with that of yeast ERG7, both in overall appearance of the protein and in the mechanistic divergence for diverse product profiles, changing a small number of amino acid substitutions on β-amyrin synthase allowed β-amyrin synthase to share the similar scaffold in its active-site cavity (Figure 6.4).

In order to examine the catalytic activity of this modified β-amyrin synthase, the site-directed mutagenesis approach was used to alter 19 critical residues, i.e., Ile-119, Phe-125, Cys-258, Tyr-259, Cys-260, Leu-262, Gly-369, Cys-370, Glu-411, Ser-412, Phe-413, Ser-415, Ala-533, Trp-534, Thr-549, Asn-611, Met-729, Met-733, and Leu-734, from β-amyrin synthase into the corresponding residues of yeast ERG7. Consequently, an artificial enzyme was designed that possessed the external structural scaffold of β-amyrin synthase and the interior yeast ERG7 cavity (Figure 6.4). In addition, to observe the local mutational effect on the variation of product profiles, the neighboring amino acid residues of the active-site cavity were classified further into nine groups in the mutagenesis experiments. The detailed classification is shown in Table 6.2.

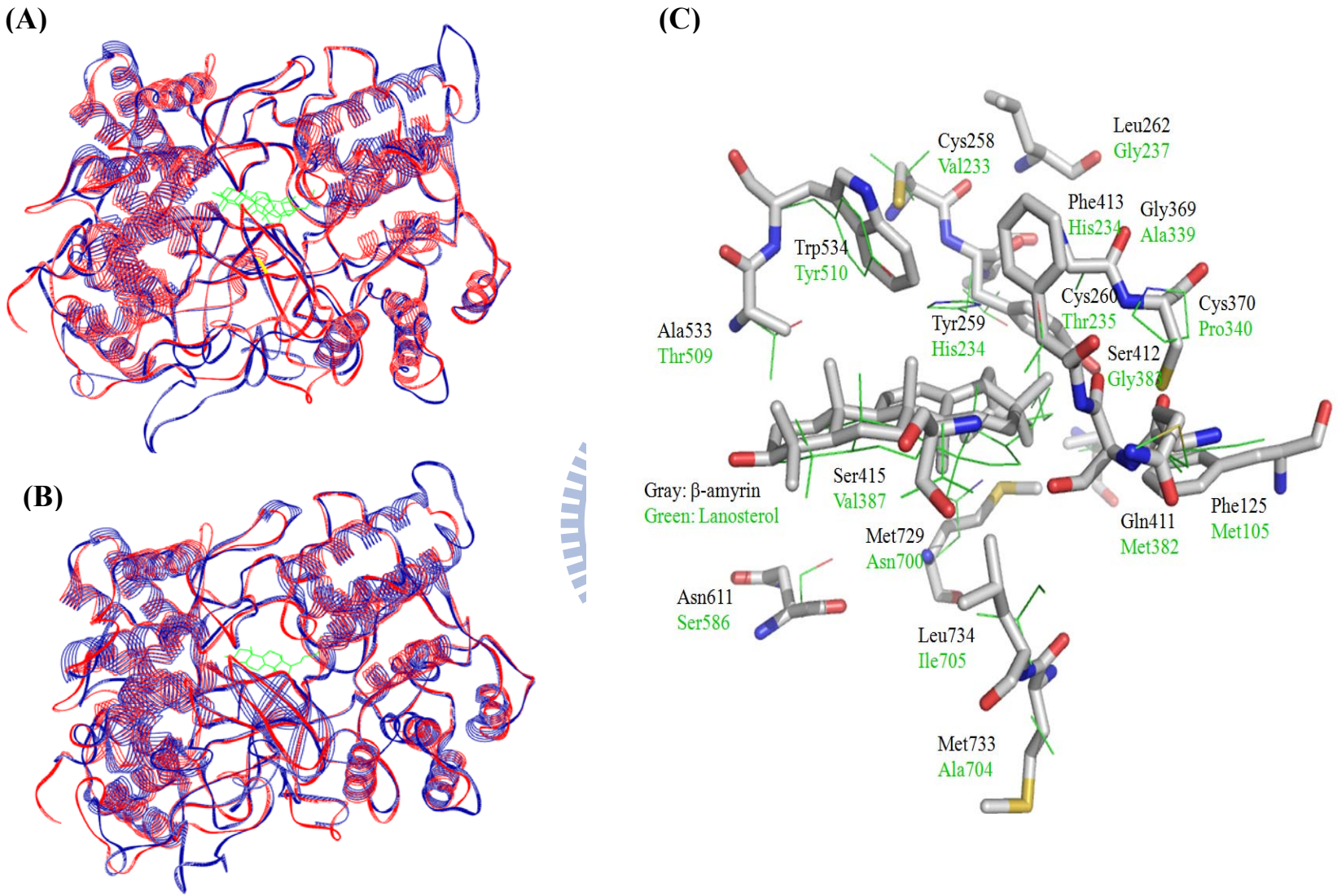


Figure 6.4. (A) Superimposition of the *S. cerevisiae ERG7* homology modeling structure (red) with *P. sativum* βAS modeling structure (blue), and the docked product, lanosterol, β-amyrin (green), are included (B) Superimposition of the *S. cerevisiae* ERG7 homology modeling structure (red) with *P. sativum* βAS mutants modeling structure (blue), and the lanosterol, is shown in green. (C)The stereo representations of *S. cerevisiae ERG7* (Green line) and *P. sativum* βAS (Gray stick), the distances between residues and the molecule less than six 6 Å are shown.

Table 6.1. Relative distances between active site residues with the docked substrate, lanosterol, in β-amyrin synthase. The residues that are gray in color were changed with the corresponding residues of yeast ERG7. The parentheses show the residue numbers of *S. cerevisiae* ERG7.

Distance	> 6 Å	> 5 Å	> 4 Å	> 3 Å	> 2.5 Å
Residues	Pro122 (Pro102)	**Ile119 (Tyr99)**	**Phe125 (Met105)**	Trp219 (Trp194)	Glu563 (Glu539)
	Cys258(Val233)	Phe124 (Phe104)	Trp257 (Trp232)	**Tyr259 (His234)**	
	Arg261 (Arg236)	Ile368 (Ile338)	**Cys260 (Thr235)**	Trp418 (Trp390)	
	Leu262 (Gly237)	**Gly369 (Ala339)**	Val263 (Val238)	Phe474 (Phe445)	
	Tyr264 (Tyr239)	Phe413 (Thr384)	**Cys370 (Pro340)**	Asp485 (Asp456)	
	Gln411(Met382)	Phe552 (Phe528)	**Ser412 (Gly383)**	**Trp534 (Tyr510)**	
	Phe413(Thr384)	Val727 (Val698)	Val483 (Val454)	Trp612 (Trp587)	
	Gly414 (Gly386)	Tyr739 (Tyr710)	Cys486 (Cys457)	Phe728 (Phe699)	
	Ser415(Val387)		**Ala533 (Thr509)**	**Leu734 (Ile705)**	
	Phe422 (Phe394)		Ile555 (Ile531)		
	Ala532 (Ala508)		Cys564 (Cys540)		
	Thr549 (Ala525)		Tyr618 (Tyr593)		
	Asn611 (Ser586)		**Met729 (Asn700)**		
	Gly613 (Gly588)		Tyr736 (Tyr707)		
	Met733 (Ala704)				

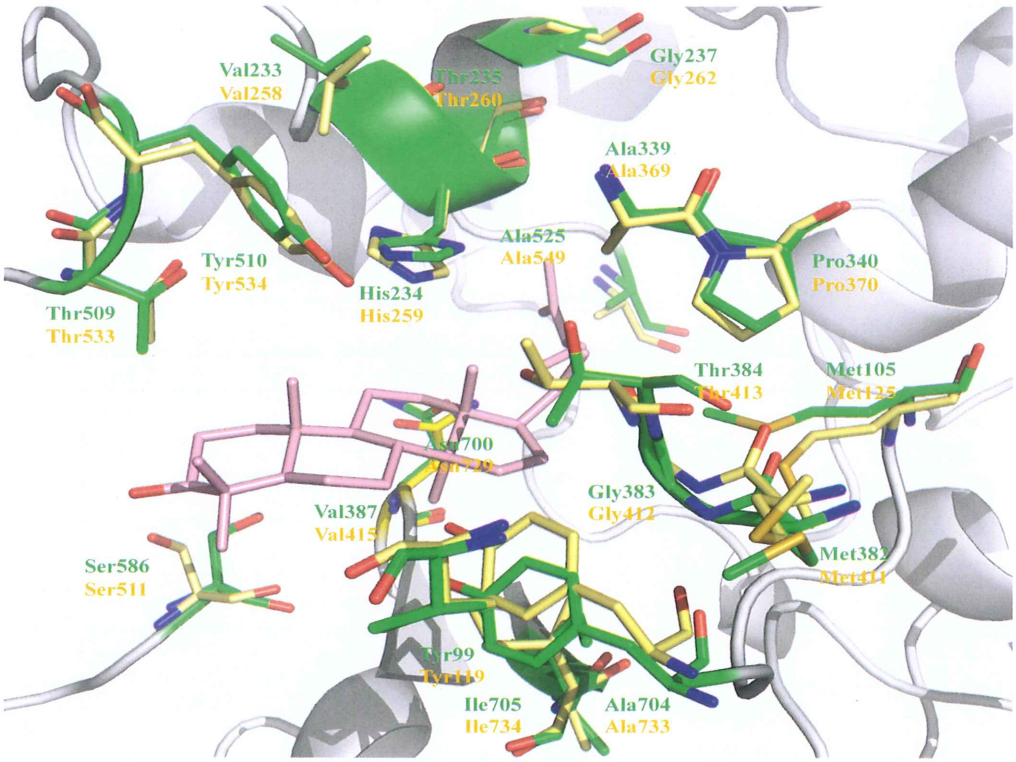


Figure 6.5. After amino acid substitutions by computational simulation, the superimposed functional orientation of the residues of "modified β-amyrin synthase" is similar to the native *S. cerevisiae* ERG7 homology modeling structure. Green shows the native *S. cerevisiae* ERG7, yellow shows the "modified β-amyrin synthase."

6.2-4 Functional analysis of artificial enzymes via the plasmid shuffle methodology and product characterization

After the construction of a mutated enzymes library, these plasmids were transformed into an *erg7*-disrupted strain, TKW14c2. The TKW14c2 strain is a suitable system for selection and analysis of β-amyrin synthase variants, which possess the ability to produce lanosterol (with a chair-boat-chair tetracyclic skeleton) and other significant compounds. The ergosterol complementation assay revealed that all of the β-amyrin synthase variants failed to generate ergosterol for yeast viability, suggesting that these mutated β-amyrin synthases have no capability of forming lanosterol. Thus, the suppression of ERG7 activity or the alteration of the product profile could be imagined in the yeast transformants with the β-amyrin synthase variants. The resulting NSL

extracts of individual mutant were thus analyzed.

After the elimination of the abundant background influence, the purified fraction with R_f value between oxidosqualene and ergosterol was analyzed by GC-MS spectra. The non-discriminating product profile, as that of the negative control strain, TKW14c2, was observed in a β-amyrin synthase mutant, "Y-all," which contained all 19 residue substitutions. This result implied that the dramatic exchange of enzymatic active site might disrupt the enzymatic activity either directly, caused by the distortion of the structure of the active site, or indirectly, from the disruption of the structurally-relative helical domain. Hence, the catalytic ability of the oxidosqualene cyclization and the production of lanosterol might be lost.

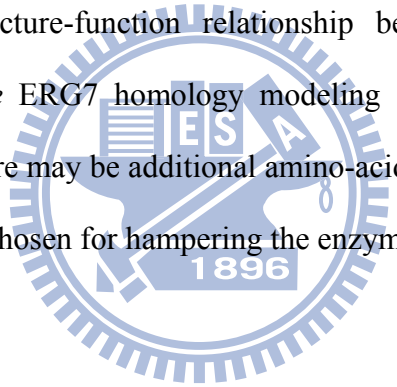
In addition, the product pattern of a series of β-amyrin synthase mutants, "Y1-Y-11," also presented the same product pattern as that of TKW14c2, except for "Y-10" (Table 6.2), which was not surprising. All of the selected residues were located in the active-site cavity of the β-amyrin synthase. These functional residues might particularly assist the substrate binding, stabilize the respective carbocationic intermediates in critical steps, or participate in the coordinative interaction with other residues for the cyclization/rearrangement cascade. Thus, substitutions at these residues might interrupt the enzymatic behavior dramatically. In addition, the homology modeling structure of β-amyrin synthase might provide some clues for explaining why "Y-10," representing an Asn611Ser mutant, possesses the wild-type enzymatic function. The Asn-611 is located below the substrate plain and at a distance of 6 Å from the substrate, suggesting that Asn-611 might possibly not be one of the important amino-acid residues involved in the overall cyclization/rearrangement cascade (Figure 6.4, Table 6.1).

Table 6.2. Product profile of *P. sativum* βAS mutants

Abbreviate	Mutation in β-amyrin synthase	Ergosterol complementary assay	Product profile
Y-all	Ile119Tyr, Phe125Met, Cys258Val, Tyr259His, Cys260Thr, Leu262Gly, Gly369Ala, Cys370Pro, Glu411Met, Ser412Gly, Phe413Thr, Ser415Val, Ala533tyr, Trp534Thr, Thr549Ala, Asn611ser, Met729Asn, Met733Ala and Leu734	–	No product
Y-1	Ile119Tyr	–	No product
Y-4	Phe125Met	–	No product
Y-5	C258V, Tyr259His, Cys260Thr, Leu262Gly,	–	No product
Y-6	Gly369Ala, Cys370Pro,	–	No product
Y-7	Glu411Met, Ser412Gly, Phe413Thr, Ser415Val	–	No product
Y-8	Ala533Tyr, Trp534Thr	–	No product
Y-9	Thr549Ala	–	No product
Y-10	Asn611Ser	–	β-amyrin
Y-11	Met729Asn, Met733Ala, Leu734	–	No product

Modification of the interior cavity of β-amyrin synthase toward the OSC active site seems a feasible strategy to acquire the desired protein function. Unfortunately, the results were unexpected, and non-distinctive product profiles were observed. Comparison with previous experimental results suggested that these two groups of enzymes (sterol-producing cyclase and triterpene alcohol-producing cyclase) should be studied using different mutagenesis methods. D-ring closured formation, the skeletal rearrangement, and the ring expansion are all conserved in sterol-producing cyclase and in triterpene-producing cyclase; the only exception is the different substrate pre-folding matter. However, comparing the primary amino acids for sterol-producing cyclase and

triterpene-producing cyclase produced a distinct result. The highly conserved pattern was observed among the triterpene-producing cyclases, whereas the sterol-producing cyclases showed a divergent conservation. This illustrates that changing the position of a single amino acid or changing the neighbor region of the triterpene cyclases is not enough to give them the synthesis ability for sterol compounds formation. In contrast, sterol-producing cyclases or other related triterpene-producing cyclases might exhibit some interesting puzzles. In the previously-mentioned results, only single-point mutations have examined their altered product specificity from original enzyme into a related homologous enzyme. In addition, recent site-saturated mutagenesis and its product characterization for various *S. cerevisiae* ERG7 mutants revealed a slight contradiction on the structure-function relationship between human OSC crystal structure and *S. cerevisiae* ERG7 homology modeling structure. This could be one reason to speculate that there may be additional amino-acid residues and other important functional residues not be chosen for hampering the enzymatic activity.



6.3 Conclusion

In addition to the above-described results of site-directed mutagenesis of β-amyrin synthase, the alanine-scanning mutagenesis approach coupled with homology modeling illustration was also used in our laboratory for the putative active-site residues of *P. sativum* βAS. However, the mutational effects of these putatively, functional residues presented a non-discriminable product profile, except for a few mutations with the generation of a sole wild-type product, β-amyrin, just like the result reported in this chapter. It is an illegitimate result. According to the result in Chapters 3-5, slight mutation modification can readily alter the catalytic properties and broaden the product

variation, especially for the OSCs. Thus, careful scrutiny required the assumption that the *E. coli*-yeast shuttle vector, pRS314, may not be a suitable agent for heterologous expression, due to its expression level. Thus, displacement of plasmid with high copy number one or insertion of the expression promoter fragment into the existing vector may provide the possibility to observe the real *in vivo* mutation phenomenon.

CHAPTER 7

Future perspective

In spite of the significant progress that has been made, many mysteries of this ingenious biotransformation from oxidosqualene to lanosterol still remain to be resolved. Three possible directions could be considered in the future to illuminate the subtle regulation of cyclase-mediate reaction:

I. Mechanism of OSC-catalyzed cyclization

Site-saturated mutagenesis provides a valuable tool for understanding how the functional residues participated in the complex cyclization reaction. Fortunately, many postulated mechanistic products were obtained that directly verify the hypothetically-proposed intermediate states. In order to further investigate the more detailed role of other functional residues in the complex OSC-mediated reaction, a systematic recombinant mutagenesis at putative active-site residues in either first or second-sphere might be further explored to establish definitively the relationship between the residue itself and the substrate. This would provide the opportunity to identify additional trapped products, and it might create a novel enzyme that has new catalytic properties.

II. Structural biology

The first crystal structure of human OSC has improved our understanding of the molecular interaction between lanosterol and the enzymatic active-site cavity. A more in-depth and detailed understanding of the role of functional residues that assist the binding of substrates, the manner in which substrates are folded, the stabilization of cationic intermediates, and the expansion of the D/E ring requires the determination of additional details about the structure of triterpene cyclase. In addition, recent

mutagenesis and analysis of resulting product profiles from various *S. cerevisiae* ERG7 mutants revealed a contradictory situation between human OSC crystal structure and the *S. cerevisiae* ERG7 homology model. Hence, three-dimensional structures of these mutated cyclases are necessary. The *pichia pastoris* expression system, which successfully elucidated the crystal structure of human OSC, might be a suitable candidate system for expression and crystallization of different triterpene cyclase enzymes.[49] Moreover, the substrate, the transition-state analogs, or the isolated abnormally-cyclized or altered deprotonation products might be in complex with the protein crystal to detailed understand the molecular interaction between enzymatic active site and the various products.

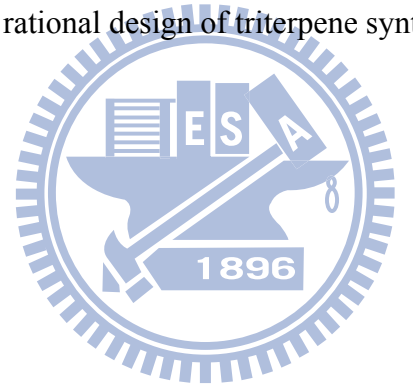
III. Determining the relationship between product distribution and the percentages associated with each product by an *in vitro* protein expression assay

Due to the intrinsic difficulty of the purification of active OSC from native tissue and the limitation of an unsuccessful expression system to acquire the active protein, the alternative modern system is considered to act as a powerful tool for providing valuable insight for an in-depth understanding of the enzymatic catalytic mechanism. Through experiments with amino acid substitution and the analysis of the resulting products, the functional role of the residue in the cyclase-mediated reaction was explained. Notably, many altered products were isolated and characterized to validate the proposed hypothetical mechanism. However, the amounts of the resulting products should be measured as accurately as possible using a novel *in vitro* protein expression assay. By adding the substrate to a reaction that used the *erg1 erg7 erg11* triple knockout yeast strain as the host cell and bore an interest in mutated plasmids, the resulting products will be abidingly acuminated because the downstream biosynthetic pathway was arrested in this *erg1 erg7 erg11* triple-knockout yeast strain. Therefore, more accurate product ratios must be obtained to reflect the true enzymatic dynamics of the cyclase

reaction. In addition, a more precise understanding of triterpene cyclase might provide information that we need to design a novel synthase for the generation of a desired compound without other accompanying by-products or some mutated enzymes that produce unnatural, but worthwhile, products.

IV. Molecular evolution of oxidosqualene cyclases for protein plasticity

Random mutagenesis, coupled with a genetic selection in an *erg7*-deficient strain, should be the convenient methodology to promote the triterpene-producing synthases gene that possesses the lanosterol-synthesis ability. Through better understanding of the critical role of individual amino-acid residues involved in each step of the chemical biotransformation, artificial cyclases that can generate desired compounds will be created in the future by the rational design of triterpene synthases.



CHAPTER 8

Experimental Section

8.1 Materials

8.1-1 Bacterial strains and molecular cloning/expression vectors

Except stated elsewhere, bacterial growth condition, agar plate preparation, recombinant DNA purification, DNA sequence determination, and agarose gel electrophoresis were performed according to standard procedures or the commercial kit.[114] Phagemid pBluescript IISK(+) and *Escherichia coli* XL1-Blue were purchased from Stratagene. Plasmid vector pRS314 was kindly provided by Dr. Botstein (Stanford University). The *ERG7* gene including the promoter sequence was cloned into pRS314 vector to create the pRS314ERG7WT plasmid. In addition, in order to facilitate the expression of extrogeneous cyclase in yeast, the promoter sequence from *ERG7* gene was fused in front of the recombinant *CAS1* gene to create the pTKP5 plasmid.[1] The haploid yeast strain CBY57[pZS11] (*ERG7*Δ*::LEU2 ade2*-101 *lys*2-801 *his3*-Δ200 *leu2*-Δ1 *trp1*-Δ63 *ura3*-52 [pZS11]) was derived from the diploid strain CBY1 (MATa/α *ERG7*Δ*::LEU2 ade2*-101 *his3*-Δ200 *lys2*-801 *trp1*-Δ63 *ura3*-52).[2]

8.1-2 Enzymes, chemicals, equipments, and reagents

(1) Enzymes: All restrictive endonucleases and modifying enzymes were purchased from New England BioLabs lnc. The pfu DNA polymerase was purchased from Stratagene. All enzymes were used according to the recommended protocol.

(2) Chemicals: The following section lists the chemicals utilized in this thesis. The manufacturers were included in the square bracket. Acetic acid [Merck], Acetic

anhydride [Sigma], Acetone [Merck], Adenine [Sigma], Agarose-LE [USB], Bacto™ Agar [DIFCO], Deoxyribonucleotide triphosphate (dNTP) 100mM Solutions [GE Healthcare], DNA 10kb Ladder [BioBasic Inc.], D-Sorbitol [Sigma], Ethanol (95% and 99%) [Merck], Ether [Merck], Ethyl acetate [Merck], [Merck], Ergosterol [Sigma], Geneticine (G418) [Gibco], Glucose [Sigma], Glycerol [Merck], Hemin chloride [Merck], Hexane [Merck], Histidine [Sigma], Lanosterol [Sigma], LB Broth (Miller) [DIFCO], Lysine [Sigma], Methanol [Merck], Methioine [Sigma], Oligonucletide Primers [BioBasic Inc.], Pyridine [Sigma], Pyrogallol [Merck], Sea sand [Merck], Silica gel [Merck], Silver nitrate [Merck], Sodium sulfate [Merck], SYBR® Green I [Roche], Squalene 99% [ACROS], TLC plate [Mecrk], Triton X-100 [Sigma], Tryptone [DIFCO], Tryptophan [Sigma], Tween 80 [Merck], Yeast Extra [DIFCO], Yeast Nitrogen Base w/o amino acid [DIFCO], Uracil [Sigma]

(3) Equipments: The following instruments were used in this thesis: ABI PRISM® 3100 Genetic Analyzer [Applied Biosystems], GeneAmp® PCR System 9700 Thermal Cycler [Applied Biosystems], Kodak Electrophoresis Documentation an Analysis System 120 [Kodak], Orbital Shaking Incubator Model-S302R [Firstek Scintific], Pulse Controller [BioRad], Rotary Vacuum Evaporator N-N Series [EYELA], Steritop™ 0.22μm Filter Unit [Millipore], 32 Karat™ System Gold® HPLC [Beckman Coulter], Agilent 6890N gas-chromatography equipped with a DB-5 column [Agilent], Agilent Technologies model 5973 MSD mass spectrometer [Agilent]

(4) Reagents

50 X ALTHMU solution:

0.2% Adenine, 0.3% Lysine, 0.2% Tryptophan, 0.2% Histidine, 0.2% Methonine, and 0.2% Uracil in ddH_2O. Sterilize and store it at 4°C.

50 X ALHMU solution:

0.2% Adenine, 0.3% Lysine, 0.2% Histidine, 0.2% Methonine, and 0.2% Uracil in

ddH_2O. Sterilize and store it at 4°C.

ALHMU/Heme/Ergosterol plate:

0.67 g Yeast Nitrogen Base and 2 g Bacto™ Agar was dissolved in 100 mL ddH_2O, and sterilized it. Add 2 mL 50 X ALHMU solutions, 4 mL 50% Glucose solution, 2 mL Heme solution, 2 mL Ergosterol supplement solution, and 100 μL G148 stock solution into the above sterile medium. The mixture was then poured and dispersed in the Petri dishes before it coagulated. All of steps are under aseptic condition. Store it at 4°C.

Ampicillin stock solution (100mg/mL):

Dissolve 1 g Ampicillin sulfate in 10 mL ddH_2O. Filter it through a 0.22 μm pore size filter and stock it at -20°C.

6 X DNA loading dye:

0.25% Bromophenol blue and 30% glycerol in ddH_2O.

20% EA developing solution:

Add 20 mL Ethyl acetate in 80 mL Hexane, and mix it.

Ergosterols supplement solution:

1 g Ergosterol was dissolved in 250 mL 95% ethanol. Mix it with 250 mL Tween 80 under the aseptic condition. Store it at 4°C under darkness.

G418 stock solution (1 g/mL):

Dissolve 500 mg G418 in 500 μL sterile ddH_2O. Store it at 4°C under darkness.

50% Glucose solution:

500 g Glucose was dissolved in 1 L ddH_2O, and sterilized it.

80% Glycerol solution:

80 mL Glycerol was added in 20 mL ddH_2O and sterilized. Store it at 4°C.

Heme solution:

0.5 g Hemin chloride was dissolved in 250 mL 0.2 N Potassium hydroxide solution. Mix it with 250 mL 95% ethanol in aseptic condition. Store it under darkness at 4°C.

LB medium:

25 g LB Broth was dissolved in 1 L ddH_2O and sterilized.

LB plate:

25 g LB Broth and 20 g Bacto™ Agar was dissolved in 1 L ddH_2O and sterilized. The sterile solution was poured and dispersed in Petri dishes before coagulation.

1 M Sorbitol solution:

182.2 g D-sorbitol was dissolved in 500 mL ddH_2O and sterilized. Store it at 4°C.

10 X SYBR Green solution:

10,000 X SYBR® Green I was diluted to 10 X with DMSO. Store it under darkness.

50 X TAE buffer:

Add 242 g Tris-base, 57.1 mL Acetic acid, and 0.5 M EDTA into 800 mL ddH_2O. Adjust the total volume into 1 L and pH value into pH 8.5. Store it at room temperature. Dilute the concentration to 1 X with ddH_2O. Adjust the pH value to 7.5~7.8 before utlization.

TLC staining solution:

40 mL concentrated H_2SO_4 was slowly added into 200 mL ethanol followed by a addtion of 12 mL Acetic acid and 16 mL Anisaldehyde. The solution was finally added with another 600 mL Enthanol with careful and slight agitation.

8.2 General experimental procedures

8.2-1 Construction of site-directed/saturated mutagenic plasmids.

For the construction of different alanine-scanning or site-directed/saturated mutagenic plasmids, pRS314ERG7WT plasmid was used as the template for the PCR reaction. Mutated plasmids were constructed by using the QuikChang Site-Directed Mutagenesis kit (Strategene, Inc., La Jolla, CA) according to the manufacturer's protocols. After PCR amplification with designed primers, recombinant plasmids were subjected to *Dpn* I

digestion to remove the parent plasmid, transforming them into XL1-Blue cells for DNA purification. The exact substitutions at corresponding position were determined using an ABI PRISM 3100 auto-sequencer.

Table 8.1. Primer sequences in the site-directed mutagenesis experiment

Oligonucleotide primers	**Sequences**
YTL-M103L1	5’- ggACCCCTCTTCATgACgATCggTTAC-3’
YTL-M103L2	5’- gTAACCgATCgTCATgAACAggggTCC-3’
YTL-G108P1	5’- AAAgggCCCATgTTCATgACAATCCCTTACgTA-3’
YTL-G108P2	5’- TACGTAAgggATTgGTCATgAACATgggCCCTTT-3’
YTL-H234Y1	5’- TgggTTTATACTCgAgCTGTTTACATT-3’
YTL-H234Y2	5’- AATgTAAACACCTCgAgTATAAACCCA-3’
YTL-I240M1	5’- CATACTCGAggTgTTTACATgCCggTC-3’
YTL-I240M2	5’- gACCggCATgTAAACACCTCgAgTATg-3’
YTL-T333S1	5’- ACggAgCTgCAgAATAgTgATTCC-3’
YTL-T333S2	5’- ggAATCACTATTCTgCAgCTCCgT-3’
YTL-P340C1	5’- ATAgCATgTgTTAACCAAgCTTTTTgC-3’
YTL-P340C2	5’-gCAAAAAgCTTggTTAACACATgCTAT-3’
YTL-G444T1	5’- AgAAAgggTgCCTggACCTTCTCAACA-3’
YTL-G444T2	5’- TgTTgAgAAggTCCAggCACCCTTTCT-3’
YTL-A525T 1	5’- CTAgCCATggAAACCTTgAATCCTACTgAAgTT-3’
YTL-A525T 2	5’- AACTTCAgTAggATTCAAggTTTCCATggCTAg-3’
YTL-I705L1	5’- TTCAACCAgAgCTgTgCACTTgAATAC-3’
YTL-I705L2	5’- gTATTCAAgTgCACAgCTgTggTTgAA-3’
YTL-PSY119Y-ApaI1	5’-gCTCAATATgCAgggCCCCCTATTTTTCATgCC-3’
YTL-PSY119Y-ApaI2	5’-ggCATGAAAAATAggggCCCTgCATATTgAgC-3’
YTL-PSYF125M-StyI1	5’-CCTCTATTTATgATgCCTCCCTTggTTTTCTg-3
YTL-PSYF125M-StyI2	5’-CAgAAAACCAAgggAggCATCATAAATAgAgg-3’

YTL-PSYC258V,Y259H,C260T,L262G-XhoI1	5’-CTAAAATgTgggTTCACACTCgAgggggTATACATgC-3’
YTL-PSYC258V,Y259H,C260T,L262G-XhoI2	5’-gCATgTATAVVVVTVgAgTgTgAACCCACATTTTAg-3’
YTL-PSYG369A, C370P-KpnI1	5’-gAATAgTCggTACCTTACTATTgCCCCTgTggAAAAgg-3’
YTL-PSYG369A, C370P-KpnI2	5’-CCTTTTCCACAggggCAATAgTAAggTACCgACTATTC-3’
YTL-PSYQ411M,S412G, F413T, S415V-MslI1	5’-gAAgATggCATgACCATgATgggTACCggTgTTCAAgAATgg-3’
YTL-PSYQ411M,S412G, F413T, S415V-MslI2	5’- CCATTCTTgAACACCggTACCCATCATggTCATgCCATCTTC-3’
YTL-PSYA533Y, W534T-PvuII1	5’-gAggTTTggCAACCTACgAgCCagCTggAgCTCAAg-3’
YTL-PSYA533Y, W534T-PvuII2	5’- CTTgAgCTCCAgCTggCTCgTAggTTgCCAAACCTC-3’
YTL-PSYT549A-EcoRI1	5’-CAATCCTgCAgAATTCTTTgCggACATTgTTg-3’
YTL-PSYT549A-EcoRI2	5’- CAACAATgTCCgCAAAgAATTCTgCAggATTg-3’
YTL-PSYN611S-PvuII1	5’-ggATggCAgCTggTATggAAgCTggggAg-3’
YTL-PSYN611S-PvuII2	5’- CTCCCCAgCTTCCATACCAgCTgCCATCC-3’
YTL-PSYM729N, M733A, L734I-dNSPI1	5’-gAgTATTCAACAAAAATTgCgCgATCCACTATCCAATg-3’
YTL-PSYM729N, M733A, L734I-dNSPI1	5’-CATTggATAgTggATCgCgCAATTTTTgTTgAATACTC-3’
YOSCT384Y-1	5’-gACCATTATgggATACAATggTgtgCAAACC-3’
YOSCT384Y-2	5’-ggTTTgCACACCATTgTATccCATAATggTC-3’
YOSCQ450H-ScaI1	5’-gggCTTCAgTACTAAAACACACggCTATACAATAgC-3
YOSCQ450H-ScaI2	5’-gCTATTgTATAgCCgTgTgTTTTAgTACTgAAgCCC-3’
YOSCV454I-PvuII1	5’-CTATACAATCgCAgATTGCACAgCTgAAgCAATTAAAgC-3’
YOSVV454I-PvuII2	5’-gCTTTAATTgCTTCAgCTgTgCAATCTgCgATTgTATAg-3’
TKPN-N	5’- ATAAgAAYGCggCCgCgAgCTACgTCAgggCCCCTAg -3’
TKOSC-SXS	5’- ACCgCgAgCTCgAgTCgACTTAAAgCgTATgTGTTTCg -3’
pYES2-outerprimer2	5’-CCTTCCTTTTCggTTAgAgCggATgTgg-3’

Table 8.2. Primer sequences in the site-saturated mutagenesis experiment

Oligonucleotide Primers	Sequences
YTL-H234X1	5'- TgggTTTATACTNNNgCTGTTTACATT-3'
YTL-H234X2	5'- AATgTAAACACCNNNAgTATAAACCCA-3'
YTL-F445X1	5'-AgAAAgggTgCCTggggCNNNTCAACA-3'
YTL-F445X2	5'- TgTTgANNNgCCCCAggCACCCTTTCT-3'

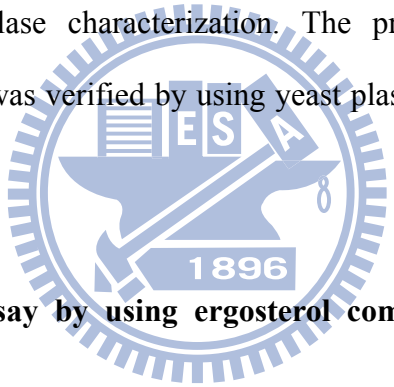
8.2-2 Preparation of competent yeast cell (CBY57 and TKW14c2)

Place the yeast TKW14c2 stock from the refrigerator into 3 mL of SD medium containing 60 μL of 50 X ALTHMU solutions, 120 μL 50% glucose solution, 60 μL of Heme solution, and 60 μL of Ergosterol supplement solution. Incubate each solution at 30 °C for three days until turbidity occurs. Transfer the yeast cells into 100 mL of SD medium containing the same nutrition additive condition and incubate at 30 °C for 12-18 hours. After OD_{600} of yeasts arrived at 1.0~1.5, the yeast cells were collected by centrifugation at 3,000 rpm for 10 min at 4 °C. Add 35 mL aseptic ddH_2O to suspend the cell pellet and centrifuge at 3,000 rpm for 10 min at 4°C. Repeat above step twice. Add 25 mL of 1 M D-sorbitol solution to suspend the cell pellet and wash it by centrifugation. Finally, add n x 50 μL (where n is the number of the transformed plasmids) of 1 M D-sorbitol to the cell pellet and suspend it on ice for 5 min. Before electroporation, 50 μL of competent cells were added into each of 1.5 mL microtubes with 5 μL of recombinant plasmids, respectively. The preparation protocol for the yeast CBY57 strain is almost the same, except the yeast cultural growth medium was changed to ALHU solution and Glucose solution.

8.2-3 Cyclase activity assay by using plasmid shuffle method in CBY57 strain

The pRS314-derived ERG7 mutated plasmids were electroporated into a yeast cyclase-deficient haploid strain CBY57[pZS11]. The pRS314 and pRS314ERG7WT

plasmids (*TRP1* centrometric plasmids with no insert) and the wild-type *S. cerevisiae* ERG7 gene were used as the negative and positive control, respectively. The yeast transformants were plated on SD+Ade+Lys+His+ 1 M sorbitol-medium plates for three days to determine the presence of both pZS11-derived and pRS314-derived plasmids. Individual colonies were chosen and grown in a 10 mL Ade+Lys+His+Ura liquid culture. Aliquots of 100 μL of each culture were simultaneously plated on the SD+Ade+Lys+His+Ura and SD+Ade+Lys+His+Ura+ 1 mg/mL 5'-Fluroorotic acid medium plates, and they were grown for another three days to elucidate the complementation effect. Colonies that grew on the non-5'-FOA plates, but not on 5'-FOA plates, were separately chosen and grown in SD+Ade+Lys+His+Ura liquid medium for mutated cyclase characterization. The presence of both pZS11 and pRS314-derived plasmids was verified by using yeast plasmid minipreps and restrictive endonuclease analysis.



8.2-4 Cyclase activity assay by using ergosterol complementation in TKW14c2 strain

The mutated plasmids were transformed into the TKW14c2 strain, a yeast *hem1/erg7* double deficient strain, by using the electroporation. Aliquots of 120 μL of the culture were plated onto SD+Ade+Lys+His+Met+Ura+Hemin+Erg+G418 plates and replated on the same plates, except without addition of Ergosterol, for selection of the cyclase complement ability. The pRS314 and pRS314ERG7WT plasmids were also transformed as negative and positive control, respectively. Different TKW14c2 [pERG7^{mutations}] transformants were then separately grown in SD+Ade+Lys +His+Met+Ura+Hemin+Erg medium at 30 °C with shaking (220 rpm) for seven to ten days for the characterization of the products.

8.2-5 Lipids extraction and column chromatography

The yeast transformants with different cyclase mutated plasmids were grown in the 500 mL SD liquid medium containing Ade+Lys+His+Met+Ura+Hemin+Erg at 30 °C with 220 rpm mixing for several days. Yeast cells were harvested by centrifugation at 6,000 rpm for 10 min. The washed cell pellets were then suspended in a solution containing 50 mL of ethanol, 50 mL of 30% KOH, and 0.1 g of Pyrogallol.[3] This reaction was refluxed at 110 °C for 3 hr. The hydrolysate was extracted three times with a total of 500 mL of petroleum ether. The organic phase was collected and concentrated by using a rotary evaporator. The non-saponifiable lipids were then dissolved in CH_2Cl_2, spotted on the TLC plates, and developed with 4:1 hexane/ethyl acetate solution. TLC plates were subjected to the staining buffer and heated until the pattern of the products appeared. In order to fractionate the different product patterns in these mutants, large scale cultural media were harvested, washed, and saponified by refluxing, as previously described. The non-saponifiable lipids (NSLs) were extracted and fractionated by using silica gel column chromatography with 19:1 hexane/ethyl acetate solution. The separated fraction was then concentrated and analyzed on the TLC plate. The same R_f value spots were collected and separated in either native form or in the acetylated form on silica gel or on the HPLC. The resultant products were analyzed by using GC, GC-MS, and different NMR spectra.

8.3-6 Acetylation modification and the alkaline hydrolysis reaction

The acetylation modification of the triterpene alcohol products was performed according to the previous literature.[115] The dry triterpene alcohols were first dissolved in 2 mL of pyridine. Then, the excess amount of 10 mM of acetic anhydride was added. The solution was stirred overnight at room temperature. The acetylation reaction was monitored by TLC analysis. After 16 hr, 5 mL of water were added to stop the reaction, and extraction with 10 mL of CH_2Cl_2 was conducted three times. The total organic

phase was collected, dried over sodium sulfate, and then evaporated in a a rotary evaporator. The acetylated products were then applied for silica chromatography and analyzed by using GC-MS and NMR spectra. The acetylated derivatives were also hydrolyzed to yield the native non-saponifiable lipid.[116] The acetylated derivatives were refluxed in the 5% KOH/methanol solution for 3 hr. After extraction with petroleum ether, the evaporated crude extracts were separated and analyzed, as previously described.

8.2-7 GC and GC-MS column chromatography

GC analyses were performed with a Hewlett-Packard model 5890 series II or Agilent 6890N chromatograph equipped with a DB-5 column (30 m x 0.25 mm I.D.; 0.25-μm film; the oven gradient was set at 50 °C for 2 min, then the temperature was increased at the rate of 20 °C/min until 300 °C was reached and sustained for 20 min; Injector: 300 °C; Interface: 250 °C; 1/40 split ratio; Carrier gas: helium with 13 psi column head pressure). GC/MS was performed on an Agilent 6890N chromatograph equipped with a DB-5MS column (30m x 0.25 mm I.D.; 0.25-μm film; Injector: 250 °C; GC-MS transfer line: 280 °C), coupled to an Agilent Technologies model 5973 MSD mass spectrometer (EI^+ energy at 70 eV, ion source temperature: 230 °C, and Scan range: 50-500). The other program parameters are the same as the GC conditions.

8.2-8 Sequence alignment and molecular modeling

The multiple sequence alignment of OSCs was generated by using CLUSTAL W program at the website (http://www.ebi.ac.uk/clustalw/). Sequence information of human OSC or *S. cerevisiae* ERG7 was extracted from the NCBI database (*H. sapiens* OSC: p48449[117], *S. cerevisiae* ERG7: p38604[118]). The crystal structure of human OSC was obtained from the Protein Data Bank (PDB entry 1W6K; structure of human OSC in complex with lanosterol[6]). Ligand and the detergent molecules in the 1W6K structure were removed before building the homology model. Mutated residue was obtained by

altering the wild-type gene. Different ligand molecules for the molecular docking experiment were generated by using Chem3D ultra software.

The genetic algorithm (GA) for protein-ligand docking software "Gold" was chosen for calculating the docking modes of different cationic intermediates into the binding site of the newly built ERG7 homology modeling structure. First, the ERG7 modeled structure file and the ligand molecule files were loaded into the "Gold," respectively. Second, the "Gold active site residue" was defined as "atom number –3684," which corresponds to the Cys-457 in the ERG7 protein. This atom is theoretically close to the center of the active site. The "active-site cavity" was then defined as a diameter of 10 Å. After submitting the ligands into the Gold software, four to ten docking modes with slightly different orientations were generated in "Gold output files." The final model was chosen according to the detailed verification and comparison with the lanosterol molecule in the human OSC crystalized structure. After the determination of the docking mode, the ligand molecule was then added into the homology model file by using the "WebLab ViewerPro" Software (Molecular Simulations, Inc.). The coordination section of the docked ligand was altered according to the HETATM records in the PDB format. The connecting bond and series numbers in the ligand were also carefully verified. The integrated file with the protein-ligand mixture was then uploaded into the "SYBYL" program and saved as a "mol 2" file. Finally, energy minimization with "SYBYL" was conducted as previously described, including hydrogen atom addition, charge alteration, and energy minimization calculation.[65] The plausible model of the OSC protein was generated based on the reference proteins and the human OSC in the PDB bank.

8.2-9 Molecular cloning of *P. sativum* β-amyrin synthase

Pea seeds were purchased from Hsin-Chu (Taiwan, ROC). Immature pea seeds were harvested on day 11 after planting, immediately frozen with liquid nitrogen, and stored

at -80 °C. The preparation of cDNA was conducted using a Poly(A) Pure^{TM} Kit with slight modification (QIAGEN, Hilden, Germany). In brief, mRNA (0.4 μg) was reverse-transcribed to produce a cDNA mixture using M-MLV RTase (Invitrogen, Carlsbad, CA, USA). Ten mM for each dNTP and 0.5 μg of oligo $(dT)_{12-18}$ in a total volume of 20 μL were used. The mixture was incubated at 65 °C for 5 min before being centrifuged at 14,000 rpm for 1 min at 4 °C. The supernatant was added to 4 μL of 5X first strand buffer, 1 μL 0.1 M DTT, and 1 μL RNaseOUT^{TM} ribonuclease inhibitor and incubated at 37 °C for 2 min. One microliter (200 units) of M-MLV RTase was added to the mixture, and the reaction tube was incubated at 37 °C for 1 hr before termination by heating at 70 °C for 15 min. The resulting cDNA mixture was directly used as a template for the following PCRs.

8.2-10 Subcloning of full-length of *P. sativum* β-amyrin synthase

A sense primer was designed for the N-terminal region, and anti-sense primer was designed for the C-terminal region according to the database sequence of *P. sativum* β-amyrin synthase. The core nucleotide sequences of three primers were synthesized in the sense (YTL-cDNA-PEP1) and anti-sense directions (YTL-cDNA-PEP2, YTL-cDNA-PEP4). The primer sequence is listed in Table 9.3. The PCR reaction was conducted with different pair of primers using *rTth* DNA polymerase (Biobasic, Inc., Toronto, Canada) with 15 mM of $MgCl_2$, 10 mM for each dNTP, 1 μg of each primer, 1.5 μL of glycerol, 2 μL of cDNA template, 5 μL of 10 X *Taq* buffer and DEPC water to the final volume of 50 μL. The reaction mixture was denaturated at 94 °C for 2 min and then for 35 cycles of denaturing at 94 °C for 1 min, annealing at 56 °C for 1 min, and extending at 72 °C for 2 min. The mixture underwent a final extension at 72 °C for 10 min to ensure the complete synthesis of the PCR product. This DNA fragment was digested with *EcoRI* and *XhoI*, ligated into the corresponding position of pYES2 (Invitrogen), and transformed into the *E. coli* XL1-Blue strain for amplification. The

insert region of the PSY coding sequence was sequenced and verified by using an ABI PRISM 3100 auto-sequencer.

8.2-11 Functional expression of *P. sativum* β-amyrin synthase

The pYES2-PSY was transformed into yeast TKW14c2 by using the electroporation method, plated on the SD+Ade+Lys+Trp+His+Met+Glucose+Hemin+Egosterol medium, and cultured at 30 °C for selecting the desired transformant. The transformant was grown in liquid culture (1 L) for 2 days at 30 °C with shaking (220 rpm). Cells were collected and resuspended in Ade+Lys+Trp+His+Met medium, supplemented with ergosterol (20 mg/mL), hemin (13 mg/mL), Tween 80 (5 mg/mL), and 2% galactose for induction at 30 °C for 10 hr. Cells were again collected and resuspended in 0.1 M potassium phosphate buffer (pH 7.0) (500 mL) supplemented with 3% glucose and hemin (13 mg/mL) and cultured for 24 hr at 30 °C. Cells were refluxed with 50 mL of 15% KOH/50% Alcohol (aq) for 3 hr, and the mixture was extracted with petroleum ether (40-60 °C) three times and concentrated by a rotary evaporator. The extract was purified by silica gel column chromatography with a 9:1 hexane/ethyl acetate mixture, and the same R_f value spots of lanosterol were collected and used for GC-MS for analysis.

Table 8.3. The primers used to clone the β-amyrin synthase from *P. sativum*. The underlined letters show the unique restriction enzyme site for construction

Oligonucleotide primers	**Sequence**
YTL-cDNA-EN1 [N-terminal]	5’- TCTAgAATTCCATATgTggA(g/A)(T/C)TgAAg-3’
YTL-cDNA-pep1 [Forward]	5’-gg(T/A)ggT(g/A)AAgA(C/A)(A/g)ATgCTTg(C/T)g-3’
YTL-cDNA-pep2 [Reverse]	5’-gTg CA (C/T) TCAACATATTCA(T/C)g-3’
YTL-cDNA-pep4 [Reverse]	5’-AT(A/g)gCT(T/C)TCCCCCCACCC(g/A)CC(g/A)TC-3’
YTL-YcDNA-XC1 [C-terminal]	5’-gATTTgTCCTCgAgCTCAAggCAAAg-3’

References

(1) Abe, I. R.; Prestwich, G.D. *Chem. Rev.* **1993**, *93*, 2189-2206.

(2) Robinson, R. *J. Soc. cem. Ind (London)* **1934**, *53*, 1062-1063.

(3) Bloch, K; Rittenberg, D. *Biol .Chem* **1942**, *145*, 625-636.

(4) Woodward, R. B.; Bloch, K. *J. Am. Chem. Soc* **1953**, *75*, 2023-2024.

(5) Cornforth, J. W.; Cornforth, R. H.; Donninger, C.; Popjak, G.; Shimizu, Y.; Ichii, S.; Forchielli, E.; Caspi, E. *J Am Chem Soc* **1965**, *87*, 3224-3228.

(6) Corey, E. J.; Russey, W. E. *J Am Chem Soc* **1966**, *88*, 4751-4752.

(7) Van Tamelen, E. E.; Willett, J. D.; Clayton, R. B.; Lord, K. E. *J Am Chem Soc* **1966**, *88*, 4752-4754.

(8) Willett, J. D.; Sharpless, K. B.; Lord, K. E.; Van Tamelen, E. E.; Clayton, R. B. *J Biol Chem* **1967**, *242*, 4182-4191.

(9) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew Chem Int Ed Engl* **2000**, *39*, 2812-2833.

(10) Ruzicka, Z. L. *Cellular and Molecular Life Sciences* **1953**, *9*, 357-367.

(11) Van Tamelen, E. E.; Leopold, E. J.; Marson, S. A.; Waespe, H. R. *J Am Chem Soc* **1982**, *104*, 6480-6481.

(12) Van Tamelen, E. E.; James, D. R. *J Am Chem Soc* **1977**, *99*, 950 - 952.

(13) Corey, E. J.; Cheng, H.; Hunter Baker, C.; Matsuda, S. P. T.; Ding, L.; Xuelei, S. *J Am Chem Soc* **1997**, *119*, 1277-1288.

(14) Abe, I. R.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189-2206.

(15) Boar, R. B.; Couchman, L. A.; Jaques, A. J; Perkins, M. J. *J. Am. Chem. Soc* **1984**, *106*, 2476-2477.

(16) Renoux, J. M.; Rohmer, M. *Eur J Biochem* **1986**, *155*, 125-132.

(17) Van Tamelen, E. E.; Sharpless, R. P.; Hanzlik, R.; Clayton, R. B.; Burlingame,A. L.; Wszolek P. C. *J Am Chem Soc* **1967**, *89*, 7150 - 7151.

(18) Corey, E. J.; Virgil, S. C.; Cheng,H.; Baker, C. H.; Matsuda,S. P. T.; Singh, V. ; Sarshar, S. *J Am Chem Soc* **1995**, *117*, 11819-11820.

(19) Hoshino, T.; Sakai, Y. *Chem. Commun.* **1998**, 1591-1592.

(20) Robin, B. B.; Alan, J. J.; Perkins, M. J. *J Am Chem Soc* **1984**, *106*, 2476-2477.

(21) Corey, E. J.; Cheng, H.; Baker, C.H.; Matsuda, S.P.T; Li, D.; Song, X. *J Am Chem Soc* **1997**, *119*, 1289-1296.

(22) Corey, E. J.; Virgil, S. C.*J Am Chem Soc* **1991**, *113*, 4025-4026.

(23) Corey, E. J.; Cheng, H. *Tetrahedron lett.* **1996**, *37*, 2709-2712.

(24) Corey, E. J.; Ortiz de, M.; Yamamoto, H. *J Am Chem Soc* **1968**, *90*, 6254 - 6255.

(25) Corey, E. J.; Virgil, S. C.; Liu, D. R.; Sarshar, S. *J Am Chem Soc* **1992**, *114*,

1524-1525.
(26) Corey, E. J.; Virgil, S. C. *J Am Chem Soc* **1991**, *113*, 4025-4026.
(27) Ourisson, G. *Pure Appl. Chem.* **1989**, *61*, 345-348.
(28) Johnson, W. S.; Steele, J. J. *J. Am. Chem. Soc* **1987**, *41*, 301-303.
(29) Johnson, W. S.; Telfer, S. J.; Cheng, S.; Schubert, U. *J. Am. Chem. Soc* **1987**, *109*, 2517-2518.
(30) Poralla, K. *Bioorg. Med. Chem. Lett* **1994**, *1994*, 285-290.
(31) Xiao, X.Y.; Prestwich, G. D. *J Am Chem Soc* **1991**, *113*, 9673-9674.
(32) Abe, I.; Prestwich, G. D. *J. Biol .Chem* **1994**, *269*, 802-804.
(33) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811-1815.
(34) Shi, Z.; Buntel, C. J.; Griffin, J. H. *Proc Natl Acad Sci U S A* **1994**, *91*, 7370-7374.
(35)Abe, I.; Bai, M.; Xiao, X. Y.; Prestwich, G. D. *Biochem Biophys Res Commun* **1992**, *187*, 32-38.
(36) Abe, I.; Prestwich, G. D. *J Biol Chem* **1994**, *269*, 802-804.
(37) Seckler, R. *Biochem Biophys Res Commun* **1986**, *134*, 975-981.
(38) Duriatti, A.; Schuber, F. *Biochem Biophys Res Commun* **1988**, *151*, 1378-1385.
(39) Carrano, L. *J. Med. Vet. Mycol.* **1995**, *33*, 53-58.
(40) Buntel, C. Ph.D. Thesis, Department of Chemistry, Stanford University, Stanford, CA **1996**.
(41) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J Mol Biol* **1999**, *286*, 175-187.
(42) Lenhart, A.; Weihofen, W. A.; Pleschke, A. E.; Schulz, G. E. *Chem Biol* **2002**, *9*, 639-645.
(43) Hoshino, T.; Kouda, M.; Abe, T.; Ohashi, S. *Biosci Biotechnol Biochem* **1999**, *63*, 2038-2041.
(44) Sato, T.; Hoshino, T. *Chem. Commun.* **1998**, 2617-2618.
(45) Sato, T.; Hoshino, T. *Chem.Commun.* **1998**, 2617-2618.
(46) Merkofer, T.; Pale-Grosdemange, C.; Wendt, K. U.; Rohmer, M.; Poralla, K. *Tetrahedron Lett.* **1999**, *40*, 2121-2124.
(47) Pale-Grosdemange, C. M.; Rohmer, M.; Poralla, K. *Tetrahedron lett.* **1999**, *40*, 6009-6012.
(48) Hoshino, T.; Sato, T. *Chem. Commun.* **1999**, 2005-2006.
(49) Thoma, R.; Schulz-Gasch, T.; D'Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. *Nature* **2004**, *432*, 118-122.
(50) Abe, I. *Nat Prod Rep* **2007**, *24*, 1311-1331.
(51) Wu, T. K.; Chang, C. H.; Liu, Y. T.; Wang, T. T. *Chem Rec* **2008**, *8*, 302-325.
(52) Connolly, J. D.; Hill, R. A. *Nat. Prod. Rep.* **2007**, *24*, 465-486.
(53) Xu, R.; Fazio, G. C.; Matsuda, S. P. *Phytochem.* **2004**, *65*, 261-291.
(54) Phillips, D. R.; Rasbery, J. M.; Bartel, B.; Matsuda, S. P. *Curr. Opin. Plant. Biol.*

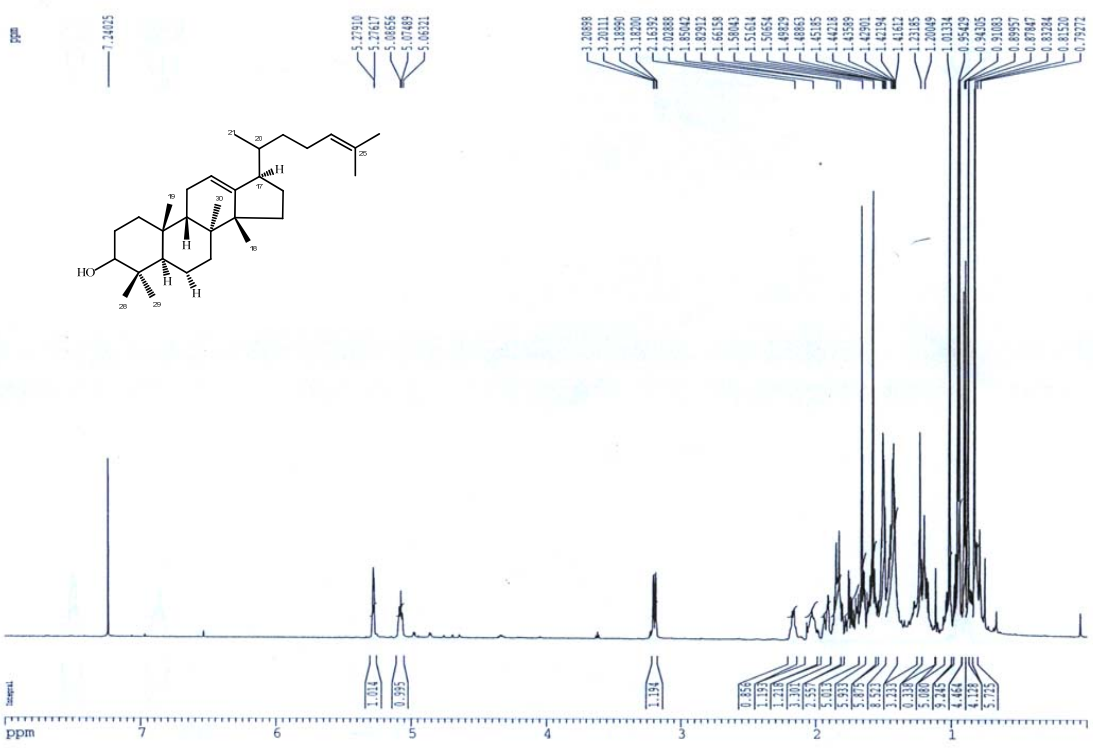
2006, *9*, 305-314.
(55) Trapp, S. C.; Croteau, R. B. *Genetics* **2001**, *158*, 811-832.
(56) Segura, M. J.; Jackson, B. E.; Matsuda, S. P. *Nat Prod Rep* **2003**, *20*, 304-317.
(57) Wu, T. K.; Griffin, J. H. *Biochemistry* **2002**, *41*, 8238-8244.
(58) Meyer, M. M.; Xu, R.; Matsuda, S. P. *Org Lett* **2002**, *4*, 1395-1398.
(59)Segura, M. J.; Lodeiro, S.; Meyer, M. M.; Patel, A. J.; Matsuda, S. P. *Org Lett* **2002**, *4*, 4459-4462.
(60) Lodeiro, S.; Schulz-Gasch, T.; Matsuda, S. P. *J Am Chem Soc* **2005**, *127*, 14132-114133.
(61) Tetsuo Kushiro, M. S.; Ebizuka, Y. *J Am Chem Soc* **1999**, *121*, 1208-1216.
(62) Tetsuo Kushiro, M. S.; Masuda, K.; Ebizuka, Y. *J Am Chem Soc* **2000**, *122*, 6816-6824.
(63) Schulz-Gasch, T. and Stahl, M. *Joural of computational chemistry* **2002**, *24*, 741-753.
(64) Griffin, J. H.; Buntel, C. J.; Siregar J. J. *Ph.D Thesis, Department of Chemistry, standford University*
(65) Chang, C. H. *Ph.D Thesis, Department of biological Chemistry, National Chiao-Tung University* **2008**.
(66) Lewis, T. A.; Taylor, F. R.; Parks, L. W. *J Bacteriol* **1985**, *163*, 199-207.
(67) Gollub, E. G.; Liu, K. P.; Dayan, J.; Adlersberg, M.; Sprinson, D. B. *J Biol Chem* **1977**, *252*, 2846-2854.
(68) Wu, T. K.; Liu, Y. T.; Chang, C. H. *Chembiochem* **2005**, *6*, 1177-1181.
(69) Wu, T. K.; Liu, Y. T.; Chang, C. H.; Yu, M. T.; Wang, H. J. *J Am Chem Soc* **2006**, 6414-6419.
(70) Wendt, K. U. *Angew Chem Int Ed Engl* **2005**, *44*, 3966-3971.
(71) Zacharias, N.; Dougherty, D. A. *Trends Pharmacol Sci* **2002**, *23*, 281-287.
(72) Cunningham, B. C.; Wells, J. A. *Science* **1989**, *244*, 1081-1085.
(73) Wu, T. K.; Chang, C. H. *Chembiochem* **2004**, *5*, 1712-1715.
(74) Wu, T. K.; Liu, Y. T.; Chiu, F. H.; Chang, C. H. *Org Lett* **2006**, *8*, 4691-4694.
(75) Wu, T. K.; Wen, H. Y.; Chang, C. H.; Liu, Y. T. *Org Lett* **2008**, *10*, 2529-2532.
(76) Wu, T. K.; Chang, C. H.; Wen, H. Y.; Liu, Y. T.; Li, W. H.; Wang, T. T.; Shie, W. S. *Org Lett* **2009**, *12*, 500-503.
(77) Wu, T. K.; Wang, T. T.; Chang, C. H.; Liu, Y. T.; Shie, W. S. *Org Lett* **2008**, *10*, 4959-4962.
(78) Wu, T. K.; Li, W. H. *The Dissertation for the Degree of Master* **2007**.
(79) Wu, T. K.; Wen, H. Y. *The Dissertation for the Degree of Master* **2007**.
(80) Wu, T. K.; Wang, T. T. *The Dissertation for the Degree of Master* **2008**.
(81) Wu, T. K.; Chang, Y. C. *The Dissertation for the Degree of Master* **2009**.

(82) Wu, T. K.; Yu, M. T. *The Dissertation for the Degree of Master* **2005**.
(83) Wu, T. K. Li, W. H.; Chang, C. H.; Wen, H. Y.; Liu, Y.T.; Chang, Y. H. *European Journal of Organic Chemistry* **2009**, *2009*, 5731-5737.
(84) Wu, T. K.; Yu, M. T.; Liu, Y. T.; Chang, C. H.; Wang, H. J.; Diau, E. W. *Org Lett* **2006**, *8*, 1319-1322.
(85) Hoshino, T.; Abe, T.; Kouda, M. *Chem. Commun.* **2000**, 441.
(86) Gerlt, J. A.; Babbitt, P. C. *Curr Opin Chem Biol* **2009**, *13*, 10-18.
(87) Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. *J Am Chem Soc* **2007**, *129*, 11213-11222.
(88) Yoshikuni, Y.; Ferrin, T. E.; Keasling, J. D. *Nature* **2006**, *440*, 1078-1082.
(89) Kushiro,T.; Shibuya, M.; Ebizuka, Y. *J Am Chem Soc* **1999**, *121*, 1208-1216.
(90) Back, K.; Chappell, J. *Proc Natl Acad Sci U S A* **1996**, *93*, 6841-6845.
(91) Ponomarenko, L. P.; Kalinovsky, A. I.; Moiseenko, O. P.; Stonik, V. A. *Comp Biochem Physiol B Biochem Mol Biol* **2001**, *128*, 53-62.
(92) Kicha, A. A.; Ivanchina, N. V.; Gorshkova, I. A.; Ponomarenko, L. P.; Likhatskaya, G. N.; Stonik, V. A. *Comp Biochem Physiol B Biochem Mol Biol* **2001**, *128*, 43-52.
(93) Cordeiro, M. L.; Djerassi, C.*J. Org. Chem.* **1990**, *55*, 2806.
(94) Kerr, R. G.; Chen, Z. *J Nat Prod* **1995**, *58*, 172-176.
(95)Whitson, E. L.; Bugni, T. S.; Chockalingam, P. S.; Concepcion, G. P.; Harper, M. K.; He, M.; Hooper, J. N.; Mangalindan, G. C.; Ritacco, F.; Ireland, C. M. *J Nat Prod* **2008**, *71*, 1213-1217.
(96) Song, Z.; Nes, W. D. *Lipids* **2007**, *42*, 15-33.
(97) Parish, E. J.; Nes, W. D. **1997**.
(98) Bellamine, A.; Mangla, A. T.; Nes, W. D.; Waterman, M. R. *Proc Natl Acad Sci U S A* **1999**, *96*, 8937-8942.
(99) Bellamine, A.; Mangla, A. T.; Dennis, A. L.; Nes, W. D.; Waterman, M. R. *J Lipid Res* **2001**, *42*, 128-136.
(100) Mitsuguchi, H.; Seshime, Y.; Fujii, I.; Shibuya, M.; Ebizuka, Y.; Kushiro, T. *J Am Chem Soc* **2009**, *131*, 6402-6411.
(101) Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Ivanova, Y.; Smith, M. L.; May, G. S.; Matsuda, S. P. *Org Lett* **2009**, *11*, 1241-1244.
(102) Mangla, A. T.; Nes, W. D. *Bioorg Med Chem* **2000**, *8*, 925-936.
(103) Nes, W. D. *Biochim Biophys Acta* **2000**, *1529*, 63-88.
(104) Morita, M.; Shibuya, M.; Lee, M. S.; Sankawa, U.; Ebizuka, Y. *Biol Pharm Bull* **1997**, *20*, 770-775.
(105) Nixon, A. E.; Firestine, S. M.; Salinas, F. G.; Benkovic, S. J. *Proc Natl Acad Sci U S A* **1999**, *96*, 3568-3571.
(106) Jurgens, C.; Strom, A.; Wegener, D.; Hettwer, S.; Wilmanns, M.; Sterner, R. *Proc*

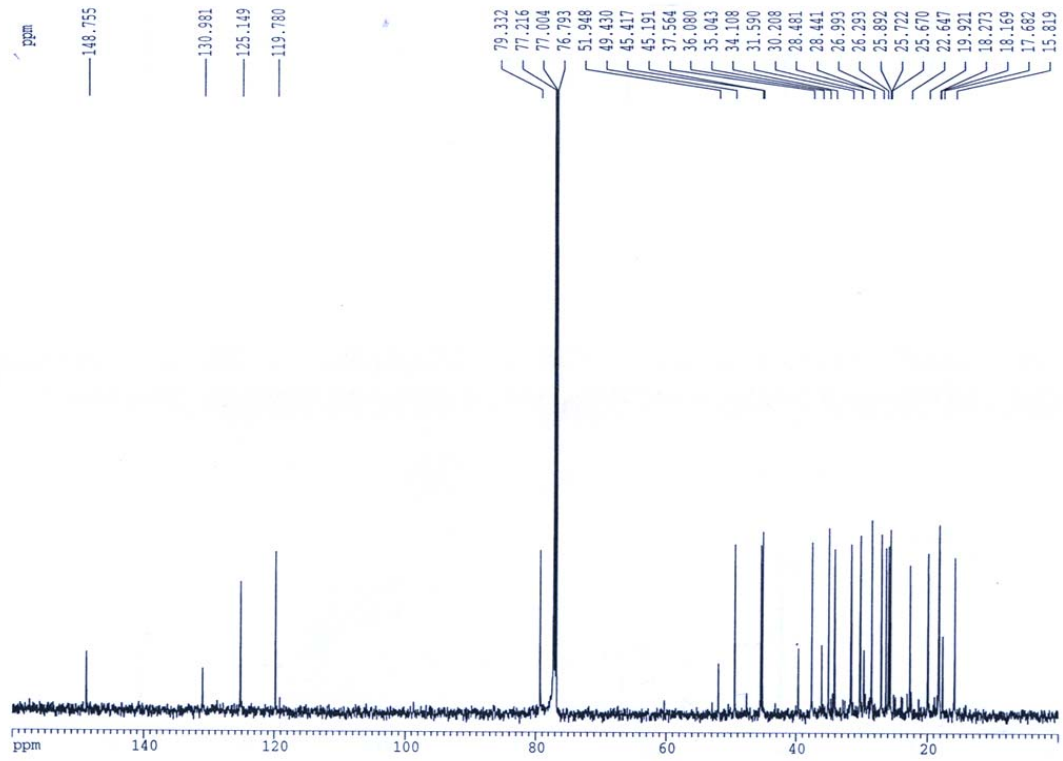
Natl Acad Sci U S A **2000**, *97*, 9925-9930.
(107) Seebeck, F. P.; Hilvert, D. *J Am Chem Soc* **2003**, *125*, 10158-10159.
(108) Lodeiro, S.; Segura, M. J.; Stahl, M.; Schulz-Gasch, T.; Matsuda, S. P. *Chembiochem* **2004**, *5*, 1581-1585.
(109) Morita, M.; Shibuya, M.; Kushiro, T.; Masuda, K.; Ebizuka, Y. *Eur J Biochem* **2000**, *267*, 3453-3460.
(110) Shibuya, M.; Zhang, H.; Endo, A.; Shishikura, K.; Kushiro, T.; Ebizuka, Y. *Eur J Biochem* **1999**, *266*, 302-307.
(111) Kawano, N.; Ichinose, K.; Ebizuka, Y. *Biol Pharm Bull* **2002**, *25*, 477-482.
(112) Hayashi, H.; Huang, P.; Inoue, K.; Hiraoka, N.; Ikeshiro, Y.; Yazaki, K.; Tanaka, S.; Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur J Biochem* **2001**, *268*, 6311-6317.
(113) Husselstein-Muller, T.; Schaller, H.; Benveniste, P. *Plant Mol Biol* **2001**, *45*, 75-92.
(114) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning: A laboratory manual 2nd ed* **1989**.
(115) Freimund, S.; Kopper, S. *Carbohydrate Research* **1998**, *308*, 195-200.
(116) Akihisa, T.; Arai, K.; Kimura, Y.; Koike, K.; Kokke, W. C.; Shibata, T.; Nikaido, T. *J. Nat. Prod.* **1999**, *62*, 265-268.
(117) Baker, C. H.; Matsuda, S. P. T.; Liu, D. R.; Corey, E. J. *Biochem Biophys Res Commun* **1995**, *213*, 154-160.
(118) Shi, Z.; Buntel, C. J.; Griffin, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7370-7374.

Appendix 3.1. Protosta-12,24-dien-3β-ol NMR characterization

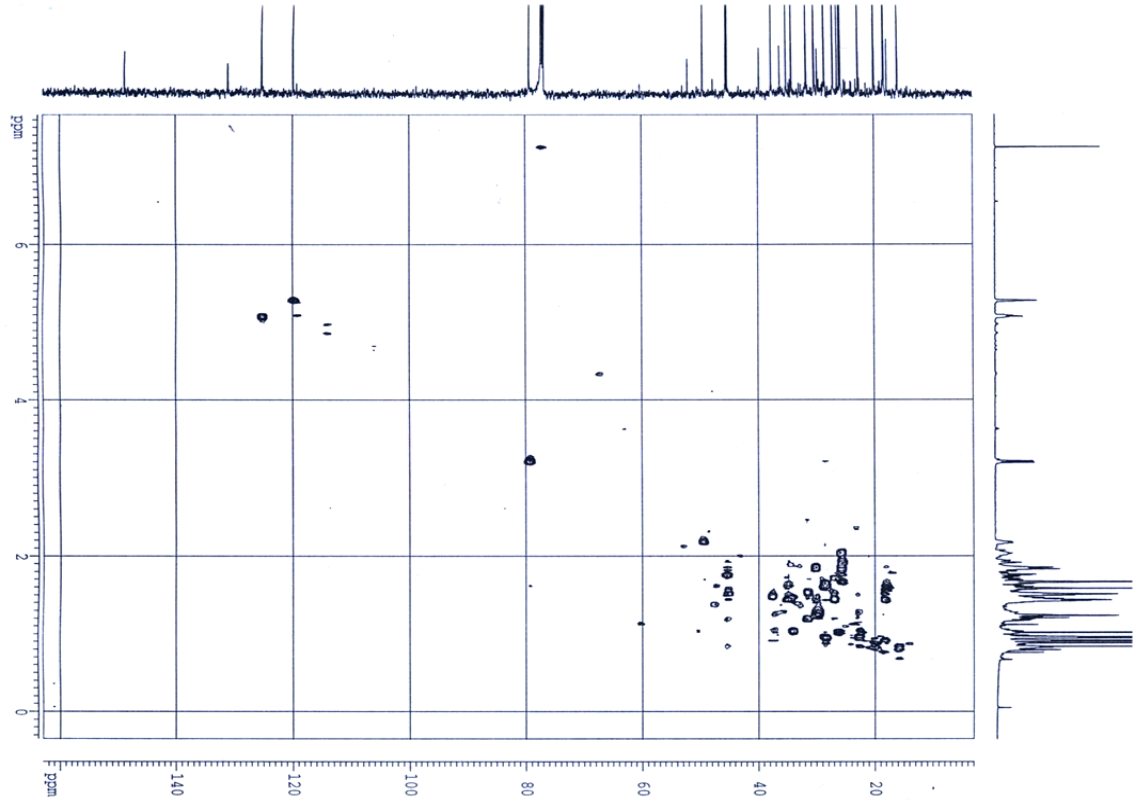
^{1}H-NMR spectrum of protosta-12,24-dien-3β-ol



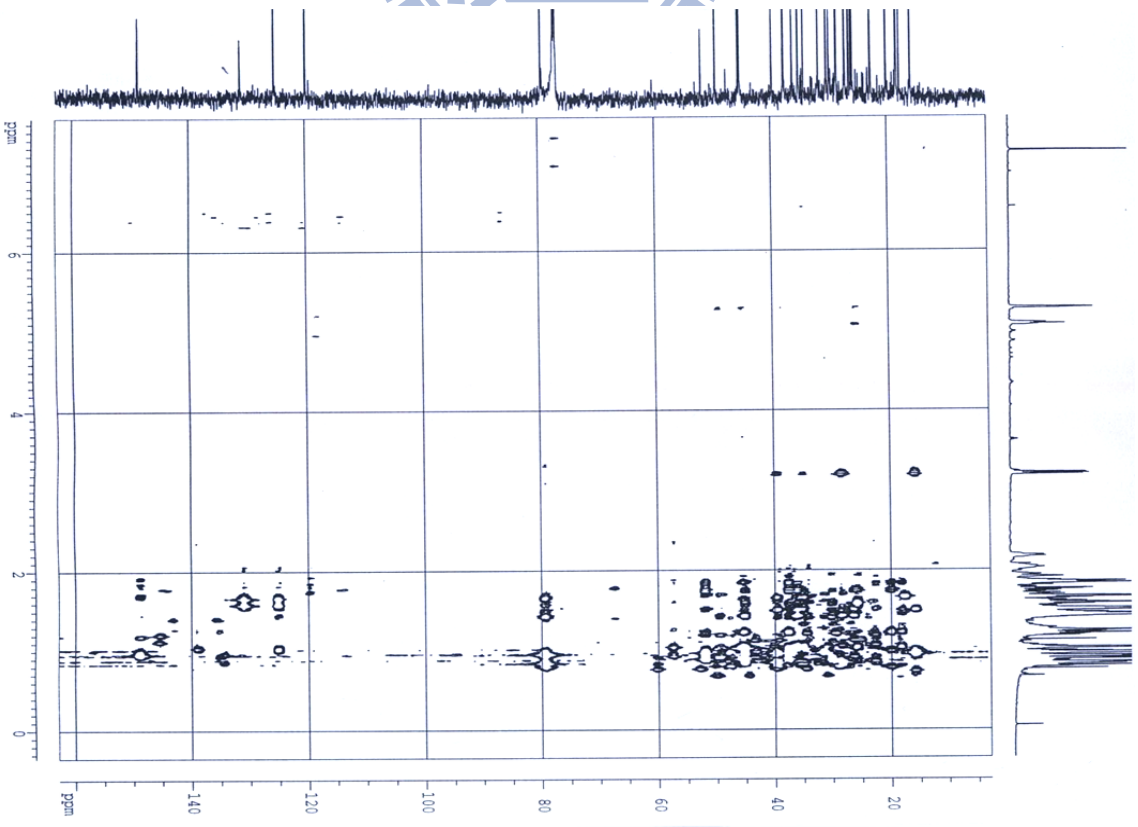
^{13}C NMR spectrum of protosta-12,24-dien-3β-ol



HSQC NMR spectrum of protosta-12,24-dien-3β-ol

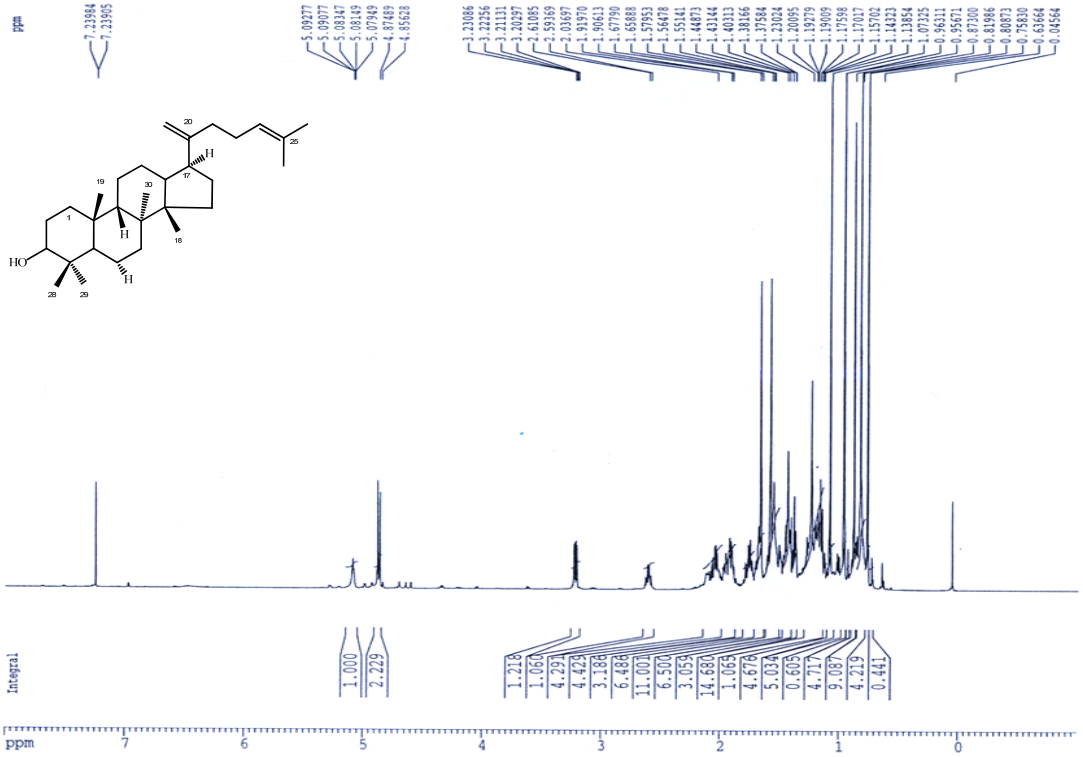


HMBC NMR spectrum of protosta-12,24-dien-3β-ol

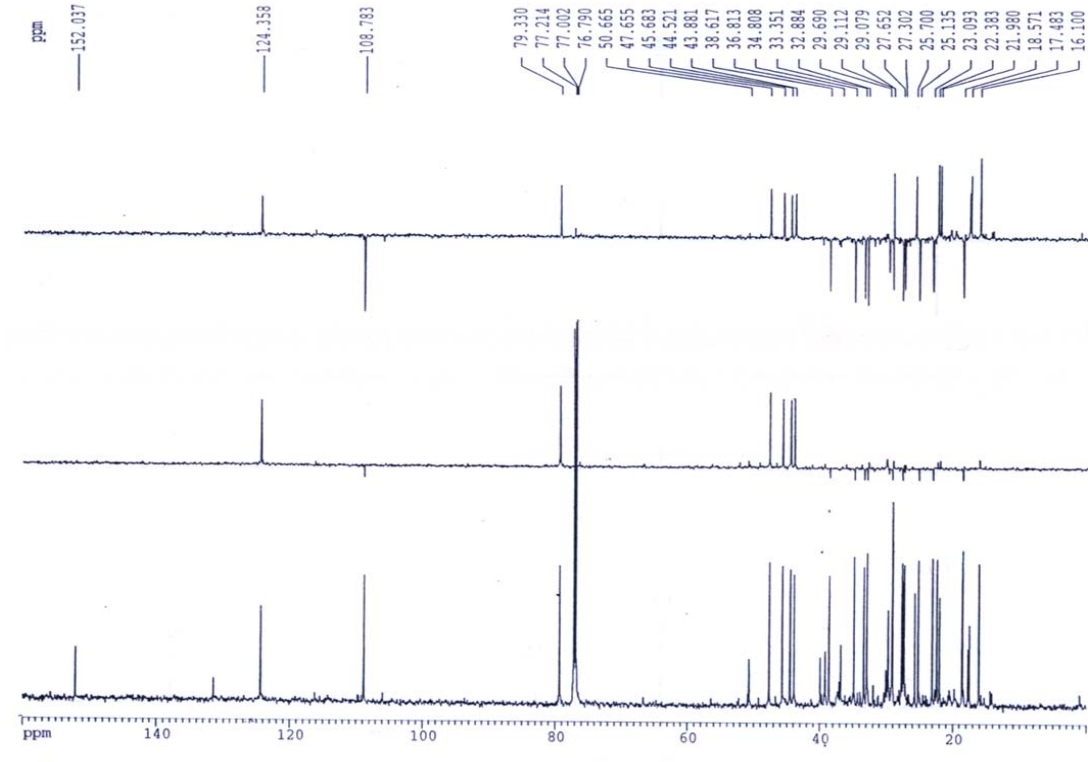


Appendix 3.2. Protosta-20,24-dien-3β-ol NMR characterization

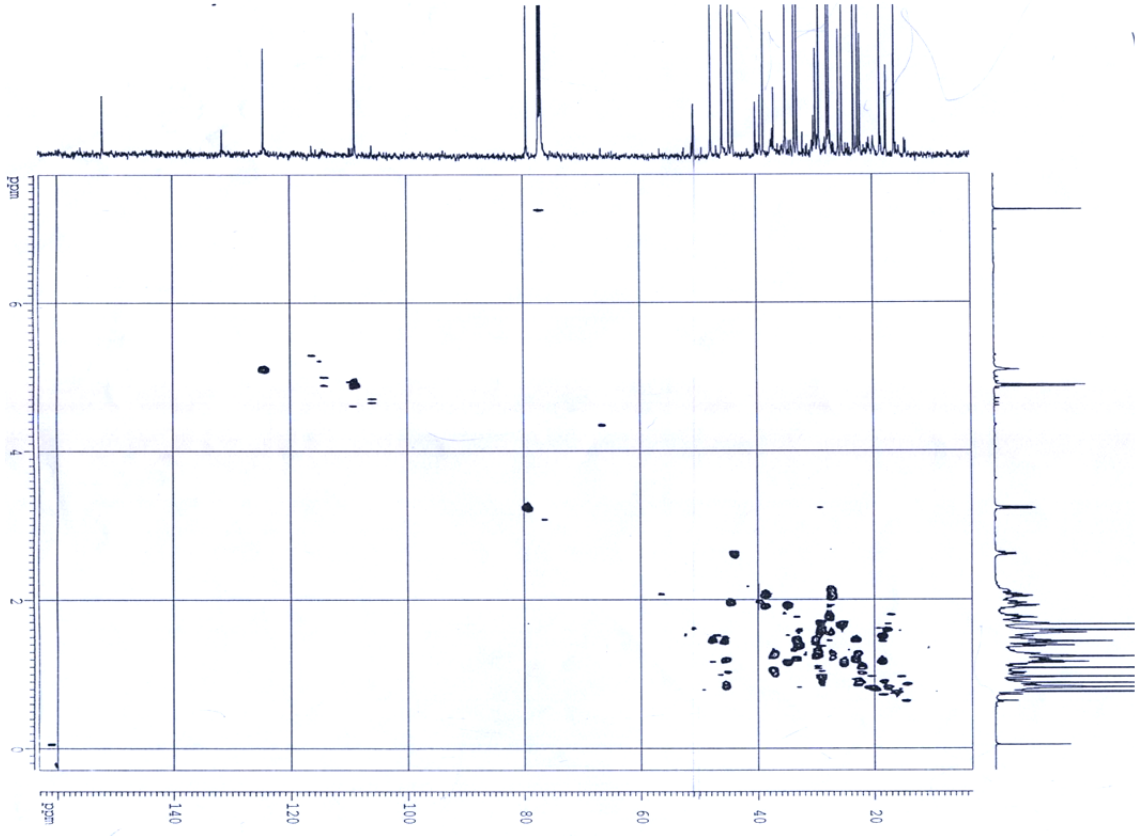
^{1}H-NMR spectrum of protosta-20,24-dien-3β-ol



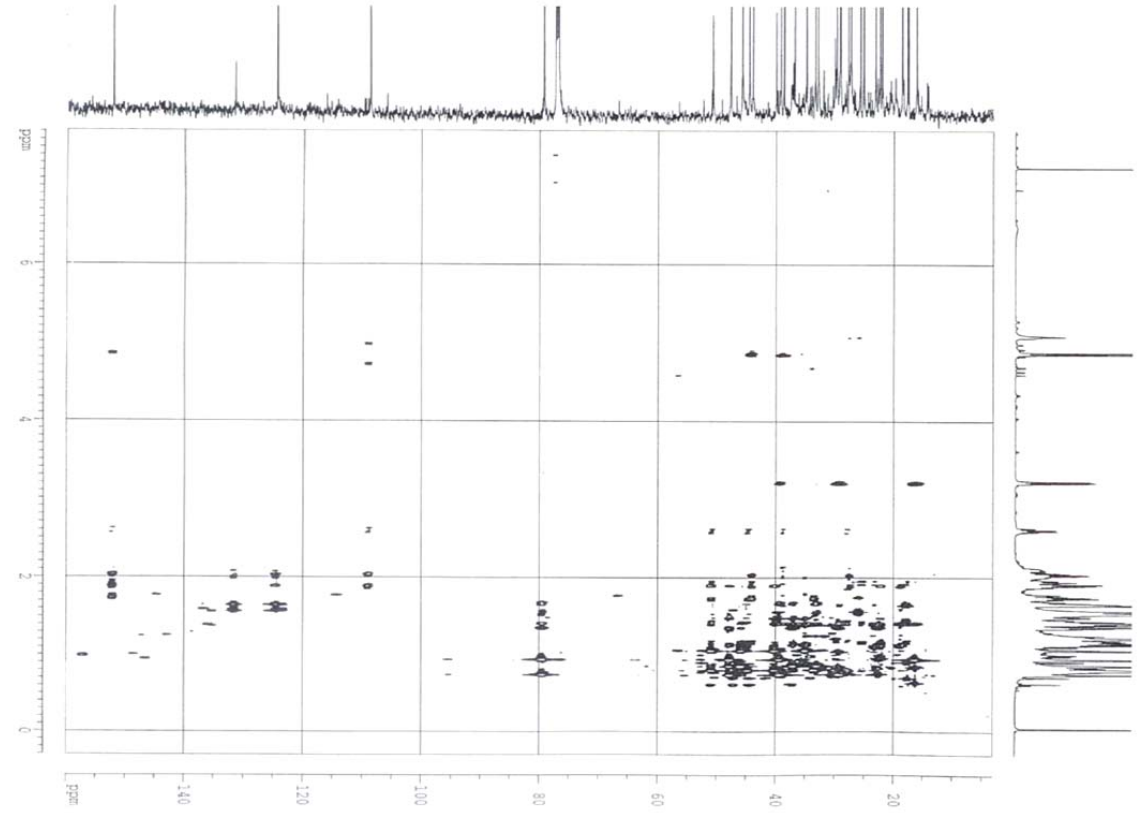
DEPT spectrum of protosta-20,24-dien-3β-ol



HSQC NMR spectrum of protosta-20,24-dien-3β-ol

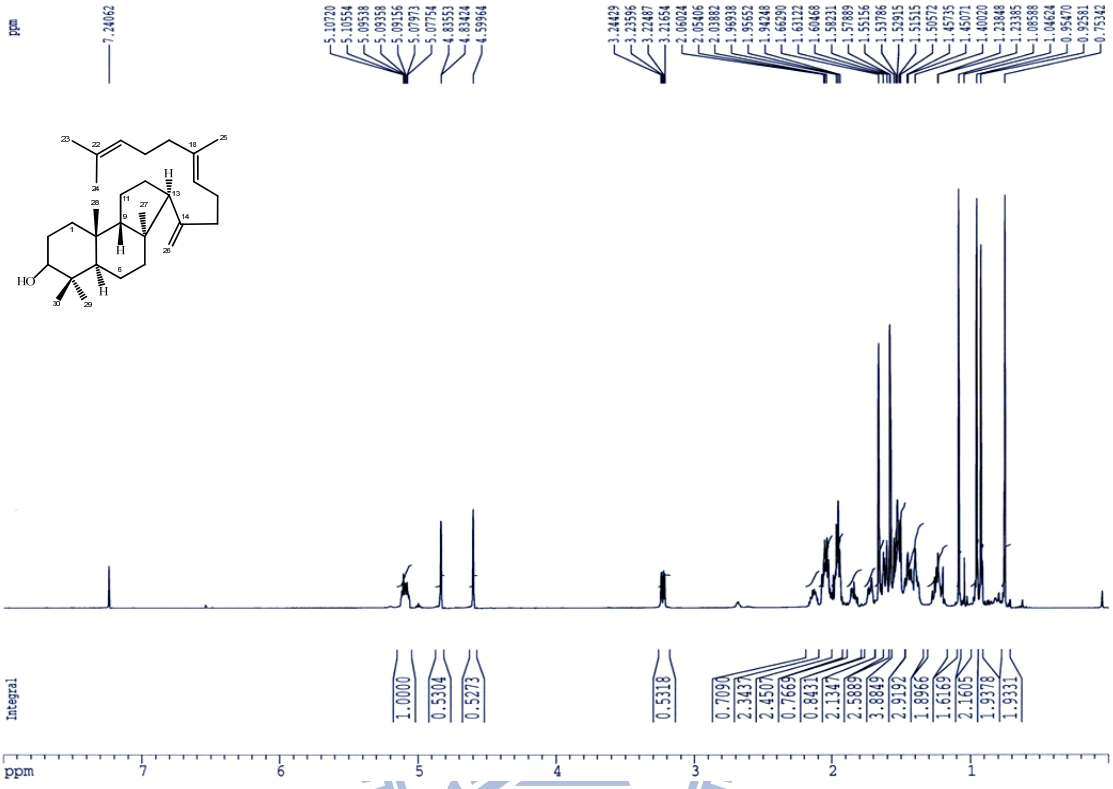


HMBC NMR spectrum of protosta-20,24-dien-3β-ol

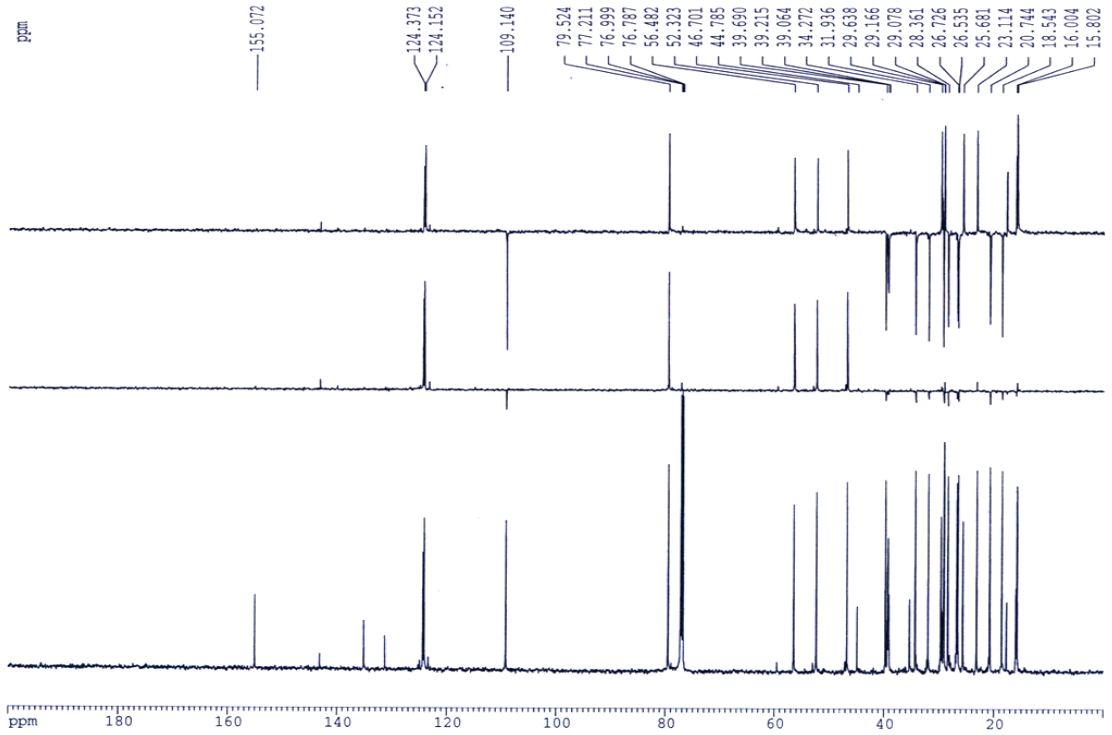


Appendix 3.3. (13*αH*)-isomalabarica-14(26),17*E*,21-trien-3β-ol NMR characterization

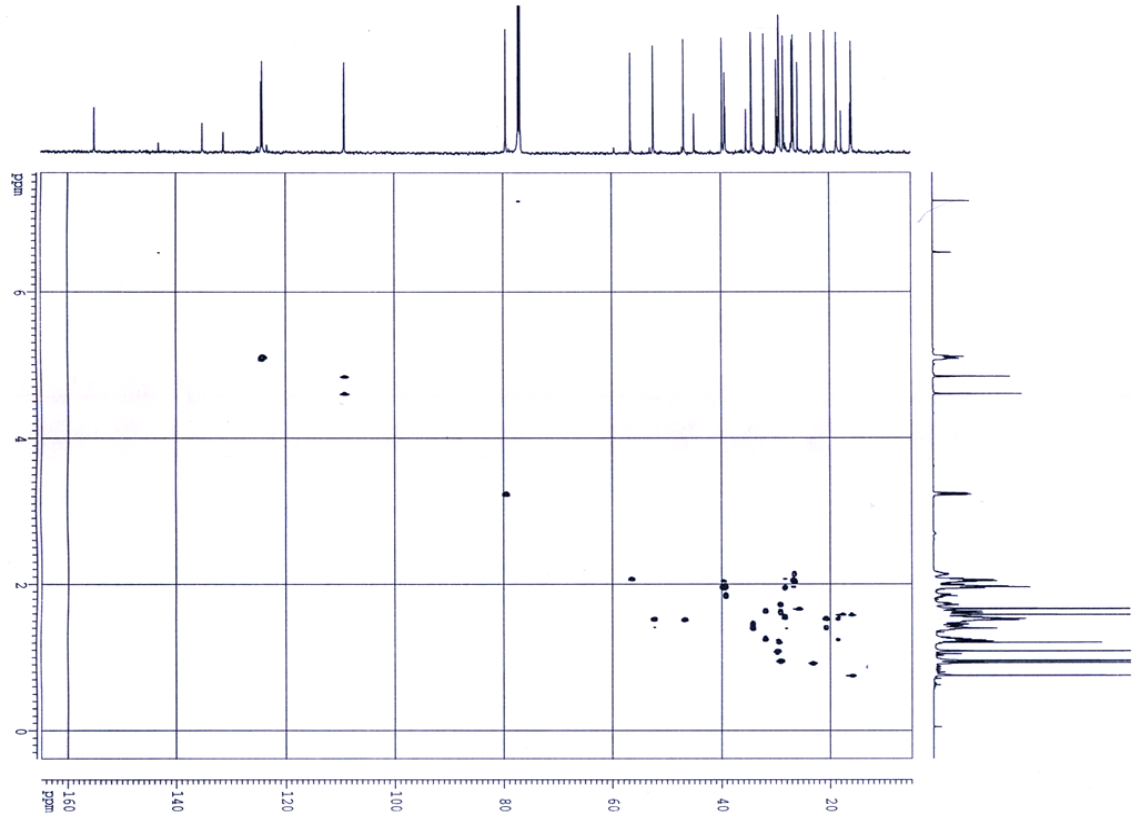
^{1}H-NMR spectrum of (13*αH*)-isomalabarica-14(26),17*E*,21-trien-3β-ol



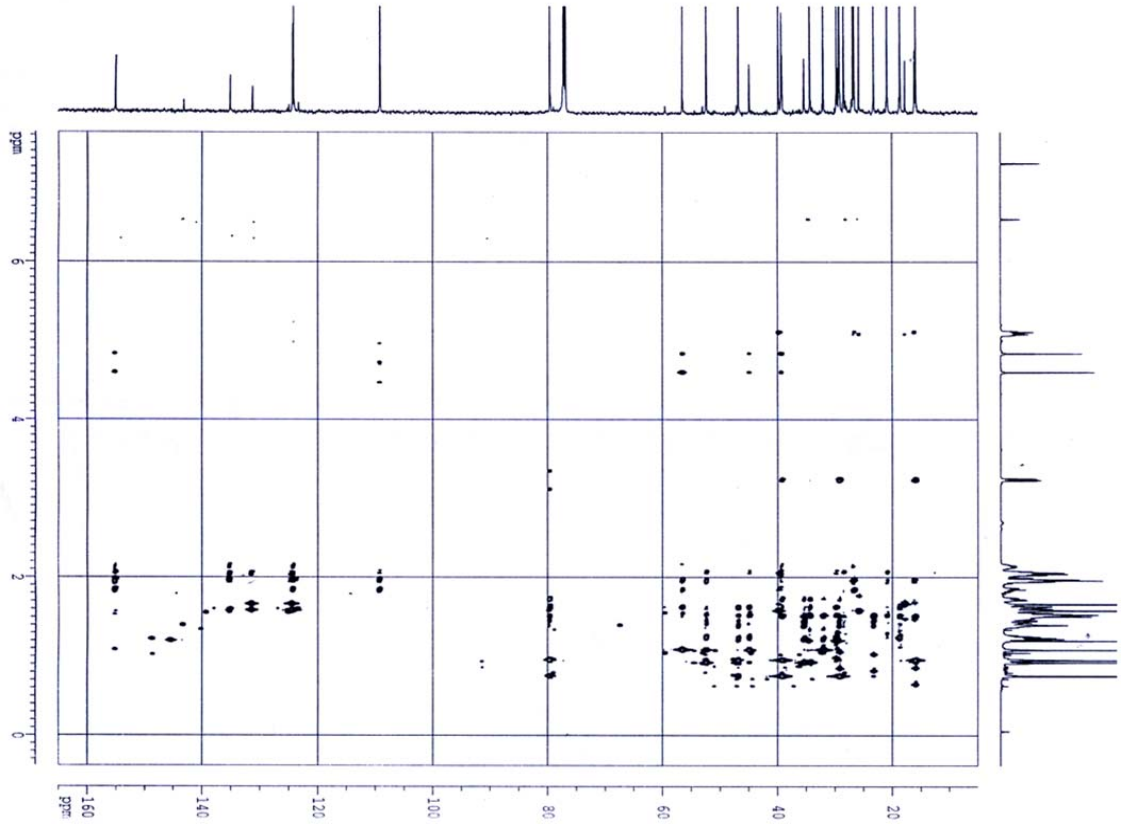
DEPT-NMR spectrum of (13*αH*)-isomalabarica-14(26),17*E*,21-trien-3β-ol



HSQC spectrum of (13*αH*)-isomalabarica-14(26),17*E*,21-trien-3β-ol

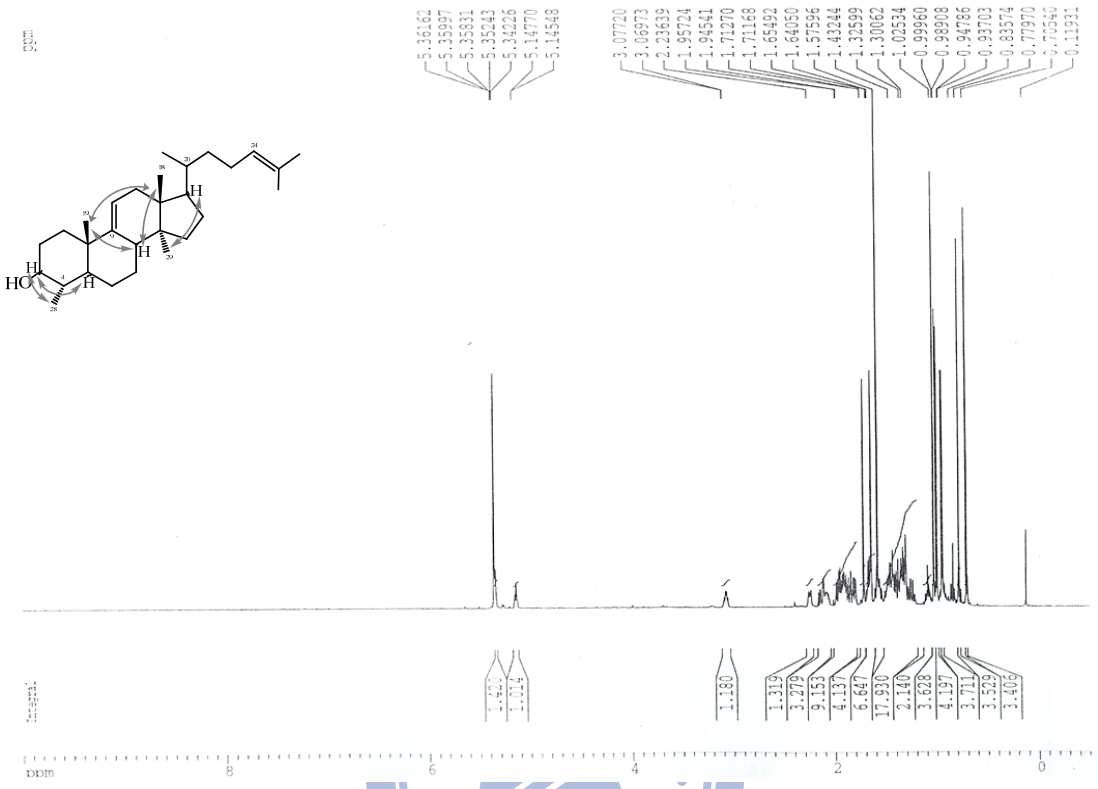


HMBC spectrum of (13*αH*)-isomalabarica-14(26),17*E*,21-trien-3β-ol

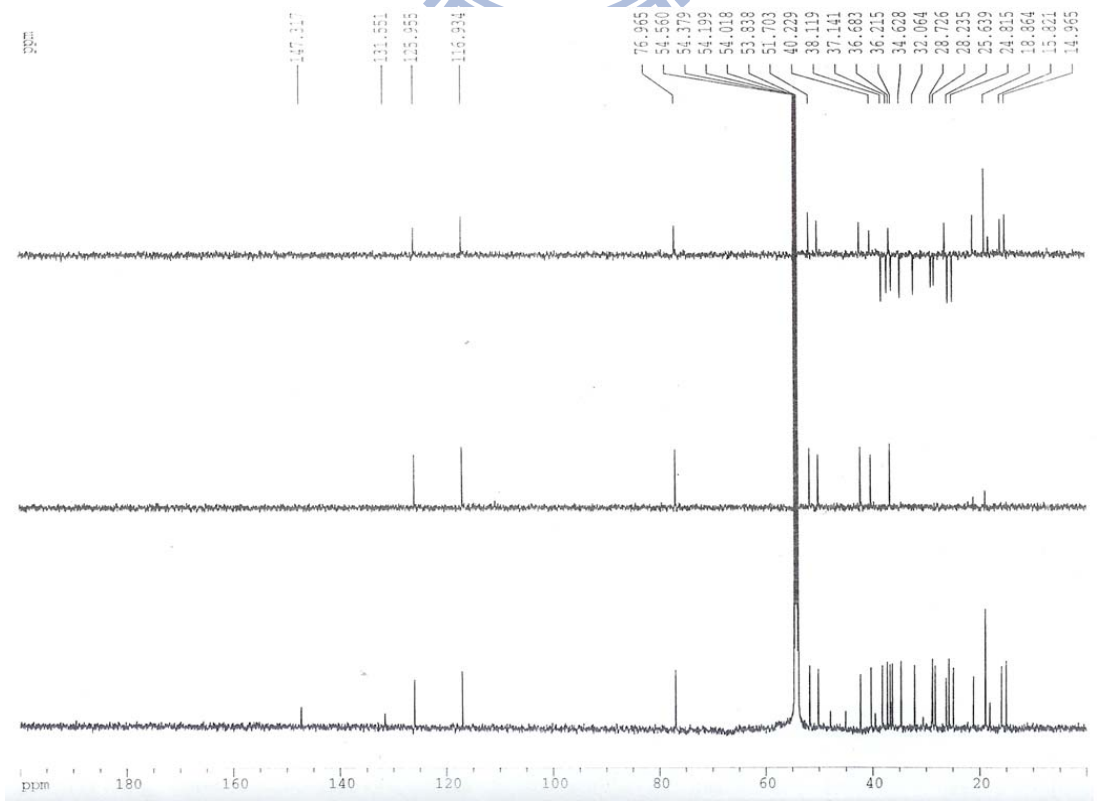


Appendix 5.1. 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol NMR characterization

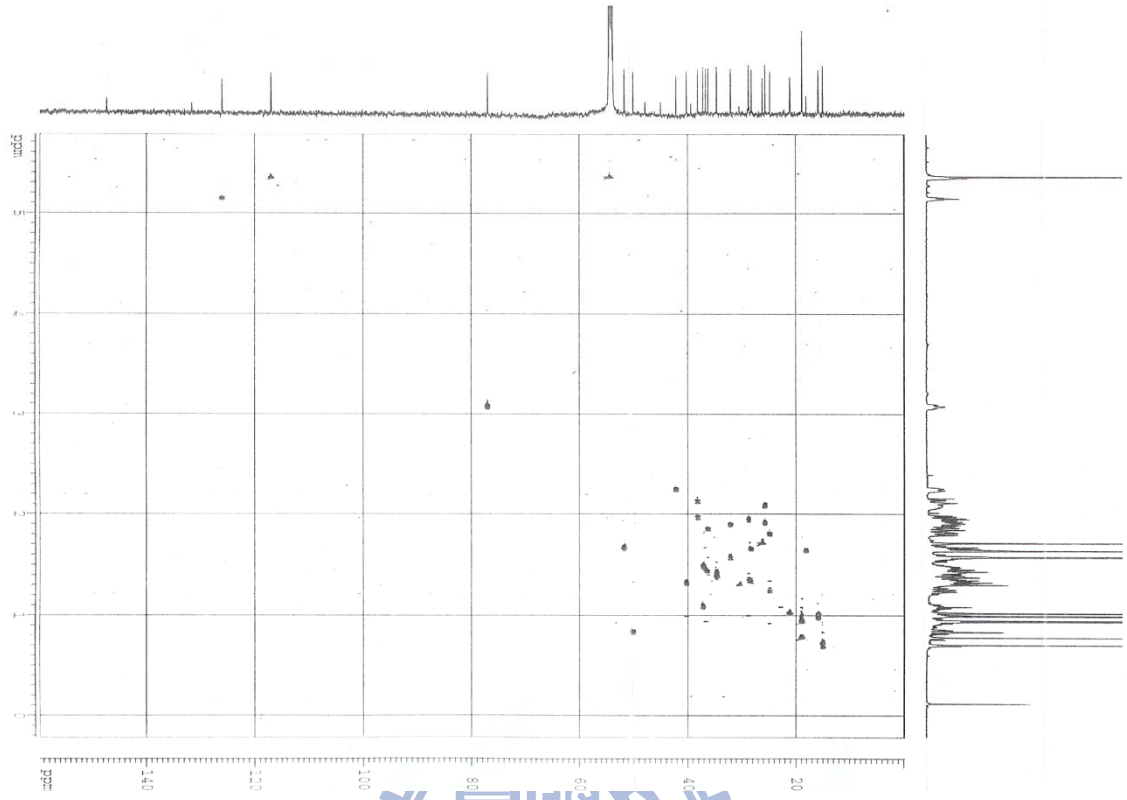
^{1}H-NMR spectrum of 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol



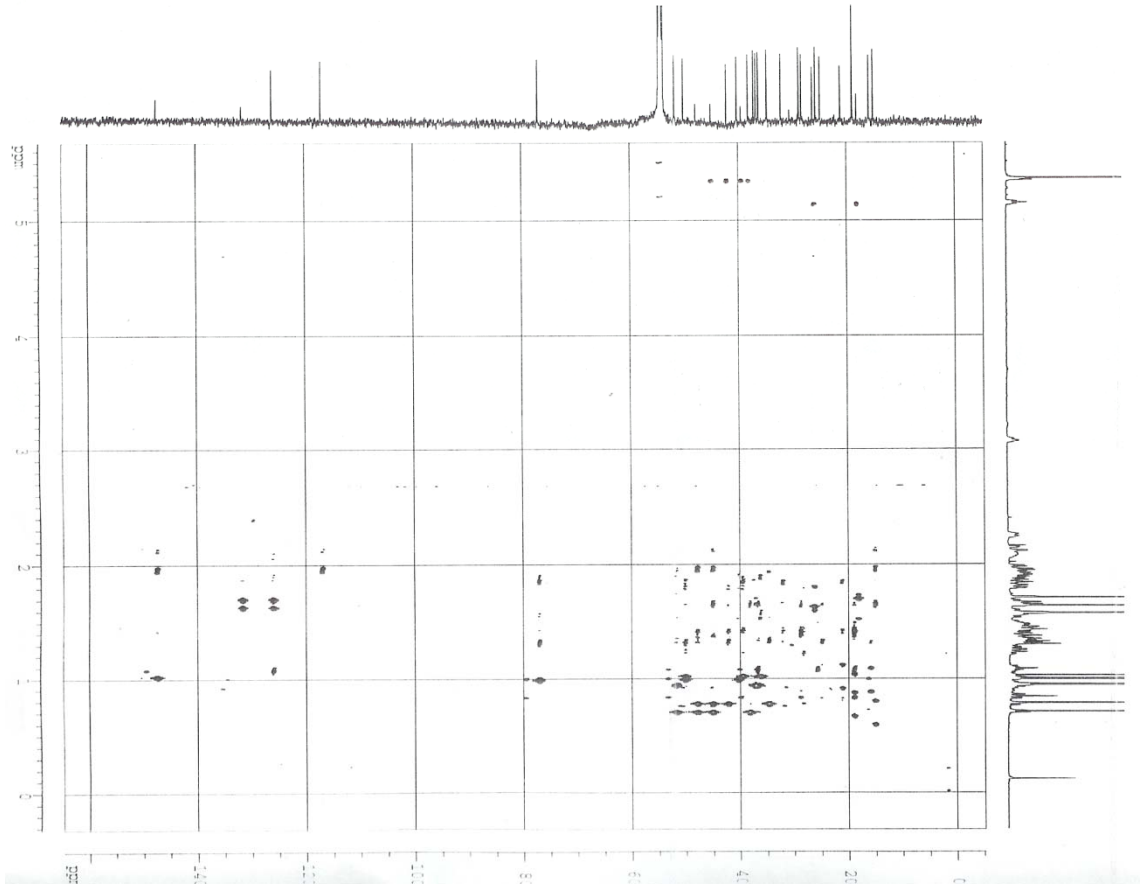
DEPT-NMR spectrum of 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol



HSQC spectrum of 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol

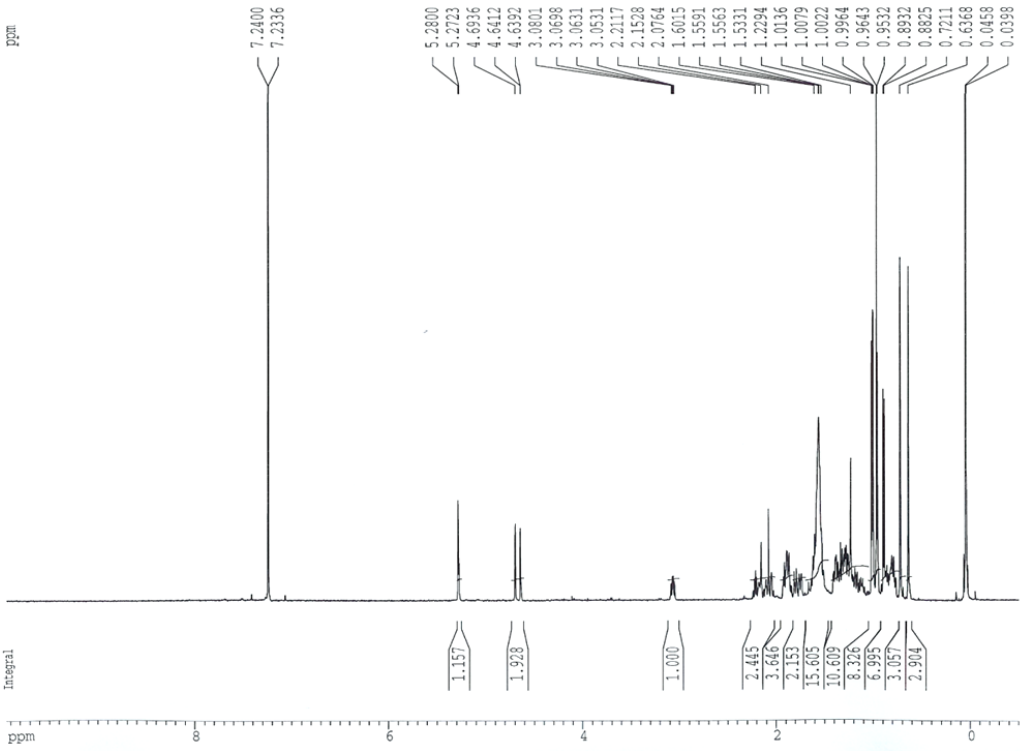


HMBC spectrum of 4α,14α-dimethyl-5α-cholest-9(11)-dien-3β-ol

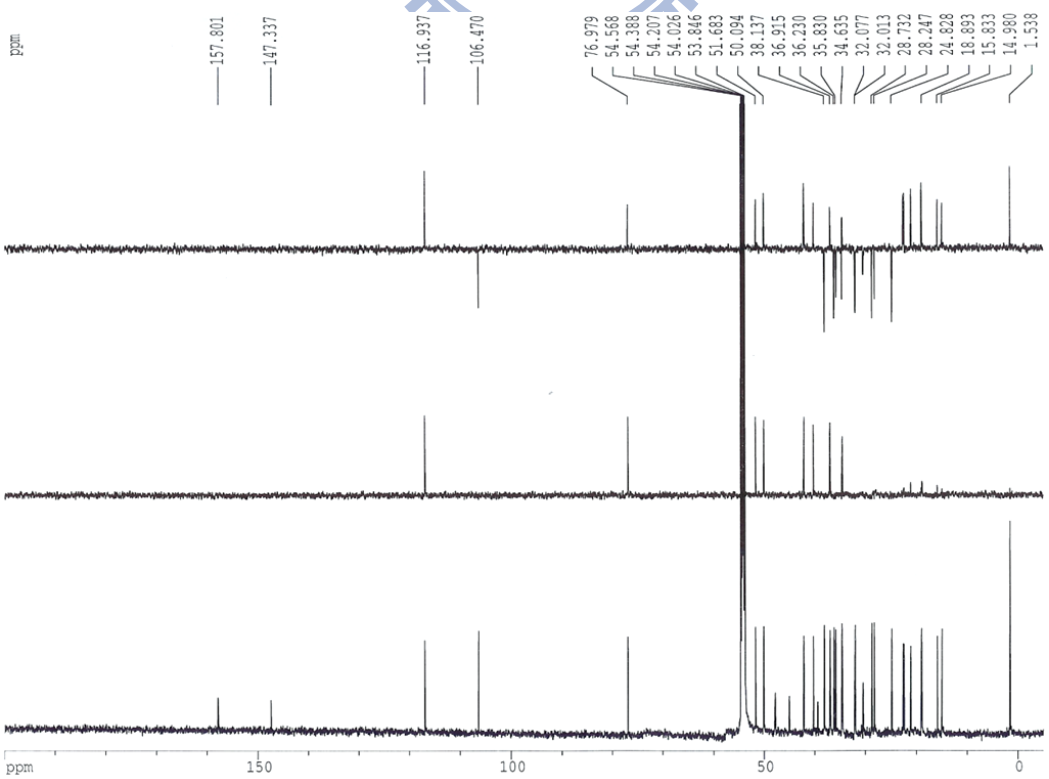


Appendix 5.2. 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol NMR characterization

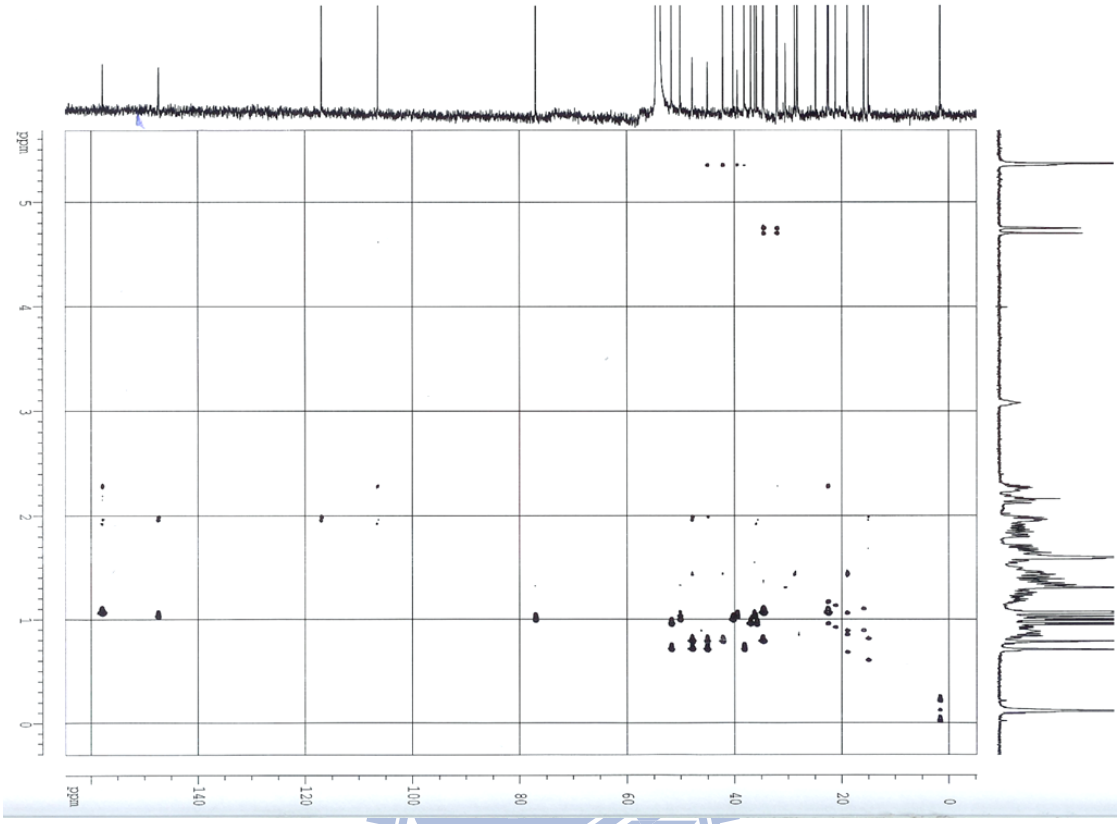
^{1}H-NMR spectrum of 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol



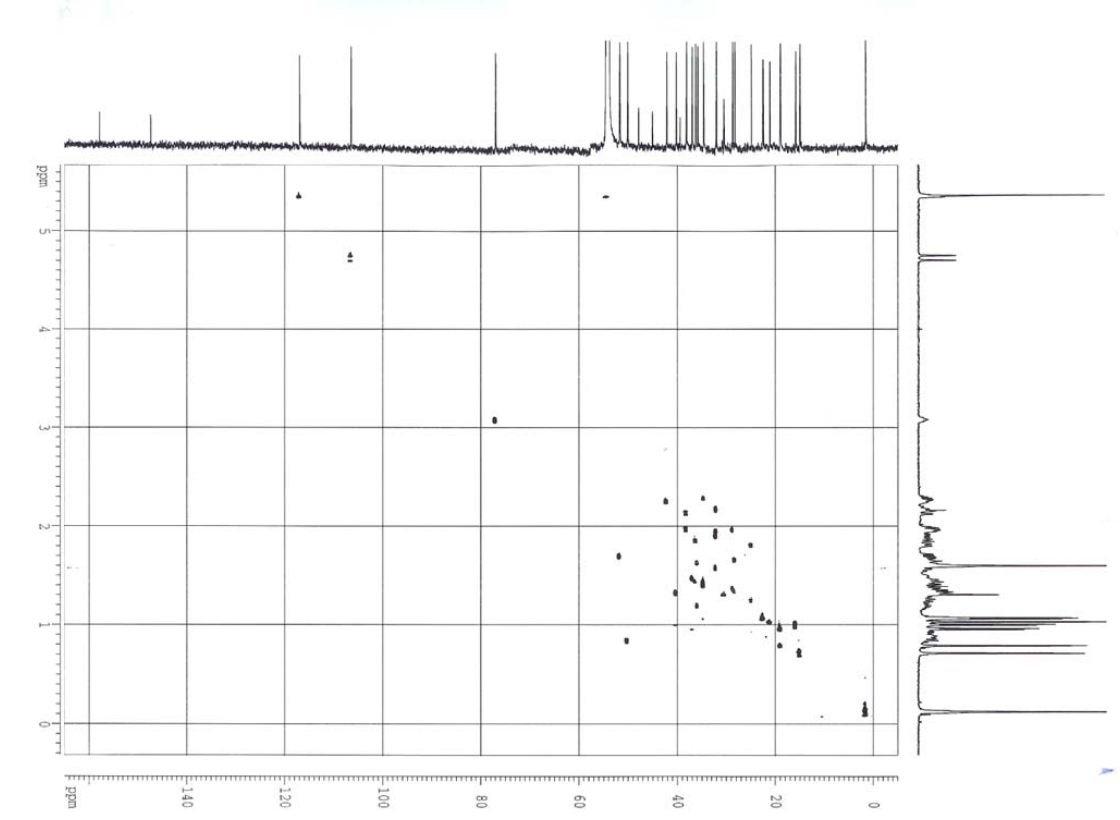
DEPT-NMR spectrum of 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol



HSQC spectrum of 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol



HMBC spectrum of 4α,14α-dimethyl-24-methylene-5α-cholest-9(11)-en-3β-ol



Yuan-Ting Liu was born in Kaohsiung County, Taiwan in 1980. In 2002, she obtained a Bachelor's degree from the Department of Medical Laboratory Science and Biotechnology at Central Taiwan University of Science and Technology, Taichung City, Taiwan. After that, she entered the Department of Biological Science and Technology at National Chiao-Tung University in Hisnchu City, and worked on her Ph.D. program in the field of molecular biology study on oxidosqualene cyclase, under the supervision of Professor Tung-Kung Wu in the Laboratory of Bioorganic and Molecular Evolution. Her research interests are mainly in the structure-function-mechanism relationships of oxidosqualene cyclases, and the molecular engineering of artificial cyclase enzymes for a novel biocatalysis. During these years, she focused on the investigation of relationships between enzymatic active-site residues and cyclization mechanisms via mutagenesis approaches, and has produced several interesting findings. The critically evolutionary divergence of cyclases from different species, especially on the enzymatic control for the stereochemistry and product selectivity, has been discussed in detailed in her recent works by using the site-directed/saturated mutagenesis as well as various structure-based rational enzyme design approaches. Currently, she is working towards a Ph.D. degree, and will receive it in this summer. Moreover, she is also fully interested in the field of molecular biology and bioorganic chemistry to elucidate the structural-functional relationships of various enzymes.